

Algebra & Number Theory

Volume 2
2008
No. 3



mathematical sciences publishers

EDITORS

MANAGING EDITOR
Bjorn Poonen
University of California
Berkeley, USA

EDITORIAL BOARD CHAIR
David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Georgia Benkart	University of Wisconsin, Madison, USA	Susan Montgomery	University of Southern California, USA
Dave Benson	University of Aberdeen, Scotland	Shigefumi Mori	RIMS, Kyoto University, Japan
Richard E. Borcherds	University of California, Berkeley, USA	Andrei Okounkov	Princeton University, USA
John H. Coates	University of Cambridge, UK	Raman Parimala	Emory University, USA
J.-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Victor Reiner	University of Minnesota, USA
Brian D. Conrad	University of Michigan, USA	Karl Rubin	University of California, Irvine, USA
Hélène Esnault	Universität Duisburg-Essen, Germany	Peter Sarnak	Princeton University, USA
Hubert Flenner	Ruhr-Universität, Germany	Michael Singer	North Carolina State University, USA
Edward Frenkel	University of California, Berkeley, USA	Ronald Solomon	Ohio State University, USA
Andrew Granville	Université de Montréal, Canada	Vasudevan Srinivas	Tata Inst. of Fund. Research, India
Joseph Gubeladze	San Francisco State University, USA	J. Toby Stafford	University of Michigan, USA
Ehud Hrushovski	Hebrew University, Israel	Bernd Sturmfels	University of California, Berkeley, USA
Craig Huneke	University of Kansas, USA	Richard Taylor	Harvard University, USA
Mikhail Kapranov	Yale University, USA	Ravi Vakil	Stanford University, USA
Yujiro Kawamata	University of Tokyo, Japan	Michel van den Bergh	Hasselt University, Belgium
János Kollár	Princeton University, USA	Marie-France Vignéras	Université Paris VII, France
Hendrik W. Lenstra	Universiteit Leiden, The Netherlands	Kei-Ichi Watanabe	Nihon University, Japan
Yuri Manin	Northwestern University, USA	Andrei Zelevinsky	Northeastern University, USA
Barry Mazur	Harvard University, USA	Efim Zelmanov	University of California, San Diego, USA

PRODUCTION

ant@mathscipub.org

Paulo Ney de Souza, Production Manager

Silvio Levy, Senior Production Editor

See inside back cover or www.jant.org for submission instructions.

Regular subscription rate for 2008: \$180.00 a year (\$120.00 electronic only).

Subscriptions, requests for back issues from the last three years and changes of subscribers address should be sent to Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840, USA.

Algebra & Number Theory, ISSN 1937-0652, at Mathematical Sciences Publishers, Department of Mathematics, University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

PUBLISHED BY



mathematical sciences publishers

<http://www.mathscipub.org>

A NON-PROFIT CORPORATION

Typeset in L^AT_EX

Copyright ©2008 by Mathematical Sciences Publishers

The intersection of a curve with a union of translated codimension-two subgroups in a power of an elliptic curve

Evelina Viada

Let E be an elliptic curve. An irreducible algebraic curve C embedded in E^g is called *weak-transverse* if it is not contained in any proper algebraic subgroup of E^g , and *transverse* if it is not contained in any translate of such a subgroup.

Suppose E and C are defined over the algebraic numbers. First we prove that the algebraic points of a transverse curve C that are close to the union of all algebraic subgroups of E^g of codimension 2 translated by points in a subgroup Γ of E^g of finite rank are a set of bounded height. The notion of closeness is defined using a height function. If Γ is trivial, it is sufficient to suppose that C is weak-transverse.

The core of the article is the introduction of a method to determine the finiteness of these sets. From a conjectural lower bound for the normalized height of a transverse curve C , we deduce that the sets above are finite. Such a lower bound exists for $g \leq 3$.

Concerning the codimension of the algebraic subgroups, our results are best possible.

1. Introduction

Let A be a semiabelian variety over $\overline{\mathbb{Q}}$ of dimension g . An irreducible algebraic subvariety V of A defined over $\overline{\mathbb{Q}}$ is *weak-transverse* if V is not contained in any proper algebraic subgroup of A , and *transverse* if it is not contained in any translate of such a subgroup.

Given an integer r with $1 \leq r \leq g$ and a subset F of $A(\overline{\mathbb{Q}})$, we define the set

$$S_r(V, F) = V(\overline{\mathbb{Q}}) \cap \bigcup_{\text{codim } B \geq r} (B + F),$$

where B runs over all semiabelian subvarieties of A of codimension at least r and

$$B + F = \{b + f : b \in B, f \in F\}.$$

MSC2000: primary 11G05; secondary 11D45, 11G50, 14K12.

Keywords: heights, diophantine approximation, elliptic curves, counting algebraic points.

Supported by the SNF (Swiss National Science Foundation).

For $r > g$, we define $S_r(V, F)$ to be the empty set. We denote the set $S_r(V, A_{\text{Tor}})$ simply by $S_r(V)$. Note that

$$S_{r+1}(V, F) \subset S_r(V, F).$$

A natural question to ask would be: For which sets F and integers r is the set $S_r(V, F)$ not Zariski-dense in V ?

Sets of this kind, for $r = g$, appear in the literature in the context of the Mordell–Lang, Manin–Mumford and Bogomolov conjectures. More recently Bombieri, Masser and Zannier [Bombieri et al. 1999] proved that $S_2(C)$ is finite for a transverse curve C in a torus. They investigated, for the first time, intersections with the union of all algebraic subgroups of a given codimension. This opens a vast number of conjectures for subvarieties of semiabelian varieties.

In this article we consider the elliptic case for curves. Let E be an elliptic curve and C an irreducible algebraic curve in E^g , both defined over $\overline{\mathbb{Q}}$. Let $\|\cdot\|$ be a seminorm on $E^g(\overline{\mathbb{Q}})$ induced by a height function. For $\varepsilon \geq 0$, we set

$$\mathbb{O}_\varepsilon = \{\xi \in E^g(\overline{\mathbb{Q}}) : \|\xi\| \leq \varepsilon\}.$$

Let $\Gamma \subseteq E^g(\overline{\mathbb{Q}})$ be a subgroup of finite rank. Define $\Gamma_\varepsilon = \Gamma + \mathbb{O}_\varepsilon$.

Conjecture 1.1. *Let $C \subset E^g$.*

- (i) *If C is weak-transverse, $S_2(C)$ is finite.*
- (ii) *If C is transverse, $S_2(C, \Gamma)$ is finite.*
- (iii) *If C is weak-transverse, there exists $\varepsilon > 0$ such that $S_2(C, \mathbb{O}_\varepsilon)$ is finite.*
- (iv) *If C is transverse, there exists $\varepsilon > 0$ such that $S_2(C, \Gamma_\varepsilon)$ is finite.*

The transversality hypothesis is crucially stronger than weak transversality. One should note carefully which hypothesis is assumed in each of the four statements.

Clearly (iv) implies (ii) by setting $\varepsilon = 0$, and similarly (iii) implies (i).

The union of all algebraic subgroups of codimension g is exactly the torsion of E^g . Then, $C \cap \Gamma_\varepsilon \subset S_g(C, \Gamma_\varepsilon) \subset S_2(C, \Gamma_\varepsilon)$. So, Conjecture 1.1(iii) implies the Bogomolov Theorem [Ullmo 1998; Zhang 1998], and (iv) implies Mordell–Lang plus Bogomolov [Poonen 1999].

Partial results related to (i) and (ii) have been proved. In [Viada 2003] we solved a weak form of (i), namely we assumed the stronger hypothesis that C is transverse. If E has CM (complex multiplication) then $S_2(C)$ is finite. If E has no CM then $S_{(g/2)+2}(C)$ is finite. In [Rémond and Viada 2003] a weak version of (ii) was presented. Again if E has CM the result is optimal. If E has no CM the codimension of the algebraic subgroups depends on Γ . In addition, we show that (i) and (ii) are equivalent. There are no trivial implications between (iii) and (iv), because of the different hypotheses on C .

These known proofs rely on Northcott’s theorem: a set is finite if and only if it has bounded height and degree. To prove that the degree is bounded one uses Siegel’s Lemma and an essentially optimal generalized Lehmer’s Conjecture. Up to a logarithmic factor, the generalized Lehmer conjecture is presently known for a point in a torus [Amoroso and David 1999] and in a CM abelian variety [David and Hindry 2000]. This method has some disadvantages: it is only known to work for transverse curves and for $\varepsilon = 0$, and a quasioptimal generalized Lehmer’s Conjecture is not likely to be proved in a near future for a general abelian variety.

In this article we introduce a different method. First, we bound the height also for weak-transverse curves.

Theorem 1.2. *There exists $\varepsilon > 0$ such that:*

- (i) *If C is weak-transverse, $S_2(C, \mathbb{O}_\varepsilon)$ has bounded height.*
- (ii) *If C is transverse, $S_2(C, \Gamma_\varepsilon)$ has bounded height.*

The proof of both statements uses a Vojta inequality, as stated in Proposition 2.1 of [Rémond and Viada 2003]. The second assertion is proved in Theorem 1.5 of the same paper. To prove the first assertion (see Section 7), we embed $S_2(C, \mathbb{O}_\varepsilon)$ into two sets associated to a transverse curve. We then manage to apply a Vojta inequality on each of these two sets.

As a second result, we prove:

Theorem 1.3. *For $r \geq 2$, the following statements are equivalent:*

- (i) *If C is weak-transverse, there exists $\varepsilon > 0$ such that $S_r(C, \mathbb{O}_\varepsilon)$ is finite.*
- (ii) *If C is transverse, there exists $\varepsilon > 0$ such that $S_r(C, \Gamma_\varepsilon)$ is finite.*

That (i) implies (ii) is elementary, but the converse implication is not as easy as the equivalence of (i) and (ii) in Conjecture 1.1. In particular we make use of Theorem 1.2 (see Section 7).

In the third instance, we show how to avoid the use of the Siegel Lemma and the generalized Lehmer Conjecture. Instead, we use Dirichlet’s Theorem and a conjectural effective version of the Bogomolov Theorem. Bogomolov’s Theorem states that the set of points of small height on a curve of genus at least 2 is finite. We define $\mu(C)$ as the supremum of the reals $\epsilon(C)$ such that $S_g(C, \mathbb{O}_{\epsilon(C)}) = C \cap \mathbb{O}_{\epsilon(C)}$ is finite. The essential minimum of C is $\mu(C)^2$. (Often in the literature the notation \mathbb{O}_ε corresponds to what we write as $\mathbb{O}_{\varepsilon^2}$; thus in the references given below the bounds are given for the essential minimum and not for its square root $\mu(C)$ as we do here.)

Nonoptimal effective lower bounds for $\mu(C)$ are given by S. David and P. Philippon [2002, Theorem 1.4; 2007, Theorem 1.6]. The lower bound we need is the elliptic analogue of [Amoroso and David 2003, Theorem 1.4], which gives a quasioptimal lower bound for the essential minimum of a variety.

The following conjecture is a weak form of [David and Philippon 2007, Conjecture 1.5(ii)] where the line bundle is fixed.

Conjecture 1.4. *Let $A = E_1 \times \cdots \times E_g$ be a product of elliptic curves defined over a number field k . Let L be the tensor product of the pullbacks of symmetric line bundles on E_i via the natural projections. Let $C \subset A$ be an irreducible transverse curve defined over $\overline{\mathbb{Q}}$. Let η be any positive real. Then there exists a constant $c(g, A, \eta) = c(g, \deg_L A, h_L(A), [k : \mathbb{Q}], \eta)$ such that, for*

$$\epsilon(C, \eta) = c(g, A, \eta)(\deg_L C)^{-1/(2(g-1))-\eta},$$

the set

$$C(\overline{\mathbb{Q}}) \cap \mathcal{O}_{\epsilon(C, \eta)}$$

is finite.

In Section 11, we prove:

Theorem 1.5. *Conjecture 1.4 implies Conjecture 1.1.*

Conjecture 1.4 can be stated for subvarieties of A . Galateau [2007] proved that such a conjecture holds for varieties of codimension 1 or 2 in a product of elliptic curves. Then, for $g \leq 3$, Conjecture 1.1 holds unconditionally.

Theorems 1.2 and 1.5 are optimal with respect to the codimension of the algebraic subgroups; see Remark 9.2.

We have already pointed out that Conjecture 1.1 implies the Bogomolov Conjecture and the Mordell–Lang plus Bogomolov Theorem. Let us emphasize that our Theorem 1.5 does not give a new proof of the Bogomolov Conjecture, as we assume such an effective result. On the other hand, it gives a new proof of the Mordell–Lang plus Bogomolov Theorem, under the assumption of Conjecture 1.4.

The proof of Theorem 1.5 is based on the observation that a union of sets is finite if and only if

- (1) the union can be taken over finitely many sets, and
- (2) all sets in the union are finite.

Showing (1) is a typical problem of Diophantine approximation. The proof relies on Dirichlet’s Theorem on the rational approximation of reals. The fact that we consider small neighborhoods enables us to move the algebraic subgroups “a bit”. So we can consider only subgroups of bounded degree, of which there are finitely many; see Proposition A, Section 12.

Step (2) takes place in the context of height theory. Its proof relies on Conjecture 1.4. The bound $\epsilon(C, \eta)$ depends on the invariants of the ambient variety and on the degree of C . A weaker dependence on the degree of C would not be enough for our application. Also the independence of the bound from the field of definition

of C proves useful. Playing on Conjecture 1.4, we produce a sharp lower bound for the essential minimum of the image of a curve under certain morphisms (see Proposition B and Section 13).

The effectiveness aspect of our method is noteworthy; the use of a Vojta inequality makes Theorem 1.2, and consequently Theorem 1.5, ineffective. Though, the rest of the method is effective. Indeed, in Section 14, we prove a weaker, but effective analogue of Theorem 1.5.

Theorem 1.6. *Assume Conjecture 1.4. If C is transverse, there exists an effective $\varepsilon > 0$ such that the set $S_2(C, \mathbb{O}_\varepsilon)$ is finite.*

A bound for the number of points of small height on the curve would then imply a bound on the cardinality of $S_2(C, \mathbb{O}_\varepsilon)$ for C transverse and ε small (Theorem 14.3).

The toric version of Theorem 1.6 was independently studied by P. Habegger in his Ph.D. thesis [2007]. He follows the idea of using a Bogomolov-type bound, proved in the toric case in [Amoroso and David 2003, Theorem 1.4]. He proves the finiteness of $S_2(C, \mathbb{O}_\varepsilon)$, for $\varepsilon > 0$ and C a transverse curve in a torus.

2. Preliminaries

Morphisms and their height. Let $(R, |\cdot|)$ be a hermitian ring, that means R is a domain and $|\cdot|$ an absolute value on R .

We denote by $M_{r,g}(R)$ the module of $r \times g$ matrices with entries in R .

For $F = (f_{ij}) \in M_{r,g}(R)$, we define the height of F as the maximum of the absolute value of its entries

$$H(F) = \max_{ij} |f_{ij}|.$$

Let E be an elliptic curve defined over a number field. The ring of endomorphism $\text{End } E$ is isomorphic either to \mathbb{Z} (if E does not have CM) or to an order in an imaginary quadratic field (if E has CM). We consider on $\text{End } E$ the standard absolute value of \mathbb{C} . This absolute value does not depend on the embedding of $\text{End } E$ in \mathbb{C} . An intrinsic definition of absolute value on $\text{End } E$ can be given using the Rosati involution.

We identify a morphism $\phi : E^g \rightarrow E^r$ with a matrix in $M_{r,g}(\text{End } E)$. The set of morphisms of height bounded by a constant is finite.

In the following, we aim to be as transparent as possible, polishing statements from technicality. Therefore, we principally present proofs for E without CM. Then $\text{End } E$ is identified with \mathbb{Z} and a morphism ϕ with an integral matrix. In the final section, we explain how to deal with the technical complication of a ring of endomorphisms of rank 2 and with a product of elliptic curves instead of a power.

Small points. On E , we fix a symmetric very ample line bundle \mathcal{L} . On E^g , we consider the bundle L which is the tensor product of the pullbacks of \mathcal{L} via the natural projections on the factors. Degrees are computed with respect to the polarization L .

Usually $E^g(\overline{\mathbb{Q}})$ is endowed with the L -canonical Néron–Tate height h' . Though, to simplify constants, we prefer to define on E^g the height of the maximum

$$h(x_1, \dots, x_g) = \max_i(h(x_i)).$$

where $h(\cdot)$ on $E(\overline{\mathbb{Q}})$ is the \mathcal{L} -canonical Néron–Tate height. The height h is the square of a norm $\|\cdot\|$ on $E^g(\overline{\mathbb{Q}}) \otimes \mathbb{R}$. For a point $x \in E^g(\overline{\mathbb{Q}})$, we write $\|x\|$ for $\|x \otimes 1\|$.

Note that $h(x) \leq h'(x) \leq gh(x)$. Hence, the two norms induced by h and h' are equivalent.

For $a \in \text{End } E$, we denote by $[a]$ the multiplication by a . For $y \in E^g(\overline{\mathbb{Q}})$ we have

$$\|[a]y\| = |a| \cdot \|y\|.$$

The height of a nonempty set $S \subset E^g(\overline{\mathbb{Q}})$ is the supremum of the heights of its elements. The norm of S is the nonnegative square root of its height.

For $\varepsilon \geq 0$, we denote

$$\mathbb{O}_\varepsilon = \mathbb{O}_{\varepsilon, E^g} = \{\xi \in E^g(\overline{\mathbb{Q}}) : \|\xi\| \leq \varepsilon\}.$$

Subgroups. Let M be a R -module. The R -rank of M is the supremum of the cardinality of a set of R -linearly independent elements of M . If M has finite rank s , a maximal free set of M is a set of s linearly independent elements of M . If M is a free R -module of rank s , we call a set of s generators of M , integral generators of M .

Note that a free \mathbb{Z} -module of finite rank is a lattice; in the literature, what we call integral generators can be called basis, and what we define as maximal free set is a basis of the vector space given by tensor product with the quotient field of R .

We say that $(M, \|\cdot\|)$ is a hermitian R -module if M is an R -module and $\|\cdot\|$ is a norm on the tensor product of M with the quotient field of R . For an element $p \in M$ we write $\|p\|$ for $\|p \otimes 1\|$.

Let E be an elliptic curve. In the following, we will simply say module for an $\text{End } E$ -module.

Note that any subgroup of $E^g(\overline{\mathbb{Q}})$ of finite rank is contained in a submodule of finite rank. Conversely, a submodule of E^g of finite rank is a subgroup of finite rank.

Let Γ be a subgroup of finite rank of $E^g(\overline{\mathbb{Q}})$. We define

$$\Gamma_\varepsilon = \Gamma + \mathbb{O}_\varepsilon.$$

The saturated module Γ_0 of the coordinates group of Γ (in short of Γ) is a submodule of $E(\overline{\mathbb{Q}})$ defined as

$$\Gamma_0 = \{\phi(y) \in E \text{ for } \phi : E^g \rightarrow E \text{ and } Ny \in \Gamma \text{ with } N \in \mathbb{Z}^*\}. \tag{2-1}$$

Note that $\Gamma_0^g = \Gamma_0 \times \cdots \times \Gamma_0$ is a submodule of E^g invariant via the image or preimage of isogenies. Furthermore, it contains Γ and it is a module of finite rank. Thus to prove finiteness statements for Γ it is enough to prove them for Γ_0^g .

We denote by s the rank of Γ_0 . Let $\gamma_1, \dots, \gamma_s$ be a maximal free set of Γ_0 . We denote the associated point of E^s by

$$\gamma = (\gamma_1, \dots, \gamma_s).$$

For $p = (p_1, \dots, p_s) \in E^s$ we define Γ_p as the saturated module of $\langle p_1, \dots, p_s \rangle$.

3. Some geometry of numbers

We present a property from the geometry of numbers and extend it to points of $E^g(\overline{\mathbb{Q}})$. The idea is that, if in \mathbb{R}^n we consider n linearly independent vectors and move them within a “small” angle, they will still be linearly independent. The norm of a linear combination of such vectors depends on the norm of these vectors, on their angles, and on the norm of the coefficients of the combination. Such estimates are frequent in the geometry of numbers.

Lemma 3.1 (compare [Schlickewei 1997, Theorem 1.1; Viada 2003, Lemma 3]). *Every hermitian free \mathbb{Z} -module of rank n admits integral generators ρ_1, \dots, ρ_n such that*

$$c_0(n) \sum_i |\alpha_i|^2 \|\rho_i\|^2 \leq \left\| \sum_i \alpha_i \rho_i \right\|^2$$

for all integers α_i , where $c_0(n)$ is a constant depending only on n .

Proof. A hermitian free \mathbb{Z} -module $(\Gamma, \|\cdot\|)$ of rank n is a lattice in the metric space $\Gamma_{\mathbb{R}}$ given by tensor product with \mathbb{R} . The proof now follows that of [Viada 2003, Lemma 3] (page 57, from line 19 onwards), with n instead of r and ρ_i instead of g_i . □

This lemma allows us to explicit the comparison constant for two norms on a finite-dimensional vector space over the quotient field of R .

Proposition 3.2. *Let $(M, \|\cdot\|)$ be a hermitian R -module, where R is a finitely generated free \mathbb{Z} -module. Let p_1, \dots, p_s be R -linearly independent elements of M . Then there exists an effective positive constant $c_1(p, \tau)$ such that*

$$c_1(p, \tau) \sum_i |b_i|_R^2 \|p_i\|^2 \leq \left\| \sum_i b_i p_i \right\|^2$$

for all $b_1, \dots, b_s \in R$, where $p = (p_1, \dots, p_s)$ and $\tau = (1, \tau_2, \dots, \tau_t)$ are integral generators of R .

Proof. The submodule of M defined by $\Gamma_{\mathbb{Z}} = \langle p_1, \dots, p_s, \dots, \tau_1 p_1, \dots, \tau_t p_s \rangle_{\mathbb{Z}}$ has rank st over \mathbb{Z} . Clearly, for $1 \leq i \leq t$ and $1 \leq j \leq s$ the elements $\tau_i p_j$ are integral generators of $\Gamma_{\mathbb{Z}}$. Consider the normed space $(M \otimes_{\mathbb{Z}} \mathbb{R}, \|\cdot\|)$, in which $\Gamma_{\mathbb{Z}}$ is embedded, and endow $\Gamma_{\mathbb{Z}}$ with the induced metric.

Apply Lemma 3.1 to $(\Gamma_{\mathbb{Z}}, \|\cdot\|)$ with $n = st$. Then, there exist integral generators ρ_1, \dots, ρ_{st} of $\Gamma_{\mathbb{Z}}$ satisfying

$$\left\| \sum_i \alpha_i \rho_i \right\|^2 \geq c_0(st) \sum_i |\alpha_i|^2 \|\rho_i\|^2 \geq c_0(st) \sum_i |\alpha_i|^2 \min_k \|\rho_k\|^2, \quad (3-1)$$

for all $\alpha_1, \dots, \alpha_{st} \in \mathbb{Z}$.

We decompose the elements $b_1, \dots, b_s \in R$ as

$$b_i = \sum_{j=1}^t \alpha_{ij} \tau_j$$

with $\alpha_{ij} \in \mathbb{Z}$. We set

$$\alpha = (\alpha_{11}, \dots, \alpha_{1t}, \dots, \alpha_{s1}, \dots, \alpha_{st}) \in \mathbb{Z}^{st}.$$

Next we write

$$p^\tau = (\tau_1 p_1, \dots, \tau_t p_1, \tau_1 p_2, \dots, \tau_t p_2, \dots, \tau_1 p_s, \dots, \tau_t p_s)^\text{T} \in \Gamma_{\mathbb{Z}}^{st},$$

$$\rho = (\rho_1, \dots, \rho_{st})^\text{T} \in \Gamma_{\mathbb{Z}}^{st},$$

where the superscript T indicates the transpose, as usual. Let $P \in \text{SL}_{st}(\mathbb{Z})$ be the base change matrix such that

$$p^\tau = P\rho.$$

Then

$$\sum_i b_i p_i = \sum_{ij} \alpha_{ij} \tau_j p_i = \alpha \cdot p^\tau = \alpha \cdot (P\rho) = (\alpha P) \cdot \rho.$$

Passing to the norms and using relation (3-1) with the coefficients $(\alpha_1, \dots, \alpha_{st}) = \alpha P$, we deduce

$$\left\| \sum_i b_i p_i \right\|^2 = \|(\alpha P) \cdot \rho\|^2 \geq c_0(st) |\alpha P|_2^2 \min_k \|\rho_k\|^2,$$

where $|\cdot|_2$ is the standard Euclidean norm. On the other hand, the triangle inequality gives

$$|b_i|_R^2 \leq \max_k |\tau_k|_R^2 \left(\sum_{j=1}^t |\alpha_{ij}| \right)^2 \leq t \max_k |\tau_k|_R^2 \sum_{j=1}^t |\alpha_{ij}|^2.$$

We deduce

$$\frac{\| \sum_i b_i p_i \|^2}{\sum_i |b_i|_R^2 \|p_i\|^2} \geq \frac{c_0(st)}{t \max_j |\tau_j|_R^2} \frac{\min_i \|\rho_i\|^2}{\max_i \|p_i\|^2} \frac{|\alpha P|_2^2}{|\alpha|_2^2}.$$

We shall still estimate $|\alpha P|_2^2/|\alpha|_2^2$ independently of α . For a linear operator A and a row vector β , there holds the classical norm relation $|\beta A|_2 \leq H(A)|\beta|_2$. For $A = P^{-1}$ and $\beta = \alpha P$, we deduce

$$\frac{|\alpha P|_2^2}{|\alpha|_2^2} \geq \frac{1}{H(P^{-1})^2}.$$

Then

$$\frac{\| \sum_i b_i p_i \|^2}{\sum_i |b_i|_R^2 \|p_i\|^2} \geq \frac{c_0(st)}{t \max_j |\tau_j|_R^2} \frac{\min_i \|\rho_i\|^2}{\max_i \|p_i\|^2} \frac{1}{H(P^{-1})^2}$$

or equivalently

$$\left\| \sum b_i p_i \right\|^2 \geq c_1(p, \tau) \sum_i |b_i|_R^2 \|p_i\|^2,$$

where

$$c_1(p, \tau) = \frac{c_0(st)}{t \max_j |\tau_j|_R^2} \frac{\min_i \|\rho_i\|^2}{\max_i \|p_i\|^2} \frac{1}{H(P^{-1})^2}. \quad \square$$

The following unsurprising proposition has some surprising implications; it allows us to prove Theorems 1.2 and 1.3.

Proposition 3.3. *Let p_1, \dots, p_s be linearly independent points of $E(\overline{\mathbb{Q}})$ and $p = (p_1, \dots, p_s)$. Let τ be a set of integral generators of $\text{End } E$. Then, there exist positive reals $c_2(p, \tau)$ and $\varepsilon_0(p, \tau)$ such that*

$$c_2(p, \tau) \sum_i |b_i|^2 \|p_i\|^2 \leq \left\| \sum_i b_i (p_i - \xi_i) - b\zeta \right\|^2$$

for all $b_1, \dots, b_s, b \in \text{End } E$ with $|b| \leq \max_i |b_i|$ and for all $\xi_1, \dots, \xi_s, \zeta \in E(\overline{\mathbb{Q}})$ with $\|\xi_i\|, \|\zeta\| \leq \varepsilon_0(p, \tau)$.

In particular $p_1 - \xi_1, \dots, p_s - \xi_s$ are linearly independent points of E .

Proof. Recall that the norm on $\text{End } E$ is compatible with the height norm on $E(\overline{\mathbb{Q}})$, that is, $\|b_i p_i\| = |b_i|_{\text{End } E} \|p_i\|$. Thus $(\text{End } E, |\cdot|)$ is a hermitian free \mathbb{Z} -module of

rank 1 if E has no CM or 2 is E has CM. Furthermore, $(E, \|\cdot\|)$ is a hermitian End E -module.

Apply Proposition 3.2 with $R = \text{End } E$, $M = E$ and $\tau = (1)$ if $\text{End } E \cong \mathbb{Z}$ or $\tau = (1, \tau_2)$ if $\text{End } E \cong \mathbb{Z} + \tau_2\mathbb{Z}$. For $b_1, \dots, b_s \in \text{End } E$, we obtain

$$\left\| \sum b_i p_i \right\|^2 \geq c_1(p, \tau) \sum_i |b_i|^2 \|p_i\|^2. \tag{3-2}$$

Let $\|\xi_i\|, \|\zeta\| \leq \varepsilon$. Since $|b| \leq \max |b_i|$ the triangle inequality implies

$$\begin{aligned} \left\| \sum_i b_i(p_i - \xi_i) - b\zeta \right\| &\geq \left\| \sum_i b_i p_i \right\| - \varepsilon \sum_i |b_i| - \varepsilon |b| \\ &\geq \left\| \sum_i b_i p_i \right\| - 2\varepsilon \sum_i |b_i|. \end{aligned}$$

Squaring and keeping in mind that $(\sum_{i=1}^s |b_i|)^2 \leq s \sum_{i=1}^s |b_i|^2$, we deduce

$$\begin{aligned} \left\| \sum_i b_i(p_i - \xi_i) - b\zeta \right\|^2 &\geq \left\| \sum_i b_i p_i \right\|^2 - 4\varepsilon \left\| \sum_i b_i p_i \right\| \sum_i |b_i| + 4\varepsilon^2 \left(\sum_i |b_i| \right)^2 \\ &\geq \left\| \sum_i b_i p_i \right\|^2 - 4s\varepsilon \left(\sum_i |b_i|^2 \right) \max_i \|p_i\|. \end{aligned}$$

Choose

$$\varepsilon \leq \varepsilon_0(p, \tau) = \frac{c_1(p, \tau)}{8s} \frac{\min_i \|p_i\|^2}{\max_i \|p_i\|}. \tag{3-3}$$

Using relation (3-2), we deduce

$$\begin{aligned} \left\| \sum_i b_i(p_i - \xi_i) - b\zeta \right\|^2 &\geq c_1(p, \tau) \sum_i |b_i|^2 \|p_i\|^2 - \frac{1}{2}c_1(p, \tau) \left(\sum_i |b_i|^2 \right) \min_i \|p_i\|^2 \\ &\geq \frac{1}{2}c_1(p, \tau) \sum_i |b_i|^2 \|p_i\|^2. \end{aligned}$$

Set, for example,

$$c_2(p, \tau) = \frac{1}{2}c_1(p, \tau), \tag{3-4}$$

where $c_1(p, \tau)$ is defined at the end of the previous proof (page 257).

The preceding relation, with $b = 0$, implies in particular that only the trivial linear combination of $p_1 - \xi_1, \dots, p_s - \xi_s$ is zero. \square

We next state a lemma that will enable us to choose a nice maximal free set of Γ_0 , the saturated module of a submodule Γ of $E(\overline{\mathbb{Q}})$ of finite rank, as defined in relation (2-1). There is nothing deep here, as we are working with finite-dimensional vector spaces.

Lemma 3.4 (Quasiorthonormality). *Let Γ_0 be the saturated module of Γ . Let s be the rank of Γ_0 . Then for any real $K > 0$, there exists a maximal free set $\gamma_1, \dots, \gamma_s$ of Γ_0 , with $\|\gamma_i\| \geq K$, such that for all $b_1, \dots, b_s \in \text{End } E$*

$$\left\| \sum_i b_i \gamma_i \right\|^2 \geq \frac{1}{9} \sum_i |b_i|^2 \|\gamma_i\|^2.$$

Proof. Recall that $\text{End } E$ is an order in an imaginary quadratic field k . Furthermore, the height norm $\|\cdot\|$ makes Γ_0 a hermitian $\text{End } E$ -module. Let Γ^{free} be a submodule of Γ_0 isomorphic to its free part. Then Γ^{free} is a k vector space of dimension s . Its tensor product with \mathbb{C} over k is a normed \mathbb{C} vector space of dimension s , and Γ^{free} is isomorphic to $\Gamma^{\text{free}} \otimes 1$. Using for instance the Gram–Schmidt orthonormalization algorithm in $\Gamma^{\text{free}} \otimes_k \mathbb{C}$, we can choose an orthonormal basis

$$v_i = g_i \otimes \rho_i.$$

So

$$\left\| \sum_i b_i v_i \right\|^2 = \sum_i |b_i|^2.$$

Decompose $\rho_i = r_{i1} + \tau r_{i2}$ for $1, \tau$ integral generators of $\text{End } E$ and $r_{ij} \in \mathbb{R}$. Choose $\delta = (2(1 + |\tau|) \max_i \|g_i\|)^{-1}$ and rationals q_{ij} such that $q_{ij} = r_{ij} + d_{ij}$ with $|d_{ij}| \leq \delta$ (use the density of the rationals).

Define

$$\gamma'_i = g_i \otimes (q_{i1} + \tau q_{i2}) = (q_{i1} + \tau q_{i2}) g_1 \otimes 1 \in \Gamma^{\text{free}} \otimes 1,$$

and

$$\delta_i = g_i \otimes (d_{i1} + \tau d_{i2}).$$

Then $v_i = \gamma'_i + \delta_i$, with $\|\delta_i\| \leq \|g_i\|(1 + |\tau|)\delta \leq \frac{1}{2}$. The triangle inequality gives

$$2 \left\| \sum_i b_i \gamma'_i \right\|^2 \geq \left\| \sum_i b_i v_i \right\|^2 - 2 \left\| \sum_i b_i \delta_i \right\|^2.$$

The orthonormality of v_i and $\|\delta_i\| \leq \frac{1}{2}$ implies that

$$2 \left\| \sum_i b_i \gamma'_i \right\|^2 \geq \sum_i |b_i|^2 - 2 \sum_i |b_i|^2 \frac{1}{4} = \frac{1}{2} \sum_i |b_i|^2.$$

Finally $\|\gamma'_i\| \leq \|v_i\| + \|\delta_i\| \leq \frac{3}{2}$, so

$$\left\| \sum_i b_i \gamma'_i \right\|^2 \geq \frac{1}{9} \sum_i |b_i|^2 \|\gamma'_i\|^2.$$

It is evident that for any integer n_0 the same relation holds:

$$\left\| \sum_i b_i n_0 \gamma'_i \right\|^2 \geq \frac{1}{9} \sum_i |b_i|^2 \|n_0 \gamma'_i\|^2.$$

Let n_0 be an integer such that $n_0 \geq 2K$. Note that

$$\|\gamma'_i\| \geq \|v_i\| - \|\delta_i\| \geq \frac{1}{2},$$

so

$$\|n_0 \gamma'_i\| \geq K.$$

Thus the maximal free set $\gamma_i = n_0 \gamma'_i$ satisfies the desired conditions. □

We cannot directly choose an orthonormal basis in Γ^{free} , because the norm has values in \mathbb{R} and not in \mathbb{Q} . What one can prove is that for any small positive real δ , there exists a maximal free set $\gamma_1, \dots, \gamma_s$ such that

$$\left\| \sum_i b_i \gamma_i \right\|^2 \geq \frac{(1-\delta)^2}{(1+\delta)^2} \sum_i |b_i|^2 \|\gamma_i\|^2.$$

4. Gauss-reduced morphisms

The aim of this section is to show that we can consider our union over Gauss-reduced algebraic subgroups, instead of over all algebraic subgroups.

Let B be an algebraic subgroup of E^g of codimension r . Then $B \subset \ker \phi_B$ for a surjective morphism $\phi_B : E^g \rightarrow E^r$. Conversely, we denote by B_ϕ the kernel of a surjection $\phi : E^g \rightarrow E^r$. Then B_ϕ is an algebraic subgroup of E^g of codimension r .

The matrices in $M_{r \times g}(\text{End } E)$ of the form

$$\phi = (aI_r | L) = \begin{pmatrix} a & \dots & 0 & a_{1,r+1} & \dots & a_{1,g} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & a & a_{r,r+1} & \dots & a_{r,g} \end{pmatrix},$$

with $H(\phi) = |a|$ and entries having no common factors (up to units), will play a key role in this work. For $r = g$, such a morphism becomes the identity, and L shall be forgotten. These matrices have three main advantages:

- The restriction of ϕ to the set $E^r \times \{0\}^{g-r}$ is just the multiplication by a .
- The image of $\mathbb{O}_\varepsilon \subset E^g$ under ϕ is contained in the image of $\mathbb{O}_{g\varepsilon} \cap (E^r \times \{0\}^{g-r})$. Similarly, the image of Γ_0^g under ϕ is contained in the image of $\Gamma_0^r \times \{0\}^{g-r}$.
- The matrix ϕ has small height compared to other matrices with same zero component of the kernel.

Definition 4.1 (Gauss-reduced morphisms). We say that a surjective morphism $\phi : E^g \rightarrow E^r$ is Gauss-reduced of rank r if the following conditions are satisfied:

- (i) There exists $a \in (\text{End } E)^*$ such that aI_r is a submatrix of ϕ , with I_r the r -identity matrix.
- (ii) $H(\phi) = |a|$.
- (iii) If there exists $f \in \text{End } E$ and $\phi' : E^g \rightarrow E^r$ such that $\phi = f\phi'$ then f is an isomorphism.

We say that an algebraic subgroup is Gauss-reduced if it is the kernel of a Gauss-reduced morphism.

Remark 4.2. If $\text{End } E \cong \mathbb{Z}$, condition (iii) simply says that the greatest common divisor of the entries of ϕ is 1 and $f = \pm 1$. Also when $\text{End } E \cong \mathbb{Z}$, we make condition (ii) more restrictive, requiring that $H(\phi) = a$, instead of $H(\phi) = |a|$; this assumption simplifies the notation. Obviously $B_\phi = B_{-\phi}$, so all lemmas below hold with this “up to units” definition of Gauss-reduced.

A morphism ϕ' given by a reordering of the rows of a morphism ϕ , has the same kernel as ϕ . Saying that aI_r is a submatrix of ϕ fixes one permutation of the rows of ϕ .

A reordering of the columns, on the other hand, corresponds to a permutation of the coordinates. Statements will be proved for Gauss-reduced morphisms of the form $\phi = (aI|L)$. For any other reordering of the columns the proofs are analogous. Since there are finitely many permutations of g columns, the finiteness statements will follow.

The following lemma is a simple useful trick to keep in mind.

Lemma 4.3. *Let $\phi : E^g \rightarrow E^r$ be Gauss-reduced of rank r .*

- (i) *For $\xi = (\xi_1, \dots, \xi_g) \in \mathbb{C}_\varepsilon$, there exists a point $\xi' = (\xi'', \{0\}^{g-r}) \in \mathbb{C}_{g\varepsilon}$ such that*

$$\phi(\xi) = \phi(\xi') = [a]\xi''.$$

- (ii) *For $y = (y_1, \dots, y_g) \in \Gamma_0^g$, there exists a point $y' = (y'', \{0\}^{g-r}) \in \Gamma_0^r \times \{0\}^{g-r}$ such that*

$$\phi(y) = \phi(y') = [a]y''.$$

Proof. Up to a reordering of the columns, the morphism ϕ has the form

$$\phi = \begin{pmatrix} a & \dots & 0 & a_{1,r+1} & \dots & a_{1,g} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \dots & a & a_{r,r+1} & \dots & a_{r,g} \end{pmatrix},$$

with $H(\phi) = |a|$.

(i) Consider a point $\xi'' \in E^r$ such that $[a]\xi'' = \phi(\xi)$. Since

$$\|\xi''\| = \frac{\|\phi(\xi)\|}{|a|} = \max_i \frac{\|\sum_j a_{ij}\xi_j\|}{|a|}$$

and $|a| = \max_{ij} |a_{ij}|$, we obtain

$$\|\xi''\| \leq g\varepsilon.$$

Define $\xi' = (\xi'', \{0\}^{g-r})$. Clearly

$$\phi(\xi') = [a]\xi'' = \phi(\xi).$$

(ii) Note that $\phi(y) \in \Gamma_0^r$. Since Γ_0 is a division group, the point y'' such that

$$[a]y'' = \phi(y),$$

belongs to Γ_0^r . Define $y' = (y'', \{0\}^{g-r})$. Then $\phi(y') = [a]y'' = \phi(y)$. □

In the next result we show that the zero components of B_ϕ , for ϕ ranging over all Gauss-reduced morphisms of rank r , are all possible abelian subvarieties of E^g of codimension r . This is proved using the classical Gauss algorithm, where the pivots have maximal absolute values.

Lemma 4.4. *Let $\psi : E^g \rightarrow E^r$ be a morphism of rank r . Then:*

(i) *For every $N \in \text{End } E^*$,*

$$B_{N\psi} \subset B_\psi + (E_{\text{Tor}}^r \times \{0\}^{g-r}).$$

(ii) *There exists a Gauss-reduced morphism $\phi : E^g \rightarrow E^r$ of rank r such that*

$$B_\psi \subset B_\phi + (E_{\text{Tor}}^r \times \{0\}^{g-r}).$$

Proof. (i) Let $b \in B_{N\psi}$. Then $N\psi(b) = 0$, so $\psi(b) = t$ with t a N -torsion point in E^r . Let ψ_1 be an invertible r -submatrix of ψ . Up to a reordering of the columns, we can suppose $\psi = (\psi_1 | \psi_2)$. Let t' be a torsion point in E^r such that $\psi_1(t') = t$. Then $\psi(b - (t', 0)) = 0$. Thus $b \in B_\psi + (E_{\text{Tor}}^r \times \{0\}^{g-r})$.

(ii) The Gauss algorithm gives an invertible integral r -matrix Δ such that, up to the order of the columns, $\Delta\psi$ is of the form

$$\Delta\psi = \begin{pmatrix} a & \dots & 0 & a_{1,r+1} & \dots & a_{1,g} \\ \vdots & & \vdots & & & \vdots \\ 0 & \dots & a & a_{r,r+1} & \dots & a_{r,g} \end{pmatrix},$$

with $H(\Delta\psi) = |a|$ (potentially there are common factors of the entries).

Let $b \in B_\psi$. Then $\psi(b) = 0$, so $\Delta\psi(x) = 0$. It follows that

$$B_\psi \subset B_{\Delta\psi}.$$

Take $N \in \text{End } E^*$ such that $N|\Delta\psi$ and such that if $f|(\Delta\psi/N)$ then f is a unit (if $\text{End } E \cong \mathbb{Z}$, then N is simply the greatest common divisor of the entries of $\Delta\psi$). Define

$$\phi = \Delta\psi/N.$$

Clearly ϕ is Gauss-reduced and $B_\psi \subset B_{\Delta\psi} = B_{N\phi}$. By part (i) of this lemma applied to $N\phi$, we conclude

$$B_\psi \subset B_\phi + (E_{\text{Tor}}^r \times \{0\}^{g-r}). \quad \square$$

Note that, in the previous lemma, a reordering of the columns of ψ or ϕ induces the same reordering of the coordinates of $E_{\text{Tor}}^r \times \{0\}^{g-r}$.

