Algebra & Number Theory

Volume 10
2016
No. 10

msp

Algebra & Number Theory

msp.org/ant

EDITORS

MANAGING EDITOR

Bjorn Poonen
Massachusetts Institute of Technology
Cambridge, USA

EDITORIAL BOARD CHAIR

David Eisenbud
University of California
Berkeley, USA

BOARD OF EDITORS

Dave Benson	University of Aberdeen, Scotland	Susan Montgomery	University of Southern California, USA
Richard E. Borcherds	University of California, Berkeley, USA	Shigefumi Mori	RIMS, Kyoto University, Japan
John H. Coates	University of Cambridge, UK	Raman Parimala	Emory University, USA
J-L. Colliot-Thélène	CNRS, Université Paris-Sud, France	Jonathan Pila	University of Oxford, UK
Brian D. Conrad	Stanford University, USA	Anand Pillay	University of Notre Dame, USA
Hélène Esnault	Freie Universität Berlin, Germany	Victor Reiner	University of Minnesota, USA
Hubert Flenner	Ruhr-Universität, Germany	Peter Sarnak	Princeton University, USA
Sergey Fomin	University of Michigan, USA	Joseph H. Silverman	Brown University, USA
Edward Frenkel	University of California, Berkeley, USA	Michael Singer	North Carolina 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
Roger Heath-Brown	Oxford University, UK	Ravi Vakil	Stanford University, USA
Craig Huneke	University of Virginia, USA	Michel van den Bergh	Hasselt University, Belgium
Kiran S. Kedlaya	Univ. of California, San Diego, USA	Marie-France Vignéras	Université Paris VII, France
János Kollár	Princeton University, USA	Kei-Ichi Watanabe	Nihon University, Japan
Yuri Manin	Northwestern University, USA	Efim Zelmanov	University of California, San Diego, USA
Philippe Michel	École Polytechnique Fédérale de Lausanne	Shou-Wu Zhang	Princeton University, USA

PRODUCTION

production@msp.org

Silvio Levy, Scientific Editor

See inside back cover or msp.org/ant for submission instructions.

The subscription price for 2016 is US $290/year for the electronic version, and $485/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscribers address should be sent to MSP.

Algebra & Number Theory (ISSN 1944-7833 electronic, 1937-0652 printed) at Mathematical Sciences Publishers, 798 Evans Hall #3840, c/o University of California, Berkeley, CA 94720-3840 is published continuously online. Periodical rate postage paid at Berkeley, CA 94704, and additional mailing offices.

ANT peer review and production are managed by EditFLOW® from MSP.

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

© 2016 Mathematical Sciences Publishers

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2053

Weight functions on Berkovich curves

Matthew Baker and Johannes Nicaise

Let C be a curve over a complete discretely valued field K. We give tropical descriptions of the weight function attached to a pluricanonical form on C and the essential skeleton of C. We show that the Laplacian of the weight function equals the pluricanonical divisor on Berkovich skeleta, and we describe the essential skeleton of C as a combinatorial skeleton of the Berkovich skeleton of the minimal *snc*-model. In particular, if C has semistable reduction, then the essential skeleton coincides with the minimal skeleton. As an intermediate step, we describe the base loci of logarithmic pluricanonical line bundles on minimal *snc*-models.

1. Introduction

We denote by R a complete discrete valuation ring with quotient field K and algebraically closed residue field k. Let X be a smooth and proper K-variety. Mustaţă and Nicaise [2015] defined the *essential skeleton* $\mathrm{Sk}(X)$ of X, which is a finite simplicial complex embedded in the Berkovich analytification X^{an} of X. It is a union of faces of the Berkovich skeleton of any strict normal crossings model of X, but it does not depend on the choice of such a model. It was proven in [Nicaise and Xu 2013] that, when k has characteristic zero and the canonical line bundle on X is semiample, the essential skeleton is a strong deformation retract of X^{an} and can be identified with the dual intersection complex of the special fiber of any minimal *dlt*-model of X over R. The definition of the essential skeleton was based on the construction of a *weight function* wt_ω on X^{an} attached to a pluricanonical form ω on X, which measures the degeneration of the pair (X, ω) locally at a point of X^{an}. The aim of the present paper is to give an explicit description of the weight function and the essential skeleton in the case where X is a curve, and to relate them to potential theory on graphs.

Let C be a smooth, proper, geometrically connected curve over K. Denote by $\mathbb{H}_0(C)$ the Berkovich analytification C^{an} minus the points of type I and IV. In Section 2 we construct a metric on $\mathbb{H}_0(C)$ using the geometry of normal crossings models of C over R. This is similar to the construction of the skeletal metric in

MSC2010: primary 14D10; secondary 14E30, 14T05.
Keywords: Berkovich spaces, degenerations of curves, tropical geometry.

the case where K is algebraically closed [Baker et al. 2013], but our metric is not invariant under base change and cannot be obtained from the skeletal metric in any direct way. Using this metric, we can speak of integral affine functions on finite subgraphs of C^{an} and Laplacians of such functions. Section 3 is the heart of the paper; here we provide combinatorial descriptions of the weight function wt_ω attached to a rational m-canonical form ω on C and of the essential skeleton of C. Our first main result, Theorem 3.2.3, states that the Laplacian of the restriction of the weight function to the Berkovich skeleton of a suitable *snc*-model of C equals the m-canonical divisor of the Berkovich skeleton, which is defined in terms of graph theory. Our second main result, Theorem 3.3.13, states that the essential skeleton of a curve C of positive genus is the subgraph of the Berkovich skeleton of the minimal *snc*-model of C obtained by contracting all the tails of rational curves. In particular, if C has semistable reduction, then the essential skeleton of C is equal to the Berkovich skeleton of its minimal *snc*-model. The proof of Theorem 3.3.13 is based on Theorem 3.3.6, which describes the base locus of the logarithmic relative pluricanonical bundles of the minimal *snc*-model of C. We also prove that, in the semistable reduction case, it suffices to look at weight functions of 2-canonical forms to recover the essential skeleton; moreover, if the essential skeleton of C is bridgeless, then canonical forms suffice (see Theorem 3.4.6). Finally, in the Appendix, we describe a different natural metric on $\mathbb{H}_0(C)$ which behaves better under (tame) base change and which is closer to the skeletal metric from [Baker et al. 2013].

1.1. *Notation.*

1.1.1. We denote by R a complete discrete valuation ring with quotient field K and algebraically closed residue field k. We assume that the valuation v_K on K is normalized, i.e., that $v_K(t) = 1$ for any uniformizer t in R, and we define an absolute value $|\cdot|_K$ on K by setting $|a|_K = \exp(-v_K(a))$ for every a in K^*. We fix an algebraic closure K^a of K. The absolute value $|\cdot|_K$ extends uniquely to an absolute value on K^a, which we still denote by $|\cdot|_K$. We write $\widehat{K^a}$ for the completion of K^a with respect to $|\cdot|_K$.

1.1.2. By a curve over K, we will mean a geometrically connected smooth proper K-variety of dimension one. For every scheme S we denote by S_{red} the maximal reduced closed subscheme. For every R-scheme $\mathscr{X}$ we set $\mathscr{X}_K = \mathscr{X} \times_R K$ and $\mathscr{X}_k = \mathscr{X} \times_R k$. If $\mathcal{L}$ is a line bundle on a scheme X and D is a Cartier divisor on X, then we write $\mathcal{L}(D)$ for the line bundle $\mathcal{L} \otimes \mathcal{O}_X(D)$, as usual.

1.1.3. We will work with the category of K-analytic spaces as defined by Berkovich [1990]. We assume a basic familiarity with the theory of analytic curves over K; see for instance [Baker et al. 2013].

2. The metric on the Berkovich analytification of a K-curve

2.1. *Metric graphs associated to curves with normal crossings.*

2.1.1. When we speak of a discrete graph G, we mean a finite connected undirected multigraph, i.e., we allow multiple loops and multiple edges between vertices. We denote the vertex set of G by $V(G)$ and the set of edges by $E(G)$. A weighted discrete graph is a couple (G, w) where G is a discrete graph and w is a function

$$w : V(G) \to \mathbb{R}.$$

2.1.2. A discrete graph G has a geometric realization Γ, which is defined as follows: We start from the set $V(G)$ and we attach one copy of the closed interval $[0, 1]$ between two vertices v_1 and v_2 for each edge of G with endpoints $\{v_1, v_2\}$. If G is endowed with a weight function w that takes values in $\mathbb{Z}_{>0}$, then we can turn the topological space Γ into a metric space by declaring that the length of every edge e between two adjacent vertices v_1 and v_2 is equal to

$$\ell(e) = \frac{1}{w(v_1)\cdot w(v_2)}. \tag{2.1.3}$$

In these definitions, we allow the possibility that $v_1 = v_2$. We call the metric space Γ the metric graph associated with (G, w).

2.1.4. Let X be a connected separated k-scheme of finite type of pure dimension one. We say that X has normal crossings if the only singular points of X_{red} are ordinary double points. We associate a weighted discrete graph $(G(X), w)$ to X as follows. The vertex set of $G(X)$ is the set of irreducible components of X and the edge set of $G(X)$ is the set of singular points of X_{red}. If e is an edge corresponding to a singular point x of X_{red}, then the end points of e are the vertices corresponding to the irreducible components of X containing x. In particular, e is a loop if and only if x is a singular point of an irreducible component of X. If v is a vertex of $G(X)$ corresponding to an irreducible component E of X, then the weight $w(v)$ is defined to be the multiplicity of X along E, i.e., the length of the local ring of X at the generic point of E. The metric graph associated with $(G(X), w)$ will be denoted by $\Gamma(X)$.

2.2. *Models with normal crossings.*

2.2.1. Let C be a curve over K. An *nc*-model of C is a regular flat proper R-scheme $\mathscr{C}$, endowed with an isomorphism of K-schemes $\mathscr{C}_K \to C$, such that the special fiber $\mathscr{C}_k$ has normal crossings. We call $\mathscr{C}$ an *snc*-model of C if, moreover, $\mathscr{C}_k$ has strict normal crossings, which means that its irreducible components (endowed with the induced reduced structure) are regular. If $\mathscr{C}$ and $\mathscr{C}'$ are *nc*-models of C, then a morphism of R-schemes $h : \mathscr{C}' \to \mathscr{C}$ is called a morphism of *nc*-models

if the morphism $h_K : \mathscr{C}'_K \to \mathscr{C}_K$ obtained by base change to K commutes with the isomorphisms to C. Morphisms of *snc*-models are defined analogously. We say that $\mathscr{C}'$ dominates $\mathscr{C}$ if there exists a morphism of *nc*-models $\mathscr{C}' \to \mathscr{C}$; such a morphism is automatically unique. We denote this property by $\mathscr{C}' \geq \mathscr{C}$. The relation $\geq$ defines a partial ordering on the set of isomorphism classes of *nc*-models of C. This partial ordering is filtered, and the *snc*-models form a cofinal subset since any *nc*-model can be transformed into an *snc*-model by blowing up at the self-intersection points of the irreducible components of the special fiber. We say that the curve C has semistable reduction if any relatively minimal *nc*-model of C has a reduced special fiber. Beware that this does not imply that the minimal *snc*-model has reduced special fiber, as blowing up at self-intersection points introduces components of multiplicity two.

2.2.2. Denote by C^{an} the Berkovich analytification of C, and let $\mathscr{C}$ be an *snc*-model of C. If E is an irreducible component of $\mathscr{C}_k$ and v denotes the corresponding vertex of the weighted discrete graph $(G(\mathscr{C}_k), w)$, then $w(v)$ is precisely the multiplicity of E in the divisor $\mathscr{C}_k$. Mustață and Nicaise [2015, §3.1] defined a canonical topological embedding of the metric graph $\Gamma(\mathscr{C}_k)$ into C^{an}, generalizing a construction by Berkovich. The image of this embedding is called the Berkovich skeleton of the model $\mathscr{C}$ and denoted by $\mathrm{Sk}(\mathscr{C})$. By [Mustață and Nicaise 2015, 3.1.5], the embedding of $\mathrm{Sk}(\mathscr{C})$ into C^{an} has a canonical continuous retraction

$$\rho_{\mathscr{C}} : C^{\mathrm{an}} \to \mathrm{Sk}(\mathscr{C}).$$

If we let $\mathscr{C}$ vary over the class of *snc*-models of C, ordered by the domination relation, then the maps $\rho_{\mathscr{C}}$ induce a homeomorphism

$$C^{\mathrm{an}} \to \varprojlim_{\mathscr{C}} \mathrm{Sk}(\mathscr{C}).$$

This is easily proven by an adaptation of the argument in [Baker et al. 2013, Theorem 5.2] (where it is assumed that the base field is algebraically closed). It is straightforward to generalize these constructions to *nc*-models, either by copying the arguments or by observing that blowing up $\mathscr{C}$ at all the self-intersection points of irreducible components of $\mathscr{C}_k$, we get an *snc*-model $\mathscr{C}'$ of C and the morphism $\mathscr{C}' \to \mathscr{C}$ induces an isometry $\Gamma(\mathscr{C}'_k) \to \Gamma(\mathscr{C}_k)$ (the effect of this operation on $\Gamma(\mathscr{C}_k)$ is that we add a vertex in the middle of every loop).

2.3. *Definition of the metric.*

2.3.1. Let C be a curve over K, and denote by $\mathbb{H}_0(C)$ the subset of C^{an} obtained by removing the points of type I and IV.

Lemma 2.3.2. *For every nc-model $\mathscr{C}$ of C, the Berkovich skeleton* $\mathrm{Sk}(\mathscr{C})$ *is contained in* $\mathbb{H}_0(C)$. *Moreover,* $\mathbb{H}_0(C)$ *is the union of the skeleta* $\mathrm{Sk}(\mathscr{C})$ *where* $\mathscr{C}$ *runs through any cofinal set of nc-models of* $\mathscr{C}$.

Proof. The points of type II on C^{an} are precisely the divisorial points in the sense of [Mustaţă and Nicaise 2015, 2.4.7], and the points of type II and III are precisely the monomial points. Thus the first part of the statement is obvious from the construction of $\mathrm{Sk}(\mathscr{C})$. The second part follows from the fact that every monomial point lies in the skeleton of some *snc*-model and the fact that, if $\mathscr{C}' \to \mathscr{C}$ is a morphism of *nc*-models of C, the skeleton $\mathrm{Sk}(\mathscr{C})$ is included in $\mathrm{Sk}(\mathscr{C}')$ [Mustaţă and Nicaise 2015, Proposition 3.1.7]. □

The following theorem explains how to define a natural metric on the set $\mathbb{H}_0(C)$.

Theorem 2.3.3. *There exists a unique metric on $\mathbb{H}_0(C)$ such that, for every nc-model $\mathscr{C}$ of C, the map*

$$\Gamma(\mathscr{C}_k) \to \mathbb{H}_0(C)$$

is an isometric embedding.

Proof. The uniqueness of the metric is obvious from Lemma 2.3.2. Thus it suffices to prove its existence. Let $\mathscr{C}$ and $\mathscr{C}'$ be *nc*-models of C such that $\mathscr{C}'$ dominates $\mathscr{C}$. Then the skeleton $\mathrm{Sk}(\mathscr{C})$ is contained in $\mathrm{Sk}(\mathscr{C}')$ by [Mustaţă and Nicaise 2015, Proposition 3.1.7], and it suffices to show that the corresponding embedding $\Gamma(\mathscr{C}_k) \to \Gamma(\mathscr{C}'_k)$ is an isometry. Since we can decompose the morphism $\mathscr{C}' \to \mathscr{C}$ into a finite composition of point blow-ups, we can assume that $\mathscr{C}' \to \mathscr{C}$ is the blow-up of $\mathscr{C}$ at a closed point x of $\mathscr{C}_k$. If x is a regular point of $(\mathscr{C}_k)_{\mathrm{red}}$ then the claim is obvious. If x is a singular point then $(G(\mathscr{C}'_k), w)$ is obtained from $(G(\mathscr{C}_k), w)$ by adding a vertex on the edge e corresponding to x and giving it weight $w(v_1)+w(v_2)$, where v_1 and v_2 are the (not necessarily distinct) endpoints of e. The lengths of the segment e in the metric graphs $\Gamma(\mathscr{C}_k)$ and $\Gamma(\mathscr{C}'_k)$ are the same, because

$$\frac{1}{w(v_1)\cdot w(v_2)} = \frac{1}{w(v_1)\cdot(w(v_1)+w(v_2))} + \frac{1}{(w(v_1)+w(v_2))\cdot w(v_2)}. \qquad \square$$

Remark 2.3.4. There is another metric on $\mathbb{H}_0(C)$ that is induced by the piecewise integral affine structure on the skeleta of *snc*-models; we will explain its construction in the Appendix. Although this second metric arises more naturally and behaves better under base change, the one we defined in Theorem 2.3.3 seems to be the correct one for the purposes of potential theory. A similar discrepancy appears in the nonarchimedean study of germs of algebraic surfaces, which is in many ways analogous to the setup we consider here, see Section 7.4.10 of [Jonsson 2015].

3. The weight function and the essential skeleton

3.1. *The weight function attached to a pluricanonical form.*

3.1.1. We fix a K-curve C. Let m be a positive integer and let ω be a nonzero rational m-canonical form on C. Thus ω is a nonzero rational section of the m-canonical line bundle $\omega_{C/K}^{\otimes m}$. As such, it defines a Cartier divisor on C, which we

denote by $\mathrm{div}_C(\omega)$. If $\mathscr{C}$ is any *snc*-model of C, we can also view ω as a rational section of the logarithmic relative m-canonical line bundle

$$\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$$

and we denote the corresponding divisor on $\mathscr{C}$ by $\mathrm{div}_{\mathscr{C}}(\omega)$. Note that the horizontal part of $\mathrm{div}_{\mathscr{C}}(\omega)$ is simply the schematic closure of $\mathrm{div}_C(\omega)$ in $\mathscr{C}$.

3.1.2. Mustaţă and the second author [2015, 4.5.4] attached to ω a so-called *weight function* wt_ω. In our setting (the case of curves) we can characterize its restriction to $\mathbb{H}_0(C)$ in the following way. Recall that the points of type II on C^{an} are precisely the divisorial points in the sense of [Mustaţă and Nicaise 2015, 2.4.7], and that the points of type II and III are precisely the monomial points.

Proposition 3.1.3. *The weight function*

$$\mathrm{wt}_\omega : \mathbb{H}_0(C) \to \mathbb{R}$$

is the unique function with the following properties for every snc-model $\mathscr{C}$ of C:

(1) *The restriction of* wt_ω *to* $\mathrm{Sk}(\mathscr{C})$ *is continuous with respect to the metric topology (which coincides with the Berkovich topology on* $\mathrm{Sk}(\mathscr{C})$*).*

(2) *Let E be an irreducible component of $\mathscr{C}_k$. We denote by N and ν the multiplicities of E in $\mathscr{C}_k$ and* $\mathrm{div}_{\mathscr{C}}(\omega)$*, respectively. If x is the divisorial point of C^{an} attached to $(\mathscr{C}, E)$ (equivalently, the vertex of* $\mathrm{Sk}(\mathscr{C})$ *corresponding to E), then*

$$\mathrm{wt}_\omega(x) = \frac{\nu}{N}.$$

Proof. It is shown in [Mustaţă and Nicaise 2015, 4.4.3] that the weight function is continuous (even piecewise affine) on $\mathrm{Sk}(\mathscr{C})$, and the description at divisorial points is part of its definition. Uniqueness is clear from Lemma 2.3.2 and the fact that the divisorial points are dense in the skeleton of every *snc*-model of C (by the proof of [Mustaţă and Nicaise 2015, 2.4.8], they correspond precisely to the points on $\Gamma(\mathscr{C})$ with rational barycentric coordinates in the sense of 3.1.2 of the same work. □

3.1.4. Beware that the weight function is not continuous with respect to the Berkovich topology on $\mathbb{H}_0(C)$ (see [Mustaţă and Nicaise 2015, Remark 4.6] for a counterexample). The explicit description of the weight function given in Theorem 3.2.3 below shows in particular that it is continuous with respect to the metric topology on $\mathbb{H}_0(C)$ (which is strictly finer than the Berkovich topology).

3.2. *The Laplacian of the weight function.*

3.2.1. By a *pair* over K, we mean a couple (C, δ) consisting of a K-curve C and a divisor δ on C. An *nc*-model of a pair (C, δ) is an *nc*-model $\mathscr{C}$ of C such that the sum of $\mathscr{C}_k$ with the schematic closure of δ is a normal crossings divisor on $\mathscr{C}$. An *snc*-model of (C, δ) is defined analogously. Note that for every point x in the

support of δ, the specialization of x to $\mathscr{C}_k$ lies in a unique irreducible component E of $\mathscr{C}_k$, and the multiplicity of $\mathscr{C}_k$ along E is equal to the degree of x over K, by the normal crossings condition. The skeleton of $(\mathscr{C}, \delta)$ is defined to be the intersection of $\mathbb{H}_0(C)$ with the convex hull in C^{an} of $\mathrm{Sk}(\mathscr{C})$ and the support of δ. We will denote it by $\mathrm{Sk}(\mathscr{C}, \delta)$. Thus we obtain $\mathrm{Sk}(\mathscr{C}, \delta)$ from $\mathrm{Sk}(\mathscr{C})$ by adding, for each point x in the support of δ, the open branch running from $\mathrm{Sk}(\mathscr{C})$ towards x. This construction is similar to the definition of the skeleton of a strictly semistable pair in [Gubler et al. 2016], but there it is assumed that K is algebraically closed and that $\mathscr{C}_k$ is reduced and has strict normal crossings. By restricting the metric on $\mathbb{H}_0(C)$ to the skeleton $\mathrm{Sk}(\mathscr{C}, \delta)$, we can view the skeleton as a metric graph with some half-open edges of infinite length. Then it makes sense to speak about a $\mathbb{Z}$-affine function f on $\mathrm{Sk}(\mathscr{C}, \delta)$ (i.e., a continuous real-valued function that is integral affine on every edge) and the Laplacian $\Delta(f)$ of such a function (the divisor on $\mathrm{Sk}(\mathscr{C}, \delta)$ whose degree at a vertex is the sum of the outgoing slopes of f).

3.2.2. Our aim is to give a combinatorial description of the weight function wt_ω on $\mathbb{H}_0(C)$ attached to a nonzero rational m-canonical form ω on C. For this description we need to introduce the m-canonical divisor of a labeled graph. Let G be a discrete graph without loops, where we allow some of the edges of G to be half-open (i.e., the edge has only one adjacent vertex and is unbounded at the other side). Assume that each vertex v of G is labeled by a couple of nonnegative integers $(N(v), g(v))$. Then the canonical divisor of G is defined by

$$K_G = \sum_{v \in V(G)} N(v)(\mathrm{val}(v) + 2g(v) - 2)v,$$

where $\mathrm{val}(v)$ denotes the valency at v, that is, the number of edges (bounded and unbounded) in G adjacent to v. When $N(v) = 1$ and $g(v) = 0$ for every vertex v, this is just the usual definition of the canonical divisor of a discrete graph. The m-canonical divisor of G is defined as m times the canonical divisor K_G.

Theorem 3.2.3. *We fix a K-curve C. Let m be a positive integer and let ω be a nonzero rational m-canonical form on C. Let δ be any divisor on C whose support contains the support of* $\mathrm{div}_C(\omega)$ *and let $\mathscr{C}$ be an snc-model for the pair (C, δ):*

(1) *The weight function* wt_ω *is $\mathbb{Z}$-affine on every edge of* $\mathrm{Sk}(\mathscr{C}, \delta)$.

(2) *For every point x in the support of δ, the weight function* wt_ω *has constant slope on the path running from* $\mathrm{Sk}(\mathscr{C})$ *to x in* C^{an}, *and this slope is equal to*

$$N(m + \deg_x(\mathrm{div}_C(\omega))),$$

where N denotes the multiplicity of the unique component in $\mathscr{C}_k$ containing the specialization of x.

(3) *The Laplacian of the restriction of* wt_ω *to* $\mathrm{Sk}(\mathscr{C}, \delta)$ *is equal to the m-canonical divisor of the graph* $\mathrm{Sk}(\mathscr{C}, \delta)$ *if we label each vertex v with* $(N(v), g(v))$,

where $N(v)$ is the multiplicity of the corresponding irreducible component in $\mathscr{C}_k$ and $g(v)$ denotes its genus.

Proof. (1) It follows from the proof of [Mustaţă and Nicaise 2015, 4.4.3] that wt_ω is $\mathbb{Z}$-affine on $\mathrm{Sk}(\mathscr{C})$, because no point in the support of $\mathrm{div}_C(\omega)$ specializes to a singular point of $(\mathscr{C}_k)_{\mathrm{red}}$. Here some care is needed, since the $\mathbb{Z}$-affine structure in [Mustaţă and Nicaise 2015] is not the same as the one induced by our metric; it corresponds to the metric one obtains by replacing the definition in (2.1.3) by

$$\ell(e) = \frac{1}{\mathrm{lcm}\{w(v_1), w(v_2)\}}.$$

Since this multiplies every edge length by an integer factor, every $\mathbb{Z}$-affine function in the sense of [Mustaţă and Nicaise 2015] is also $\mathbb{Z}$-affine with respect to the metric we use; we will come back to this point in the Appendix. The fact that wt_ω is also $\mathbb{Z}$-affine on the unbounded edges of $\mathrm{Sk}(\mathscr{C}, \delta)$ is a consequence of (2).

(2) Let x be a closed point of C. We can compute the slope of wt_ω on the path running from $\mathrm{Sk}(\mathscr{C})$ to x as follows. Denote by E the unique irreducible component of $\mathscr{C}_k$ containing the specialization x_k of x. Denote by N the multiplicity of E in $\mathscr{C}_k$ and by ν the multiplicity of E in $\mathrm{div}_{\mathscr{C}}(\omega)$. Let $h : \mathscr{C}' \to \mathscr{C}$ be the blow-up at x_k. Then $\mathscr{C}'$ is again an *snc*-model of (C, δ) and its skeleton $\mathrm{Sk}(\mathscr{C}')$ is obtained from $\mathrm{Sk}(\mathscr{C})$ by adding a closed interval I in the direction of x. The length of this interval is $1/N^2$, since the exceptional component E' of the blow-up has multiplicity N in $\mathscr{C}'_k$. Moreover, the multiplicity of E' in $\mathrm{div}_{\mathscr{C}'}(\omega)$ is equal to

$$\nu + m + \deg_x(\mathrm{div}_C(\omega))),$$

because

$$\omega^{\otimes m}_{\mathscr{C}'/R}(\mathscr{C}'_{k,\mathrm{red}}) = (h^*\omega^{\otimes m}_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})) \otimes \mathcal{O}_{\mathscr{C}'}(mE')$$

as submodules of the pushforward of $\omega^{\otimes m}_{C/K}$ to $\mathscr{C}'$. Thus if we denote by v and v' the vertices of $\mathrm{Sk}(\mathscr{C}')$ corresponding to E and E', respectively, then $\mathrm{wt}_\omega(v) = \nu/N$ and

$$\mathrm{wt}_\omega(v') = (m + \nu + \deg_x(\mathrm{div}_C(\omega)))/N.$$

Since v and v' are precisely the endpoints of I, we see that wt_ω has slope

$$N(m + \deg_x(\mathrm{div}_C(\omega)))$$

on I if we orient I from v to v'. Replacing $\mathscr{C}$ by $\mathscr{C}'$ and repeating the argument, we conclude that wt_ω has constant slope

$$N(m + \deg_x(\mathrm{div}_C(\omega)))$$

along the whole path from v to x.

(3) It remains to compute the Laplacian $\Delta(\mathrm{wt}_\omega)$ of wt_ω on $\mathrm{Sk}(\mathscr{C}, \delta)$. Let v_0 be a vertex of $\mathrm{Sk}(\mathscr{C})$ corresponding to an irreducible component E_0 of $\mathscr{C}_k$. Denote by $x_1, \ldots, x_a$ the points in the support of δ that specialize to a point in E_0, and by $y_1, \ldots, y_b$ the intersection points of E_0 with the other irreducible components of $\mathscr{C}_k$.

For each $i \in \{1, \dots, b\}$ we denote by E_i the unique irreducible component of $\mathscr{C}_k$ intersecting E_0 at y_i. For each $i \in \{0, \dots, b\}$ we write v_i and N_i for the multiplicities of E_i in $\mathrm{div}_{\mathscr{C}}(\omega)$ and $\mathscr{C}_k$, respectively. Then the edges of $\mathrm{Sk}(\mathscr{C})$ adjacent to v_0 correspond precisely to the points $y_1, \dots, y_b$, and the unbounded edges of $\mathrm{Sk}(\mathscr{C}, \delta)$ adjacent to v_0 are precisely the paths from v_0 to the points $x_1, \dots, x_a$. We have already computed the slopes of wt_ω along these unbounded edges, and taking into account the edge lengths of $\mathrm{Sk}(\mathscr{C})$ we find that the degree of $\Delta(\mathrm{wt}_\omega)$ at v_0 is equal to

$$\sum_{i=1}^{a}(N_0 m + \deg_{x_i}(\mathrm{div}_C(\omega))) + \sum_{j=1}^{b}(v_j N_0 - v_0 N_j).$$

This is nothing but

$$m N_0 a + N_0(E_0 \cdot (\mathrm{div}_{\mathscr{C}}(\omega) - \frac{v_0}{N_0}\mathscr{C}_k)) = m N_0 a + N_0(E_0 \cdot \mathrm{div}_{\mathscr{C}}(\omega)).$$

By adjunction, the restriction of the line bundle $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ to E_0 is precisely

$$\omega_{E_0/k}(y_1 + \cdots + y_b)^{\otimes m}.$$

By computing the degree of this line bundle we find that the degree of $\Delta(\mathrm{wt}_\omega)$ at v_0 is equal to

$$m N_0(a + b + 2g(E_0) - 2),$$

where $g(E_0)$ denotes the genus of E_0. By definition, this is exactly the degree of the m-canonical divisor of $\mathrm{Sk}(\mathscr{C}, \delta)$ at v_0. □

3.2.4. We can use Theorem 3.2.3 to describe the Laplacian of the restriction of the weight function to the skeleton of *any snc*-model $\mathscr{C}$ of C. Beware that the weight function is not necessarily affine on the edges of $\mathrm{Sk}(\mathscr{C})$, only piecewise affine. The Laplacian of such a function is still defined, but it is no longer supported on the vertices of $\mathrm{Sk}(\mathscr{C})$, in general. We denote by $(\rho_{\mathscr{C}})_*$ the map on divisors induced by linearity from the retraction map $\rho_{\mathscr{C}} : C^{\mathrm{an}} \to \mathrm{Sk}(\mathscr{C})$. We have

$$(\rho_{\mathscr{C}})_*(x) = \deg(x) \cdot \rho_{\mathscr{C}}(x)$$

for every type I point x of C^{an}.

Corollary 3.2.5. *Let $\mathscr{C}$ be any snc-model of C. We denote by f the restriction of* wt_ω *to* $\mathrm{Sk}(\mathscr{C})$ *and by $mK_{\mathrm{Sk}(\mathscr{C})}$ the m-canonical divisor of* $\mathrm{Sk}(\mathscr{C})$, *where we label each vertex of* $\mathrm{Sk}(\mathscr{C})$ *by its multiplicity and genus as before. Then*

$$\Delta(f) = m K_{\mathrm{Sk}(\mathscr{C})} - (\rho_{\mathscr{C}})_*(\mathrm{div}_C(\omega)).$$

In particular, if ω is regular, then $\Delta(f) \leq m K_{\mathrm{Sk}(\mathscr{C})}$.

Proof. We can always dominate $\mathscr{C}$ by an *snc*-model $\mathscr{C}'$ of the pair $(C, \mathrm{div}_C(\omega))$. If we denote by f' the restriction of wt_ω to $\mathrm{Sk}(\mathscr{C}')$, then it follows easily from Theorem 3.2.3 that

$$\Delta(f') = m K_{\mathrm{Sk}(\mathscr{C}')} - (\rho_{\mathscr{C}'})_*(\mathrm{div}_C(\omega)).$$

Denote by $\rho : \mathrm{Sk}(\mathscr{C}') \to \mathrm{Sk}(\mathscr{C})$ the map of metric graphs obtained by restricting $\rho_{\mathscr{C}}$ to $\mathrm{Sk}(\mathscr{C}')$. Since the fibers of ρ are metric trees, it is straightforward to check that $\Delta(f) = \rho_*(\Delta(f'))$. On the other hand, we also have that

$$\rho_*(\rho_{\mathscr{C}'})_*(\mathrm{div}_C(\omega)) = (\rho_{\mathscr{C}})_*(\mathrm{div}_C(\omega)),$$

and by factoring $\mathscr{C}' \to \mathscr{C}$ into point blow-ups, one sees that $\rho_*(K_{\mathrm{Sk}(\mathscr{C}')}) = K_{\mathrm{Sk}(\mathscr{C})}$. Thus the formula is valid for $\mathscr{C}$, as well. □

Example 3.2.6. Let C be an elliptic curve over K of Kodaira–Néron reduction type II (see [Silverman 1994, IV§8]) and let ω be a generator for the relative canonical line bundle of the minimal regular model of C. Let $\mathscr{C}$ be the minimal *snc*-model of C. Then the special fiber of $\mathscr{C}$ is of the form

$$\mathscr{C}_k = E_1 + 2E_2 + 3E_3 + 6E_4,$$

where each component E_i is a rational curve, E_4 intersects each other component in precisely one point, and there are no other intersection points. The skeleton $\mathrm{Sk}(\mathscr{C})$ consists of four vertices $v_1, \ldots, v_4$ corresponding to the components $E_1, \ldots, E_4$. These are joined by three edges of respective lengths $\ell(v_1v_4) = 1/6$, $\ell(v_2v_4) = 1/12$ and $\ell(v_3v_4) = 1/18$. Moreover,

$$\mathrm{div}_{\mathscr{C}}(\omega) = E_1 + 2E_2 + 3E_3 + 5E_4$$

and the weight function wt_ω is affine on $\mathrm{Sk}(\mathscr{C})$ with values 1, 1, 1, 5/6 at the vertices v_1, v_2, v_3, v_4, respectively. Direct computation shows that

$$\Delta\,\mathrm{wt}_\omega = 6v_4 - v_1 - 2v_2 - 3v_3,$$

which is also the canonical divisor of $\mathrm{Sk}(\mathscr{C})$ (labeled with multiplicities and genera).

Remark 3.2.7. It is worth noting that Corollary 3.2.5 uniquely determines wt_ω up to an additive constant as a function on $\mathbb{H}_0(C)$, and that this description of wt_ω does not require K to be discretely valued (if we replace *snc*-models by semistable models). This gives us a way to define wt_ω for any curve C over any nontrivially valued nonarchimedean field K and any nonzero rational m-canonical form ω on C. M. Temkin [2014] has recently discovered a different way to extend the definition of wt_ω to the nondiscretely valued setting, and his method works in any dimension.

3.3. *The essential skeleton.*

3.3.1. Let C be a K-curve of genus $g(C) \geq 1$ and let ω be a nonzero regular m-canonical form on C, for some positive integer m. Then it is easy to deduce from the properties of the weight function wt_ω in Theorem 3.2.3 that this function is bounded below, and that its locus of minimal values is a union of closed faces of $\mathrm{Sk}(\mathscr{C})$ for any *snc*-model $\mathscr{C}$ of $(C, \mathrm{div}_C(\omega))$. Corollary 3.2.5 shows that this remains true for any *snc*-model $\mathscr{C}$ of C (one needs to observe that the weight function is *concave* on every edge of $\mathrm{Sk}(\mathscr{C})$ because its Laplacian is nonpositive at each point

in the interior of an edge), see [Mustaţă and Nicaise 2015, Theorem 4.7.5] for a more general statement. The locus of minimal values of wt_ω was called the Kontsevich–Soibelman skeleton of the pair (C, ω) in [Mustaţă and Nicaise 2015, 4.7.1] and denoted by $\mathrm{Sk}(C, \omega)$. The essential skeleton $\mathrm{Sk}(C)$ is the union of the Kontsevich–Soibelman skeleta $\mathrm{Sk}(C, \omega)$ over all the nonzero regular pluricanonical forms ω on C, see [Mustaţă and Nicaise 2015, Definition 4.10]. The aim of this section is to compare the essential skeleton $\mathrm{Sk}(C)$ to the skeleton $\mathrm{Sk}(\mathscr{C})$ of the minimal *snc*-model $\mathscr{C}$ of C.

3.3.2. We first recall the description of $\mathrm{Sk}(C, \omega)$, the Kontsevich–Soibelman skeleton, in terms of an *snc*-model $\mathscr{C}$ of C (see [Mustaţă and Nicaise 2015, Theorem 4.7.5] for a more general result; in our setting, it can also be easily deduced from Proposition 3.1.3 and Theorem 3.2.3). We write

$$\mathscr{C}_k = \sum_{i \in I} N_i E_i$$

and we denote by ν_i the multiplicity of E_i in $\mathrm{div}_{\mathscr{C}}(\omega)$, for every $i \in I$. We say that the vertex of $\mathrm{Sk}(\mathscr{C})$ corresponding to a component E_j, $j \in I$, is *ω-essential* if

$$\frac{\nu_j}{N_j} = \min\left\{ \frac{\nu_i}{N_i} \;\middle|\; i \in I \right\}.$$

We say that an edge in $\mathrm{Sk}(\mathscr{C})$ is ω-essential if its adjacent vertices are ω-essential and the point of $\mathscr{C}_k$ corresponding to the edge is not contained in the closure of $\mathrm{div}_C(\omega)$ (i.e., the horizontal part of $\mathrm{div}_{\mathscr{C}}(\omega)$). Then $\mathrm{Sk}(C, \omega)$ is the union of the ω-essential faces of $\mathrm{Sk}(\mathscr{C})$. Note however that, by its very definition, $\mathrm{Sk}(C, \omega)$ does not depend on the choice of a particular model $\mathscr{C}$.

3.3.3. In order to determine the essential skeleton $\mathrm{Sk}(C)$, we will need a description of the base locus of the logarithmic pluricanonical bundle on the minimal *snc*-model of C. Let $\mathscr{C}$ be any *snc*-model of C. We label the vertices of the skeleton $\mathrm{Sk}(\mathscr{C})$ by the multiplicities and genera of the corresponding irreducible components of $\mathscr{C}_k$. We define a *tail* in $\mathrm{Sk}(\mathscr{C})$ as a connected subchain with successive vertices $v_0, \ldots, v_n$ where v_n has valency one in $\mathrm{Sk}(\mathscr{C})$, v_i has valency 2 in $\mathrm{Sk}(\mathscr{C})$ for $1 \leq i < n$, and v_i has genus zero for $1 \leq i \leq n$. We say that the tail is *maximal* if v_0 has valency at least 3 in $\mathrm{Sk}(\mathscr{C})$ or v_0 has positive genus. The vertex v_0 is called the starting point of the maximal tail and v_n is called its end point. We call the components of $\mathscr{C}_k$ corresponding to the vertices $v_1, \ldots, v_n$ *inessential* components of $\mathscr{C}_k$. The *combinatorial skeleton* of $\mathrm{Sk}(\mathscr{C})$ is the subspace that we obtain by replacing every maximal tail by its starting point. Thus in Example 3.2.6, the combinatorial skeleton of $\mathrm{Sk}(\mathscr{C})$ consists only of the vertex v_4. Note that contracting maximal tails may create new ones, but we do *not* repeat the operation to contract those. For instance, if C is an elliptic K-curve of reduction type I_n^* (see [Silverman 1994, IV§8]) and $\mathscr{C}$ is its minimal *snc*-model, then the combinatorial skeleton of $\mathrm{Sk}(\mathscr{C})$ is the

subchain formed by the $n+1$ vertices of multiplicity two. Note that $\mathscr{C}_k$ can never consist entirely of inessential components, by our assumption that $g(C) \geq 1$ (this follows from basic intersection theory and adjunction, see for instance [Nicaise 2013, Lemma 3.1.2]). We also observe that, if C has semistable reduction and $\mathscr{C}$ is its minimal *snc*-model, there are no inessential components in $\mathscr{C}_k$ because the end point of a tail would correspond to a rational (-1)-curve, which contradicts the minimality of $\mathscr{C}$.

3.3.4. We will need a technical lemma on two-dimensional log regular schemes. We refer to [Kato 1989] for the basic theory of log schemes, and to [Kato 1994] for the theory of log regular schemes.

Lemma 3.3.5. *Let A be a normal Noetherian local ring of dimension 2 and let $D = D_0 + D_1$ be a reduced Weil divisor on $X = \operatorname{Spec} A$ with prime components D_0 and D_1. We define a log scheme X^+ by endowing X with the divisorial log structure induced by D. Assume that X^+ is log regular. Then D_0 and D_1 are $\mathbb{Q}$-Cartier, and $D_0 \cdot D_1 \leq 1$ with equality if and only if A is regular.*

Proof. We denote by $\mathcal{M}$ the multiplicative monoid consisting of the elements of A that are invertible on $X \setminus D$, and we consider the characteristic monoid

$$\overline{\mathcal{M}} = \mathcal{M}/A^{\times}.$$

By the log regularity assumption, D_0 and D_1 are regular and $\overline{\mathcal{M}}$ is a toric monoid of dimension 2. In particular, its groupification $\overline{\mathcal{M}}^{\mathrm{gp}}$ is a rank two lattice. The ring A is regular if and only if the monoid $\overline{\mathcal{M}}$ is generated by two elements, that is, $\overline{\mathcal{M}} \cong \mathbb{N}^2$.

Let e_0 and e_1 be the primitive generators of the one-dimensional faces of $\overline{\mathcal{M}}$. Then $e_0 \bigwedge e_1$ generates

$$m \cdot \bigwedge^2 (\overline{\mathcal{M}}^{\mathrm{gp}}),$$

for a unique positive integer m (in other words, m is the absolute value of the determinant of (e_0, e_1)), and $m = 1$ if and only if A is regular. Since the fan of the log scheme $\operatorname{Spec} A$ is canonically isomorphic with $\operatorname{Spec} \overline{\mathcal{M}}$ by [Kato 1994, Proposition 10.1], we know that (up to renumbering), D_i is the zero locus in $\operatorname{Spec} A$ of the prime ideal $\overline{\mathcal{M}} \setminus \mathbb{N}e_i$ of $\overline{\mathcal{M}}$, for $i = 0, 1$ (by which we mean the zero locus of its inverse image in $\mathcal{M}$). Moreover, any representative $\widetilde{e}_i$ of e_i in $\mathcal{M}$ is a regular local parameter on D_i, and the characteristic monoid at the generic point of D_i is $\overline{\mathcal{M}}/\mathbb{N}e_{1-i}$, see the proof of [Eriksson et al. 2015, Proposition 4.3.2(1)] for a similar computation. It follows that $mD_i = \operatorname{div}(\widetilde{e}_{1-i})$, so that D_1 and D_2 are $\mathbb{Q}$-Cartier, and $mD_1 \cdot D_2 = 1$. Thus

$$D_1 \cdot D_2 = 1/m \leq 1,$$

with equality if and only if A is regular. □

Theorem 3.3.6. *Let C be a K-curve of genus $g(C) \geq 1$ and let $\mathscr{C}$ be its minimal snc-model. If m is a sufficiently divisible positive integer, then the base locus of the line bundle $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ on $\mathscr{C}$ is the union of the inessential components of $\mathscr{C}_k$.*

Proof. Let m be a positive integer. Using adjunction, one sees that the line bundle $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ has negative degree or degree 0 on each rational curve in $\mathscr{C}_k$ that intersects the other components in precisely one point or two points, respectively. It follows at once that the union of the inessential components in $\mathscr{C}_k$ is contained in the base locus of $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$. We will show there are no other points in the base locus if m is sufficiently divisible. If $\mathscr{C}_{k,\mathrm{red}}$ is either an elliptic curve or a loop of rational curves, then C has genus one and $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ is trivial for some $m > 0$ by [Liu et al. 2004, Lemma 5.7 and Theorem 6.6]. Thus we can discard these cases in the remainder of the proof.

We can choose a reduced divisor H on $\mathscr{C}$ with the following properties:

- The divisor H does not contain any prime component of $\mathscr{C}_k$ (in other words, H is horizontal) and $H + \mathscr{C}_k$ is a divisor with strict normal crossings.
- We have $H \cdot E = 1$ if E is a prime component of $\mathscr{C}_k$ that corresponds to the end point of a maximal tail in $\mathrm{Sk}(\mathscr{C})$, and $H \cdot E = 0$ for every other prime component of $\mathscr{C}_k$.

We denote by S^+ the scheme $S = \operatorname{Spec} R$ endowed with its standard log structure (the divisorial log structure induced by the closed point of S) and by $\mathscr{C}^+$ the log scheme we obtain by endowing $\mathscr{C}$ with the divisorial log structure associated with the divisor $\mathscr{C}_k + H$. Then $\mathscr{C}^+$ is log regular in the sense of [Kato 1994] because $\mathscr{C}_k + H$ has strict normal crossings.

By Lipman's generalization [1969, Theorem 27.1] of Artin's contractibility criterion, any chain of rational curves in $\mathscr{C}_k$ can be contracted to a rational singularity. In particular, there exists a morphism $h : \mathscr{C} \to \mathscr{D}$ of normal proper R-models of C that contracts precisely the rational components of $\mathscr{C}_k$ that meet the rest of the special fiber in exactly one or two points. We endow $\mathscr{D}$ with the divisorial log structure associated with $\mathscr{D}_k + h_* H$ and denote the resulting log scheme by $\mathscr{D}^+$. It follows from [Ito and Schröer 2015, §3] that $\mathscr{D}^+$ is still log regular (this is the reason why we added the horizontal divisor H). The morphism h induces a morphism of log schemes $h : \mathscr{C}^+ \to \mathscr{D}^+$, and this morphism is log étale since it is a composition of log blow-ups.

We consider the canonical line bundle

$$\omega_{\mathscr{C}^+/S^+} = \det \Omega^1_{\mathscr{C}^+/S^+}$$

on $\mathscr{C}$. It follows easily from [Eriksson et al. 2015, Proposition 3.3.4] that $\omega_{\mathscr{C}^+/S^+}$ is isomorphic to $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}} + H)$. We can copy the proofs of [Eriksson et al. 2015,

3.3.2 and Proposition 3.3.6] to show that the coherent sheaf $\Omega^1_{\mathscr{D}^+/S^+}$ on $\mathscr{D}$ is perfect, so that we can define the canonical line bundle

$$\omega_{\mathscr{D}^+/S^+} = \det \Omega^1_{\mathscr{D}^+/S^+}$$

on $\mathscr{D}$ (the results in [Eriksson et al. 2015] were formulated for $H = 0$ but the arguments carry over immediately). Since h is log étale, [Eriksson et al. 2015, Proposition 3.3.6] also implies that we have a canonical isomorphism $\omega_{\mathscr{C}^+/S^+} \cong h^*\omega_{\mathscr{D}^+/S^+}$.

By Lemma 3.3.5, the divisor h_*H is $\mathbb{Q}$-Cartier. Thus by choosing m sufficiently divisible, we can assume that mh_*H is Cartier. We will prove that the line bundle $\omega^{\otimes m}_{\mathscr{D}^+/S^+}(-mh_*H)$ on $\mathscr{D}$ is ample. This implies that its pullback to $\mathscr{C}$ is semiample (that is, some tensor power is generated by its global sections). But this pullback is isomorphic to $\omega^{\otimes m}_{\mathscr{C}^+/S^+}(-h^*h_*mH)$, which is a subbundle of

$$\omega^{\otimes m}_{\mathscr{C}^+/S^+}(-mH) \cong \omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$$

that coincides with $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ away from the inessential components of $\mathscr{C}_k$ (note that, for every closed point x of h_*H, the inverse image $h^{-1}(x)$ is a maximal tail of inessential components in $\mathscr{C}_k$).

Thus it is enough to show that $\omega^{\otimes m}_{\mathscr{D}^+/S^+}(-mh_*H)$ is ample. By [Liu 2002, Chapter 7, Proposition 5.5], it suffices to show that it has positive degree on every prime component E of $\mathscr{D}_k$. By adjunction, the restriction of $\omega_{\mathscr{D}^+/S^+}$ to E is isomorphic to $\omega_{E/k}(F)$ where F is the reduced divisor on E supported on the intersection points of E with the other components of $\mathscr{D}_k + h_*H$. Note that either E has positive genus, or F consists of at least three points including at least two intersections points of E with the other components of $\mathscr{D}_k$, since we contracted all the other components in $\mathscr{C}_k$. Therefore, we only need to show that $h_*H_0 \cdot E < 1$ for every prime component H_0 of H. This follows from Lemma 3.3.5 (note that $\mathscr{D}$ is singular at every point of $h_*H \bigcap \mathscr{D}_k$ by minimality of $\mathscr{C}$). □

Remark 3.3.7. In the language of [Nicaise and Xu 2013], the proof of Theorem 3.3.6 can also be interpreted as follows: The model $\mathscr{C}'$ for C that we obtain from the minimal *snc*-model $\mathscr{C}$ by contracting all the inessential components in the special fiber is a minimal *dlt*-model of C. Even for curves, minimal *dlt*-models are not unique, because we can construct a new one by blowing up an intersection point of two components in the special fiber (in the language of the minimal model program, the minimality of a *dlt*-model only expresses that the logarithmic relative canonical line bundle is semiample). However, the set of isomorphism classes of minimal *dlt*-models has a unique minimal element with respect to the dominance relation (defined as in 2.2.1), and this is precisely the isomorphism class of $\mathscr{C}'$. Beware that such a unique minimal isomorphism class need no longer exist if we replace C by a K-variety of dimension ≥ 2.

3.3.8. If C has semistable reduction, we can be more precise; we will show in Theorem 3.3.11 that the logarithmic 2-canonical line bundle on the minimal *snc*-model of C is generated by global sections. This follows at once from Theorem 7 in [Lee 2005], which states that $\omega_{\mathscr{C}_{\min}/R}^{\otimes m}$ is generated by global sections if $m \geq 2$ and $\mathscr{C}_{\min}$ is the minimal regular model of a curve C of genus $g \geq 2$ (recall that the minimal regular model of a curve with semistable reduction coincides with its minimal *nc*-model). Although we only deal with the semistable case, we feel that our alternative proof of Theorem 3.3.11 is still interesting, because it uses a different method and it is substantially simpler than the proof of the more general result in [Lee 2005]. It does not seem possible to deduce Theorem 3.3.6 from the semiampleness of $\omega_{\mathscr{C}_{\min}/R}$ in a direct way, because of the discrepancy between the minimal regular model and the minimal *nc*-model of C if C does not have semistable reduction. We start by proving two elementary lemmas.

Lemma 3.3.9. *Let $\mathscr{X}$ be a regular flat proper R-scheme of relative dimension one and let $\mathcal{L}$ be a line bundle on $\mathscr{X}$. Let E be an irreducible component of multiplicity N in $\mathscr{X}_k$ and let a be an integer in $\{1, \ldots, N\}$. If the restriction of $\mathcal{L}((1-a)E)$ to E has negative degree, then*

$$H^0(aE, \mathcal{L}|_{aE}) = 0.$$

Proof. We prove this by induction on a. The case $a = 1$ is obvious. Assume that $a > 1$ and that the property holds for $a-1$. If $\mathcal{L}((1-a)E)$ has negative degree on E then the same holds for $\mathcal{L}((b-a)E)$ for all $b \geq 1$ because $E^2 \leq 0$. We consider the short exact sequence

$$0 \longrightarrow \mathcal{L}|_{aE} \otimes \mathcal{I} \longrightarrow \mathcal{L}|_{aE} \longrightarrow \mathcal{L}|_{(a-1)E} \longrightarrow 0,$$

where $\mathcal{I}$ is the ideal sheaf of $(a-1)E$ in aE. By our induction hypothesis, it suffices to show that

$$H^0(aE, \mathcal{L}|_{aE} \otimes \mathcal{I}) = 0.$$

This follows from the isomorphism of $\mathcal{O}_{aE}$-modules

$$\mathcal{L}|_{aE} \otimes \mathcal{I} \cong \mathcal{L}((1-a)E)|_E$$

on E. □

Lemma 3.3.10. *Let $\mathscr{X}$ be a regular flat proper R-scheme of relative dimension one and let $\mathcal{L}$ be a line bundle on $\mathscr{X}$. Let D be a reduced connected divisor supported on $\mathscr{X}_k$. Suppose that the restriction of $\mathcal{L}$ to each component in D has nonpositive degree, and that this degree is negative for at least one component. Then*

$$H^0(D, \mathcal{L}|_D) = 0.$$

Proof. This follows easily by induction on the number r of irreducible components of D. If $r = 1$ the result is obvious. Suppose that $r > 1$ and let E be a component of D on which $\mathcal{L}$ has negative degree. Then every section of $\mathcal{L}$ on D vanishes on E,

so it is also a section of $\mathcal{L}(-E)$ on D. The line bundle $\mathcal{L}(-E)$ has negative degree on each irreducible component of D that intersects E, so we can apply the induction hypothesis to this line bundle and to every connected component of $D - E$. □

Theorem 3.3.11. *Let C be a K-curve of genus $g(C) \geq 1$ with semistable reduction and let $\mathscr{C}$ be its minimal snc-model. Then the logarithmic 2-canonical line bundle $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes 2}$ on $\mathscr{C}$ is generated by its global sections.*

Proof. It will be convenient to start from the minimal *nc*-model $\mathscr{C}'$ of C instead of the minimal *snc*-model $\mathscr{C}$. We will show that $\omega_{\mathscr{C}'/R}^{\otimes 2}$ is generated by global sections. This implies the desired result; the line bundle $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})$ is isomorphic to the pullback of $\omega_{\mathscr{C}'/R} \cong \omega_{\mathscr{C}'/R}(\mathscr{C}'_k)$ through the morphism $g : \mathscr{C}' \to \mathscr{C}$, since $\mathscr{C}'_k$ is reduced and $\mathscr{C}$ is a composition of log blow-ups if we endow both models with the divisorial log structure associated with their special fibers. We can assume that C has genus at least 2, since otherwise $\omega_{\mathscr{C}'/R}$ is trivial.

Let x be a closed point of $\mathscr{C}'_k$ and denote by $h : \mathscr{D} \to \mathscr{C}'$ the blow-up of $\mathscr{C}'$ at x and by E_0 the exceptional curve of h. Then $\omega_{\mathscr{C}'/R}^{\otimes 2}$ is globally generated at x if and only if the morphism

$$H^0(\mathscr{D}, h^*\omega_{\mathscr{C}'/R}^{\otimes 2}) \to H^0(E_0, h^*\omega_{\mathscr{C}'/R}^{\otimes 2}|_{E_0})$$

is surjective. To prove surjectivity, it suffices to show that $H^1(\mathscr{D}, h^*\omega_{\mathscr{C}'/R}^{\otimes 2}(-E_0))$ vanishes. By Serre duality, this is equivalent to showing that $H^0(\mathscr{D}_k, \mathcal{L}) = 0$, with

$$\mathcal{L} = (\omega_{\mathscr{D}/R} \otimes (h^*\omega_{\mathscr{C}'/R}^{-2})(E_0))|_{\mathscr{D}_k} \cong (\omega_{\mathscr{D}/R}^{-1}(3E_0))|_{\mathscr{D}_k}.$$

We write

$$\mathscr{D}_k = N_0 E_0 + \sum_{i=1}^{r} E_i,$$

where N_0 is one or two, depending on whether x is a regular or singular point of $\mathscr{C}'_k$.

We first observe that the restriction of $\mathcal{L}$ to N_0E_0 has no nonzero global sections. Because the restriction of the line bundle $\mathcal{L}((1 - N_0)E_0)$ to E_0 has negative degree we can apply Lemma 3.3.9. Thus every section of $\mathcal{L}$ on $\mathscr{D}_k$ is also a section of

$$\mathcal{L}' = (\omega_{\mathscr{D}/R}^{-1}((3 - N_0)E_0))|_{\mathscr{D}_k}.$$

Note that $\mathcal{L}'$ has degree -1 on E_0 if $N_0 = 1$ and degree 0 if $N_0 = 2$. Next, we consider any component $E_i \neq E_0$ in $\mathscr{D}_k$. By the adjunction formula, the degree of $\mathcal{L}'$ on E_i is given by

$$\deg(\mathcal{L}'|_{E_i}) = 2 - 2p_a(E_i) + E_i^2 + (3 - N_0)E_0 \cdot E_i. \tag{3.3.12}$$

By the projection formula, $E_i^2 = h(E_i)^2 - \delta$, where

- $\delta = 0$ if x does not lie on $h(E_i)$,
- $\delta = 1$ if x is a regular point of $h(E_i)$ (then $E_0 \cdot E_i = 1$),
- $\delta = 4$ if x is a self-intersection point of $h(E_i)$ (then $N_0 = 2$ and $E_i \cdot E_0 = 2$).

Thus if $E_i^2+(3-N_0)E_0\cdot E_i$ is positive, we have $\delta=1$ and $h(E_i)^2=0$, which means that $\mathscr{C}'_k=h(E_i)$ and $p_a(E_i)=p_a(h(E_i))\geq 2$ by our assumption on the genus of C. Note also that $E_i^2+(3-N_0)E_0\cdot E_i=0$ implies that $\delta=1$ and $h(E_i)^2=-1$, and thus $p_a(E_i)=p_a(h(E_i))>0$, since otherwise $h(E_i)$ would be an exceptional curve on $\mathscr{C}'$, contradicting minimality.

It follows that the number in (3.3.12) is negative, unless

- $p_a(E_i)=1$ and $E_i^2+(3-N_0)E_0\cdot E_i=0$, or
- $p_a(E_i)=0$ and $E_i^2+(3-N_0)E_0\cdot E_i\in\{-2,-1\}$.

If $p_a(E_i)=1$ and $E_i^2+(3-N_0)E_0\cdot E_i=0$ then $h(E_i)$ is a (-1)-curve of arithmetic genus one and x is a point on $h(E_i)$ that does not lie on any other component of $\mathscr{C}'_k$. Similarly, if $p_a(E_i)=0$ and $E_i^2+(3-N_0)E_0\cdot E_i=-1$ then $h(E_i)$ must be a regular rational (-2)-curve and x is a point on $h(E_i)$ that does not lie on any other component of $\mathscr{C}'_k$. Finally, if $p_a(E_i)=0$ and $E_i^2+(3-N_0)E_0\cdot E_i=-2$ then $h(E_i)$ contains x or $h(E_i)$ is a regular rational curve of self-intersection number -2.

From these observations, we can deduce the following properties:

- The divisor $\mathscr{D}_k$ contains at most one component E_i on which $\mathcal{L}'$ has positive degree. In that case, this degree equals one, and $h(E_i)$ is a regular rational (-2)-curve and it is the only component of $\mathscr{C}'_k$ that contains x. Then each connected component of $\mathscr{D}_k-N_0E_0-E_i$ contains a curve on which $\mathcal{L}'$ has negative degree, since such a component cannot consist entirely of regular rational (-2)-curves. It follows from Lemma 3.3.10 that $\mathcal{L}'$ has no nonzero global sections on $\mathscr{D}_k-N_0E_0-E_i$. Then $\mathcal{L}$ has no nonzero global sections on $\mathscr{D}_k$, because every section vanishes at the two intersection points of E_i with $\mathscr{D}_k-N_0E_0$.
- Assume that $\mathcal{L}'$ has nonpositive degree on every component of $\mathscr{D}_k$. The divisor $\mathscr{D}_k$ contains at least one component E_i on which $\mathcal{L}'$ has negative degree, if x lies on only one component of $\mathscr{C}'_k$ then we can take $E_i=E_0$. In the other case, all components of $\mathscr{D}_k$ on which $\mathcal{L}'$ has degree zero are regular rational curves that intersect the rest of $\mathscr{D}_k$ in precisely two points, and $\mathscr{D}_k$ cannot consist entirely of such curves because of our assumption that $g(C)\geq 2$. Thus Lemma 3.3.10 again implies that $\mathcal{L}'$ has no nonzero global sections on $\mathscr{D}_{k,\mathrm{red}}$, so that $H^0(\mathscr{D}_k,\mathcal{L})=0$.

This concludes the proof. □

We are now ready to compare the essential skeleton of a K-curve C of positive genus to the Berkovich skeleton $\mathrm{Sk}(\mathscr{C})$ of its minimal *snc*-model $\mathscr{C}$. Recall from 3.3.3 that the *combinatorial skeleton* of $\mathrm{Sk}(\mathscr{C})$ is the subspace that we obtain by replacing every maximal tail by its starting point.

Theorem 3.3.13. *Let C be a K-curve of genus $g(C)\geq 1$ and let $\mathscr{C}$ be its minimal snc-model.*

(1) *The essential skeleton* $\mathrm{Sk}(C)$ *is equal to the combinatorial skeleton of* $\mathrm{Sk}(\mathscr{C})$ *(as a subspace of* C^{an}*). In particular,* $\mathrm{Sk}(C)$ *is a strong deformation retract of* C^{an}.

(2) *If* C *has semistable reduction, then* $\mathrm{Sk}(C) = \mathrm{Sk}(\mathscr{C})$. *Moreover,*

$$\mathrm{Sk}(C) = \bigcup_{\omega} \mathrm{Sk}(C, \omega),$$

where ω *runs through the set of nonzero regular* 2*-canonical forms on* C.

Proof. (1) It follows from Corollary 3.2.5 that, for every nonzero regular pluricanonical form ω on C, the weight function wt_ω is strictly increasing along every tail of $\mathrm{Sk}(\mathscr{C})$ if we orient the tail from its starting point to its end point. Thus the essential skeleton $\mathrm{Sk}(C)$ is contained in the combinatorial skeleton of $\mathrm{Sk}(\mathscr{C})$. The converse inclusion is a consequence of Theorem 3.3.6; we choose a positive integer m such that the base locus of $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ is the union of inessential components of $\mathscr{C}_k$. If x is a singular point of $\mathscr{C}_{k,\mathrm{red}}$ that does not lie on an inessential component and ω is a global section of $\omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}$ that does not vanish at x, then the weight function wt_ω vanishes on the edge of $\mathrm{Sk}(\mathscr{C})$ and it is nonnegative on the whole skeleton $\mathrm{Sk}(\mathscr{C})$, so that the edge belongs to $\mathrm{Sk}(C, \omega)$. Thus the combinatorial skeleton of $\mathrm{Sk}(\mathscr{C})$ is equal to

$$\bigcup_{\omega} \mathrm{Sk}(C, \omega),$$

where ω runs through any basis of the R-module

$$H^0(\mathscr{C}, \omega_{\mathscr{C}/R}(\mathscr{C}_{k,\mathrm{red}})^{\otimes m}).$$

(2) As we have already observed in 3.3.3, the special fiber of $\mathscr{C}$ does not contain any inessential components. Therefore, the combinatorial skeleton of $\mathrm{Sk}(\mathscr{C})$ is equal to $\mathrm{Sk}(\mathscr{C})$ and thus also to the essential skeleton $\mathrm{Sk}(C)$ by point (1). The proof of (1), together with Theorem 3.3.11, shows that 2-canonical forms ω suffices to generate the whole essential skeleton $\mathrm{Sk}(C)$. □

3.4. *The subset of the essential skeleton cut out by canonical forms.*

3.4.1. Let C be a K-curve of genus $g \geq 1$ and denote by $\mathscr{C}$ its minimal *snc*-model. We assume that $\mathscr{C}_k$ is reduced. Looking at the definition of the essential skeleton in 3.3.1, it is natural to ask which part of the essential skeleton we recover by taking the union of the Kontsevich–Soibelman skeleta $\mathrm{Sk}(C, \omega)$ where ω runs through the set of nonzero canonical (rather than pluricanonical) forms on C. In this section, we will show that one obtains the union of all the closed *nonbridge edges* and all the vertices of positive genus of the skeleton $\mathrm{Sk}(\mathscr{C})$. Recall that a *bridge* in a graph G is an edge that is not contained in any nontrivial cycle, or equivalently, that is contained in every spanning tree.

3.4.2. Let $G = G(\mathscr{C}_k)$ be the dual graph of the special fiber $\mathscr{C}_k$, and let $\nu : \widetilde{\mathscr{C}}_k \to \mathscr{C}_k$ be a normalization morphism. The set $V(G)$ of vertices v of G is in bijection with the set of connected components C_v of $\widetilde{\mathscr{C}}_k$. Let $E(G)$ denote the set of edges of G, each endowed with a fixed (but arbitrary) orientation. If x is the singular point of $\mathscr{C}_k$ corresponding to an edge e, then the choice of an orientation on e amounts to choosing a point y in $\nu^{-1}(x)$; if the oriented edge $\vec{e}$ points towards the vertex v, then we take y to be the unique point of $\nu^{-1}(x)$ lying on C_v.

3.4.3. By the cohomological flatness of $\mathscr{C} \to \mathrm{Spec}(R)$ and Grothendieck–Serre duality, the module $H^1(\mathscr{C}, \omega_{\mathscr{C}/R})$ is free, so that

$$H^0(\mathscr{C}_k, \omega_{\mathscr{C}/R}) \otimes k \cong H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k}).$$

We can identify $H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k})$ with the space of *Rosenlicht differentials* on $\mathscr{C}_k$. A Rosenlicht differential ω is, by definition, the data of a meromorphic differential ω_v on C_v for each $v \in V(G)$ such that:

(1) Each ω_v has at worst logarithmic poles at the inverse images under ν of the singular points of $\mathscr{C}_k$, and is regular everywhere else.

(2) If x is a singular point of $\mathscr{C}_k$ and $\nu^{-1}(x) = \{y_1, y_2\}$, then the residues of ω at y_1 and y_2 sum to zero.

Given $\omega \in H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k})$ and an oriented edge $\vec{e} \in E(G)$, let $\mathrm{res}_{\vec{e}}(\omega)$ be the residue of ω at the point of $\widetilde{\mathscr{C}}_k$ corresponding to $\vec{e}$. By the residue theorem, the sum

$$\mathrm{res}(\omega) := \sum_{e \in E(G)} \mathrm{res}_{\vec{e}}(\omega)(\vec{e})$$

belongs to $H_1(G, k)$, so that we obtain a morphism of k-vector spaces

$$\mathrm{res} : H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k}) \to H_1(G, k),$$

which is called the residue map.

Lemma 3.4.4. *The residue map fits into a short exact sequence of k-vector spaces*:

$$0 \longrightarrow \oplus_{v \in V(G)} H^0(C_v, \omega_{C_v/k}) \xrightarrow{\alpha} H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k}) \xrightarrow{\mathrm{res}} H_1(G, k) \longrightarrow 0.$$

Proof. By the definition of Rosenlicht differentials and the residue map, the kernel of res is equal to $\bigoplus_{v \in V(G)} H^0(C_v, \omega_{C_v/k})$. Surjectivity of the residue map now follows by a dimension count, since

$$\dim_k H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k}) = \dim_k H_1(G, k) + \sum_{v \in V(G)} \dim_k H^0(C_v, \omega_{C_v/k}) = g. \qquad \square$$

Lemma 3.4.5. *Let ω be a regular canonical form on C and let v be a vertex of genus zero of* $\mathrm{Sk}(\mathscr{C})$. *Then v belongs to* $\mathrm{Sk}(C, \omega)$ *if and only if some edge adjacent to v belongs to* $\mathrm{Sk}(C, \omega)$.

Proof. The "if" part follows from the fact that $\mathrm{Sk}(C,\omega)$ is closed, so we only need to prove the converse implication. We denote by f the restriction of wt_ω to $\mathrm{Sk}(\mathscr{C})$. Assume that v lies in $\mathrm{Sk}(C,\omega)$, that is, f reaches its minimal value at v. By Corollary 3.2.5 and the assumption that v has genus zero, the degree of $\Delta(f)$ at v is strictly less than the valency of v in $\mathrm{Sk}(\mathscr{C})$. Since f has integer slopes, this means that at least one of the outgoing slopes of f from v must be zero, so that the corresponding edge also lies in $\mathrm{Sk}(C,\omega)$. □

Theorem 3.4.6. *If C is a K-curve of genus $g \geq 1$ whose minimal snc-model $\mathscr{C}$ over R is semistable, then the union $S = \bigcup_\omega \mathrm{Sk}(C,\omega)$, as ω runs through the set of nonzero global sections of $\omega_{C/K}$, is equal to the union of all the closed nonbridge edges and all the vertices of positive genus of* $\mathrm{Sk}(\mathscr{C})$.