Taking intersections with the algebraic points of our curve, part (ii) of the previous lemma translates immediately as

Lemma 4.5. *Let $C \subset E^g$ be an algebraic curve (transverse or not). For any real $\varepsilon \geq 0$*

$$S_r(C, (\Gamma_0^g)_\varepsilon) = \bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk}(\phi)=r}} C(\overline{\mathbb{Q}}) \cap (B_\phi + (\Gamma_0^g)_\varepsilon).$$

Proof. By definition

$$S_r(C, (\Gamma_0^g)_\varepsilon) \supseteq \bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk}(\phi)=r}} C(\overline{\mathbb{Q}}) \cap (B_\psi + (\Gamma_0^g)_\varepsilon).$$

On the other hand, by Lemma 4.4(ii), we see that

$$C(\overline{\mathbb{Q}}) \cap (B_\psi + (\Gamma_0^g)_\varepsilon) \subset C(\overline{\mathbb{Q}}) \cap (B_\phi + (E_{\text{Tor}}^r \times \{0\}^{g-r}) + (\Gamma_0^g)_\varepsilon),$$

with ϕ Gauss-reduced of rank r . Moreover $(E_{\text{Tor}}^r \times \{0\}^{g-r}) \subset \mathcal{O}_\varepsilon \subset (\Gamma_0^g)_\varepsilon$. □

5. Relation between transverse and weak-transverse curves

We discuss here how we can associate to a couple (C, Γ) , where C is a transverse curve and Γ a subgroup of finite rank, a weak-transverse curve C' and vice versa. There are properties which are easier for C and others for C' . Using this association, we will try to gain advantages from both situations.

From transverse to weak-transverse. Let C be transverse in E^g . If Γ has rank 0, we set $C' = C$. If $\text{rk } \Gamma \geq 1$, consider the saturated module Γ_0 of rank s associated to Γ , as defined in relation (2-1). Let $\gamma_1, \dots, \gamma_s$ be a maximal free set of Γ_0 . We denote the associated point of E^s by

$$\gamma = (\gamma_1, \dots, \gamma_s).$$

We define

$$C' = C \times \gamma.$$

Since C is transverse and the γ_i are $\text{End } E$ -linearly independent, the curve C' is weak-transverse. For suppose to the contrary that C' were contained in an algebraic subgroup B_ϕ of codimension 1, with $\phi = (a_1, \dots, a_{g+s})$. Take a point $y_1 \in E$ such that $a_1 y_1 = \sum_{i=g+1}^{g+s} a_i \gamma_{i-g}$ and define $y = (y_1, 0, \dots, 0) \in E^g$. Then $C \subset B_{\phi_1} + y$ with $\phi_1 = (a_1, \dots, a_g)$, contradicting that C is transverse.

From weak-transverse to transverse. Let C' be weak-transverse in E^n . If C' is transverse, we set $C = C'$ and $\Gamma = 0$. Suppose that C' is not transverse. Let H_0 be the abelian subvariety of smallest dimension g such that $C' \subset H_0 + p$ for $p \in H_0^\perp(\overline{\mathbb{Q}})$ and let H_0^\perp be the orthogonal complement of H_0 with respect to the canonical polarization. Then E^n is isogenous to $H_0 \times H_0^\perp$. Furthermore H_0 is isogenous to E^g and H_0^\perp is isogenous to E^s , where $s = n - g$. Let j_0, j_1 and j_2 be such isogenies. We fix the isogeny

$$j = (j_1 \times j_2) \circ j_0 : E^n \rightarrow H_0 \times H_0^\perp \rightarrow E^g \times E^s,$$

which sends H_0 to $E^g \times 0$ and H_0^\perp to $0 \times E^s$. Then

$$j(C') \subset (E^g \times 0) + j(p),$$

with $j(p) = (0, \dots, 0, p_1, \dots, p_s)$.

We consider the natural projection on the first g coordinates

$$\pi : E^g \times E^s \rightarrow E^g, \quad j(C') \mapsto \pi(j(C')).$$

We define

$$C = \pi(j(C')) \text{ and } \Gamma = \langle p_1, \dots, p_s \rangle^g.$$

Since H_0 has minimal dimension, the curve C is transverse in E^g .

Note that

$$j(C') = C \times (p_1, \dots, p_s).$$

In addition $j(C')$ is weak-transverse, because C' is. Therefore, $\langle p_1, \dots, p_s \rangle$ has rank s ; indeed if $\sum_{i=1}^s a_i p_i = 0$, then $j(C') \subset B_\phi$ for $\phi = (\{0\}^g, a_1, \dots, a_s)$.

Weak-transverse up to an isogeny. Statements on boundedness of heights or finiteness of sets are invariant under an isogeny of the ambient variety. Namely, given an isogeny j of E^g , Theorems 1.2 and 1.5 hold for a curve if and only if they hold for its image via j . Thus, the previous discussion shows that without loss of generality, we can assume that a weak-transverse curve C' in E^n is of the form

$$C' = C \times p,$$

where

- (i) C is transverse in E^g ,
- (ii) $p = (p_1, \dots, p_s) \in E^s$ is such that the module $\langle p_1, \dots, p_s \rangle$ has rank s , and
- (iii) $n = g + s$.

This simplifies the setting for weak-transverse curves.

Implying Mordell–Lang plus Bogomolov for curves. Note that

$$S_g(C, \mathbb{O}_\varepsilon) = C \cap \mathbb{O}_\varepsilon \quad \text{and} \quad S_g(C(\Gamma_0^g)_\varepsilon) = C \cap (\Gamma_0^g)_\varepsilon.$$

Moreover $S_2(C, \cdot) \supset S_g(C, \cdot)$. This immediately shows that Conjecture 1.1 implies the Bogomolov Theorem for weak-transverse curves and the Mordell–Lang and Bogomolov Theorems for transverse curves. We want to show that Conjecture 1.1 implies these theorems for all curves of genus ≥ 2 .

In E^g a curve of genus 2 is a translate of an elliptic curve isogenous to E . If C is not transverse, then $C \subsetneq H_0 + p$ with H_0 an algebraic subgroup of minimal dimension satisfying such inclusion. Let $\pi : E^g \rightarrow E^g/H_0^\perp$ be the natural projection and let $\psi : E^g/H_0^\perp \rightarrow E^k$ be an isogeny. Then $\|\psi\pi(x)\| \ll \|x\|$. In E^k , consider the transverse curve $C' = \psi\pi(C - p)$ and $\Gamma' = \psi\pi(\Gamma, \Gamma_p)$. Note that $\psi\pi(\text{Tor}_{E^g}) \subset \text{Tor}_{E^k}$. Then

$$S_g(C, (\Gamma_0^g)_\varepsilon) \subset \pi|_C^{-1} S_k(C', (\Gamma_0^g)_{\varepsilon'}).$$

The map $\pi|_C^{-1}$ has finite fiber. Applying Conjecture 1.1 to $C' \subset E^k$ we deduce that $S_g(C, (\Gamma_0^g)_\varepsilon)$ is finite.

Note that such a proof works only for $S_g(C, \cdot)$, because the projection $\psi\pi(B) \subset E^k$ of an algebraic subgroup B of E^g of codimension r may not have codimension r in E^k . It could even be all of E^k .

6. Quasispecial morphisms

Just as Gauss-reduced morphisms play a key role for transverse curves, quasispecial morphisms play a key role for weak-transverse curves. In particular, for small ε , quasispecial morphisms are enough to cover the whole of $S_r(C \times p, \mathbb{O}_\varepsilon)$; this is Lemma 6.2 below.

To motivate quasispecialness, suppose that $C \times p$ is weak-transverse in E^{g+s} with C transverse in E^g . A point of $C \times p$ is of the form (x, p) . The last s -coordinates are constant and just the x varies. This two parts must be treated differently. Saying that a morphism $\tilde{\phi} = (\phi|\phi')$ is quasispecial ensures that the rank of ϕ is maximal (note that ϕ acts on x). In particular, this allows us to apply the Gauss algorithm on the first g columns of $\tilde{\phi}$.

Definition 6.1 (Quasispecial morphism). A surjective morphism $\tilde{\phi} : E^{g+s} \rightarrow E^r$ is quasispecial if there exist $N \in \text{End } E^*$, morphisms $\phi : E^g \rightarrow E^r$ and $\phi' : E^s \rightarrow E^r$ such that

- (i) $\tilde{\phi} = (N\phi|\phi')$,
- (ii) $\phi = (aI_r|L)$ is Gauss-reduced of rank r , and
- (iii) if there exists $f \in \text{End } E$ and $\tilde{\phi}' : E^{s+s} \rightarrow E^r$ such that $\tilde{\phi} = f\tilde{\phi}'$, then f is an isomorphism.

We do not require that $\tilde{\phi}$ be Gauss-reduced; the fact is that $H(\phi')$ might not be controlled by $NH(\phi)$. This extra condition will define special morphisms (see Definition 10.1).

Lemma 6.2. *Let $C \times p$ be weak-transverse in E^{s+s} with C transverse in E^s . Then, there exists $\varepsilon > 0$ such that*

$$S_r(C \times p, \mathbb{O}_\varepsilon) \subset \bigcup_{\substack{\tilde{\phi} \text{ quasispecial} \\ \text{rk } \tilde{\phi}=r}} (C(\overline{\mathbb{Q}}) \times p) \cap (B_{\tilde{\phi}} + \mathbb{O}_\varepsilon).$$

We can choose $\varepsilon \leq \varepsilon_0(p, \tau)$, where $\varepsilon_0(p, \tau)$ is as in Proposition 3.3.

Proof. Take $(x, p) \in S_r(C \times p, \mathbb{O}_\varepsilon)$. Then $(x, p) \in (C(\overline{\mathbb{Q}}) \times p) \cap (B_{\tilde{\psi}} + \mathbb{O}_\varepsilon)$ for a morphism $\tilde{\psi} = (\psi|\psi') : E^{s+s} \rightarrow E^r$ of rank r . In other words, there exists a point $(\xi, \xi') \in \mathbb{O}_\varepsilon$ such that

$$\tilde{\psi}((x, p) + (\xi, \xi')) = 0.$$

First, we show that ψ has rank r . Suppose, on the contrary, that the rank of ψ were less than r . Then a linear combination of the rows of ψ is trivial, namely

$$(\lambda_1, \dots, \lambda_r)\psi = 0.$$

Since $\psi(x + \xi) + \psi'(p + \xi') = 0$, the same linear combination of the r coordinates of $\psi'(p + \xi')$ is trivial, namely

$$(\lambda_1, \dots, \lambda_r)\psi'(p + \xi') = 0.$$

Apply Proposition 3.3 with $(b_1, \dots, b_s) = (\lambda_1, \dots, \lambda_r)\psi'$, $(\xi_1, \dots, \xi_s) = -\xi'$, $\zeta = 0$ and $b = 0$. This implies that, if $\varepsilon \leq \varepsilon_0(p, \tau)$, then the points $p_1 + \xi'_1, \dots, p_s + \xi'_s$ are linearly independent. It follows that

$$(\lambda_1, \dots, \lambda_r)\psi' = 0.$$

Hence, the rank of $\tilde{\psi}$ would be less than r , contradicting the fact that the rank of $\tilde{\psi}$ is r .

Since the rank of ψ is r , we can apply the Gauss algorithm using pivots in ψ of maximal absolute values in ψ (clearly we cannot require that they have maximal absolute values in $\tilde{\psi}$). Let Δ be an invertible matrix, given by the Gauss algorithm, such that $\Delta\tilde{\psi} = (\phi_1|\phi_2)$ with fI_r a submatrix of ϕ_1 .

We next get rid of possible common factors. Take $N_1, n_1 \in \text{End } E^*$ such that $N_1 | \phi_1$ and $n_1 | \Delta \tilde{\psi}$. Further suppose that, if $f | (\phi_1/N_1)$ or $f | (\Delta \tilde{\psi}/n_1)$ then f is a unit of $\text{End } E$ (if $\text{End } E \cong \mathbb{Z}$, then N_1 is the greatest common divisor of the entries of ϕ_1 and n_1 the greatest common divisor of the entries of $\Delta \tilde{\psi}$). Then

$$\Delta \tilde{\psi} = n_1(N\phi | \phi')$$

with $N = N_1/n_1$, $\phi = \phi_1/N_1$ and $\phi' = \phi_2/n_1$. We define

$$\tilde{\phi} = (N\phi | \phi').$$

Clearly $\tilde{\phi}$ is quasispecial. In addition

$$B_{\tilde{\psi}} \subset B_{\Delta \tilde{\psi}} = B_{n_1 \tilde{\phi}}.$$

By Lemma 4.4(i), with $\psi = \tilde{\phi}$ and $N = n_1$, we deduce that

$$B_{\tilde{\psi}} \subset B_{\tilde{\phi}} + E_{\text{Tor}}^r \times \{0\}^{g+s-r}.$$

Since $(x, p) \in B_{\tilde{\psi}} + \mathbb{O}_\varepsilon$, we obtain $(x, p) \in B_{\tilde{\phi}} + \mathbb{O}_\varepsilon$ with $\tilde{\phi}$ quasispecial. □

7. Estimates for the height: the proof of Theorem 1.2

As mentioned, Theorem 1.2(ii) is part of Theorem 1.5 in [Rémond and Viada 2003]. In this section, we adapt the proof given there to part (i) of Theorem 1.2.

In view of Section 5, we can assume, without loss of generality, that a weak-transverse curve C' in E^n has the form

$$C' = C \times p,$$

where C and p satisfy conditions (i)–(iii) on page 265.

Definition 7.1. Let p be a point in E^s and ε a nonnegative real. We define G_p^ε as the set of points $\theta \in E^2$ for which there exist a matrix $A \in M_{2,s}(\text{End } E)$, an element $a \in \text{End } E$ with $0 < |a| \leq H(A)$, points $\xi \in E^s$ and $\zeta \in E^2$ of norm at most ε such that

$$[a]\theta = A(p + \xi) + [a]\zeta.$$

We identify G_p^ε with the subset $G_p^\varepsilon \times \{0\}^{g-2}$ of E^g .

Recall that Γ_p is the saturated module of the coordinates of p .

Now we embed $S_2(C \times p, \mathbb{O}_\varepsilon)$ in two sets related to the transverse curve C . We then use the Vojta inequality on these new sets.

Lemma 7.2. *The natural projection on the first g coordinates,*

$$E^g \times E^s \rightarrow E^g, \quad (x, y) \mapsto x,$$

defines an injection

$$S_2(C \times p, \mathbb{C}_{\varepsilon/2gs}) \hookrightarrow S_2(C, (\Gamma_p^g)_\varepsilon) \cup \bigcup_{\substack{\phi: E^g \rightarrow E^2 \\ \text{Gauss-reduced}}} C(\overline{\mathbb{Q}}) \cap B_\phi + G_p^\varepsilon.$$

Proof. Let $(x, p) \in S_2(C \times p, \mathbb{C}_{\varepsilon/2gs})$. By Lemma 6.2, $(x, p) \in B_{\tilde{\phi}} + \mathbb{C}_{\varepsilon/2gs}$, with $\tilde{\phi} = (N\phi|\phi') : E^{g+s} \rightarrow E^2$ quasispecial of rank 2. Hence

$$\tilde{\phi}((x, p) + (\xi, \xi')) = 0,$$

for $(\xi, \xi') \in \mathbb{C}_{\varepsilon/2gs}$. We can write the equality as

$$N\phi(x) + N\phi(\xi) + \phi'(p + \xi') = 0.$$

By the definition of quasispecialness ϕ is Gauss-reduced, so

$$\phi = (aI_2|L).$$

By Lemma 4.3(i) applied to ϕ and ξ , we can assume that

$$\xi = (\xi_1, \xi_2, 0, \dots, 0) \in \mathbb{C}_{\varepsilon/2s}.$$

Suppose first that $NH(\phi) \geq H(\tilde{\phi})$. Let ζ be a point in $E^2 \times \{0\}^{g-2}$ such that

$$N[a]\zeta = (\phi'(\xi'), 0, \dots, 0).$$

Then

$$\|\zeta\| = \frac{\|\phi'(\xi')\|}{NH(\phi)} \leq \frac{\varepsilon}{2}.$$

Let y be a point in $E^2 \times \{0\}^{g-2}$ such that

$$N[a]y = (\phi'(p), 0, \dots, 0).$$

Since Γ_p is saturated, $y \in \Gamma_p^2 \times \{0\}^{g-2}$. Then

$$N\phi(x + \xi + \zeta + y) = 0$$

with $y + \xi + \zeta \in \Gamma_p^g + \mathbb{C}_\varepsilon$. So

$$x \in S_2(C, (\Gamma_p^g)_\varepsilon).$$

Now suppose that $NH(\phi) < H(\tilde{\phi})$ or, equivalently, $NH(\phi) < H(\phi')$. Let θ' be a point in E^2 such that

$$N[a]\theta' = \phi'(p + \xi') + N[a] \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix},$$

and $\theta = (\theta', \{0\}^{g-r})$. Then $\theta \in G_p^\varepsilon$. Moreover

$$\begin{aligned} N\phi(x + \theta) &= N\phi(x) + N\phi(\theta) \\ &= N\phi(x) + N[a]\theta' = N\phi(x) + \phi'(p + \xi') + N[a] \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} \\ &= N\phi(x) + N\phi(\xi) + \phi'(p + \xi') \\ &= \tilde{\phi}((x, p) + (\xi, \xi')) = 0. \end{aligned}$$

Thus $x \in B_{N\phi} + G_p^\varepsilon$, and by Lemma 4.4(i),

$$x \in B_\phi + (E_{\text{Tor}}^2 \times \{0\}^{g-2}) + G_p^\varepsilon.$$

Note that $G_p^\varepsilon + (E_{\text{Tor}}^2 \times \{0\}^{g-2}) \subset G_p^\varepsilon$. Hence,

$$x \in C(\overline{\mathbb{Q}}) \cap B_\phi + G_p^\varepsilon. \quad \square$$

Lemma 7.3 (Counterpart to [R mond and Viada 2003, Lemma 3.2]). *For $\phi : E^g \rightarrow E^2$ Gauss-reduced of rank 2, we have the set inclusion*

$$(B_\phi + G_p^\varepsilon) \subset \{P + \theta : P \in B_\phi, \theta \in G_p^\varepsilon \text{ and } \max(\|\theta\|, \|P\|) \leq 2g\|P + \theta\|\}.$$

Proof. Take $x \in (B_\phi + G_p^\varepsilon)$ with $\phi = (aI_r|L)$ Gauss-reduced of rank 2. Then $x = P + \theta$ with $P \in B_\phi$ and $\theta \in G_p^\varepsilon$ and $\phi(x - \theta) = 0$. By definition $G_p^\varepsilon \subset E^2 \times \{0\}^{g-2}$, so $\phi(\theta) = [a]\theta$. Then

$$\|\theta\| = \frac{\|\phi(\theta)\|}{H(\phi)} = \frac{\|\phi(x)\|}{H(\phi)} \leq g\|x\|.$$

So

$$\|P\| = \|x - \theta\| \leq (g + 1)\|x\| = (g + 1)\|P + \theta\|. \quad \square$$

Lemma 3.3(1) of [R mond and Viada 2003] is a statement on the morphism; therefore it holds with no need for any remarks.

Lemma 7.4 (Counterpart to [R mond and Viada 2003, Lemma 3.3(2)]). *There exists an effective $\varepsilon_2 > 0$ such that, for all $\varepsilon \leq \varepsilon_2$, any sequence of elements in G_p^ε admits a subsequence in which every two elements θ, θ' satisfy*

$$\left\| \frac{\theta}{\|\theta\|} - \frac{\theta'}{\|\theta'\|} \right\| \leq \frac{1}{16gc_1},$$

where c_1 depends on C and is as defined in [R mond and Viada 2003, Proposition 2.1].

Proof. We decompose two elements θ and θ' in a given sequence of elements of G_p^ε as

$$[a]\theta = A(p + \xi) + [a]\zeta, \quad [a']\theta' = A'(p + \xi') + a'\zeta',$$

with $A, A' \in M_{2,s}(\text{End } E)$ and $0 < |a| \leq H(A), 0 < |a'| \leq H(A')$. Define y and y' such that

$$[a]y = A(p) \quad \text{and} \quad [a']y' = A'(p).$$

Since the sphere of radius 1 is compact in $(\langle p_1, \dots, p_s \rangle \times \langle p_1, \dots, p_s \rangle) \otimes \mathbb{R}$, we can extract a subsequence such that, for any two elements y and y' ,

$$\left\| \frac{y}{\|y\|} - \frac{y'}{\|y'\|} \right\| \leq \frac{1}{48gc_1}.$$

Note that

$$\left\| \frac{\theta}{\|\theta\|} - \frac{y}{\|y\|} \right\| = \left\| \frac{A(\xi) + [a]\zeta}{A(p + \xi) + [a]\zeta} \right\|$$

and

$$\left\| \frac{y}{\|\theta\|} - \frac{y}{\|y\|} \right\| = \left| \frac{\|A(p)\| - \|A(p + \xi) + [a]\zeta\|}{\|A(p + \xi) + [a]\zeta\|} \right| \leq \left\| \frac{A(\xi) + [a]\zeta}{A(p + \xi) + [a]\zeta} \right\|,$$

and the same relations for primed variables. We deduce that

$$\begin{aligned} \left\| \frac{\theta}{\|\theta\|} - \frac{\theta'}{\|\theta'\|} \right\| &\leq \left\| \frac{y}{\|y\|} - \frac{y'}{\|y'\|} \right\| + \left\| \frac{y}{\|\theta\|} - \frac{y}{\|y\|} \right\| + \left\| \frac{y'}{\|\theta'\|} - \frac{y'}{\|y'\|} \right\| \\ &\quad + \left\| \frac{\theta}{\|\theta\|} - \frac{y}{\|\theta\|} \right\| + \left\| \frac{\theta'}{\|\theta'\|} - \frac{y'}{\|\theta'\|} \right\| \\ &\leq \left\| \frac{y}{\|y\|} - \frac{y'}{\|y'\|} \right\| + 2 \left\| \frac{A(\xi) + [a]\zeta}{A(p + \xi) + [a]\zeta} \right\| + 2 \left\| \frac{A'(\xi') + [a']\zeta'}{A'(p + \xi') + [a']\zeta'} \right\|. \end{aligned}$$

Choose

$$\varepsilon \leq \varepsilon_2 = \min(\varepsilon_0(p, \tau), \varepsilon'_0(p, \tau)), \tag{7-1}$$

where $\varepsilon_0(p, \tau)$ is defined in (3-3), $c_2(p, \tau)$ is defined in (3-4) and

$$\varepsilon'_0(p, \tau) = \frac{c_2(p, \tau)^{1/2} \min \|p_i\|}{96(s + 1)c_1}.$$

Note that $\|A(p + \xi) + [a]\zeta\| = \|A_k(p + \xi) + a\zeta_k\|$ for $k = 1$ or 2 and $A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}$. Proposition 3.3 applied with $b_1, \dots, b_s = A_k, \xi = -\xi, \zeta = -\zeta_k$ and $b = a$, implies

$$\|A(p + \xi) + [a]\zeta\| \geq H(A)c_2(p, \tau)^{1/2} \min \|p_i\|$$

(same relation with $'$).

It follows that

$$\begin{aligned} & \left\| \frac{\theta}{\|\theta\|} - \frac{\theta'}{\|\theta'\|} \right\| \\ & \leq \frac{1}{48gc_1} + \varepsilon \frac{2H(A)(s+1)}{H(A)c_2(p, \tau)^{1/2} \min \|p_i\|} + \varepsilon \frac{2H(A')(s+1)}{H(A')c_2(p, \tau)^{1/2} \min \|p_i\|} \\ & \leq \frac{1}{48gc_1} + \frac{1}{48gc_1} + \frac{1}{48gc_1}, \end{aligned}$$

where in the last inequality we use $\varepsilon \leq \varepsilon'_0(p, \tau)$. □

We are ready to conclude.

Proof of Theorem 1.2(i). In view of Lemma 7.2, we shall prove that there exists $\varepsilon > 0$ such that $S_2(C, (\Gamma_p^g)_\varepsilon)$ and $\bigcup_{\substack{\phi: E^g \rightarrow E^2 \\ \text{Gauss-reduced}}} C(\overline{\mathbb{Q}}) \cap B_\phi + G_p^\varepsilon$ have bounded height.

By Theorem 1.2(ii), there exists $\varepsilon_1 > 0$ such that for $\varepsilon \leq \varepsilon_1$, the first set has bounded height.

It remains to show that there exists $\varepsilon_2 > 0$ such that, for $\varepsilon \leq \varepsilon_2$, the second set has bounded height. The proof follows, step by step, the proof of [Rémond and Viada 2003, Theorem 1.5]. In view of Lemma 7.3 and 7.4, all conditions for the proof of that theorem are satisfied. The proof is then exactly equal to the one in [Rémond and Viada 2003, p. 1927–1928]. □

Remark 7.5. In Theorem 1.5 of [Rémond and Viada 2003] we showed that for $\varepsilon_1 = 1/(2^g c_1)$, the set $S_2(C, \Gamma_{\varepsilon_1})$ has bounded height. The constant c_1 depends on the invariants of the curve C . This constant is defined in Proposition 2.1 of the same reference and it is effective. On the other hand, the height of $S_2(C, \Gamma_{\varepsilon_1})$ is bounded by a constant which is not known to be effective, unless Γ has rank 0.

For $C \times p$, we have shown that for

$$\varepsilon'_2 = \frac{\min(1, c_2(p, \tau)) \min \|p_i\|^2}{2^8 g (s+1)^2 \max \|p_i\| c_1}$$

the set $S_2(C \times p, \mathbb{O}_{\varepsilon'_2})$ has bounded height; see relation (7-1) and Lemma 7.2. As in the previous case, the height of $S_2(C \times p, \mathbb{O}_{\varepsilon'_2})$ is bounded by a constant which, in general, is not known to be effective.

8. Summary of notation

We stop to recapitulate and fix the notations for the rest of the article.

For simplicity, we assume that $\text{End } E \cong \mathbb{Z}$. In this case the saturated module of a group coincides with its division group. According to Remark 4.2, we use $H(\phi) = a$ in the definition of a Gauss-reduced morphism and $N \in \mathbb{N}^*$ in the definition of quasispecialness.

- Let E be an elliptic curve without CM over $\overline{\mathbb{Q}}$.
- Let C be a transverse curve in E^g over $\overline{\mathbb{Q}}$.
- Let

$$\phi = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_r \end{pmatrix} = \begin{pmatrix} a & \dots & 0 & L_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & a & L_r \end{pmatrix}$$

be a Gauss-reduced morphism of rank $1 \leq r \leq g$, with $L_i \in \mathbb{Z}^{g-r}$ and $H(\phi) = a$.

- Let Γ be a subgroup of finite rank of $E^g(\overline{\mathbb{Q}})$.
- Let Γ_0 be the division group of Γ and s its rank (the definition is given in relation (2-1)).
- Choose $\varepsilon_1 > 0$ so that $S_2(C, (\Gamma_0^g)_{\varepsilon_1})$ has bounded height; the definition is consistent in view of Theorem 1.2 (ii).
- Let K_1 be the norm of $S_2(C, (\Gamma_0^g)_{\varepsilon_1})$.
- Let $\gamma = (\gamma_1, \dots, \gamma_s)$ be a point of $E^s(\overline{\mathbb{Q}})$ such that $\gamma_1, \dots, \gamma_s$ is a maximal free set of Γ_0 satisfying the conditions of Lemma 3.4 with $K = 3gK_1$. Namely, for all integers b_i

$$\frac{1}{9} \sum_i |b_i|^2 \|\gamma_i\|^2 \leq \left\| \sum_i b_i \gamma_i \right\|^2 \tag{8-1}$$

and

$$\min_i \|\gamma_i\| \geq 3gK_1. \tag{8-2}$$

- Let $C \times \gamma$ be the associated weak-transverse curve in E^{g+s} .
- Let $\tilde{\phi} = (N\phi|\phi') : E^{g+s} \rightarrow E^r$ be a quasispecial morphism with $N \in \mathbb{N}^*$.
- Choose $\varepsilon_2 > 0$ so that $S_2(C \times \gamma, \mathbb{O}_{\varepsilon_2})$ has bounded height; this definition is consistent in view of Theorem 1.2(i).
- Let K_2 be the norm of $S_2(C \times \gamma, \mathbb{O}_{\varepsilon_2})$.
- Let $p = (p_1, \dots, p_s) \in E^s$ be a point such that the rank of $\langle p_1, \dots, p_s \rangle$ is s .
- Let Γ_p be the division group of $\langle p_1, \dots, p_s \rangle$ (in short the division group of p).
- Let c_p and ε_p be the constants $(c_2(p, \tau))^{1/2}$ and $\varepsilon_0(p, \tau)$ defined in Proposition 3.3 for the point p and $\tau = 1$ (please note the square root in c_p).
- Let $C \times p$ be the associated weak-transverse curve in E^{g+s} .
- Choose $\varepsilon_3 > 0$ so that $S_2(C \times p, \mathbb{O}_{\varepsilon_3})$ has bounded height; the definition is consistent in view of Theorem 1.2 (i).
- Let K_3 be the norm of $S_2(C \times p, \mathbb{O}_{\varepsilon_3})$.

9. Equivalence of the strong statements: the proof of Theorem 1.3

The following theorem implies Theorem 1.3 immediately; in addition it gives explicit inclusions. Once more, we emphasize that we need to assume that $S_r(C \times p, \mathbb{O}_\varepsilon)$ has bounded height in order to embed it in a set of the type $S_r(C, \Gamma_{\varepsilon'})$. Therefore we assume $r \geq 2$ and $\varepsilon \leq \varepsilon_3$ in part (ii).

Theorem 9.1. *Let $\varepsilon \geq 0$.*

(i) *The map $x \rightarrow (x, \gamma)$ defines an injection*

$$S_r(C, \Gamma_\varepsilon) \hookrightarrow S_r(C \times \gamma, \mathbb{O}_\varepsilon).$$

(Recall that γ is a maximal free set of Γ_0 .)

(ii) *For $2 \leq r$ and $\varepsilon \leq \min(\varepsilon_p, \varepsilon_3)$, the map $(x, p) \rightarrow x$ defines an injection*

$$S_r(C \times p, \mathbb{O}_\varepsilon) \hookrightarrow S_r(C, (\Gamma_p^g)_{\varepsilon K_4}),$$

where $K_4 = (g + s) \max\left(1, \frac{g(K_3 + \varepsilon)}{c_p \min_i \|p_i\|}\right)$. (Recall that Γ_p is the division group of p .)

Proof. (i) Let $x \in S_r(C, \Gamma_\varepsilon)$. There exists a surjective $\phi : E^g \rightarrow E^r$ and points $y \in \Gamma$ and $\xi \in \mathbb{O}_\varepsilon$ such that

$$\phi(x + y + \xi) = 0.$$

Since $\gamma = (\gamma_1, \dots, \gamma_s)$ is a maximal free set of Γ_0 , there exists a positive integer N and a matrix $G \in M_{r,s}(\mathbb{Z})$ such that

$$[N]y = G\gamma.$$

We define

$$\tilde{\phi} = (N\phi| \phi G).$$

Then $\tilde{\phi}((x, \gamma) + (\xi, 0)) = N\phi(x + \xi) + \phi G(\gamma) = N\phi(x + \xi + y) = 0$, so

$$(x, \gamma) \in S_r(C \times \gamma, \mathbb{O}_\varepsilon).$$

(ii) Take $(x, p) \in S_r(C \times p, \mathbb{O}_\varepsilon)$. Thanks to Lemma 6.2, the assumption $\varepsilon \leq \varepsilon_p$ implies

$$(x, p) \in (B_{\tilde{\phi}} + \mathbb{O}_\varepsilon)$$

with $\tilde{\phi} = (N\phi| \phi')$ quasispecial. Hence

$$\tilde{\phi}((x, p) + (\xi, \xi')) = 0$$

for $(\xi, \xi') \in \mathbb{O}_\varepsilon$. Equivalently,

$$N\phi(x + \xi) = -\phi'(p + \xi'). \tag{9-1}$$

By the definition of quasispecialness, ϕ is Gauss-reduced of rank r . Let

$$\phi = \begin{pmatrix} \phi_1 \\ \vdots \\ \phi_r \end{pmatrix} = \begin{pmatrix} a & \dots & 0 & L_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & a & L_r \end{pmatrix},$$

with $L_i \in \mathbb{Z}^{g-r}$ and $H(\phi) = a$.

Since Γ_p is the division group of p , the point y' defined by

$$N[a]y' = \phi'(p)$$

belongs to Γ_p^r .

Let ζ' be a point of E^r such that

$$N[a]\zeta' = \tilde{\phi}(\xi, \xi').$$

We define

$$y = (y', 0, \dots, 0) \in \Gamma_p^r \times \{0\}^{g-r}, \quad \zeta = (\zeta', 0, \dots, 0) \in E^r \times \{0\}^{g-r}.$$

We have

$$N\phi(y) = N[a]y' = \phi'(p)N\phi(\zeta) = N[a]\zeta' = \tilde{\phi}(\xi, \xi').$$

It follows that

$$N\phi(x + y + \zeta) = N\phi(x) + \phi'(p) + \tilde{\phi}(\xi, \xi') = \tilde{\phi}((x, p) + (\xi, \xi')) = 0.$$

Thus

$$x \in C(\overline{\mathbb{Q}}) \cap (B_{N\phi} + \Gamma_p^g + \mathbb{O}_{\|\zeta\|}).$$

In order to finish the proof, we shall prove

$$\|\zeta\| \leq \varepsilon K_4.$$

By the definition of ζ we see that

$$\begin{aligned} \|\zeta\| = \|\zeta'\| &= \frac{\|\tilde{\phi}(\xi, \xi')\|}{Na} \leq (g + s) \frac{\max(H(\phi'), Na)}{Na} \|(\xi, \xi')\| \\ &\leq (g + s) \frac{\max(H(\phi'), Na)}{Na} \varepsilon. \end{aligned}$$

We claim that

$$\frac{\max(H(\phi'), Na)}{Na} \leq \frac{K_4}{g + s}.$$

Let $\phi' = (b_{ij})$. We shall prove that $H(\phi') = \max_{ij} |b_{ij}| \leq \frac{K_4}{g+s} Na$. Let $|b_{kl}| = H(\phi')$. Consider the k -th row of the system (9-1)

$$N\phi_k(x) + N\phi_k(\xi) = -\sum_j b_{kj}(p_j + \xi'_j).$$

The triangle inequality gives

$$\frac{\|\phi_k(x)\|}{a} + \frac{\|\phi_k(\xi)\|}{a} \geq \frac{\|\sum_j b_{kj}(p_j + \xi'_j)\|}{Na}. \tag{9-2}$$

Since $\varepsilon \leq \varepsilon_3$ and $r \geq 2$, we have $(x, p) \in S_2(C \times p, \mathbb{O}_{\varepsilon_3})$, which has norm K_3 . Hence

$$\|x\| \leq \|(x, p)\| \leq K_3.$$

Since $a = H(\phi)$, we see that

$$\frac{\|\phi_k(x)\|}{a} \leq (g-r+1)K_3 \quad \text{and} \quad \frac{\|\phi_k(\xi)\|}{a} \leq (g-r+1)\varepsilon.$$

Substituting in (9-2),

$$(g-r+1)(K_3 + \varepsilon) \geq \frac{\|\sum_j b_{kj}(p_j + \xi'_j)\|}{Na}.$$

Recall that $\varepsilon \leq \varepsilon_p$. Hence, Proposition 3.3 with $(b_1, \dots, b_s) = (b_{k1}, \dots, b_{ks})$, $(\xi_1, \dots, \xi_s) = -\xi'$ and $\zeta = 0$, implies that

$$(g-r+1)(K_3 + \varepsilon) \geq \frac{1}{Na} \left(c_p^2 \sum_j |b_{kj}|^2 \|p_j\|^2 \right)^{1/2} \geq \frac{c_p H(\phi')}{Na} \min_i \|p_i\|.$$

Whence

$$H(\phi') \leq \frac{K_4}{g+s} Na. \tag{□}$$

The inclusion in Theorem 9.1(ii) has been proved only for a set $S_r(C \times p, \mathbb{O}_\varepsilon)$ known to have bounded height. If the norm K_3 of $S_r(C \times p, \mathbb{O}_\varepsilon)$ goes to infinity, the set $(\Gamma_p^g)_{\varepsilon K_4}$ tends to be the whole of E^g .

Remark 9.2. We would like to show that our Theorems 1.2 and 1.5 are optimal. Take $\Gamma = \langle (y_1, 0, \dots, 0) \rangle$, where y_1 is a nontorsion point in $E(\overline{\mathbb{Q}})$. Since C is transverse, the projection π_1 of $C(\overline{\mathbb{Q}})$ on the first factor $E(\overline{\mathbb{Q}})$ is surjective. Let $x_n \in C(\overline{\mathbb{Q}})$ such that $\pi_1(x_n) = ny_1$. So $x_n - n(y_1, 0, \dots, 0)$ has first coordinate zero, and belongs to the algebraic subgroup $0 \times E^{g-1}$. Then, for all $n \in \mathbb{N}$ we have

$$x_n \in B_{\phi=(1,0,\dots,0)} + \Gamma.$$

This shows that $x_n \in S_1(C, \Gamma)$, so $S_1(C, \Gamma)$ does not have bounded height. By Theorem 9.1(i), neither does $S_1(C \times y_1)$.

10. Special morphisms and an important inclusion

We can actually show a stronger inclusion than the one in Theorem 9.1(i). The set $S_r(C, \Gamma_\varepsilon)$ can be included in a subset of $S_r(C \times \gamma, \mathbb{O}_\varepsilon)$, namely the subset defined by special morphisms.

Definition 10.1 (Special morphisms). A surjective morphism $\tilde{\phi} : E^{g+s} \rightarrow E^r$ is special if $\tilde{\phi} = (N\phi|\phi')$ is quasispecial and satisfies the further condition

$$H(\tilde{\phi}) = NH(\phi).$$

Equivalently, $\tilde{\phi}$ is special if and only if

- (i) $\tilde{\phi}$ is Gauss-reduced, and
- (ii) $H(\tilde{\phi})I_r$ is a submatrix of the matrix consisting of the first g columns of $\tilde{\phi}$.

Proof of the equivalence of the two definitions. That the first definition implies the second is clear. For the converse, take the decomposition $\tilde{\phi} = (A|\phi')$, with $A \in M_{r \times g}(\mathbb{Z})$ and $\phi' \in M_{r \times s}(\mathbb{Z})$. Let N be the greatest common divisor of the entries of A . Define $\phi = A/N$ and $a = H(\tilde{\phi})/N$. Then $\phi = (aI_r|L')$ is Gauss-reduced and $\tilde{\phi} = (N\phi|\phi')$. □

A nice remark is that the obstruction to showing unconditionally that $S_r(C \times p, \mathbb{O}_\varepsilon)$ is included in $S_r(C, (\Gamma_p^g)_{\varepsilon'})$ is exactly due to the nonspecial morphisms. Sets of the form

$$(C(\overline{\mathbb{Q}}) \times p) \cap (B_{\tilde{\phi}} + \mathbb{O}_\varepsilon)$$

not having bounded height can be included in $S_r(C, (\Gamma_p^g)_{\varepsilon'})$ if $\tilde{\phi}$ is special; indeed in general

$$\varepsilon' = c(g, s) \frac{H(\tilde{\phi})}{H(A)} \varepsilon$$

for any $\tilde{\phi} = (A|\phi')$.

Proposition 10.2. *Let $2 \leq r$ and $\varepsilon \leq \min(\varepsilon_1, K_1/g)$. The map $x \rightarrow (x, \gamma)$ defines an injection*

$$\bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk } \phi=r}} C(\overline{\mathbb{Q}}) \cap (B_\phi + (\Gamma_0^g)_\varepsilon) \hookrightarrow \bigcup_{\substack{\tilde{\phi}=(N\phi|\phi') \text{ special} \\ \text{rk } \tilde{\phi}=r}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_\varepsilon).$$

Proof. Let $x \in C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_0^g + \mathbb{O}_\varepsilon)$, with ϕ Gauss-reduced of rank r . Equivalently, there exist $y \in \Gamma_0^g$ and $\xi \in \mathbb{O}_\varepsilon \subset E^g$ such that

$$\phi(x + y + \xi) = 0.$$

Since $\gamma_1, \dots, \gamma_s$ is a maximal free set of Γ_0 , there exists an integer N and a matrix $G \in M_{r,s}(\mathbb{Z})$ such that

$$[N]y = G(\gamma).$$

Let n be the greatest common divisor of the entries of $(N\phi|\phi G)$. We define

$$\tilde{\phi} = \frac{1}{n}(N\phi|\phi G).$$

Clearly

$$(N\phi|\phi G)((x, \gamma) + (\xi, 0)) = N\phi(x) + \phi G(\gamma) + N\phi(\xi) = N\phi(x + y + \xi) = 0.$$

Thus

$$n\tilde{\phi}((x, \gamma) + (\xi, 0)) = 0. \tag{10-1}$$

Equivalently,

$$(x, \gamma) \in (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{n\tilde{\phi}} + \mathcal{O}_\varepsilon).$$

By Lemma 4.4(i) with $\psi = \tilde{\phi}$ and $N = n$, it follows

$$(x, \gamma) \in (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathcal{O}_\varepsilon).$$

We next show that $\tilde{\phi}$ is special. By assumption, the morphism ϕ is Gauss-reduced. By the definition of $\tilde{\phi}$, the greatest common divisor of its entries is 1. In order to conclude that $\tilde{\phi}$ is special, we still have to show that

$$H(\tilde{\phi}) = Na$$

or equivalently

$$H(\phi') \leq Na.$$

The proof is similar to the last part of the proof of Theorem 9.1(ii).

Let $\phi' = (b_{ij}) = \phi G$. Let $|b_{kl}| = \max_{ij} |b_{ij}| = H(\phi')$. Let ϕ_k be the k -th row of ϕ . Consider the k -th row of the system (10-1):

$$nN(\phi_k(x) + \phi_k(\xi)) = -n \sum_j b_{kj} \gamma_j.$$

Then

$$\frac{\|\phi_k(x)\|}{a} + \frac{\|\phi_k(\xi)\|}{a} \geq \frac{1}{Na} \left\| \sum_j b_{kj} \gamma_j \right\|.$$

Clearly $x \in S_r(C, (\Gamma_0^g)_\varepsilon)$. Since $\varepsilon \leq \varepsilon_1$, we have $x \in S_2(C, (\Gamma_0^g)_{\varepsilon_1})$, which has norm bounded by K_1 . So

$$\|x\| \leq K_1.$$

Since $H(\phi_k) \leq H(\phi) = a$,

$$\frac{\|\phi_k(x)\|}{a} \leq (g - r + 1)K_1.$$

Furthermore,

$$\frac{\|\phi_k(\xi)\|}{a} \leq (g - r + 1)\varepsilon.$$

Then

$$(g - 1)(K_1 + \varepsilon) \geq \frac{1}{Na} \left\| \sum_j b_{kj} \gamma_j \right\|.$$

From relations (8-1) with $(b_1, \dots, b_s) = (b_{k1}, \dots, b_{ks})$ and (8-2), we deduce

$$(g - 1)(K_1 + \varepsilon) \geq \frac{1}{Na} \left(\frac{1}{9} \sum_j |b_{kj}|^2 \|\gamma_j\|^2 \right)^{1/2} \geq \frac{H(\phi')}{3Na} \min_j \|\gamma_j\| \geq \frac{H(\phi')}{3Na} 3gK_1.$$

We assumed that $\varepsilon \leq K_1/g$, so $H(\phi') \leq Na$. □

This inclusion is important; the Bogomolov-type bounds are given for intersections with \mathbb{O}_ε and not with Γ_ε . Actually there exist bounds for ε such that $C \cap \Gamma_\varepsilon$ is finite. They are deduced using the Bogomolov-type bounds and their dependence on the degree of the curve is not sharp enough for our purpose. To overcome this obstacle and solve the problem with Γ_ε , we use Proposition 10.2 and the Bogomolov-type bounds for $C \times \gamma$ intersected with $B_{\tilde{\phi}} + \mathbb{O}_\varepsilon$, where $\tilde{\phi}$ is special of rank 2.