Proof. Multiplying a nonzero canonical form ω with $a \in K^\times$ shifts the weight function wt_ω by $v_K(a)$ and does not affect $\mathrm{Sk}(C,\omega)$. Moreover, since $\mathscr{C}_k$ is reduced, wt_ω takes integer values at the vertices of $\mathrm{Sk}(\mathscr{C})$. Thus in the definition of S, we only need to consider canonical forms ω whose minimal value on $\mathrm{Sk}(\mathscr{C})$ equals zero (recall from 3.3.1 that this minimal value is always reached at a vertex).

Now it is clear from the definition of the weight function that wt_ω vanishes at an edge, or vertex, of $\mathrm{Sk}(\mathscr{C})$ if and only if ω generates $\omega_{\mathscr{C}/R}(\mathscr{C}_k)$ at the corresponding point of $\mathscr{C}_k$ or at the generic point of the corresponding irreducible component of $\mathscr{C}_k$, respectively. Thus, in order to find the faces of $\mathrm{Sk}(\mathscr{C})$ that lie in S, we need to determine which singular points and irreducible components of $\mathscr{C}_k$ lie in the base locus of

$$\omega_{\mathscr{C}/R}(\mathscr{C}_k) \cong \omega_{\mathscr{C}/R}.$$

For this aim, we can use Rosenlicht differentials; a point of $\mathscr{C}_k$ lies in the base locus of $\omega_{\mathscr{C}/R}$ if and only if it lies in the base locus of $\omega_{\mathscr{C}_k/k}$ on $\mathscr{C}_k$, by the surjectivity of the reduction map

$$H^0(\mathscr{C}, \omega_{\mathscr{C}/R}) \to H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k}).$$

Using the morphism α in Lemma 3.4.4, we can find an element of $H^0(\mathscr{C}_k, \omega_{\mathscr{C}_k/k})$ that generates $\omega_{\mathscr{C}_k/k}$ at the generic point of every component of positive genus of $\mathscr{C}_k$. In particular, all the vertices of positive genus of $\mathrm{Sk}(\mathscr{C})$ belong to S. By Lemma 3.4.5, it now suffices to determine which edges of $\mathrm{Sk}(\mathscr{C})$ lie in S. The residue theorem immediately implies that a bridge never belongs to S, while the surjectivity of the residue map in Lemma 3.4.4 shows that every nonbridge edge lies in S. This concludes the proof. □

Remark 3.4.7. By Theorem 3.3.13, the essential skeleton $\mathrm{Sk}(C)$ is always connected, but it is easy to use the proof of Theorem 3.4.6 to produce examples of a curve C and a nonzero canonical form ω such that $\mathrm{Sk}(C,\omega)$ is disconnected (for instance, when $\mathrm{Sk}(\mathscr{C})$ is a chain with vertices of positive genus).

3.5. *An alternate approach to computing the essential skeleton of a maximally degenerate semistable curve.*

3.5.1. Let C be a K-curve of genus $g \geq 1$ and denote by $\mathscr{C}$ its minimal *snc*-model. There is an elegant way to prove Theorems 3.3.13(2) and 3.4.6 using potential theory on metric graphs if we assume that C is a *maximally degenerate* K-curve. This assumption is common in tropical geometry; it means that $\mathscr{C}_k$ is reduced and that all the irreducible components of $\mathscr{C}_k$ are rational curves. This implies that the metric graph $\mathrm{Sk}(\mathscr{C})$ still has genus g. The proofs yield some additional information about the structure of $\mathrm{Sk}(C, \omega)$ for certain explicit 2-canonical forms ω. They also have the advantage that they can be extended to the nondiscretely valued setting (see Remark 3.2.7).

3.5.2. For background on potential theory on metric graphs, see for instance [Baker 2008]. We recall that a *tropical rational function* on a metric graph Γ is a real-valued continuous piecewise affine function on Γ with integral slopes, and that the divisor of such a function is defined by $\mathrm{div}(f) = -\Delta(f)$. In other words, the degree of $\mathrm{div}(f)$ at a point of Γ is the sum of the *incoming* slopes of f. Two divisors on Γ are called *equivalent* if they differ by the divisor of a tropical rational function. We begin with a combinatorial lemma needed for our alternate proof of Theorem 3.4.6.

Lemma 3.5.3. *Let G be a discrete graph without loops and denote by Γ the metric graph associated with G. Let T be a spanning tree of Γ, let e be an edge of Γ not contained in T, and let $Z(T, e)$ be the unique cycle in $T \cup e$. Let D be an effective divisor on Γ which is equivalent to the canonical divisor K_G and whose support contains a point p_i from the relative interior of each edge $e_i \neq e$ contained in the complement of T. Finally, let f be a tropical rational function on Γ with $\mathrm{div}(f) = D - K_G$. Then the locus of points $p \in \Gamma$ where f achieves its minimum value is equal to $Z(T, e)$.*

Proof. Let $\mathrm{Sk}(f)$ be the locus of $p \in \Gamma$ at which f attains its minimum value. For each $p \in \mathrm{Sk}(f)$, f can be strictly increasing in at most $\mathrm{val}(p) - 2$ tangent directions, since it has slope at least 1 in each such direction and nonnegative slope in every other direction and the total sum of outgoing slopes of f at p is at most

$$\deg_p K_G = \mathrm{val}(p) - 2.$$

Thus there are at least two tangent directions at p along which f is constant. It follows that every connected component of $\mathrm{Sk}(f)$ is a graph in which every vertex has valency at least 2. However, $\mathrm{Sk}(f)$ cannot contain any of the points p_i, since the sum of the outgoing slopes of f at p_i is equal to $-\deg_{p_i} D < 0$. Thus $\mathrm{Sk}(f) \subset T \cup e$, and the only possible cycle in $\mathrm{Sk}(f)$ is $Z(T, e)$. Hence, $\mathrm{Sk}(f) = Z(T, e)$. □

We obtain the following strengthening of Theorem 3.4.6 in this context (it can also be deduced directly from Lemma 3.4.4 and the proof of Theorem 3.4.6):

Proposition 3.5.4. *Assume that C is maximally degenerate. If e is a nonbridge edge of $\mathrm{Sk}(\mathscr{C})$, then there exists a nonzero canonical form $\omega \in H^0(C, \omega_{C/K})$ such that $\mathrm{Sk}(C, \omega)$ is a simple cycle with e in its support.*

Proof. Since e is not a bridge, there exists a spanning tree T of $\mathrm{Sk}(\mathscr{C})$ not containing e. Let $e = e_0, e_1, \ldots, e_{g-1}$ be the edges of $\mathrm{Sk}(\mathscr{C})$ not contained in T, and choose a type II point p_i in the relative interior of e_i for every i in $\{1, \ldots, g-1\}$ (type II points are the divisorial points in the terminology of [Mustaţă and Nicaise 2015]). We set $D_0 = p_1 + \cdots + p_{g-1}$. We would like to find a divisor $\widetilde{D}_0$ on C such that $(\rho_{\mathscr{C}})_*(\widetilde{D}_0) = D_0$. Unfortunately, this is not possible, since only the vertices of $\mathrm{Sk}(\mathscr{C})$ lift to K-rational points of C.

This issue can be solved in the following way. Let K' be a finite Galois extension of K whose degree $n = [K' : K]$ is not divisible by the characteristic of k. We denote by R' the valuation ring of K'. Set $C' = C \times_K K'$ and let $\mathscr{C}'$ be the minimal resolution of $\mathscr{C} \times_R R'$. Then it is well known, and easy to see, that $\mathscr{C}'$ is the minimal *snc*-model of C', and $\mathscr{C}'_k$ is reduced. Moreover, the projection morphism $\pi : (C')^{\mathrm{an}} \to C^{\mathrm{an}}$ induces a homeomorphism $\mathrm{Sk}(\mathscr{C}') = \pi^{-1}(\mathrm{Sk}(\mathscr{C})) \to \mathrm{Sk}(\mathscr{C})$, and $\mathrm{Sk}(\mathscr{C}')$ is obtained from $\mathrm{Sk}(\mathscr{C})$ by subdividing each edge into n edges. Now we choose each point p_i to be a vertex of $\mathrm{Sk}(\mathscr{C}')$ in the relative interior of e_i. Then we can find a divisor $\widetilde{D}_0$ on C' such that $(\rho_{\mathscr{C}'})_*(\widetilde{D}_0) = D_0$.

Since $H^0(C', \omega_{C'/K'})$ has dimension g and $\widetilde{D}_0$ has degree $g-1$, there exists a nonzero $\omega' \in H^0(C', \omega_{C'/K'}(-\widetilde{D}_0))$. Let f be the restriction of $\mathrm{wt}_{\omega'}$ to $\mathrm{Sk}(\mathscr{C}')$. By Corollary 3.2.5, we have

$$\mathrm{div}(f) = (\rho_{\mathscr{C}'})_*(\mathrm{div}_{C'}(\omega')) - K_{\mathrm{Sk}(\mathscr{C}')}.$$

If we set $D = (\rho_{\mathscr{C}'})_*(\mathrm{div}_{C'}(\omega'))$, then $D \geq D_0$ by construction. Now it follows from Lemma 3.5.3 that $\mathrm{Sk}(C', \omega') = \mathrm{Sk}(f)$ is a simple cycle that contains e.

It remains to produce a nonzero element ω of $H^0(C, \omega_{C/K})$ such that $\mathrm{Sk}(C, \omega) = \mathrm{Sk}(C', \omega')$. Multiplying ω' with a suitable element of $(K')^\times$, we can assume that the minimal value of $\mathrm{wt}_{\omega'}$ on $\mathrm{Sk}(\mathscr{C}')$ is equal to 0. We denote by $\omega \in H^0(C, \omega_{C/K})$ the trace of ω' with respect to the Galois extension K'/K. Then it is easy to see that $\mathrm{wt}_\omega = \mathrm{wt}_{\omega \otimes_K K'} \geq \mathrm{wt}_{\omega'}$ on $\mathrm{Sk}(\mathscr{C})$. It is also clear that every singular point x of $\mathscr{C}'_k$ is fixed under the action of $\mathrm{Gal}(K'/K)$. Thus the logarithmic residues at x of the conjugates of ω' are all equal, and their sum is nonzero if and only if the logarithmic residue of ω' at x is nonzero. It follows that an edge of $\mathrm{Sk}(\mathscr{C}')$ lies in the zero locus of $\mathrm{wt}_{\omega'}$ if and only if it lies in the zero locus of wt_ω. Since $\mathrm{Sk}(C', \omega')$ is a union of edges, it follows that

$$\mathrm{Sk}(C, \omega) = \mathrm{Sk}(C', \omega'). \qquad \square$$

Remark 3.5.5. The statement and proof of both Lemma 3.4.5 and Proposition 3.5.4 are closely related to Lemma 3.2 and Proposition 3.3, respectively, of [Jensen and Payne 2016].

We now show that if e is a bridge edge of $\mathrm{Sk}(\mathscr{C})$, then there exists a 2-canonical form ω such that $\mathrm{Sk}(C,\omega)$ contains e, providing a new proof of Theorem 3.3.13(2) in the present context.

Lemma 3.5.6. *Let G be a discrete graph without loops and denote by Γ the metric graph associated with G. We assume that G has no 1-valent vertices. Choose any maximal chain B of bridge edges in Γ. We denote by v_1, v_2 the endpoints of B. Let T be a spanning tree in Γ. Let D be an effective divisor on Γ equivalent to $2K_G$ satisfying the following properties*:

(1) *The support of D contains a point from the relative interior of each edge contained in the complement of T.*

(2) $D \geq K_G - (v_1) - (v_2)$.

Finally, let f be a tropical rational function on Γ with $\mathrm{div}(f) = D - 2K_G$. Then the locus of points $p \in \Gamma$ where f achieves its minimum value is equal to B.

Proof. Let $\mathrm{Sk}(f)$ be the locus of $p \in \Gamma$ at which f attains its minimum value. We can argue in the same way as in the proof of Lemma 3.5.3. By condition (2), for each $p \neq v_1, v_2$ in $\mathrm{Sk}(f)$ there are at least two tangent directions at p along which f is constant, and if $p \in \{v_1, v_2\}$ there is at least one such direction. Thus every connected component of $\mathrm{Sk}(f)$ is a graph in which every vertex different from v_1, v_2 has valency at least two, and it cannot be equal to $\{v_1\}$ or $\{v_2\}$. On the other hand, by condition (1), the set $\mathrm{Sk}(f)$ cannot contain any cycles. It follows that $\mathrm{Sk}(f) = B$. □

Proposition 3.5.7. *Assume that C is maximally degenerate. Let B be any maximal chain of bridge edges of $\mathrm{Sk}(\mathscr{C})$. Then there exists a nonzero 2-canonical form $\omega \in H^0(C, \omega_{C/K}^{\otimes 2})$ such that $\mathrm{Sk}(C,\omega) = B$.*

Proof. We can assume that $g \geq 2$ since in the genus one case $\mathrm{Sk}(\mathscr{C})$ is a cycle and does not contain any bridges. Since $\mathscr{C}$ is the minimal *snc*-model of C and $\mathscr{C}_k$ is reduced, $\mathrm{Sk}(\mathscr{C})$ has no 1-valent vertices. We set $\Gamma = \mathrm{Sk}(\mathscr{C})$. We choose a spanning tree T of Γ. We define K', C', and $\mathscr{C}'$ as in the proof of Proposition 3.5.4. Then, by the same arguments as in that proof, it suffices to find a nonzero element $\omega' \in H^0(C', \omega_{C'/K'}^{\otimes 2})$ such that $\mathrm{Sk}(C',\omega') = B$. We can find an effective divisor $\widetilde{D}_0$ on C' of degree $3g-4 = g+(2g-4)$ over K' such that $D_0 = (\rho_{\mathscr{C}'})_*(\widetilde{D}_0)$ satisfies properties (1) and (2) from the statement of Lemma 3.5.6. Since the space $H^0(C', \omega_{C'/K'}^{\otimes 2})$ has dimension $3g-3$ by Riemann–Roch, there exists a nonzero 2-canonical form $\omega' \in H^0(C, \omega_{C'/K'}^{\otimes 2})$ with $\mathrm{div}_{C'}(\omega') \geq \widetilde{D}_0$. We set $D = (\rho_{\mathscr{C}'})_*(\mathrm{div}_{C'}(\omega'))$. Then $D \geq D_0$ by construction. Let f be the restriction of $\mathrm{wt}_{\omega'}$ to Γ. By Theorem 3.2.3, we have

$$\mathrm{div}(f) = D - 2K_\Gamma.$$

The result now follows from Lemma 3.5.6. □

Appendix: The stable metric on $\mathbb{H}_0(C)$

A.1. *Definition of the stable metric.*

A.1.1. Let C be a K-curve. The metric on $\mathbb{H}_0(C)$ defined in Theorem 2.3.3 was well-suited for the description of the Laplacian of the weight function in Theorem 3.2.3, but it does not behave well under extensions of the base field K. We will now define an alternative metric on $\mathbb{H}_0(C)$, which we call the *stable* metric, which has better properties with respect to base change. In particular, if k has characteristic zero, one can compare it to the skeletal metric from [Baker et al. 2013] (see Proposition A.2.3).

A.1.2. We first put a metric on the geometric realization Γ of a weighted discrete graph (G, w) by replacing the formula in (2.1.3) by

$$\ell(e) = \frac{1}{\mathrm{lcm}\{w(v_1), w(v_2)\}}.$$

Now the same arguments as in Section 2.3 show that this definition induces a unique metric on $\mathbb{H}_0(C)$ such that, for every *snc*-model $\mathscr{C}$ of C, the embedding

$$\Gamma(\mathscr{C}_k) \to \mathbb{H}_0(C)$$

is an isometry onto $\mathrm{Sk}(\mathscr{C})$. We call this metric the stable metric on $\mathbb{H}_0(C)$. Note that, if $\mathscr{C}_k$ is reduced, the stable metric on $\mathrm{Sk}(\mathscr{C})$ coincides with the one defined in Theorem 2.3.3.

A.1.3. By [Mustaţă and Nicaise 2015, §3.2], the skeleton $\mathrm{Sk}(\mathscr{C})$ of an *nc*-model $\mathscr{C}$ of C carries a natural $\mathbb{Z}$-affine structure. If e is an edge of $\mathrm{Sk}(\mathscr{C})$ with endpoints v_1 and v_2, then a $\mathbb{Z}$-affine function

$$f : e \setminus \{v_1, v_2\} \to \mathbb{R}$$

is a function of the form

$$(x_1, x_2) \mapsto ax_1/N_1 + bx_2/N_2 + c,$$

where a, b, c are integers, $N_1 = w(v_1)$, $N_2 = w(v_2)$, and x_1 and $x_2 = 1 - x_1$ are barycentric coordinates on $e \setminus \{v_1, v_2\} \cong]0, 1[$ such that the limit of x_1 at v_1 is 1 and the limit of x_2 at v_2 is 1 (beware that we are not excluding the possibility $v_1 = v_2$). This definition is motivated by the following fact: if $h \neq 0$ is a rational function on C, then

$$\mathrm{Sk}(\mathscr{C}) \to \mathbb{R}, \quad x \mapsto -\ln|h(x)|$$

is continuous and piecewise $\mathbb{Z}$-affine, and this function is affine on an edge e if and only if the point of $\mathscr{C}_k$ corresponding to e does not belong to the horizontal part of the divisor $\mathrm{div}_{\mathscr{C}}(h)$ on $\mathscr{C}$ (see [Mustaţă and Nicaise 2015, Proposition 3.2.2]). Moreover, if e is an edge of $\mathrm{Sk}(\mathscr{C})$ that is not a loop, then every $\mathbb{Z}$-affine function on $e \setminus \{v_1, v_2\}$ can be written as

$$x \mapsto -\ln|h(x)|,$$

for some rational function $h \neq 0$ on C (simply consider a monomial with suitable integer exponents in the local equations for the components corresponding to the vertices adjacent to e).

A.1.4. The $\mathbb{Z}$-affine structure on $\mathrm{Sk}(\mathscr{C})$ induces the stable metric on $\mathrm{Sk}(\mathscr{C}) = \Gamma(\mathscr{C}_k)$, in the following sense: the length of e is equal to

$$\inf_f \{|\lim_0 f - \lim_1 f|\},$$

where f runs through the set of injective $\mathbb{Z}$-affine functions

$$f : e \setminus \{v_1, v_2\} \to \mathbb{R},$$

and where $\lim_i f$ denotes the limit of f at i for $i = 0, 1$, where we choose any homeomorphism to identify $e \setminus \{v_1, v_2\}$ with the open interval $]0, 1[$.

To see this, note that this infimum is equal to the smallest positive element of the set

$$\{a/N_1 - b/N_2 \mid a, b \in \mathbb{Z}\},$$

which is precisely

$$\frac{\gcd(N_1, N_2)}{N_1 N_2} = \frac{1}{\mathrm{lcm}(N_1, N_2)}.$$

Thus our definition of the length of e is the unique one such that the affine functions on $e \setminus \{v_1, v_2\}$ are precisely the differentiable functions with constant integer slope whose value at v_1 is a multiple of $1/N_1$.

A.2. *Comparison with the skeletal metric.*

A.2.1. The set

$$\mathbb{H}_0(C \times_K \widehat{K^a}) = (C \times_K \widehat{K^a})^{\mathrm{an}} \setminus \{\text{points of type I and IV}\}$$

carries a natural metric, which was called the *skeletal metric* in [Baker et al. 2013]. Its construction is described in detail in [Baker et al. 2013, §5.3]. We will now compare it to the metric we defined on $\mathbb{H}_0(C)$, in the case where k has characteristic zero.

A.2.2. Let C be a K-curve and let $\mathscr{C}$ be an *snc*-model for C. An irreducible component of $\mathscr{C}_k$ is called *principal* if it has positive genus or it is a rational curve that intersects the rest of $\mathscr{C}_k$ in at least three points. A principal vertex of $\mathrm{Sk}(\mathscr{C})$ is a vertex corresponding to a principal component in $\mathscr{C}_k$.

Proposition A.2.3. *Assume that k has characteristic zero. Let C be a K-curve and let $\mathscr{C}$ be an snc-model of C. Denote by π the canonical projection $C \times_K \widehat{K^a} \to C$. Then the corestriction*

$$\pi_{\mathscr{C}} : \pi^{-1}(\mathrm{Sk}(\mathscr{C})) \to \mathrm{Sk}(\mathscr{C})$$

of π to $\mathrm{Sk}(\mathscr{C})$ is a local isometry over the complement of the principal vertex set of $\mathrm{Sk}(\mathscr{C})$. Moreover, if $\mathscr{C}$ is semistable, then $\pi_{\mathscr{C}}$ is an isometry.

Proof. This can be deduced in a rather straightforward way from the results in Sections 1 and 4 of Chapter 3 in [Halle and Nicaise 2012]. Since the arguments are somewhat tedious and the result is not needed in this paper, we omit the proof. □

Acknowledgements

The authors would like to thank Mattias Jonnson for helpful comments on an earlier version of this text. Nicaise was supported by the ERC Starting Grant MOTZETA (project 306610) of the European Research Council.

References

[Baker 2008] M. Baker, "Specialization of linear systems from curves to graphs", *Algebra Number Theory* **2**:6 (2008), 613–653. MR 2448666 Zbl 1162.14018

[Baker et al. 2013] M. Baker, S. Payne, and J. Rabinoff, "On the structure of non-Archimedean analytic curves", pp. 93–121 in *Tropical and non-Archimedean geometry*, edited by O. Amini et al., Contemp. Math. **605**, Amer. Math. Soc., Providence, 2013. MR 3204269 Zbl 1320.14040

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, 1990. MR 1070709 Zbl 0715.14013

[Eriksson et al. 2015] D. Eriksson, L. H. Halle, and J. Nicaise, "A logarithmic interpretation of Edixhoven's jumps for Jacobians", *Adv. Math.* **279** (2015), 532–574. MR 3345191 Zbl 06435582

[Gubler et al. 2016] W. Gubler, J. Rabinoff, and A. Werner, "Skeletons and tropicalizations", *Adv. Math.* **294** (2016), 150–215. MR 3479562 Zbl 06567870

[Halle and Nicaise 2012] L. H. Halle and J. Nicaise, "Néron models and base change", preprint, 2012. Zbl 06518678 arXiv 1209.5556

[Ito and Schröer 2015] H. Ito and S. Schröer, "Wild quotient surface singularities whose dual graphs are not star-shaped", *Asian J. Math.* **19**:5 (2015), 951–986. MR 3431685 Zbl 1333.14007

[Jensen and Payne 2016] D. Jensen and S. Payne, "Tropical independence, II: The maximal rank conjecture for quadrics", *Algebra Number Theory* **10**:8 (2016), 1601–1640. MR 3556794 Zbl 06640297

[Jonsson 2015] M. Jonsson, "Dynamics of Berkovich spaces in low dimensions", pp. 205–366 in *Berkovich spaces and applications*, edited by A. Ducros et al., Lecture Notes in Math. **2119**, Springer, 2015. MR 3330767 Zbl 06463429

[Kato 1989] K. Kato, "Logarithmic structures of Fontaine–Illusie", pp. 191–224 in *Algebraic analysis, geometry, and number theory* (Baltimore, MD, 1988), edited by J.-I. Igusa, Johns Hopkins Univ. Press, Baltimore, MD, 1989. MR 1463703 Zbl 0776.14004

[Kato 1994] K. Kato, "Toric singularities", *Amer. J. Math.* **116**:5 (1994), 1073–1099. MR 1296725 Zbl 0832.14002

[Lee 2005] J. Lee, "Relative canonical sheaves of a family of curves", *J. Algebra* **286**:2 (2005), 341–360. MR 2128021 Zbl 1073.14013

[Lipman 1969] J. Lipman, "Rational singularities, with applications to algebraic surfaces and unique factorization", *Inst. Hautes Études Sci. Publ. Math.* 36 (1969), 195–279. MR 0276239 Zbl 0181.48903

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxf. Grad. Texts Math. **6**, Oxford University Press, 2002. MR 1917232 Zbl 0996.14005

[Liu et al. 2004] Q. Liu, D. Lorenzini, and M. Raynaud, "Néron models, Lie algebras, and reduction of curves of genus one", *Invent. Math.* **157**:3 (2004), 455–518. MR 2092767 Zbl 1060.14037

[Mustaţă and Nicaise 2015] M. Mustaţă and J. Nicaise, "Weight functions on non-Archimedean analytic spaces and the Kontsevich–Soibelman skeleton", *Algebr. Geom.* **2**:3 (2015), 365–404. MR 3370127 Zbl 1322.14044

[Nicaise 2013] J. Nicaise, "Geometric criteria for tame ramification", *Math. Z.* **273**:3-4 (2013), 839–868. MR 3030680 Zbl 1329.11067

[Nicaise and Xu 2013] J. Nicaise and C. Xu, "The essential skeleton of a degeneration of algebraic varieties", preprint, 2013. To appear in Amer. J. Math. arXiv 1307.4041

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR 1312368 Zbl 0911.14015

[Temkin 2014] M. Temkin, "Metrization of differential pluriforms on Berkovich analytic spaces", preprint, 2014. arXiv 1410.3079

Communicated by Joseph H. Silverman
Received 2015-10-28 Revised 2016-06-24 Accepted 2016-10-02

mbaker@math.gatech.edu *School of Mathematics, Georgia Institute of Technology, 686 Cherry Street, Atlanta, GA 30332-0160, United States*

j.nicaise@imperial.ac.uk *Department of Mathematics, KU Leuven, Celestijnenlaan 200B, 3001 Heverlee, Belgium*

Current address: *Department of Mathematics, Imperial College, South Kensington Campus, London, SW7 2AZ, United Kingdom*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2081

Nonvanishing of Dirichlet L-functions

Rizwanur Khan and Hieu T. Ngo

We show that for at least $\frac{3}{8}$ of the primitive Dirichlet characters χ of large prime modulus, the central value $L\left(\frac{1}{2}, \chi\right)$ does not vanish.

1. Introduction

The zeros of L-functions on the critical line are as important in number theory as they are mysterious. At the real point on the critical line (the central point), an L-function is expected to vanish only for either a good reason or a trivial reason. A good reason is when the central value has some arithmetic significance which explains why it may vanish. For example, the central value of the L-function attached to an elliptic curve over a number field is expected to vanish if and only if the elliptic curve has positive rank (according to the Birch and Swinnerton-Dyer conjecture). A trivial reason is when the functional equation implies that the central value is zero. For instance, the L-function of any odd Hecke–Maass form f has functional equation $L\left(\frac{1}{2}, f\right) = -L\left(\frac{1}{2}, f\right)$ at the central point. In all other cases, the most extensive success in proving the nonvanishing of L-functions has been achieved through the use of mollifiers. For notable examples of the mollifier method, see [Kowalski et al. 2000a; 2000b; Iwaniec and Sarnak 2000; Soundararajan 2000] as well as the works discussed below.

In this paper, we study the classical nonvanishing problem of primitive Dirichlet L-functions. It is conjectured that $L\left(\frac{1}{2}, \chi\right) \neq 0$ for every primitive Dirichlet character χ. Consider for each odd prime p the family of L-functions

$$\{L(s, \chi) : \chi \text{ is primitive modulo } p\};$$

this family has size $p-2$. Viewing $L\left(\frac{1}{2}, \chi\right)$ as a statistical object, we would like to understand its distribution as $p \to \infty$. One way to get a handle on the distribution is through understanding the moments of $L\left(\frac{1}{2}, \chi\right)$, but currently only moments of small order are known. Nevertheless, this is enough to make some progress in proving that a positive proportion of the family is nonvanishing.

MSC2010: 11M20.
Keywords: L-functions, Dirichlet characters, nonvanishing central value, mollifier.

Asymptotic expressions for the first and second moments of $L\left(\frac{1}{2}, \chi\right)$ are well known. By a classical result of Paley [1931], we have

$$\frac{1}{p} \sideset{}{^\star}\sum_{\chi \bmod p} L\left(\tfrac{1}{2}, \chi\right) \sim 1,$$
$$\frac{1}{p} \sideset{}{^\star}\sum_{\chi \bmod p} \left|L\left(\tfrac{1}{2}, \chi\right)\right|^2 \sim \log p,$$

where $\sum^\star$ restricts the summation to the primitive characters. The discrepancy between the first and second moments indicates fluctuations in the sizes of the central values. Using these moments and the Cauchy–Schwarz inequality, one can only infer that at least 0% of the family is nonvanishing, since

$$\frac{1}{p} \sideset{}{^\star}\sum_{\substack{\chi \bmod p \\ L(1/2,\chi)\neq 0}} 1 \geq \frac{\left|\frac{1}{p}\sum^*_{\chi \bmod p} L\left(\frac{1}{2}, \chi\right)\right|^2}{\frac{1}{p}\sum^*_{\chi \bmod p}\left|L\left(\frac{1}{2}, \chi\right)\right|^2} \gg \frac{1}{\log p}.$$

The mollifier method is used to remedy this situation. The origin of the method traces back to the works of Bohr and Landau [1914] and Selberg [1942] on zeros of the Riemann zeta function. The starting idea is to introduce a quantity $M(\chi)$, called the "mollifier", which, on average, approximates the inverses of the supposedly nonvanishing values $L\left(\frac{1}{2}, \chi\right)$. The goal is to choose a mollifier such that the mollified first and second moments are comparable; that is,

$$\frac{1}{p} \sideset{}{^\star}\sum_{\chi \bmod p} L\left(\tfrac{1}{2}, \chi\right) M(\chi) \asymp 1,$$
$$\frac{1}{p} \sideset{}{^\star}\sum_{\chi \bmod p} \left|L\left(\tfrac{1}{2}, \chi\right) M(\chi)\right|^2 \asymp 1.$$

From this, a positive nonvanishing proportion can be inferred:

$$\frac{1}{p} \sideset{}{^\star}\sum_{\substack{\chi \bmod p \\ L(1/2,\chi)\neq 0}} 1 \geq \frac{1}{p} \sideset{}{^\star}\sum_{\substack{\chi \bmod p \\ L(1/2,\chi)M(\chi)\neq 0}} 1 \geq \frac{\left|\frac{1}{p}\sum^*_{\chi \bmod p} L\left(\frac{1}{2}, \chi\right) M(\chi)\right|^2}{\frac{1}{p}\sum^*_{\chi \bmod p}\left|L\left(\frac{1}{2}, \chi\right) M(\chi)\right|^2} \gg 1. \quad (1\text{-}1)$$

Balasubramanian and Murty [1992] were the first to do this; however, their mollifier was inefficient and they obtained only a very small positive proportion of nonvanishing.

Next came the work of Iwaniec and Sarnak [1999], who introduced a systematic technique that has since served as a model for other families of L-functions. Iwaniec

and Sarnak took the mollifier

$$M(\chi) = \sum_{m \le M} \frac{y_m \chi(m)}{m^{1/2}}, \tag{1-2}$$

where $M = p^\theta$ is the mollifier length and (y_m) is a sequence of real numbers satisfying $y_m \ll p^\epsilon$. They established the asymptotics of the mollified first and second moments for $\theta < \frac{1}{2}$ and found that the choice of coefficients which maximizes the ratio in (1-1) is essentially

$$y_m = \mu(m) \frac{\log(M/m)}{\log M}, \tag{1-3}$$

yielding a nonvanishing proportion of

$$\frac{1}{p} \sum_{\substack{\chi \bmod p \\ L(1/2,\chi) \neq 0}}^{\star} 1 \ge \frac{\theta}{1+\theta}.$$

This can be taken as close to $\frac{1}{3}$ as possible on letting θ approach $\frac{1}{2}$. Computing the mollified moments for larger values of θ would result in a higher proportion of nonvanishing, but this appears to be very difficult to do. The problem seems to have been attempted by Bettin, Chandee, and Radziwiłł. In [Bettin et al. 2015], these authors solved the parallel problem for the Riemann zeta function, by obtaining the asymptotics as $T \to \infty$ of

$$\int_T^{2T} \left|\zeta\left(\tfrac{1}{2}+it\right)\right|^2 \left| \sum_{m \le M} \frac{y_m}{m^{(1/2)+it}} \right|^2 dt,$$

where $M = T^\theta$, for values of θ slightly larger than $\frac{1}{2}$. However with regard to the problem for Dirichlet L-functions, the authors remarked, "Our proof would not extend to give an asymptotic formula in this case, and additional input is needed."

Shortly after the work of Iwaniec and Sarnak, in their study of the nonvanishing of high derivatives of Dirichlet L-functions, Michel and VanderKam [2000] used the "twisted" mollifier

$$M(\chi) = \sum_{m \le M} \frac{y_m \chi(m)}{m^{1/2}} + \frac{\bar{\tau}_\chi}{p^{1/2}} \sum_{m \le M} \frac{y_m \bar{\chi}(m)}{m^{1/2}}, \tag{1-4}$$

where $M = p^\theta$, y_m is as in (1-3), and τ_χ is the Gauss sum as defined in their paper. Heuristically, this is a better mimic of $L\left(\frac{1}{2}, \chi\right)^{-1}$ because the approximate functional equation of $L\left(\frac{1}{2}, \chi\right)$ essentially consists of a sum of two Dirichlet polynomials, one multiplied by a Gauss sum. A similar two-piece mollifier was first used by Soundararajan [1995] in the context of the Riemann zeta function. Michel and

VanderKam [2000] proved for $\theta < \frac{1}{4}$ a nonvanishing proportion of

$$\frac{1}{p} \sum_{\substack{\chi \bmod p \\ L(1/2,\chi)\neq 0}}^{\star} 1 \geq \frac{2\theta}{1+2\theta},$$

recovering the $\frac{1}{3}$ proportion of Iwaniec and Sarnak [1999]. For this method too, computing the mollified moments for larger θ would result in a higher proportion of nonvanishing.

The nonvanishing problem was stuck at the proportion $\frac{1}{3}$ for ten years until Bui [2012] dexterously proved a nonvanishing proportion of 0.3411. His breakthrough was not to increase the length of any existing mollifier but to use an ingenious new two-piece mollifier. Bui [2012, page 1857] commented that "There are two different approaches to improve the results in this and other problems involving mollifiers. One can either extend the length of the Dirichlet polynomial or use some 'better' mollifiers. The former is certainly much more difficult." We take the former, more difficult approach.

Our first idea to attack the nonvanishing problem is to increase the length of the Michel–VanderKam mollifier. This may be a somewhat unexpected avenue because previous attempts at lengthening mollifiers have, as far as we are aware, been directed at the Iwaniec–Sarnak mollifier. Our second idea is to establish an estimate for a trilinear sum of Kloosterman sums with general coefficients (Lemma 3.2). To prove this, we appeal to some work of Fouvry, Ganguly, Kowalski and Michel [Fouvry et al. 2014]. These authors proved best possible estimates for sums of products of Kloosterman sums to prime moduli by using powerful algebro-geometric methods (this work built on [Fouvry et al. 2004] and was later generalized in [Fouvry et al. 2015]). We stress that although the deepest part of our proof comes from [Fouvry et al. 2014], it is not clear how this work is related to the nonvanishing problem. We figure out this relationship.

Before stating our result, it should be said that the works [Iwaniec and Sarnak 1999; Michel and VanderKam 2000; Bui 2012] actually treat general moduli, while we are restricting to prime moduli, which is arguably the most interesting case.

Theorem 1.1. *Let $\epsilon > 0$ be arbitrary. For all primes p large enough in terms of ϵ, there are at least $\left(\frac{3}{8} - \epsilon\right)$ of the primitive Dirichlet characters $\chi \pmod p$ for which $L\left(\frac{1}{2}, \chi\right) \neq 0$.*

The significance of our work is that we show for the first time how to increase the length of a classical mollifier in this context. An interesting open problem that remains is to increase the length of the Iwaniec–Sarnak mollifier. Our nonvanishing proportion $\frac{3}{8}$ improves upon that of Bui for prime moduli. For general moduli, Bui's nonvanishing proportion 0.3411 is still the best known.

Throughout the paper, we use the standard convention that ϵ denotes an arbitrarily small positive constant which may differ from one occurrence to the next, and that the implied constants in the various estimates depend on ϵ.

2. The work of Michel and VanderKam

We briefly summarize the mollifier method of Michel and VanderKam [2000], setting the ground for our further discussion.

Let the mollifier $M(\chi)$ be given by (1-4), where the mollifier length is $M = p^\theta$ and the real mollifying coefficients y_m are given by (1-3). Michel and VanderKam asymptotically evaluated the mollified first moment

$$\frac{2}{p}\sum_{\chi \bmod p}^{+} L\left(\tfrac{1}{2},\chi\right)M(\chi)$$

for $\theta < \frac{1}{2}$, where $\sum^+$ restricts the summation to the even primitive characters, of which there are about $\frac{p}{2}$. The evaluation for the odd primitive characters is entirely similar. They evaluated the mollified second moment

$$\frac{2}{p}\sum_{\chi \bmod p}^{+} \left|L\left(\tfrac{1}{2},\chi\right)M(\chi)\right|^2 = \frac{4}{p}\sum_{\chi \bmod p}^{+} \left|L\left(\tfrac{1}{2},\chi\right)\right|^2 \left|\sum_{m\le M}\frac{y_m\chi(m)}{m^{1/2}}\right|^2 + \frac{4}{p}\sum_{\chi \bmod p}^{+} \left|L\left(\tfrac{1}{2},\chi\right)\right|^2 \frac{\tau_\chi}{p^{1/2}}\left(\sum_{m\le M}\frac{y_m\chi(m)}{m^{1/2}}\right)^2 \tag{2-1}$$

for $\theta < \frac{1}{4}$; see [Michel and VanderKam 2000, Equation (10)] for the above identity. An asymptotic for the first sum on the right-hand side of (2-1) is derived for $\theta < \frac{1}{2}$, as was done by Iwaniec and Sarnak [1999], but the second sum is more difficult and could only be handled for $\theta < \frac{1}{4}$. In the end, the main terms of the mollified moments of Michel and VanderKam yield a nonvanishing proportion of $2\theta/(1+2\theta)$, by taking $P_0(t) = t$ in [Michel and VanderKam 2000, Section 7].

Let us concentrate on the second sum on the right-hand side of (2-1). Recall the standard approximate functional equation (see for example [Michel and VanderKam 2000, Equation (3)]):

$$\left|L\left(\tfrac{1}{2},\chi\right)\right|^2 = 2\sum_{n_1,n_2\ge 1}\frac{\chi(n_1)\bar{\chi}(n_2)}{(n_1n_2)^{1/2}}V\left(\frac{n_1n_2}{p}\right), \tag{2-2}$$

where

$$V(x) = \frac{1}{2\pi i}\int_{(2)}\frac{\Gamma\left(\frac{s}{2}+\frac{1}{4}\right)^2}{\Gamma\left(\frac{1}{4}\right)^2}(\pi x)^{-s}\frac{ds}{s}.$$

By moving the line of integration, one shows that $V(x) \ll_c x^{-c}$ for any $c > 0$, whence the sum in (2-2) is essentially supported on $n_1 n_2 \le p^{1+\epsilon}$. Therefore,

$$\frac{4}{p}\sideset{}{^+}\sum_{\chi \bmod p} \left|L\left(\tfrac{1}{2}, \chi\right)\right|^2 \frac{\tau_\chi}{p^{1/2}}\left(\sum_{m\le M}\frac{y_m \chi(m)}{m^{1/2}}\right)^2$$
$$= \sum_{\substack{n_1,n_2\ge 1\\ m_1,m_2\le M}} \frac{y_{m_1}y_{m_2}}{(n_1n_2m_1m_2)^{1/2}} V\left(\frac{n_1n_2}{p}\right)\frac{4}{p}\sideset{}{^+}\sum_{\chi \bmod p}\frac{\tau_\chi}{p^{1/2}}\chi(n_1m_1m_2)\bar{\chi}(n_2). \quad (2\text{-}3)$$

By [Michel and VanderKam 2000, Equation (17)] or [Iwaniec and Sarnak 1999, Equation (3.4)], for $(n, p) = 1$ we have

$$\sideset{}{^+}\sum_{\chi \bmod p} \tau_\chi \chi(n) = p\cos\left(\frac{2\pi\bar{n}}{p}\right) + O(1),$$

so that (2-3) equals

$$\frac{4}{p^{1/2}}\,\mathrm{Re}\sum_{\substack{n_1,n_2\ge 1\\ m_1,m_2\le M\\ (n_1n_2m_1m_2,p)=1}} \frac{y_{m_1}y_{m_2}}{(n_1n_2m_1m_2)^{1/2}} V\left(\frac{n_1n_2}{p}\right) e\left(\frac{n_2\overline{n_1m_1m_2}}{p}\right) + O\left(\frac{M}{p^{1-\epsilon}}\right) \quad (2\text{-}4)$$

for any $\epsilon > 0$, where $e(x) = e^{2\pi i x}$ and $\bar{n}$ denotes the multiplicative inverse of n modulo p for $(n, p) = 1$. The terms with $m_1m_2 = 1$ contain a main term of (2-3); see [Michel and VanderKam 2000, Section 6]. Consider the rest of the terms in dyadic intervals. Let

$$\mathcal{B}(M_1, M_2, N_1, N_2) = \frac{1}{(pM_1M_2N_1N_2)^{1/2}}$$
$$\times \sum_{\substack{n_1,n_2\ge 1\\ M_1\le m_1\le 2M_1\\ M_2\le m_2\le 2M_2\\ (n_1n_2m_1m_2,p)=1}} y_{m_1}y_{m_2}\, e\left(\frac{n_2\overline{n_1m_1m_2}}{p}\right) V\left(\frac{n_1n_2}{p}\right) f_1\left(\frac{n_1}{N_1}\right) f_2\left(\frac{n_2}{N_2}\right) \quad (2\text{-}5)$$

for $1 \le M_1, M_2 \le \frac{1}{2}M$, $M_1M_2 \ge 2$, $1 \le N_1N_2 \le p^{1+\epsilon}$, and any fixed smooth functions f_1, f_2 compactly supported on the positive reals. Michel and VanderKam [2000, Equations (24) and (27)] proved the bounds

$$\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^\epsilon\left(\frac{M^2N_1}{pN_2}\right)^{1/2} \quad (2\text{-}6)$$

and

$$\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^\epsilon\left(\frac{M^2N_2}{N_1}\right)^{1/2} + \frac{M}{p^{1-\epsilon}}. \quad (2\text{-}7)$$

These bounds together yield $\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{-\epsilon}$, provided that $M \le p^{(1/4)-\epsilon}$. Thus the contribution to (2-4) of the terms with $m_1 m_2 \ge 2$ is $O(p^{-\epsilon})$ for $\theta < \frac{1}{4}$.

In the next section we will show how to improve the bound (2-7), in the ranges where (2-6) is not useful. This together with (2-6) will imply that

$$\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{-\epsilon}$$

for larger values of θ, thereby extending the asymptotics of Michel and VanderKam.

3. Proof of Theorem 1.1

To get the bounds (2-6) and (2-7), Michel and VanderKam obtained cancellation in only the (n_1, n_2)-sums of $\mathcal{B}(M_1, M_2, N_1, N_2)$. On the other hand, we use the (m_1, m_2)-sums to our advantage. To set this up, we first prove some estimates for averages of products of Kloosterman sums. Let

$$S(a, b; c) = \sum_{\substack{x \bmod c \\ x\bar{x} \equiv 1 \bmod c}} e\Big(\frac{ax + b\bar{x}}{c}\Big)$$

denote the Kloosterman sum. The following lemma is a consequence of a result from [Fouvry et al. 2014].

Lemma 3.1. *For $B < p$, we have*

$$\sum_{1 \le b_1, b_2, b_3, b_4 \le B} \Big| \sum_{h \bmod p} S(h, \bar{b}_1; p) S(h, \bar{b}_2; p) S(h, \bar{b}_3; p) S(h, \bar{b}_4; p) \Big| \ll B^4 p^{5/2} + B^2 p^3. \quad (3\text{-}1)$$

Proof. Write the left-hand side of (3-1) as

$$\sum_{b_1, b_2, b_3, b_4 \le B} = \sum_{\substack{b_1, b_2, b_3, b_4 \le B \\ (b_1, b_2, b_3, b_4) \in \mathfrak{D}}} + \sum_{\substack{b_1, b_2, b_3, b_4 \le B \\ (b_1, b_2, b_3, b_4) \notin \mathfrak{D}}}$$

where $\mathfrak{D}$ is the set of tuples (b_1, b_2, b_3, b_4) such that no component b_i is distinct from the others. Note that $|\mathfrak{D}| \ll B^2$.

On the one hand, it follows from the Weil bound for Kloosterman sums that

$$\sum_{\substack{b_1, b_2, b_3, b_4 \le B \\ (b_1, b_2, b_3, b_4) \in \mathfrak{D}}} \Big| \sum_{h \bmod p} S(h, \bar{b}_1; p) S(h, \bar{b}_2; p) S(h, \bar{b}_3; p) S(h, \bar{b}_4; p) \Big| \ll B^2 p^3.$$

On the other hand, if $(b_1, b_2, b_3, b_4) \notin \mathfrak{D}$, then in the language of [Fouvry et al. 2014, Definition 3.1], $(\bar{b}_1, \bar{b}_2, \bar{b}_3, \bar{b}_4)$ is not in "mirror configuration". Thus [Fouvry

et al. 2014, Proposition 3.2] asserts that

$$\sum_{h \bmod p} S(h, \bar{b}_1; p)S(h, \bar{b}_2; p)S(h, \bar{b}_3; p)S(h, \bar{b}_4; p) \ll p^{5/2},$$

saving a factor of $p^{1/2}$ over Weil's bound. So

$$\sum_{\substack{b_1,b_2,b_3,b_4 \le B \\ (b_1,b_2,b_3,b_4) \notin \mathfrak{D}}} \left| \sum_{h \bmod p} S(h, \bar{b}_1; p)S(h, \bar{b}_2; p)S(h, \bar{b}_3; p)S(h, \bar{b}_4; p) \right| \ll B^4 p^{5/2}.$$

The lemma follows. □

Let now

$$\mathcal{S} = \sum_{\substack{1 \le |n| \le N \\ 1 \le a \le A \\ 1 \le b \le B}} x_n y_a z_b S(n, \overline{ab}; p),$$

where the coefficients satisfy $x_n, y_a, z_b \ll p^{\epsilon}$, $y_a = 0$ for $p|a$, and $z_b = 0$ for $p|b$.

Lemma 3.2. *For $NA \le \frac{p}{2}$ and $B < p$, we have*

$$\mathcal{S} \ll p^{\epsilon} N^{3/4} A^{3/4} (Bp^{5/8} + B^{1/2} p^{3/4}).$$

Proof. On applying the Cauchy–Schwarz inequality, we infer

$$|\mathcal{S}|^2 \ll p^{\epsilon} NA \sum_{\substack{|n| \le N \\ a \le A}} \left| \sum_{b \le B} z_b S(n\bar{a}, \bar{b}; p) \right|^2.$$

Hence

$$|\mathcal{S}|^2 \ll p^{\epsilon} NA \sum_{h \bmod p} \nu(h) \left| \sum_{b \le B} z_b S(h, \bar{b}; p) \right|^2, \tag{3-2}$$

where

$$\nu(h) = \sum_{\substack{|n| \le N \\ a \le A \\ n\bar{a} \equiv h \bmod p}} 1.$$

On applying Cauchy–Schwarz to (3-2), we find that

$$|\mathcal{S}|^4 \ll p^{\epsilon} N^2 A^2 \Big(\sum_{h \bmod p} \nu(h)^2 \Big) \Big(\sum_{h \bmod p} \left| \sum_{b \le B} z_b S(h, \bar{b}; p) \right|^4 \Big). \tag{3-3}$$

Observe that

$$\sum_{h \bmod p} \nu(h)^2 = \sum_{\substack{|n_1|, |n_2| \le N \\ a_1, a_2 \le A \\ n_1\bar{a}_1 \equiv n_2\bar{a}_2 \bmod p}} 1 = \sum_{\substack{|n_1|, |n_2| \le N \\ a_1, a_2 \le A \\ n_1 a_2 \equiv n_2 a_1 \bmod p}} 1.$$

Since $NA \le \frac{p}{2}$ by assumption, it follows that

$$\sum_{h \bmod p} \nu(h)^2 = \sum_{\substack{n_1 a_2 = n_2 a_1 \\ |n_1|,|n_2| \le N \\ a_1, a_2 \le A}} 1 \ll p^{\epsilon} NA.$$

Therefore, (3-3) becomes

$$|\mathcal{S}|^4 \ll p^{\epsilon} N^3 A^3 \sum_{b_1,b_2,b_3,b_4 \le B} \left| \sum_{h \bmod p} S(h, \bar{b}_1; p) S(h, \bar{b}_2; p) S(h, \bar{b}_3; p) S(h, \bar{b}_4; p) \right|.$$

Finally, we apply Lemma 3.1 to conclude that

$$|\mathcal{S}|^4 \ll p^{\epsilon} N^3 A^3 (B^4 p^{5/2} + B^2 p^3),$$

and the lemma is proved. □

We are now in a position to prove a new bound for our nonvanishing problem.

Lemma 3.3. *For $N_1/N_2 > p^{\epsilon} M$ and $M < p^{1-\epsilon}$, we have*

$$\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{\epsilon} \left(\frac{N_2 M^3}{N_1 p^3} \right)^{1/4} \left(p^{5/8} + \frac{p^{3/4}}{M^{1/2}} \right) + \frac{M}{p^{1-\epsilon}}. \tag{3-4}$$

Proof. In (2-5), separate n_1 into residue classes modulo p and apply the Poisson summation formula to get

$$\begin{aligned} &\mathcal{B}(M_1, M_2, N_1, N_2) \\ &= \frac{1}{(pM_1M_2N_1N_2)^{1/2}} \frac{N_1}{p} \sum_{\substack{k \in \mathbb{Z} \\ n_2 \ge 1, (n_2,p)=1 \\ M_1 \le m_1 \le 2M_1 \\ M_2 \le m_2 \le 2M_2}} y_{m_1} y_{m_2} S(kn_2, \overline{m_1 m_2}; p) f_2\left(\frac{n_2}{N_2}\right) F(k), \end{aligned} \tag{3-5}$$

where

$$F(k) = \int_{-\infty}^{\infty} f_1(x) V\left(\frac{xN_1n_2}{p}\right) e\left(\frac{-xkN_1}{p}\right) dx.$$

Repeatedly integrating by parts, we find that $F(k) \ll_c (kN_1/p)^{-c}$ for any $c > 0$. Thus, the k-sum may be restricted to $|k| \le p^{1+\epsilon}/N_1$.

The contribution to (3-5) of the terms with $k = 0$ is

$$\begin{aligned} &\frac{1}{(pM_1M_2N_1N_2)^{1/2}} \frac{N_1}{p} \sum_{\substack{n_2 \ge 1, (n_2,p)=1 \\ M_1 \le m_1 \le 2M_1 \\ M_2 \le m_2 \le 2M_2}} y_{m_1} y_{m_2} S(0, \overline{m_1 m_2}; p) f_2\left(\frac{n_2}{N_2}\right) F(0) \\ &\qquad \ll \frac{(N_1N_2M_1M_2)^{1/2}}{p^{(3/2)-\epsilon}} \ll \frac{M}{p^{1-\epsilon}}, \end{aligned}$$

on using that the Ramanujan sum $S(0, \overline{m_1 m_2}; p)$ equals -1. This is the last term in (3-4). The contribution of the terms with $|k| > 0$ is bounded using Lemma 3.2, by putting

$$n = kn_2, \quad x_n = f_2(n_2/N_2)F(k) \text{ if } (n_2, p) = 1, \quad x_n = 0 \text{ if } p|n_2, \quad N = N_2 p^{1+\epsilon}/N_1,$$
$$a = m_1, \quad y_a = y_{m_1}, \quad A = 2M_1,$$
$$b = m_2, \quad z_b = y_{m_2}, \quad B = 2M_2.$$

Note that the conditions of Lemma 3.2, namely $B < p$ and $NA \le \frac{p}{2}$, are satisfied by the assumptions that $M < p^{1-\epsilon}$ and that $N_1/N_2 > p^{\epsilon} M$. The bound (3-4) follows. □

Finally, we sum up the work done to arrive at the following power-saving result.

Lemma 3.4. *We have* $\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{-\epsilon}$ *for* $M < p^{(3/10)-\epsilon}$.

Proof. Assume first that $M < p^{(1/3)-\epsilon}$. If $N_1/N_2 \le p^{\epsilon} M$, it follows from (2-6) that $\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{-\epsilon}$, whence the lemma follows.

We therefore suppose that $N_1/N_2 > p^{\epsilon} M$. Now since the conditions of Lemma 3.3 are met, we have the bound (3-4). In this bound, we may suppose $N_2/N_1 < M^2/p^{1-\epsilon}$, since otherwise by (2-6), we have $\mathcal{B}(M_1, M_2, N_1, N_2) \ll p^{-\epsilon}$. Thus, (3-4) becomes

$$\mathcal{B}(M_1, M_2, N_1, N_2) \ll \frac{M^{5/4}}{p^{1-\epsilon}} \left(p^{5/8} + \frac{p^{3/4}}{M^{1/2}} \right) + p^{-(1/6)+\epsilon}.$$

The bound is $O(p^{-\epsilon})$ precisely when $M \ll p^{(3/10)-\epsilon}$. The lemma follows. □

Proof of Theorem 1.1. By Lemma 3.4, the nonvanishing proportion $2\theta/(1+2\theta)$ of Michel and VanderKam is valid for any $\theta < \frac{3}{10}$. On letting θ approach $\frac{3}{10}$, we infer that the nonvanishing proportion is at least $\frac{3}{8} - \epsilon$ for any $\epsilon > 0$. □

References

[Balasubramanian and Murty 1992] R. Balasubramanian and V. K. Murty, "Zeros of Dirichlet L-functions", *Ann. Sci. École Norm. Sup.* (4) **25**:5 (1992), 567–615. MR 1191737 Zbl 0771.11033

[Bettin et al. 2015] S. Bettin, V. Chandee, and M. Radziwiłł, "The mean square of the product of the Riemann zeta-function with Dirichlet polynomials", *J. Reine Angew. Math.* (online publication February 2015).

[Bohr and Landau 1914] H. Bohr and E. Landau, "Sur les zéros de la fonction $\zeta(s)$ de Riemann", *C. R. Acad. Sci., Paris* **158** (1914), 106–110. Zbl 45.0716.02

[Bui 2012] H. M. Bui, "Non-vanishing of Dirichlet L-functions at the central point", *Int. J. Number Theory* **8**:8 (2012), 1855–1881. MR 2978845 Zbl 1292.11093

[Fouvry et al. 2004] É. Fouvry, P. Michel, J. Rivat, and A. Sárközy, "On the pseudorandomness of the signs of Kloosterman sums", *J. Aust. Math. Soc.* **77**:3 (2004), 425–436. MR 2099811 Zbl 1063.11023

[Fouvry et al. 2014] É. Fouvry, S. Ganguly, E. Kowalski, and P. Michel, "Gaussian distribution for the divisor function and Hecke eigenvalues in arithmetic progressions", *Comment. Math. Helv.* **89**:4 (2014), 979–1014. MR 3284303 Zbl 1306.11079

[Fouvry et al. 2015] É. Fouvry, E. Kowalski, and P. Michel, "A study in sums of products", *Philos. Trans. A* **373**:2040 (2015), 20140309, 26. MR 3338119

[Iwaniec and Sarnak 1999] H. Iwaniec and P. Sarnak, "Dirichlet L-functions at the central point", pp. 941–952 in *Number theory in progress* (Zakopane-Kościelisko, 1997), vol. 2, de Gruyter, Berlin, 1999. MR 1689553 Zbl 0929.11025

[Iwaniec and Sarnak 2000] H. Iwaniec and P. Sarnak, "The non-vanishing of central values of automorphic L-functions and Landau–Siegel zeros", *Israel J. Math.* **120**:1 (2000), 155–177. MR 1815374 Zbl 0992.11037

[Kowalski et al. 2000a] E. Kowalski, P. Michel, and J. VanderKam, "Mollification of the fourth moment of automorphic L-functions and arithmetic applications", *Invent. Math.* **142**:1 (2000), 95–151. MR 1784797 Zbl 1054.11026

[Kowalski et al. 2000b] E. Kowalski, P. Michel, and J. VanderKam, "Non-vanishing of high derivatives of automorphic L-functions at the center of the critical strip", *J. Reine Angew. Math.* **526** (2000), 1–34. MR 1778299 Zbl 1020.11033

[Michel and VanderKam 2000] P. Michel and J. VanderKam, "Non-vanishing of high derivatives of Dirichlet L-functions at the central point", *J. Number Theory* **81**:1 (2000), 130–148. MR 1743500 Zbl 1001.11032

[Paley 1931] R. E. A. C. Paley, "On the k-analogues of some theorems in the theory of the Riemann ζ-function", *Proc. London Math. Soc.* **S2-32**:1 (1931), 273–311. MR 1575993 Zbl 0002.01601

[Selberg 1942] A. Selberg, "On the zeros of Riemann's zeta-function", *Skr. Norske Vid. Akad. Oslo I.* **1942**:10 (1942), 59. MR 0010712 Zbl 68.0161.01

[Soundararajan 1995] K. Soundararajan, "Mean-values of the Riemann zeta-function", *Mathematika* **42**:1 (1995), 158–174. MR 1346680 Zbl 0830.11032

[Soundararajan 2000] K. Soundararajan, "Nonvanishing of quadratic Dirichlet L-functions at $s = \frac{1}{2}$", *Ann. of Math.* (2) **152**:2 (2000), 447–488. MR 1804529 Zbl 0964.11034

Communicated by Andrew Granville
Received 2015-12-13 Revised 2016-08-18 Accepted 2016-09-17

rizwanur.khan@qatar.tamu.edu *Science Program, Texas A&M University at Qatar, PO Box 23874, Doha, Qatar*

trunghieu.ay@gmail.com *Science Program, Texas A&M University at Qatar, PO Box 23874, Doha, Qatar*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2093

Every integer greater than 454 is the sum of at most seven positive cubes

Samir Siksek

A long-standing conjecture states that every positive integer other than

15, 22, 23, 50, 114, 167, 175, 186, 212,
231, 238, 239, 303, 364, 420, 428, 454

is a sum of at most seven positive cubes. This was first observed by Jacobi in 1851 on the basis of extensive calculations performed by the famous computationalist Zacharias Dase. We complete the proof of this conjecture, building on previous work of Linnik, Watson, McCurley, Ramaré, Boklan, Elkies, and many others.

1. Historical introduction

In 1770, Edward Waring stated in his *Meditationes Algebraicæ*,

> *Omnis integer numerus vel est cubus, vel e duobus, tribus, 4, 5, 6, 7, 8, vel novem cubis compositus, ...*

Waring's assertion can be concisely reformulated as the assertion that "every positive integer is the sum of nine nonnegative cubes". Henceforth, *by a cube we shall mean a nonnegative cube*. In the 19th century, numerical experimentation led to refinements of Waring's assertion for sums of cubes. As noted by Dickson [1927], "*At the request of Jacobi, the famous computer Dase constructed a table showing the least number of positive cubes whose sum is any* $p < 12\,000$". In an influential Crelle paper, Jacobi [1851] made a series of observations based on Dase's table: every positive integer other than 23 and 239 is the sum of eight cubes, every integer greater than 454 is the sum of seven cubes, and every integer greater than 8 042 is the sum of six cubes. Jacobi believed that every sufficiently large integer is the sum of five cubes, whilst recognizing that the cutoff point must be far beyond Dase's table, and he wondered if the same is true for sums of four cubes. He noted that integers equivalent to 4, 5 (mod 9) cannot be sums of three cubes. Later

The author is supported by an EPSRC Leadership Fellowship EP/G007268/1, and EPSRC *LMF: L-Functions and Modular Forms* Programme Grant EP/K034383/1.
MSC2010: 11P05.
Keywords: Waring, cubes, sums of cubes.

computations by Romani [1982] convincingly suggest that every integer greater than 1 290 740 is the sum of five cubes, and by Deshouillers et al. [2000] that every integer greater than 7 373 170 279 850 is the sum of four cubes.

Progress towards proving these observations of Waring, Jacobi and others has been exceedingly slow. Maillet [1895] showed that twenty-one cubes are enough to represent every positive integer. At the heart of Maillet's proof is an idea crucial to virtually all future developments; the identity $(r+x)^3+(r-x)^3 = 2r^3+6rx^2$ allows one to reformulate the problem of representing an integer as the sum of a (certain number of) cubes in terms of representing a related integer as the sum of (a smaller number of) squares. Exploiting this idea, Wieferich [1908] proved Waring's assertion (Wieferich's proof had a mistake that was corrected by Kempner [1912]). In fact, the theoretical part of Wieferich's proof showed that all integers exceeding 2.25×10^9 are sums of nine cubes. Completing the proof required appealing to a table of von Sterneck [1903] (who extended Dase's table to 40 000), and applying what is now known as *the greedy algorithm* to reach the bound.

Soon thereafter, Landau [1908] showed that every sufficiently large integer is the sum of eight cubes. This was made effective by Baer [1913], who showed that every integer greater than or equal to $14.1 \times 233^6 \approx 2.26 \times 10^{15}$ is the sum of eight cubes. Dickson [1939] completed the proof of Jacobi's observation that all positive integers other than 23 and 239 are sums of eight cubes. Remarkably, Dickson's proof relied on extending von Sterneck's table to 123 000 (with the help of his assistant, Miss Evelyn Garbe) and then applying the greedy algorithm to reach Baer's bound.

Linnik [1943] showed that every sufficiently large integer is the sum of seven cubes. A substantially simpler proof (though still ineffective) was given by Watson [1951]. Linnik's seven cubes theorem was first made effective by McCurley [1984], who showed that it is true for integers greater than $\exp(\exp(13.94))$. Ramaré [2005] improved this to $\exp(205\,000)$ and finally to $\exp(524) \approx 3.72 \times 10^{227}$ [Ramaré 2007]. This bound is way beyond computer searches combined with the greedy algorithm. In [Deshouillers et al. 2000], it is shown that every integer between 1 290 741 and 10^{16} is a sum of five cubes. As observed in [Ramaré 2007], combining this with the greedy algorithm [Bertault et al. 1999, Lemma 3], we can easily deduce that every integer $455 \le N \le \exp(78.7) \approx 1.51 \times 10^{34}$ is the sum of seven cubes.

There has been a number of partial results concerning sums of seven cubes. Bertault et al. [1999] show that every nonnegative integer which is a cubic residue modulo 9 and an invertible cubic residue modulo 37 is a sum of 7 cubes. Boklan and Elkies [2009] show that every multiple of 4 greater than 454 is the sum of seven cubes, whilst Elkies [2010] shows the same for integers equivalent to $2 \pmod 4$.

In this paper we complete the proof of Jacobi's seven cubes conjecture, building on the aforementioned great works.

Theorem 1. *Every positive integer other than*

$$15, 22, 23, 50, 114, 167, 175, 186, 212, 231, 238, 239, 303, 364, 420, 428, 454$$

is the sum of seven cubes.

An online supplement contains Magma scripts implementing the algorithms that support the proof.

2. The main criterion

Let $\mathscr{K} = \exp(524)$ and $\mathscr{K}' = \exp(78.7)$. By results found in [Ramaré 2007] and [Deshouillers et al. 2000], it is sufficient to prove that every integer $\mathscr{K}' \le N \le \mathscr{K}$ is the sum of seven cubes.

Results from [Boklan and Elkies 2009] and [Elkies 2010] allow us to restrict ourselves to odd integers N (our method can certainly be adapted to deal with even integers, but restricting ourselves to odd integers brings coherence to our exposition). In this section we give a criterion (Proposition 2.2) for all odd integers N in a range $K_1 \le N \le K_2$ to be sums of seven cubes. Most of the remainder of the paper is devoted to showing that this criterion holds for each of the ranges $\left(\frac{9}{10}\right)^{n+1}\mathscr{K} \le N \le \left(\frac{9}{10}\right)^{n}\mathscr{K}$ with $0 \le n \le 4226$. This will complete the proof of Theorem 1 as

$$\left(\frac{9}{10}\right)^{4227}\mathscr{K} \approx 1.42 \times 10^{34} \quad \text{and} \quad \mathscr{K}' \approx 1.51 \times 10^{34}.$$

Theorem 2 (Gauss, Legendre). *Let $k \ge 0$ be an even integer. There exist integers x, y, z such that*

$$x^2 + x + y^2 + y + z^2 + z = k. \tag{1}$$

Proof. Dividing by 2 we see that this is in fact the famous theorem, due to Gauss, that every nonnegative integer is the sum of three triangular numbers. Alternatively, we can rewrite (1) as

$$(2x+1)^2 + (2y+1)^2 + (2z+1)^2 = 4k + 3. \tag{2}$$

As k is even, $4k + 3 \equiv 3 \pmod 8$; by a theorem of Legendre, every positive integer equivalent to $3 \pmod 8$ is the sum of three odd squares. □

Throughout this section m will denote a positive integer satisfying the conditions

(i) m is a squarefree,

(ii) $3 \mid m$,

(iii) every prime divisor of $m/3$ is $\equiv 5 \pmod 6$.

Observe that $m \equiv 3 \pmod 6$. Moreover, for any integer N, there is a unique integer $t \in [0, m)$ such that $N \equiv 8t^3 \pmod m$. Our starting point is a modified version of Lemma 3 of [Watson 1951].

Lemma 2.1. *Let $0 < K_1 < K_2$ be real numbers. Let m be a positive integer satisfying (i)–(iii) above. Let ε_m and δ_m be real numbers satisfying*

(iv) $0 \le \varepsilon_m < \delta_m \le 1$,

(v) $K_1 \ge \left(8\delta_m^3 + \frac{1}{36}\right)m^3 + 3m/4$,

(vi) $K_2 \le \left(8\varepsilon_m^3 + \frac{1}{18}\right)m^3 + m/2$.

Let $K_1 \le N \le K_2$ be an odd integer. Suppose $N \equiv 8t^3 \pmod m$ with $t \in [\varepsilon_m \cdot m, \delta_m \cdot m)$. Then N is the sum of seven nonnegative cubes.