11. Proof of Theorem 1.5: Structure

Sections 12 and 13 below will develop the core of the proof of Theorem 1.5. In Proposition A we show that the union can be taken over finitely many sets, while in Proposition B we show that each set in the union is finite.

We prefer to present first the proof of Theorem 1.5 assuming Propositions A and B, and then to prove them. We hope that, knowing a priori the aim of sections 12 and 13, the reader gets the right inspiration to handle the proofs.

Proof of Theorem 1.5. Assuming Conjecture 1.4, we prove Conjecture 1.1(iv). In view of Theorem 1.3, part (iii) is also proved. Parts (i) and (ii) are then obtained by setting $\varepsilon = 0$.

Choose

$$\begin{aligned} n &= 2(g + s) - 3, \\ \delta_1 &= \frac{\min(\varepsilon_4, \varepsilon_2)}{(g + s)^2}, \quad \text{where } \varepsilon_4 \text{ is as in Proposition B,} \\ \delta &= \delta_1 M'^{-1-1/(2n)}, \quad \text{where } M' = \max(2, \lceil K_2/\delta_1 \rceil^2)^n. \end{aligned}$$

Recall that $\Gamma_\delta \subset (\Gamma_0^g)_\delta$. Apply Lemma 4.5, replacing ε by δ . Then

$$S_2(C, \Gamma_\delta) \subset \bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk } \phi=2}} C(\overline{\mathbb{Q}}) \cap (B_\phi + (\Gamma_0^g)_\delta).$$

Observing that $\delta < \delta_1 < \min(\varepsilon_1, K_1/g)$ and applying Proposition 10.2 (again with $\varepsilon = \delta$) we obtain an injection

$$\bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk } \phi=2}} C(\overline{\mathbb{Q}}) \cap (B_\phi + (\Gamma_0^g)_\delta) \hookrightarrow \bigcup_{\substack{\tilde{\phi}=(N\phi|\phi') \text{ special} \\ \text{rk } \tilde{\phi}=2}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_\delta).$$

Note that $\delta = \delta_1 M'^{-(1+1/(2n))}$ and $\delta_1 \leq \varepsilon_2$. Then, Proposition A(ii) in Section 12 below, with $\varepsilon = \delta_1$, $r = 2$ (and n already defined as $2(g+s) - 4 + 1$), shows that

$$\bigcup_{\substack{\tilde{\phi} \text{ special} \\ \text{rk } \tilde{\phi}=2}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_\delta)$$

is a subset of

$$\bigcup_{\substack{\tilde{\phi} \text{ special} \\ H(\tilde{\phi}) \leq M' \\ \text{rk } \tilde{\phi}=2}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{(g+s)\delta_1/H(\tilde{\phi})^{1+1/(2n)}}). \tag{11-1}$$

Observe that in (11-1), $\tilde{\phi}$ ranges over finitely many morphisms, because $H(\tilde{\phi})$ is bounded by M' .

We have chosen $\delta_1 \leq \varepsilon_4/(g+s)^2$. Proposition B(ii) in Section 13 below with $\varepsilon = (g+s)\delta_1$, implies that for all $\tilde{\phi} = (N\phi|\phi')$ special of rank 2, the set

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{(g+s)\delta_1/H(\tilde{\phi})^{1+1/(2n)}})$$

is finite. Note that $H(\phi) \leq H(\tilde{\phi})$, thus also the sets

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{(g+s)\delta_1/H(\tilde{\phi})^{1+1/(2n)}})$$

appearing in (11-1) are finite.

It follows that, the set $S_2(C, \Gamma_\delta)$ is contained in the union of finitely many finite sets. So it is finite. □

Despite our proof relying on Dirichlet’s Theorem and a Bogomolov-type bound, a direct use of these two theorems is not sufficient to prove Theorem 1.5. Using Dirichlet’s Theorem in a more natural way, one can prove that, for $r \geq 2$,

$$S_r(C, \Gamma_\varepsilon) \subset \bigcup_{\substack{H(\phi) \leq M(\varepsilon) \\ \text{rk } \phi=r}} C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_\varepsilon).$$

On the other hand, a direct use of the Bogomolov type bound shows that

$$C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathcal{O}_{\varepsilon/H(\phi)^2})$$

is finite, for ϕ of rank at least 2. Even if we forget Γ , the discrepancy between ε and $\varepsilon/H(\phi)^2$ does not look encouraging, and it took us a long struggle to overcome the problem. In Propositions A and B, we succeed in overcoming the mismatch; in both statements we obtain neighborhoods of radius $\varepsilon/H(\phi)^{1+1/(2n)}$.

Warning: One might think that, since we consider only morphisms ϕ such that $H(\phi) \leq M$, it might be enough to choose $\varepsilon' = \varepsilon/M^2$. However, $M = M(\varepsilon)$ is an unbounded function of ε as ε tends to 0.

12. Proof of Theorem 1.5:

The box principle and the reduction to a finite union

In Lemma 12.2, we approximate a Gauss-reduced morphism by a Gauss-reduced morphism of bounded height. Such an approximation allows us to restrict our attention to unions over finitely many algebraic subgroups, instead of over all algebraic subgroups; this is Proposition A, already mentioned. We start by recalling Dirichlet’s Theorem on the rational approximation of reals.

Theorem 12.1 (Dirichlet, 1842; see [Schmidt 1980, p. 24, Theorem 1]). *Suppose that $\alpha_1, \dots, \alpha_n$ are real numbers and that $Q \geq 2$ is an integer. There exist integers f, f_1, \dots, f_n such that*

$$1 \leq f < Q^n \quad \text{and} \quad |\alpha_i f - f_i| \leq \frac{1}{Q} \quad \text{for } 1 \leq i \leq n. \quad (12-1)$$

Lemma 12.2. *Let $Q \geq 2$ be an integer. Let $\phi = (aI_r | L) \in M_{r \times g}(\mathbb{Z})$ be Gauss-reduced. There exists a Gauss-reduced $\psi = (fI_r | L') \in M_{r \times g}(\mathbb{Z})$ such that*

- (i) $H(\psi) = f \leq Q^{rg-r^2+1}$ and
- (ii) $\left| \frac{\psi}{f} - \frac{\phi}{a} \right| \leq Q^{-1/2} f^{-1 - \frac{1}{2(rg-r^2+1)}}$.

Here the norm $|\cdot|$ of a matrix is the maximum of the absolute values of its entries.

Proof. If $a \leq Q^{rg-r^2+1}$ no approximation is needed, since ϕ itself satisfies the conclusion. So we can assume that

$$\phi = \begin{pmatrix} a & \dots & 0 & L_1 \\ \vdots & & \vdots & \vdots \\ 0 & \dots & a & L_r \end{pmatrix}$$

is a Gauss-reduced morphism such that $H(\phi) = a > Q^{rg-r^2+1}$. Consider the

element

$$\alpha = \left(1, \frac{L_1}{a}, \dots, \frac{L_r}{a} \right) = (\alpha_1, \alpha_2, \dots, \alpha_{rg-r^2+1}) \in \mathbb{R}^{rg-r^2+1}.$$

Set $n = rg - r^2 + 1$. Apply Dirichlet's Theorem to α to select integers f, f_1, \dots, f_n satisfying (12-1); they can be assumed to have greatest common divisor 1. Define

$$w = \frac{1}{f}(f_1, \dots, f_n) = \frac{1}{f}(f_1, L'_1, \dots, L'_r),$$

with $L'_i \in \mathbb{Z}^{g-r}$. We claim that

$$f_1 = f, \quad |f_i| \leq f.$$

Indeed, (12-1) for $i = 1$ yields

$$\left| \frac{f_1}{f} - 1 \right| \leq \frac{1}{Qf},$$

so $|f - f_1| < 1$. Since f and f_1 are integers, we must have $f = f_1$. Similarly, by (12-1) for $i = 2, \dots, n$, we have $|f_i/f - \alpha_i| \leq 1/(Qf)$. This implies that $|f_i| \leq f + 1/Q$. We deduce that $|f_i| \leq f$.

It follows that

$$\psi = \begin{pmatrix} f & \dots & 0 & L'_1 \\ \vdots & & \vdots & \\ 0 & \dots & f & L'_r \end{pmatrix}$$

is a Gauss-reduced morphism of rank r with $H(\psi) = f$.

Relation (12-1) immediately gives

$$f \leq Q^n$$

and

$$\left| \frac{\psi}{f} - \frac{\phi}{a} \right| \leq \frac{1}{Qf} \leq \frac{1}{Q^{1/2} f^{1+1/(2n)}},$$

where in the last inequality we have used the inequality $Q^{1/2} \geq f^{1/(2n)}$. □

At last we can prove our first main proposition; the union can be taken over finitely many algebraic subgroups. If ϕ has large height and B_ϕ is close to x , with x in a set of bounded height, then there exists ψ with height bounded by a constant such that B_ψ is also close to x . One shall be careful that, in the following inclusions, on the left-hand side we consider a neighborhood of B_ϕ of fixed radius, while on the right-hand side the neighborhood becomes smaller as the height of ψ grows. This is a crucial gain with respect to the simpler approximation (obtained by a direct use of Dirichlet's Theorem) where the neighborhoods have constant radius on both sides.

Proposition A. *Assume $r \geq 2$.*

(i) *If $0 < \varepsilon \leq \varepsilon_1$, then*

$$\bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk } \phi=r}} C(\overline{\mathbb{Q}}) \cap (B_\phi + (\Gamma_0^g)_{\varepsilon/M^{1+1/(2n)}}) \subset \bigcup_{\substack{\psi \text{ Gauss-reduced} \\ \text{rk } \psi=r \\ H(\psi) \leq M}} C(\overline{\mathbb{Q}}) \cap (B_\psi + (\Gamma_0^g)_{g\varepsilon/H(\psi)^{1+1/(2n)}}),$$

where $n = rg - r^2 + 1$ and $M = \max(2, \lceil K_1/\varepsilon \rceil^2)^n$.

(ii) *If $0 < \varepsilon \leq \varepsilon_2$, then*

$$\bigcup_{\substack{\tilde{\phi} \text{ special} \\ \text{rk } \tilde{\phi}=r}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{\varepsilon/M^{1+1/(2n)}}) \subset \bigcup_{\substack{\tilde{\psi} \text{ special} \\ \text{rk } \tilde{\psi}=r \\ H(\tilde{\psi}) \leq M'}} (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\psi}} + \mathbb{O}_{(g+s)\varepsilon/H(\tilde{\psi})^{1+1/(2n)}}),$$

where $n = r(g + s) - r^2 + 1$ and $M' = \max(2, \lceil K_2/\varepsilon \rceil^2)^n$.

Proof. (i) Let $\phi = (aI_r|L)$ be Gauss-reduced of rank r .

First consider the case $H(\phi) \leq M$. Then $\varepsilon/M^{1+1/(2n)} \leq \varepsilon/H(\phi)^{1+1/(2n)}$. Obviously

$$C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_0^g + \mathbb{O}_{\varepsilon/M^{1+1/(2n)}}) \subset C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_0^g + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}})$$

is contained in the right-hand side.

Secondly consider the case $H(\phi) > M$. We shall show that there exists ψ Gauss-reduced with $H(\psi) \leq M$ such that

$$C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_0^g + \mathbb{O}_{\varepsilon/M^{1+1/(2n)}}) \subset C(\overline{\mathbb{Q}}) \cap (B_\psi + \Gamma_0^g + \mathbb{O}_{g\varepsilon/H(\psi)^{1+1/(2n)}}).$$

We fix $Q = \max(2, \lceil K_1/\varepsilon \rceil^2)$. Recall that $n = rg - r^2 + 1$. By Lemma 12.2, there exists a Gauss-reduced morphism

$$\psi = \begin{pmatrix} f & \dots & 0 & L'_1 \\ \vdots & & \vdots & \\ 0 & \dots & f & L'_r \end{pmatrix}$$

such that

$$H(\psi) = f \leq M$$

and

$$\left| \frac{\psi}{f} - \frac{\phi}{a} \right| \leq \frac{1}{Q^{1/2} f^{1+1/(2n)}}. \tag{12-2}$$

Let $x \in C(\overline{\mathbb{Q}}) \cap (B_\phi + \Gamma_0^g + \mathbb{O}_{\varepsilon/M^{1+1/(2n)}})$. Then there exist $y \in \Gamma_0^g$ and $\xi \in \mathbb{O}_{\varepsilon/M^{1+1/(2n)}}$ such that

$$\phi(x - y - \xi) = 0.$$

We want to show that there exist $y' \in \Gamma_0^g$ and $\xi' \in \mathbb{O}_{g\varepsilon/f^{1+1/(2n)}}$ such that

$$\psi(x - y' - \xi') = 0.$$

Let y'' be a point such that

$$[a]y'' = \phi(y).$$

Since Γ_0 is a division group, $y'' \in \Gamma_0^r$. We define

$$y' = (y'', 0) \in \Gamma_0^r \times \{0\}^{g-r},$$

whence

$$\psi(y') = [f]y''.$$

Let ξ'' be a point such that

$$[f]\xi'' = \psi(x - y'),$$

and define $\xi' = (\xi'', 0)$. Then

$$\psi(\xi') = [f]\xi'' = \psi(x - y') \quad \text{and} \quad \psi(x - y' - \xi') = 0.$$

It follows that

$$x \in C(\overline{\mathbb{Q}}) \cap (B_\psi + \Gamma_0^g + \mathbb{O}_{\|\xi'\|}).$$

In order to finish the proof, we are going to prove that

$$\|\xi'\| \leq \frac{g\varepsilon}{f^{1+1/(2n)}}.$$

By definition

$$\|\xi'\| = \|\xi''\| = \frac{\|\psi(x - y')\|}{f}.$$

Consider the equivalence

$$\begin{aligned} a\psi(x - y') &= a\psi(x) - a\psi(y') = a\psi(x) - a[f]y'' \\ &= a\psi(x) - f\phi(y) = a\psi(x) - f\phi(x) + f\phi(\xi). \end{aligned}$$

Then

$$\|\xi'\| = \frac{1}{af} \|f\phi(\xi) - f\phi(x) + a\psi(x)\| \leq \frac{1}{a} \|\phi(\xi)\| + \frac{1}{af} \|a\psi(x) - f\phi(x)\|.$$

We estimate separately each norm on the right.

On the one hand,

$$\frac{1}{a} \|\phi(\xi)\| \leq (g - r + 1) \|\xi\| \leq \frac{(g - 1)\varepsilon}{M^{1+1/(2n)}} \leq \frac{(g - 1)\varepsilon}{f^{1+1/(2n)}},$$

because $\|\xi\| \leq \varepsilon/M^{1+1/(2n)}$ and $f \leq M$.

On the other hand, since the rank of ϕ is at least 2 and $\varepsilon \leq \varepsilon_1$, we have $x \in S_2(C, (\Gamma_0^g)_{\varepsilon_1})$, which has norm K_1 . Thus

$$\|x\| \leq K_1.$$

Using relation (12-2) and that $Q \geq \lceil K_1/\varepsilon \rceil^2$, it follows that

$$\begin{aligned} \frac{1}{af} \|a\psi(x) - f\phi(x)\| &\leq \left| \frac{\psi}{f} - \frac{\phi}{a} \right| \|x\| \leq \frac{1}{Q^{1/2} f^{1+1/(2n)}} \|x\| \\ &\leq \frac{\varepsilon \|x\|}{K_1 f^{1+1/(2n)}} \leq \frac{\varepsilon}{f^{1+1/(2n)}}. \end{aligned}$$

We obtain

$$\|\xi'\| \leq \frac{(g - 1)\varepsilon}{f^{1+1/(2n)}} + \frac{\varepsilon}{f^{1+1/(2n)}} \leq \frac{g\varepsilon}{f^{1+1/(2n)}},$$

concluding the proof of part (i) of the proposition.

(ii) We fix $Q = \max(2, \lceil K_2/\varepsilon \rceil^2)$. Let $\tilde{\phi} = (N\phi|\phi') : E^{g+s} \rightarrow E^r$ be special. From conditions (i) and (ii) of Definition 10.1 we know that

$$\tilde{\phi} = (bI_r|*)$$

is Gauss-reduced and $H(\tilde{\phi}) = b$.

As in part (i) of the proof, if $H(\tilde{\phi}) \leq M'$ then $\varepsilon/M'^{1+1/(2n)} \leq \varepsilon/H(\tilde{\phi})^{1+1/(2n)}$ and the set

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{\varepsilon/M'^{1+1/(2n)}})$$

is contained in the right-hand side.

Now, suppose that $H(\tilde{\phi}) > M'$. Recall that $n = r(g + s) - r^2 + 1$. By Lemma 12.2 (applied with $\phi = \tilde{\phi}$ and $\psi = \tilde{\psi}$) there exists a Gauss-reduced $\tilde{\psi} = (fI_r|*)$ such that $H(\tilde{\psi}) = f \leq M'$ and

$$\left| \frac{\tilde{\phi}}{b} - \frac{\tilde{\psi}}{f} \right| \leq \frac{1}{Q^{1/2} f^{1+1/(2n)}}. \tag{12-3}$$

Then $\tilde{\psi}$ is special, since it satisfies (i) and (ii) in Definition 10.1.

The proof is now similar to that of part (i). We want to show that, if $\tilde{\phi}((x, \gamma) + \xi)$ vanishes for $\xi \in \mathbb{O}_{\varepsilon/M'^{1+1/(2n)}}$, then $\tilde{\psi}((x, \gamma) + \xi')$ vanishes for $\xi' \in \mathbb{O}_{(g+s)\varepsilon/H(\tilde{\psi})^{1+1/(2n)}}$.

Let ξ'' be a point in E^r such that

$$[f]\xi'' = -\tilde{\psi}(x, \gamma).$$

We define $\xi' = (\xi'', \{0\}^{g-r+s})$. Then

$$\tilde{\psi}((x, \gamma) + (\xi', 0)) = 0.$$

It follows that

$$(x, \gamma) \in (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\psi}} + \mathbb{O}_{\|\xi'\|}),$$

where $\tilde{\psi}$ is special and $H(\tilde{\psi}) \leq M'$.

It remains to prove that

$$\|\xi'\| \leq \frac{(g+s)\varepsilon}{H(\tilde{\psi})^{1+1/(2n)}}.$$

Obviously

$$b\tilde{\psi}(x, \gamma) = f(\tilde{\phi}(x, \gamma) - \tilde{\phi}(x, \gamma)) + b\tilde{\psi}(x, \gamma).$$

According to the definition of ξ' ,

$$\begin{aligned} \|\xi'\| = \|\xi''\| &= \frac{\|\tilde{\psi}(x, \gamma)\|}{f} = \frac{1}{bf} \|f(\tilde{\phi}(x, \gamma) - \tilde{\phi}(x, \gamma)) + b\tilde{\psi}(x, \gamma)\| \\ &\leq \frac{1}{b} \|\tilde{\phi}(x, \gamma)\| + \frac{1}{bf} \|b\tilde{\psi}(x, \gamma) - f\tilde{\phi}(x, \gamma)\|. \end{aligned}$$

We estimate the two norms on the right.

On the one hand,

$$\frac{\|\tilde{\phi}(x, \gamma)\|}{b} = \frac{\|\tilde{\phi}(\xi)\|}{b} \leq (g-r+1+s)\|\xi\| \leq \frac{(g-r+1+s)\varepsilon}{M'^{1+1/(2n)}} \leq \frac{(g-r+1+s)\varepsilon}{f^{1+1/(2n)}},$$

where in the last inequality we have used that $f \leq M'$.

On the other hand, by the definition of ε_2 , we know that the norm of the set $S_2(C \times \gamma, \mathbb{O}_{\varepsilon_2})$ is bounded by K_2 . Since $\varepsilon \leq \varepsilon_2$, we have $(x, \gamma) \in S_2(C \times \gamma, \mathbb{O}_{\varepsilon_2})$. Therefore

$$\|(x, \gamma)\| \leq K_2.$$

Using relation (12-3) and the inequality $Q \geq \lceil K_2/\varepsilon \rceil^2$, we estimate

$$\begin{aligned} \frac{1}{bf} \|b\tilde{\psi}(x, \gamma) - f\tilde{\phi}(x, \gamma)\| &\leq \left| \frac{\tilde{\phi}}{b} - \frac{\tilde{\psi}}{f} \right| \|(x, \gamma)\| \leq \frac{\|(x, \gamma)\|}{Q^{1/2} f^{1+1/(2n)}} \\ &\leq \frac{\varepsilon \|(x, \gamma)\|}{(K_2) f^{1+1/(2n)}} \leq \frac{\varepsilon}{f^{1+1/(2n)}}. \end{aligned}$$

Since $r \geq 2$, we conclude that

$$\|\xi'\| \leq \frac{(g-1+s)\varepsilon}{f^{1+1/(2n)}} + \frac{\varepsilon}{f^{1+1/(2n)}} = \frac{(g+s)\varepsilon}{H(\tilde{\psi})^{1+1/(2n)}}. \quad \square$$

13. Proof of Theorem 1.5:

The essential minimum and the finiteness of each intersection

Up until now we have used, several times, the boundedness of the height of our sets. In this section we often use the fact that we are working with a curve.

In the following, we set

$$n = 2(g + s) - 3.$$

We would like to use Conjecture 1.4 to provide $\varepsilon > 0$ such that, for all ϕ Gauss-reduced of rank $r = 2$, the set

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_\phi + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}) \tag{13-1}$$

is finite. This set is simply

$$\phi_{|C \times \gamma}^{-1} (\phi(C \times \gamma) \cap \phi(\mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}})).$$

Further

$$\phi(\mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}) \subset \mathbb{O}_{g\varepsilon/H(\phi)^{1/(2n)}},$$

because if $\zeta \in \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}$ then $\|\phi(\zeta)\| \leq gH(\phi)\|\zeta\| \leq g\varepsilon H(\phi)^{-1/(2n)}$. Thus, the set (13-1) is contained in the preimage of

$$\phi(C \times \gamma) \cap \mathbb{O}_{g\varepsilon/H(\phi)^{1/(2n)}}.$$

If we can ensure that there exists $\varepsilon > 0$ such that, for all morphisms ϕ Gauss-reduced of rank $r = 2$,

$$g\varepsilon H(\phi)^{-1/(2n)} < \mu(\phi(C \times \gamma)), \tag{13-2}$$

then the set (13-1) is finite.

The direct use of a Bogomolov-type bound, even an optimal one, is not successful in the following sense: For a curve $X \subset E^s$ and any $\eta > 0$, Conjecture 1.4 provides an invariant $\varepsilon(X, \eta)$ such that $\varepsilon(X, \eta) < \mu(X)$. To ensure (13-2), we could naively require that

$$g\varepsilon H(\phi)^{-1/(2n)} \leq \varepsilon(\phi(C \times \gamma), \eta)$$

for all ϕ Gauss-reduced of rank $r = 2$. But this can be fulfilled only for $\varepsilon = 0$.

We need to throw new light on the problem in order to prove (13-2); via some isogenies, we construct a helping curve D and then we relate its essential minimum to $C \times \gamma$. We then apply Conjecture 1.4 to D . In this way we manage to provide a good lower bound for the essential minimum of $C \times \gamma$. We take advantage of the fact that $\mu([b]C) = b\mu(C)$, while $\varepsilon([b]C, \eta) = \varepsilon(C, \eta)/b^{1/(g-1)+2\eta}$ for any positive integers b .

Let

$$\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \begin{pmatrix} a & 0 & L_1 \\ 0 & a & L_2 \end{pmatrix}$$

be a Gauss-reduced morphism of rank 2 with $H(\phi) = a$. We introduce the notation $\bar{x} = (x_3, \dots, x_g)$, and recall that $n = 2(g + s) - 3$.

We define

$$a_0 = \lfloor a^{1/(2n)} \rfloor.$$

We associated to the morphism ϕ an isogeny

$$\Phi : E^g \rightarrow E^g, \quad (x_1, \dots, x_g) \mapsto (a_0\phi(x), x_3, \dots, x_g).$$

We then relate it to the isogenies

$$\begin{aligned} A : E^g &\rightarrow E^g, & (x_1, \dots, x_g) &\mapsto (x_1, x_2, ax_3, \dots, ax_g), \\ A_0 : E^g &\rightarrow E^g, & (x_1, \dots, x_g) &\mapsto (x_1, x_2, a_0x_3, \dots, a_0x_g), \\ L : E^g &\rightarrow E^g, & (x_1, \dots, x_g) &\mapsto (x_1 + L_1(\bar{x}), x_2 + L_2(\bar{x}), x_3, \dots, x_g). \end{aligned}$$

Definition 13.1 (Helping curve). We define the curve D to be an irreducible component of $A_0^{-1}LA^{-1}(C)$, where $(\cdot)^{-1}$ simply means the inverse image.

The obvious relation

$$[a_0a]D = \Phi(C)$$

is going to play a key role in the following.

We need to estimate degrees, since the Bogomolov-type bound depends on the degree of the curve.

Lemma 13.2. (i) *The degree of the curve $\phi(C)$ in E^2 is bounded by $6ga^2 \deg C$.*
 (ii) *The degree of the curve D in E^g is bounded by $12g^2a_0^{2(g-2)}a^{2(g-1)} \deg C$.*

Proof. (i) Consider

$$\deg \phi(C) = \sum_{i=1}^2 \phi(C) \cdot H_i,$$

where H_i is the coordinate divisor given by $3x_i = 0$. The intersection number $\phi(C) \cdot H_i$ is bounded by the degree of the morphism $\phi_{i|C} : C \rightarrow E$. Recall that $\phi = \begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$. By Bézout’s Theorem, $\deg \phi_{i|C}$ is at most $3ga^2 \deg C$; see [Viada 2003, p. 61]. Therefore

$$\deg \phi(C) \leq 6ga^2 \deg C.$$

(ii) Let X be a generic transverse curve in E^g . By [Hindry 1988, Lemma 6(i)], we deduce that

$$\deg A^{-1}(X) \leq 2a^{2(g-2)} \deg X, \quad \deg A_0^{-1}(X) \leq 2a_0^{2(g-2)} \deg X.$$

To estimate the degree of $L(X)$, we proceed as in part (i). We write

$$\deg L(X) = \sum_{i=1}^g L(X) \cdot H_i,$$

where H_i is given by $3x_i = 0$. The intersection number $L(X) \cdot H_i$ is bounded by the degree of the morphism $L'_{iX} : X \rightarrow E$, where L'_i is the i th. row of L . By Bézout's Theorem, $\deg L'_{iX}$ is at most $3ga^2 \deg X$. Therefore

$$\deg L(X) \leq 3g^2 a^2 \deg X.$$

We conclude that

$$\begin{aligned} \deg D &\leq \deg A_0^{-1} L A^{-1}(C) \leq 2a_0^{2(g-2)} \deg L A^{-1}(C) \\ &\leq 6g^2 a_0^{2(g-2)} a^2 \deg A^{-1}(C) \leq 12g^2 a_0^{2(g-2)} a^2 a^{2(g-2)} \deg C. \quad \square \end{aligned}$$

The next proposition is a lower bound for the essential minimum of the image of a curve under Gauss-reduced morphisms. It reveals the dependence on the height of the morphism. While the first bound is an immediate application of Conjecture 1.4, the second estimate is subtle. Our lower bound for $\mu(\Phi(C + y))$ grows with $H(\phi)$. On the contrary, the Bogomolov-type lower bound $\epsilon(\Phi(C + y))$ goes to zero as $(a_0 H(\phi))^{-1/(g-1)-\eta}$ — a nice gain.

Potentially, this suggests an interesting question; to investigate the behavior of the essential minimum under a general morphism.

Proposition 13.3. *Assume Conjecture 1.4 and take $y \in E^g(\overline{\mathbb{Q}})$ and $\eta > 0$. Then:*

(i)
$$\mu(\phi(C + y)) > \epsilon_1(C, \eta) a^{-(1+2\eta)},$$

where $\epsilon_1(C, \eta)$ is an effective constant depending on C and η . (Recall that $a = H(\phi)$.)

(ii)
$$\mu(\Phi(C + y)) > \epsilon_2(C, \eta) a_0^{1/(g-1)-8(g+s)(g-1)\eta},$$

where $\epsilon_2(C, \eta)$ is an effective constant depending on C , g and η . (Recall that $a_0 = \lfloor a^{1/(2n)} \rfloor$.)

Proof. Recall the Bogomolov-type bound given in Conjecture 1.4: for a transverse irreducible curve X in E^g over $\overline{\mathbb{Q}}$ and any $\eta > 0$,

$$\epsilon(X, \eta) = \frac{c(g, E, \eta)}{\deg X^{1/(2 \operatorname{codim} X) + \eta}} < \mu(X).$$

(i) Observe that $\phi(C) \subset E^2$ has codimension 1.

Let $q' = \phi(y)$. So $\phi(C + y) = \phi(C) + q'$. Since C is irreducible, transverse and defined over $\overline{\mathbb{Q}}$, so is $\phi(C) + q'$. Conjecture 1.4 gives

$$\mu(\phi(C + y)) = \mu(\phi(C) + q') > \epsilon(\phi(C) + q', \eta) = \frac{c(2, E, \eta)}{(\deg(\phi(C) + q'))^{1/2+\eta}}.$$

Degrees are preserved by translations; hence Lemma 13.2(i) implies that

$$\deg(\phi(C) + q') = \deg \phi(C) \leq 9ga^2 \deg C.$$

It follows that

$$\epsilon(\phi(C) + q', \eta) \geq \frac{c(2, E, \eta)}{(9ga^2 \deg C)^{1/2+\eta}}.$$

Define

$$\epsilon_1(C, \eta) = \frac{c(2, E, \eta)}{(9g \deg C)^{1/2+\eta}}.$$

Then

$$\mu(\phi(C + y)) \geq \frac{\epsilon_1(C, \eta)}{a^{1+2\eta}}.$$

(ii) Let $q \in E^g$ be a point such that $[a_0a]q = \Phi(y)$. Then

$$\Phi(C + y) = [a_0a](A_0^{-1}LA^{-1}(C) + q) = [a_0a](D + q).$$

Therefore

$$\mu(\Phi(C + y)) = (a_0a)\mu(D + q). \tag{13-3}$$

We now estimate $\mu(D + q)$ using Conjecture 1.4. The curve $D + q$ is irreducible by the definition of D . Since C is transverse and defined over $\overline{\mathbb{Q}}$, so is $D + q$. Thus

$$\mu(D + q) > \epsilon(D + q, \eta) = \frac{c(g, E, \eta)}{\deg(D + q)^{1/(2(g-1))+\eta}}.$$

Translations by a point preserve degrees, so Lemma 13.2(ii) gives

$$\deg(D + q) = \deg D \leq 12g^2a_0^{2(g-2)}a^{2(g-1)} \deg C.$$

Then

$$\epsilon(D + q, \eta) \geq \frac{c(g, E, \eta)}{(12g^2 \deg C)^{1/(2(g-1))+\eta}} (a_0^{2(g-2)}a^{2(g-1)})^{-\frac{1}{2(g-1)}-\eta}.$$

Define

$$\epsilon_2(C, \eta) = \frac{c(g, E, \eta)}{(12g^2 \deg C)^{1/(2(g-1))+\eta}}.$$

So

$$\mu(D + q) \geq \epsilon_2(C, \eta)a_0^{-1+\frac{1}{g-1}-2(g-2)\eta} a^{-1-2(g-1)\eta}.$$

Substitute into (13-3), to obtain

$$\mu(\Phi(C + y)) > \epsilon_2(C, \eta) a_0^{\frac{1}{g-1} - 2(g-2)\eta} a^{-2(g-1)\eta}.$$

Recall that a_0 is the integral part of $a^{1/(2n)}$, where $n = 2(g+s) - 3$. So $2a_0 \geq a^{1/(2n)}$ and

$$a^{2(g-1)\eta} \leq (2a_0)^{4n(g-1)\eta}.$$

Further, $2(g-2) + 4n(g-1) \leq 8(g+s)(g-1)$, so

$$\mu(\Phi(C + y)) > \epsilon_2(C, \eta) a_0^{1/(g-1) - 8(g+s)(g-1)\eta}. \quad \square$$

We now come to our second main proposition: each set in the union is finite. The proof of (i) case (1) below is delicate. In general $\mu(\pi(C)) \leq \mu(C)$, for π a projection on some factors. We shall rather find a kind of reverse inequality. On a set of bounded height this will be possible.

Proposition B. *Assume Conjecture 1.4. There exists $\epsilon_4 > 0$ with the following properties:*

- (i) *For $\epsilon \leq \epsilon_4$, for all $y \in \Gamma_0^2 \times \{0\}^{g-2}$ and for all Gauss-reduced morphisms ϕ of rank 2, the set*

$$(C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathcal{O}_{\epsilon/H(\phi)}^{1+1/(2n)})$$

is finite.

- (ii) *For $\epsilon \leq \epsilon_4/(g+s)$ and for all special morphisms $\tilde{\phi} = (N\phi|\phi')$ of rank 2, the set*

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathcal{O}_{\epsilon/H(\phi)}^{1+1/(2n)})$$

is finite.

(Recall that $n = 2(g+s) - 3$.)

Proof. (i) Choose

$$\eta \leq \eta_0 = \frac{1}{2^4(g+s)(g-1)^2}.$$

Define

$$m = \max\left(2, \left(\frac{K_1}{\epsilon_2(C, \eta)}\right)^{\frac{g-1}{1-8(g+s)(g-1)^2\eta}}\right),$$

and choose

$$\epsilon \leq \min\left(\epsilon_1, \frac{K_1}{g}, \frac{\epsilon_1(C, \eta)}{gm^{4n}}\right),$$

where $\epsilon_1(C, \eta)$ and $\epsilon_2(C, \eta)$ are as in Proposition 13.3.

Recall that $H(\phi) = a$. We distinguish two cases:

- (1) $a_0 = \lfloor a^{1/(2n)} \rfloor \geq m$,

$$(2) a_0 = \lfloor a^{1/(2n)} \rfloor \leq m.$$

Case (1): $a_0 \geq m$. Let $x + y \in (C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathbb{O}_{\varepsilon/a^{1+1/(2n)}})$, where

$$y = (y_1, y_2, 0, \dots, 0) \in \Gamma_0^2 \times \{0\}^{g-2}.$$

Then

$$\phi(x + y) = \phi(\xi)$$

for $\|\xi\| \leq \varepsilon/a^{1+1/(2n)}$.

We have chosen $\varepsilon \leq \varepsilon_1$, so $x \in S_2(C, (\Gamma_0^g)_{\varepsilon_1})$ which is a set of norm K_1 . Then

$$\|x\| \leq K_1.$$

Recall that $\Phi(z_1, \dots, z_g) = (a_0\phi(z), z_3, \dots, z_g)$. So

$$\Phi(x + y) = (a_0\phi(x + y), x_3, \dots, x_g) = (a_0\phi(\xi), x_3, \dots, x_g).$$

Therefore

$$\|\Phi(x + y)\| = \|(a_0\phi(\xi), x_3, \dots, x_g)\| \leq \max(a_0\|\phi(\xi)\|, \|x\|).$$

Since $\|\xi\| \leq \varepsilon a^{-(1+1/(2n))}$, $a_0 \leq a^{1/(2n)}$ and $\varepsilon \leq K_1/g$, we have

$$a_0\|\phi(\xi)\| \leq a_0(g - r + 1) \frac{\varepsilon}{a^{1/(2n)}} \leq K_1.$$

Also $\|x\| \leq K_1$. Thus

$$\|\Phi(x + y)\| \leq K_1.$$

In view of the hypothesis $a_0 \geq m$, we have

$$K_1 \leq \varepsilon_2(C, \eta) a_0^{\frac{1}{g-1} - 8(g+s)(g-1)\eta}.$$

In Proposition 13.3(ii) we have proved that

$$\varepsilon_2(C, \eta) a_0^{\frac{1}{g-1} - 8(g+s)(g-1)\eta} < \mu(\Phi(C + y)).$$

So

$$\|\Phi(x + y)\| \leq K_1 < \mu(\Phi(C + y)).$$

We deduce that $\Phi(x + y)$ belongs to the finite set

$$\Phi(C + y) \cap \mathbb{O}_{K_1}.$$

The morphism $C + y \rightarrow \Phi(C + y)$ has finite fiber. We can conclude that since $\varepsilon \leq \min(\varepsilon_1, K_1/g)$, for every ϕ Gauss-reduced of rank 2 with $a_0 = \lfloor a^{1/(2n)} \rfloor \geq m$, the set

$$(C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}})$$

is finite.

Case (2): $a_0 \leq m$. Let $x + y \in (C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathbb{O}_{\varepsilon/a^{1+1/(2n)}})$, where $y \in \Gamma_0^2 \times \{0\}^{g-2}$.

Then

$$\phi(x + y) = \phi(\xi)$$

for $\|\xi\| \leq \varepsilon/a^{1+1/(2n)}$. However we have chosen $\varepsilon \leq \varepsilon_1(C, \eta)/gm^{4n}$. Hence

$$\|\phi(x + y)\| = \|\phi(\xi)\| \leq \frac{g\varepsilon}{a^{1/(2n)}} \leq \frac{\varepsilon_1(C, \eta)}{m^{4n}a^{1/(2n)}}.$$

We are working under the hypothesis $a_0 = \lfloor a^{1/(2n)} \rfloor \leq m$ and $m \geq 2$, so $a < (2a_0)^{2n} \leq m^{4n}$. Furthermore, $\eta \leq \eta_0 < \frac{1}{4n}$ implies that $a^{2\eta} < a^{1/(2n)}$. Thus

$$a^{1+2\eta} < m^{4n}a^{1/(2n)}.$$

And consequently

$$\|\phi(x + y)\| \leq \frac{\varepsilon_1(C, \eta)}{m^{4n}a^{1/(2n)}} < \frac{\varepsilon_1(C, \eta)}{a^{1+2\eta}}.$$

In Proposition 13.3(i) we proved that

$$\frac{\varepsilon_1(C, \eta)}{a^{1+2\eta}} < \mu(\phi(C + y)).$$

We deduce that $\phi(x + y)$ belongs to the finite set

$$\phi(C + y) \cap \mathbb{O}_{\varepsilon_1(C, \eta)m^{-4n}a^{-1/(2n)}}.$$

The morphism $C + y \rightarrow \phi(C + y)$ has finite fiber. Since $\varepsilon \leq \varepsilon_1(C, \eta)/(gm^{4n})$, we conclude that for all ϕ Gauss-reduced of rank 2 with $a_0 = \lfloor a^{1/(2n)} \rfloor \leq m$, the set

$$(C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}})$$

is finite.

For the curve C , define

$$\varepsilon(C) = \min(\varepsilon_1(C, \eta_0), \varepsilon_2(C, \eta_0)).$$

Note that

$$\left(\frac{\varepsilon(C)}{gK_1}\right)^{8(g+s)g} \leq \frac{\varepsilon_1(C, \eta)}{gm^{4n}}.$$

Thus, we could for instance choose

$$\varepsilon_4 = \min\left(\varepsilon_1, \frac{K_1}{g}, \left(\frac{\varepsilon(C)}{gK_1}\right)^{8(g+s)g}\right).$$

Proof of (ii). We want to show that, for every $\tilde{\phi} = (N\phi|\phi')$ special of rank 2, there exists ϕ Gauss-reduced of rank 2 and $y \in \Gamma_0^2 \times \{0\}^{g-2}$ such that the map

$(x, \gamma) \rightarrow x + y$ defines an injection

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}) \hookrightarrow (C(\overline{\mathbb{Q}}) + y) \cap (B_{\phi} + \mathbb{O}_{(g+s)\varepsilon/H(\phi)^{1+1/(2n)}}). \tag{13-4}$$

We then apply part (i) of this proposition; if $(g + s)\varepsilon \leq \varepsilon_4$, then

$$(C(\overline{\mathbb{Q}}) + y) \cap (B_{\phi} + \mathbb{O}_{(g+s)\varepsilon/H(\phi)^{1+1/(2n)}})$$

is finite. So if $\varepsilon \leq \varepsilon_4/(g + s)$, the set

$$(C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}})$$

is finite too.

Let us prove the inclusion (13-4). Let $\tilde{\phi} = (N\phi|\phi')$ be special of rank 2. By definition of special $\phi = (aI_r|L)$ is Gauss-reduced of rank 2. Let

$$(x, \gamma) \in (C(\overline{\mathbb{Q}}) \times \gamma) \cap (B_{\tilde{\phi}} + \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}).$$

Then, there exists $\xi \in \mathbb{O}_{\varepsilon/H(\phi)^{1+1/(2n)}}$ such that

$$\tilde{\phi}((x, \gamma) + \xi) = 0.$$

Equivalently,

$$N\phi(x) + \phi'(\gamma) + \tilde{\phi}(\xi) = 0.$$

Let $y' \in E^2$ be a point such that

$$N[ay'] = \phi'(\gamma).$$

Since Γ_0 is a division group,

$$y = (y', 0, \dots, 0) \in \Gamma_0^2 \times \{0\}^{s-2}$$

and

$$N\phi(y) = N[ay'] = \phi'(\gamma).$$

Therefore

$$N\phi(x + y) + \tilde{\phi}(\xi) = 0.$$

Let $\xi'' \in E^2$ be a point such that

$$N[a]\xi'' = \tilde{\phi}(\xi).$$

We define $\xi' = (\xi'', \{0\}^{s-2})$. Then

$$N\phi(\xi') = N[a]\xi'' = \tilde{\phi}(\xi),$$

and

$$N\phi(x + y + \xi') = 0.$$

Since $\tilde{\phi}$ is special $H(\tilde{\phi}) = Na$. Furthermore $\|\xi\| \leq \varepsilon/a^{1+1/(2n)}$. We deduce

$$\|\xi'\| = \|\xi''\| = \frac{\|\tilde{\phi}(\xi)\|}{Na} \leq \frac{(g+s)\varepsilon}{a^{1+1/(2n)}}.$$

In conclusion

$$N\phi(x + y + \xi') = 0$$

with $\|\xi'\| \leq (g+s)\varepsilon/a^{1+1/(2n)}$ and $y \in \Gamma_0^2 \times \{0\}^{g-2}$. Equivalently,

$$(x + y) \in (C(\overline{\mathbb{Q}}) + y) \cap (B_{N\phi} + \mathbb{O}_{(g+s)\varepsilon/H(\phi)^{1+1/(2n)}}).$$

By Lemma 4.4(i), with $\psi = \phi$, we deduce that

$$(x + y) \in (C(\overline{\mathbb{Q}}) + y) \cap (B_\phi + \mathbb{O}_{(g+s)\varepsilon/H(\phi)^{1+1/(2n)}}),$$

with $y \in \Gamma_0^2 \times \{0\}^{g-2}$ and ϕ Gauss-reduced of rank 2.

This proves relation (13-4) and concludes the proof. □

14. The effectiveness aspect

An effective weak height bound. We give an effective bound for the height of $S_1(C, \mathbb{O}_\varepsilon)$ for C transverse.

Theorem 14.1. *Let C be transverse. For every real $\varepsilon \geq 0$, the norm of the set $S_1(C, \mathbb{O}_\varepsilon)$ is bounded by $K_0 \max(1, \varepsilon)$, where K_0 is an effective constant depending on the degree and the height of C .*

Proof. If $x \in S_1(C, \mathbb{O}_\varepsilon)$, there exist $\phi : E^g \rightarrow E$ and $\xi \in \mathbb{O}_\varepsilon$ such that $\phi(x - \xi) = 0$. Now the proof follows that of [Viada 2003, Theorem 1, p. 55], where we replace \hat{h} by h , y by ϕ , p by x and $\hat{h}(y(p)) = 0$ by $h(\phi(x)) = c_0(\deg \phi)h(\xi)$ with $h(\xi) \leq \varepsilon^2$. □

The strong hypotheses and an effective weak theorem.

Proof of Theorem 1.6. The proof is similar to the proof of Theorem 1.5 given in Section 11.

Theorem 14.1 implies that for $r \geq 1$ the norm of the set $S_r(C, \mathbb{O}_1)$ is bounded by an effective constant K_0 . Define

$$\eta_0 = \frac{1}{2^4 g^2},$$

$$\epsilon(C) = \min(\epsilon_1(C, \eta_0), \epsilon_2(C, \eta_0)), \quad \text{where } \epsilon_1, \epsilon_2 \text{ are as in Proposition 13.3,}$$

$$\delta_1 = \frac{1}{g} \min\left(1, \frac{K_0}{g}, \left(\frac{\epsilon(C)}{gK_0}\right)^{8g^2}\right),$$

$$\delta = \delta_1 M^{-1 - \frac{1}{2(2g-3)}}, \quad \text{where } M = \max\left(2, \left\lceil \frac{K_0}{\delta_1} \right\rceil^2\right)^{2g-3}.$$

In Section 12, Proposition A(i) with $\Gamma = 0$, $\varepsilon_1 = 1$ and $K_1 = K_0$, we have shown that

$$\bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk}(\phi)=2}} C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathbb{O}_\delta) \subset \bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk}(\phi)=2 \ H(\phi) \leq M}} C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathbb{O}_{g\delta_1/H(\phi)^{1+1/(2(2g-3))}}).$$

In Section 13, Proposition B(i) with $y = 0$, $s = 0$ and $n = 2g - 3$, $K_1 = K_0$, we have shown that for all ϕ Gauss-reduced of rank 2, the set

$$C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathbb{O}_{g\delta_1/H(\phi)^{1+1/(2(2g-3))}})$$

is finite. It follows that

$$\bigcup_{\substack{\phi \text{ Gauss-reduced} \\ \text{rk}(\phi)=2}} C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathbb{O}_\delta)$$

is finite. By Lemma 4.5(i) we deduce that $S_2(C, \mathbb{O}_\delta)$ is finite. This shows that Theorem 1.6 holds for

$$\varepsilon \leq \frac{1}{g^{4g}} \min(1, K_0^{-1})^{4g} \min\left(1, K_0, \left(\frac{\varepsilon(C)}{gK_0}\right)^{8g^2}\right)^{4g}. \quad \square$$

An effective bound for the cardinality of the sets. We have just shown that for C transverse, ε can be made effective. The purpose of this section is to indicate an effective bound for the cardinality of $S_2(C, \mathbb{O}_\varepsilon)$, under the following conjecture:

Conjecture 14.2 (S. David; personal communication). *Let C be a transverse curve in A . There exist constants c' and c'' , each depending on $g, \deg_L A, h_L(A), [k : \mathbb{Q}]$, such that, for*

$$\varepsilon(C) = \frac{c'}{(\deg_L V)^{1/(2 \text{codim } V)}} \quad \text{and} \quad \Theta(C) = c''(\deg_L C)^g,$$

the cardinality of $C(\overline{\mathbb{Q}}) \cap \mathbb{O}_{\varepsilon(C)}$ is bounded by $\Theta(C)$.

This is the abelian analogue to [Amoroso and David 2003, Conjecture 1.2].

Theorem 14.3. *Let C be transverse. Assume that Conjecture 14.2 holds. Then, there exists an effective $\varepsilon > 0$ such that the cardinality of $S_2(C, \mathbb{O}_\varepsilon)$ is bounded by an effective constant.*

Proof. Let δ and δ_1 be as defined in the previous proof.

By Proposition A(i) in Section 12 we deduce that

$$S_2(C, \mathbb{O}_\delta) \subset \bigcup_{\substack{\phi \text{ Gauss-reduced} \\ H(\phi) \leq M}} C(\overline{\mathbb{Q}}) \cap (B_\phi + \mathbb{O}_{g\delta_1/H(\phi)^{1+1/(2(2g-3))}}).$$

Note that, for any curve D and positive integers n , the cardinality of $[n]D \cap \mathbb{O}_{n\epsilon(D)}$ is still $\Theta(D)$. Going through the proofs of Proposition B(i) in Section 13, we see that

$$\#S_2(C, \mathbb{O}_\delta) \leq \sum_{H(\phi) \leq M} \#(\phi|_C^{-1}(\phi(C) \cap \mathbb{O}_{\epsilon(\phi(C))})),$$

where $\phi|_C : C \rightarrow \phi(C)$ is the restriction of ϕ to C . Recall from [Viada 2003, p. 61] that the fiber of $\phi|_C$ has cardinality at most $3gH(\phi)^2 \leq 3gM^2$. We set

$$\Delta_{\max} = \max_{H(\phi) \leq M} \#(\phi(C) \cap \mathbb{O}_{\epsilon(\phi(C))}).$$

We deduce

$$\#S_2(C, \mathbb{O}_\delta) \leq 3gM^3 \Delta_{\max}.$$

By Lemma 13.2(i), $\deg \phi(C) \leq (3gH(\phi))^2 \deg C$. Conjecture 14.2 implies that

$$\Delta_{\max} \leq (3gH(\phi))^{2g} \Theta(C),$$

with $\Theta(C)$ explicitly given. We conclude that

$$\#S_2(C, \mathbb{O}_\delta) \leq (3g)^{2g+1} M^{2g+3} \Theta(C). \tag{14-1}$$

By Theorem 14.1 the constant K_0 is effective. So M is also effective. Thus the bound (14-1) is effective, for C transverse. □

Similar computations imply a bound for the cardinality of $S_2(C, \Gamma_\delta)$.

For $\delta \leq \varepsilon_4(g+s)^{-2} M'^{-1-1/(4g+4s-6)}$ we obtain

$$\#S_2(C, \Gamma_\delta) \leq c_1(g) M'^{c_2(g,s)} \Theta(C).$$

Here $c_1(g)$ (and $c_2(g, s)$) are effective constants depending only on g (and s). The number M' depends explicitly on C, g and K_2 , while ε_4 depends explicitly on C, g, s and K_1 . In view of Theorem 9.1, the above bound also implies a bound for the cardinality of $S_2(C \times \gamma, \mathbb{O}_{\delta/(g+s)K_4})$.

However, Theorem 1.2 does not give effective K_1 or K_2 . Consequently neither M' nor ε_4 are effective. An effective estimate for K_1 or K_2 would imply an effective Mordell Conjecture. This gives an indication of the difficulty to extend effective height proofs from transverse curves to weak-transverse curves.

15. Final remarks

The CM case. The proofs in 2–7 hold whether or not E has CM. Since Conjecture 1.4 is stated for any E , Proposition B holds unchanged for E with CM.

We can extend Proposition A to Gauss-reduced $\phi \in M_{r,g}(\mathbb{Z} + \tau\mathbb{Z})$ as follows. Decompose $\phi = \phi_1 + \tau\phi_2$ for $\phi_i \in M_{r,g}(\mathbb{Z})$, then let the morphism $\psi = (\phi_1|\phi_2)$ act on $(x, \tau x) + (y, \tau y) + (\xi, \tau\xi)$ for $x \in S_r(C, (\Gamma_0^g)_\varepsilon)$, $y \in \Gamma_0^g$ and $\xi \in \mathbb{O}_\varepsilon$. Apply Proposition A to ψ . Constants will depend on τ .

From powers to products. In a power there are more algebraic subgroups than in a product where not all the factors are isogenous. If we consider a product of non-CM elliptic curves, then the matrix of a morphism ϕ is simply an integral matrix where the entries corresponding to nonisogenous factors are zeros. So nothing changes with respect to our proofs.

If the curve is in a product of elliptic curves in general, we extend the definition of Gauss-reduced, introducing constants $c_1(\tau)$ and $c_2(\tau)$, such that the element a on the diagonal has norm satisfying $c_1(\tau)H(\phi) \leq |a| \leq c_2(\tau)H(\phi)$.

We leave the details to the reader.

Acknowledgments

My deepest thanks go to the referee for his precious comments, which encouraged me to drastically improve a first manuscript.

I kindly thank Francesco Amoroso and Özlem Imamoglu for inspiring me to pursue research in mathematics.

I thank the Centro de Giorgi and in particular Umberto Zannier for inviting me to the Diophantine Geometry Semester in 2005. I gave there a talk where the core material of this paper was presented.

I thank the SNF for the financial support which allows me to work serenely near my family.

It is to my husband Robert that I am most thankful; without him I would have given up long ago and I would not know how nice it is to be happy.

References

- [Amoroso and David 1999] F. Amoroso and S. David, “Le problème de Lehmer en dimension supérieure”, *J. Reine Angew. Math.* **513** (1999), 145–179. MR 2001a:11116 Zbl 1011.11045
- [Amoroso and David 2003] F. Amoroso and S. David, “Minoration de la hauteur normalisée dans un tore”, *J. Inst. Math. Jussieu* **2**:3 (2003), 335–381. MR 2004m:11101 Zbl 1041.11048
- [Bombieri et al. 1999] E. Bombieri, D. Masser, and U. Zannier, “Intersecting a curve with algebraic subgroups of multiplicative groups”, *Internat. Math. Res. Notices* **20** (1999), 1119–1140. MR 2001c:11081 Zbl 0938.11031
- [David and Hindry 2000] S. David and M. Hindry, “Minoration de la hauteur de Néron-Tate sur les variétés abéliennes de type C. M”, *J. Reine Angew. Math.* **529** (2000), 1–74. MR 2001j:11054 Zbl 0993.11034
- [David and Philippon 2002] S. David and P. Philippon, “Minorations des hauteurs normalisées des sous-variétés de variétés abéliennes, II”, *Comment. Math. Helv.* **77**:4 (2002), 639–700. MR 2004a:11055 Zbl 1030.11026
- [David and Philippon 2007] S. David and P. Philippon, “Minorations des hauteurs normalisées des sous-variétés des puissances des courbes elliptiques”, *Int. Math. Res. Pap.* **3** (2007), ID rpm006. MR 2355454 Zbl 05238017
- [Galateau 2007] A. Galateau, *Problème de Bogomolov sur les variétés abéliennes*, Ph.D. thesis, Université de Paris VI, 2007, Available at <http://tel.archives-ouvertes.fr/tel-00203583>.

- [Habegger 2007] P. Habegger, *Heights and multiplicative relations on algebraic varieties*, Ph.D. thesis, Universität Basel, 2007, Available at http://pages.unibas.ch/diss/2007/DissB_7940.htm.
- [Hindry 1988] M. Hindry, “Autour d’une conjecture de Serge Lang”, *Invent. Math.* **94**:3 (1988), 575–603. MR 89k:11046 Zbl 0638.14026
- [Poonen 1999] B. Poonen, “Mordell–Lang plus Bogomolov”, *Invent. Math.* **137**:2 (1999), 413–425. MR 2001c:11070 Zbl 0995.11040
- [Rémond and Viada 2003] G. Rémond and E. Viada, “Problème de Mordell–Lang modulo certaines sous-variétés abéliennes”, *Int. Math. Res. Not.* 35 (2003), 1915–1931. MR 2004h:11054 Zbl 1072.11038
- [Schlickewei 1997] H. P. Schlickewei, “Lower bounds for heights on finitely generated groups”, *Monatsh. Math.* **123**:2 (1997), 171–178. MR 98d:11077 Zbl 0973.11067
- [Schmidt 1980] W. M. Schmidt, *Diophantine approximation*, Lecture Notes in Mathematics **785**, Springer, Berlin, 1980. MR 81j:10038 Zbl 0421.10019
- [Ullmo 1998] E. Ullmo, “Positivité et discrétion des points algébriques des courbes”, *Ann. of Math.* (2) **147**:1 (1998), 167–179. MR 99e:14031 Zbl 0934.14013
- [Viada 2003] E. Viada, “The intersection of a curve with algebraic subgroups in a product of elliptic curves”, *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)* **2**:1 (2003), 47–75. MR 2004c:11099 Zbl 05019603
- [Zhang 1998] S.-W. Zhang, “Equidistribution of small points on abelian varieties”, *Ann. of Math.* (2) **147**:1 (1998), 159–165. MR 99e:14032 Zbl 0991.11034

Communicated by Bjorn Poonen

Received 2007-04-12

Revised 2008-04-02

Accepted 2008-04-04

viada@math.ethz.ch

*Université de Fribourg Suisse, Pérolles,
Département de Mathématiques, 23 Chemin du Musée,
CH-1700 Fribourg, Suisse*

Diophantine subsets of function fields of curves

János Kollár

We consider diophantine subsets of function fields of curves and show, roughly speaking, that they are either very small or very large. In particular, this implies that the ring of polynomials $k[t]$ is not a diophantine subset of the field of rational functions $k(t)$ for many fields k .

Let R be a commutative ring. A subset $D \subset R$ is called *diophantine* if there are polynomials

$$F_i(x, y_1, \dots, y_n) \in R[x, y_1, \dots, y_n]$$

such that the system of equations

$$F_i(r, y_1, \dots, y_n) = 0 \quad \forall i$$

has a solution $(y_1, \dots, y_n) \in R^n$ if and only if $r \in D$.

Equivalently, if there is a (possibly reducible) algebraic variety X_R over R and a morphism $\pi : X_R \rightarrow \mathbb{A}^1_R$ such that $D = \pi(X_R(R))$. In this situation we call

$$\text{dioph}(X_R, \pi) := \pi(X_R(R)) \subset R$$

the *diophantine set* corresponding to X_R and π .

A characterization of diophantine subsets of \mathbb{Z} was completed in connection with Hilbert's 10th problem, but a description of diophantine subsets of \mathbb{Q} is still not known. In particular, it is not known if \mathbb{Z} is a diophantine subset of \mathbb{Q} or not. (See [Poonen 2003] or the volume [Denef et al. 2000] for surveys and many recent results.)

In this paper we consider analogous questions where $R = k(t)$ is a function field of one variable and k is an uncountable *large field* of characteristic 0. That is, for any k -variety Y with a smooth k -point, $Y(k)$ is Zariski dense. Examples of uncountable large fields are

- (1) \mathbb{C} or any uncountable algebraically closed field,
- (2) \mathbb{R} or any uncountable real closed field, and

MSC2000: primary 11U05; secondary 14G25, 14M20, 14G27.

Keywords: diophantine set, rationally connected variety.

Partial financial support was provided by the NSF under grant number DMS-0500198.

(3) $\mathbb{Q}_p, \mathbb{Q}((x))$ or the quotient field of any uncountable local Henselian domain.

Roughly speaking, we show that for such fields, a diophantine subset of $k(t)$ is either very small or very large. The precise result is somewhat technical, but here are two easy-to-state consequences which serve as motivating examples.

Corollary 1. *Let k be an uncountable large field of characteristic 0. Then $k[t]$ is not a diophantine subset of $k(t)$.*

Corollary 2. *Let k be an uncountable large field of characteristic 0 and*

$$K_2 \supset K_1 \supset k(t)$$

finite field extensions. Then K_1 is a diophantine subset of K_2 if and only if $K_1 = K_2$.

The latter gives a partial answer to a question of Bogomolov: When is a subfield $K_1 \subset K_2$ diophantine in K_2 ?

It is possible that both of these corollaries hold for any field k . Unfortunately, my method says nothing about countable fields. The geometric parts of the proof for Propositions 12 and 13 (see 16 and 20) work for any uncountable field, but the last step 23 uses in an essential way that k is large.

We use two ways to measure how large a diophantine set is.

3 (Diophantine dimension and polar sets). Let B be a smooth, projective, irreducible curve over k . One can think of a rational function $f \in k(B)$ as a section of the first projection $\pi_1 : B \times \mathbb{P}^1 \rightarrow B$. This establishes a one-to-one correspondence

$$k(B) \cup \{\infty\} \leftrightarrow \{ \text{sections of } \pi_1 : B \times \mathbb{P}^1 \rightarrow B \}.$$

Any section $\sigma : B \rightarrow B \times \mathbb{P}^1$ can be identified with its image, which gives a point in the Chow variety of curves of $B \times \mathbb{P}^1$. This gives an injection

$$k(B) \cup \{\infty\} \hookrightarrow \text{Chow}_1(B \times \mathbb{P}^1).$$

Let U be a countable (disjoint) union of k -varieties and $D \subset U(k)$ a subset. Define the *diophantine dimension* of D over k as the smallest $n \in \{-1, 0, 1, \dots, \infty\}$ such that D is contained in a countable union of irreducible k -subvarieties of U of dimension $\leq n$. It is denoted by $\text{d-dim}_k D$. Note that $\text{d-dim}_k D = -1$ if and only if $D = \emptyset$ and $\text{d-dim}_k D \leq 0$ if and only if D is countable.

In particular, we can talk about the diophantine dimension of

$$\text{dioph}(X, f) \subset k(B) \subset \text{Chow}_1(B \times \mathbb{P}^1).$$

For $f \in k(B)$, let $\text{pole}(f)$ denote its divisor of poles. For $D \subset k(B)$, set

$$\text{Pole}_n(D) := \{ \text{pole}(f) : f \in D \text{ and } \deg \text{pole}(f) = n \}.$$

I think of $\text{Pole}_n(D)$ as a subset of the \bar{k} -points of the n -th symmetric power $S^n B$.

Taking each point with multiplicity $r \geq 1$ gives embeddings $S^m B \rightarrow S^{rm} B$, whose image I denote by $r \cdot S^m B$.

With these definitions, the main result is the following theorem illustrating the “very small or very large” dichotomy.

Theorem 4. *Let k be an uncountable large field of characteristic 0 and B a smooth, projective, irreducible curve over k . Let $X_{k(B)}$ be a (possibly reducible) algebraic variety of dimension n over $k(B)$ and $\pi_{k(B)} : X_{k(B)} \rightarrow \mathbb{A}^1_{k(B)}$ a morphism. Then either*

- (1) $\text{d-dim}_k \text{dioph}(X_{k(B)}, \pi_{k(B)}) \leq n$, or
- (2) $\text{d-dim}_k \text{dioph}(X_{k(B)}, \pi_{k(B)}) = \infty$ and there is a 0-cycle $P_a \in S^a B$ and $r > 0$ such that for every $m > 0$ there is a smooth, irreducible k -variety D_m and a morphism $\rho_m : D_m \rightarrow S^{a+rm} B$ such that
 - (a) $D_m(k) \neq \emptyset$,
 - (b) $\text{Pole}_{a+rm}(\text{dioph}(X_{k(B)}, \pi_{k(B)})) \supset \rho_m(D_m(k))$, and
 - (c) the Zariski closure of $\rho_m(D_m(k))$ contains $P_a + r \cdot S^m B \subset S^{a+rm} B$.

5 (Proof of the Corollaries). In trying to write a subset $D \subset k(B)$ as

$$D = \text{dioph}(X_{k(B)}, \pi_{k(B)}),$$

we do not have an a priori bound on $\dim X_{k(B)}$; thus the assertion

$$\text{d-dim}_k \text{dioph}(X_{k(B)}, \pi_{k(B)}) = \infty$$

is hard to use. The Corollaries 1 and 2 both follow from the more precise results about the distribution of poles.

If $B = \mathbb{P}^1$, then a rational function with at least 2 poles on \mathbb{P}^1 is not a polynomial; thus Theorem 4 implies Corollary 1.

Next consider Corollary 2. Let $K_1 = k(B_1) \subsetneq K_2 = k(B_2)$ be a degree $d > 1$ extension of function fields of smooth, projective, irreducible k -curves. By Riemann–Roch, any zero cycle of degree $\geq 2g(B_1)$ defined over k is the polar set of some $f \in k(B_1)$. Pulling back gives a map $j : S^m B_1 \rightarrow S^{md} B_2$; thus

$$\text{Pole}_n(K_1) = \begin{cases} j((S^m B_1)(k)) & \text{if } n = md \geq 2dg(B_1), \text{ and} \\ \emptyset & \text{if } d \nmid n. \end{cases}$$

If $b_1 \neq b_2 \in B_2$ map to the same point of B_1 , then a 0-cycle in $j(S^m B_1)$ contains either both b_1 and b_2 or neither. Thus the Zariski closed set $j(S^m B_1)$ never contains a set of the form $P_a + r \cdot S^m B_2$. By (2.c) of Theorem 4, this shows that K_1 is not diophantine in K_2 , proving Corollary 2.

Example 6. (1) The bound n in (1) of Theorem 4 is actually sharp, as shown by the following.

Note first that any $k(t)$ -solution of $x^3 + y^3 = 1$ is constant. Set

$$X_n := (x_1^3 + y_1^3 = \cdots = x_n^3 + y_n^3 = 1) \subset \mathbb{A}^{2n}$$

and

$$\pi : (x_1, y_1, \dots, x_n, y_n) \mapsto x_1 + x_2 t + \cdots + x_n t^{n-1}.$$

Then $\dim X = n$ and for $k = \mathbb{C}$ or $k = \mathbb{R}$, $\text{dioph}(X_n, \pi)$ is the set of all degree $\leq n - 1$ polynomials.

Using similar constructions one can see that any (finite dimensional) k -algebraic subset of $k(t)$ is diophantine when k is algebraically closed or real closed. These are the “small” diophantine subsets of $k(t)$.

(2) The somewhat unusual looking condition about the Zariski closure of D_m in (2.c) of Theorem 4 is also close to being optimal. For $g \in k(t)$ and $r > 0$ consider the diophantine set

$$L_{g,r} := \{f \in k(t) : \exists h \text{ such that } f = gh^r\}.$$

Then, up to some lower dimensional contribution coming from possible cancellations between poles and zeros of g and h^r ,

$$\text{Pole}_n(L_{g,r}) = \begin{cases} \text{pole}(g) + r \cdot (S^m B)(k) & \text{if } n = \text{deg pole}(g) + rm, \text{ and} \\ \emptyset & \text{otherwise.} \end{cases}$$

7. If $k = \mathbb{C}$ then our proof shows that in case (2) of Theorem 4 there is a finite set $P \subset B(\mathbb{C})$ such that for every $p \in B(\mathbb{C}) \setminus P$ there is an $f_p \in \text{dioph}(X, \pi)$ with a pole at p .

If $k = \mathbb{R}$, then we guarantee many poles, but one may not get any real poles. To get examples, note that $h \in \mathbb{R}(t)$ is everywhere nonnegative on \mathbb{R} if and only if h is a sum of two squares. Thus for any $g \in \mathbb{R}(t)$, the set

$$L_1(g) := \{f \in \mathbb{R}(t) : f(t) \leq g(t) \forall t \in \mathbb{R}\}$$

is diophantine. $L_1(g)$ is infinite dimensional but if $g \in \mathbb{R}[t]$ then no element of $L_1(g)$ has a real pole.

From the point of view of our proof a more interesting example is the diophantine set

$$L_2(g) := \{f \in \mathbb{R}(t) : \exists c \in \mathbb{R}, f^2(t) \leq c^2 \cdot g^2(t) \forall t \in \mathbb{R}\}.$$

The elements of $L_2(g)$ are unbounded everywhere yet no element of $L_2(g)$ has a pole in \mathbb{R} if g is a polynomial.

This leads to the following question.

Question 8. Is $\mathbb{R}[t]_{\mathbb{R}}$, the set of all rational functions without poles in \mathbb{R} , diophantine?

There should be some even stronger variants of the “very small or very large” dichotomy, especially over \mathbb{C} . As a representative case, I propose the following.

Conjecture 9. *Let $D \subset \mathbb{C}(t)$ be a diophantine subset which contains a Zariski open subset of $\mathbb{C}[t]$. (Meaning, for instance, that D contains a Zariski open subset of the space of degree $\leq n$ polynomials for infinitely many n .) Then $\mathbb{C}(t) \setminus D$ is finite.*

In connection with Bogomolov’s question, I would hazard the following:

Conjecture 10. *Let k be a large field and $K_1 \subset K_2$ function fields of k -varieties. Then K_1 is diophantine in K_2 if and only if K_1 is algebraically closed in K_2 .*

11. The proof of Theorem 4 relies on the theory of rational curves on algebraic varieties. A standard reference is [Kollár 1996], but nonexperts may prefer the more introductory [Araujo and Kollár 2003].

The proof is divided into three steps.

First we show that if $d\text{-dim}_k \text{dioph}(X_{k(B)}, \pi) \geq n + 1$ then there is a rationally connected (see 18) subvariety $Z_{k(B)} \subset X_{k(B)}$ such that $\pi|_Z$ is nonconstant and $Z_{k(B)}$ has a smooth $k(B)$ -point. This relies on the *bend and break* method of [Mori 1979]. In a similar context it was first used in [Graber et al. 2005].

Then we show, using the *deformation of combs* technique developed in [Kollár et al. 1992; Kollár 1996; 2004; Graber et al. 2003], that for any such $Z_{k(B)}$, there are infinitely many k -varieties S_m and maps $S_m \times B \dashrightarrow Z_{k(B)}$ which give injections $S_m(k) \hookrightarrow Z_{k(B)}(k(B))$.

Both of these steps are geometric, but the statements are formulated to work over an arbitrary field L .

Finally, if k is a large field, then each $S_m(k)$ is “large”, which shows that $Z_{k(B)}(k(B))$ is “very large”.

For all three steps it is better to replace $\pi : X_{k(B)} \rightarrow \mathbb{A}^1_{k(B)}$ with a morphism of k -varieties $f : X \rightarrow B \times \mathbb{P}^1$.

Proposition 12. *Let L be any field and B a smooth, projective, irreducible curve over L . Let $f : X \rightarrow B \times \mathbb{P}^1$ be an L -variety of dimension $n + 1$ and consider the corresponding diophantine set $\text{dioph}(X_{L(B)}, f) \subset L(B)$. Then either*

- (1) $d\text{-dim}_L \text{dioph}(X_{L(B)}, f) \leq n$, or
- (2) *there is a subvariety $Z \subset X$ such that*
 - (a) $Z \rightarrow B \times \mathbb{P}^1$ *is dominant,*
 - (b) *the generic fiber of $Z \rightarrow B$ is rationally connected, and*
 - (c) *there is a rational section $\sigma : B \dashrightarrow Z$ whose image is not contained in $\text{Sing } Z$.*

Proposition 13. *Let L be an infinite field and B a smooth, projective, irreducible curve over L . Let $f : Z \rightarrow B \times \mathbb{P}^1$ be a smooth, projective L -variety such that*

- (1) $Z \rightarrow B \times \mathbb{P}^1$ is dominant,
- (2) the generic fiber of $Z \rightarrow B$ is separably rationally connected, and
- (3) there is a section $\sigma : B \rightarrow Z$.

Then, for some $r > 0$ and for all $m > 0$ in an arithmetic progression, there are

- (4) a smooth, irreducible L -variety S_m with an L -point, and
- (5) a dominant rational map $\sigma_m : S_m \times B \dashrightarrow Z$ which commutes with projection to B ,

such that the Zariski closure of the image of $f \circ \sigma_m : S_m \dashrightarrow \text{Chow}_1(B \times \mathbb{P}^1)$ contains

$$[f \circ \sigma(B)] + r[\{b_1\} \times \mathbb{P}^1] + \cdots + r[\{b_m\} \times \mathbb{P}^1] \quad \text{for every } b_i \in B(\bar{L}).$$

14 (Spaces of sections). Let L be any field, B a smooth, projective, irreducible curve over L and $f : X \rightarrow B$ a projective morphism. A section of f (defined over some $L' \supset L$) can be identified with the corresponding L' -point in the Chow variety of 1-cycles $\text{Chow}_1(X)$. All sections $\Sigma(X/B)$ defined over \bar{L} form an open set of $\text{Chow}_1(X)$. Indeed, if H is an ample line bundle on B of degree d then a 1-cycle C is a section if and only if C is irreducible (an open condition) and $(C \cdot f^*H) = d$ (an open and closed condition). This procedure realizes $X_{k(B)}(k(B))$ as the set of k -points of a countable union of algebraic k -varieties $\Sigma(X/B) = \bigcup_i \Sigma_i$.

The choice of the Σ_i is not canonical. Given $X \rightarrow B$, we get “natural” irreducible components, but for fixed generic fiber $X_{k(B)}$, these components depend on the choice of X . Any representation gives, however, the same constructible sets. We usually make a further decomposition. Since every variety is a finite set-theoretic union of locally closed smooth subvarieties, we may choose the Σ_i such that each one is smooth and irreducible.

As an explicit example, consider $B = \mathbb{P}^1$. Then $k(B) \cong k(t)$ and every $f \in k(t)$ can be uniquely written (up to scalars) as

$$f = \frac{a_0 + a_1 t + \cdots + a_n t^n}{b_0 + b_1 t + \cdots + b_n t^n},$$

where the nominator and the denominator are relatively prime and at least one of a_n or b_n is nonzero. For any n , all such f form an open subset

$$\Sigma_n \subset \mathbb{P}(a_0 : a_1 : \cdots : a_n : b_0 : b_1 : \cdots : b_n) \cong \mathbb{P}^{2n+1}.$$

15 (Very dense subsets). Let U be an irreducible variety over a field L . We say that a subset $D \subset U(\bar{L})$ is *Zariski very dense* if D is not contained in a countable union of L -subvarieties $V_i \subsetneq U$.

It is easy to see that for any D , there are countably many closed, irreducible L -subvarieties $W_i \subset U$ such that $D \subset \bigcup_i W_i(\bar{L})$ and $D \cap W_i(\bar{L})$ is Zariski very dense in W_i for every i . There is a unique irredundant choice of these W_i .

16 (Proof of Proposition 12). Write $X = \bigcup X_i$ as a finite set-theoretic union of locally closed, smooth, connected varieties. If Proposition 12 holds for each X_i then it also holds for X ; thus we may assume that X is smooth and irreducible. Let $X' \supset X$ be a smooth compactification such that f extends to $f' : X' \rightarrow B \times \mathbb{P}^1$.

As before, there are countably many disjoint, irreducible, smooth L -varieties $\bigcup_i \Sigma_i = \Sigma(X'/B)$ and morphisms $u_i : B \times \Sigma_i \rightarrow X'$ commuting with projection to B giving all \bar{L} -sections of f' . As in 15, there are countably many disjoint, irreducible, smooth L -varieties $S_i \subset \Sigma(X'/B)$ such that each $S_i(L)$ is Zariski very dense in S_i and the L -sections of $X' \rightarrow B$ are exactly given by $\bigcup_i S_i(L)$.

Composing u_i with f' , we obtain maps

$$f'_* : S_i \rightarrow \Sigma \subset \text{Chow}_1(B \times \mathbb{P}^1).$$

There are 2 distinct possibilities. Either

- (1) $\dim_L f'_*(S_i) \leq n$ whenever $u_i(B \times S_i) \cap X \neq \emptyset$, or
- (2) there is an i_0 such that $\dim_L f'_*(S_{i_0}) \geq n + 1$ and $u_{i_0}(B \times S_{i_0}) \cap X \neq \emptyset$.

In the first case $\text{dioph}(X', f')$ is contained in the union of the constructible sets $f'_*(S_i)$, thus we have (1) of Proposition 12. This is always the case if L is countable.

In the second case we construct Z as required by (2) of Proposition 12 using only the existence of $u_{i_0} : B \times S_{i_0} \rightarrow X$. Set $S := S_{i_0}$ and $u := u_{i_0}$. We can replace X' by a desingularization of the closure of the image $u(B \times S)$. By shrinking S we may assume that u lifts to $u : B \times S \rightarrow X'$.

For a point $x \in X'$ let $S_x \subset S$ be the subvariety parametrizing those sections that pass through x .

Let us now fix $b \in B(\bar{L})$ such that $u(\{b\} \times S)$ is dense in X'_b and let x run through X'_b , the fiber of X' over b . Since every section intersects X'_b ,

$$S = \bigcup_{x \in X'_b} S_x \quad \text{and so} \quad f'_*(S) = \bigcup_{x \in X'_b} f'_*(S_x).$$

By assumption $\dim_L f'_*(S) \geq n + 1$ and $\dim_L X'_b = n$; hence $\dim_L f'_*(S_x) \geq 1$ for general $x \in X'_b(\bar{L})$. In particular, there is a 1-parameter family of sections $C_x \subset S_x$ such that

$$f' \circ u : B \times C \rightarrow X \rightarrow B \times \mathbb{P}^1$$

is a nonconstant family of sections passing through the point $f'(x)$.

By 17, this leads to a limit 1-cycle of the form

$$A + \{b\} \times \mathbb{P}^1 + (\text{other fibers of } \pi_1)$$

where A is a section of $\pi_1 : B \times \mathbb{P}^1 \rightarrow B$.

Correspondingly, we get a limit 1-cycle in X' of the form

$$A_x + R_x + (\text{other rational curves})$$

where A_x is a section of $X' \rightarrow B$ which dominates A and R_x is a connected union of rational curves which dominates $\{b\} \times \mathbb{P}^1$. Note also that $x \in R_x$.

Thus we conclude that for general $x \in X'_b(\bar{L})$, there is a connected union of rational curves $x \in R_x \subset X'_b$ which dominates $\{b\} \times \mathbb{P}^1$.

As in 19, let us take the relative MRC-fibration $f' : X' \dashrightarrow W' \dashrightarrow B$.

For very general $x \in X'(\bar{L})$ let X'_x be the fiber of w containing x . By 19, X'_x is closed in X' and every rational curve in X' that intersects X'_x is contained in X'_x . In particular, $R_x \subset X'_x$ and hence X'_x dominates $\{b\} \times \mathbb{P}^1$.

Let now $p \in S(L)$ be a general point and $C \subset X'$ the corresponding section. By assumption $S(L)$ is Zariski dense in S ; hence we may assume that w is smooth at the generic point of C . Let $Z' \subset w^{-1}(w(C))$ be the unique irreducible component that dominates C and $Z = Z' \cap X$. It satisfies all the required properties. \square

17 (Bend-and-break for sections [Mori 1979; Kollár 1996, § II.5; Graber et al. 2005, Lemma 3.2]). Let $h : Y \rightarrow B$ be a proper morphism onto a smooth projective curve B . Let C be a smooth curve and $u : B \times C \rightarrow Y$ a nonconstant family of sections passing through a fixed point $y \in Y$.

Then C can not be a proper curve and for a suitable point $c \in \bar{C} \setminus C$ the corresponding limit 1-cycle is of the form

$$\Sigma_y = A_y + R_y,$$

where A_y is a section of h (which need not pass through y) and R_y is a nonempty union of rational curves contained in finitely many fibers of h . Furthermore, $A_y + R_y$ is connected and $y \in R_y$.

This holds whether we take the limit in the Chow variety of 1-cycles, in the Hilbert scheme or in the space of stable maps.

18 (Rationally connected varieties [Kollár et al. 1992; Kollár 1996, Chapter IV; Araujo and Kollár 2003, Section 7]). Let k be a field and $K \supset k$ an uncountable algebraically closed field. A smooth projective k -variety X is called *rationally connected* or *RC* if for every pair of points $x_1, x_2 \in X(K)$ there is a K -morphism $f : \mathbb{P}^1 \rightarrow X$ such that $f(0) = x_1$ and $f(\infty) = x_2$. X is called *separably rationally connected* or *SRC* if for every point $x \in X(K)$ there is a K -morphism $f : \mathbb{P}^1 \rightarrow X$

such that $f(0) = x$ and f^*T_X is an ample vector bundle. (That is, a sum of positive degree line bundles.) Furthermore, $f : \mathbb{P}^1 \rightarrow X$ can be taken to be an embedding if $\dim X \geq 3$. It is known that SRC implies RC and the two notions are equivalent in characteristic 0.

We may not have any rational curves over k , but we can work with the universal family of these maps $f : \mathbb{P}^1 \rightarrow X$. Thus, if $\dim X \geq 3$ and $p \in X$ is a k -point, then there is an irreducible, smooth k -variety U and a k -morphism $G : U \times \mathbb{P}^1 \rightarrow X$ such that

- (1) $G(U \times \{0\}) = p$,
- (2) $G_u : \{u\} \times \mathbb{P}^1 \rightarrow X_{\bar{k}}$ is an embedding for every $u \in U(\bar{k})$, and
- (3) $G^*_u T_X$ is ample for every $u \in U(\bar{k})$.

By [Kollár 1999, Theorem 1.4], if k is large then we can choose U such that $U(k) \neq \emptyset$.

19 (MRC fibrations [Kollár et al. 1992; Kollár 1996, §IV.5]). Let $K \supset k$ be as above. Let X be a smooth projective k -variety and $g : X \rightarrow S$ a k -morphism. There is a unique (up to birational maps) factorization

$$g : X \overset{w}{\dashrightarrow} W \overset{h}{\dashrightarrow} S$$

such that

- (1) for general $p \in W(K)$, the fiber $w^{-1}(p)$ is closed in X and rationally connected, and
- (2) for very general $p \in W(K)$ (that is, for p in a countable intersection of dense open subsets) every rational curve in $X(K)$ which intersects $w^{-1}(p)$ and maps to a point in S is contained in $w^{-1}(p)$.

The map $w : X \dashrightarrow W$ is called the (relative) *maximal rationally connected fibration* or *MRC fibration* of $X \rightarrow S$. Note that if X contains very few rational curves (for example, if X is an Abelian variety or a K3 surface) then $X = W$.

20 (Proof of Proposition 13). Here we essentially reverse the procedure of the first part. Instead of degenerating a 1-parameter family of sections to get a 1-cycle consisting of a section plus rational curves, we start with a section, add to it suitably chosen rational curves and prove that this 1-cycle can be written as the limit of sections in many different ways.

We assume that Z is smooth, projective. If necessary, we take its product with \mathbb{P}^3 to achieve that $\dim Z \geq 4$. This changes the space of sections $\Sigma(Z/B)$ but it does not change the image of $\Sigma(Z/B)$ in $L(B)$.

Apply 18 to $X = Z_{L(B)}$ and the point $p = \sigma(B)$ to get

$$G : U_{L(B)} \times \mathbb{P}^1 \rightarrow Z_{L(B)}.$$

Next replace $U_{L(B)}$ by an L -variety $\tau : U \rightarrow B$ such that G extends to

$$g : U \times \mathbb{P}^1 \rightarrow Z.$$

By shrinking U if necessary, we may assume that for general $b \in B(\bar{L})$, the corresponding

$$g_b : U_b \times \mathbb{P}^1 \rightarrow Z_b$$

is a family of smooth rational curves passing through $\sigma(b)$ and $g^*_{b,u}T_{Z_b}$ is ample for every $u \in U_b$ where $g_{b,u}$ is the restriction of g_b to $\{u\} \times \mathbb{P}^1$.

Given distinct points $b_1, \dots, b_m \in B(\bar{L})$, let $B(b_1, \dots, b_m)$ be the comb assembled from B and m copies of \mathbb{P}^1 where we attach \mathbb{P}^1_i to B at b_i (see 21).

By [Kollár 2004, Theorem 16], there are $b_1, \dots, b_{m_0} \in B(\bar{L})$ and an embedding

$$\sigma(g_1, \dots, g_{m_0}) : B(b_1, \dots, b_{m_0}) \rightarrow Z$$

given by σ on B and by $g_i := g_{b_i, u_i}$ on \mathbb{P}^1_i for some $u_i \in U_{b_i}$ such that the image, denoted by $B(g_1, \dots, g_{m_0}) \subset Z$ is defined over L and its normal bundle is as positive as one wants. In particular, by 22, $B(g_1, \dots, g_{m_0})$ gives a smooth point of the Hilbert scheme of Z . Furthermore, for any further distinct points b_{m_0+1}, \dots, b_m and g_i for $i = m_0 + 1, \dots, m$, the resulting

$$\sigma(g_1, \dots, g_m) : B(b_1, \dots, b_m) \rightarrow Z$$

also gives a smooth point of the Hilbert scheme of Z .

Let S_m denote the smooth locus of the corresponding L -irreducible component of the Hilbert scheme of Z . $B(g_1, \dots, g_m)$ gives an L -point of S_m ; hence S_m is geometrically irreducible. By 22 the general point of S_m corresponds to a section of f , the universal family is a product over a dense open subset of S_m and we have a dominant rational map $\sigma_m : S_m \times B \dashrightarrow Z$.

For a given m , it is not always possible to choose b_{m_0+1}, \dots, b_m such that the set b_1, \dots, b_m is defined over L . To achieve this, choose a generically finite and dominant map $\rho : U \dashrightarrow \mathbb{A}_L^{\dim U}$. For general $c \in \mathbb{A}^{\dim U}(L)$, its preimage $\rho^{-1}(c)$ gives $\deg \rho$ general points in U which are defined over L . Thus we can choose b_1, \dots, b_m to be defined over L whenever $m - m_0$ is a multiple of $\deg \rho$.