Proof. Write $m = 6r + 3$. Let

$$k = \frac{N - 8t^3}{m} - (r^2 + r + 1). \tag{3}$$

The quantity k is an integer as $N \equiv 8t^3 \pmod m$, and even as $(N - 8t^3)/m$ and $r^2 + r + 1$ are both odd. Observe that

$$\begin{aligned}
k &> \frac{N - 8\delta_m^3 \cdot m^3}{m} - (r^2 + r + 1) && \text{(since } t < \delta_m \cdot m\text{)} \\
&\ge \frac{K_1 - 8\delta_m^3 \cdot m^3}{m} - (r^2 + r + 1) && \text{(since } N \ge K_1\text{)} \\
&= \frac{K_1 - 8\delta_m^3 \cdot m^3}{m} - \frac{m^2}{36} - \frac{3}{4} && \text{(substituting } r = (m-3)/6\text{)} \\
&\ge 0 && \text{(by (v)).}
\end{aligned}$$

As k is nonnegative and even, by the Gauss–Legendre theorem, there exist integers x, y, z satisfying (1). We shall make use of the identity

$$\begin{aligned}
(r+1+x)^3 + (r-x)^3 + (r+1+y)^3 + (r-y)^3 + (r+1+z)^3 + (r-z)^3 \\
= (6r+3)(r^2 + r + 1 + x^2 + x + y^2 + y + z^2 + z).
\end{aligned} \tag{4}$$

From the definition of k in (3) and the fact that $m = 6r + 3$, we see that $N - 8t^3$ is equal to the right-hand side of the identity (4). Hence

$$N = (r+1+x)^3 + (r-x)^3 + (r+1+y)^3 + (r-y)^3 + (r+1+z)^3 + (r-z)^3 + (2t)^3.$$

To complete the proof it is enough to show that these cubes are nonnegative, or equivalently that

$$-r - 1 \le x, y, z \le r.$$

This is equivalent to showing that

$$-(2r+1) \le 2x + 1, 2y + 1, 2z + 1 \le 2r + 1.$$

Now $(2y+1)^2, (2z+1)^2 \geq 1$ and so from (2), we have $(2x+1)^2 \leq 4k+1$. It is therefore enough to show that $4k+1 \leq (2r+1)^2$ or equivalently that $k \leq r^2+r$. The following inequalities complete the proof:

$$\begin{aligned} k - r^2 - r &= \frac{N - 8t^3}{m} - (2r^2+2r+1) && \text{(from (3))} \\ &\leq \frac{N - 8\varepsilon_m^3 \cdot m^3}{m} - (2r^2+2r+1) && \text{(since } t \geq \varepsilon_m \cdot m\text{)} \\ &\leq \frac{K_2 - 8\varepsilon_m^3 \cdot m^3}{m} - (2r^2+2r+1) && \text{(since } N \leq K_2\text{)} \\ &\leq \frac{K_2 - 8\varepsilon_m^3 \cdot m^3}{m} - \frac{m^2}{18} - \frac{1}{2} && \text{(substituting } r = (m-3)/6\text{)} \\ &\leq 0 && \text{(by (vi)).} \end{aligned}$$

□

This simpleminded lemma has one serious flaw. The inequality $K_1 < K_2$ and the conditions (iv)–(vi) together imply that

$$\delta_m^3 < \varepsilon_m^3 + \frac{1}{288} - \frac{1}{32m^2}.$$

In particular, this forces the interval $[\varepsilon_m \cdot m, \delta_m \cdot m)$ to have length less than $m/\sqrt[3]{288} \approx 0.15m$. On the other hand, the integer t appearing in the lemma (which is the cube root of $N/8$ modulo m) can be any integer in the interval $[0, m)$. Thus the lemma only treats a small fraction of the odd integers $K_1 \leq N \leq K_2$. Our key innovation over the works mentioned in the introduction is to use not just one value of m, but many of them simultaneously. Each value of m will give some information about those odd integers $K_1 \leq N \leq K_2$ that cannot be expressed as sums of seven cubes; collecting this information will allow us to deduce a contradiction.

Let x be a real number and m be a positive integer. Define the *quotient* and *remainder* obtained on dividing x by m as

$$\mathrm{Q}(x,m) = \lfloor x/m \rfloor, \quad \mathrm{R}(x,m) = x - \mathrm{Q}(x,m) \cdot m.$$

In particular, $\mathrm{R}(x,m)$ belongs to the half-open interval $[0,m)$. If $x \in \mathbb{Z}$ then $\mathrm{R}(x,m)$ is the usual remainder on dividing by m and $x \equiv \mathrm{R}(x,m) \pmod m$. Let ε_m and δ_m be real numbers satisfying $0 \leq \varepsilon_m < \delta_m \leq 1$. Define

$$\begin{aligned} \mathrm{Bad}(m, \varepsilon_m, \delta_m) &= \{x \in \mathbb{R} : \mathrm{R}(x,m) \in [0,m) \setminus [\varepsilon_m \cdot m, \delta_m \cdot m)\} \\ &= \bigcup_{k=-\infty}^{\infty} km + ([0,m) \setminus [\varepsilon_m \cdot m, \delta_m \cdot m)). \end{aligned} \tag{5}$$

The reader will observe, in Lemma 2.1, if N is not the sum of seven cubes, then $t \in \mathrm{Bad}(m, \varepsilon_m, \delta_m)$, which explains our choice of the epithet "bad". Given a set of

positive integers $\mathcal{W}$, and sequences $\underset{\sim}{\varepsilon} = (\varepsilon_m)_{m\in\mathcal{W}}$, $\underset{\sim}{\delta} = (\delta_m)_{m\in\mathcal{W}}$ of real numbers satisfying $0 \le \varepsilon_m < \delta_m \le 1$ for all $m \in \mathcal{W}$, we define

$$\mathrm{Bad}(\mathcal{W}, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta}) = \bigcap_{m\in\mathcal{W}} \mathrm{Bad}(m, \varepsilon_m, \delta_m). \tag{6}$$

To make the notation less cumbersome, we usually regard the values ε_m and δ_m as implicit, and write $\mathrm{Bad}(m)$ for $\mathrm{Bad}(m, \varepsilon, \delta)$, and $\mathrm{Bad}(\mathcal{W})$ for $\mathrm{Bad}(\mathcal{W}, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta})$.

Proposition 2.2. *Let $0 < K_1 < K_2$ be real numbers. Let $\mathcal{W}$ be a nonempty finite set of integers such that every element $m \in \mathcal{W}$ satisfies conditions (i)–(iii). Suppose moreover, that for each $m \in \mathcal{W}$, there are real numbers ε_m and δ_m satisfying conditions (iv)–(vi). Let $M = \mathrm{lcm}(\mathcal{W})$. Let $\mathfrak{S} \subset [0, 1]$ be a finite set of rational numbers a/q (here $\gcd(a, q) = 1$) with denominators q bounded by $\sqrt[3]{M/2K_2}$. Suppose that*

$$\mathrm{Bad}(\mathcal{W}) \cap [0, M) \subseteq \bigcup_{a/q\in\mathfrak{S}} \left(\frac{a}{q}M - \frac{\sqrt[3]{M/16}}{q}, \frac{a}{q}M + \frac{\sqrt[3]{M/16}}{q} \right). \tag{7}$$

Then every odd integer $K_1 \le N \le K_2$ is the sum of seven nonnegative cubes.

Proof. Let N be an odd integer satisfying $K_1 \le N \le K_2$. It follows from assumptions (i)–(iii) that $M = \mathrm{lcm}(\mathcal{W})$ is squarefree and divisible only by 3 and primes equivalent to 5 (mod 6). Thus there exists a unique integer $T \in [0, M)$ such that

$$N \equiv 8T^3 \pmod{M}. \tag{8}$$

Suppose N is not the sum of seven cubes. Then, by Lemma 2.1, for each $m \in \mathcal{W}$, we have $\mathrm{R}(T, m) \in [0, m) \setminus [\varepsilon_m \cdot m, \delta_m \cdot m)$. Thus $T \in \mathrm{Bad}(\mathcal{W}) \cap [0, M)$. By (7) there is some rational $a/q \in \mathfrak{S}$ such that

$$-\frac{\sqrt[3]{M/16}}{q} < T - \frac{a}{q}M < \frac{\sqrt[3]{M/16}}{q},$$

or equivalently

$$-\frac{M}{2} < 8(qT - aM)^3 < \frac{M}{2}.$$

Moreover, the denominator q is bounded by $\sqrt[3]{M/2K_2}$ and so

$$q^3 N \le \frac{MN}{2K_2} \le \frac{M}{2},$$

as $N \le K_2$. Hence

$$|q^3 N - 8(qT - aM)^3| < M.$$

However, by (8), we have $q^3N - 8(qT - aM)^3 \equiv 0 \pmod{M}$. Thus $q^3N = 8(aT - aM)^3$. It follows that N is a perfect cube, and so is certainly the sum of seven nonnegative cubes. □

We shall mostly apply Proposition 2.2 with the parameter choices given by the following lemma.

Lemma 2.3. *Let $K \geq 10^5$. Let $K_1 = 9K/10$, $K_2 = K$. Let*

$$\tfrac{263}{100}K^{\frac{1}{3}} \leq m \leq \tfrac{292}{100}K^{\frac{1}{3}}. \tag{9}$$

Then conditions (iv)–(vi) are satisfied with $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$.

3. Plan for the paper

The rest of the paper is devoted to understanding and computing the intersections $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ appearing in Proposition 2.2. Section 4 collects various properties of remainders and bad sets that are used throughout. Section 5 provides justification, under a plausible assumption, that the intersection $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ should be decomposable as in (7). Section 6 gives an algorithm (Algorithm 1) which takes as input a finite set of positive integers $\mathcal{W}$ and an interval $[A, B)$ and returns the intersection $\mathrm{Bad}(\mathcal{W}) \cap [A, B)$. We also give a heuristic analysis of the algorithm and its running time. Section 7 introduces the concept of a "tower", which is a sequence

$$\mathcal{W}_0 \subseteq \mathcal{W}_1 \subseteq \mathcal{W}_2 \subseteq \cdots \subseteq \mathcal{W}_r = \mathcal{W}. \tag{10}$$

Letting $M_i = \mathrm{lcm}(\mathcal{W}_i)$, we prove the recursive formula for computing $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ in terms of $\mathrm{Bad}(\mathcal{W}_{i-1}) \cap [0, M_{i-1})$. This recursive formula together with Algorithm 1 is the basis for a much more efficient algorithm (Algorithm 2) for computing $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ given in Section 7.

In Section 8 we let M^* be the product of all primes $p \leq 167$ that are equivalent to 5 (mod 6), and

$$\mathcal{W}^* = \{m \mid M^* : 265 \times 10^9 \leq m \leq 290 \times 10^9\}. \tag{11}$$

We use a tower and Algorithm 2 to compute $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$. The actual computation consumed about 18,300 hours of CPU time.

Section 9 is devoted to proving Theorem 1 for $N \geq \left(\frac{9}{10}\right)^{3998} \cdot \mathscr{K} \approx 4.28 \times 10^{44}$, where $\mathscr{K} = \exp(524)$. The approach is to divide the interval $\left(\frac{9}{10}\right)^{3998}\mathscr{K} \leq N \leq \mathscr{K}$ into subintervals $\left(\frac{9}{10}\right)^{n+1}\mathscr{K} \leq N \leq \left(\frac{9}{10}\right)^{n}\mathscr{K}$ with $0 \leq n \leq 3997$, and apply Proposition 2.2 and Lemma 2.3 to prove that all odd integers in the interval $\left(\frac{9}{10}\right)^{n+1}\mathscr{K} \leq N \leq \left(\frac{9}{10}\right)^{n}\mathscr{K}$ are sums of seven nonnegative cubes. Indeed, we show that given $0 \leq n \leq 3997$, there is some suitable positive κ such that the elements of $\mathcal{W}_0 = \kappa \cdot \mathcal{W}^*$ satisfy conditions (i)–(iii) (with $K_1 = \left(\frac{9}{10}\right)^{n+1}\mathscr{K}$ and

$K_2 = \left(\frac{9}{10}\right)^n \mathscr{K}$) and that moreover, $\mathrm{Bad}(\mathcal{W}_0) = \kappa\,\mathrm{Bad}(\mathcal{W}^*)$. Thus the results of the huge computation of Section 8 are recycled 3998 times; on top of this $\mathcal{W}_0$ we construct a tower and continue until we have found a set $\mathcal{W}$ that satisfies the hypotheses of Proposition 2.2, thereby proving Theorem 1 for $N \geq \left(\frac{9}{10}\right)^{3998} \cdot \mathscr{K}$. The CPU time for the computations described in Section 9 was around 10,000 hours.

The proof of Theorem 1 is completed in Section 10 where a modified strategy is needed to handle the "small" ranges $\left(\frac{9}{10}\right)^{n+1} \mathscr{K} \leq N \leq \left(\frac{9}{10}\right)^n \mathscr{K}$ with $3998 \leq n \leq 4226$. Although these intervals are small (and few) compared to those handled in Section 9, we are unable to recycle the computation of Section 8. This makes the computations far less efficient, though still practical. The CPU time for the computations described in Section 10 was around 2,750 hours.

4. Some properties of remainders and bad sets

Lemma 4.1. *Let m and κ be positive integers with $\kappa | m$. Then for any real x we have*

$$\mathrm{Q}\left(\frac{x}{\kappa}, \frac{m}{\kappa}\right) = \mathrm{Q}(x, m), \quad \mathrm{R}\left(\frac{x}{\kappa}, \frac{m}{\kappa}\right) = \frac{1}{\kappa}\,\mathrm{R}(x, m).$$

Let κ be a positive integer. For a set $X \subset \mathbb{R}$ we denote $\kappa X = \{\kappa x : x \in X\}$.

Lemma 4.2. *Let m and κ be positive integers. Let $0 \leq \varepsilon < \delta \leq 1$ be real numbers. Then*

$$\mathrm{Bad}(\kappa m, \varepsilon, \delta) = \kappa \cdot \mathrm{Bad}(m, \varepsilon, \delta).$$

Let $\mathcal{W}$ be a set of positive integers and for $m \in \mathcal{W}$ let $0 \leq \varepsilon_m < \delta_m \leq 1$ be real numbers. Let

$$\mathcal{W}' = \kappa \cdot \mathcal{W}, \quad \underset{\sim}{\varepsilon} = (\varepsilon_m)_{m \in \mathcal{W}}, \quad \underset{\sim}{\delta} = (\delta_m)_{m \in \mathcal{W}}, \quad \underset{\sim}{\varepsilon}' = (\varepsilon_{m/\kappa})_{m \in \mathcal{W}'} \quad \underset{\sim}{\delta}' = (\delta_{m/\kappa})_{m \in \mathcal{W}'}.$$

Then

$$\mathrm{Bad}(\mathcal{W}', \underset{\sim}{\varepsilon}', \underset{\sim}{\delta}') = \kappa \cdot \mathrm{Bad}(\mathcal{W}, \underset{\sim}{\varepsilon}, \underset{\sim}{\delta}).$$

Proof. By (5) and Lemma 4.1,

$$\begin{aligned} x \in \mathrm{Bad}(\kappa m, \varepsilon, \delta) &\iff \mathrm{R}(x, \kappa m) \in [0, \kappa m) \setminus [\varepsilon \cdot \kappa m, \delta \cdot \kappa m) \\ &\iff 1/\kappa\,\mathrm{R}(x, \kappa m) \in [0, m) \setminus [\varepsilon \cdot m, \delta \cdot m) \\ &\iff \mathrm{R}(x/\kappa, m) \in [0, m) \setminus [\varepsilon \cdot m, \delta \cdot m) \\ &\iff x/\kappa \in \mathrm{Bad}(m, \varepsilon, \delta) \\ &\iff x \in \kappa \cdot \mathrm{Bad}(m, \varepsilon, \delta). \end{aligned}$$

This proves the first part of the lemma. The second part now follows from (6). □

Lemma 4.3. *Given positive integers $M_1 \mid M_2$, we define the "natural" map*

$$\pi_{M_2,M_1} : [0, M_2) \longrightarrow [0, M_1), \quad x \mapsto \mathrm{R}(x, M_1).$$

Then π_{M_2,M_1} is surjective, and for any $T \subseteq [0, M_1)$,

$$\pi^{-1}_{M_2,M_1}(T) = \bigcup_{k=0}^{(M_2/M_1)-1} (k \cdot M_1 + T).$$

Lemma 4.4. *Let $\mathcal{W}_1, \mathcal{W}_2$ be sets of positive integers with $\mathcal{W}_1 \subseteq \mathcal{W}_2$, $M_i = \mathrm{lcm}(\mathcal{W}_i)$, and $\pi = \pi_{M_2,M_1}$. Write $\mathcal{U} = \mathcal{W}_2 \setminus \mathcal{W}_1$. Then $\pi(\mathrm{Bad}(\mathcal{W}_2)) \subseteq \mathrm{Bad}(\mathcal{W}_1)$ and*

$$\mathrm{Bad}(\mathcal{W}_2) \cap [0, M_2) = \pi^{-1}(\mathrm{Bad}(\mathcal{W}_1) \cap [0, M_1)) \cap \mathrm{Bad}(\mathcal{U}).$$

Proof. Let $x \in \mathbb{R}$ and let $y = \pi(x) = \mathrm{R}(x, M_1)$. If $m \in \mathcal{W}_1$ then $\mathrm{R}(y, m) = \mathrm{R}(x, m)$, as $m \mid M_1$. Observe that

$$\begin{aligned} x \in \mathrm{Bad}(\mathcal{W}_2) &\iff \mathrm{R}(x, m) \notin [\varepsilon_m \cdot m, \delta_m \cdot m) \text{ for all } m \in \mathcal{W}_2 \\ &\implies \mathrm{R}(x, m) \notin [\varepsilon_m \cdot m, \delta_m \cdot m) \text{ for all } m \in \mathcal{W}_1 \\ &\iff \mathrm{R}(y, m) \notin [\varepsilon_m \cdot m, \delta_m \cdot m) \text{ for all } m \in \mathcal{W}_1 \\ &\iff y \in \mathrm{Bad}(\mathcal{W}_1). \end{aligned}$$

This shows that $\pi(\mathrm{Bad}(\mathcal{W}_2)) \subseteq \mathrm{Bad}(\mathcal{W}_1)$. The rest of the lemma easily follows. □

5. Gaps and ripples

We will soon give an algorithm for computing the intersection

$$\mathrm{Bad}(\mathcal{W}) \cap [0, M) = \Big(\bigcap_{m \in \mathcal{W}} \mathrm{Bad}(m) \Big) \cap [0, M), \quad M = \mathrm{lcm}(\mathcal{W}),$$

given a set $\mathcal{W}$ that satisfies the conditions of Proposition 2.2. The statement of Proposition 2.2 (notably (7)) suggests that we are expecting this intersection to be concentrated in small intervals around aM/q for certain a/q with relatively small denominators q. In this section we provide an explanation for this. The situation is easier to analyze if we make choices of parameters as in Lemma 2.3. Thus for this section we fix the choices $\varepsilon_m = 0$, $\delta_m = \frac{1}{10}$, and hence $\mathrm{Bad}(m) = \mathrm{Bad}\big(m, 0, \frac{1}{10}\big)$. We suppose that the elements $m \in \mathcal{W}$ belong to an interval of the form

$$\tfrac{263}{100} L \le m \le \tfrac{292}{100} L, \tag{12}$$

for some $L > 0$ (see Lemma 2.3). In fact, we show that if q is large, and if the residues of the integers aM/m are regularly distributed modulo q (in a sense that will be made precise), then the intersection $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ contains no points in a certain explicitly given neighborhood of aM/q. Likewise we show, for certain a/q

with q small, that $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ does contain some points near aM/q. We stress that the material in this section does not form part of our proof of Theorem 1. It does however explain the results of our computations that do form part of the proof of Theorem 1, and it lends credibility to them.

We fix the following notation throughout this section:

- L is a positive real number.
- $\mathcal{W}$ is a nonempty set of positive integers that belong to the interval (12).
- $M = \mathrm{lcm}(\mathcal{W})$.

Ripples.

Proposition 5.1. *Suppose $M \geq 2000L$. Let $a/q \in [0, 1)$ be a fraction in simplest form with $1 \leq q \leq 9$ and $0 \leq a \leq q-1$. For $0 \leq k \leq 9-q$ let*

$$\psi_k = \frac{292}{100}\Big(\frac{k}{q} + \frac{1}{10}\Big), \quad \Psi_k = \frac{263}{100} \cdot \frac{k+1}{q}. \tag{13}$$

Then $\psi_k < \Psi_k$ and

$$\bigcup_{k=0}^{9-q} \Big(\frac{a}{q}M + \psi_k \cdot L, \frac{a}{q}M + \Psi_k \cdot L\Big) \subseteq \mathrm{Bad}(\mathcal{W}) \cap [0, M). \tag{14}$$

This recipe gives 103 *disjoint intervals contained in* $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ *of total length $\xi \cdot L$ where*

$$\xi = \tfrac{261707}{10500} \approx 24.9.$$

We shall informally refer to the union of intervals (14) as a *ripple emanating from aM/q* in the positive direction. The reader will easily modify the proof below to show, under similar hypotheses, that there are ripples emanating from the aM/q in the negative direction.

Proof. It is easy to check that $\psi_k < \Psi_k$ for $q \leq 9$ and $0 \leq k \leq 9-q$. The assumption $M \geq 2000L$ ensures that the 103 intervals are contained in $[0, M)$ and are disjoint, so it is enough to show that the intervals are contained in $\mathrm{Bad}(\mathcal{W})$. Let α be a real number belonging to the interval $\psi_k \cdot L < \alpha < \Psi_k \cdot L$. We would like to show that $aM/q + \alpha \in \mathrm{Bad}(m)$ for all $m \in \mathcal{W}$. Let $m \in \mathcal{W}$. It follows from (12) and (13) that

$$\Big(\frac{k}{q} + \frac{1}{10}\Big)m \leq \psi_k \cdot L < \alpha < \Psi_k \cdot L \leq \frac{k+1}{q}m. \tag{15}$$

As $m \mid M$ we can write $aM = um$ with $u \in \mathbb{Z}$. Now $u = bq + s$ where $0 \leq s \leq q-1$. Thus

$$\frac{a}{q}M = bm + \frac{s}{q}m.$$

From (15),

$$bm + \frac{k+s}{q}m + \frac{m}{10} < \frac{a}{q}M + \alpha < bm + \frac{k+s+1}{q}m.$$

Let $k+s = qt+v$ where $0 \le v \le q-1$. Hence

$$(b+t)m + \Big(\frac{v}{q} + \frac{1}{10}\Big)m < \frac{a}{q}M + \alpha < (b+t)m + \frac{v+1}{q}m.$$

Observe that

$$\frac{1}{10} \le \frac{v}{q} + \frac{1}{10} < \frac{v+1}{q} \le 1,$$

as $q \le 9$ and $0 \le v \le q-1$. Thus $\mathrm{Q}(aM/q+\alpha, m) = b+t$ and

$$\frac{m}{10} < \mathrm{R}\Big(\frac{aM}{q} + \alpha, m\Big) < m.$$

This shows that $aM/q + \alpha \in \mathrm{Bad}(m)$ as required. □

In the above proposition we showed the existence of ripples emanating from aM/q for $q \le 9$. There can also be ripples emanating for aM/q for larger values of q if the sequence of residues $\overline{aM/m}$ in $\mathbb{Z}/q\mathbb{Z}$ contains large gaps as illustrated by the following proposition.

Proposition 5.2. *Let $a/q \in (0,1)$ be a rational number in simplest form with $q \ge 11$ and $1 \le a \le q-1$. Let $(q-10)/10 < d < q-1$ be an integer, and let s be a nonnegative integer satisfying*

$$s < q - d - 1, \quad s < \tfrac{263}{290}(10d + 10 - q). \tag{16}$$

Suppose

$$\overline{s+1}, \overline{s+2}, \ldots, \overline{s+d} \notin \{\overline{aM/m} : m \in \mathcal{W}\} \subseteq \mathbb{Z}/q\mathbb{Z}. \tag{17}$$

Let

$$\pi = \frac{292}{100} \cdot \frac{s}{q}, \quad \Pi = \frac{263}{100}\Big(\frac{s+d+1}{q} - \frac{1}{10}\Big).$$

Then $\pi < \Pi$ and

$$\Big(\frac{a}{q}M - \Pi \cdot L, \frac{a}{q}M - \pi \cdot L\Big) \subseteq \mathrm{Bad}(\mathcal{W}).$$

Proof. Let $m \in \mathcal{W}$, and recall that $m \mid M$. Thus aM/m is an integer, and hence so is $\mathrm{R}(aM/m, q)$. By assumption (17),

$$\mathrm{R}(aM/m, q) \ne s+1, s+2, \ldots, s+d.$$

Thus $\mathrm{R}(aM/m, q) \notin (s, s+d+1)$. By Lemma 4.1,

$$\mathrm{R}(aM/q, m) = \mathrm{R}(aM, qm) \cdot 1/q = \mathrm{R}(aM/m, q) \cdot m/q.$$

Thus

$$\mathrm{R}(aM/q, m) \notin \Big(\frac{s}{q}m, \frac{s+d+1}{q}m\Big). \tag{18}$$

The condition $d > (q-10)/10$ implies that

$$\frac{s}{q} < \frac{s+d+1}{q} - \frac{1}{10}.$$

Let α belong to the interval

$$\frac{s}{q}m < \alpha < \Big(\frac{s+d+1}{q} - \frac{1}{10}\Big)m. \tag{19}$$

We claim that

$$\mathrm{R}(aM/q - \alpha, m) \notin [0, m/10).$$

Suppose otherwise; then we can write

$$\frac{a}{q}M - \alpha = bm + r,$$

where $0 \le r < m/10$. Thus

$$bm + \frac{s}{q}m < bm + \alpha \le \frac{a}{q}M < bm + \alpha + \frac{m}{10} < bm + \frac{s+d+1}{q}m,$$

as α satisfies (19). This contradicts (18), and establishes our claim. In fact we have shown that if α belongs to the interval (19), then $aM/q - \alpha \in \mathrm{Bad}(m)$.

Suppose now that α belongs to the interval $\pi \cdot L < \alpha < \Pi \cdot L$ (the second inequality in (16) ensures $\pi < \Pi$). To prove the proposition, all we have to show is that α satisfies the inequalities in (19) for all $m \in \mathcal{W}$. However, these follow straightforwardly from the fact that all $m \in \mathcal{W}$ belong to the interval (12). □

A few remarks are in order concerning Proposition 5.2 and its proof:

- For simplicity we have only constructed the first interval in a ripple emanating from aM/q in the negative direction. If inequalities (16) are satisfied with a significant margin, then it is possible to construct more intervals belonging to this ripple. Likewise, with a suitable modification of the assumptions one can also construct a ripple in the positive direction.
- The first inequality in (16) is imposed merely for simplicity; if it does not hold one can also construct ripples emanating from aM/q after suitably modifying the second inequality in (16).
- The one indispensable assumption in Proposition 5.2 is the existence of a sequence

$$\overline{s+1}, \overline{s+2}, \ldots, \overline{s+d}$$

of consecutive residues belonging to $(\mathbb{Z}/q\mathbb{Z}) \setminus \{\overline{aM/m} : m \in \mathcal{W}\}$ of length d that is roughly larger than $q/10$. We shall show below that if there is no such sequence, then $\mathrm{Bad}(\mathcal{W})$ contains no elements in a neighborhood of aM/q.

Gaps. Let $a/q \in [0, 1]$ be a rational in simplest form, and let

$$\Phi_{a/q} : \mathcal{W} \to \mathbb{Z}/q\mathbb{Z}, \quad m \mapsto \overline{a(M/m)}.$$

In view of the above, define the *defect* $d(\mathcal{W}, a/q)$ of $\mathcal{W}$ with respect to a/q as the length of the longest sequence $\overline{s+1}, \overline{s+2}, \ldots, \overline{s+d}$ belonging to $(\mathbb{Z}/q\mathbb{Z}) \setminus \Phi_{a/q}(\mathcal{W})$. As $\mathcal{W} \neq \varnothing$, we have $d(\mathcal{W}, a/q) < q$. For example, if $\Phi_{a/q}$ is surjective then $d(\mathcal{W}, a/q) = 0$, and if $\Phi_{a/q}(\mathcal{W}) = (\mathbb{Z}/q\mathbb{Z})^*$ then $d(\mathcal{W}, a/q) = 1$.

Lemma 5.3. *With notation as above, let $d = d(\mathcal{W}, a/q)$. Let $x \in \mathbb{R}$. Then there is some element $m \in \mathcal{W}$ and an integer k such that*

$$\left| x - \frac{aM}{qm} - k \right| \leq \frac{d+1}{2q}.$$

Proof. Let $u \in \mathbb{Z}$ satisfy $|u - qx| \leq \frac{1}{2}$. We first suppose that d is even. Consider the sequence

$$\overline{u - d/2}, \overline{u - d/2 + 1}, \overline{u - d/2 + 2}, \ldots, \overline{u + d/2}$$

of $d+1$ elements of $\mathbb{Z}/q\mathbb{Z}$. By the definition of d, one of these equals $\Phi_{a/q}(m)$ for some $m \in \mathcal{W}$. Thus there is some integer k such that

$$\left| u - \frac{aM}{m} - kq \right| \leq \frac{d}{2}.$$

As $|u - qx| \leq \frac{1}{2}$, the result follows.

Now suppose that d is odd and $qx \geq u$ (the case $qx < u$ is similar). Consider the sequence

$$\overline{u - (d-1)/2}, \overline{u - (d-1)/2 + 1}, \overline{u - (d-1)/2 + 2}, \ldots, \overline{u + (d+1)/2}$$

which again has $d+1$ elements, and so there is some $m \in \mathcal{W}$ and some integer k such that

$$u - \frac{d-1}{2} \leq \frac{aM}{m} + kq \leq u + \frac{d+1}{2}.$$

Since $0 \leq qx - u \leq \frac{1}{2}$, the lemma follows. □

Lemma 5.4. *Let*

$$m^* = \frac{38398}{13875} \cdot L,$$

Then for all $m \in \mathcal{W}$,

$$\left| \frac{L}{m} - \frac{L}{m^*} \right| \leq \frac{725}{38398}.$$

Proof. By (12), the quantity L/m belongs to the interval $\left[\frac{100}{292}, \frac{100}{263}\right]$. We have chosen m^* so that L/m^* is the midpoint of the interval. The lemma follows as $\frac{725}{38398}$ is half the length of the interval. □

Proposition 5.5. *With notation as above, let $d = d(\mathcal{W}, a/q)$ and suppose that $d < (q-10)/10$. Let*

$$\mu = \frac{38398}{725}\Big(\frac{1}{20} - \frac{d+1}{2q}\Big). \tag{20}$$

Then

$$\Big(\frac{a}{q}M - \mu L, \frac{a}{q}M + \mu L\Big) \cap \mathrm{Bad}(\mathcal{W}) = \varnothing.$$

A few words are perhaps appropriate to help the reader appreciate the content of the proposition. We shall suppose that $q > 11$. If $\#\mathcal{W}$ is large compared to q, then we expect that $\Phi_{a/q}$ is close to being surjective which forces d to be small. If that is the case then μ should be close to $\frac{38398}{725\times 20} \approx 2.64$. Suppose now that $\#\mathcal{W}$ is large, but that q is much larger. Suppose also that the residues in the image $\Phi_{a/q}(\mathcal{W})$ are "randomly" distributed in $\mathbb{Z}/q\mathbb{Z}$. The quantity d measures how large the gaps between these residues in the image can be, and we expect that d should be around $q/\#\mathcal{W}$. We therefore expect that

$$\mu \approx \frac{38398}{725}\Big(\frac{1}{20} - \frac{1}{2\cdot\#\mathcal{W}}\Big).$$

We see that μ should be positive if $\mathcal{W}$ has much more than 10 elements.

Proof of Proposition 5.5. The assumption $d < (q-10)/10$ ensures that μ is positive. Let $y \in (aM/q - \mu L, aM/q + \mu L)$. We would to like to show that there is some $m \in \mathcal{W}$ such that $y \notin \mathrm{Bad}(m)$.

Write $y = aM/q + \beta$ where $|\beta| < \mu L$. Letting $x = \frac{1}{20} - \frac{\beta}{m^*}$ in Lemma 5.3, we deduce the existence of some integer k and some element $m \in \mathcal{W}$ such that

$$\Big|\frac{\beta}{m^*} + \frac{aM}{qm} + k - \frac{1}{20}\Big| \le \frac{d+1}{2q}.$$

Thus

$$\Big|\frac{\beta}{m} + \frac{aM}{qm} + k - \frac{1}{20}\Big| \le \frac{d+1}{2q} + \Big|\frac{\beta}{m^*} - \frac{\beta}{m}\Big|.$$

Using $|\beta| < \mu L$, Lemma 5.4 and the definition of μ in (20), we see that

$$\Big|\frac{\beta}{m} + \frac{aM}{qm} + k - \frac{1}{20}\Big| < \frac{1}{20}.$$

Thus $y = aM/q + \beta$ belongs to the interval $-km + (0, m/10)$, showing that $y \notin \mathrm{Bad}(m)$ as required. □

6. A first approach to computing Bad($\mathcal{W}$)

In this section $\mathcal{W}$ is a finite set of positive integers m. Associated to each $m \in \mathcal{W}$ are real numbers $0 \le \varepsilon_m < \delta_m < 1$. We shall write $\underset{\sim}{\varepsilon} = (\varepsilon_m)_{m\in\mathcal{W}}$ and $\underset{\sim}{\delta} = (\delta_m)_{m\in\mathcal{W}}$.

Lemma 6.1. *Let $A < B$ be real numbers. For $m \in \mathcal{W}$, let*

$$q_m = \mathrm{Q}(A, m) \quad \textit{and} \quad r_m = \mathrm{R}(A, m).$$

(a) *Suppose $r_n \in [\varepsilon_n \cdot n, \delta_n \cdot n)$ for some $n \in \mathcal{W}$. Write $A' = \min((q_n + \delta_n) \cdot n, B)$. Then*

$$\mathrm{Bad}(\mathcal{W}) \cap [A, B) = \mathrm{Bad}(\mathcal{W}) \cap [A', B).$$

(b) *Suppose $r_m \notin [\varepsilon_m \cdot m, \delta_m \cdot m)$ for all $m \in \mathcal{W}$. Define*

$$A_m = \begin{cases} (q_m+\varepsilon_m) \cdot m & \text{if } r_m < \varepsilon_m \cdot m, \\ (q_m+1+\varepsilon_m) \cdot m & \text{if } r_m \ge \delta_m \cdot m, \end{cases}, \quad A' = \min\left(B, \min(A_m)_{m\in\mathcal{W}}\right). \tag{21}$$

Then

$$\mathrm{Bad}(\mathcal{W}) \cap [A, B) = (\mathrm{Bad}(\mathcal{W}) \cap [A', B)) \cup [A, A').$$

Proof. Suppose $n \in \mathcal{W}$ satisfies $r_n \in [\varepsilon_n \cdot n, \delta_n \cdot n)$, and let A' be as in (a). By (5) we have

$$(q_n \cdot n + [\varepsilon_n \cdot n, \delta_n \cdot n)) \cap \mathrm{Bad}(n) = \varnothing.$$

Observe that $[A, A') \subseteq q_n \cdot n + [\varepsilon_n \cdot n, \delta_n \cdot n)$ and $[A, A') \subseteq [A, B)$. Part (a) follows.

Suppose now that $r_m \notin [\varepsilon_m \cdot m, \delta_m \cdot m)$ for all $m \in \mathcal{W}$, and let A' be as in (b). It is easy to check that $\mathrm{R}(A'', m) \notin [\varepsilon_m \cdot m, \delta_m \cdot m)$ for all $A'' \in [A, A_m)$. From this we see that $[A, A') \subseteq \bigcap_{m\in\mathcal{W}} \mathrm{Bad}(m, \varepsilon_m, \delta_m) = \mathrm{Bad}(\mathcal{W})$. Part (b) follows. □

Lemma 6.1 immediately leads us to the following algorithm.