Let us now consider $f_*(S_m) \subset \text{Chow}_1(B \times \mathbb{P}^1)$. By construction, it contains

$$f_*(B(g_1, \dots, g_m)) = A + r(\{b_1\} \times \mathbb{P}^1) + \dots + r(\{b_m\} \times \mathbb{P}^1)$$

where $A = f \circ \sigma(B)$ is a section of $B \times \mathbb{P}^1$ and $r \geq 1$ is the common (geometric) degree of the maps

$$f \circ g_{b,u} : \{u\} \times \mathbb{P}^1 \rightarrow \{b\} \times \mathbb{P}^1 \subset B \times \mathbb{P}^1.$$

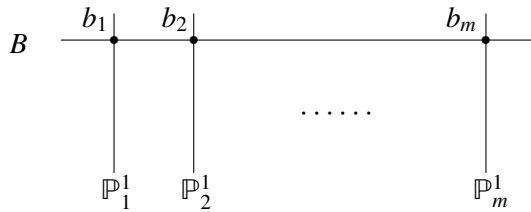
The combs $B(g_1, \dots, g_m) \subset Z$ have some obvious deformations where we keep B fixed and vary the points b_i and the maps g_i with them. By construction, the points b_i can vary independently. The images of these 1-cycles in $B \times \mathbb{P}^1$ are of the form

$$A + r(\{b'_1\} \times \mathbb{P}^1) + \dots + r(\{b'_m\} \times \mathbb{P}^1),$$

where the b'_i vary independently. □

21 (Combs). A *comb* assembled from B and m copies of \mathbb{P}^1 attached at the points b_i is a curve $B(b_1, \dots, b_m)$ obtained from B and $\{b_1, \dots, b_m\} \times \mathbb{P}^1$ by identifying the points $b_i \in B$ and $(b_i, 0) \in \{b_i\} \times \mathbb{P}^1$.

A comb (with m teeth) can be pictured as below:



22 (Deformation of reducible curves). Let X be a smooth projective variety and $C \subset X$ a connected curve with ideal sheaf I_C . Assume that C has only nodes as singularities. By the smoothness criterion of the Hilbert scheme [Grothendieck 1962, exposé 221, p. 21], if $H^1(C, (I_C/I_C^2)^*) = 0$ then $[C]$ is a smooth point of the Hilbert scheme $\text{Hilb}(X)$ and there is a unique irreducible component $H_C \subset \text{Hilb}(X)$ containing $[C]$. Let $U_C \rightarrow H_C$ be the universal family.

If, in addition, $(I_C/I_C^2)^*$ is generated by global sections, then a general point of H_C corresponds to a smooth curve and the natural map $U_C \rightarrow X$ is dominant.

23 (Proof of Theorem 4). Let us start with the $k(B)$ -variety $X_{k(B)}$. We can write it as the generic fiber of a quasiprojective k -variety $X \rightarrow B$ and extend $\pi_{k(B)}$ to $f : X \rightarrow B \times \mathbb{P}^1$. If (1) of Theorem 4 fails then using Proposition 12 we obtain $Z \subset X$. Take a compactification \bar{Z} and a resolution $Z_1 \rightarrow \bar{Z}$ such that the composite map $Z_1 \dashrightarrow B \times \mathbb{P}^1$ is a morphism. Next apply Proposition 13 to $Z_1 \rightarrow B \times \mathbb{P}^1$. We obtain, for every m in an arithmetic progression, a dominant family of sections $\sigma_m : S_m \times B \rightarrow Z_1$.

There is a dense open subset $D_m \subset S_m$ such that for every $q \in D_m(\bar{k})$,

- (1) the corresponding section $\sigma_m(\{q\} \times B) \subset Z_1$ intersects Z , and
- (2) the corresponding rational function $f \circ \sigma_m : \{q\} \times B \rightarrow \mathbb{P}^1$ has exactly $a + rm$ poles where a is the number of poles of $f \circ \sigma(B)$.

Thus the composite map

$$\rho_m := \text{pole} \circ f \circ \sigma_m : D_m \rightarrow S^{a+rm} B$$

is defined. The condition (2.b) of Theorem 4 holds by construction and the Zariski closure of $\rho_m(D_m)$ contains $P_{a+r} \cdot S^m B \subset S^{a+rm} B$ by (5) of Proposition 13, where P_a denotes the polar divisor of the section σ , that is, the 0-cycle

$$f \circ \sigma(B) \cap (B \times \{\infty\}).$$

S_m has a smooth k -point by Proposition 13 and k -points are Zariski dense since k is a large field. Thus $D_m(k)$ is Zariski dense in D_m . This implies both (2.a) and (2.c) of Theorem 4.

Finally, $D_m(k)$ is Zariski very dense in D_m by Lemma 24 and

$$\text{d-dim}_k \text{dioph}(X, \pi) \geq \text{d-dim}_k \rho_m(D_m(k)) = \text{dim}_k \rho_m(D_m) \geq \text{dim}_k S^m B = m,$$

where the middle equality holds by Lemma 24. Thus $\text{d-dim}_k \text{dioph}(X, \pi) = \infty$.

The only remaining issue is that our m runs through an arithmetic progression and is not arbitrary. If the progression is $m = b + m'c$ then

$$a + r(b + m'c) = (a + rb) + (rc)m',$$

so by changing $a \mapsto a + rb, r \mapsto rc$ we get Theorem 4. □

Lemma 24. *Let X be a smooth and irreducible variety over a large field k such that $X(k) \neq \emptyset$. Then $X(k)$ is not contained in the union of fewer than $|k|$ subvarieties $X_\lambda \subsetneq X$. In particular, if k is uncountable then $X(k)$ is Zariski very dense in X .*

Proof. Assume to the contrary that $X(k) \neq \emptyset$ but $X(k) \subset \bigcup_{\lambda \in \Lambda} X_\lambda$ where $|\Lambda| < |k|$ and $X_\lambda \neq X$.

If $\text{dim } X \geq 2$, then pick $p \in X(k)$ and let $\{H_t : t \in \mathbb{P}_k^1\}$ be a general pencil of hypersurface sections of X passing through p . Since $|\Lambda| < |k|$, there is an H_t such that H_t is smooth at p and $H_t \not\subset X_\lambda$ for every λ . Thus $H_t(k) \subset \bigcup_{\lambda \in \Lambda} (H_t \cap X_\lambda)$ is a lower dimensional counterexample. Thus it is enough to prove Lemma 24 when X is a curve. Then lower dimensional k -subvarieties are just points, thus we need to show that $|X(k)| = |k|$.

Let $\{H_t : t \in \mathbb{P}^1\}$ be a linear system on $X \times X$ which has (p, p) as a base point and whose general member is smooth at (p, p) . Since k is large, each H_t contains a k -point different from (p, p) . Thus $|X(k)| = |X(k) \times X(k)| \geq |k|$. □

Acknowledgments

I thank F. Bogomolov, K. Eisenträger, B. Poonen and J. Starr for useful conversations and the referee for several insightful suggestions.

References

- [Araujo and Kollár 2003] C. Araujo and J. Kollár, “Rational curves on varieties”, pp. 13–68 in *Higher dimensional varieties and rational points* (Budapest, 2001), edited by K. Böröczky, Jr. et al., Bolyai Soc. Math. Stud. **12**, Springer, Berlin, 2003. MR 2004k:14049 Zbl 1080.14521
- [Denef et al. 2000] J. Denef, L. Lipshitz, T. Pheidas, and J. Van Geel (editors), *Hilbert’s tenth problem: relations with arithmetic and algebraic geometry* (Ghent, Belgium, 1999), Contemporary Mathematics **270**, American Mathematical Society, Providence, RI, 2000. MR 2001g:00018 Zbl 0955.00034
- [Graber et al. 2003] T. Graber, J. Harris, and J. Starr, “Families of rationally connected varieties”, *J. Amer. Math. Soc.* **16**:1 (2003), 57–67. MR 2003m:14081 Zbl 01832408
- [Graber et al. 2005] T. Graber, J. Harris, B. Mazur, and J. Starr, “Rational connectivity and sections of families over curves”, *Ann. Sci. École Norm. Sup.* (4) **38**:5 (2005), 671–692. MR 2006j:14044 Zbl 05009844
- [Grothendieck 1962] A. Grothendieck, *Fondements de la géométrie algébrique*. [Extraits du Séminaire Bourbaki, 1957–1962], Secrétariat mathématique, Paris, 1962. MR 26 #3566 Zbl 0239.14002
- [Kollár 1996] J. Kollár, *Rational curves on algebraic varieties*, Ergebnisse der Mathematik (3) **32**, Springer, Berlin, 1996. MR 98c:14001 Zbl 0877.14012
- [Kollár 1999] J. Kollár, “Rationally connected varieties over local fields”, *Ann. of Math.* (2) **150**:1 (1999), 357–367. MR 2000h:14019 Zbl 0976.14016
- [Kollár 2004] J. Kollár, “Specialization of zero cycles”, *Publ. Res. Inst. Math. Sci.* **40**:3 (2004), 689–708. MR 2005i:14007 Zbl 1081.14008
- [Kollár et al. 1992] J. Kollár, Y. Miyaoka, and S. Mori, “Rationally connected varieties”, *J. Algebraic Geom.* **1**:3 (1992), 429–448. MR 93i:14014 Zbl 0780.14026
- [Mori 1979] S. Mori, “Projective manifolds with ample tangent bundles”, *Ann. of Math.* (2) **110**:3 (1979), 593–606. MR 81j:14010 Zbl 0423.14006
- [Poonen 2003] B. Poonen, “Hilbert’s tenth problem and Mazur’s conjecture for large subrings of \mathbb{Q} ”, *J. Amer. Math. Soc.* **16**:4 (2003), 981–990. MR 2004f:11145 Zbl 1028.11077

Communicated by Bjorn Poonen

Received 2007-09-26

Revised 2008-01-16

Accepted 2008-02-14

kollar@math.princeton.edu

*Department of Mathematics, Princeton University,
Princeton, NJ 08544-1000, United States*

Symmetric obstruction theories and Hilbert schemes of points on threefolds

Kai Behrend and Barbara Fantechi

In an earlier paper by one of us (Behrend), Donaldson–Thomas type invariants were expressed as certain weighted Euler characteristics of the moduli space. The Euler characteristic is weighted by a certain canonical \mathbb{Z} -valued constructible function on the moduli space. This constructible function associates to any point of the moduli space a certain invariant of the singularity of the space at the point.

Here we evaluate this invariant for the case of a singularity that is an isolated point of a \mathbb{C}^* -action and that admits a symmetric obstruction theory compatible with the \mathbb{C}^* -action. The answer is $(-1)^d$, where d is the dimension of the Zariski tangent space.

We use this result to prove that for any threefold, proper or not, the weighted Euler characteristic of the Hilbert scheme of n points on the threefold is, up to sign, equal to the usual Euler characteristic. For the case of a projective Calabi–Yau threefold, we deduce that the Donaldson–Thomas invariant of the Hilbert scheme of n points is, up to sign, equal to the Euler characteristic. This proves a conjecture of Maulik, Nekrasov, Okounkov and Pandharipande.

Introduction

The first purpose of this paper is to introduce *symmetric obstruction theories*. In a nutshell, these are obstruction theories for which the space of infinitesimal deformations is the dual of the space of infinitesimal obstructions.

As an example of an obstruction theory, consider the intersection of two smooth varieties V , W inside an ambient smooth variety M . The intersection X carries an obstruction theory. This is the 2-term complex of vector bundles

$$E = [\Omega_M|_X \xrightarrow{\text{res}_V - \text{res}_W} \Omega_V|_X \oplus \Omega_W|_X]$$

considered as an object of the derived category $D(X)$ of X , taking up degrees -1 and 0 . We see that infinitesimal deformations of X are classified by $h^0(E^\vee) = T_X$,

MSC2000: 00A05.

Keywords: symmetric obstruction theories, Hilbert schemes, Calabi–Yau threefolds, C^* actions, S^1 actions, Donaldson–Thomas invariants, MNOP conjecture.

the sheaf of derivations on X . Moreover, the obstructions to the smoothness of X are contained in $h^1(E^\vee)$, which is called the *obstruction sheaf*, notation $ob = h^1(E^\vee)$. Note that $h^0(E^\vee)$ is intrinsic to X , but $h^1(E^\vee)$ is not. In fact, if X is smooth, all obstructions vanish, but $h^1(E^\vee)$ may be nonzero, although it is always a vector bundle, in this case.

This obstruction theory E is *symmetric*, if M is a complex symplectic manifold, i.e., hyperkähler, and V, W are Lagrangian submanifolds. In fact, the symplectic form σ induces a homomorphism $T_X \rightarrow \Omega_M$, defined by $v \mapsto \sigma(v, -)$. The fact that V and W are Lagrangian, i.e., equal to their own orthogonal complements with respect to σ , implies that there is an exact sequence

$$0 \longrightarrow T_X \longrightarrow \Omega_M|_X \longrightarrow \Omega_V|_X \oplus \Omega_W|_X \longrightarrow \Omega_X \longrightarrow 0.$$

Assuming for simplicity that X is smooth and hence this is an exact sequence of vector bundles, we see that $ob = h^1(E^\vee) = \Omega_X$ and hence T_X is, indeed, dual to ob .

In more abstract terms, what makes an obstruction theory E symmetric is a nondegenerate symmetric bilinear form of degree 1

$$\beta : E \overset{L}{\otimes} E \longrightarrow \mathbb{C}_X[1].$$

If M is an arbitrary smooth scheme, then Ω_M is a symplectic manifold in a canonical way, and the graph of any closed 1-form ω is a Lagrangian submanifold. Thus the scheme theoretic zero locus $X = Z(\omega)$ of ω is an example of the above, the second Lagrangian being the zero section.

As a special case of this, we may consider the Jacobian locus $X = Z(df)$ of a regular function on a smooth variety M . It is endowed with a canonical symmetric obstruction theory. In Donaldson–Thomas theory, where the moduli space is heuristically the critical locus of the holomorphic Chern–Simons functional, there is a canonical symmetric obstruction theory; see [Thomas 2000].

Unfortunately, we are unable to prove that every symmetric obstruction theory is locally given as the zero locus of a closed 1-form on a smooth scheme, even though we see no reason why this should not be true.

The best we can do is to show that the most general local example of a symmetric obstruction theory is the zero locus of an *almost closed* 1-form on a smooth scheme. A form ω is almost closed if its differential $d\omega$ vanishes on the zero locus $Z(\omega)$.

For the applications we have in mind we also need equivariant versions of all of the above, in the presence of a \mathbb{G}_m -action.

Weighted Euler characteristics and \mathbb{G}_m -actions. In [Behrend 2005] a new (as far as we can tell) invariant of singularities was introduced. For a singularity (X, P) the notation was

$$\nu_X(P).$$

The function ν_X is a constructible \mathbb{Z} -valued function on any Deligne–Mumford stack X . In [Behrend 2005], the following facts were proved about ν_X :

- If X is smooth at P , then $\nu_X(P) = (-1)^{\dim X}$.
- $\nu_X(P) \nu_Y(Q) = \nu_{X \times Y}(P, Q)$.
- If $X = Z(df)$ is the singular locus of a regular function f on a smooth variety M , then

$$\nu_X(P) = (-1)^{\dim M} (1 - \chi(F_P)),$$

where F_P is the Milnor fibre of f at P .

- Let X be a projective scheme endowed with a symmetric obstruction theory. The associated Donaldson–Thomas type invariant (or virtual count) is the degree of the associated virtual fundamental class. In this case, $\nu_X(P)$ is the contribution of the point P to the Donaldson–Thomas type invariant, in the sense that

$$\#\text{vir}(X) = \chi(X, \nu_X) = \sum_{n \in \mathbb{Z}} n \chi(\{\nu_X = n\}).$$

We define the weighted Euler characteristic of X to be

$$\tilde{\chi}(X) = \chi(X, \nu_X).$$

The last property shows the importance of $\nu_X(P)$ for the calculation of Donaldson–Thomas type invariants.

In this paper we calculate the number $\nu_X(P)$ for certain kinds of singularities. In fact, we will assume that X admits a \mathbb{G}_m -action and a symmetric obstruction theory, which are compatible with each other. Moreover, we assume P to be an isolated fixed point for the \mathbb{G}_m -action. We prove that

$$\nu_X(P) = (-1)^{\dim T_X|_P}, \tag{1}$$

in this case.

We get results of two different flavors from this:

- If the scheme X admits a globally defined \mathbb{G}_m -action with isolated fixed points and around every fixed point admits a symmetric obstruction theory compatible with the \mathbb{G}_m -action we obtain

$$\tilde{\chi}(X) = \sum_P (-1)^{\dim T_X|_P}, \tag{2}$$

the sum extending over the fixed points of the \mathbb{G}_m -action. This is because nontrivial \mathbb{G}_m -orbits do not contribute, the Euler characteristic of \mathbb{G}_m being zero, and ν_X being constant on such orbits.

• If X is projective, with globally defined \mathbb{G}_m -action and symmetric obstruction theory, these two structures being compatible, we get

$$\#^{\text{vir}}(X) = \tilde{\chi}(X) = \sum_P (-1)^{\dim T_X|_P}. \tag{3}$$

An example. It may be worth pointing out how to prove (1) in a special case. Assume the multiplicative group \mathbb{G}_m acts on affine n -space \mathbb{A}^n in a linear way with nontrivial weights $r_1, \dots, r_n \in \mathbb{Z}$, so that the origin P is an isolated fixed point. Let f be a regular function on \mathbb{A}^n , which is invariant with respect to the \mathbb{G}_m -action. This means that $f(x_1, \dots, x_n)$ is of degree zero, if we assign to x_i the degree r_i . Let $X = Z(df)$ be the scheme-theoretic critical set of f . The scheme X inherits a \mathbb{G}_m -action. It also carries a symmetric obstruction theory which is compatible with the \mathbb{G}_m -action.

Assume that $f \in (x_1, \dots, x_n)^3$. This is not a serious restriction. It ensures that $T_X|_P = T_{\mathbb{A}^n}|_P$ and hence that $\dim T_X|_P = n$.

Let $\epsilon \in \mathbb{R}$, $\epsilon > 0$ and $\eta \in \mathbb{C}$, $\eta \neq 0$ be chosen such that the Milnor fiber of f at the origin may be defined as

$$F_P = \{x \in \mathbb{C}^n \mid f(x) = \eta \text{ and } |x| < \epsilon\}.$$

It is easy to check that F_P is invariant under the S^1 -action on \mathbb{C}^n induced by our \mathbb{G}_m -action. Moreover, the S^1 -action on F_P has no fixed points. This implies immediately that $\chi(F_P) = 0$ and hence that $\nu_X(P) = (-1)^n$.

Even though we consider this example $(Z(df), P)$ to be the prototype of a singularity admitting compatible \mathbb{G}_m -actions and symmetric obstruction theories, we cannot prove that every such singularity is of the form $(Z(df), P)$. We can only prove that a singularity with compatible \mathbb{G}_m -action and symmetric obstruction theory looks like $(Z(\omega), P)$, where ω is an almost closed \mathbb{G}_m -invariant 1-form on \mathbb{A}^n , rather than the exact invariant 1-form df . This is why the proof of (1) is more involved, in the general case. Rather than the Milnor fiber, we use the expression of $\nu_X(P)$ as a linking number, found in Proposition 4.22 of [Behrend 2005].

Lagrangian intersections. One amusing application of (3) is the following formula. Assume M is a complex symplectic manifold with a Hamiltonian \mathbb{C}^* -action, all of whose fixed points are isolated. Let V and W be invariant Lagrangian submanifolds. Assume their intersection is compact. Finally, assume that the Zariski tangent space of the intersection at every fixed point is even-dimensional. Then we can express the intersection number as an Euler characteristic:

$$\deg([V] \cap [W]) = \chi(V \cap W).$$

Hilbert schemes. Our result is a powerful tool for computing weighted Euler characteristics. It is a replacement for the lacking additivity of $\tilde{\chi}$ over strata.

As an example of the utility of (1), we will show in this paper that

$$\tilde{\chi}(\text{Hilb}^n Y) = (-1)^n \chi(\text{Hilb}^n Y), \tag{4}$$

for every smooth scheme Y of dimension 3.

In particular, if Y is projective and Calabi–Yau (i.e., has a chosen trivialization $\omega_Y = \mathbb{O}_Y$), we get that

$$\#^{\text{vir}}(\text{Hilb}^n Y) = (-1)^n \chi(\text{Hilb}^n Y),$$

where $\#^{\text{vir}}$ is the virtual count à la Donaldson–Thomas [2000]. This latter formula was conjectured by Maulik, Nekrasov, Okounkov and Pandharipande in [Maulik et al. 2003]. Using the McMahon function $M(t) = \prod_{n=1}^{\infty} (1 - t^n)^{-n}$, we can also express this result as

$$\sum_{n=0}^{\infty} \#^{\text{vir}}(\text{Hilb}^n Y) t^n = M(-t)^{\chi(Y)}.$$

The strategy for proving (4) is as follows. We first consider the open Calabi–Yau threefold \mathbb{A}^3 . We exploit a suitable \mathbb{G}_m -action on \mathbb{A}^3 to prove (4) for $Y = \mathbb{A}^3$, using Formula (2). At this point, we can drop all Calabi–Yau assumptions.

Let F_n be the *punctual* Hilbert scheme. It parameterizes subschemes of \mathbb{A}^3 of length n which are entirely supported at the origin. Let ν_n be the restriction of $\nu_{\text{Hilb}^n \mathbb{A}^3}$ to F_n . Formula (4) for $Y = \mathbb{A}^3$ is equivalent to

$$\chi(F_n, \nu_n) = (-1)^n \chi(F_n). \tag{5}$$

Finally, using more or less standard stratification arguments, we express $\tilde{\chi}(\text{Hilb}^n Y)$ in terms of $\chi(F_n, \nu_n)$. This uses the fact that the punctual Hilbert scheme of Y at a point P is isomorphic to F_n . Then (5) implies (4).

Conventions. We will work over the field of complex numbers. All stacks will be of Deligne–Mumford type. All schemes and stacks will be of finite type over \mathbb{C} . Hence the derived category $D_{\text{qcoh}}(\mathbb{O}_X)$, of complexes of sheaves of \mathbb{O}_X -modules with quasicohherent cohomology is equivalent to the derived category $D(\text{Qcoh-}\mathbb{O}_X)$ of the category of quasicohherent \mathbb{O}_X -modules (see Proposition 3.7 in Exposé II of SGA6). To fix ideas, we will denote by $D(X)$ the latter derived category and call it the derived category of X . We will often write $E \otimes F$ instead of $E \overset{L}{\otimes} F$, for objects E, F of $D(X)$.

Let X be a Deligne–Mumford stack. We will write L_X for the cutoff at -1 of the cotangent complex of X . Thus, if $U \rightarrow X$ is étale and $U \rightarrow M$ a closed immersion into a smooth scheme M , we have, *canonically*,

$$L_X|_U = [I/I^2 \rightarrow \Omega_M|_X],$$

where I is the ideal sheaf of U in M and we think of the homomorphism $I/I^2 \rightarrow \Omega_M|_X$ of coherent sheaves on U as a complex concentrated in the interval $[-1, 0]$. We will also call L_X the cotangent complex of X , and hope the reader will forgive this abuse of language. The cotangent complex L_X is an object of $D(X)$.

We will often use homological notation for objects in the derived category. This means that $E_n = E^{-n}$, for a complex $\cdots \rightarrow E^i \rightarrow E^{i+1} \rightarrow \cdots$ in $D(X)$.

For a complex of sheaves E , we denote the cohomology sheaves by $h^i(E)$.

Let us recall a few sign conventions: If $E = [E_1 \xrightarrow{\alpha} E_0]$ is a complex concentrated in the interval $[-1, 0]$, then

$$E^\vee = [E_0^\vee \xrightarrow{-\alpha^\vee} E_1^\vee]$$

is a complex concentrated in the interval $[0, 1]$. Thus the shifted dual $E^\vee[1]$ is given by

$$E^\vee[1] = [E_0^\vee \xrightarrow{\alpha^\vee} E_1^\vee]$$

and concentrated, again, in the interval $[-1, 0]$.

If $\theta : E \rightarrow F$ is a homomorphism of complexes concentrated in the interval $[-1, 0]$, such that $\theta = (\theta_1, \theta_0)$, then the shifted dual $\theta^\vee[1] : F^\vee[1] \rightarrow E^\vee[1]$ is given by $\theta^\vee[1] = (\theta_0^\vee, \theta_1^\vee)$.

Suppose $E = [E_1 \xrightarrow{\alpha} E_0]$ and $F = [F_1 \xrightarrow{\beta} F_0]$ are complexes concentrated in the interval $[-1, 0]$ and $\theta : E \rightarrow F$ and $\eta : E \rightarrow F$ homomorphisms of complexes. Then a homotopy from η to θ is a homomorphism $h : E_0 \rightarrow F_1$ such that $h \circ \alpha = \theta_1 - \eta_1$ and $\beta \circ h = \theta_0 - \eta_0$.

1. Symmetric obstruction theories

Nondegenerate symmetric bilinear forms.

Definition 1.1. Let X be a scheme or a Deligne–Mumford stack. Let $E \in D_{coh}^b(\mathbb{C}_X)$ be a perfect complex. A *nondegenerate symmetric bilinear form of degree 1* on E is a morphism

$$\beta : E \otimes^L E \longrightarrow \mathbb{C}_X[1]$$

in $D(X)$, which is

(i) *symmetric*, which means that

$$\beta(e \otimes e') = (-1)^{\deg(e) \deg(e')} \beta(e' \otimes e);$$

(ii) *nondegenerate*, which means that β induces an isomorphism

$$\theta : E \longrightarrow E^\vee[1].$$

Remark 1.2. The isomorphism $\theta : E \rightarrow E^\vee[1]$ determines β as the composition

$$E \otimes E \xrightarrow{\theta \otimes \text{id}} E^\vee[1] \otimes E \xrightarrow{\text{tr}[1]} \mathbb{O}_X[1].$$

Symmetry of β is equivalent to the condition

$$\theta^\vee[1] = \theta.$$

Usually, we will find it more convenient to work with θ , rather than β . Thus we will think of a nondegenerate symmetric bilinear form of degree 1 on E as an isomorphism $\theta : E \rightarrow E^\vee[1]$, satisfying $\theta^\vee[1] = \theta$.

Remark 1.3. Above, we have defined nondegenerate symmetric bilinear forms of degree 1. One can generalize the definition to any degree $n \in \mathbb{Z}$. Only the case $n = 1$ will interest us in this paper.

Example 1.4. Let F be a vector bundle on X and let $\alpha : F \rightarrow F^\vee$ a symmetric bilinear form. Define the complex $E = [F \rightarrow F^\vee]$, by putting F^\vee in degree 0 and F in degree -1 . Then $E^\vee[1] = E$. Define $\theta = (\theta_1, \theta_0)$ by $\theta_1 = \text{id}_F$ and $\theta_0 = \text{id}_{F^\vee}$:

$$\begin{array}{ccc} E & = & [F \xrightarrow{\alpha} F^\vee] \\ \downarrow \theta & & \begin{array}{ccc} 1 \downarrow & & \downarrow 1 \\ & & \end{array} \\ E^\vee[1] & = & [F \xrightarrow{\alpha} F^\vee] \end{array}$$

Then E is a perfect complex with perfect amplitude contained in $[-1, 0]$. Moreover, θ is a nondegenerate symmetric bilinear form on E . Note that θ is an isomorphism, and hence the form it defines is nondegenerate, whether or not α is nondegenerate.

Example 1.5. Let f be a regular function on a smooth variety M . The Hessian of f defines a symmetric bilinear form on $T_M|_X$, where $X = Z(df)$. So there is an induced symmetric bilinear form on the complex $[T_M|_X \rightarrow \Omega_M|_X]$.

Lemma 1.6. *Let E be a complex of vector bundles on X , concentrated in the interval $[-1, 0]$. Let $\theta : E \rightarrow E^\vee[1]$ be a homomorphism of complexes. Assume that $\theta^\vee[1] = \theta$, as morphisms in the derived category. Then Zariski-locally on the scheme X (or étale locally on the stack X) we can represent the derived category morphism given by θ as a homomorphism of complexes (θ_1, θ_0) :*

$$\begin{array}{ccc} E & = & [E_1 \xrightarrow{\alpha} E_0] \\ \downarrow \theta & & \begin{array}{ccc} \theta_1 \downarrow & & \downarrow \theta_0 \\ & & \end{array} \\ E^\vee[1] & = & [E_0^\vee \xrightarrow{\alpha^\vee} E_1^\vee] \end{array}$$

such that $\theta_1 = \theta_0^\vee$.

Proof. Let us use notation $\theta = (\psi_1, \psi_0)$. Then the equality of derived category morphisms $\theta^\vee[1] = \theta$ implies that, locally, $\theta^\vee[1] = (\psi_0^\vee, \psi_1^\vee)$ and $\theta = (\psi_1, \psi_0)$ are homotopic. So let $h : E_0 \rightarrow E_0^\vee$ be a homotopy:

$$\begin{aligned} h\alpha &= \psi_1 - \psi_0^\vee, \\ \alpha^\vee h &= \psi_0 - \psi_1^\vee. \end{aligned}$$

Now define

$$\begin{aligned} \theta_0 &= \frac{1}{2}(\psi_0 + \psi_1^\vee), \\ \theta_1 &= \frac{1}{2}(\psi_1 + \psi_0^\vee). \end{aligned}$$

One checks, using h , that (θ_1, θ_0) is a homomorphism of complexes, and as such, homotopic to (ψ_1, ψ_0) . Thus (θ_1, θ_0) represents the derived category morphism θ , and has the required property. \square

The next lemma shows that for amplitude 1 objects, every nondegenerate symmetric bilinear form locally looks like the one given in Example 1.4. Again, locally means étale locally, but in the scheme case Zariski locally.

Lemma 1.7. *Suppose that $A \in D_{coh}^b(\mathbb{C}_X)$ is of perfect amplitude contained in $[-1, 0]$, and that $\eta : A \rightarrow A^\vee[1]$ is an isomorphism satisfying $\eta^\vee[1] = \eta$. Then we can locally represent A by a homomorphism of vector bundles $\alpha : E \rightarrow E^\vee$ satisfying $\alpha^\vee = \alpha$ and the isomorphism η by the identity.*

Proof. Start by representing the derived category object A by an actual complex of vector bundles $\alpha : A_1 \rightarrow A_0$, and the morphism $\eta : A \rightarrow A^\vee[1]$ by an actual homomorphism of complexes (η_1, η_0) . Then pick a point $P \in X$ and lift a basis of $\text{cok}(\alpha)(P)$ to A_0 . replace A_0 by the free \mathbb{C}_X -module on this bases, and pull back to get a quasiisomorphic complex.

Now any representative of η has, necessarily, that η_0 is an isomorphism in a neighborhood of P . By Lemma 1.6, we can assume that $\eta_1 = \eta_0^\vee$. Then both η_0 and η_1 are isomorphisms of vector bundles. Now use η_0 to identify A_0 with A_1^\vee . \square

Isometries.

Definition 1.8. Consider perfect complexes A and B , endowed with nondegenerate symmetric forms $\theta : A \rightarrow A^\vee[1]$ and $\eta : B \rightarrow B^\vee[1]$. An isomorphism $\Phi : B \rightarrow A$, such that the diagram

$$\begin{array}{ccc} B & \xrightarrow{\Phi} & A \\ \eta \downarrow & & \downarrow \theta \\ B^\vee[1] & \xleftarrow{\Phi^\vee[1]} & A^\vee[1] \end{array}$$

commutes in $D(X)$, is called an *isometry* $\Phi : (B, \eta) \rightarrow (A, \theta)$.

Note that because η and θ are isomorphisms, the condition on Φ is equivalent to $\Phi^{-1} = \Phi^\vee[1]$, if we use η and θ to identify A with B .

We include the following lemma on the local structure of isometries for the information of the reader. Since we do not use it in the sequel, we omit the (lengthy) proof.

Lemma 1.9. *Let A and B be perfect, of amplitude contained in $[-1, 0]$. Suppose $\theta : A \rightarrow A^\vee[1]$ and $\eta : B \rightarrow B^\vee[1]$ are nondegenerate symmetric forms. Let $\Phi : B \rightarrow A$ be an isometry.*

Suppose that (A, θ) and (B, η) are represented as in Example 1.4 or Lemma 1.7. Thus,

$$A = [E \xrightarrow{\alpha} E^\vee] \quad \text{and} \quad B = [F \xrightarrow{\beta} F^\vee],$$

for vector bundles E and F on X . Moreover, θ and η are the respective identities.

Assume that $\text{rk}(F) = \text{rk}(E)$. Then, étale locally in X (Zariski locally if X is a scheme), we can find a vector bundle isomorphism

$$\phi : F \longrightarrow E,$$

such that $\alpha \circ \phi = \phi^{\vee-1} \circ \beta$, and $(\phi, \phi^{\vee-1})$ represents Φ :

$$\begin{array}{ccc} B & = & [F \xrightarrow{\beta} F^\vee] \\ \Phi \downarrow & & \downarrow \phi \qquad \qquad \downarrow \phi^{\vee-1} \\ A & = & [E \xrightarrow{\alpha} E^\vee] \end{array}$$

In particular, (ϕ^{-1}, ϕ^\vee) represents $\Phi^\vee[1]$.

Symmetric obstruction theories. Recall from [Behrend and Fantechi 1997] that a *perfect obstruction theory* for the scheme (or Deligne–Mumford stack) X is a morphism $\phi : E \rightarrow L_X$ in $D(X)$, where E is perfect, of amplitude in $[-1, 0]$, we have $h^0(\phi) : h^0(E) \rightarrow \Omega_X$ is an isomorphism and $h^{-1} : h^{-1}(E) \rightarrow h^{-1}(L_X)$ is onto.

We denote the coherent sheaf $h^1(E^\vee)$ by *ob* and call it the *obstruction sheaf* of the obstruction theory. It contains in a natural way the obstructions to the smoothness of X . Even though we do not include E in the notation, *ob* is by no means an intrinsic invariant of X .

Any perfect obstruction theory for X induces a virtual fundamental class $[X]^{\text{vir}}$ for X . We leave the obstruction theory out of the notation, even though $[X]^{\text{vir}}$ depends on it. The virtual fundamental class is an element of $A_{\text{rk } E}(X)$, the Chow group of algebraic cycles modulo rational equivalence. The degree of $[X]^{\text{vir}}$ is equal to the rank of E .

Definition 1.10. Let X be a Deligne–Mumford stack. A *symmetric obstruction theory* for X is a triple (E, ϕ, θ) where $\phi : E \rightarrow L_X$ is a perfect obstruction theory for X and $\theta : E \rightarrow E^\vee[1]$ a nondegenerate symmetric bilinear form.

We will often refer to such an E as a symmetric obstruction theory, leaving the morphisms ϕ and θ out of the notation.

Remark 1.11. It is shown in [Behrend 2005] that for symmetric obstruction theories, the virtual fundamental class is intrinsic to X , namely it is the degree zero Aluffi class of X .

Proposition 1.12. *Every symmetric obstruction theory has expected dimension zero.*

Proof. Recall that the expected dimension of $E \rightarrow L_X$ is the rank of E . If $E \rightarrow L_X$ is symmetric, we have $\text{rk } E = \text{rk}(E^\vee[1]) = -\text{rk } E^\vee = -\text{rk } E$ and hence $\text{rk } E = 0$. \square

By this proposition, the following definition makes sense.

Definition 1.13. Assume X is proper and we have given a symmetric obstruction theory for X . We define the *virtual count* of X to be the number

$$\#\text{vir}(X) = \deg[X]^{\text{vir}} = \int_{[X]^{\text{vir}}} 1.$$

If X is a scheme (or an algebraic space), the virtual count is an integer. In general it may be a rational number.

Proposition 1.14. *For a symmetric obstruction theory $E \rightarrow L_X$, the obstruction sheaf is canonically isomorphic to the sheaf of differentials:*

$$ob = \Omega_X.$$

Proof. We have $ob = h^1(E^\vee) = h^0(E^\vee[1]) = h^0(E) = \Omega_X$. \square

Corollary 1.15. *For a symmetric obstruction theory,*

$$h^{-1}(E) = \mathcal{H}om(\Omega_X, \mathbb{O}_X) = T_X.$$

Proof. We always have $h^{-1}(E) = ob^\vee$. \square

Definition 1.16. Let E and F be symmetric obstruction theories for X . An *isomorphism* of symmetric obstruction theories is an isometry $\Phi : E \rightarrow F$ commuting with the maps to L_X .

Remark 1.17. Let $f : X \rightarrow X'$ be an étale morphism of Deligne–Mumford stacks, and suppose that X' has a symmetric obstruction theory E' . Then f^*E' is naturally a symmetric obstruction theory for X .

Conversely, if we are given symmetric obstruction theories E for X and E' for X' , we will say that the morphism f is compatible with the obstruction theories if E is isomorphic to f^*E' as symmetric obstruction theory.

Remark 1.18. If X and X' are Deligne–Mumford stacks with symmetric obstruction theories E and E' , then $p_1^*E \oplus p_2^*E'$ is naturally a symmetric obstruction theory for $X \times X'$. We call it the product symmetric obstruction theory.

Example 1.19. Let M be smooth and ω a closed 1-form on M . Let $X = Z(\omega)$ be the scheme-theoretic zero locus of ω . Consider ω as a linear epimorphism $\omega^\vee : T_M \rightarrow I$, where I is the ideal sheaf of X in M . Let us denote the restriction to X of the composition of ω^\vee and $d : I \rightarrow \Omega_M$ by $\nabla\omega$. It is a linear homomorphism of vector bundles $\nabla\omega : T_M|_X \rightarrow \Omega_M|_X$. Because ω is closed, $\nabla\omega$ is symmetric and, as we have seen in Example 1.4, defines a symmetric bilinear form on the complex $E = [T_M|_X \rightarrow \Omega_M|_X]$.

The morphism $\phi : E \rightarrow L_X$ as in the diagram

$$\begin{array}{ccc} E & = & [T_M|_X \xrightarrow{\nabla\omega} \Omega_M|_X] \\ \phi \downarrow & & \omega^\vee \downarrow \qquad \qquad \downarrow 1 \\ L_X & = & [I/I^2 \xrightarrow{d} \Omega_M|_X] \end{array}$$

makes E into a symmetric obstruction theory for X . In particular, note that Example 1.5 gives rise to a symmetric obstruction theory on the Jacobian locus of a regular function.

Let us remark that for the symmetry of $\nabla\omega$ and hence the symmetry of the obstruction theory given by ω , it is sufficient that ω be *almost closed*, which means that $d\omega \in I\Omega_M^2$.

A remark on the lci case. We will show that the existence of a symmetric obstruction theory puts strong restrictions on the singularities X can have.

For the following proposition, it is important to recall that we are working in characteristic zero.

Proposition 1.20. *Let $E \rightarrow L_X$ be a perfect obstruction theory, symmetric or not. A criterion for the obstruction sheaf to be locally free is that X be a reduced local complete intersection.*

Proof. As the claim is local, we may assume that E has a global resolution $E = [E_1 \rightarrow E_0]$, that $X \hookrightarrow M$ is embedded in a smooth scheme M (with ideal sheaf I) and that $E \rightarrow L_X$ is given by a homomorphism of complexes $[E_1 \rightarrow E_0] \rightarrow [I/I^2 \rightarrow \Omega_M|_X]$. We may even assume that $E_0 \rightarrow \Omega_M|_X$ is an isomorphism of vector bundles.

Under the assumption that X is a reduced local complete intersection, I/I^2 is locally free and that $I/I^2 \rightarrow \Omega_M|_X$ is injective. Then a simple diagram chase proves that we have a short exact sequence of coherent sheaves

$$0 \longrightarrow h^{-1}(E) \longrightarrow E_1 \longrightarrow I/I^2 \longrightarrow 0.$$

Hence, $h^{-1}(E)$ is a subbundle of E_1 and $ob = h^{-1}(E)^\vee$. In particular, ob is locally free. (We always have $h^{-1}(E) = ob^\vee$; the converse is generally false.) \square

Corollary 1.21. *If X is a reduced local complete intersection and admits a symmetric obstruction theory, then X is smooth.*

Proof. Because $ob = \Omega_X$, the sheaf Ω_X is locally free. This implies that X is smooth. \square

Examples.

Lagrangian intersections. Let M be an algebraic symplectic manifold and V, W two Lagrangian submanifolds. Let X be the scheme-theoretic intersection. Then X carries a canonical symmetric obstruction theory.

To see this, note first of all that for a Lagrangian submanifold $V \subset M$, the normal bundle is equal to the cotangent bundle, $N_{V/M} = \Omega_V$. The isomorphism is given by $v \mapsto \sigma(v, -)$, where σ is the symplectic form, which maps $N_{V/M} = T_M/T_V$ to $\Omega_V = T_V^\vee$. It is essentially the definition of Lagrangian, that this map is an isomorphism of vector bundles on V .

Next, note that the obstruction theory for X as an intersection of V and W can be represented by the complex

$$E = [\Omega_M \xrightarrow{\text{res}_V - \text{res}_W} \Omega_V \oplus \Omega_W]|_X.$$

The shifted dual is

$$E^\vee[1] = [T_V \oplus T_W \longrightarrow T_M]|_X.$$

Define $\theta : T_M \rightarrow \Omega_V \oplus \Omega_W$ as the canonical map $T_M \rightarrow N_{V/M} \oplus N_{W/M}$ given by the projections, multiplied by the scalar factor $\frac{1}{2}$. Then (θ^\vee, θ) defines a morphism of complexes $E^\vee[1] \rightarrow E^\vee$:

$$\begin{array}{ccc} E^\vee[1] & = & [T_V \oplus T_W \longrightarrow T_M] \\ \downarrow & & \downarrow \theta^\vee \quad \downarrow \theta \\ E & = & [\Omega_M \longrightarrow \Omega_V \oplus \Omega_W] \end{array}$$

One checks that (θ^\vee, θ) is a quasiisomorphism. Since $(\theta^\vee, \theta)^\vee[1] = (\theta^\vee, \theta)$, this morphism of complexes defines a symmetric bilinear form on $E^\vee[1]$, hence on E . Thus E is a symmetric obstruction theory on X .

Sheaves on Calabi–Yau threefolds. Let Y be an integral proper 3-dimensional Gorenstein Deligne–Mumford stack (for example a projective threefold). By the Gorenstein assumption, Y admits a dualizing sheaf ω_Y , which is a line bundle over Y , also called the *canonical bundle*. Let $\omega_Y \rightarrow \mathbb{O}_Y$ be a nonzero homomorphism, giving rise to the short exact sequence

$$0 \longrightarrow \omega_Y \longrightarrow \mathbb{O}_Y \longrightarrow \mathbb{O}_D \longrightarrow 0,$$

so that D is an anticanonical divisor on Y . In fact, D is a Cartier divisor. Of course, D may be empty (this case we refer to as the Calabi–Yau case). Finally, choose an arbitrary line bundle L on Y . Often we are only interested in the case $L = \mathbb{O}_Y$.

Now let us define a certain moduli stack \mathfrak{M} of sheaves on Y . For an arbitrary \mathbb{C} -scheme S , let $\mathfrak{M}(S)$ be the groupoid of pairs (\mathcal{E}, ϕ) . Here \mathcal{E} is a sheaf of $\mathbb{O}_{Y \times S}$ -modules, such that

- (i) \mathcal{E} coherent,
- (ii) \mathcal{E} is flat over S ,
- (iii) \mathcal{E} is perfect as an object of the derived category of $Y \times S$, i.e., locally admits finite free resolutions, (by Cor. 4.6.1 of Exp. III of SGA 6, this is a condition which may be checked on the fibres of $\pi : Y \times S \rightarrow S$).

The second component of the pair (\mathcal{E}, ϕ) is an isomorphism $\phi : \det \mathcal{E} \rightarrow L$ of line bundles on $Y \times S$. Note that the determinant $\det \mathcal{E}$ is well-defined, by Condition (iii) on \mathcal{E} .

We require two more conditions on \mathcal{E} , namely that for every point $s \in S$, denoting the fibre of \mathcal{E} over s by \mathcal{E}_s , we have

- (iv) \mathcal{E}_s is simple, i.e., $\kappa(s) \rightarrow \text{Hom}(\mathcal{E}_s, \mathcal{E}_s)$ is an isomorphism,
- (v) the map induced by the trace $R\mathcal{H}om(\mathcal{E}_s, \mathcal{E}_s) \rightarrow \mathbb{O}_{Y_s}$ is an isomorphism in a neighborhood of D_s .

The last condition (v) is empty in the Calabi–Yau case. It is, for example, satisfied if \mathcal{E}_s is locally free of rank 1 in a neighborhood of D .

We let X be an open substack of \mathfrak{M} which is algebraic (for example, fix the Hilbert polynomial and pass to stable objects, but we do not want to get more restrictive than necessary). Then X is a Deligne–Mumford stack. We will now construct a symmetric obstruction theory for X .

For this, denote the universal sheaf on $Y \times X$ by \mathcal{E} and the projection $Y \times X \rightarrow X$ by π . Consider the trace map $R\mathcal{H}om(\mathcal{E}, \mathcal{E}) \rightarrow \mathbb{O}$ and let \mathcal{F} be its shifted cone, so that we obtain a distinguished triangle in $D(\mathbb{O}_{Y \times X})$:

$$\begin{array}{ccc}
 & \mathbb{O} & \\
 +1 \swarrow & & \nwarrow tr \\
 \mathcal{F} & \longrightarrow & R\mathcal{H}om(\mathcal{E}, \mathcal{E})
 \end{array}$$

Note that \mathcal{F} is self-dual: $\mathcal{F}^\vee = \mathcal{F}$, canonically.

Lemma 1.22. *The complex*

$$E = R\pi_* R\mathcal{H}om(\mathcal{F}, \omega_Y)[2]$$

is an obstruction theory for X .

Proof. This is well-known deformation theory; see [Thomas 2000]. \square

The homomorphism $\omega_Y \rightarrow \mathbb{O}_Y$ induces an *isomorphism*

$$R\pi_*(\mathcal{F} \otimes \omega_Y) \longrightarrow R\pi_*\mathcal{F},$$

because the cone of this homomorphism is $R\pi_*(\mathcal{F} \otimes \mathbb{O}_D)$ and $\mathcal{F} \otimes \mathbb{O}_D = 0$, by Assumption (v), above. Dualizing and shifting, we obtain an isomorphism

$$(R\pi_*\mathcal{F})^\vee[-1] \longrightarrow (R\pi_*(\mathcal{F} \otimes \omega_Y))^\vee[-1].$$

Exploiting the fact that \mathcal{F} is self-dual, we may rewrite this as

$$(R\pi_*\mathcal{F})^\vee[-1] \longrightarrow (R\pi_* R\mathcal{H}om(\mathcal{F}, \omega_Y))^\vee[-1],$$

or in other words

$$(R\pi_*\mathcal{F})^\vee[-1] \longrightarrow E^\vee[1]. \quad (6)$$

Now, relative Serre duality for the morphism $\pi : Y \times X \rightarrow X$ applied to \mathcal{F} states that

$$R\pi_* R\mathcal{H}om(\mathcal{F}, \omega_Y[3]) = (R\pi_*\mathcal{F})^\vee,$$

or in other words

$$E = (R\pi_*\mathcal{F})^\vee[-1].$$

Thus, we see that (6) is, in fact, an isomorphism

$$\theta : E \longrightarrow E^\vee[1].$$

Lemma 1.23. *The isomorphism $\theta : E \rightarrow E^\vee[1]$ satisfies the symmetry property $\theta^\vee[1] = \theta$.*

Proof. This is just a derived version of the well-known fact that $\text{tr}(AB) = \text{tr}(BA)$, for endomorphisms A, B of a free module. \square

Lemma 1.24. *The complex E has perfect amplitude contained in the interval $[-1, 0]$.*

Proof. Perfection is clear. To check the interval, note that by symmetry of E it suffices to check that the interval is $[-1, \infty]$. We have seen that $E = R\pi_*\mathcal{F}[2]$. So the interval is no worse than $[-2, \infty]$. But $h^{-2}(E) = 0$, by Assumption (iv), above. \square

Corollary 1.25. *The Deligne–Mumford stack X admits, in a natural way, a symmetric obstruction theory, namely*

$$E = R\pi_* R\mathcal{H}om(\mathcal{F}, \omega_Y)[2] = R\pi_* \mathcal{F}[2].$$

We call this obstruction theory the *Donaldson–Thomas* obstruction theory.

In the next two propositions we mention two special cases. The first was originally introduced in [Thomas 2000], where the symmetry was pointed out, too.

Proposition 1.26. *Let Y be a smooth projective threefold with trivial canonical bundle, and let X be the fine moduli stack of stable sheaves on Y of rank $r > 0$, with fixed determinant L and with Chern classes c_2, c_3 . Then X admits a symmetric obstruction theory.*

Proof. In fact, every trivialization $\omega_Y = \mathcal{O}_Y$ induces a symmetric obstruction theory. □

Proposition 1.27. *Let Y be a smooth projective threefold and D an effective anti-canonical divisor on Y . Let X' be the scheme of torsion-free rank 1 sheaves with trivial determinant and fixed Chern classes c_2, c_3 . Recall that such sheaves can be identified with ideal sheaves. Let $X \subset X'$ be the open subscheme consisting of ideal sheaves which define a subscheme of Y whose support is disjoint from D . Then X admits a symmetric obstruction theory.*

For example, $\text{Hilb}^n(Y \setminus D)$, the Hilbert scheme of length n subschemes of $Y \setminus D$ admits a symmetric obstruction theory.

Proof. Again, we would like to point out that every homomorphism $\omega_Y \rightarrow \mathcal{O}_Y$ defining D gives rise to a symmetric obstruction theory on X . Even though the compactification is used in its construction, this symmetric obstruction theory does not depend on which compactification is chosen. □

Stable maps to Calabi–Yau threefolds.

Proposition 1.28. *Let Y be a Calabi–Yau threefold and let X be the open locus in the moduli stack of stable maps parameterizing immersions of smooth curves. Then the Gromov–Witten obstruction theory of X is symmetric, in a natural way.*

Proof. Let $\pi : C \rightarrow X$ be the universal curve and $f : C \rightarrow Y$ the universal map. Let F be the kernel of $f^* \Omega_Y \rightarrow \Omega_C$, which is a vector bundle of rank 2 on C . The Gromov–Witten obstruction theory on X is $E = R\pi_*(F \otimes \omega_{C/X})[1]$. By Serre duality for $\pi : C \rightarrow X$, we have $E^\vee[1] = R\pi_*(F^\vee)[1]$.

As F is of rank 2, we have $F = F^\vee \otimes \det F$. Because Y is Calabi–Yau, we have $\det F \otimes \omega_{C/X} = \mathcal{O}_C$. Putting these two facts together, we get $F \otimes \omega_{C/X} = F^\vee$ and hence $E = E^\vee[1]$. □

2. Equivariant symmetric obstruction theories

A few remarks on equivariant derived categories. Let X be a scheme with an action of an algebraic group G . Let $(\text{Qcoh-}\mathbb{O}_X)^G$ denote the abelian category of G -equivariant quasicoherent \mathbb{O}_X -modules. Thus, an object of $(\text{Qcoh-}\mathbb{O}_X)^G$ is a quasicoherent \mathbb{O}_X -module F together with descent data to the quotient stack $[X/G]$, in other words an isomorphism between p^*F and σ^*F satisfying the cocycle condition. Here p and σ are projection and action maps $X \times G \rightarrow X$, respectively. Denote by $D(X)^G$ the derived category of $(\text{Qcoh-}\mathbb{O}_X)^G$. Note that \mathbb{O}_X is an object of $D(X)^G$, in a natural way.

There is the forgetful functor $D(X)^G \rightarrow D(X)$, which maps a complex of equivariant sheaves to its underlying complex of sheaves. It is an exact functor.

To simplify matters, let us make two assumptions:

- (a) X admits a G -equivariant ample invertible sheaf $\mathcal{O}(1)$,
- (b) G is a diagonalizable group, i.e., $G = \text{Spec } \mathbb{C}[W]$ is the spectrum of the group ring of a finitely generated abelian group W . Then W is canonically identified with the character group of G .

The affine case. If $X = \text{Spec } A$ is affine, A is W -graded. A G -equivariant \mathbb{O}_X -module is the same thing as a W -graded A -module.

We call a W -graded A -module *quasifree*, if it is free as an A -module on a set of homogeneous generators. Any quasifree W -graded A -module is isomorphic to a direct sum of shifted copies of A . Quasifree W -graded A -modules are projective objects in the abelian category $(\text{Qcoh-}\mathbb{O}_X)^G$ of W -graded A -modules. Hence this category has enough projective objects.

The global case. Let F be a G -equivariant \mathbb{O}_X -module. We can shift F by any character $w \in W$ of G . We denote the shift by $F[w]$. Every G -equivariant quasicoherent \mathbb{O}_X -module F can be written as a quotient of sheaf of the form

$$\bigoplus_{i \in I} \mathcal{O}(n_i)[w_i]. \quad (7)$$

Thus, every G -equivariant quasicoherent \mathbb{O}_X -module admits left resolutions consisting of objects of form (7). More generally, every bounded above complex in $D(X)^G$ can be replaced by a bounded above complex of objects of type (7). These resolutions are G -equivariant.

Since objects of the form (7) are locally free as \mathbb{O}_X -modules (forgetting the G -structure), we can use these resolutions to compute the derived functors of \otimes and $\mathcal{H}om(-, F)$. Thus we see that for G -equivariant quasicoherent \mathbb{O}_X -modules E, F the quasicoherent \mathbb{O}_X -modules $\mathcal{T}or_i(E, F)$ and $\mathcal{E}xt^i(E, F)$ are again G -equivariant. More generally, we see that for bounded above objects E, F of $D(X)^G$, the objects $E \overset{L}{\otimes} F$ and $R \mathcal{H}om(E, F)$ are again in $D(X)^G$.

For a G -equivariant sheaf E , we write $E^\vee = \mathcal{H}om(E, \mathbb{O}_X)$. For a bounded above object E of $D(X)^G$, we write $E^\vee = R\mathcal{H}om(E, \mathbb{O}_X)$.

Let $\{U_i\}$ be an invariant affine open cover. Let F be a G -equivariant quasi-coherent \mathbb{O}_X -module. Then, the Čech resolution $\mathcal{C}^\bullet(\{U_i\}, F)$ is a right resolution of F by G -equivariant quasicoherent \mathbb{O}_X -modules. It is an acyclic resolution for the global section functor, showing that the cohomology groups $H^i(X, F)$ are W -graded. More generally, let $f : X \rightarrow Y$ be a G -equivariant morphism. Then we see that $R^i f_* F$ are G -equivariant quasicoherent \mathbb{O}_Y -modules.

Moreover, if E is a bounded below complex in $D(X)^G$, we can construct the associated Čech complex $\mathcal{C}(\{U_i\}, E)$, which is a double complex. Passing to the associated single complex, we see that we may replace E by a bounded below complex of G -equivariant \mathbb{O}_X -modules which are acyclic for f_* , for any G -equivariant morphism $f : X \rightarrow Y$. Thus we see that the functor $Rf : D(X) \rightarrow D(Y)$ passes to a functor $Rf : D(X)^G \rightarrow D(Y)^G$.

The cotangent complex. If X is a G -scheme as above, the sheaf of Kähler differentials Ω_X and its dual $T_X = \Omega_X^\vee$ are G -equivariant.

We can use the equivariant ample line bundle L to construct a G -equivariant embedding $X \hookrightarrow M$ into a smooth G -scheme M . The cotangent complex $I/I^2 \rightarrow \Omega_M|_X$ is then naturally an object of $D(X)^G$. The usual proof that L_X is a canonically defined object of $D(X)$ works equivariantly and proves that L_X is a canonically defined object of $D(X)^G$. By canonically defined, we mean that any two constructions are related by a canonical isomorphism.

Perfect objects. We call an object E in $D(X)^G$ *perfect* (of perfect amplitude in the interval $[m, n]$), if its underlying object of $D(X)$ is perfect (of perfect amplitude in the interval $[m, n]$).

Remark 2.1. If X is a scheme and E in $D(X)$ is a perfect complex, of perfect amplitude contained in $[m, n]$, then we can write E locally as a complex

$$[E^m \rightarrow \dots \rightarrow E^n]$$

of free \mathbb{O}_X -modules contained in the interval $[m, n]$. This is essentially because if $E \rightarrow E''$ is an epimorphism of locally free coherent sheaves, the kernel is again locally free coherent.

In the equivariant context, we have to forgo this convenient fact. Suppose E in $D(X)^G$ is perfect, again of amplitude contained in $[m, n]$. We can, as we saw above, write E as a bounded above complex of sheaves of form (7), all of them coherent, i.e., with finite indexing set I . But when we cut off this infinite complex to fit into the interval $[m, n]$, we end up with a G -equivariant quasicoherent sheaf which is locally free coherent as an \mathbb{O}_X -module without the G -structure, but which is not locally quasifree and not locally projective in the category $(\text{Qcoh-}\mathbb{O}_X)^G$.

Symmetric equivariant obstruction theories.

Definition 2.2. Let X be a scheme with a G -action. An *equivariant perfect obstruction theory* is a morphism $E \rightarrow L_X$ in the category $D(X)^G$, which is a perfect obstruction theory as a morphism in $D(X)$. (This definition is originally due to Graber and Pandharipande [1999].)

A *symmetric equivariant obstruction theory*, (or an *equivariant symmetric obstruction theory*) is a pair $(E \rightarrow L_X, E \rightarrow E^\vee[1])$ of morphisms in the category $D(X)^G$, such that $E \rightarrow L_X$ is an (equivariant) perfect obstruction theory and $\theta : E \rightarrow E^\vee$ is an isomorphism satisfying $\theta^\vee[1] = \theta$.

This is more than requiring that the obstruction theory be equivariant and symmetric, separately, as we can see in the following example.

Example 2.3. Let $\omega = \sum_i^n f_i dx_i$ be an almost closed 1-form on \mathbb{A}^n . Recall from Example 1.19 that ω defines a symmetric obstruction theory

$$H(\omega) = [T_M|_X \xrightarrow{\nabla\omega} \Omega_M|_X]$$

on the zero locus X of ω .

Define a \mathbb{G}_m -action on \mathbb{A}^n by setting the degree of x_i to be r_i , where $r_i \in \mathbb{Z}$. Assume that each f_i is homogeneous with respect to these degrees and denote the degree of f_i by n_i . Then the zero locus X of ω inherits a \mathbb{G}_m -action.

If we let \mathbb{G}_m act on T_M by declaring the degree of $\partial/\partial x_i$ to be equal to n_i , then $H(\omega)$ is \mathbb{G}_m -equivariant as well as the morphism $H(\omega) \rightarrow L_X$. Thus $H(\omega)$ is an equivariant obstruction theory.

But note that $H(\omega)$ is not equivariant symmetric. This is because the identity on $H(\omega)$ (which is θ in this case) is not \mathbb{G}_m -equivariant if we consider it as a homomorphism

$$H(\omega) \rightarrow H(\omega)^\vee[1].$$

Unless $n_i = -r_i$, because then the degree of $\partial/\partial x_i$ is equal to its degree as the dual of dx_i .

In the case $n_i = -r_i$, the form $\omega = \sum f_i dx_i$ is an invariant element of $\Gamma(M, \Omega_M)$, or an equivariant homomorphism $\mathcal{O}_M \rightarrow \Omega_M$. In this case we do get an equivariant symmetric obstruction theory.

The equivariant Donaldson–Thomas obstruction theory. Let G be a diagonalizable group as above. Consider a projective threefold Y , endowed with a linear G -action. Consider a G -equivariant nonzero homomorphism $\omega_Y \rightarrow \mathcal{O}_Y$, defining the G -invariant anticanonical Cartier divisor D .

Proposition 2.4. *Let X be as in Proposition 1.27. Then the Donaldson–Thomas obstruction theory of Corollary 1.25 on X is G -equivariant symmetric.*

Proof. Let X' be the compactification of X as in Proposition 1.27. Let \mathcal{E} be the universal sheaf on $Y \times X$ and $Z \subset Y \times X$ be the universal subscheme. We have an exact sequence

$$0 \longrightarrow \mathcal{E} \longrightarrow \mathcal{O}_{Y \times X} \longrightarrow \mathcal{O}_Z \longrightarrow 0.$$

Let $\pi : Y \times X \rightarrow X$ be the projection. Note that \mathcal{E} and \mathcal{O}_Z are G -equivariant. This follows directly from the universal mapping property of \mathcal{E} .

The standard ample invertible sheaf on X' is $\det \pi_* (\mathcal{O}_Z(n))$, for n sufficiently large. It is G -equivariant, as all ingredients in its construction are. Hence X admits an equivariant ample invertible sheaf.

Next, notice that all the constructions involved in producing the obstruction theory $E = R\pi_* R\mathcal{H}om(\mathcal{F}, \omega_Y)[2]$ work equivariantly. Hence the symmetric obstruction theory is equivariant.

To prove that it is equivariant symmetric, we just need to remark that the bilinear form θ is induced from $\omega \rightarrow \mathcal{O}_Y$, which is equivariant, and that Serre duality is equivariant, because it is natural. \square

Local structure in the \mathbb{G}_m -case. Let $G = \mathbb{G}_m$. We will prove that Example 2.3 describes the unique example of a symmetric \mathbb{G}_m -equivariant obstruction theory, at least locally around a fixed point.

Lemma 2.5. *Let X be an affine \mathbb{G}_m -scheme with a fixed point P . Let n denote the dimension of $T_X|_P$, the Zariski tangent space of X at P . Then there exists an invariant affine open neighborhood X' of P in X , a smooth \mathbb{G}_m -scheme M of dimension n and an equivariant closed embedding $X' \hookrightarrow M$*

Proof. Let A be the affine coordinate ring of X . The \mathbb{G}_m -action induces a grading on A . Let \mathfrak{m} be the maximal ideal given by the point P . We can lift an eigenbasis of $\mathfrak{m}/\mathfrak{m}^2$ to homogeneous elements x_1, \dots, x_n of \mathfrak{m} . Choose homogeneous elements y_1, \dots, y_m in \mathfrak{m} in such a way that $x_1, \dots, x_n, y_1, \dots, y_m$ is a set of generators of A as a \mathbb{C} -algebra. This defines a closed embedding $X \hookrightarrow \mathbb{A}^{n+m}$, which is equivariant if we define a \mathbb{G}_m -action on \mathbb{A}^{n+m} in a suitable, obvious, way.

We have thus written A as a quotient of $\mathbb{C}[x, y]$. Let I denote the corresponding homogeneous ideal in $\mathbb{C}[x, y]$. Then we have

$$\mathfrak{m}/\mathfrak{m}^2 = (x, y)/(I + (x, y)^2).$$

Since this \mathbb{C} -vector space is generated by x_1, \dots, x_n , we have, in fact,

$$y_i \in I + (x, y)^2 + (x),$$

for $i = 1, \dots, m$. We can therefore find homogeneous elements $f_1, \dots, f_m \in I$ such that

$$y_i - f_i \in (x, y)^2 + (x) \quad \text{and} \quad \deg f_i = \deg y_i,$$

for all $i = 1, \dots, m$. Let $g \in \mathbb{C}[x, y]$ be the determinant of the Jacobian matrix $(\partial f_i / \partial y_j)$. We see that g is homogeneous of degree 0 and that $g(0, 0) = 1$. Let $U \subset \mathbb{A}^{n+m}$ be the affine open subset where g does not vanish. This is an invariant subset containing P . Let $Z \subset \mathbb{A}^{n+m}$ be the closed subscheme defined by (f_1, \dots, f_m) . It carries an induced \mathbb{G}_m -action. The intersection $M = Z \cap U$ is a smooth scheme of dimension n .

As (f_1, \dots, f_m) is contained in I , X is a closed subscheme of Z . Let $X' = X \cap U$. □

Proposition 2.6. *Let X be an affine \mathbb{G}_m -scheme with a fixed point P and let $n = \dim T_X|_P$. Furthermore, let X be endowed with a symmetric equivariant obstruction theory $E \rightarrow L_X$. Then there exists an invariant affine open neighborhood X' of P in X , an equivariant closed embedding $X' \hookrightarrow M$ into a smooth \mathbb{G}_m -scheme M of dimension n and an invariant almost closed 1-form ω on M such that $X = Z(\omega)$. We can further construct an equivariant isometry $E \rightarrow H(\omega)$ commuting with the maps to L_X , but it will not be necessary for the purposes of this paper.*

Proof. We apply Lemma 2.5, to obtain the equivariant closed embedding $X' \hookrightarrow M$. Write X for X' . Let A be the affine coordinate ring of X and I the ideal of $\Gamma(\mathbb{C}_M)$ defining X .

Consider the object E of $D(X)^{\mathbb{G}_m}$. We can represent E by an infinite complex $[\dots \rightarrow E_1 \rightarrow E_0]$ of finitely generated quasifree A -modules.

Because quasifree modules are projective, if E is represented by a bounded complex of quasifree modules as above and $E \rightarrow F$ is a morphism in $D(X)^G$, then $E \rightarrow F$ can be represented by an actual morphism of complexes, without changing E .

Thus we have morphisms of complexes of graded modules

$$[\dots \rightarrow E_1 \rightarrow E_0] \rightarrow [I/I^2 \rightarrow \Omega_M|_X]$$

and

$$\theta : [\dots \rightarrow E_1 \rightarrow E_0] \rightarrow [E_0^\vee \rightarrow E_1^\vee \rightarrow \dots].$$

We can represent the equality of derived category morphisms $\theta^\vee[1] = \theta$ by a homotopy between $\theta^\vee[1]$ and θ , because E is a bounded above complex of quasifrees. Then, as in the proof of Lemma 1.6, we can replace θ_0 by $\frac{1}{2}(\theta_0 + \theta_1^\vee)$ and θ_1 by $\frac{1}{2}(\theta_1 + \theta_0^\vee)$, without changing the homotopy class of θ . Then $\theta_1 = \theta_0^\vee$.

Now we can replace E_1 by $\text{cok}(E_2 \rightarrow E_1)$ and E_1^\vee by $\ker(E_1^\vee \rightarrow E_2^\vee)$. Because of the perfection of E , both $\text{cok}(E_2 \rightarrow E_1)$ and $\ker(E_1^\vee \rightarrow E_2^\vee)$ are projective A -modules (after forgetting the grading), which are, moreover, dual to each other.

Thus we have now represented E by a complex $[E_1 \rightarrow E_0]$ of equivariant vector bundles and $E \rightarrow L_X$ and $\theta : E \rightarrow E^\vee[1]$ by equivariant morphisms of complexes.

Moreover, $\theta = (\theta_0^\vee, \theta_0)$, for an equivariant morphism of vector bundles $\theta_0 : E_0 \rightarrow E_1^\vee$.

Now we remark that we may assume that the rank of E_0 is equal to n . Simply lift a homogeneous basis of $\Omega_X|_P$ to E_0 and replace E_0 by the quasifree module on these n elements of E_0 . Then pass to an invariant open neighborhood of P over which both $E_0 \rightarrow \Omega_M|_X$ and $\theta_0 : E_0 \rightarrow E_1^\vee$ are isomorphisms. Use these isomorphisms to identify. Then our obstruction theory is given by an equivariant homomorphism

$$\begin{array}{ccc} [T_M|_X & \xrightarrow{\alpha} & \Omega_M|_X \\ \phi \downarrow & & \downarrow \text{id} \\ [I/I^2 & \longrightarrow & \Omega_M|_X \end{array}$$

such that $\alpha^\vee = \alpha$. Note that ϕ is necessarily surjective.

As we may assume that $\Omega_M|_X$ and hence $T_M|_X$ is given by a quasifree A -module, we may lift ϕ to an equivariant epimorphism $T_X \rightarrow I$. This gives the invariant 1-form ω . □

3. The main theorem

Preliminaries on linking numbers. Here our dimensions are all *real* dimensions.

We work with orbifolds. Orbifolds are differentiable stacks of Deligne–Mumford type, which means that they are representable by Lie groupoids $X_1 \rightrightarrows X_0$, where source and target maps $X_1 \rightarrow X_0$ are étale (i.e., local diffeomorphisms) and the diagonal $X_1 \rightarrow X_0 \times X_0$ is proper. If a compact Lie group G acts with finite stabilizers on a manifold X , the quotient stack $[X/G]$ is an orbifold.

All our orbifolds will tacitly assumed to be *oriented*, which means that any presenting groupoid $X_1 \rightrightarrows X_0$ is oriented, i.e., X_0 and X_1 are oriented and all structure maps (in particular source and target $X_1 \rightarrow X_0$) preserve orientations.

Given an orbifold X , presented by the groupoid $X_1 \rightrightarrows X_0$, with proper diagonal $X_1 \rightarrow X_0 \times X_0$, the image of the diagonal is a closed equivalence relation on X_0 . The quotient is the *coarse moduli space* of X .

We call an orbifold *compact*, if its coarse moduli space is compact. More generally, we call a morphism $f : X \rightarrow Y$ of orbifolds *proper*, if the induced map on coarse moduli spaces is proper.

To fix ideas, let $H^*(X)$ denote de Rham cohomology of the orbifold X . For the definition and basic properties of this cohomology theory, see [Behrend 2004]. Note that homotopy invariance holds: the projection $X \times \mathbb{R} \rightarrow X$ induces an isomorphism $H^*(X) \rightarrow H^*(X \times \mathbb{R})$.

If $f : X \rightarrow Y$ is a proper morphism of orbifolds, there exists a wrong way map $f_! : H^i(X) \rightarrow H^{i-d}(Y)$, where $d = \dim X - \dim Y$ is the relative dimension of

f . If Y is the point, then we also denote $f_!$ by \int_X . We will need the following properties of $f_!$:

(i) Functoriality: $(g \circ f)_! = g_! \circ f_!$.

(ii) Naturality: if $v : V \subset Y$ is an open suborbifold and $u : U \subset X$ the inverse image of U under $f : X \rightarrow Y$, we have $v^* \circ f_! = g_! \circ u^*$, where $g : U \rightarrow V$ is the restriction of f .

(iii) Projection formula: $f_!(f^*(\alpha) \cup \beta) = \alpha \cup f_!(\beta)$.

(iv) Poincaré duality: if X is a compact orbifold, the pairing $\int_X \alpha \cup \beta$ between $H^i(X)$ and $H^{n-i}(X)$ is a perfect pairing of finite dimensional \mathbb{R} -vector spaces ($n = \dim X$).

(v) Long exact sequence: if $\iota : Z \subset X$ is a closed suborbifold with open complement U , there is a long exact sequence ($c = \dim X - \dim Z$)

$$\dots \xrightarrow{\partial} H^{i-c}(Z) \xrightarrow{\iota_!} H^i(X) \longrightarrow H^i(U) \xrightarrow{\partial} H^{i-c+1}(Z) \longrightarrow \dots$$

In the situation of (v), we call $\text{cl}(Z) = \iota_!(1) \in H^c(X)$ the *class* of Z .

We could use any other cohomology theory with characteristic zero coefficients which satisfies these basic properties.

Remark 3.1. Let $T \subset \mathbb{R}$ be an open interval containing the points 0 and 1. Let Z and X be a compact orbifolds and $h : Z \times T \rightarrow X$ a differentiable morphism of orbifolds such that $h_0 : Z \times \{0\} \rightarrow X$ and $h_1 : Z \times \{1\} \rightarrow X$ are isomorphisms onto closed suborbifolds Z_0 and Z_1 of X . We call h a *differentiable homotopy* between Z_0 and Z_1 . It is not difficult to see, using Poincaré duality and homotopy invariance, that the existence of such an h implies that $\text{cl}(Z_0) = \text{cl}(Z_1) \in H^*(X)$.

Linking numbers and S^1 -actions. Let A and B be closed submanifolds, both of dimension p , of a compact manifold S of dimension $2p+1$. Assume that $H^{p+1}(S) = H^p(S) = 0$ and that $A \cap B = \emptyset$. For simplicity, assume also that p is odd.

Under these assumptions we can define the *linking number* $L_S(A, B)$ as follows. By our assumption, the boundary map $\partial : H^p(S \setminus B) \rightarrow H^0(B)$ is an isomorphism. Let $\beta \in H^p(S \setminus B)$ be the unique element such that $\partial\beta = 1 \in H^0(B)$. Via the inclusion $A \rightarrow S \setminus B$ we restrict β to A and set

$$L_S(A, B) = \int_A \beta.$$

Now assume A' is another closed submanifold of S of dimension p , and $A' \cap B = \emptyset$, too. Thus $L_S(A', B)$ is defined. We wish to compare $L_S(A', B)$ with $L_S(A, B)$.

Suppose $h : Z \times T \rightarrow S$ is a differentiable homotopy between A and A' , as in Remark 3.1. It is an obvious, well-known fact, that if the image of h is entirely contained in $S \setminus B$, then $L_S(A', B) = L_S(A, B)$. We wish to show that in the presence of an S^1 -action, $L_S(A', B) = L_S(A, B)$, even if $h(Z \times T)$ intersects B .

Proposition 3.2. *Let S^1 act on S with finite stabilizers. Assume that A, A' and B are S^1 -invariant. Finally, assume that there exists an S^1 -equivariant homotopy $h : T \times Z \rightarrow S$ from A to A' . Then $L_S(A', B) = L_S(A, B)$.*

Proof. The condition that h be equivariant means that S^1 acts on Z with finite stabilizers and that h is equivariant, i.e. $h(t, \gamma \cdot z) = \gamma \cdot h(t, z)$, for all $\gamma \in S^1$ and $(t, z) \in T \times Z$.

We form the quotient orbifold $X = [S/S^1]$, which is compact of dimension $2p$. It comes together with a principal S^1 -bundle $\pi : S \rightarrow X$. Let $\tilde{A}, \tilde{A}', \tilde{B}$ and \tilde{Z} be the quotient orbifolds obtained from A, A', B and Z . The homotopy h descends to a differentiable homotopy $h : T \times \tilde{Z} \rightarrow X$ between \tilde{A} and \tilde{A}' , proving that $\text{cl}(\tilde{A}) = \text{cl}(\tilde{A}') \in H^{p+1}(X)$. This conclusion is all we need the homotopy h for.

Next we will construct, for a fixed B , an element $\eta \in H^{p-1}(X)$, such that

$$L_S(A, B) = \int_X \eta \cup \text{cl}(\tilde{A}),$$

for any A , such that $A \cap B = \emptyset$. This will conclude the proof of the proposition.

In fact, let $\beta \in H^p(S \setminus B)$, such that $\partial\beta = 1 \in H^0(B)$. The S^1 -bundle $S \setminus B \rightarrow X \setminus \tilde{B}$ induces a homomorphism $\pi_! : H^p(S \setminus B) \rightarrow H^{p-1}(X \setminus \tilde{B})$. Note that the restriction $H^{p-1}(X) \rightarrow H^{p-1}(X \setminus \tilde{B})$ is an isomorphism, since the codimension of \tilde{B} in X is $p + 1$. Thus, there exists a unique $\eta \in H^{p-1}(X)$, such that

$$\eta|_{X \setminus \tilde{B}} = \pi_!\beta.$$

Hence

$$L_S(A, B) = \int_A \beta = \int_{\tilde{A}} \pi_!\beta = \int_{\tilde{A}} \eta = \int_X \eta \cup \text{cl}(\tilde{A}),$$

as claimed. The last equality follows from naturality of the wrong way maps and the projection formula. □

The proof of $v_X(P) = (-1)^n$. We return to the convention that dimensions are complex dimensions.

Let X be a scheme with a \mathbb{G}_m -action. Let $P \in X$ be a fixed point of this action. The point P is called an *isolated* fixed point, if 0 is not a weight of the induced action of \mathbb{G}_m on the Zariski tangent space $T_X|_P$.

Proposition 3.3. *Let M be a smooth scheme on which \mathbb{G}_m acts with an isolated fixed point $P \in M$. Let ω be an invariant (homogeneous of degree zero) almost closed 1-form on M and $X = Z(\omega)$. Assume $P \in X$. Then*

$$v_X(P) = (-1)^{\dim M}.$$

Proof. We will use the expression of $v_X(P)$ as a linking number from Proposition 4.22 of [Behrend 2005]. We choose étale homogeneous coordinates x_1, \dots, x_n

for M around P and the induced étale coordinates $x_1, \dots, x_n, p_1, \dots, p_n$ of Ω_M . Since the linking number in question is defined inside a sufficiently small sphere in Ω_M around P (and is a topological invariant), we may as well assume that $M = \mathbb{C}^n$ and P is the origin. Of course, ω is then a 1-form holomorphic (instead of algebraic) at the origin. We write $\omega = \sum_{i=1}^n f_i dx_i$.

As in [ibid.], for $\eta \in \mathbb{C}, \eta \neq 0$, we write Γ_η for the graph of the section $\frac{1}{\eta}\omega$ of Ω_M . It is defined as a subspace of Ω_M by the equations $\eta p_i = f_i(x)$. It is oriented so that $M \rightarrow \Gamma_\eta$ is orientation preserving.

For $t \in \mathbb{R}$, we write Δ_t for the subspace of Ω_M defined by the equations $tp_i = \bar{x}_i$. We orient Δ_1 in such a way that the map $\mathbb{C}^n \rightarrow \Delta_1$ given by $(x_1, \dots, x_n) \mapsto (x_1, \dots, x_n, \bar{x}_1, \dots, \bar{x}_n)$ preserves orientation. Then we orient all other Δ_t by continuity. This amounts to the same as saying that the map $(p_1, \dots, p_n) \mapsto (t\bar{p}_1, \dots, t\bar{p}_n, p_1, \dots, p_n)$ from \mathbb{C}^n to Δ_t preserves orientation up to a factor of $(-1)^n$.

Proposition 4.22 of [ibid.] says that for sufficiently small $\epsilon > 0$ there exists $\eta \neq 0$ such that $\Gamma'_\eta = \Gamma_\eta \cap S_\epsilon$ is a manifold disjoint from $\Delta'_1 = \Delta_1 \cap S_\epsilon$ and

$$v_X(P) = L_{S_\epsilon}(\Delta'_1, \Gamma'_\eta).$$

Here S_ϵ is the sphere of radius ϵ centered at the origin P in Ω_M . It has dimension $4n - 1$. Let us fix ϵ and η .

The given $\mathbb{G}_m = \mathbb{C}^*$ -action on M induces an action on $\Omega_M = \mathbb{C}^{2n}$. Let us denote the degree of x_i by r_i . Then the degrees of p_i and f_i are both equal to $-r_i$. By restricting to $S^1 \subset \mathbb{C}^*$, we get an induced S^1 -action on S_ϵ . This action has finite stabilizers, because none of the r_i vanish, P being an isolated fixed point of the \mathbb{G}_m -action. Note that Γ'_η is an S^1 -invariant submanifold of S_ϵ .

Consider the map from $\mathbb{R} \times S^{2n-1} \rightarrow S_\epsilon$ given by

$$(t, p_1, \dots, p_n) \mapsto \frac{\epsilon}{\sqrt{1+t^2}}(t\bar{p}_1, \dots, t\bar{p}_n, p_1, \dots, p_n).$$

This map is an S^1 -equivariant homotopy between the invariant submanifolds $\Delta'_0 = \Delta_0 \cap S_\epsilon$ and Δ'_1 .

The fact that Δ'_1 is disjoint from Γ'_η follows from the fact that ω is almost closed, as explained in [ibid.]. The fact that Δ'_0 is disjoint from Γ'_η is trivial: Δ_0 is (up to orientation) the fiber of the vector bundle $\Omega_M \rightarrow M$ over the origin and Γ_η is the graph of a section. But there is no reason (at least none apparent to the authors) why there shouldn't exist values of t other than 0 or 1, for which $\Delta'_t = \Delta_1 \cap S_\epsilon$ intersects Γ_η .

Still, Proposition 3.2 implies that

$$L_{S_\epsilon}(\Delta'_1, \Gamma'_\eta) = L_{S_\epsilon}(\Delta'_0, \Gamma'_\eta).$$

Let us denote the fiber of Ω_M over the origin by $\overline{\Delta}_0$, and its intersection with S_ϵ by $\overline{\Delta}'_0$. By the correspondence between linking numbers and intersection numbers (see [Fulton 1984], Example 19.2.4), we see that $L_{S_\epsilon}(\overline{\Delta}'_0, \Gamma'_\eta)$ is equal to the intersection number of $\overline{\Delta}_0$ with Γ_η at the origin. This number is 1, as the section Γ_η intersects the fiber $\overline{\Delta}_0$ transversally.

Since the orientations of Δ_0 and $\overline{\Delta}_0$ differ by $(-1)^n$, we conclude that

$$v_X(P) = L_{S_\epsilon}(\Delta'_1, \Gamma'_\eta) = L_{S_\epsilon}(\Delta'_0, \Gamma'_\eta) = (-1)^n L_{S_\epsilon}(\overline{\Delta}'_0, \Gamma'_\eta) = (-1)^n,$$

which is what we set out to prove. □

Theorem 3.4. *Let X be an affine \mathbb{G}_m -scheme with an isolated fixed point P . Assume that X admits an equivariant symmetric obstruction theory. Then*

$$v_X(P) = (-1)^{\dim T_X|_P}.$$

Proof. Let $n = \dim T_X|_P$. By Proposition 2.6, we can assume that X is embedded equivariantly in a smooth scheme M of dimension n and that X is the zero locus of an invariant almost closed 1-form on M . Note that the embedding $X \hookrightarrow M$ identifies $T_X|_P$ with $T_M|_P$, so that P is an isolated point of the \mathbb{G}_m -action on M . Thus Proposition 3.3 implies that $v_X(P) = (-1)^n$. □

Corollary 3.5. *Let X be a \mathbb{G}_m -scheme such that all fixed points are isolated and every fixed point admits an invariant affine open neighborhood over which there exists an equivariant symmetric obstruction theory. Then we have*

$$\tilde{\chi}(X) = \sum_P (-1)^{\dim T_X|_P},$$

the sum extending over the fixed points. Moreover, if $Z \subset X$ is an invariant locally closed subscheme, we have

$$\tilde{\chi}(Z, X) = \sum_{P \in Z} (-1)^{\dim T_X|_P},$$

the sum extending over the fixed points in Z .

Proof. The product property of v implies that v_X is constant on nontrivial \mathbb{G}_m -orbits. The Euler characteristic of a scheme on which \mathbb{G}_m acts without fixed points is zero. These two facts imply that only the fixed points contribute to $\tilde{\chi}(X) = \chi(X, v_X)$. □

Corollary 3.6. *Let X be a projective scheme with a linear \mathbb{G}_m -action. Let X be endowed with an equivariant symmetric obstruction theory. Assume all fixed points of \mathbb{G}_m on X are isolated. Then we have*

$$\#\text{vir}(X) = \sum_P (-1)^{\dim T_X|_P},$$

the sum extending over the fixed points of \mathbb{G}_m on X .

Proof. We use the fact that X can be equivariantly embedded into a smooth scheme to prove that every fixed point has an invariant affine open neighborhood. Thus Corollary 3.5 applies. The main result of [Behrend 2005], Theorem 4.18, says that $\#^{\text{vir}} = \tilde{\chi}(X)$. \square

Application to Lagrangian intersections. Let M be an algebraic symplectic manifold with a Hamiltonian \mathbb{G}_m -action. Assume all fixed points are isolated. Let V and W be invariant Lagrangian submanifolds, X their intersection.

Proposition 3.7. *We have*

$$\tilde{\chi}(X) = \sum_{P \in X} (-1)^{\dim T_X|_P},$$

the sum extending over all fixed points inside X .

Proof. One checks that the action of \mathbb{G}_m being Hamiltonian, i.e., that \mathbb{G}_m preserves the symplectic form, implies that the symmetric obstruction theory on X is equivariant symmetric. \square

Proposition 3.8. *Assume X is compact. Then*

$$\deg([V] \cap [W]) = \sum_{P \in X} (-1)^{\dim T_X|_P},$$

the sum extending over the fixed points contained in X .

Proof. Note that, in fact, the virtual number of points of X is the intersection number of V and W . \square

Corollary 3.9. *Assume that X is compact and that $\dim T_X|_P$ is even, for all fixed points P . Then we have*

$$\deg([V] \cap [W]) = \chi(X).$$

4. Hilbert schemes of threefolds

The threefold \mathbb{A}^3 . Let $T = \mathbb{G}_m^3$ be the standard 3-dimensional torus with character group \mathbb{Z}^3 . Let T_0 be the kernel of the character $(1, 1, 1)$. Thus,

$$T_0 = \{(t_1, t_2, t_3) \in T \mid t_1 t_2 t_3 = 1\}.$$

We let T act in the natural way on \mathbb{A}^3 . Write coordinates on \mathbb{A}^3 as x, y, z , then, as elements of the affine coordinate ring $\mathbb{C}[x, y, z]$ of \mathbb{A}^3 , the weight of x is $(1, 0, 0)$, the weight of y is $(0, 1, 0)$ and the weight of z is $(0, 0, 1)$.

We choose on \mathbb{A}^3 the standard 3-form $dx \wedge dy \wedge dz$ to fix a Calabi–Yau structure. The torus T_0 acts by automorphisms of \mathbb{A}^3 preserving the Calabi–Yau structure.

by Proposition 2.4 we obtain a T_0 -equivariant symmetric obstruction theory on $X = \text{Hilb}^n \mathbb{A}^3$.

Lemma 4.1. (a) *The T_0 -action on X has a finite number of fixed points. These correspond to monomial ideals in $\mathbb{C}[x, y, z]$.*

(b) *If I is such an ideal, the T_0 -action on the Zariski tangent space to X at I has no invariant subspace.*

(c) *If I is such an ideal and d is the dimension of the Zariski tangent space to X at I , we have $(-1)^d = (-1)^n$, in other words, the integer d has the same parity as n .*

Proof. (a) Since the T_0 -action on \mathbb{A}^3 has the origin as unique fixed point, any invariant subscheme must be supported at the origin. Let $I \subset \mathbb{C}[x, y, z]$ be the corresponding ideal; I must be generated by eigenvectors of the torus action on the polynomial ring. Any eigenvector can be written uniquely in the form $m g(xyz)$ where m is a monomial and $g \in \mathbb{C}[t]$ is a polynomial with $g(0) \neq 0$. However, since the ideal is supported at the origin, the zero locus of $g(xyz)$ is disjoint from the zero locus of I , and so by Hilbert's Nullstellensatz, the monomial m is also in I . Hence every T_0 -invariant ideal is generated by monomials.