Algorithm 1. To compute $\mathrm{Bad}(\mathcal{W}) \cap [A, B)$ as a disjoint union of intervals $\bigcup_{I\in\mathscr{I}} I$.
Input: A, B, $\mathcal{W}$, $\underset{\sim}{\varepsilon}$, $\underset{\sim}{\delta}$.

Initialize $\mathscr{I} \leftarrow \varnothing$.
Repeat until $A = B$:

(a) Loop through the elements $m \in \mathcal{W}$ computing $q_m = \mathrm{Q}(A, m)$, $r_m = \mathrm{R}(A, m)$.

(b) If there is some $n \in \mathcal{W}$ such that $\varepsilon_n \cdot n \le r_n < \delta_n \cdot n$ then for any such n set

$$A \leftarrow \min((q_n + \delta_n) \cdot n, B)$$

and go back to (a).

(c) Otherwise, let A' be as in (21). Set $\mathscr{I} \leftarrow \mathscr{I} \cup \{[A, A')\}$ and then $A \leftarrow A'$. Go back to (a).

Output: $\mathscr{I}$.

A heuristic analysis of Algorithm 1 and its running time. Let $x \in [0, M)$ and recall that $\mathrm{R}(x, m) \in [0, m)$. Moreover, $x \in \mathrm{Bad}(m)$ if and only if $\mathrm{R}(x, m) \in [0, m) \setminus [\varepsilon_m \cdot m, \delta_m \cdot m)$. Thus the "probability" that x belongs to $\mathrm{Bad}(m)$ is $1 - (\delta_m - \varepsilon_m)$. Assuming "independence of events" we expect that the total length of intervals produced by Algorithm 1 is

$$(B - A) \cdot \prod_{m \in \mathcal{W}} (1 - \delta_m + \varepsilon_m). \tag{22}$$

To analyze the running time, we shall suppose parameter choices as in Lemma 2.3: namely $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$ for all $m \in \mathcal{W}$. Moreover, we shall suppose that the elements of $m \in \mathcal{W}$ belong to an interval (12) for some large positive L. By the above, the expected total length of the intervals produced by Algorithm 1 is $(B - A) \cdot 0.9^{\#\mathcal{W}}$. Moreover, we suppose that $\mathcal{W}$ is sufficiently large so that the length of the output should be negligible compared to $B - A$; this should mean that step (c) is relatively rare. We will estimate the expected number of times we loop through steps (a), (b). Note that in step (b), A is increased by $0.1 \cdot n - r_n$. The remainder $r_n = \mathrm{R}(A, n)$ belongs to $[0, 0.1 \cdot n)$. We regard the increase as a product $(0.1 - r_n/n) \cdot n$. Treating r_n/n as a random variable uniformly distributed in $[0, 0.1)$ and n as a random variable uniformly distributed in interval (12), we see that the expected increase is $0.05 \cdot (2.63 + 2.92)L/2 = 0.13875 \cdot L$. A standard probability theory argument that we omit tells us that the expected number of times the algorithm loops through steps (a), (b) is roughly

$$\frac{B-A}{0.13875L} \approx \frac{7(B-A)}{L}.$$

We now suppose that K is very large, and we would like to compute the intersection $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ for some set $\mathcal{W}$ where we hope that the hypotheses of Proposition 2.2 and Lemma 2.3 are satisfied. In particular, we take $L = K^{\frac{1}{3}}$. The number of steps should be around $7M/K^{\frac{1}{3}}$. We have to choose $\mathcal{W}$ so that $M = \mathrm{lcm}(\mathcal{W})$ is much larger than K (see (7) and just above it). Thus the number of steps to compute $\mathrm{Bad}(\mathcal{W})$ is much greater than $K^{\frac{2}{3}}$. For $K = \exp(524)$, the expected number of steps is larger than 10^{150}, which makes the computation entirely impractical.

7. A refined approach to computing $\mathrm{Bad}(\mathcal{W})$: The tower

In this section we let $\mathcal{W}$ be a set of positive integers with $M = \mathrm{lcm}(\mathcal{W})$. Let $M_0, M_1, M_2, \ldots, M_r$ be positive integers such that $M_i \mid M_{i+1}$ and $M_r = M$. Write $p_i = M_{i+1}/M_i$. In our later computations the p_i will be primes, but we need not assume that yet. Let

$$\mathcal{W}_i = \{m \in \mathcal{W} : m \mid M_i\}.$$

We suppose that $M_i = \mathrm{lcm}(\mathcal{W}_i)$. Write $\mathcal{U}_i = \mathcal{W}_{i+1} \setminus \mathcal{W}_i$. Recall (Lemmas 4.3 and 4.4) that we have natural surjections $\pi_{M_j,M_i} : [0, M_j) \to [0, M_i)$ whenever $j \geq i$, and that these restrict to give maps (not necessarily surjections) $\mathrm{Bad}(\mathcal{W}_j) \to \mathrm{Bad}(\mathcal{W}_i)$. For ease of notation we shall denote π_{M_j,M_i} simply by $\pi_{j,i}$. We shall refer to the sequence of inclusions (10) as a *tower leading up to* $\mathrm{Bad}(\mathcal{W})$, and use this to compute $\mathrm{Bad}(\mathcal{W})$.

Lemma 7.1. *Let $0 \leq i \leq r-1$. Suppose $\mathscr{I}_i$ is a finite set of disjoint subintervals of $[0, M_i)$ such that*

$$\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i) = \bigcup_{I \in \mathscr{I}_i} I.$$

Then

$$\mathrm{Bad}(\mathcal{W}_{i+1}) \cap [0, M_{i+1}) = \bigcup_{I \in \mathscr{I}_i} \bigcup_{k=0}^{p_i - 1} ((k \cdot M_i + I) \cap \mathrm{Bad}(\mathcal{U}_i)).$$

Proof. This is immediate from Lemmas 4.3 and 4.4. □

Lemma 7.1 immediately leads us to the following algorithm.

Algorithm 2. The following computes a finite set $\mathscr{I} = \mathscr{I}_r$ of subintervals of $[0, M)$ such that $\mathrm{Bad}(\mathcal{W}) \cap [0, M) = \bigcup_{I \in \mathscr{I}} I$.
Input: $\mathcal{W}_0, \ldots, \mathcal{W}_r = \mathcal{W}$, $\underset{\sim}{\varepsilon}$, $\underset{\sim}{\delta}$.

Initialize: $\mathscr{I}_0$ to be the set of disjoint intervals whose union is $\mathrm{Bad}(\mathcal{W}_0) \cap [0, M_0)$, which is computed using Algorithm 1.
Initialize: $i \leftarrow 0$.
Repeat until $i = r$:

(a) $\mathscr{I}_{i+1} \leftarrow \varnothing$.

(b) for $I \in \mathscr{I}_i$ and $k \in \{0, \ldots, p_i - 1\}$, compute, using Algorithm 1, a finite set $\mathscr{I}'$ of subintervals of $[0, M_{i+1})$ such that $(k \cdot M_i + I) \cap \mathrm{Bad}(\mathcal{U}_i) = \bigcup_{I' \in \mathscr{I}'} I'$; let $\mathscr{I}_{i+1} \leftarrow \mathscr{I}_{i+1} \cup \mathscr{I}'$.

(c) $i \leftarrow i + 1$.

Output: $\mathscr{I} = \mathscr{I}_r$.

A heuristic analysis of Algorithm 2 and its running time. We shall suppose, as in Lemma 2.3, that $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$ for all $m \in \mathcal{W}$. Write $n_i = \#\mathcal{W}_i$. We assume that the elements of $\mathcal{W}_i, \mathcal{U}_i$ belong to an interval of the form $\left[\frac{263L}{100}, \frac{292L}{100}\right]$ for some large L. By our previous analysis, we expect that we can compute $\mathscr{I}_0$ in roughly $7M_0/L$ steps. The total length $\ell(\mathscr{I}_0)$ of the intervals in $\mathscr{I}_0$ should roughly be $0.9^{n_0} M_0$. In Step (b) of the algorithm, we will replace each $I \in \mathscr{I}_0$ with p_0 intervals of the same length, and then apply Algorithm 1 to each. Thus we expect that the number of steps to compute $\mathscr{I}_1$ to be roughly

$$\frac{7p_0 \cdot 0.9^{n_0} \cdot M_0}{L} \approx \frac{7M_1 \cdot 0.9^{n_0}}{L}.$$

The total length of the intervals in $\mathscr{I}_1$ should be roughly $M_1 \cdot 0.9^{n_1}$. It is now apparent that the total number of steps should be around

$$(7/L) \cdot (M_0 + M_1 \cdot 0.9^{n_0} + M_2 \cdot 0.9^{n_1} + \cdots + M_r \cdot 0.9^{n_{r-1}}).$$

8. A large computation

Let M^* be the product of all primes $p \leq 167$ that are $\equiv 5 \pmod 6$, and $\mathcal{W}^*$ be as in (11). In this section we compute $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$, using a tower and Algorithm 2. As explained in Section 3, the result of this computation will be reused again and again in Section 9. Let

$$M_0 = 5 \times 11 \times 17 \times 23 \times 29 \times 41 \times 47 \times 53 \times 59 \times 71 \times 83 \times 89,$$

which is the product of the primes < 100 that are $\equiv 5 \pmod 6$. Let

$$M_1 = 101 \cdot M_0, \quad M_2 = 107 \cdot M_1, \quad M_3 = 113 \cdot M_2, \quad M_4 = 131 \cdot M_3,$$
$$M_5 = 137 \cdot M_4, \quad M_6 = 149 \cdot M_5, \quad M^* = M_7 = 167 \cdot M_6.$$

We let

$$\mathcal{W}_i = \{m \mid M_i : 265 \times 10^9 \leq m \leq 290 \times 10^9\}.$$

Thus $\mathcal{W}_0 \subseteq \cdots \subseteq \mathcal{W}_7 = \mathcal{W}^*$. We checked that $M_i = \mathrm{lcm}(\mathcal{W}_i)$. Table 1 gives the cardinalities of the $\mathcal{W}_i$. We use this tower and Algorithm 2 to compute $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$. By our heuristic in the previous section, the number of steps needed for this computation should very roughly be equal to 6.0×10^{10}, which is the sum of the entries of the table's third column. It appears from this estimate that the computation can be done in reasonable time.

We wrote simple implementations of Algorithms 1 and 2 for the computer algebra system Magma [Bosma et al. 1997]. We divided the interval $[0, M_0)$ into 59000 subintervals of equal length and ran our program on each of these intervals $[A_{k-1}, A_k)$ successively computing $\mathrm{Bad}(\mathcal{W}_i) \cap \pi_{i,0}^{-1}([A_{k-1}, A_k))$ for $i = 0, \ldots, 7$. Our computation was distributed over 59 processors (on a 64 core machine with 2500MHz AMD Opteron Processors). Note that

$$\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i) = \bigcup_{k=1}^{59000} \mathrm{Bad}(\mathcal{W}_i) \cap \pi_{i,0}^{-1}([A_{k-1}, A_k));$$

thus our computation gives us a decomposition of $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ as a union of disjoint intervals. The total CPU time for the computation was around 18,300 hours, but as we distributed it over 59 processors, it was over in less than two weeks.

i	$\log_{10}(M_i)$	$n_i = \#\mathcal{W}_i$	$7M_i 0.9^{n_i-1}/10^{11}$
0	18.3	16	1.4×10^9
1	20.3	38	2.6×10^9
2	22.3	83	2.7×10^{10}
3	24.4	149	2.7×10^{10}
4	26.5	250	3.4×10^9
5	28.6	401	1.1×10^7
6	30.8	620	2.0×10^2
7	33.0	911	3.2×10^{-6}

Table 1. The M_i and the $\mathcal{W}_i$ are given at the beginning of Section 8. The third column gives an estimate for the number of steps needed to compute $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ from $\mathrm{Bad}(\mathcal{W}_{i-1}) \cap [0, M_{i-1})$ according to the heuristic analysis at the end of Section 7.

Lemma 8.1. *There are sequences* $(B_j)_{j=1}^{854}$ *and* $(C_j)_{j=1}^{854}$ *contained in* $[0, M^*]$ *such that*

$$B_1 < C_1 < B_2 < C_2 < \cdots < B_{854} < C_{854}$$

and

$$\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*) = \bigcup_{j=1}^{854} [B_j, C_j),$$

with total length $\sum_{j=1}^{854} (C_j - B_j) = 20382195221000\frac{6}{10}$.

Proof. As indicated by Table 2, our computation gives $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$ as a union of 861 intervals disjoint subintervals of $[0, M^*)$. Among these there are 7 pairs of the form $[\alpha, \beta) \cup [\beta, \gamma)$, where the values of β are of the form $\beta' \cdot M^*/59000$ with

$$\beta' = 7375, 14750, 22125, 29500, 36875, 44250, 51625.$$

These subdivisions are clearly a result of our original subdivision of interval $[0, M_0^*)$ into 59000 subintervals of equal length. We simply replace the pairs $[\alpha, \beta) \cup [\beta, \gamma)$ with $[\alpha, \gamma)$ so that $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$ is expressed as a union of 854 intervals. This simplification of course preserves the total length of intervals. □

Remarks and sanity checks. Our computations are done with exact arithmetic. The reader will note by looking back at Algorithms 1 and 2 (and recalling that all $\varepsilon_m = 0$ and $\delta_m = m/10$) that the end points of the intervals encountered will be rationals with denominators that are divisors of 10, except for the A_k appearing in our original subdivision which have denominators that are divisors of 59000. As a

i	$\#\mathscr{I}_i$	$\ell_i = \ell(\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i))$	ℓ_i / M_i	0.9^{n_i}
0	23 458 002	$365\,300\,497\,739\,376\,385\frac{8}{10}$	1.85×10^{-1}	1.85×10^{-1}
1	553 209 618	$3\,625\,384\,986\,862\,035\,664\frac{4}{10}$	1.82×10^{-2}	1.82×10^{-2}
2	1 106 375 245	$3\,313\,998\,145\,602\,553\,709\frac{1}{10}$	1.56×10^{-4}	1.59×10^{-4}
3	209 982 392	$350\,826\,426\,611\,537\,217\frac{1}{10}$	1.46×10^{-7}	1.52×10^{-7}
4	1 062 201	$1\,076\,402\,154\,947\,217\frac{8}{10}$	3.41×10^{-12}	3.64×10^{-12}
5	904	$20\,663\,973\,893\,432\frac{1}{10}$	4.78×10^{-16}	4.48×10^{-19}
6	870	$20\,504\,346\,087\,851\frac{7}{10}$	3.19×10^{-18}	4.27×10^{-29}
7	861	$20\,382\,195\,221\,000\frac{6}{10}$	1.90×10^{-20}	2.07×10^{-42}

Table 2. Some details for the computation described Section 8. The second column gives $\#\mathscr{I}_i$, where $\mathscr{I}_i$ is a disjoint collection of intervals $\bigcup \mathscr{I}_i = \mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$. The third column gives the total length ℓ_i of these intervals. The fourth column gives the ratio ℓ_i / M_i. According to the heuristic at the end of Section 6, this ratio should approximately equal 0.9^{n_i} which is given in the last column (here $n_i = \#\mathcal{W}_i$ as in Table 1). We explain the discrepancy between the last two columns in the remarks on page 2111.

check on our computations, we verify that our results for $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$ are consistent with Proposition 5.1. The set $\mathcal{W}^*$ satisfies

$$\min(\mathcal{W}^*) = 265024970473 \quad \text{and} \quad \max(\mathcal{W}^*) = 289916573827.$$

We take $L = \min(\mathcal{W}^*) \cdot \frac{100}{263}$. It turns out that $L > \max(\mathcal{W}^*) \cdot \frac{100}{292}$. Thus $\mathcal{W}^*$ is contained in the interval (12) for this value of L. Proposition 5.1 yields a total of 103 intervals of the form $(aM^*/q + \psi_k \cdot L, aM^*/q + \Psi_k \cdot L)$ that must be contained in $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$. We checked that each of these is contained in one of the 854 intervals produced by our computation. It is instructive to compare the fourth and fifth columns of Table 2. According to our heuristic, the total length $\ell(\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i))$ should be around $M_i \cdot 0.9^{n_i}$ (with $n_i = \#\mathcal{W}_i$) and therefore we expect the two columns to be roughly the same. From the table, we see that this heuristic is remarkably accurate for $0 \le i \le 4$, and extremely inaccurate for $i \ge 5$. An explanation for this is provided by the ripples. The total length of the intervals contained in $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ produced by Proposition 5.1 is $\approx 24.9L$. Now

$$\frac{24.9L}{M_5} = 5.8 \times 10^{-17}, \quad \frac{24.9L}{M_6} = 4.0 \times 10^{-19}, \quad \frac{24.9L}{M_7} = 2.3 \times 10^{-21},$$

which does provide an explanation for the discrepancy between the two columns. Proposition 5.2 (with $\mathcal{W}^*$ and M^* in place of $\mathcal{W}$ and M) produces 172 intervals with $11 \le q \le 100$ with total length $\approx 17.8L$. We checked that each of these is also contained in one of the 854 intervals produced by our computation.

According to the overall philosophy of Section 5, the set $\mathrm{Bad}(\mathcal{W}^*) \cap [0, M^*)$ should be concentrated in short intervals around rational multiples $(a/q) \cdot M^*$ with q small. To test this, we computed, using continued fractions, the best rational approximation to $(B_i + C_i)/(2M^*)$ with denominator at most 10^{20}, for $1 \le i \le 854$. The largest denominator we found was 42.

The reader is probably wondering, given that we are employing 59 processors, why we have subdivided $[0, M_0)$ into 59, 000 intervals instead of 59 intervals. This was done purely for memory management reasons. A glance at Table 2 will show the reader that there is an explosion of intervals at levels $i = 1, 2, 3$. By dividing $[0, M_0)$ into 59, 000 subintervals, we only need to store roughly $\frac{1}{59000}$-th of the intervals appearing at levels i at any one time per processor, and so only need to store around $\frac{1}{1000}$-th of these intervals in the memory at any one time.

9. Proof of Theorem 1 for $N \ge \left(\frac{9}{10}\right)^{3998} \cdot \exp(524) \approx 4.28 \times 10^{44}$

The reader might at this point find it helpful to review the first paragraph of Section 2 as well as the plan in Section 3. Let $\mathscr{K} = \exp(524)$. In this section we prove Theorem 1 for $N \ge \left(\frac{9}{10}\right)^{3998} \mathscr{K}$. We shall divide the interval $\left(\frac{9}{10}\right)^{3998} \mathscr{K} \le N \le \mathscr{K}$ into subintervals $\left(\frac{9}{10}\right)^{n+1} \mathscr{K} \le N \le \left(\frac{9}{10}\right)^{n} \mathscr{K}$ with $0 \le n \le 3997$. We apply Proposition 2.2 and Lemma 2.3 to prove that all odd integers in the interval $\left(\frac{9}{10}\right)^{n+1} \mathscr{K} \le N \le \left(\frac{9}{10}\right)^{n} \mathscr{K}$ are sums of seven nonnegative cubes.

Lemma 9.1. *Let $0 \le n \le 3997$. Let $K = \left(\frac{9}{10}\right)^{n} \cdot \mathscr{K}$. There exists an integer κ that satisfies*:

(a) *κ is squarefree.*

(b) *$3 \mid \kappa$.*

(c) *$\kappa/3$ is divisible only by primes $q \equiv 5 \pmod 6$ that satisfy $q > 167$.*

(d) *κ belongs to the interval*

$$\frac{263}{265} \cdot \frac{K^{\frac{1}{3}}}{10^{11}} \le \kappa \le \frac{292}{290} \cdot \frac{K^{\frac{1}{3}}}{10^{11}}. \tag{23}$$

Proof. We proved the lemma using a Magma script. Let I_1, I_2 be the lower and upper bounds for κ in (23). If $I_2 < 10^7$ then our script uses brute enumeration of integers in the interval $[I_1, I_2]$ to find a suitable κ. Otherwise, the script takes τ to be a product of consecutive primes $\equiv 5 \pmod 6$ starting with 173 up to a certain bound,

and keeps increasing the bound until $I_2/\tau < 10^7$. It then loops through the integers $I_{\frac{1}{3}}\tau \le \mu \le I_{\frac{2}{3}}\tau$ until it finds one such that $\kappa = 3\mu\tau$ satisfies conditions (a), (b), (c). □

Remark. For $n = 3998$, the interval in (23) is $7481.6\ldots \le \kappa \le 7590.5\ldots$, which is too short for the existence of a suitable κ. This is also the case for most values of n that are ≥ 3998.

Lemma 9.2. *Let $0 \le n \le 3997$ and let κ be as in Lemma 9.1. Let $\mathcal{W}^*$ and M^* be as in Lemma 8.1. Let*

$$\mathcal{W}_0 = \{\kappa \cdot m^* : m^* \in \mathcal{W}^*\} \quad \text{and} \quad M_0 = \operatorname{lcm}(\mathcal{W}_0) = \kappa M^*.$$

Let $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$ for all $m \in \mathcal{W}_0$. Then $m \in \mathcal{W}_0$ satisfy the conditions (i)–(vi) of Section 2, where

$$K_1 = \left(\tfrac{9}{10}\right)^{n+1} \cdot \mathscr{K} \quad \text{and} \quad K_2 = \left(\tfrac{9}{10}\right)^{n} \cdot \mathscr{K}.$$

Moreover,

$$\operatorname{Bad}(\mathcal{W}_0) \cap [0, M_0) = \bigcup_{j=1}^{854} [\kappa \cdot B_j, \kappa \cdot C_j), \tag{24}$$

where the B_j and C_j are as in Lemma 8.1.

Proof. All $m^* \in \mathcal{W}^*$ are squarefree and divisible only by primes $q \le 167$ satisfying $q \equiv 5 \pmod 6$. Thus conditions (i)–(iii) of Section 2 are satisfied by $m \in \mathcal{W}_0$. As we are taking $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$, to verify conditions (iv)–(vi) we may apply Lemma 2.3. For this we need only check that (9) holds for $m \in \mathcal{W}_0$, where $K = K_2$. This immediately follows from (23) and the fact that $\mathcal{W}^* \subset [265 \times 10^9, 290 \times 10^9]$.

Finally, by Lemma 4.2,

$$\operatorname{Bad}(\mathcal{W}_0) \cap [0, M_0) = \kappa \cdot (\operatorname{Bad}(\mathcal{W}^*) \cap [0, M^*)).$$

Lemma 8.1 completes the proof. □

Our Magma script for proving Theorem 1 in the range $K_1 \le N \le K_2$ proceeds as follows. We inductively construct a tower $\mathcal{W}_0 \subset \mathcal{W}_1 \subset \mathcal{W}_2 \subset \cdots$. Observe that

$$\frac{\ell(\operatorname{Bad}(\mathcal{W}_0) \cap [0, M_0))}{M_0} = \frac{\ell(\operatorname{Bad}(\mathcal{W}^*) \cap [0, M^*))}{M^*} \approx 1.90 \times 10^{-20},$$

thus the computation of the previous section has already substantially depleted the interval $[0, M_0)$. Given $\mathcal{W}_i$, and M_i, we let p_i be the smallest prime $\equiv 5 \pmod 6$ that does not divide M_i and let $M_{i+1} = p_i M_i$. The script then writes down positive integers m belonging to the interval (9), such that $m \mid M_{i+1}$ and $3p_i \mid m$. It is not necessary or practical to find all such integers, but we content ourselves with finding around $3\log(p_i)/\log(0.9^{-1})$ of them; we explain this choice shortly. These m will form the set $\mathcal{U}_i$ and we take $\mathcal{W}_{i+1} = \mathcal{W}_i \cup \mathcal{U}_i$. The script then

applies our implementation of Algorithm 2 to compute $\mathrm{Bad}(\mathcal{W}_{i+1}) \cap [0, M_{i+1})$ as a union of disjoint intervals. Our heuristic analysis of Algorithm 2 suggests that $\ell(\mathrm{Bad}(\mathcal{W}_{i+1}) \cap [0, M_{i+1}))$ should roughly equal $p_i \cdot 0.9^{\#\mathcal{U}_i} \cdot \ell(\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i))$. We desire the total length of the intervals to decrease in each step of the tower, so we should require $\#\mathcal{U}_i > \log(p_i)/\log(0.9^{-1})$. Experimentation suggests that requiring $\#\mathcal{U}_i \approx 3\log(p_i)/\log(0.9^{-1})$ provides good control of both the total length of $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ and the number of intervals that make it up. Our script continues to build the tower and compute successive $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ until it finds $\mathcal{W} = \mathcal{W}_i$ and $M = M_i$ that satisfy (7) for some set of rationals $\mathfrak{S} \subset [0, 1]$ with denominators bounded by $\sqrt[3]{M/2K}$. Specifically, once $M_i > 2K$, for each of the disjoint intervals $[\alpha, \beta)$ that make up $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$, the script uses continued fractions to compute the best rational approximation a/q to $(\alpha + \beta)/2M$ with $q \le \sqrt[3]{M/2K}$, and then checks whether

$$[\alpha, \beta) \subseteq (aM/q - \sqrt[3]{M/16}/q, aM/q + \sqrt[3]{M/16}/q).$$

The script continues constructing the tower until this criterion is satisfied for all the intervals making up $\mathrm{Bad}(\mathcal{W})$. It then follows from Proposition 2.2 that all odd integers in the range $\mathscr{K} \cdot \left(\frac{9}{10}\right)^{n+1} \le N \le \mathscr{K} \cdot \left(\frac{9}{10}\right)^{n}$ are sums of seven nonnegative cubes. We again distributed the computation among 59 processors on the aforementioned machine, with each processor handling an appropriate portion of the range $0 \le n \le 3997$. The script succeeded in finding an appropriate $\mathcal{W}$ for all n in this range. The entire CPU time was around 10,000 hours, but as the computation was distributed among 59 processors the actual time was around 7 days.

We give more details for the case $n = 0$. Thus $K = \mathscr{K} = \exp(524)$, and we would like to show, using Proposition 2.2 that all odd integers $9K/10 \le N \le K$ are sums of seven nonnegative cubes. The routine described in the proof of Lemma 9.1 gives the following suitable value for κ:

$$\kappa = 3 \times 173 \times 179 \times 191 \times 197 \times 227 \times 233 \times 239 \times 251 \times 257 \times 263 \times 269 \times 281 \times 293 \times 311 \times 317 \times 347 \times 353 \times 359 \times 383 \times 389 \times 401 \times 419 \times 431 \times 443 \times 207443.$$

Table 3 gives some of the details for the computation. We take $\mathcal{W} = \mathcal{W}_{48}$. Then $\#\mathcal{W} = \#\mathcal{W}_0 + \sum \#\mathcal{U}_i = 9943$, and

$$\ell(\mathrm{Bad}(\mathcal{W}) \cap [0, M)) = 1245937137395549638824015714140403151401411370898968055175937887691670913319978\tfrac{1}{2} \approx 1.25 \times 10^{78}.$$

In comparison,

$$M = M_{48} \approx 1.64 \times 10^{235} \quad \text{and} \quad K = 3.72 \times 10^{227}.$$

i	p_{i-1}	$\#\mathcal{U}_{i-1}$	N	ℓ_i/M_i	i	p_{i-1}	$\#\mathcal{U}_{i-1}$	N	ℓ_i/M_i
0	–	–	854	1.90×10^{-20}	25	761	189	729	9.93×10^{-90}
1	449	174	775	3.73×10^{-23}	26	773	190	729	1.28×10^{-92}
2	461	175	745	7.94×10^{-26}	27	797	191	729	1.61×10^{-95}
3	467	176	740	1.70×10^{-28}	28	809	191	729	1.99×10^{-98}
4	479	176	735	3.54×10^{-31}	29	821	192	729	2.43×10^{-101}
5	491	177	732	7.20×10^{-34}	30	827	192	729	2.93×10^{-104}
6	503	178	730	1.42×10^{-36}	31	839	192	729	3.50×10^{-107}
7	509	178	730	2.80×10^{-39}	32	857	193	729	4.08×10^{-110}
8	521	179	730	5.38×10^{-42}	33	863	193	729	4.73×10^{-113}
9	557	181	730	9.65×10^{-45}	34	881	194	729	5.36×10^{-116}
10	563	181	731	1.71×10^{-47}	35	887	194	729	6.05×10^{-119}
11	569	181	730	3.01×10^{-50}	36	911	195	729	6.64×10^{-122}
12	587	182	729	5.13×10^{-53}	37	929	195	729	7.14×10^{-125}
13	593	182	729	8.64×10^{-56}	38	941	195	729	7.59×10^{-128}
14	599	183	729	1.44×10^{-58}	39	947	196	729	8.02×10^{-131}
15	617	183	729	2.34×10^{-61}	40	953	196	729	8.41×10^{-134}
16	641	185	729	3.64×10^{-64}	41	971	196	729	8.66×10^{-137}
17	647	185	729	5.63×10^{-67}	42	977	197	729	8.87×10^{-140}
18	653	185	729	8.62×10^{-70}	43	983	197	729	9.02×10^{-143}
19	659	185	729	1.31×10^{-72}	44	1013	198	729	8.91×10^{-146}
20	677	186	729	1.93×10^{-75}	45	1019	198	729	8.74×10^{-149}
21	683	186	729	2.83×10^{-78}	46	1031	198	729	8.48×10^{-152}
22	701	187	729	4.04×10^{-81}	47	1049	199	729	8.08×10^{-155}
23	719	188	729	5.61×10^{-84}	48	1061	199	729	7.62×10^{-158}
24	743	189	729	7.55×10^{-87}					

Table 3. details for the computation for the case $n = 0$. For each $i \geq 1$, our script computes $\mathrm{Bad}(\mathcal{W}_i)$ as a disjoint union of subintervals of $[0, M_i)$. The number of intervals, N, is given in the fourth column. The fifth column gives, to 3 significant figures, the ratio ℓ_i/M_i where $\ell_i = \ell(\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i))$.

The set $\mathfrak{S}$ as in (7) turns out be precisely the set of 171 rationals $a/q \in [0, 1]$ with denominators q belonging to

$$1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 12, 13, 14, 15, 16, 18, 19, 21, 24, 26, 28, 30, 36, 42.$$

As a check on our results, we apply Proposition 5.2 to show that there is an interval close to $(a/42) \cdot M$ for $1 \leq a \leq 41$ with $\gcd(a, 42) = 1$. Our $\mathcal{W}$ and M satisfy the hypotheses of Section 5 with $L = K^{\frac{1}{3}}$. Note that $3 \mid m \mid M$ for all $m \in \mathcal{W}$. As M is squarefree, we have $3 \nmid (M/m)$. Moreover, all the prime divisors of $M/3$

are $\equiv 5 \pmod 6$. It follows that $\gcd(aM/m, 42) = 1$ for all $m \in \mathcal{W}$. Let $q = 42$ and $s = d = 5$ in Proposition 5.2; hypothesis (16) is trivially satisfied. Now $s+1, \dots, s+d$ are the integers 6, 7, 8, 9, 10 and none of these are coprime to 42. Thus condition (17) is also satisfied. By Proposition 5.2, for each $1 \le a \le 41$ with $\gcd(a, 42) = 1$ we have

$$\left(\tfrac{a}{42}M - \tfrac{4471}{10500} \cdot K^{\frac{1}{3}}, \tfrac{a}{42}M - \tfrac{73}{210} \cdot K^{\frac{1}{3}}\right) \subseteq \mathrm{Bad}(\mathcal{W}). \tag{25}$$

One of the 729 intervals that make up $\mathrm{Bad}(\mathcal{W})$ is $[u, v)$ where the end points u, v are

$$u = 389517364042358487471334903242152096024666487365329397553740030752245815701505789686613824871153976672579237296943737371206769063932017310777324617938079775100516093623104146032249096179399599141014593742168620464205667747229312339206610/3,$$

$$v = 389517364042358487471334903242152096024666487365329397553740030752245815701505789686613824871153976672579237296943737371206769063932017310777324617938079775101078697615391607190077739469387928238665618669912989320140106379011502569660,$$

and we checked that the interval in (25) with $a = 1$ is contained in $[u, v)$. It is also interesting to note how close the two intervals are in length: the ratio of the lengths of the two intervals is

$$\frac{\left(\frac{4471}{10500} - \frac{73}{210}\right) \cdot K^{\frac{1}{3}}}{v - u} \approx 0.9994$$

which illustrates how remarkably accurate our Proposition 5.2 is.

10. Completing the proof of Theorem 1

It remains to apply Proposition 2.2 to the intervals $\left(\frac{9}{10}\right)^{n+1}\mathcal{K} \le N \le \left(\frac{9}{10}\right)^{n}\mathcal{K}$ with $3998 \le n \le 4226$. We write $K = K_2 = \left(\frac{9}{10}\right)^{n}\mathcal{K}$ and $K_1 = \left(\frac{9}{10}\right)^{n+1}\mathcal{K}$. It is no longer practical to use the choices in Lemma 2.3 as the interval in (9) is too short to contain many squarefree m whose prime divisors are 3 and small primes $\equiv 5 \pmod 6$. The interval in (9) is a result of imposing the uniform choices $\varepsilon_m = 0$ and $\delta_m = \frac{1}{10}$. Instead we consider integers m satisfying conditions (i)–(iii) of Section 2 but belonging to the (much larger) interval

$$\tfrac{12}{5}K^{\frac{1}{3}} \le m \le \tfrac{16}{5}K^{\frac{1}{3}}. \tag{26}$$

For each such m we take $\varepsilon_m = \varepsilon'/1000$ and $\delta_m = \delta'/1000$ where ε', δ' are integers with ε' and δ' respectively as small and as large as possible such that the conditions (v), (vi) of Section 2 are satisfied. We only keep those values of m for which

$$0 \le \varepsilon_m < \delta_m \le 1 \quad \text{and} \quad \delta_m - \varepsilon_m \ge \tfrac{1}{20}; \tag{27}$$

an elementary though lengthy analysis in fact shows that the inequalities in (27) together with (v) and (vi) force m to belong to the interval (26). Note that the set $\mathrm{Bad}(m, \varepsilon_m, \delta_m)$ has "relative density" $1-\delta_m+\varepsilon_m$ in $\mathbb{R}$; the restriction $\delta_m - \varepsilon_m \geq \frac{1}{20}$ ensures that this relative density is not too close to 1, and that therefore m makes a significant contribution to depleting the intervals in Algorithms 1 and 2.