(b) Let us write $A = \mathbb{C}[x, y, z]$. The tangent space in question is $\text{Hom}(I, A/I)$. We will prove that none of the weights $w = (w_1, w_2, w_3)$ of T on $\text{Hom}_A(I, A/I)$ can satisfy $w_1, w_2, w_3 < 0$ or $w_1, w_2, w_3 \geq 0$. In particular, none of these weights can be an integer multiple of $(1, 1, 1)$.

This will suffice, in view of the following elementary fact: Let $w_1, \dots, w_n \in \mathbb{Z}^3$ be characters of T . If none of the w_i is an integer multiple of $(1, 1, 1)$, there exists a one-parameter subgroup $\lambda : \mathbb{G}_m \hookrightarrow T_0$, such that $w_i \circ \lambda \neq 0$, for all $i = 1, \dots, n$.

Suppose, then, that $\phi : I \rightarrow A/I$ is an eigenvector of T with weight (w_1, w_2, w_3) , with $w_1 \geq 0, w_2 \geq 0$ and $w_3 \geq 0$. Then for a monomial $x^a y^b z^c \in I$ we have $\phi(x^a y^b z^c) \equiv x^{a+w_1} y^{b+w_2} z^{c+w_3} \pmod{I}$, which vanishes in A/I , proving that $\phi = 0$.

Now suppose $\phi : I \rightarrow A/I$ is an eigenfunction whose weights satisfy $w_1 < 0, w_2 < 0$ and $w_3 < 0$. Let a be the smallest integer such that $x^a \in I$. Then let b be the smallest integer such that $x^{a-1} y^b \in I$. Finally, let c be the smallest integer such that $x^{a-1} y^{b-1} z^c \in I$. Then if a monomial $x^r y^s z^t$ is in I , it follows that $r \geq a, s \geq b$ or $t \geq c$.

We have

$$\phi(x^a y^b z^c) = x z^c \phi(x^{a-1} y^b) \equiv x z^c x^{a-1+w_1} y^{b+w_2} \equiv x^{a+w_1} y^{b+w_2} z^c \pmod{I}.$$

We also have

$$\phi(x^a y^b z^c) = x y \phi(x^{a-1} y^{b-1} z^c) \equiv x^{a+w_1} y^{b+w_2} z^{c+w_3} \pmod{I}.$$

We conclude that

$$x^{a+w_1} y^{b+w_2} z^c - x^{a+w_1} y^{b+w_2} z^{c+w_3} \in I.$$

Since the ideal I is monomial, each of these two monomials is in I . But the latter one cannot be in I .

(c) This is an immediate consequence of [Maulik et al. 2003], Theorem 2 in § 4.10. In fact, this theorem states that if w_1, \dots, w_d are the weights of T on the tangent space V ,

$$\frac{\prod_{i=1}^d (-w_i)}{\prod_{i=1}^d w_i} = (-1)^n$$

inside the field of rational functions on T . □

Proposition 4.2. *For any T_0 -invariant locally closed subset Z of $\text{Hilb}^n \mathbb{A}^3$ we have*

$$\tilde{\chi}(Z, \text{Hilb}^n \mathbb{A}^3) = (-1)^n \chi(Z).$$

Proof. Since there are only finitely many fixed points of T_0 on X , we can use the fact mentioned in the proof of Lemma 4.1 to find a one-parameter subgroup $\mathbb{G}_m \rightarrow T_0$ with respect to which all weights of all tangent spaces at all fixed points are nonzero. Thus, all \mathbb{G}_m -fixed points are isolated. Because $\text{Hilb}^n \mathbb{A}^3$ admits an equivariant embedding into projective space (see the proof of Proposition 2.4), every fixed point has an invariant affine open neighborhood.

The symmetric obstruction theory on $\text{Hilb}^n_{(n)} \mathbb{A}^3$ is equivariant symmetric with respect to the induced \mathbb{G}_m -action. We can therefore apply Corollary 3.5. We obtain:

$$\tilde{\chi}(Z, \text{Hilb}^n \mathbb{A}^3) = \sum_{P \in Z} (-1)^n,$$

where the sum extends over fixed points P contained in Z . Since $\chi(Z) = \#\{P \in Z, P \text{ fixed}\}$, the result follows. □

Let F_n denote the closed subset of $\text{Hilb}^n \mathbb{A}^3$ consisting of subschemes supported at the origin. Let ν_n be the restriction of the canonical constructible function $\nu_{\text{Hilb}^n \mathbb{A}^3}$ to F_n . Thus $\tilde{\chi}(F_n, \text{Hilb}^n \mathbb{A}^3) = \chi(F_n, \nu_n)$. Note that all T_0 -fixed points of $\text{Hilb}^n \mathbb{A}^3$ are contained in F_n .

Let $M(t) = \prod_{n=1}^{\infty} (1 - t^n)^{-n}$ be the McMahon function. It is the generating series for 3-dimensional partitions. Hence, if we write $M(t) = \sum_{n=0}^{\infty} p_n t^n$, then p_n denotes the number of monomial ideals I in $A = \mathbb{C}[x, y, z]$, such that $\dim_{\mathbb{C}} A/I = n$. The number p_n is the number of T_0 -fixed points in F_n or $\text{Hilb}^n \mathbb{A}^3$. Thus, $p_n = \chi(F_n) = \chi(\text{Hilb}^n \mathbb{A}^3)$.

Corollary 4.3. *We have*

$$\chi(F_n, \nu_n) = (-1)^n \chi(F_n) = (-1)^n p_n,$$

and hence

$$\sum_{n=0}^{\infty} \chi(F_n, \nu_n) t^n = M(-t).$$

Weighted Euler characteristics of Hilbert schemes. Let Y be a smooth threefold, and $n > 0$ an integer. Consider the Hilbert scheme of n points on Y , denoted $\text{Hilb}^n Y$. The scheme $\text{Hilb}^n Y$ is connected, smooth for $n \leq 3$ and singular otherwise, and reducible for large enough n .

Let us denote by ν_Y the canonical constructible function on $\text{Hilb}^n Y$. Our goal is to calculate

$$\tilde{\chi}(\text{Hilb}^n Y) = \chi(\text{Hilb}^n Y, \nu_Y).$$

Let us start with a useful general lemma on Hilbert schemes.

Lemma 4.4. *Let $f : Y \rightarrow Y'$ be a morphism of projective schemes and $Z \subset Y$ a closed subscheme. Assume that f is étale in a neighborhood of Z and that the composition $Z \rightarrow Y'$, which we will denote by $f(Z)$, is a closed immersion of schemes.*

Let X be the Hilbert scheme of Y which contains Z and P the point of X corresponding to Z . Let X' be the Hilbert scheme of Y' which contains $f(Z)$. Then there exists an open neighborhood U of P in X and an étale morphism $\phi : U \rightarrow X'$, which sends a subscheme $\tilde{Z} \rightarrow Y$ to the composition $\tilde{Z} \rightarrow Y'$.

Proof. For the existence of the open set U and the morphism ϕ , see for example Proposition 6.1, Chapter I of [Kollár 1996]. The fact that ϕ is étale in a neighborhood of P follows from a direct application of the formal criterion. \square

The closed stratum. We start by recalling the standard stratification of $\text{Hilb}^n Y$. The strata are indexed by partitions of n . Let $\alpha = (\alpha_1, \dots, \alpha_r)$ be a length r partition of n , i.e., $\alpha_1 \geq \alpha_2 \geq \dots \geq \alpha_r \geq 1$ and $\sum_{i=1}^r \alpha_i = n$. Let $\text{Hilb}_\alpha^n Y$ be the locus of subschemes whose support consists of r distinct points with multiplicities $\alpha_1, \dots, \alpha_r$. The closed stratum is $\text{Hilb}_{(n)}^n Y$. It corresponds to subschemes supported at a single point. To fix ideas, we will endow all strata with the reduced scheme structure.

Lemma 4.5. *For any threefold Y there is a natural morphism*

$$\pi_Y : \text{Hilb}_{(n)}^n Y \rightarrow Y.$$

Proof. This is a part of the Hilbert–Chow morphism $\text{Hilb}^n Y \rightarrow S^n Y$ to the symmetric product. A proof that this is a morphism of schemes can be found, for example, in [Lehn 2004]. \square

Note that F_n is the fiber of $\pi_{\mathbb{A}^3}$ over the origin.

Lemma 4.6. *We have a canonical isomorphism*

$$\text{Hilb}_{(n)}^n \mathbb{A}^3 = \mathbb{A}^3 \times F_n. \tag{8}$$

Moreover, $\nu_{\mathbb{A}^3} = p^* \nu_n$, where $p : \text{Hilb}_{(n)}^n \mathbb{A}^3 \rightarrow F_n$ is the projection given by (8).

Proof. Consider the action of the group \mathbb{A}^3 on itself by translations. We get an induced action of \mathbb{A}^3 on $\text{Hilb}^n \mathbb{A}^3$. Use this action to translate a subscheme supported at a point P to a subscheme supported at the origin. Obtain the morphism $p : \text{Hilb}_{(n)}^n \mathbb{A}^3 \rightarrow F_n$ in this way. The product morphism $\pi_{\mathbb{A}^3} \times p : \text{Hilb}_{(n)}^n \mathbb{A}^3 \rightarrow \mathbb{A}^3 \times F_n$ is an isomorphism.

It is a formal consequence of the general properties of the canonical constructible function, that it is constant on orbits under a group action. This implies the claim about $v_{\mathbb{A}^3}$. □

Lemma 4.7. *Consider an étale morphism of threefolds $\phi : Y \rightarrow Y'$.*

(a) *Let $U \subset \text{Hilb}^n Y$ be the open subscheme parameterizing subschemes $Z \subset Y$, which satisfy: if P and Q are distinct points in the support of Z , then $\phi(P) \neq \phi(Q)$. There is an étale morphism $\tilde{\Phi} : U \rightarrow \text{Hilb}^n Y'$ sending a subscheme of Y to its image under ϕ .*

$$\begin{array}{ccccc}
 \text{Hilb}_{(n)}^n Y & \longrightarrow & U & \longrightarrow & \text{Hilb}^n Y \\
 \Phi \downarrow & & \downarrow \tilde{\Phi} & & \\
 \text{Hilb}_{(n)}^n Y' & \longrightarrow & \text{Hilb}^n Y' & &
 \end{array}$$

(b) *The restriction of $\tilde{\Phi}$ to $\text{Hilb}_{(n)}^n Y$ induces a cartesian diagram of schemes*

$$\begin{array}{ccc}
 \text{Hilb}_{(n)}^n Y & \xrightarrow{\Phi} & \text{Hilb}_{(n)}^n Y' \\
 \pi_Y \downarrow & \square & \downarrow \pi_{Y'} \\
 Y & \xrightarrow{\phi} & Y'
 \end{array}$$

Proof. The existence and étaleness of $\tilde{\Phi}$ follows immediately from Lemma 4.4, applied to quasiprojective covers of Y and Y' . Part (b) is clear. □

Let $\phi : Y \rightarrow Y'$ be an étale morphism with induced morphism $\Phi : \text{Hilb}_{(n)}^n Y \rightarrow \text{Hilb}_{(n)}^n Y'$. By Lemma 4.7, the morphism Φ extends to open neighborhoods in $\text{Hilb}^n Y$ and $\text{Hilb}^n Y'$, respectively. The extension $\tilde{\Phi}$ is étale. Thus, we see that

$$\Phi^*(v_{Y'}) = v_Y.$$

Proposition 4.8. *Every étale morphism $\phi : Y \rightarrow \mathbb{A}^3$ induces an isomorphism $\text{Hilb}_{(n)}^n Y = Y \times F_n$. The constructible function $v_Y|_{\text{Hilb}_{(n)}^n Y}$ is obtained by pulling back v_n via the induced projection $\text{Hilb}_{(n)}^n Y \rightarrow F_n$.*

Proof. Combine Lemmas 4.6 and 4.7(b) with each other. □

Corollary 4.9. *The morphism $\pi_Y : \text{Hilb}_{(n)}^n Y \rightarrow Y$ is a Zariski-locally trivial fibration with fiber F_n . More precisely, there exists a Zariski open cover $\{U_i\}$ of Y , such*

that for every i , we have

$$(\pi_Y^{-1}(U_i), \nu_Y) = (U_i, 1) \times (F_n, \nu_n).$$

This is a product of schemes with constructible functions on them.

Proof. Every point of Y admits étale coordinates, defined in a Zariski open neighborhood. □

Reduction to the closed stratum. From now on the threefold Y will be fixed and we denote $\text{Hilb}_\alpha^n Y$ by X_α^n and $\text{Hilb}^n Y$ by X^n .

Lemma 4.10. *Let $\alpha = (\alpha_1, \dots, \alpha_r)$ be a partition of n .*

(a) *Let V be the open subscheme of $\prod_{i=1}^r X^{\alpha_i}$ parameterizing r -tuples of subschemes with pairwise disjoint support. Then there is a morphism $f_\alpha : V \rightarrow X^n$ mapping (Z_1, \dots, Z_r) to $Z = \bigcup_i Z_i$. The morphism f_α is étale. Its image U is open and contains X_α^n . Let $Z_\alpha = f_\alpha^{-1} X_\alpha^n$:*

$$\begin{array}{ccccc} Z_\alpha & \longrightarrow & V & \longrightarrow & \prod_i X^{\alpha_i} \\ \text{Galois} \downarrow & & \square & & \downarrow f_\alpha \\ X_\alpha^n & \longrightarrow & U & \longrightarrow & X^n \end{array}$$

Moreover, the induced morphism $Z_\alpha \rightarrow X_\alpha^n$ is a Galois cover with Galois group G_α , where G_α is the automorphism group of the partition α .

(b) *The scheme Z_α is contained in $\prod_i X_{(\alpha_i)}^{\alpha_i}$ and has therefore a morphism $Z_\alpha \rightarrow Y^r$. There is a cartesian diagram*

$$\begin{array}{ccc} Z_\alpha & \longrightarrow & \prod_i X_{(\alpha_i)}^{\alpha_i} \\ \downarrow & & \downarrow \\ Y_0^r & \longrightarrow & Y^r \end{array}$$

where Y_0^r is the open subscheme in Y^r consisting of r -tuples with pairwise disjoint entries.

Proof. The existence of f_α and the fact that it is étale follows from Lemma 4.4 applied to the étale map $\prod_{i=1}^r Y \rightarrow Y$ and the subscheme $Z_1 \sqcup \dots \sqcup Z_r \subset \prod_{i=1}^r Y$. All other facts are also straightforward to prove. □

Theorem 4.11. *Let Y be a smooth scheme of dimension 3. Then for all $n > 0$*

$$\tilde{\chi}(\text{Hilb}^n Y) = (-1)^n \chi(\text{Hilb}^n Y).$$

This implies

$$\sum_{n=0}^{\infty} \tilde{\chi}(\text{Hilb}^n Y) t^n = M(-t)^{\chi(Y)}.$$

Proof. By formal properties of $\tilde{\chi}$ as proved in [Behrend 2005], we can calculate as follows, using Lemma 4.10(a):

$$\begin{aligned} \tilde{\chi}(X^n) &= \sum_{\alpha \vdash n} \tilde{\chi}(X_\alpha^n, X^n) = \sum_{\alpha \vdash n} \tilde{\chi}(X_\alpha^n, U) = \sum_{\alpha \vdash n} |G_\alpha| \tilde{\chi}(Z_\alpha, V) \\ &= \sum_{\alpha \vdash n} |G_\alpha| \tilde{\chi}\left(Z_\alpha, \prod_i X^{\alpha_i}\right). \end{aligned}$$

By Lemma 4.10(b) and Corollary 4.9, $Z_\alpha \rightarrow Y_0^{\ell(\alpha)}$ is a Zariski-locally trivial fibration with fiber $\prod_i F_{\alpha_i}$. Here we have written $\ell(\alpha)$ for the length r of the partition $(\alpha_1, \dots, \alpha_r)$. We conclude:

$$\tilde{\chi}\left(Z_\alpha, \prod_i X^{\alpha_i}\right) = \chi(Y_0^{\ell(\alpha)}) \prod_i \chi(F_{\alpha_i}, \nu_{\alpha_i})$$

Together with Corollary 4.3 this gives:

$$\tilde{\chi}(X^n) = (-1)^n \sum_{\alpha \vdash n} |G_\alpha| \chi(Y_0^{\ell(\alpha)}) \prod_i \chi(F_{\alpha_i}). \tag{9}$$

Using the exact same arguments with the constant function 1 in place of ν gives the same answer, except without the sign $(-1)^n$. This proves our first claim. The second one follows then directly from the result of [Cheah 1996], which says that $\sum_{n=0}^\infty \chi(\text{Hilb}^n Y) t^n = M(t)^{\chi(Y)}$. □

The dimension zero MNOP conjecture. We now prove Conjecture 1 of [Maulik et al. 2003]. A proof of this result was also announced by J. Li at the workshop on Donaldson–Thomas invariants in Urbana–Champaign in March 2005.

Theorem 4.12. *Let Y be a projective Calabi–Yau threefold. Then, for the virtual count of $\text{Hilb}^n Y$ with respect to the Donaldson–Thomas obstruction theory, we have*

$$\#\text{vir}(\text{Hilb}^n Y) = (-1)^n \chi(\text{Hilb}^n Y).$$

In other words:

$$\sum_{n=0}^\infty \#\text{vir}(\text{Hilb}^n Y) t^n = M(-t)^{\chi(Y)}.$$

Proof. By Theorem 4.18 of [Behrend 2005], we have

$$\#\text{vir}(\text{Hilb}^n Y) = \tilde{\chi}(\text{Hilb}^n Y).$$

Thus the result follows from Theorem 4.11. □

Acknowledgments

We thank Jim Bryan and Lothar Göttsche for helpful discussions. We would especially like to thank Bumsig Kim and the Korea Institute for Advanced Study for the warm hospitality during a stay in Seoul, where the main results of this paper were proved.

References

- [Behrend 2004] K. Behrend, “Cohomology of stacks”, pp. 249–294 in *Intersection theory and moduli*, edited by E. Arbarello et al., ICTP Lecture Notes **19**, Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2004. MR 2006i:14016 Zbl 1081.58003
- [Behrend 2005] K. Behrend, “Donaldson–Thomas type invariants via microlocal geometry”, preprint, 2005. arXiv math.AG/0507523
- [Behrend and Fantechi 1997] K. Behrend and B. Fantechi, “The intrinsic normal cone”, *Invent. Math.* **128**:1 (1997), 45–88. MR 98e:14022 Zbl 0909.14006
- [Cheah 1996] J. Cheah, “On the cohomology of Hilbert schemes of points”, *J. Algebraic Geom.* **5**:3 (1996), 479–511. MR 97b:14005 Zbl 0889.14001
- [Fulton 1984] W. Fulton, *Intersection theory*, Ergebnisse der Mathematik (3) **2**, Springer, Berlin, 1984. MR 85k:14004 Zbl 0541.14005
- [Graber and Pandharipande 1999] T. Graber and R. Pandharipande, “Localization of virtual classes”, *Invent. Math.* **135**:2 (1999), 487–518. MR 2000h:14005 Zbl 0953.14035
- [Kollár 1996] J. Kollár, *Rational curves on algebraic varieties*, Ergebnisse der Mathematik (3) **32**, Springer, Berlin, 1996. MR 98c:14001 Zbl 0877.14012
- [Lehn 2004] M. Lehn, “Lectures on Hilbert schemes”, pp. 1–30 in *Algebraic structures and moduli spaces*, edited by J. Hurtubise and E. Markman, CRM Proc. Lecture Notes **38**, Amer. Math. Soc., Providence, RI, 2004. MR 2005g:14010 Zbl 1076.14010
- [Maulik et al. 2003] D. Maulik, N. Nekrasov, A. Okounkov, and R. Pandharipande, “Gromov–Witten theory and Donaldson–Thomas theory, I”, preprint, 2003. arXiv math.AG/0312059
- [Thomas 2000] R. P. Thomas, “A holomorphic Casson invariant for Calabi–Yau 3-folds, and bundles on $K3$ fibrations”, *J. Differential Geom.* **54**:2 (2000), 367–438. MR 2002b:14049 Zbl 1034.14015

Communicated by Andrei Okounkov

Received 2007-10-04 Accepted 2007-11-05

behrend@math.ubc.ca

*University of British Columbia, 1984 Mathematics Road,
Vancouver, BC V6T 1Z2, Canada*

fantechi@sissa.it

SISSA, Via Beirut 4, 34014 Trieste, Italy

Minimal γ -sheaves

Manuel Blickle

In a seminal work Lyubeznik [1997] introduces a category F -finite modules in order to show various finiteness results of local cohomology modules of a regular ring R in positive characteristic. The key notion on which most of his arguments rely is that of a *generator* of an F -finite module. This may be viewed as an R finitely generated representative for the generally nonfinitely generated local cohomology modules. In this paper we show that there is a functorial way to choose such an R -finitely generated representative, called the *minimal root*, thereby answering a question that was left open in Lyubeznik's work. Indeed, we give an equivalence of categories between F -finite modules and a category of certain R -finitely generated modules with a certain Frobenius operation which we call minimal γ -sheaves.

As immediate applications we obtain a globalization result for the parameter test module of tight closure theory and a new interpretation of the generalized test ideals of Hara and Takagi [2004] which allows us to easily recover the rationality and discreteness results for F -thresholds of Blickle et al. [2008].

1. Introduction

Let R be a regular ring of positive characteristic $p > 0$. We denote by $\sigma : R \rightarrow R$ the Frobenius map which sends $r \in R$ to its p -th power r^p . We assume that R is F -finite, which means that the Frobenius map σ is a finite morphism. In order to show various finiteness results for the generally nonfinitely generated local cohomology modules $H_J^i(R)$, with J being some ideal of R , Lyubeznik [1997] observes that if one enlarges the ring R by adjoining a new, noncommutative variable representing the Frobenius morphism σ then all local cohomology modules (on which the Frobenius acts in a natural way) are finitely generated over the resulting ring

$$R[\sigma] = \frac{R\{\sigma\}}{\langle \sigma r - r^p \sigma \mid r \in R \rangle}.$$

MSC2000: 13A35.

Keywords: positive characteristic, D-module, F-module, Frobenius operation.

During the preparation of this article the author was supported by the *DFG Schwerpunkt Komplexe Geometrie*. Some part of the research was done while the author was visiting the Institute Mittag-Leffler, Djursholm, Sweden. Their hospitality and financial support are greatly appreciated.

Note that the datum of a module over this ring is the same as giving an R -module \mathcal{M} together with a p -linear map $\sigma_{\mathcal{M}} : \mathcal{M} \rightarrow \mathcal{M}$ representing the action of σ . By adjointness of restriction and extension, the p -linear map $\sigma_{\mathcal{M}}$ is equivalent to an R -linear map $\theta : \sigma^* \mathcal{M} \stackrel{\text{def}}{=} \mathcal{M} \otimes_{\sigma} R \rightarrow \mathcal{M}$. This leads to the key definition:

Definition 1.1. A finitely generated $R[\sigma]$ -module \mathcal{M} is called a *unit* module if the structural morphism $\theta : \sigma^* \mathcal{M} \rightarrow \mathcal{M}$ is an isomorphism.

Lyubeznik [1997] shows many good properties of the category of finitely generated unit $R[\sigma]$ -modules.¹ Most notably, they form an abelian category in which every object has finite length. By observing that local cohomology modules are finitely generated unit $R[\sigma]$ -modules he is able to conclude many strong finiteness results about them. A systematic study of a globalized version of finitely generated unit modules is undertaken by Emerton and Kisin [2004]. There it is shown that the category of locally finitely generated $R[\sigma]$ -modules is (derived) equivalent to the category of étale sheaves of \mathbb{F}_p -vector spaces on $\text{Spec } R$.

A prominent role in Lyubeznik’s as well as in Emerton and Kisin’s study of finitely generated unit $R[\sigma]$ -modules has the concept of a *generator* or *root*. A *root* of a finitely generated unit $R[\sigma]$ -module \mathcal{M} is a finitely generated R -submodule $M \subseteq \mathcal{M}$ such that the structural map θ_M induces an inclusion $\gamma : M \subseteq \sigma^* M$ and that $\bigcup \sigma^{e^*} M = \mathcal{M}$. Hence one may view the root $\gamma : M \rightarrow \sigma^* M$ as a coherent representative of the finitely generated unit $R[\sigma]$ -module \mathcal{M} . Generalizing slightly we define:

Definition 1.2. A finitely generated R -module M together with an R -linear morphism $\gamma : M \rightarrow \sigma^* M$ is called a γ -*module*, or γ -sheaf.

The Frobenius iterates of the map γ form a directed system

$$M \xrightarrow{\gamma} \sigma^* M \xrightarrow{\sigma^* \gamma} \sigma^{2^*} M \xrightarrow{\sigma^{2^*} \gamma} \dots ,$$

the limit of which we denote by $\text{Gen}M$. One checks easily that γ induces a map $\text{Gen}M \rightarrow \sigma^* \text{Gen}M$ which is an isomorphism. By inverting this isomorphism, $\text{Gen}M$ becomes a unit $R[\sigma]$ -module. It is shown in [Emerton and Kisin 2004] that a finitely generated unit $R[\sigma]$ -module \mathcal{M} is precisely a module which is isomorphic to $\text{Gen}M$ for some γ -module (M, γ) .

Of course, different γ -modules may generate isomorphic unit $R[\sigma]$ -modules. The obvious question whether there is a unique minimal (in an appropriate sense, see Definition 2.7) γ -module that generates a given unit $R[\sigma]$ -module has remained open for a long time. In the case that R is complete, a positive answer was given already in [Lyubeznik 1997, Theorem 3.5]. In [Blickle 2004] this is extended to

¹In [Lyubeznik 1997] this category is called F -finite modules. We follow here the notation of Blickle and Böckle [\geq 2008] which in turn is taken from the monograph [Emerton and Kisin 2004].

the case that R is local (at least if R is F -finite). The purpose of this article is to show this in general, that is, for any F -finite regular ring R (see Theorem 2.24). A notable point in the proof is that it does not rely on the hard finiteness result [Lyubeznik 1997, Theorem 4.2], but only on the (easier) local case of it which is in some sense proven here *en passant* (see Remark 2.14). Our main result is the following.

Theorem 2.27. *Let R be a regular ring of positive characteristic $p > 0$ such that the Frobenius map is finite. Let \mathcal{M} be a finitely generated unit $R[\sigma]$ -module, then there is a unique minimal γ -sheaf M such that $\text{Gen}M \cong \mathcal{M}$.*

Moreover, the functor Gen induces an equivalence of abelian categories between finitely generated unit $R[\sigma]$ -modules and minimal γ -sheaves.

Our approach to prove this result is not the most direct one imaginable since we essentially develop a theory of minimal γ -sheaves from scratch (Section 2). The benefit is that after this is established, the result on the existence of minimal roots naturally appears as a byproduct. For this reason, it is important to isolate the key point in the argument: For a fixed coherent γ -sheaf M , the order of nilpotency of quotients of M is universally bounded. This is the statement of Proposition 2.11 (local case) and of the main result Theorem 2.22 (general case). The proof of this comes down to checking that decreasing sequences of γ -subsheaves of a fixed coherent γ -sheaf are eventually constant. This is achieved using the Chevalley lemma in the local case, or, via duality, by invoking a key result of Hartshorne and Speiser [1977]. The main difficulty (and the achievement in this paper) however lies in reducing the general case to the local case.

The quite explicit nature of our proof allows us to draw a series of interesting consequences. In particular, the connection to generalized test ideals [Hara and Takagi 2004] which appeared in computing the simplest examples of minimal gamma sheaves is quite surprising at first. In Section 3 we also give some applications of the result on the existence of minimal γ -sheaves. First, we show that the category of minimal γ -sheaves is equivalent to the category γ -crystals of Blickle and Böckle [\geq 2008]. We show that a notion from tight closure theory, namely the parameter test module, is a global object (Proposition 3.3). Statements of this type are notoriously hard in the theory of tight closure, particularly in the light of recent evidence that localization for tight closure might fail in general. Furthermore, we give a concrete description of minimal γ -sheaves in a very simple case (Proposition 3.5), relating it to the generalized test ideals studied in [Blickle et al. 2008]. This viewpoint also recovers (and slightly generalizes, with new proofs) the main results on the rationality and discreteness of jumping numbers of Blickle et al. [2008] and the results on generators of certain D -modules of Alvarez-Montaner et al. [2005]. A similar generalization, however using slightly different (but related, see

Remark 2.12) methods, was recently obtained independently by Katzman et al. [2007].

We are pleased to recently have learned that Carl Miller obtained the existence of minimal roots in the case of dimension 1 independently in his PhD thesis [Miller 2007]. He uses the existence of minimal roots in this case as the main tool to obtain a lower bound for the Euler characteristic of a p -torsion étale sheaf on a smooth characteristic p curve, thereby answering a question of Pink (who considered the case where there is no higher wild ramification). The link between the p -torsion sheaf and the coherent minimal γ -sheaf is obtained via the Riemann–Hilbert type correspondence of Emerton and Kisin [2004]. I expect that the results in the present paper could be of great use for similar applications to the study of p -torsion sheaves via coherent sheaves. For example, starting with a suitable p -torsion étale sheaf F , one can now uniquely associate a *coherent* γ -sheaf \mathcal{M} with F , and one may, as in the work of Miller, use invariants of \mathcal{M} , such as its degree, as an invariant for F .

The ideas in this paper have two sources. Firstly, the ongoing project of Blickle and Böckle [≥ 2008] lead to a systematic study of γ -sheaves (the notation γ -sheaf is chosen to remind of the notion of a **generator** introduced in [Lyubeznik 1997]). Secondly, insight gained from the D -module theoretic viewpoint on generalized test ideals developed in [Blickle et al. 2008] jointly with Mircea Mustața and Karen Smith leads to the observation that these techniques can be successfully applied to study γ -sheaves.

Notation. Throughout we fix a *regular* scheme X over a field $k \supseteq \mathbb{F}_q$ of characteristic $p > 0$ (with $q = p^e$ fixed). We further assume that X is F -finite, that is, the Frobenius morphism $\sigma : X \rightarrow X$, which is given by sending $f \in \mathcal{O}_X$ to f^q , is a finite morphism.² In general, σ is affine. This allows to reduce in many arguments below to the case that X itself is affine and I will do so if convenient. We will use without further mention that because X is regular, the Frobenius morphism $\sigma : X \rightarrow X$ is flat such that σ^* is an exact functor (see [Kunz 1969]).

2. Minimal γ -sheaves

We begin with recalling the notion of γ -sheaves and nilpotence.

Definition 2.1. A γ -sheaf on X is a pair (M, γ_M) consisting of a quasicohherent \mathcal{O}_X -module M and a \mathcal{O}_X -linear map $\gamma : M \rightarrow \sigma^*M$. A γ -sheaf is called *coherent* if its underlying sheaf of \mathcal{O}_X -modules is coherent.

²It should be possible to replace the assumption of F -finiteness to saying that if X is a k -scheme with k a field such that the relative Frobenius $\sigma_{X/k}$ is finite. This would extend the results given here to desirable situations such as X of finite type over a field k with $[k : k^q] = \infty$. The interested reader should have no trouble to adjust our treatment to this case.

A γ -sheaf (M, γ) is called *nilpotent* (of order n) if

$$\gamma^n \stackrel{\text{def}}{=} \sigma^{n*} \gamma \circ \sigma^{(n-1)*} \gamma \circ \dots \circ \sigma^* \gamma \circ \gamma = 0$$

for some $n > 0$. A γ -sheaf is called *locally nilpotent* if it is the union of nilpotent γ subsheaves.

Maps of γ -sheaves are maps of the underlying \mathbb{C}_X -modules such that the obvious diagram commutes. We denote the category of coherent γ -sheaves on X by $\mathbf{Coh}_\gamma(X)$. The following proposition summarizes some properties of γ -sheaves; for proofs and more details see [Blickle and Bökke \geq 2008].

Proposition 2.2. (a) *The set of γ -sheaves forms an abelian category.*

(b) *The coherent, nilpotent and locally nilpotent γ -sheaves are abelian subcategories, also closed under extension.*

Proof. The point in the first statement is that the \mathbb{C}_X -module kernel, cokernel and extension of (maps of) γ -sheaves naturally carry the structure of a γ -sheaf. This is really easy to verify so we only give the construction of the γ -structure on the kernel as an illustration. Recall that we assume that X is regular such that σ is flat, hence σ^* is an exact functor. A morphism $\varphi : M \rightarrow N$ of γ -sheaves is a commutative diagram

$$\begin{array}{ccc} M & \xrightarrow{\varphi} & N \\ \gamma_M \downarrow & & \downarrow \gamma_N \\ \sigma^* M & \xrightarrow{\sigma^* \varphi} & \sigma^* N \end{array}$$

from which we obtain the induced map $\ker \varphi \rightarrow \ker(\sigma^* \varphi)$. Since σ^* is exact, the natural map $\sigma^*(\ker \varphi) \rightarrow \ker(\sigma^* \varphi)$ is an isomorphism. Hence the composition

$$\ker \varphi \rightarrow \ker(\sigma^* \varphi) \xrightarrow{\cong} \sigma^*(\ker \varphi)$$

equips $\ker \varphi$ with a natural structure of a γ -sheaf.

The second part of Proposition 2.2 is also easy to verify so we leave it to the reader (see the proof of Lemma 2.3 below). □

Lemma 2.3. *A morphism $\varphi : M \rightarrow N$ of γ -sheaves is called nil-injective (respectively, nil-surjective, nil-isomorphism) if its kernel (respectively, cokernel, both) is locally nilpotent.*

(a) *If M (respectively, N) is coherent and φ is nil-injective (respectively, nil-surjective) then $\ker \varphi$ (respectively, $\text{coker } \varphi$) is nilpotent.*

(b) *Kernel and cokernel of φ are nilpotent (of order n and m respectively) if and only if there is, for some $k \geq 0$ ($k = n + m$), a map $\psi : N \rightarrow \sigma^{k*} M$ such that $\gamma_M^k = \psi \circ \varphi$ and $\gamma_N^k = \sigma^{k*}(\varphi) \circ \psi$.*

(c) If N is nilpotent of order $\leq n$ (that is, $\gamma_N^n = 0$) and $N' \subseteq N$ contains the kernel of γ_N^i for $1 \leq i \leq n$, then N/N' is nilpotent of order $\leq n - i$.

Proof. The first statement is clear since X is noetherian. For the second statement consider the diagram obtained from the exact sequence $0 \rightarrow K \rightarrow M \rightarrow N \rightarrow C \rightarrow 0$:

$$\begin{array}{ccccccccc}
 0 & \longrightarrow & K & \longrightarrow & M & \longrightarrow & N & \longrightarrow & C & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow & & \swarrow \psi & & \downarrow & & \downarrow 0 \\
 0 & \longrightarrow & \sigma^{n*} K & \longrightarrow & \sigma^{n*} M & \longrightarrow & \sigma^{n*} N & \longrightarrow & \sigma^{n*} C & \longrightarrow & 0 \\
 & & \downarrow 0 & & \downarrow & & \swarrow \psi & & \downarrow & & \downarrow \\
 0 & \longrightarrow & \sigma^{(n+m)*} K & \longrightarrow & \sigma^{(n+m)*} M & \longrightarrow & \sigma^{(n+m)*} N & \longrightarrow & \sigma^{(n+m)*} C & \longrightarrow & 0.
 \end{array}$$

If there is a ψ as indicated, then clearly the leftmost and rightmost vertical arrows of the first row are zero, that is, K and C are nilpotent. Conversely, let $K = \ker \varphi$ be nilpotent of degree n and $C = \text{coker } \varphi$ be nilpotent of degree m . Then the top right vertical arrow and the bottom left vertical arrow are zero. The fact that the top right arrow is zero allows to define ψ as follows: Take $n \in N$, map it down to $\sigma^{n*} N$. Since its image to the right is zero, take any preimage from the left and map that element down in the diagram to $\sigma^{(n+m)*} M$. This procedure defines $\psi(n)$. To show that it is well defined and to see that the two relevant triangles commute is not difficult by using that the bottom left vertical arrow is zero.

For the last part, consider the short exact sequence $0 \rightarrow N' \rightarrow N \rightarrow N/N' \rightarrow 0$ and the diagram one obtains by considering $\sigma^{(n-i)*}$ and σ^{n*} of this sequence:

$$\begin{array}{ccccccccc}
 0 & \longrightarrow & N' & \longrightarrow & N & \longrightarrow & N/N' & \longrightarrow & 0 \\
 & & \downarrow & & \downarrow \gamma^{n-i} & & \downarrow & & \\
 0 & \longrightarrow & \sigma^{(n-i)*} N' & \longrightarrow & \sigma^{(n-i)*} N & \longrightarrow & \sigma^{(n-i)*} (N/N') & \longrightarrow & 0 \\
 & & \downarrow 0 & & \downarrow \sigma^{(n-i)*} \gamma^i & & \downarrow & & \\
 0 & \longrightarrow & \sigma^{n*} N' & \longrightarrow & \sigma^{n*} N & \longrightarrow & \sigma^{n*} (N/N') & \longrightarrow & 0.
 \end{array}$$

The composition of the middle vertical map is γ_N^n which is zero by assumption. To conclude that the top right vertical arrow is zero one uses the fact that

$$\sigma^{(n-i)*} N' \supseteq \sigma^{(n-i)*} \ker \gamma^i = \ker(\sigma^{(n-i)*} \gamma^i).$$

With this it is an easy diagram chase to conclude that the top right vertical map is zero. □

Lemma 2.4. *Let $M \xrightarrow{\varphi} N$ be a map of γ -sheaves. Let $N' \subseteq N$ be such that N/N' is nilpotent (hence $N' \subseteq N$ is a nil-isomorphism). Then $M/(\varphi^{-1}N')$ is also nilpotent.*

Proof. $M/(\varphi^{-1}N')$ injects into N/N' . Since the latter is nilpotent, so is the former. □

If (M, γ) is a γ -sheaf, then σ^*M is naturally a γ -sheaf with structural map $\sigma^*\gamma$. Furthermore, the map $\gamma : M \rightarrow \sigma^*M$ is then a map of γ -sheaves which is a nil-isomorphism, that is, kernel and cokernel are nilpotent. We can iterate this process to obtain a directed system

$$M \xrightarrow{\gamma} \sigma^*M \xrightarrow{\sigma^*\gamma} \sigma^{2*}M \xrightarrow{\sigma^{2*}\gamma} \dots$$

whose limit we denote by $\text{Gen}M$. Clearly $\text{Gen}M$ is a γ -sheaf whose structural map $\gamma_{\text{Gen}M}$ is injective. In fact, it is an isomorphism since clearly $\sigma^*\text{Gen}M \cong \text{Gen}M$. One observes that $\text{Gen}M = 0$ if and only if M is locally nilpotent. Note that even if M is coherent, $\text{Gen}M$ is generally not coherent. Furthermore, let \bar{M} be the image of M under the natural map $M \rightarrow \text{Gen}M$. Then, if M is coherent, so is \bar{M} and the map $M \twoheadrightarrow \bar{M}$ is a nil-isomorphism. Since \bar{M} is a γ -submodule of $\text{Gen}M$ whose structural map is injective, the structural map $\bar{\gamma}$ of \bar{M} is injective as well.

Proposition 2.5. *The operation that assigns to each γ -sheaf M its image \bar{M} in $\text{Gen}M$ is an end-exact functor (preserves exactness only at the end of sequences) from $\mathbf{Coh}_\gamma(X)$ to $\mathbf{Coh}_\gamma(X)$. The kernel*

$$M^\circ = \bigcup \ker \gamma_M^i$$

of the natural map $M \rightarrow \bar{M}$ is the maximal locally nilpotent subsheaf of M .

Proof. The verification of the statement about M° is left to the reader. One has a functorial map between the exact functors $\text{id} \rightarrow \text{Gen}$. An easy diagram chase shows that the image of such a functorial map is an end-exact functor (see for example [Katz 1996, 2.17 Appendix 1]). In the concrete situation we are in, one can also verify this directly: right exactness is clear since \bar{M} is a quotient of M . On the other hand, if $M \subseteq N$ is a γ -subsheaf, then $N^\circ \cap M \subseteq M^\circ$ since the former is clearly locally nilpotent. It follows that $\bar{M} \subseteq \bar{N}$. □

Such γ -submodules with injective structural map enjoy a certain minimality property with respect to nilpotent subsheaves:

Lemma 2.6. *Let (M, γ) be a γ -sheaf. The structural map γ_M is injective if and only if M does not have a nontrivial nilpotent subsheaf.*

Proof. Assume that the structural map of M is injective. This implies that the structural map of any γ -subsheaf of M is injective. But a γ -sheaf with injective structural map is nilpotent if and only if it is zero.

Conversely, $\ker \gamma_M$ is a nil-potent subsheaf of M . If γ_M is not injective it is nontrivial. \square

2.1. Definition of minimal γ -sheaves.

Definition 2.7. A coherent γ -sheaf M is called *minimal* if the following two conditions hold:

- (a) M does not have nontrivial nilpotent subsheaves;
- (b) M does not have nontrivial nilpotent quotients.

We denote by $\mathbf{Min}_\gamma(X)$ the subcategory of all γ -sheaves consisting of the minimal ones.

A simple consequence of the definition is

Lemma 2.8. *Let M be a γ -sheaf. If M satisfies (a) then any γ -subsheaf of M also satisfies (a). If M satisfies (b), so does any quotient.*

Proof. Immediate from the definition. \square

As the preceding Lemma 2.6 shows, (a) is equivalent to the condition that the structural map γ_M is injective. We give a concrete description of the second condition.

Proposition 2.9. *For a coherent γ -sheaf M , the following conditions are equivalent.*

- (a) M does not have nontrivial nilpotent quotients.
- (b) *For any map of γ -sheaves $\varphi : N \rightarrow M$, if $\gamma_M(M) \subseteq (\sigma^*\varphi)(\sigma^*N)$ (as subsets of σ^*M) then φ is surjective.*

Proof. I begin with showing the easy direction that (a) implies (b): Note that the condition $\gamma_M(M) \subseteq (\sigma^*\varphi)(\sigma^*N)$ in (b) precisely says that the induced structural map on the cokernel of $N \rightarrow M$ is the zero map, thus in particular $M/\varphi(N)$ is a nilpotent quotient of M . By assumption on M , $M/\varphi(N) = 0$ and hence $\varphi(N) = M$.

Let $\pi : M \twoheadrightarrow C$ be such that C is nilpotent. Let $N \subseteq M$ be its kernel. We have to show that $N = M$. The proof is by induction on the order of nilpotency of C (simultaneously for all C). If $C = M/N$ is nilpotent of order 1 this means precisely that $\gamma(M) \subseteq \sigma^*N$; hence by (b) we have $N = M$ as claimed. Now let N be such that the nilpotency order of $C \stackrel{\text{def}}{=} M/N$ is equal to $n \geq 2$. Consider the γ -submodule $N' = \pi^{-1}(\ker \gamma_C)$ of M . This N' clearly contains N and we have that $M/N' \cong C/(\ker \gamma_C)$. By the previous Lemma 2.3 we conclude that the nilpotency order of M/N' is $\leq n - 1$. Thus by induction $N' = M$. Hence

$$M/N = N'/N \cong \ker \gamma_C$$

is of nilpotency order 1. Again by the base case of the induction we conclude that $M = N$. \square

By replacing N by its image $\varphi(N)$ in M in item (b) of Proposition 2.9 it follows that it would be enough to consider such φ which are injective.

Corollary 2.10. *A coherent γ -sheaf M is minimal if and only if the following two conditions hold.*

- (a) *The structural map of M is injective.*
- (b) *If $N \subseteq M$ is a subsheaf such that $\gamma(M) \subseteq \sigma^*N$ then $N = M$.*

The conditions in this corollary are essentially the definition of a *minimal root* of a finitely generated unit $R[\sigma]$ -module in [Lyubeznik 1997]. The finitely generated unit $R[\sigma]$ -module generated by (M, γ) is of course $\text{Gen}M$. Lyubeznik shows in the case that R is a complete regular ring, that minimal roots exist. In [Blickle 2004, Theorem 2.10] I showed how to reduce the local case to the complete case if R is F -finite. For convenience we give a streamlined argument of the result in the local case in the language of γ -sheaves.

2.2. Minimal γ -sheaves over local rings. The difficult part in establishing the existence of a minimal root is to satisfy condition (b) of Definition 2.7. The point is to universally bound the order of nilpotency of any nilpotent quotient of a fixed γ -sheaf M .

Proposition 2.11. *Let (R, \mathfrak{m}) be regular, local and F -finite. Let M be a coherent γ -sheaf and $\{N_i\}_{i \in I}$ be a collection of γ -subsheaves which is closed under finite intersections and such that M/N_i is nilpotent for all i . Then $M/\bigcap N_i$ is nilpotent.*

Proof. Note that if N and N' are γ -subsheaves of M such that M/N and M/N' are nilpotent of order n and n' , then $M/(N \cap N')$ is nilpotent of order $\max\{n, n'\}$. Hence, with $\{N_i\}$ the family of all $N \subseteq M$ such that M/N is nilpotent, Proposition 2.11 may be equivalently stated:

The order of nilpotency of any nilpotent quotient of M is universally bounded.

By faithful flatness of completion (together with the fact that completion commutes with Frobenius), *order of nilpotency* of quotients of M is preserved by completion. Therefore we may reduce to the case that R is complete.

Let us hence assume that R is complete, local, regular and F -finite. Since R is via σ a free R -module of finite rank, σ^* is nothing but tensorisation with a free R -module of finite rank. Such an operation commutes with the formation of inverse limits such that $\sigma^*\bigcap N_i = \bigcap(\sigma^*N_i)$ and hence $\bigcap N_i$ is a γ -subsheaf of M . Clearly we may replace M by $M/\bigcap N_i$ such that we have $\bigcap N_i = 0$. We may further replace M by its image \overline{M} in $\text{Gen}M$. Thus we may assume that M has injective structural map $\gamma : M \subseteq \sigma^*M$. We have to show that $M = 0$.

By the Artin–Rees Lemma (applied to $M \subseteq \sigma^* M$) there exists $t \geq 0$ such that for all $s > t$,

$$M \cap \mathfrak{m}^s \sigma^* M \subseteq \mathfrak{m}^{s-t} (M \cap \mathfrak{m}^t \sigma^* M) \subseteq \mathfrak{m}^{s-t} M .$$

By Chevalley’s Theorem in the version of [Lyubeznik 1997, Lemma 3.3], for some $s \gg 0$ (in fact $s \geq t + 1$ will suffice) we find N_i with $N_i \subseteq \mathfrak{m}^s M$. Possibly increasing s we may assume that $N_i \not\subseteq \mathfrak{m}^{s+1} M$ (unless, of course $N_i = 0$ in which case $M/N_i = M$ is nilpotent $\Rightarrow M = 0$ since γ_M is injective, and we are done). Combining these inclusions we get

$$\begin{aligned} N_i &\subseteq \sigma^* N_i \cap M \subseteq \sigma^*(\mathfrak{m}^s M) \cap M \\ &\subseteq (\mathfrak{m}^s)^{[q]} \sigma^* M \cap M \subseteq \mathfrak{m}^{sq} \sigma^* M \cap M \\ &\subseteq \mathfrak{m}^{sq-t} M . \end{aligned}$$

But since $sq - t \geq s + 1$ for our choice of $s \geq t + 1$ this is a contradiction (to the assumption $N_i \neq 0$) and the result follows. \square

Remark 2.12. An alternative way to prove this result is to use Matlis duality and then invoke a result of Hartshorne and Speiser [1977, Proposition 1.11]. Their result states that if U is a cofinite $R[\sigma]$ -module then the subset

$$U_{\text{nil}} = \{ u \in U \mid \sigma^n(u) = 0 \text{ for some } n \}$$

is annihilated by a fixed power of σ , that is, there is $k \geq 1$ such that $\sigma^k(U_{\text{nil}}) = 0$. If one applies this to the Matlis dual $U = M^\vee$ of M and the union of its σ -nilpotent submodules $(M/N_i)^\vee$ in the above statement, an alternative proof is obtained. This approach via the Hartshorne–Speiser result is used in [Katzman et al. 2007] to study F -thresholds and hence appears to be directly related to the observations we make in Section 3.3 below.

Corollary 2.13. *Let R be regular, local and F -finite and M a coherent γ -sheaf. Then M has a nil-isomorphic subsheaf without nonzero nilpotent quotients (that is, satisfying (b) of the definition of minimality). In particular, M is nil-isomorphic to a minimal γ -sheaf.*

Proof. Let N_i be the collection of all nil-isomorphic subsheaves of M . Since M is coherent each M/N_i is indeed nilpotent, say of order $\leq n_i$. Since

$$M/(N_i \cap N_j) \subseteq M/N_i \oplus M/N_j ,$$

it follows that $M/(N_i \cap N_j)$ is nilpotent of order $\leq \max\{n_i, n_j\}$. Hence the collection of nil-isomorphic subsheaves of M is closed under intersection which allows to apply Proposition 2.11 to conclude that $M/\bigcap N_i$ is nilpotent. Hence $N \stackrel{\text{def}}{=} \bigcap N_i$ is the unique smallest nil-isomorphic subsheaf of M . It is clear that N cannot have

nonzero nilpotent quotients (since the kernel would be a strict subsheaf of N , nil-isomorphic to M , by Proposition 2.2 (b)).

By first replacing M by \overline{M} we can also achieve that condition (a) of the definition of minimality holds. As condition (a) passes to subsheaves, the smallest nil-isomorphic subsheaf of \overline{M} is the sought after minimal γ -sheaf which is nil-isomorphic to M . \square

Remark 2.14. Essentially the same argument as in the proof of Proposition 2.11 shows the following: If R is local and M is a coherent γ -sheaf over R with injective structural map, then any descending chain of γ -submodules of M stabilizes. This was shown (with essentially the same argument) in [Lyubeznik 1997] and implies immediately that γ -sheaves with injective structural map satisfy DCC.

If one tries to reduce the general case of Corollary 2.13 (that is, R not local) to the local case just proven, one encounters the problem of having to deal with the behavior of the infinite intersection $\bigcap N_i$ under localization. This is a source of troubles I do not know how to deal with directly. The solution to this is to take a detour and realize this intersection in a fashion such that each term functorially depends on M and furthermore that this functorial construction commutes with localization. This is explained in the following section.

2.3. $D_X^{(1)}$ -modules and Frobenius descent. Let D_X denote the sheaf of differential operators on X . This is a sheaf of rings on X which locally, on each affine subvariety $\text{Spec } R$, is described as

$$D_R = \bigcup_{i=0}^{\infty} D_R^{(i)}$$

where $D_R^{(i)}$ is the subset of $\text{End}_{\mathbb{F}_q}(R)$ consisting of the operators which are linear over R^{q^i} , the subring of q^i -th powers of elements of R . In particular

$$D_R^{(0)} \cong R, \quad \text{and} \quad D_R^{(1)} = \text{End}_{R^q}(R).$$

Clearly, R itself becomes naturally a left $D_R^{(i)}$ -module. Now denote by $R^{(1)}$ the $D_R^{(1)}$ - R -bimodule R which has this left $D_R^{(1)}$ -module structure and the right R -module structure via Frobenius, that is, for $r \in R^{(1)}$ and $x \in R$ we have $r \cdot x = x^q r$. With this notation we may view $D_R^{(1)} = \text{End}_R^r(R^{(1)})$ as the right R -linear endomorphisms of $R^{(1)}$. Thus we have

$$\sigma^*(_) = R^{(1)} \otimes_R _ : R\text{-mod} \rightarrow D_R^{(1)}\text{-mod},$$

which makes σ^* into an equivalence of categories from R -modules to $D_R^{(1)}$ -modules (because, since σ is flat and R is F -finite, $R^{(1)}$ is a locally free right R -module of

finite rank). Its inverse functor is given by

$$\sigma^{-1}(_) = \text{Hom}_R^r(R^{(1)}, R) \otimes_{D_R^{(1)}} _ : D_R^{(1)}\text{-mod} \rightarrow R\text{-mod}. \tag{2-1}$$

For details see [Alvarez-Montaner et al. 2005, Section 2.2]. I want to point out that these constructions commute with localization at arbitrary multiplicative sets. Let S be a multiplicative set of R .³ We have

$$\begin{aligned} S^{-1}D_R^{(1)} &= S^{-1}\text{End}_R^r(R^{(1)}) \\ &= \text{End}_{S^{-1}R}^r((S^{[q]})^{-1}R^{(1)}) = \text{End}_{S^{-1}R}^r((S^{-1}R)^{(1)}) \\ &= D_{S^{-1}R}^{(1)}. \end{aligned}$$

Furthermore we have for an $D_R^{(1)}$ -module M :

$$\begin{aligned} S^{-1}(\sigma^{-1}M) &= S^{-1}(\text{Hom}_R^r(R^{(1)}, R) \otimes_{D_R^{(1)}} M) \\ &= S^{-1}\text{Hom}_R^r(R^{(1)}, R) \otimes_{S^{-1}D_R^{(1)}} S^{-1}M \\ &= \text{Hom}_{S^{-1}R}^r((S^{-1}R)^{(1)}, S^{-1}R) \otimes_{D_{S^{-1}R}^{(1)}} S^{-1}M \\ &= \sigma^{-1}(S^{-1}M). \end{aligned}$$

These observations are summarized in the following proposition.

Proposition 2.15. *Let X be F -finite and regular. Let U be an open subset (more generally, U is locally given on $\text{Spec } R$ as $\text{Spec } S^{-1}R$ for some (sheaf of) multiplicative sets on X). Then*

$$(D_X^{(1)})|_U = D_U^{(1)}$$

and for any sheaf of $D_X^{(1)}$ -modules M one has that

$$\begin{aligned} (\sigma^{-1}M)|_U &= (\text{Hom}^r(\mathbb{O}_X^{(1)}, \mathbb{O}_X) \otimes_{D_X^{(1)}} M)|_U \\ &\cong \text{Hom}^r(\mathbb{O}_U^{(1)}, \mathbb{O}_U) \otimes_{D_U^{(1)}} M|_U = \sigma^{-1}(M|_U) \end{aligned}$$

as \mathbb{O}_U -modules.

2.4. A criterion for minimality. The Frobenius descent functor σ^{-1} can be used to define an operation on γ -sheaves which assigns to a γ -sheaf M its smallest γ -subsheaf N with the property that M/N has the trivial ($= 0$) γ -structure. This is the opposite of what the functor σ^* does: $\gamma : M \rightarrow \sigma^*M$ is a map of γ sheaves such that $\sigma^*M/\gamma(M)$ has trivial γ -structure.

We define the functor σ_γ^{-1} from γ -sheaves to γ -sheaves as follows. Let $M \xrightarrow{\gamma} \sigma^*M$ be a γ -sheaf. Then $\gamma(M)$ is an \mathbb{O}_X -submodule of the $D_X^{(1)}$ -module σ^*M .

³Since $S^{-1}R = (S^{[q]})^{-1}R$ we may assume that $S \subseteq R^q$. This implies that S is in the center of $D_R^{(1)}$ such that localization in this noncommutative ring along S is harmless. With this I mean that we may view the localization of the left R -module $D_R^{(1)}$ at S^{-1} in fact as the localization of $D_R^{(1)}$ at the central multiplicative set $(S^{[q]})^{-1}$

Denote by $D_X^{(1)}\gamma(M)$ the $D_X^{(1)}$ -submodule of σ^*M generated by $\gamma(M)$. To this inclusion of $D_X^{(1)}$ -modules

$$D_X^{(1)}\gamma(M) \subseteq \sigma^*M,$$

we apply the Frobenius descent functor $\sigma^{-1} : D_X^{(1)}\text{-mod} \rightarrow \mathbb{C}_X\text{-mod}$ defined above in (2-1) and use that $\sigma^{-1} \circ \sigma^* = \text{id}$ to define

$$\sigma_\gamma^{-1}M \stackrel{\text{def}}{=} \sigma^{-1}(D_X^{(1)}\gamma(M)) \subseteq \sigma^{-1}\sigma^*M = M.$$

In general one has

$$\sigma_\gamma^{-1}(\sigma^*M) = \sigma^{-1}D_X^{(1)}\sigma^*(\gamma)(\sigma^*M) = \gamma(M)$$

since $\sigma^*(\gamma)(\sigma^*M)$ already is a $D_X^{(1)}$ -subsheaf of the $D_X^{(2)}$ -module $\sigma^*(\sigma^*M) = \sigma^{2*}M$.

By construction,

$$\sigma_\gamma^{-1}M \subseteq M \xrightarrow{\gamma} \gamma(M) \subseteq D_X^{(1)}\gamma(M) = \sigma^*\sigma^{-1}D_X^{(1)}\gamma(M) = \sigma^*\sigma_\gamma^{-1}M$$

such that $\sigma_\gamma^{-1}M$ is a γ -subsheaf of M .

Furthermore, the quotient $M/\sigma_\gamma^{-1}M$ has zero structural map. One makes the following observation.

Lemma 2.16. *Let M be a γ -sheaf. Then $\sigma_\gamma^{-1}M$ is the smallest subsheaf N of M such that $\sigma^*N \supseteq \gamma(M)$.*

Proof. Clearly $\sigma^{-1}M$ satisfies this condition. Let N be as in the statement of the Lemma. Then σ^*N is a $D_X^{(1)}$ -subsheaf of σ^*M containing $\gamma(M)$. Hence $D_X^{(1)}\gamma(M) \subseteq \sigma^*N$. Applying σ^{-1} we see that $\sigma^{-1}M \subseteq N$. \square

Therefore, the result of the lemma could serve as an alternative definition of σ_γ^{-1} (one would have to show that the intersection of all such N has again the property that $\gamma(M) \subseteq \sigma^* \bigcap N$ but this follows since σ^* commutes with inverse limits since it is locally just tensorisation with a free module of finite rank). The following lemma is the key point in our reduction to the local case. It is an immediate consequence of Proposition 2.15. Nevertheless we include here a proof using only the characterization of Lemma 2.16. Hence one may avoid the appearance of $D^{(1)}$ -modules in this paper altogether but I believe it to be important to explain where the ideas for the arguments originated, so $D^{(1)}$ -modules are still there.

Lemma 2.17. *Let M be a γ -sheaf and let $S \subseteq \mathbb{C}_X$ be multiplicative set. Then $S^{-1}(\sigma_\gamma^{-1}M) = \sigma_\gamma^{-1}(S^{-1}M)$.*

Proof. This follows from Proposition 2.15. However, this can also be proven using only the characterization in Lemma 2.16. By this we have

$$\sigma^*(S^{-1}(\sigma_\gamma^{-1}M)) = S^{-1}(\sigma^*(\sigma_\gamma^{-1}M)) \supseteq S^{-1}\gamma(M) = \gamma(S^{-1}M), \tag{2-2}$$

which implies that $\sigma_\gamma^{-1}(S^{-1}M) \subseteq S^{-1}(\sigma_\gamma^{-1}M)$ because $\sigma_\gamma^{-1}(S^{-1}M)$ is smallest (by Lemma 2.16) with respect to the inclusion shown in (2-2). To show the converse inclusion, we consider the localization map $\varphi : M \rightarrow S^{-1}M$ and $N \subseteq S^{-1}M$ is a submodule such that $\gamma(S^{-1}M) \subseteq \sigma^*N$. Consider the diagram

$$\begin{array}{ccc} M/\varphi^{-1}N & \hookrightarrow & S^{-1}M/N \\ \downarrow \gamma & & \downarrow \gamma \\ \sigma^*M/\sigma^*(\varphi^{-1}N) & \hookrightarrow & \sigma^*(S^{-1}M)/\sigma^*N \end{array}$$

of which the horizontal arrows are injections (using the exactness of σ^*). By assumption on N , the right vertical arrow is zero, hence also the left vertical arrow. This implies that $\gamma(M) \subseteq \sigma^*(\varphi^{-1}N)$. By the characterization of Lemma 2.16 one concludes that $\sigma_\gamma^{-1}M \subseteq \varphi^{-1}N$ and hence $S^{-1}\sigma_\gamma^{-1}M \subseteq N$. Applying this with $N = \sigma_\gamma^{-1}(S^{-1}M)$ our claim follows. \square

Proposition 2.18. *Let M be a γ -sheaf. Then $\sigma_\gamma^{-1}M = M$ if and only if M has no proper nilpotent quotient (that is, M satisfies condition (b) of the definition of minimality).*

If M is coherent, the condition on $x \in X$ that the inclusion $\sigma_\gamma^{-1}(M_x) \subseteq M_x$ is equality is an open condition on X .

Proof. One direction is clear since $M/\sigma_\gamma^{-1}M$ is a nilpotent quotient of M . We use the characterization in Proposition 2.9. For this let $N \subseteq M$ be such that $\gamma(M) \subseteq \sigma^*N$. As $\sigma_\gamma^{-1}M$ was the smallest subsheaf with this property we obtain $\sigma_\gamma^{-1}M \subseteq N \subseteq M$. Since $M = \sigma_\gamma^{-1}M$ by assumption it follows that $N = M$. Hence, by Proposition 2.9, M does not have nontrivial nilpotent quotients.

By Lemma 2.17, σ_γ^{-1} commutes with localization which means that $\sigma_\gamma^{-1}(M_x) = (\sigma_\gamma^{-1}M)_x$. Hence the second statement follows simply since both M and $\sigma_\gamma^{-1}M$ are coherent (and equality of two coherent modules via a given map is an open condition). \square

Lemma 2.19. *The assignment $M \mapsto \sigma_\gamma^{-1}M$ is an end-exact functor on γ -sheaves.*

Proof. Formation of the image of the functorial map $\text{id} \xrightarrow{\gamma} \sigma^*$ of exact functors is end-exact (see for example [Katz 1996, 2.17 Appendix 1]). If M is a $D_X^{(1)}$ -module and $A \subseteq B$ are \mathbb{O}_X -submodules of M then

$$D_X^{(1)}A \subseteq D_X^{(1)}B.$$

If $M \twoheadrightarrow N$ is a surjection of $D^{(1)}$ -modules which induces a surjection on \mathbb{O}_X -submodules $A \twoheadrightarrow B$ then, clearly, $D_X^{(1)}A$ surjects onto $D_X^{(1)}B$. Now one concludes by observing that σ^{-1} is an exact functor. \square

Lemma 2.20. *Let $N \subseteq M$ be an inclusion of γ -sheaves such that $\sigma^{n*}N \supseteq \gamma^n(M)$ (that is, the quotient is nilpotent of order $\leq n$). Then*

$$\sigma^{(n-1)*}(N \cap \sigma_\gamma^{-1}M) \supseteq \gamma^{n-1}(\sigma_\gamma^{-1}M).$$

Proof. Consider the γ -subsheaf $M' = (\gamma^{n-1})^{-1}(\sigma^{(n-1)*}N)$ of M . One has

$$\sigma^*M' = (\sigma^* \gamma^{n-1})^{-1}(\sigma^{n*}N) \supseteq \gamma(M)$$

by the assumption that $\gamma^n(M) \subseteq \sigma^{n*}N$. Since $\sigma_\gamma^{-1}M$ is minimal with respect to this property we have $\sigma_\gamma^{-1}M \subseteq (\gamma^{n-1})^{-1}(\sigma^{(n-1)*}N)$. Applying γ^{n-1} we conclude that $\gamma^{n-1}(\sigma_\gamma^{-1}M) \subseteq \sigma^{(n-1)*}N$. Since $\sigma_\gamma^{-1}M$ is a γ -sheaf we have

$$\gamma^{n-1}(\sigma_\gamma^{-1}M) \subseteq \sigma^{(n-1)*}(\sigma_\gamma^{-1}M)$$

such that the claim follows. □

2.5. Existence of minimal γ -sheaves. For a given γ -sheaf M we can iterate the functor σ_γ^{-1} to obtain a decreasing sequence of γ -subsheaves,

$$\dots \subseteq M_3 \subseteq M_2 \subseteq M_1 \subseteq M \xrightarrow{\gamma} \sigma^*M \rightarrow \dots$$

where $M_i = \sigma_\gamma^{-1}M_{i-1}$. Note that each inclusion $M_i \subseteq M_{i-1}$ is a nil-isomorphism.

Proposition 2.21. *Let M be a coherent γ -sheaf. Then the following conditions are equivalent.*

- (a) *M has a nil-isomorphic γ -subsheaf \underline{M} which does not have nontrivial nilpotent quotients (that is, \underline{M} satisfies condition (b) in the definition of minimal γ -sheaf).*
- (b) *M has a unique smallest nil-isomorphic subsheaf (equivalently, M has a (unique) maximal nilpotent quotient).*
- (c) *For some $n \geq 0$, $M_n = M_{n+1}$.*
- (d) *There is $n \geq 0$ such that for all $m \geq n$, $M_m = M_{m+1}$.*

Proof. (a) \Rightarrow (b): Let $\underline{M} \subseteq M$ be the nil-isomorphic subsheaf of part (a) and let $N \subseteq M$ be another nil-isomorphic subsheaf of M . By Lemma 2.4 it follows that $\underline{M} \cap N$ is also nil-isomorphic to M . In particular $\underline{M}/(\underline{M} \cap N)$ is a nilpotent quotient of \underline{M} and hence must be trivial. Thus $N \subseteq \underline{M}$ which shows that \underline{M} is the smallest nil-isomorphic subsheaf of M .

(b) \Rightarrow (c): Let N be this smallest subsheaf as in (b). Since each M_i is nil-isomorphic to M , it follows that $N \subseteq M_i$ for all i . Let n be the order of nilpotency of the quotient M/N , that is, $\gamma^n(M) \subseteq \sigma^{n*}N$. Repeated application (n times) of Lemma 2.20 yields that $M_n \subseteq N$. Hence we get $N \subseteq M_{n+1} \subseteq M_n \subseteq N$ which implies that $M_{n+1} = M_n$.

(c) \Rightarrow (d) is clear.

(d) \Rightarrow (a) is clear by Proposition 2.18. □

This characterization enables us to show the existence of minimal γ -sheaves by reducing to the local case which we proved above.

Theorem 2.22. *Let M be a coherent γ -sheaf. There is a unique nil-isomorphic subsheaf \underline{M} of M which does not have nontrivial nilpotent quotients.*

Remark 2.23. The following proof shows that in the notation of Proposition 2.21, $\underline{M} = M_n$ for $n \gg 0$.

Proof. By Proposition 2.21 it is enough to show that the sequence M_i is eventually constant. Let U_i be the subset of X consisting of all $x \in X$ on which

$$(M_i)_x = (M_{i+1})_x (= (\sigma_\gamma^{-1} M_i)_x).$$

By Proposition 2.18 U_i is an open subset of X (in this step I use the key observation Proposition 2.15) and $(M_i)|_{U_i} = (M_{i+1})|_{U_i}$. By the functorial construction of the M_i 's the equality $M_i = M_{i+1}$ for one i implies equalities for all bigger i . It follows that the sets U_i form an increasing sequence of open subsets of X whose union is X itself by Corollary 2.13 and Proposition 2.21. Since X is noetherian, $X = U_i$ for some i . Hence $M_i = M_{i+1}$ so the claim follows by Proposition 2.21. □

Theorem 2.24. *Let M be a coherent γ -sheaf. Then there is a functorial way to assign to M a minimal γ -sheaf M_{\min} in the nil-isomorphism class of M .*

Proof. We may first replace M by the nil-isomorphic quotient \overline{M} which satisfies condition (a) of Definition 2.7. Then replace \overline{M} by its minimal nil-isomorphic submodule $\underline{\overline{M}}$ which also satisfies condition (b) of Definition 2.7 (and condition (a) because (a) is passed to submodules). Thus the assignment

$$M \mapsto M_{\min} \stackrel{\text{def}}{=} \underline{\overline{M}}$$

is a functor since it is a composition of the functors $M \mapsto \overline{M}$ and $M \mapsto \underline{M}$. □

Proposition 2.25. *If $\varphi: M \rightarrow N$ is a nil-isomorphism, then $\varphi_{\min}: M_{\min} \rightarrow N_{\min}$ is an isomorphism.*

Proof. Clearly, φ_{\min} is a nil-isomorphism. Since $\ker \varphi_{\min}$ is a nilpotent subsheaf of M_{\min} , we have by Definition 2.7 (a) that $\ker \varphi_{\min} = 0$. Since $\text{coker } \varphi_{\min}$ is a nilpotent quotient of N_{\min} it must be zero by Definition 2.7 (b). □

Corollary 2.26. *Let \mathcal{M} be a finitely generated unit $\mathbb{O}_X[\sigma]$ -module. Then M has a unique minimal root in the sense of [Lyubeznik 1997].*

Proof. Let M be any root of \mathcal{M} , that is, M is a coherent γ -sheaf such that γ_M is injective and $\text{Gen}M \cong \mathcal{M}$. Then $M_{\min} = \underline{M}$ is a minimal nil-isomorphic γ -subsheaf of M by Theorem 2.24. By Corollary 2.10 it follows that M_{\min} is the sought after minimal root of \mathcal{M} . Uniqueness is clear since the intersection of any two roots is again a root. \square

Note that the only assumption needed in this result is that X is F -finite and regular. In particular it does not rely on the finite-length result [Lyubeznik 1997, Theorem 3.2] which assumes that R is of finite type over a regular local ring (however, in [Lyubeznik 1997] F -finiteness is not assumed).

Theorem 2.27. *Let X be regular and F -finite. Then the functor*

$$\text{Gen} : \mathbf{Min}_\gamma(X) \rightarrow \text{finitely generated unit } \mathbb{C}_X[\sigma]\text{-modules}$$

is an equivalence of categories.

Proof. The preceding corollary shows that Gen is essentially surjective. The induced map on Hom sets is injective since a map of minimal γ -sheaves f is zero if and only if its image is nilpotent (since minimal γ -sheaves do not have nilpotent submodules) which is the condition that $\text{Gen}(f) = 0$. It is surjective since any map between $g : \text{Gen}(M) \rightarrow \text{Gen}(N)$ is obtained from a map of γ -sheaves $M \rightarrow \sigma^{e*}N$ for some $e \gg 0$. But this induces a map $M = M_{\min} \rightarrow (\sigma^{e*}N)_{\min} = N_{\min} = N$. \square

3. Applications and Examples

In this section we discuss some further examples and applications of the results on minimal γ -sheaves we obtained so far.

3.1. γ -crystals. The purpose of this section is to quickly explain the relationship of minimal γ -sheaves to γ -crystals which were introduced in [Blickle and Böckle ≥ 2008]. The category of γ -crystals is obtained by inverting nil-isomorphisms in $\mathbf{Coh}_\gamma(X)$. In [Blickle and Böckle ≥ 2008] it is shown that the resulting category is abelian. One has a natural functor

$$\mathbf{Coh}_\gamma(X) \twoheadrightarrow \mathbf{Crys}_\gamma(X)$$

whose fibers we may think of consisting of nil-isomorphism classes of M . Note that the objects of $\mathbf{Crys}_\gamma(X)$ are the same as those of $\mathbf{Coh}_\gamma(X)$; however a morphism between γ -crystals $M \rightarrow N$ is represented by a left-fraction, that is, a diagram of γ -sheaves $M \leftarrow M' \rightarrow N$ where the arrow \leftarrow is a nil-isomorphism.

On the other hand we just constructed the subcategory of minimal γ -sheaves $\mathbf{Min}_\gamma(X) \subseteq \mathbf{Coh}_\gamma(X)$ and showed that there is a functorial splitting $M \mapsto M_{\min}$ of this inclusion. An immediate consequence of Proposition 2.25 is that if M and N

are in the same nil-isomorphism class, then $M_{\min} \cong N_{\min}$. The verification of this may be reduced to considering the situation

$$M \leftarrow M' \Rightarrow N$$

with both maps nil-isomorphisms in which case Proposition 2.25 shows that $M_{\min} \cong M'_{\min} \cong N_{\min}$. One has the following Proposition.

Proposition 3.1. *Let X be regular and F -finite. Then the composition*

$$\mathbf{Min}_\gamma(X) \hookrightarrow \mathbf{Coh}_\gamma(X) \twoheadrightarrow \mathbf{Crys}_\gamma(X)$$

is an equivalence of categories whose inverse is given by sending a γ -crystal represented by the γ -sheaf M to the minimal γ -sheaf M_{\min} .

Proof. The existence of M_{\min} shows that $\mathbf{Min}_\gamma(X) \rightarrow \mathbf{Crys}_\gamma(X)$ is essentially surjective. It remains to show that $\mathrm{Hom}_{\mathbf{Min}_\gamma}(M, N) \cong \mathrm{Hom}_{\mathbf{Crys}_\gamma}(M, N)$. A map $\varphi: M \rightarrow N$ of minimal γ -sheaves is zero in \mathbf{Crys}_γ if and only if $\mathrm{img} \varphi$ is nilpotent. But $\mathrm{img} \varphi$ is a subsheaf of the minimal γ -sheaf N , which by Definition 2.7 (a) has no nontrivial nilpotent subsheaves. Hence $\mathrm{img} \varphi = 0$ and therefore $\varphi = 0$. This shows that the map on Hom sets is injective. The surjectivity follows again by functoriality of $M \mapsto M_{\min}$. □

Corollary 3.2. *Let X be regular and F -finite. The category of minimal γ -sheaves $\mathbf{Min}_\gamma(X)$ is an abelian category. If $\varphi: M \rightarrow N$ is a morphism then*

$$\ker_{\min} \varphi = (\ker \varphi)_{\min} = \underline{\ker \varphi} \quad \text{and} \quad \mathrm{coker}_{\min} \varphi = (\mathrm{coker} \varphi)_{\min} = \overline{\mathrm{coker} \varphi}.$$

Proof. Since $\mathbf{Min}_\gamma(X)$ is equivalent to $\mathbf{Crys}_\gamma(X)$ and since the latter is abelian, so is $\mathbf{Min}_\gamma(X)$. Since \ker and coker in $\mathbf{Crys}_\gamma(X)$ are represented by the kernel and cokernel of the underlying coherent sheaf the statement about \ker and coker in $\mathbf{Min}_\gamma(X)$ follows, where overline and underline are as defined in Proposition 2.5 and Proposition 2.21. □

3.2. The parameter test module. We give an application to the theory of tight closure. In [Blickle 2004, Proposition 4.5], it was shown that the parameter test module τ_{ω_A} is the unique minimal root of the intersection homology unit module $\mathcal{L} \subseteq H_I^{n-d}(R)$ if $A = R/I$ is the quotient of the regular local ring R (where $\dim R = n$ and $\dim A = d$). Locally, the parameter test module τ_{ω_A} is defined as the Matlis dual of

$$H_m^d(A)/0_{H_m^d(A)}^*$$

where $0_{H_m^d(A)}^*$ is the tight closure of zero in $H_m^d(A)$. The fact that we are now able to construct minimal γ -sheaves globally allows us to give a global candidate for the parameter test module.

Proposition 3.3. *Let $A = R/I$ be equi-dimensional of dimension d where R is regular and F -finite. Then there is a submodule*

$$L \subseteq \omega_A = \text{Ext}^{n-d}(R/I, R)$$

such that for each $x \in \text{Spec } A$ we have $L_x \cong \tau_{\omega_x}$.

Proof. Let $\mathcal{L} \subseteq H_I^{n-d}(R)$ be the unique smallest submodule of $H_I^{n-d}(R)$ which agrees with $H_I^{n-d}(R)$ on all smooth points of $\text{Spec } A$. By [Blickle 2004, Theorem 4.1] \mathcal{L} exists and is a unit \mathbb{O}_X -submodule of $H_I^{n-d}(R)$. Let L be a minimal generator of \mathcal{L} , that is, a coherent minimal γ -sheaf such that $\text{Gen } L = \mathcal{L}$ which exists due to Theorem 2.22. Because of Proposition 2.15 it follows that L_x is also a minimal γ -sheaf and $\text{Gen } L_x \cong \mathcal{L}_x$. But from [Blickle 2004, Proposition 4.5] we know that the unique minimal root of \mathcal{L}_x is $\tau_{\omega_{A_x}}$, the parameter test module of A_x . It follows that $L_x \cong \tau_{\omega_{A_x}}$ by uniqueness. To see that $L \subseteq \text{Ext}^{n-d}(R/I, R)$ we just observe that $\text{Ext}^{n-d}(R/I, R)$ with the map induced by $\sigma^*(R/I) = R/I^{[q]} \rightarrow R/I$ is a γ -sheaf which generates $H_I^{n-d}(R)$. Furthermore, the map

$$\text{Ext}^{n-d}(R/I, R) \rightarrow \sigma^* \text{Ext}^{n-d}(R/I, R)$$

is injective since it is so locally and in this case the map is dual to the surjection $\sigma^* H_m^d(R/I) \rightarrow H_m^d(R/I)$ ($d = \dim R/I$) via local duality for the local ring (R_m, \mathfrak{m}) . Hence by minimality of L we have the desired inclusion. \square

3.3. Test ideals and minimal γ -sheaves. We consider now the simplest example of a γ -sheaf, namely that of a free rank one R -module $M (\cong R)$. That means that via the identification $R \cong \sigma^* R$ the structural map

$$\gamma : M \cong R \xrightarrow{f \cdot} R \cong \sigma^* R \cong \sigma^* M$$

is given by multiplication with an element $f \in R$. It follows that γ^e is given by multiplication by $f^{1+q+\dots+q^{e-1}}$ under the identification of $\sigma^{e*} R \cong R$. It is an easy exercise to observe that $\text{Gen } M \cong R_f$ with its usual unit $R[F]$ -structure.

We will show that the minimal γ -subsheaf of the just described γ -sheaf M can be expressed in terms of generalized test ideals. We recall from [Blickle et al. 2008, Lemma 2.1] that the test ideal of a principal ideal (f) of exponent $\alpha = \frac{m}{q^e}$ is given by

$$\tau(f^\alpha) = \text{the smallest ideal } J \text{ such that } f^m \in J^{[q^e]}.$$

By Lemma 2.2 of op. cit. $\tau(f^\alpha)$ can also be characterized as σ^{-e} of the $D^{(e)}$ -module generated by f^m . We set as a shorthand

$$J_e = \tau(f^{(1+q+q^2+\dots+q^{e-1})/q^e})$$

and repeat the definition

$$J_e = \text{the smallest ideal } J \text{ of } R \text{ such that } f^{1+q+q^2+\dots+q^{e-1}} \in J^{[q^e]}$$

and we set $J_0 = R$. Further recall from Section 2.5 that

$$M_e = \text{smallest ideal } I \text{ of } R \text{ such that } f \cdot M_{e-1} \subseteq I^{[q]}$$

with $M_0 = M$.

Lemma 3.4. *For all $e \geq 0$ one has $J_e = M_e$.*

Proof. The equality is true for $e = 0, 1$ by definition. We first show the inclusion $J_e \subseteq M_e$ by induction on e .

$$\begin{aligned} M_e^{[p^e]} &\supseteq (f \cdot M_{e-1})^{[q^{e-1}]} = (f^{q^{e-1}} M_{e-1}^{[q^{e-1}]}) \\ &= (f^{q^{e-1}} J_{e-1}^{[q^{e-1}]}) \supseteq f^{q^{e-1}} \cdot f^{1+q+q^2+\dots+q^{e-2}} \\ &= f^{1+q+q^2+\dots+q^{e-1}}, \end{aligned}$$

and since J_e is minimal with respect to this inclusion we have $J_e \subseteq M_e$.

Now we show for all $e \geq 1$ that $f \cdot J_{e-1} \subseteq J_e^{[q]}$. The definition of J_e implies that

$$f^{1+q+\dots+q^{e-2}} \in (J^{[q^e]} : f^{q^{e-1}}) = (J^{[q]} : f)^{[q^{e-1}]}$$

which implies that $J_{e-1} \subseteq (J^{[q]} : f)$ by minimality of J_{e-1} . Hence $f \cdot J_{e-1} \subseteq J^{[q]}$. Now, we can show the inclusion $M_e \subseteq J_e$ by observing that by induction one has

$$J_e^{[q]} \supseteq f \cdot J_{e-1} \supseteq f \cdot M_{e-1}.$$

which implies by minimality of M_e that $M_e \subseteq J_e$. □

This shows that the minimal γ -sheaf M_{\min} , which is equal to M_e for $e \gg 0$ by Proposition 2.21, is just the test ideal $\tau(f^{(1+q+q^2+\dots+q^{e-1})/q^e})$ for $e \gg 0$. As a consequence we have:

Proposition 3.5. *Let M be the γ -sheaf given by*

$$R \xrightarrow{f \cdot} R \cong \sigma^* R.$$

Then $M_{\min} = \tau(f^{(1+q+q^2+\dots+q^{e-1})/q^e})$ for $q \gg 0$. In particular, $M_{\min} \supseteq \tau(f^{\frac{1}{q-1}})$ and the F -pure-threshold of $f \geq \frac{1}{q-1}$ if and only if M is minimal.

Proof. For $e \gg 0$ the increasing sequence of rational numbers $(1 + q + q^2 + \dots + q^{e-1})/q^e$ approaches $\frac{1}{q-1}$. Hence

$$M_e = \tau(f^{(1+q+q^2+\dots+q^{e-1})/q^e}) \supseteq \tau(f^{\frac{1}{q-1}})$$

for all e . If M is minimal, then all M_e are equal and hence the multiplier ideals $\tau(f^\alpha)$ must be equal to R for all $\alpha \in [0, \frac{1}{q-1})$. In particular, the F -pure-threshold

of $f \geq \frac{1}{q-1}$. Conversely, if the F -pure threshold is less than $\frac{1}{q-1}$, then for some e we must have that

$$\tau\left(f^{(1+q+q^2+\dots+q^{e-1})/q^e}\right) \neq \tau\left(f^{(1+q+q^2+\dots+q^e)/q^{e+1}}\right)$$

so $M_e \neq M_{e+1}$ which implies that $M \neq M_1$. So M is not minimal. \square

Remark 3.6. After replacing f by f^r , this also shows that $\frac{r}{q-1}$ is not an accumulation point of F -thresholds of f for any f in an F -finite regular ring. In [Blickle et al. 2008] this was shown for R essentially of finite type over a local ring since our argument there depended on [Lyubeznik 1997, Theorem 4.2]. Even though D -modules appear in the present article, they only do so by habit of the author; as remarked before, they can easily be avoided.

Remark 3.7. Of course, for $r = q - 1$ this recovers (and slightly generalizes) the main result in [Alvarez-Montaner et al. 2005].

Remark 3.8. I expect that this descriptions of minimal roots can be extended to a more general setting using the modifications of generalized test ideals to modules as introduced in the preprint [Takagi and Takahashi 2007].

References

- [Alvarez-Montaner et al. 2005] J. Alvarez-Montaner, M. Blickle, and G. Lyubeznik, “Generators of D -modules in positive characteristic”, *Math. Res. Lett.* **12**:4 (2005), 459–473. MR 2006m:13024 Zbl 02207823
- [Blickle 2004] M. Blickle, “The intersection homology D -module in finite characteristic”, *Math. Ann.* **328**:3 (2004), 425–450. MR 2005a:14005 Zbl 1065.14006
- [Blickle and Böckle \geq 2008] M. Blickle and G. Böckle, “Cartier crystals”. In preparation.
- [Blickle et al. 2008] M. Blickle, M. Mustařa, and K. E. Smith, “ F -thresholds of hypersurfaces”, *Trans AMS* (2008). To appear.
- [Emerton and Kisin 2004] M. Emerton and M. Kisin, “The Riemann–Hilbert correspondence for unit F -crystals”, *Astérisque* 293 (2004), vi+257. MR 2005e:14027 Zbl 1056.14025
- [Hara and Takagi 2004] N. Hara and S. Takagi, “On a generalization of test ideals”, *Nagoya Math. J.* **175** (2004), 59–74. MR 2005g:13009 Zbl 02138455
- [Hartshorne and Speiser 1977] R. Hartshorne and R. Speiser, “Local cohomological dimension in characteristic p ”, *Ann. of Math. (2)* **105**:1 (1977), 45–79. MR 56 #353 Zbl 0362.14002
- [Katz 1996] N. M. Katz, *Rigid local systems*, Annals of Mathematics Studies **139**, Princeton University Press, Princeton, NJ, 1996. MR 97e:14027 Zbl 0864.14013
- [Katzman et al. 2007] M. Katzman, G. Lyubeznik, and W. Zhang, “On the discreteness and rationality of jumping coefficients”, preprint, 2007.
- [Kunz 1969] E. Kunz, “Characterizations of regular local rings for characteristic p ”, *Amer. J. Math.* **91** (1969), 772–784. MR 40 #5609 Zbl 0188.33702
- [Lyubeznik 1997] G. Lyubeznik, “ F -modules: applications to local cohomology and D -modules in characteristic $p > 0$ ”, *J. Reine Angew. Math.* **491** (1997), 65–130. MR 99c:13005 Zbl 0904.13003

[Miller 2007] C. Miller, *Cohomology of p -torsion sheaves on characteristic- p curves*, PhD Thesis, UC Berkeley, 2007.

[Takagi and Takahashi 2007] S. Takagi and R. Takahashi, “ D -modules over rings of F -finite representation type”, preprint, 2007.

Communicated by Craig Huneke

Received 2007-12-10 Revised 2008-02-13 Accepted 2008-03-02

manuel.blickle@uni-due.de

*Mathematik Essen, Universität Duisburg-Essen,
45117 Essen, Germany*
www.mabli.org

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in ANT are usually in English, but articles written in other languages are welcome.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use L^AT_EX but submissions in other varieties of T_EX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibT_EX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@mathscipub.org with details about how your graphics were generated.

White Space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 2 No. 3 2008

The intersection of a curve with a union of translated codimension-two subgroups in a power of an elliptic curve	249
EVELINA VIADA	
Diophantine subsets of function fields of curves	299
JÁNOS KOLLÁR	
Symmetric obstruction theories and Hilbert schemes of points on threefolds	313
KAI BEHREND AND BARBARA FANTECHI	
Minimal γ -sheaves	347
MANUEL BLICKLE	



1937-0652(2008)2:3;1-D