We choose a prime $q \equiv 5 \pmod 6$, depending on K, and let

$$M_0 = 3 \cdot 5 \cdot 11 \cdots q,$$

which is the product of 3 and the primes $\leq q$ that are $\equiv 5 \pmod 6$. Let $\mathcal{W}_0$ be the set of positive integers dividing M_0 and satisfying the above conditions. We found experimentally that for each n in the above range it is always possible to choose q so that

$$M_0 = \mathrm{lcm}(\mathcal{W}_0), \qquad \prod_{m \in \mathcal{W}_0} (1 - \delta_m + \varepsilon_m) \leq \tfrac{1}{5}, \quad \text{and} \quad \log_{10}(M_0/K^{\frac{1}{3}}) \leq 7.5.$$

The inequality $\prod_{m\in\mathcal{W}_0}(1 - \delta_m + \varepsilon_m) \leq \frac{1}{5}$ indicates that $\ell(\mathrm{Bad}(\mathcal{W}_0) \cap [0, M_0))$ should heuristically be at most $M_0/5$ which means that this is a good first step at depleting the interval $[0, M_0)$. The other inequality indicates that we can compute $\mathrm{Bad}(\mathcal{W}_0) \cap [0, M_0)$ in a reasonable number of steps, according to the heuristic, following Algorithm 1. We let p_0 be the first prime $\equiv 5 \pmod 6$ that is $> q$, and p_1 be the next such prime and so on. We let $M_{i+1} = p_i M_i$ and construct a tower as before. We stop once $\mathrm{Bad}(\mathcal{W}_i) \cap [0, M_i)$ satisfies the criterion of Proposition 2.2. Our Magma script succeeded in doing this for all n in the range $3998 \leq n \leq 4226$. The total CPU time was around 2750 hours, but the computation was spread over 59 processors so the actual time was less than 2 days.

We give a few of details for the computation for the value $n = 4226$. The final M is the product of 3 and the primes $p \equiv 5 \pmod 6$ that are ≤ 227. The final $\mathcal{W}$ has 8083 elements. It turns out that $\mathrm{Bad}(\mathcal{W}) \cap [0, M)$ consists of 305 intervals and that $\ell(\mathrm{Bad}(\mathcal{W}) \cap [0, M))/M \approx 2.24 \times 10^{-32}$.

Acknowledgements

The programs that accompany this paper are available at http://tinyurl.com/zlaeweo. It is a pleasure to thank Alex Bartel, Tim Browning, John Cremona, Roger Heath-Brown and Trevor Wooley for stimulating discussions. We are grateful to the referee for a careful reading of the paper and for pointing out several corrections.

References

[Baer 1913] W. S. Baer, *Beiträge zum Waringschen Problem*, dissertation, University of Göttingen, 1913. Zbl

[Bertault et al. 1999] F. Bertault, O. Ramaré, and P. Zimmermann, "On sums of seven cubes", *Math. Comp.* **68**:227 (1999), 1303–1310. MR Zbl

[Boklan and Elkies 2009] K. D. Boklan and N. D. Elkies, "Every multiple of 4 except 212, 364, 420, and 428 is the sum of seven cubes", preprint, 2009. arXiv

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl

[Deshouillers et al. 2000] J.-M. Deshouillers, F. Hennecart, and B. Landreau, "7 373 170 279 850", *Math. Comp.* **69**:229 (2000), 421–439. MR

[Dickson 1927] L. E. Dickson, "Extensions of Waring's Theorem on Nine Cubes", *Amer. Math. Monthly* **34**:4 (1927), 177–183. MR Zbl

[Dickson 1939] L. E. Dickson, "All integers except 23 and 239 are sums of eight cubes", *Bull. Amer. Math. Soc.* **45** (1939), 588–591. MR Zbl

[Elkies 2010] N. D. Elkies, "Every even number greater than 454 is the sum of seven cubes", preprint, 2010. arXiv

[Jacobi 1851] C. G. J. Jacobi, "Über die Zusammensetzung der Zahlen aus ganzen positiven Cuben; nebst einer Tabelle für die kleinste Cubenanzahl, aus welcher jede Zahl bis 12000 zusammengesetzt werden kann", *J. Reine Angew. Math.* **42** (1851).

[Kempner 1912] A. Kempner, "Bemerkungen zum Waringschen Problem", *Math. Ann.* **72**:3 (1912), 387–399. MR Zbl

[Landau 1908] E. Landau, "Über eine Anwendung der Primzahltheorie auf das Waringsche Problem in der elementaren Zahlentheorie", *Math. Ann.* **66**:1 (1908), 102–105. MR

[Linnik 1943] U. V. Linnik, "On the representation of large numbers as sums of seven cubes", *Mat. Sbornik* (*N.S.*) **12(54)**:2 (1943), 218–224. MR Zbl

[Maillet 1895] E. Maillet, "Sur la décomposition d'un nombre entier en une somme de cubes d'entiers positifs", *Assoc. Franç. Bordeaux Notes Mém.* **24** (1895), 242–247. JFM

[McCurley 1984] K. S. McCurley, "An effective seven cube theorem", *J. Number Theory* **19**:2 (1984), 176–183. MR Zbl

[Ramaré 2005] O. Ramaré, "An explicit seven cube theorem", *Acta Arith.* **118**:4 (2005), 375–382. MR Zbl

[Ramaré 2007] O. Ramaré, "An explicit result of the sum of seven cubes", *Manuscripta Math.* **124**:1 (2007), 59–75. MR Zbl

[Romani 1982] F. Romani, "Computations concerning Waring's problem for cubes", *Calcolo* **19**:4 (1982), 415–431. MR Zbl

[von Sterneck 1903] R. D. von Sterneck, "Über die kleinste Anzahl Kuben, aus welchen jede Zahl bis 40000 zusammengesetzt werden kann", *Akad. Wiss. Wien, Math.-Natur. Kl. Sitz.* (*IIa*) **112** (1903), 1627–1666. Zbl

[Watson 1951] G. L. Watson, "A proof of the seven cube theorem", *J. London Math. Soc.* (2) **26** (1951), 153–156. MR Zbl

[Wieferich 1908] A. Wieferich, "Beweis des Satzes, daß sich eine jede ganze Zahl als Summe von höchstens neun positiven Kuben darstellen läßt", *Math. Ann.* **66**:1 (1908), 95–101. MR JFM

Communicated by Roger Heath-Brown
Received 2016-01-06 Revised 2016-08-21 Accepted 2016-09-23

samir.siksek@gmail.com *Mathematics Institute, University of Warwick, Coventry, CV4 7AL, United Kingdom*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2121

Constructible isocrystals

Bernard Le Stum

We introduce a new category of coefficients for p-adic cohomology called constructible isocrystals. Conjecturally, the category of constructible isocrystals endowed with a Frobenius structure is equivalent to the category of perverse holonomic arithmetic $\mathcal{D}$-modules. We prove here that a constructible isocrystal is completely determined by any of its geometric realizations.

Introduction

The relation between topological invariants and differential invariants of a manifold is always fascinating. We may first recall de Rham's theorem, which implies the existence of an isomorphism

$$\mathrm{H}^i_{\mathrm{dR}}(V) \simeq \mathrm{Hom}(\mathrm{H}_i(V), \mathbb{C})$$

on any complex analytic manifold V. The nonabelian version is an equivalence of categories

$$\mathrm{MIC}(V) \simeq \mathrm{Rep}_{\mathbb{C}}(\pi_1(V, x))$$

between coherent modules endowed with an integrable connection and finite-dimensional representations of the fundamental group. The analogous result holds on a smooth complex algebraic variety X if we stick to regular connections (see

MSC2010: 14F30.
Keywords: constructible isocrystal, overconvergent isocrystal, rigid cohomology, p-adic cohomology, module with connection.

[Deligne 1970] or Bernard Malgrange's lecture in [Borel et al. 1987]). It has been generalized by Masaki Kashiwara [1984] to an equivalence

$$\mathrm{D}^{\mathrm{b}}_{\mathrm{reg,hol}}(X) \simeq \mathrm{D}^{\mathrm{b}}_{\mathrm{cons}}(X^{\mathrm{an}})$$

between the categories of bounded complexes of $\mathcal{D}_X$-modules with regular holonomic cohomology and bounded complexes of $\mathbb{C}_{X^{\mathrm{an}}}$-modules with constructible cohomology.

Both categories come with a so-called t-structure but these t-structures do not correspond under this equivalence. Actually, they define a new t-structure on the other side that may be called *perverse*. The notion of a perverse sheaf on X^{an} has been studied for some time now (see [Borel et al. 1987], for example). On the $\mathcal{D}$-module side, however, this notion only appeared in the recent article [Kashiwara 2004], even if he does not give it a name (we call it perverse but it might as well be called constructible; see [Abe 2013]). In any case, he shows that the perverse t-structure on $\mathrm{D}^{\mathrm{b}}_{\mathrm{reg,hol}}(X)$ is given by

$$\begin{cases} \mathrm{D}^{\leq 0} : \operatorname{codim} \operatorname{supp} \mathcal{H}^n(\mathcal{F}^\bullet) \geq n & \text{for } n \geq 0, \\ \mathrm{D}^{\geq 0} : \mathcal{H}^n_Z(\mathcal{F}^\bullet) = 0 & \text{for } n < \operatorname{codim} Z. \end{cases}$$

In particular, if we call *perverse* a complex of $\mathcal{D}_X$-modules satisfying both conditions, there exists an equivalence of categories

$$\mathrm{D}^{\mathrm{perv}}_{\mathrm{reg,hol}}(X) \simeq \mathrm{Cons}(X^{\mathrm{an}})$$

between the categories of perverse (complexes of) $\mathcal{D}_X$-modules with regular holonomic cohomology and constructible $\mathbb{C}_{X^{\mathrm{an}}}$-modules.

In a handwritten note called "Cristaux discontinus", Pierre Deligne gave an algebraic interpretation of the right-hand side of this equivalence. More precisely, he introduces the notion of a constructible procoherent crystal and proves an equivalence

$$\mathrm{Cons}_{\mathrm{reg,procoh}}(X/\mathbb{C}) \simeq \mathrm{Cons}(X^{\mathrm{an}})$$

between the categories of regular constructible procoherent crystals and constructible $\mathbb{C}_{X^{\mathrm{an}}}$-modules.

By composition, we obtain what may be called the *Deligne–Kashiwara correspondence*:

$$\mathrm{Cons}_{\mathrm{reg,procoh}}(X/\mathbb{C}) \simeq \mathrm{D}^{\mathrm{perv}}_{\mathrm{reg,hol}}(X).$$

It would be quite interesting to give an algebraic construction of this equivalence but this is not our purpose here. Actually, we would like to describe an arithmetic analog.

Let K be a p-adic field with discrete valuation ring $\mathcal{V}$ and perfect residue field k. Let $X \hookrightarrow P$ be a locally closed embedding of an algebraic k-variety into a formal $\mathcal{V}$-scheme. Assume for the moment that P is smooth and quasicompact, and that

the locus of X at infinity inside P has the form $D \cap \overline{X}$, where D is a divisor in P. We may consider the category $\mathrm{D}^{\mathrm{b}}(X \subset P/K)$ of bounded complexes of $\mathcal{D}_P^{\dagger}({}^{\dagger}D)_{\mathbb{Q}}$-modules on P with support on $\overline{X}$ (see [Berthelot 2002], for example). On the other hand, we may also consider the category of overconvergent isocrystals on $(X \subset P/K)$. Daniel Caro proved [2009] that there exists a fully faithful functor

$$\mathrm{sp}_+ : \mathrm{Isoc}^{\dagger}_{\mathrm{coh}}(X \subset P/K) \to \mathrm{D}^{\mathrm{b}}_{\mathrm{coh}}(X \subset P/K)$$

(the index coh simply means overconvergent isocrystals in Berthelot's sense — see below). This is the first step towards an overconvergent Deligne–Kashiwara correspondence. Note that this construction is extended to a slightly more general situation by Tomoyuki Abe and Caro [2013] and was already known to Pierre Berthelot [1996b, Proposition 4.4.3] in the case $\overline{X} = P_k$.

In [Le Stum 2014], we defined a category, which we may denote $\mathrm{MIC}^{\dagger}_{\mathrm{cons}}(P/K)$, of convergent constructible ∇-modules on P_K when P is a geometrically connected smooth proper curve over $\mathcal{V}$, as well as a category $\mathrm{D}^{\mathrm{perv}}(P/K)$ of perverse (complexes of) $\mathcal{D}^{\dagger}_{P\mathbb{Q}}$-modules on P, and we built a functor

$$\mathrm{R}\widetilde{\mathrm{sp}}_* : \mathrm{MIC}^{\dagger}_{\mathrm{cons}}(P/K) \to \mathrm{D}^{\mathrm{perv}}_{\mathrm{coh}}(P/K).$$

Actually, we proved the overconvergent Deligne–Kashiwara correspondence in this situation: this functor induces an equivalence of categories

$$\mathrm{R}\widetilde{\mathrm{sp}}_* : F\text{-}\mathrm{MIC}^{\dagger}_{\mathrm{cons}}(P/K) \simeq F\text{-}\mathrm{D}^{\mathrm{perv}}_{\mathrm{hol}}(P/K)$$

between (convergent) constructible F-∇-modules on P_K and perverse holonomic F-$\mathcal{D}^{\dagger}_{P\mathbb{Q}}$-modules on P. Note that this is compatible with Caro's sp_+ functor.

In order to extend this theorem to a higher dimension, it is necessary to develop a general theory of *constructible* (*overconvergent*) *isocrystals*. One could try to mimic Berthelot's original definition and let $\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(X \subset Y \subset P/K)$ be the category of $j_X^{\dagger}\mathcal{O}_{]Y[}$-modules $\mathcal{F}$ endowed with an overconvergent connection which are only "constructible" and not necessarily coherent (here X is open in Y and Y is closed in P). It means that there exists a locally finite covering of X by locally closed subvarieties Z such that $j_Z^{\dagger}\mathcal{F}$ is a coherent $j_Z^{\dagger}\mathcal{O}_{]Y[}$-module. It would then be necessary to show that the definition is essentially independent of P as long as P is smooth and Y proper, and that they glue when there does not exist any global geometric realization.

We choose here an equivalent but different approach with built-in functoriality. I introduced in [Le Stum 2011] the overconvergent site of the algebraic variety X and showed that we can identify the category of locally finitely presented modules on this site with the category of overconvergent isocrystals in the sense of Berthelot. Actually, we can define a broader category of overconvergent isocrystals (without any finiteness condition) and call an overconvergent isocrystal E *constructible* when

there exists a locally finite covering of X by locally closed subvarieties Y such that $E_{|Y}$ is locally finitely presented. Note that K may be any nontrivial complete ultrametric field and that there exists a relative theory (over some base O). We denote by $\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(X/O)$ the category of constructible overconvergent isocrystals on X/O. We expect a "Grothendieck's six operations formalism" for overconvergent F-isocrystals and, more generally, that all usual properties of constructible coefficients will hold in our context.

As usual, when we are given a crystalline solution to a coefficient problem, it is necessary to be able to give an interpretation in terms of modules with an integrable connection. Here, one may define a category $\mathrm{MIC}^{\dagger}_{\mathrm{cons}}(X, V/O)$ of constructible modules endowed with an overconvergent connection on any "geometric realization" V of X/O, as in Berthelot's approach. We will prove (Theorem 4.12 below) that, when $\mathrm{Char}(K) = 0$, there exists an equivalence of categories

$$\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(X/O) \simeq \mathrm{MIC}^{\dagger}_{\mathrm{cons}}(X, V/O).$$

As a corollary, we obtain that the later category is essentially independent of the choice of the geometric realization (and that they glue when there does not exist such a geometric realization). Note that this applies in particular to the case of the curve P above which "is" a geometric realization of P_k so that

$$\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(P_k/K) = \mathrm{MIC}^{\dagger}_{\mathrm{cons}}(P/K).$$

In Section 1, we briefly present the overconvergent site and review some material that will be needed afterwards. In Section 2, we study some functors between overconvergent sites that are associated to locally closed embeddings. We do a little more that what is necessary for the study of constructible isocrystals, hoping that this will be useful in the future. In Section 3, we introduce overconvergent isocrystals and explain how one can construct and deconstruct them. In Section 4, we show that constructible isocrystals may be interpreted in terms of modules with integrable connections.

Notation and conventions

Throughout this article, K denotes a nontrivial complete ultrametric field with valuation ring $\mathcal{V}$ and residue field k.

An *algebraic variety* over k is a scheme over k that admits a locally finite covering by schemes of finite type over k. A *formal scheme* over $\mathcal{V}$ always admits a locally finite covering by π-adic formal schemes of finite presentation over $\mathcal{V}$. An *analytic variety* over K is a strictly analytic K-space in the sense of [Berkovich 1993], for example. We will use the letters $X, Y, Z, U, C, D, \ldots$ to denote algebraic varieties over k, P, Q, S for formal schemes over $\mathcal{V}$ and V, W, O for analytic varieties over K.

An analytic variety over K is said to be *good* if it is locally affinoid. This is the case, for example, if V is affinoid, proper or algebraic, or more generally if V is an open subset of such a variety. Note that in Berkovich's original definition [1990] all analytic varieties were good.

As usual, we will write $\mathbb{A}^1$ and $\mathbb{P}^1$ for the affine and projective lines. We will also use $\mathbb{D}(0, 1^{\pm})$ for the open or closed disc of radius 1.

1. The overconvergent site

We briefly recall the definition of the overconvergent site from [Le Stum 2011]. An object is made of

(1) a locally closed embedding $X \hookrightarrow P$ of an algebraic variety (over k) into a formal scheme (over $\mathcal{V}$) and

(2) a morphism $\lambda : V \to P_K$ of analytic varieties (over K).

We denote this object by $X \subset P \xleftarrow{\mathrm{sp}} P_K \leftarrow V$ and call it an *overconvergent variety*. Here, sp denotes the *specialization* map and we also introduce the notion of a *tube* of X in V:

$$]X[_V := \lambda^{-1}(\mathrm{sp}^{-1}(X)).$$

We call the overconvergent variety *good* if any point of $]X[_V$ has an affinoid neighborhood in V. It makes it simpler to assume from the beginning that all overconvergent varieties are good since the important theorems can only hold for those (and bad overconvergent varieties play no role in the theory). But, on the other hand, most constructions can be carried out without this assumption.

We define a *formal morphism* between overconvergent varieties as a triple of compatible morphisms:

$$\begin{array}{ccccccc} X' & \hookrightarrow & P' & \leftarrow & P'_K & \leftarrow & V' \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle v_K} & & \downarrow{\scriptstyle u} \\ X & \hookrightarrow & P & \leftarrow & P_K & \leftarrow & V \end{array}$$

Such a formal morphism induces a continuous map

$$]f[_u :]X'[_{V'} \to]X[_V$$

between the tubes.

Actually, the notion of a formal morphism is too rigid to reflect the true nature of the algebraic variety X and it is necessary to make invertible what we call a *strict neighborhood*, which we define now: it is a formal morphism as above such that f is an isomorphism $X' \simeq X$ and u is an open immersion that induces an isomorphism between the tubes $]X'[_{V'} \simeq]X[_V$. Formal morphisms admit a calculus

of right fractions with respect to strict neighborhoods, and the quotient category is the *overconvergent site* $\mathrm{An}^{\dagger}_{/\mathcal{V}}$. Roughly speaking, we allow the replacement of V by any neighborhood of $]X[_V$ in V and we make the role of P secondary (only existence is required).

Since we call our category a site, we must endow it with a topology which is actually defined by the pretopology of families of formal morphisms

$$\begin{array}{ccccccc} X & \hookrightarrow & P_i & \leftarrow & P_{iK} & \leftarrow & V_i \\ \| & & \downarrow{\scriptstyle v_i} & & \downarrow{\scriptstyle v_{iK}} & & \downarrow \\ X & \hookrightarrow & P & \leftarrow & P_K & \leftarrow & V \end{array}$$

in which V_i is open in V and $]X[_V \subset \bigcup V_i$ (this is a *standard* site).

Since the formal scheme plays a very loose role in the theory, we usually denote by (X, V) an overconvergent variety and write (f, u) for a morphism.

We use the general formalism of *restricted* category (also called *localized* or *comma* or *slice* category) to define relative overconvergent sites. First of all, we define an *overconvergent presheaf* as a presheaf (of sets) T on $\mathrm{An}^{\dagger}_{/\mathcal{V}}$. If we are given an overconvergent presheaf T, we may consider the restricted site $\mathrm{An}^{\dagger}_{/T}$. An object is a section s of T on some overconvergent variety (X, V) but we like to see s as a morphism from (the presheaf represented by) (X, V) to T. We will then say that (X, V) is a (*overconvergent*) *variety over* T. A morphism between varieties over T is just a morphism of overconvergent varieties which is compatible with the given sections. The above pretopology is still a pretopology on $\mathrm{An}^{\dagger}_{/T}$ and we denote by $T_{\mathrm{An}^{\dagger}}$ the corresponding topos. As explained by David Zureick-Brown [2010; 2014], one may as well replace $\mathrm{An}^{\dagger}_{/T}$ by any fibered category over $\mathrm{An}^{\dagger}_{/\mathcal{V}}$. This is necessary if one wishes to work with algebraic stacks instead of algebraic varieties.

As a first example, we can apply our construction to the case of a representable sheaf $T := (X, V)$. Another very important case is the following: we are given an overconvergent variety (C, O) and an algebraic variety X over C. Then, we define the overconvergent sheaf X/O as follows: a section of X/O is a variety (X', V') over (C, O) with a given factorization $X' \to X \to C$ (this definition extends immediately to algebraic spaces — or even algebraic stacks if one is ready to work with fibered categories). Alternatively, if we are actually given a variety (X, V) over (C, O), we may also consider the overconvergent presheaf X_V/O: a section is a variety (X', V') over (C, O) with a given factorization $X' \to X \to C$ which extends to *some* factorization $(X', V') \to (X, V) \to (C, O)$. Note that we only require the *existence* of the second factorization. In other words, X_V/O is the image presheaf of the natural map $(X, V) \to X/O$. An important theorem (more precisely Corollary 2.5.12 in [Le Stum 2011]) states that, if we work only with *good* overconvergent varieties, then there exists an isomorphism of topos

$(X_V/O)_{\mathrm{An}^\dagger} \simeq (X/O)_{\mathrm{An}^\dagger}$ when we start from a *geometric* situation

$$\begin{array}{ccccccc} X & \hookrightarrow & P & \leftarrow & P_K & \leftarrow & V \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle v_K} & & \downarrow{\scriptstyle u} \\ C & \hookrightarrow & S & \leftarrow & S_K & \leftarrow & O \end{array} \tag{1}$$

with P proper and smooth around X over S and V a neighborhood of the tube of X in $P_K \times_{S_K} O$ (and (C, O) is good).

If we are given a morphism of overconvergent presheaves $v : T' \to T$, we will also say that T' is a (*overconvergent*) *presheaf over* T. It will induce a morphism of topos $v_{\mathrm{An}^\dagger} : T'_{\mathrm{An}^\dagger} \to T_{\mathrm{An}^\dagger}$. We will often drop the index $\mathrm{An}^\dagger$ and keep writing v instead of $v_{\mathrm{An}^\dagger}$. Also, we will usually write the inverse image of a sheaf $\mathcal{F}$ as $\mathcal{F}_{|T'}$ when there is no ambiguity about v. Note that there will exist a triple of adjoint functors $v_!$, v^{-1}, v_* with $v_!$ exact.

For example, any morphism $(f, u) : (Y, W) \to (X, V)$ of overconvergent varieties will give rise to a morphism of topos

$$(f, u)_{\mathrm{An}^\dagger} : (Y, W)_{\mathrm{An}^\dagger} \to (X, V)_{\mathrm{An}^\dagger}.$$

It will also induce a morphism of overconvergent presheaves $f_u : Y_W/O \to X_V/O$ giving rise to a morphism of topos

$$f_{u\mathrm{An}^\dagger} : (Y_W/O)_{\mathrm{An}^\dagger} \to (X_V/O)_{\mathrm{An}^\dagger}.$$

Finally, if (C, O) is an overconvergent variety, then any morphism $f : Y \to X$ of algebraic varieties over C induces a morphism of overconvergent presheaves $f : Y/O \to X/O$ giving rise to a morphism of topos

$$f_{\mathrm{An}^\dagger} : (Y/O)_{\mathrm{An}^\dagger} \to (X/O)_{\mathrm{An}^\dagger}.$$

If we are given an overconvergent variety (X, V), there exists a *realization* map (morphism of topos)

$$(X, V)_{\mathrm{An}^\dagger} \xrightarrow{\varphi}]X[_{V\mathrm{an}}, \quad (X, V') \leftarrow\!\shortmid]X[_{V'},$$

where $]X[_{V\mathrm{an}}$ denotes the category of sheaves (of sets) on the analytic variety $]X[_V$ (which has a section ψ). Now, if T is any overconvergent presheaf and (X, V) is a variety over T, then there exists a canonical morphism $(X, V) \to T$. Therefore, if $\mathcal{F}$ is a sheaf on T, we may consider its restriction $\mathcal{F}_{|(X,V)}$, which is a sheaf on (X, V). We define the *realization* of $\mathcal{F}$ on (X, V) as

$$\mathcal{F}_{X,V} := \varphi_{V*}(\mathcal{F}_{|(X,V)})$$

(we shall simply write $\mathcal{F}_V$ in practice unless we want to emphasize the role of X). As one might expect, the sheaf $\mathcal{F}$ is completely determined by its realizations $\mathcal{F}_V$

and the transition morphisms $]f[_u^{-1}\mathcal{F}_V \to \mathcal{F}_{V'}$ obtained by functoriality whenever $(f,u):(X',V')\to(X,V)$ is a morphism over T.

We will need below the following result:

Proposition 1.1. *If we are given a* ***cartesian*** *diagram of overconvergent presheaves (with a representable upper map)*

$$\begin{array}{ccc} (X',V') & \xrightarrow{(f,u)} & (X,V) \\ \downarrow{\scriptstyle s'} & & \downarrow{\scriptstyle s} \\ T' & \xrightarrow{v} & T \end{array}$$

and $\mathcal{F}'$ is a sheaf on T', then

$$(v_*\mathcal{F}')_V =]f[_{u*}\mathcal{F}'_{V'}.$$

Proof. Since the diagram is cartesian, we have (this is formal)

$$s^{-1}v_*\mathcal{F}' = (f,u)_*(s')^{-1}\mathcal{F}'.$$

It follows that

$$\begin{aligned}(v_*\mathcal{F}')_V &= \varphi_{V*}s^{-1}v_*\mathcal{F}' = \varphi_{V*}(f,u)_*(s')^{-1}\mathcal{F}' \\ &=]f[_{u*}\varphi_{V'*}(s')^{-1}\mathcal{F}' =]f[_{u*}\mathcal{F}'_{V'}.\end{aligned}$$

□

If (X,V) is an overconvergent variety, we will denote by $i_X :]X[_V \hookrightarrow V$ the inclusion map. Then, if T is an overconvergent presheaf, we define the structural sheaf of $\mathrm{An}^\dagger_{/T}$ as the sheaf $\mathcal{O}^\dagger_T$ whose realization on any (X,V) is $i_X^{-1}\mathcal{O}_V$. An $\mathcal{O}^\dagger_T$-module E will also be called a (*overconvergent*) *module* on T. As it was the case for sheaves of sets, the module E is completely determined by its realizations E_V and the transition morphisms

$$]f[_u^\dagger E_V := i_{X'}^{-1}u^*i_{X*}E_V \to E_{V'} \tag{2}$$

obtained by functoriality whenever $(f,u):(X',V')\to(X,V)$ is a morphism over T. A module on T is called an (*overconvergent*) *isocrystal* if all the transition maps (2) are actually isomorphisms (used to be called a crystal in [Le Stum 2011]). We will denote by

$$\mathrm{Isoc}^\dagger(T) \subset \mathcal{O}^\dagger_T\text{-Mod}$$

the full subcategory made of all isocrystals on T (used to be denoted by $\mathrm{Cris}^\dagger(T)$ in [Le Stum 2011]). Be careful that inclusion is only right exact in general.

If we are given a morphism of overconvergent presheaves $v : T' \to T$ then the functors $v_!$, v^{-1}, v_* preserve modules (we use the same notation $v_!$ for sheaves of sets and abelian groups; this should not create any confusion) and v^{-1} preserves isocrystals.

One can show that a module on T is locally finitely presented if and only if it is an isocrystal with coherent realizations. We will denote their category by $\mathrm{Isoc}^{\dagger}_{\mathrm{coh}}(T)$ (be careful that it only means that the realizations are coherent: $\mathcal{O}_T^{\dagger}$ is not a coherent ring in general). In the case $T = X/S_K$ and $\mathrm{Char}(K) = 0$, this is equivalent to Berthelot's original definition [1996a, Definition 2.3.6] of an overconvergent isocrystal.

Back to our examples, it is not difficult to see that, when (X, V) is an overconvergent variety, the realization functor induces an equivalence of categories

$$\mathrm{Isoc}^{\dagger}(X, V) \simeq i_X^{-1}\mathcal{O}_V\text{-Mod}$$

between isocrystals on (X, V) and $i_X^{-1}\mathcal{O}_V$-modules. Now, if (X, V) is a variety over an overconvergent variety (C, O) and

$$p_1, p_2 : (X, V \times_O V) \to (X, V)$$

denote the projections, we define an *overconvergent stratification* on an $i_X^{-1}\mathcal{O}_V$-module $\mathcal{F}$ as an isomorphism

$$\epsilon :]p_2[^{\dagger}\mathcal{F} \simeq]p_1[^{\dagger}\mathcal{F}$$

that satisfies the cocycle condition on triple products and the normalization condition along the diagonal. They form an additive category $\mathrm{Strat}^{\dagger}(X, V/O)$ with cokernels and tensor products. It is even an abelian category when V is universally flat over O in a neighborhood of $]X[_V$. In any case, the realization functor will induce an equivalence

$$\mathrm{Isoc}^{\dagger}(X_V/O) \simeq \mathrm{Strat}^{\dagger}(X, V/O).$$

We may also consider, for $n \in \mathbb{N}$, the n-th infinitesimal neighborhood $V^{(n)}$ of V in $V \times_O V$. Then, a (*usual*) *stratification* on an $i_X^{-1}\mathcal{O}_V$-module $\mathcal{F}$ is a compatible family of isomorphisms

$$\epsilon^{(n)} : i_X^{-1}\mathcal{O}_{V^{(n)}} \otimes_{i_X^{-1}\mathcal{O}_V} \mathcal{F} \simeq \mathcal{F} \otimes_{i_X^{-1}\mathcal{O}_V} i_X^{-1}\mathcal{O}_{V^{(n)}}$$

that satisfy the cocycle condition on triple products and the normalization condition along the diagonal. Again, they form an additive category $\mathrm{Strat}(X, V/O)$ with cokernels and tensor products, and even an abelian category when V is smooth over O in a neighborhood of $]X[_V$. There exists an obvious faithful functor

$$\mathrm{Strat}^{\dagger}(X, V/O) \to \mathrm{Strat}(X, V/O). \tag{3}$$

Note that, *a priori*, different overconvergent stratifications might give rise to the same usual stratification (and of course many usual stratifications will *not* extend at all to an overconvergent one). Finally, a *connection* on an $i_X^{-1}\mathcal{O}_V$-module $\mathcal{F}$ is an

$\mathcal{O}_O$-linear map

$$\nabla : \mathcal{F} \to \mathcal{F} \otimes_{i_X^{-1}\mathcal{O}_V} i_X^{-1}\Omega^1_V$$

that satisfies the Leibniz rule. Integrability is defined as usual. They form an additive category $\mathrm{MIC}(X, V/O)$ and there exists again a faithful functor

$$\mathrm{Strat}(X, V/O) \to \mathrm{MIC}(X, V/O) \tag{4}$$

(∇ is induced by $\epsilon^{(1)} - \sigma$, where σ switches the factors in $V \times_O V$). When V is smooth over O in a neighborhood of $]X[_V$ and $\mathrm{Char}(K) = 0$, then the functor (4) is an equivalence. Actually, both categories are then equivalent to the category of $i_X^{-1}\mathcal{D}_{V/O}$-modules. In general, we will denote by $\mathrm{MIC}^\dagger(X, V/O)$ the image of the composition of the functors (3) and (4) and then call the connection *overconvergent* (and add an index coh when we consider only coherent modules). Thus, there exists a realization functor

$$\mathrm{Isoc}^\dagger(X_V/O) \to \mathrm{MIC}^\dagger(X, V/O) \tag{5}$$

which is faithful and essentially surjective (but not an equivalence in general). In practice, we are interested in isocrystals on X/O, where (C, O) is an overconvergent variety and X is an algebraic variety over C. We can localize in order to find a geometric realization V for X over O such as (1) and work directly on (X, V): there exists an equivalence of categories

$$\mathrm{Isoc}^\dagger(X/O) \simeq \mathrm{Isoc}^\dagger(X_V/O)$$

that may be composed with (5) in order to get the realization functor

$$\mathrm{Isoc}^\dagger(X/O) \to \mathrm{MIC}^\dagger(X, V/O).$$

In [Le Stum 2011], we proved that, when $\mathrm{Char}(K) = 0$, it induces an equivalence

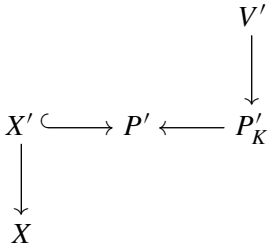
$$\mathrm{Isoc}^\dagger_{\mathrm{coh}}(X/O) \simeq \mathrm{MIC}^\dagger_{\mathrm{coh}}(X, V/O)$$

(showing in particular that the right-hand side is independent of the choice of the geometric realization and that they glue). We will extend this below to what we call *constructible isocrystals*.

2. Locally closed embeddings

In this section, we fix an algebraic variety X over k. Recall that a (overconvergent) variety over $X/\mathcal{M}(K)$ (we will simply write X/K in the future) is a pair made of an overconvergent variety (X', V') and a morphism $X' \to X$. In other words,

it is a diagram



in which P' is a formal scheme.

We also fix a presheaf T over X/K. For example, T could be (the presheaf represented by) an overconvergent variety (X', V') over X/K. Also, if (C, O) is an overconvergent variety and X is an algebraic variety over C, then we may consider the sheaf $T := X/O$ (see Section 1). Finally, if we are given a morphism of overconvergent varieties $(X, V) \to (C, O)$, then we could set $T := X_V/O$ (see Section 1 again).

Finally, we also fix an open immersion $\alpha : U \hookrightarrow X$ and denote by $\beta : Z \hookrightarrow X$ the embedding of a closed complement. Actually, in the beginning, we consider more generally a locally closed embedding $\gamma : Y \hookrightarrow X$.

Definition 2.1. The *restriction* of T to Y is the inverse image

$$T_Y := (Y/K) \times_{(X/K)} T$$

of T over Y/K. We will still denote by $\gamma : T_Y \hookrightarrow T$ the corresponding map. When $\mathcal{F}$ is a sheaf on T, the *restriction* of T to Y is $\mathcal{F}_{|Y} := \gamma^{-1}\mathcal{F}$.

For example, if $T = (X', V')$ is a variety over X/K, then $T_Y = (Y', V')$, where Y' is the inverse image of Y in X'. Also, if (C, O) is an overconvergent variety, X is an algebraic variety over C and $T = X/O$, then $T_Y = Y/O$. Finally, if we are given a morphism of overconvergent varieties $(X, V) \to (C, O)$ and $T = X_V/O$, then we will have $T_Y = Y_V/O$.

If (X, V) is an overconvergent variety, we may consider the morphism of overconvergent varieties $(\gamma, \mathrm{Id}_V) : (Y, V) \hookrightarrow (X, V)$. We will then denote by $]\gamma[_V :]Y[_V \hookrightarrow]X[_V$, or simply $]\gamma[$ if there is no ambiguity, the corresponding map on the tubes. Recall that $]\gamma[$ is the inclusion of an analytic domain. This is an open immersion when γ is a closed embedding and vice versa (we use Berkovich topology).

The next result generalizes Proposition 3.1.10 of [Le Stum 2011].

Proposition 2.2. *Let (X', V') be an overconvergent variety over T and $\gamma' : Y' \hookrightarrow X'$ be the inclusion of the inverse image of Y inside X'. If $\mathcal{F}$ is a sheaf on T_Y, then*

$$(\gamma_*\mathcal{F})_{X',V'} =]\gamma'[_*\mathcal{F}_{Y',V'}.$$

Proof. Using [Le Stum 2011, Corollary 2.4.15], this follows from Proposition 1.1. □

Since we will use it in some of our examples, we should also mention that $\mathrm{R}^i\gamma_* E = 0$ for $i > 0$ when E is an isocrystal with coherent realizations. This follows from the fact that, with the notation of the proposition, $]\gamma'[$ is a quasi-Stein map.

We can work out very simple examples right now. We will do our computations on the overconvergent variety

$$\mathbb{P}^1_{k/K} := \mathbb{P}^1_k \hookrightarrow \widehat{\mathbb{P}}^1_{\mathcal{V}} \leftarrow \mathbb{P}^{1,\mathrm{an}}_K.$$

We consider first the open immersion $\alpha : \mathbb{A}^1_k \hookrightarrow \mathbb{P}^1_k$ and the structural sheaf $\mathcal{O}^\dagger_{\mathbb{A}^1_k/K}$. If we let $i : \mathbb{D}(0, 1^+) \hookrightarrow \mathbb{P}^{1,\mathrm{an}}_K$ denote the inclusion map, we have

$$\mathrm{R}\Gamma\big(\mathbb{P}^1_{k/K}, \alpha_*\mathcal{O}^\dagger_{\mathbb{A}^1_k/K}\big) = \mathrm{R}\Gamma\big(\mathbb{P}^{1,\mathrm{an}}_K, i_*i^{-1}\mathcal{O}_{\mathbb{P}^{1,\mathrm{an}}_K}\big) = K[t]^\dagger := \bigcup_{\lambda>1} K\{t/\lambda\}$$

(functions with radius of convergence (strictly) bigger than one at the origin).

On the other hand, if we start from $\beta : \infty \hookrightarrow \mathbb{P}^1_k$ and let $j : \mathbb{D}(\infty, 1^-) \hookrightarrow \mathbb{P}^{1,\mathrm{an}}_K$ denote the inclusion map, we have

$$\mathrm{R}\Gamma\big(\mathbb{P}^1_{k/K}, \beta_*\mathcal{O}^\dagger_{\infty/K}\big) = \mathrm{R}\Gamma\big(\mathbb{P}^{1,\mathrm{an}}_K, j_*j^{-1}\mathcal{O}_{\mathbb{P}^{1,\mathrm{an}}_K}\big) = K[1/t]^{\mathrm{an}} := \bigcap_{\lambda>1} K\{\lambda/t\}$$

(functions with radius of convergence at least one at infinity).

The following is immediate from Proposition 2.2:

Corollary 2.3. (1) $\gamma^{-1}_{\mathrm{An}^\dagger} \circ \gamma_{\mathrm{An}^\dagger *} = \mathrm{Id}$, *and*

(2) *if* $\gamma' : Y' \hookrightarrow X$ *is another locally closed embedding with* $Y \cap Y' = \varnothing$, *then*

$$\gamma^{-1}_{\mathrm{An}^\dagger} \circ \gamma'_{\mathrm{An}^\dagger *} = 0.$$

Alternatively, one may say that if $\mathcal{F}$ is a sheaf on T_Y, we have

$$(\gamma_*\mathcal{F})_{|Y} = \mathcal{F} \quad \text{and} \quad (\gamma_*\mathcal{F})_{|Y'} = 0.$$

The first assertion of the corollary means that $\gamma_{\mathrm{An}^\dagger}$ is an embedding of topos (direct image is fully faithful). Actually, from the fact that Y is a subobject of X in the category of varieties, one easily deduces that T_Y is a subobject of T in the overconvergent topos and $\gamma_{\mathrm{An}^\dagger}$ is therefore an *open* immersion of topos. Note also that the second assertion applies in particular to open and closed complements (both ways): in particular, these functors *cannot* be used to glue along open and closed complements. We will need some refinement.

We focus now on the case of an *open* immersion $\alpha : U \hookrightarrow X$ which gives rise to a closed embedding on the tubes.

Proposition 2.4. *The functor* $\alpha_{\mathrm{An}^\dagger *} : T_{U,\mathrm{An}^\dagger} \to T_{\mathrm{An}^\dagger}$ *is exact and preserves isocrystals.*

Proof. This is not trivial but can be proved exactly as in Corollary 3.1.12 and Proposition 3.3.15 of [Le Stum 2011] (which is the case $T = X/O$). □

The following definition is related to rigid cohomology with compact support (recall that $\beta : Z \hookrightarrow X$ denotes the embedding of a closed complement of U):

Definition 2.5. If $\mathcal{F}$ is a sheaf of abelian groups on T, then

$$\underline{\Gamma}_U \mathcal{F} = \ker(\mathcal{F} \to \beta_* \mathcal{F}_{|Z})$$

is the subsheaf of *sections of $\mathcal{F}$ with support in U.*

If we denote by $\mathcal{U}$ the *closed* subtopos of $T_{\mathrm{An}^\dagger}$ which is the complement of the open topos $T_{Z,\mathrm{An}^\dagger}$, then $\underline{\Gamma}_U$ is the same thing as the functor $\mathcal{H}^0_{\mathcal{U}}$ of sections with support in $\mathcal{U}$. With this in mind, the first two assertions of the next proposition below are completely formal. One may also show that the functor $\mathcal{F} \mapsto \mathcal{F}/\beta_! \beta^{-1} \mathcal{F}$ is an exact left adjoint to $\underline{\Gamma}_U$; it follows that $\underline{\Gamma}_U$ preserves injectives.

Actually, we shall use the open/closed formalism only in the classical situation. Recall (see [Iversen 1986, Section II.6], for example, for these kinds of things) that if $i : W \hookrightarrow V$ is a closed embedding of topological spaces, then i_* has a right adjoint $i^!$ (and one usually sets $\underline{\Gamma}_W := i_* i^!$) which commutes with direct images. If (X, V) is an overconvergent variety, we know that $]\alpha[:]U[\hookrightarrow]X[$ is a closed embedding and we may therefore consider the functors $]\alpha[^!$ and $\underline{\Gamma}_{]U[}$.

Proposition 2.6. (1) *The functor $\underline{\Gamma}_U$ is left exact and preserves modules.*

(2) *If $\mathcal{F}$ is a sheaf of abelian groups on T, then there exists a distinguished triangle*

$$\mathrm{R}\underline{\Gamma}_U \mathcal{F} \to \mathcal{F} \to \mathrm{R}\beta_* \mathcal{F}_{|Z} \to .$$

(3) *If (X', V') is a variety over T and $\alpha' : U' \hookrightarrow X'$ denotes the immersion of the inverse image of U into X', we have*

$$(\mathrm{R}\underline{\Gamma}_U E)_{V'} = \mathrm{R}\underline{\Gamma}_{]U'[_{V'}} E_{V'}$$

*for any **isocrystal** E on T.*

Proof. The first assertion follows immediately from the fact that all the functors involved (β^{-1}, β_* and ker) do have these properties. The second assertion results from the fact that the map $\mathcal{F} \to \beta_* \mathcal{F}_{|Z}$ is surjective when $\mathcal{F}$ is an injective sheaf (this is formal). In order to prove the last assertion, it is sufficient to remember (this is a standard fact) that there exists a distinguished triangle

$$\mathrm{R}\underline{\Gamma}_{]U'[_{V'}} E_{V'} \to E_{V'} \to \mathrm{R}]\beta'[_*]\beta[^{-1} E_{V'} \to,$$

where $\beta' : Z' \hookrightarrow X'$ denotes the inverse image of the inclusion of a closed complement of U. Since E is an isocrystal, we have $(E_{|Z})_{Z',V'} =]\beta'[^{-1} E_{X',V'}$. □

Note that the second assertion means that there exists an exact sequence

$$0 \to \underline{\Gamma}_U \mathcal{F} \to \mathcal{F} \to \beta_* \mathcal{F}_{|Z} \to \mathrm{R}^1 \underline{\Gamma}_U \mathcal{F} \to 0$$

and that $\mathrm{R}^i \beta_* \mathcal{F}_{|Z} = \mathrm{R}^{i+1} \underline{\Gamma}_U \mathcal{F}$ for $i > 0$. We can do the exercise with $\alpha : \mathbb{A}^1_k \hookrightarrow \mathbb{P}^1_k$ and $\beta : \infty \hookrightarrow \mathbb{P}^1_k$ as above. We obtain

$$\mathrm{R}\Gamma\big(\mathbb{P}^1_{k/K}, \mathrm{R}\underline{\Gamma}_{\mathbb{A}^1_k} \mathcal{O}^\dagger_{\mathbb{P}^1_k/K}\big) = [K \to K[1/t]^{\mathrm{an}}] = (K[1/t]^{\mathrm{an}}/K)[-1].$$

Since realization does not commute with the inverse image in general, we need to introduce a new functor. Recall that in order to define a sheaf on T, it is sufficient (and even equivalent) to give a compatible family of sheaves on the tubes $]X'[_{V'}$ for all (X', V') over T.

Lemma 2.7. *If $\mathcal{F}$ is a sheaf on T, then the assignment*

$$(X', V') \mapsto \big(j_U^\dagger \mathcal{F}\big)_{V'} :=]\alpha'[_*]\alpha'[^{-1}\mathcal{F}_{V'},$$

where $\alpha' : U' \hookrightarrow X'$ denotes the immersion of the inverse image of U into X', defines a sheaf on T.

Proof. We give ourselves a morphism $(f, u) : (X'', V'') \to (X', V')$ over T, we denote by $g : U'' \to U'$ the map induced by f on the inverse images of U into X' and X'', respectively, and by $\alpha'' : U'' \hookrightarrow X''$ the inclusion map. We consider the cartesian diagram (forgetful functor to algebraic varieties is left exact)

$$\begin{array}{ccc} (U'', V'') & \hookrightarrow & (X'', V'') \\ \downarrow{\scriptstyle (g,u)} & & \downarrow{\scriptstyle (f,u)} \\ (U', V') & \hookrightarrow & (X', V') \end{array}$$

which gives rise to a cartesian diagram (tube is left exact)

$$\begin{array}{ccc}]U''[_{V''} & \hookrightarrow &]X''[_{V''} \\ \downarrow{\scriptstyle]g[_u} & & \downarrow{\scriptstyle]f[_u} \\ ]U'[_{V'} & \hookrightarrow &]X'[_{V'} \end{array}$$

Since $]\alpha'[$ is a closed embedding, we have $]f[_u^{-1} \circ]\alpha'[_* =]\alpha''[_* \circ]g[_u^{-1}$ and there exists a canonical map

$$]f[_u^{-1}]\alpha'[_*]\alpha'[^{-1}\mathcal{F}_{V'} =]\alpha''[_*]g[_u^{-1}]\alpha'[^{-1}\mathcal{F}_{V'} =]\alpha''[_*]\alpha''[^{-1}]f[_u^{-1}\mathcal{F}_{V''} \to]\alpha''[_*]\alpha''[^{-1}\mathcal{F}_{V''}. \quad \square$$

Definition 2.8. If $\mathcal{F}$ is a sheaf on T, then $j_U^\dagger \mathcal{F}$ is the sheaf of *overconvergent sections of $\mathcal{F}$ around U*.

Proposition 2.9. (1) *The functor $j_U^\dagger$ is exact and preserves isocrystals.*

(2) *If E is an isocrystal on T, we have $j_U^\dagger E = \alpha_* \alpha^{-1} E$.*

Proof. Exactness can be checked on realizations. But, if (X', V') is a variety over T and $\alpha' : U' \hookrightarrow X'$ denotes the immersion of the inverse image of U in X', then we know the exactness of $]\alpha'[_*$ (because $]\alpha'[$ is a closed embedding) and $]\alpha'[^{-1}$. The second part of the first assertion is a consequence of the second assertion which follows from the fact that $(\alpha^{-1}E)_{V'} =]\alpha'[^{-1} E_{V'}$ when E is an isocrystal. □

Note that the canonical map $j_U^\dagger \mathcal{F} \to \alpha_* \alpha^{-1} \mathcal{F}$ is still bijective when $\mathcal{F}$ is a sheaf of *Zariski type* (see Definition 4.6.1[1] of [Le Stum 2011]) but there are important concrete situations where equality fails, as we shall see right now.

In order to exhibit a counterexample, we let again $\alpha : \mathbb{A}^1_k \hookrightarrow \mathbb{P}^1_k$ and $\beta : \infty \hookrightarrow \mathbb{P}^1_k$ denote the inclusion maps and consider the sheaf $\mathcal{F} := \beta_* \mathcal{O}^\dagger_{\infty/K}$, which is *not* an isocrystal (and not even of Zariski type). Since $\alpha^{-1} \circ \beta_* = 0$, we have $\alpha_* \alpha^{-1} \mathcal{F} = 0$. Now, let us denote by $i_\xi : \xi \hookrightarrow \mathbb{P}^{1,\mathrm{an}}_K$ the inclusion of the generic point of the unit disc (corresponding to the Gauss norm) and let $i : \mathbb{D}(0, 1^+) \hookrightarrow \mathbb{P}^{1,\mathrm{an}}_K$ and $j : \mathbb{D}(\infty, 1^-) \hookrightarrow \mathbb{P}^{1,\mathrm{an}}_K$ be the inclusion maps as above. Let

$$\mathcal{R} := \left\{ \sum_{n \in \mathbb{Z}} a_n t^n : \begin{cases} \exists \lambda > 1, \lambda^n a_n \to 0 \text{ for } n \to +\infty \\ \forall \lambda > 1, \lambda^n a_n \to 0 \text{ for } n \to -\infty \end{cases} \right\}$$

be the *Robba ring* (functions that converge on some open annulus of outer radius one at infinity). Then, one easily sees that

$$\left(j^\dagger_{\mathbb{A}^1_k} \beta_* \mathcal{O}^\dagger_{\infty/K}\right)_{\mathbb{P}^1_{k/K}} = i_* i^{-1} j_* \mathcal{O}_{\mathbb{D}(0,1^-)} = i_{\xi *} \mathcal{R}$$

so that $j^\dagger_{\mathbb{A}^1_k} \mathcal{F} \neq 0$. This computation also shows that

$$\mathrm{R}\Gamma\left(\mathbb{P}^1_{k/K}, j^\dagger_{\mathbb{A}^1_k} \beta_* \mathcal{O}^\dagger_{\infty/K}\right) = \mathcal{R}.$$

We now turn to the study of the *closed* embedding $\beta : Z \hookrightarrow X$, which requires some care (as we just experienced, the direct image of an isocrystal need not be an isocrystal).

The following definition has to do with cohomology with support in a closed subset.

Definition 2.10. For any sheaf of abelian groups $\mathcal{F}$ on T,

$$\underline{\Gamma}^\dagger_Z \mathcal{F} := \ker(\mathcal{F} \to \alpha_* \mathcal{F}_{|U})$$

is the subsheaf of *overconvergent sections of $\mathcal{F}$ with support in Z*.

[1] The comment following Definition 4.6.1 in [Le Stum 2011] is not correct and Lemma 4.6.2 is only valid for an *open* immersion.

We will do some examples below when we have more material at our disposal.

As above, if we denote by $\mathcal{Z}$ the closed subtopos of $T_{\mathrm{An}^\dagger}$ which is the complement of the open topos $T_{U,\mathrm{An}^\dagger}$, then $\underline{\Gamma}^\dagger_Z$ is the same thing as the functor $\mathcal{H}^0_{\mathcal{Z}}$ of sections with support in $\mathcal{Z}$. This is the approach taken by David Zureick-Brown [2010; 2014] in order to define cohomology with support in Z on the overconvergent site. The next proposition is completely formal if one uses Zureick-Brown's approach. Also, as above, one may prove that $\underline{\Gamma}^\dagger_Z$ preserves injectives because the functor $\mathcal{F} \mapsto \mathcal{F}/\alpha_!\alpha^{-1}\mathcal{F}$ is an exact left adjoint.

Proposition 2.11. (1) *The functor $\underline{\Gamma}^\dagger_Z$ is left exact and preserves modules.*

(2) *If $\mathcal{F}$ is an abelian sheaf on T, then there exists a distinguished triangle*

$$0 \to \mathrm{R}\underline{\Gamma}^\dagger_Z\mathcal{F} \to \mathcal{F} \to \alpha_*\mathcal{F}_{|U} \to .$$

We will also show below that $\underline{\Gamma}^\dagger_Z$ preserves isocrystals.

Proof. As in the proof of Proposition 2.6, the first assertion follows from the fact that all the functors involved (and the kernel as well) are left exact and preserve overconvergent modules. Similarly the second one is a formal consequence of the definition because α_* and α^{-1} both preserve injectives (they both have an exact left adjoint) and the map $\mathcal{F} \to \alpha_*\mathcal{F}_{|U}$ is an epimorphism when $\mathcal{F}$ is injective (standard). □

Note that the last assertion of the proposition means that there exists an exact sequence

$$0 \to \underline{\Gamma}^\dagger_Z\mathcal{F} \to \mathcal{F} \to \alpha_*\mathcal{F}_{|U} \to \mathrm{R}^1\underline{\Gamma}^\dagger_Z\mathcal{F} \to 0$$

and that $\mathrm{R}^i\underline{\Gamma}^\dagger_Z\mathcal{F} = 0$ for $i > 1$.

Before going any further, we want to stress the fact that β^{-1} has an adjoint $\beta_!$ on the left in the category of all modules (or abelian groups or even sets with a light modification) but $\beta_!$ does not preserve isocrystals in general. Actually, we always have $(\beta_!\mathcal{F})_{X',V'} = 0$ unless the morphism $X' \to X$ factors through Z (recall that we use the coarse topology on the algebraic side). Again, the workaround consists in working directly with the realizations. If $j : W \hookrightarrow V$ is an open immersion of topological spaces, then j^{-1} has an adjoint $j_!$ on the left also (on sheaves of abelian groups or sheaves of sets with a light modification). This is an exact functor that commutes with inverse images (see [Iversen 1986, Section II.6] again). Now, if (X, V) is an overconvergent variety, then $]\beta[:]Z[\hookrightarrow]X[$ is an *open* immersion and we may consider the functor $]\beta[_!$.

In the next lemma again, we use realizations and transition maps in order to define a sheaf.

Lemma 2.12. *If $\mathcal{F}$ is a sheaf (of sets or abelian groups) on T_Z, then the assignment*

$$(X', V') \mapsto (\beta_\dagger\mathcal{F})_{X',V'} :=]\beta'[_!\mathcal{F}_{Z',V'},$$

where $\beta' : Z' \hookrightarrow X'$ denotes the embedding of the inverse image of Z into X', defines a sheaf on T. Moreover, if E is an isocrystal on T_Z, then $\beta_\dagger E$ is an isocrystal on T.

Proof. As above, we consider a morphism $(f, u) : (X'', V'') \to (X', V')$ over T. We denote by $h : Z'' \to Z'$ the map induced by f on the inverse images of Z into X' and X'', respectively, and by $\beta'' : Z'' \hookrightarrow X''$ the inclusion map. We have a cartesian diagram

$$\begin{array}{ccc} (Z'', V'') & \hookrightarrow & (X'', V'') \\ \downarrow{\scriptstyle (h,u)} & & \downarrow{\scriptstyle (f,u)} \\ (Z', V') & \hookrightarrow & (X', V') \end{array}$$

giving rise to a cartesian diagram

$$\begin{array}{ccc}]Z''[_{V''} & \hookrightarrow &]X''[_{V''} \\ \downarrow{\scriptstyle]h[_u} & & \downarrow{\scriptstyle]f[_u} \\ ]Z'[_{V'} & \hookrightarrow &]X'[_{V'} \end{array}$$

It follows that there exists a canonical map

$$]f[_u^{-1}]\beta'[_!\mathcal{F}_{V'} =]\beta''[_!]h[_u^{-1}\mathcal{F}_{V'} \to]\beta''[_!\mathcal{F}_{V''}$$

as asserted. We consider now an isocrystal E and we want to show that

$$]f[_u^{\dagger}]\beta'[_!E_{V'} \simeq]\beta''[_!E_{V''}.$$

This immediately follows from the equality (which is formal)

$$i_{X''}^{-1}\mathcal{O}_{V''} \otimes_{i_{X''}^{-1}u^{-1}\mathcal{O}_{V'}}]\beta''[_!]h[_u^{-1}E_{V'} =]\beta''[_!\big(i_{Z''}^{-1}\mathcal{O}_{V''} \otimes_{i_{Z''}^{-1}u^{-1}\mathcal{O}_{V'}}]h[_u^{-1}E_{V'}\big). \quad \square$$

Definition 2.13. The sheaf $\beta_\dagger\mathcal{F}$ is the *overconvergent direct image* of $\mathcal{F}$.

Note that there exist two flavors of $\beta_\dagger$: for sheaves of sets and for sheaves of abelian groups. Whichever we consider should be clear from the context.

Proposition 2.14. (1) *If $\mathcal{F}$ is a sheaf on T_Z, then*:

(a) $(\beta_\dagger\mathcal{F})_{|Z} = \mathcal{F}$.

(b) $(\beta_\dagger\mathcal{F})_{|U} = 0$.

(c) *If E is an isocrystal on T, then*

$$\mathcal{H}om(\beta_\dagger\mathcal{F}, E) = \beta_*\,\mathcal{H}om(\mathcal{F}, \beta^{-1}E). \tag{6}$$

(d) *There exists a short exact sequence*

$$0 \to \beta_\dagger\mathcal{F} \to \beta_*\mathcal{F} \to j_U^\dagger\beta_*\mathcal{F} \to 0. \tag{7}$$

(2) *The functor $\beta_\dagger$ is fully faithful, exact, and preserves isocrystals, and the induced functor*

$$\beta_\dagger : \mathrm{Isoc}^\dagger(T_Z) \to \mathrm{Isoc}^\dagger(T)$$

is left adjoint to

$$\beta^{-1} : \mathrm{Isoc}^\dagger(T) \to \mathrm{Isoc}^\dagger(T_Z).$$

Proof. As usual, if (X', V') is a variety over T, then we denote by $\alpha' : U' \hookrightarrow X'$ and $\beta' : Z' \hookrightarrow X'$ the inclusions of the inverse images of U and Z, respectively.

When (X', V') is an overconvergent variety over T_Z, then we will have $]\beta'[= \mathrm{Id}$, and when (X', V') is an overconvergent variety over T_U then $]\beta'[= \varnothing$. We obtain the first two assertions. When E is an isocrystal on T, we have an isomorphism (this is standard)

$$\mathcal{H}om(]\beta'[_!\mathcal{F}_{V'}, E_{V'}) =]\beta'[_*\,\mathcal{H}om(\mathcal{F}_{V'},]\beta'[^{-1}E_{V'}),$$

from which the third assertion follows. Also, there exists a short exact sequence

$$0 \to]\beta'[_!\mathcal{F}_{Z',V'} \to]\beta'[_*\mathcal{F}_{Z',V'} \to]\alpha'[_*]\alpha'[^{-1}]\beta'[_*\mathcal{F}_{Z',V'} \to 0$$

which provides the fourth assertion.

Full faithfulness and exactness of $\beta_\dagger$ follow from the full faithfulness and exactness of $]\beta'[_!$ for all (X', V'). The fact that $\beta_\dagger$ preserves isocrystals was proved in Lemma 2.12. The last assertion may be obtained by taking global sections of the equality (6). □

We can also mention that there exists a distinguished triangle

$$\beta_\dagger\mathcal{F} \to \mathrm{R}\beta_*\mathcal{F} \to j_U^\dagger \mathrm{R}\beta_*\mathcal{F} \to .$$

Now, we prove that the exact sequence (7) is universal:

Proposition 2.15. *If $\mathcal{F}'$ and $\mathcal{F}''$ are modules on T_Z and T_U, respectively, then any extension*

$$0 \to \beta_\dagger\mathcal{F}' \to \mathcal{F} \to \alpha_*\mathcal{F}'' \to 0$$

is a pull-back of the fundamental extension (7) through a unique morphism

$$\alpha_*\mathcal{F}'' \to j_U^\dagger\beta_*\mathcal{F}'.$$

Proof. We know that $\beta^{-1}\alpha_*\mathcal{F}'' = 0$ and it follows that

$$\mathrm{Hom}(\alpha_*\mathcal{F}'', \beta_*\mathcal{F}') = \mathrm{Hom}(\beta^{-1}\alpha_*\mathcal{F}'', \mathcal{F}') = 0.$$

This being true for any sheaves, we see that, actually, $\mathrm{R}\,\mathrm{Hom}(\alpha_*\mathcal{F}'', \mathrm{R}\beta_*\mathcal{F}') = 0$. It formally follows that $\mathrm{R}^i\,\mathrm{Hom}(\alpha_*\mathcal{F}'', \beta_*\mathcal{F}') = 0$ for $i \leq 1$. As a consequence, we

obtain a canonical isomorphism

$$\mathrm{Hom}(\alpha_*\mathcal{F}'', j_U^\dagger\beta_*\mathcal{F}') \simeq \mathrm{Ext}(\alpha_*\mathcal{F}'', \beta_\dagger\mathcal{F}').$$

This is exactly the content of our assertion. □

We should observe that we always have $\mathrm{Hom}(\alpha_*\mathcal{F}'', \beta_\dagger\mathcal{F}') = 0$. However, it is *not* true that $\mathrm{Ext}(\alpha_*\mathcal{F}'', \beta_\dagger\mathcal{F}') = 0$ in general. This can happen because $\beta_\dagger$ does not preserve injectives (although it is exact).

The overconvergent direct image is related to overconvergent support as follows:

Proposition 2.16. *If E is an isocrystal on T, then*

$$\underline{\Gamma}_Z^\dagger E = \beta_\dagger E_{|Z}$$

and, for all $i > 0$, $\mathrm{R}^i\underline{\Gamma}_Z^\dagger E = 0$.

Proof. Recall from Proposition 2.11 that there exists an exact sequence

$$0 \to \underline{\Gamma}_Z^\dagger E \to E \to \alpha_* E_{|U} \to \mathrm{R}^1\underline{\Gamma}_Z^\dagger E \to 0$$

and that $\mathrm{R}^i\underline{\Gamma}_Z^\dagger E = 0$ for $i > 1$. Now, let (X', V') be a variety over T. Denote by $\beta' : Z' \hookrightarrow X'$, $\alpha' : U' \hookrightarrow X'$ the embeddings of the inverse images of Z and U into X'. There exists a short exact sequence (standard again)

$$0 \to]\beta'[_!]\beta'[^{-1}E_{V'} \to E_{V'} \to]\alpha'[_*]\alpha'[^{-1}E_{V'} \to 0.$$

Since E is an isocrystal, we have $(\alpha_* E_{|U})_{V'} =]\alpha'[_*]\alpha'[^{-1}E_{V'}$. It follows that $(\mathrm{R}^1\underline{\Gamma}_Z^\dagger E)_{V'} = 0$ and we also see that

$$(\underline{\Gamma}_Z^\dagger E)_{V'} =]\beta'[_!]\beta'[^{-1}E_{V'} = (\beta_\dagger E_{|Z})_{V'}.$$ □

Note that the proposition is still valid for sheaves of Zariski type and not merely for isocrystals. Be careful however that $\beta_\dagger E \neq \underline{\Gamma}_Z^\dagger\beta_* E$ in general, even when E is an isocrystal on T_Z. With our favorite example in mind, we have

$$\underline{\Gamma}_Z^\dagger\beta_*\mathcal{O}_{\infty/K}^\dagger = \beta_*\mathcal{O}_{\infty/K}^\dagger \neq \beta_\dagger\mathcal{O}_{\infty/K}^\dagger,$$

as our computations below will show.

Corollary 2.17. *The functor $\underline{\Gamma}_Z^\dagger$ preserves isocrystals, and the induced functor*

$$\underline{\Gamma}_Z^\dagger : \mathrm{Isoc}^\dagger(T) \to \mathrm{Isoc}^\dagger(T)$$

is exact. Moreover, if E is an isocrystal on T, then there exists a short exact sequence

$$0 \to \underline{\Gamma}_Z^\dagger E \to E \to j_U^\dagger E \to 0.$$

We might as well write this last short exact sequence as

$$0 \to \beta_\dagger E_{|Z} \to E \to \alpha_* E_{|U} \to 0.$$

As promised above, we can do an example and consider the closed embedding $\beta : \infty \hookrightarrow \mathbb{P}^1_k$ again. We compute

$$\beta_\dagger \mathcal{O}^\dagger_{\infty/K} = \underline{\Gamma}^\dagger_\infty \mathcal{O}^\dagger_{\mathbb{P}^1_k/K}.$$

We have

$$\mathrm{R}\Gamma\big(\mathbb{P}^1_{k/K}, \beta_\dagger \mathcal{O}^\dagger_{\infty/K}\big) = [K \to K[t]^\dagger] = (K[t]^\dagger/K)[-1].$$

We can also remark that the (long) exact sequence obtained by applying $\mathrm{R}\Gamma(\mathbb{P}^1_{k/K}, -)$ to the fundamental short exact sequence

$$0 \to \beta_\dagger \mathcal{O}^\dagger_{\infty/K} \to \beta_* \mathcal{O}^\dagger_{\infty/K} \to j^\dagger_U \beta_* \mathcal{O}^\dagger_{\infty/K} \to 0$$

reads

$$0 \to K[1/t]^{\mathrm{an}} \to \mathcal{R} \to K[t]^\dagger/K \to 0. \tag{8}$$

Corollary 2.18. (1) *The functors α_* and α^{-1} induce an equivalence between isocrystals on T_U and isocrystals on T such that $\underline{\Gamma}^\dagger_Z E = 0$ (or $j^\dagger_U E = E$).*

(2) *The functors $\beta_\dagger$ and β^{-1} induce an equivalence between isocrystals on T_Z and isocrystals on T such that $\underline{\Gamma}^\dagger_Z E = E$ (or $j^\dagger_U E = 0$).*

Proof. If E'' is an isocrystal on T_U, then $\alpha_* E''$ is an isocrystal on T and therefore

$$\underline{\Gamma}^\dagger_Z \alpha_* E'' = \beta_\dagger \beta^{-1} \alpha_* E'' = 0.$$

Conversely, if E is an isocrystal on T such that $\underline{\Gamma}^\dagger_Z E = 0$, then $E = j^\dagger_U E = \alpha_* \alpha^{-1} E$. This shows the first part.

Now, if E' is an isocrystal on T_Z, then $\beta_\dagger E'$ is an isocrystal on T and therefore

$$\underline{\Gamma}^\dagger_Z \beta_\dagger E' = \beta_\dagger \beta^{-1} \beta_\dagger E' = \beta_\dagger E'.$$

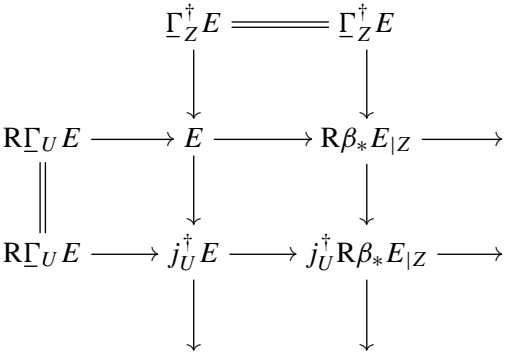
Conversely, if E is an isocrystal on T such that $\underline{\Gamma}^\dagger_Z E = E$, then $E = \beta_\dagger \beta^{-1} E$. □

We can also make the functor of sections with support in an open subset come back into the picture:

Corollary 2.19. *If E is an isocrystal on T, then there exists a distinguished triangle*

$$\mathrm{R}\underline{\Gamma}_U E \to j^\dagger_U E \to j^\dagger_U \mathrm{R}\beta_* E_{|Z} \to .$$

Proof. There exists actually a commutative diagram of distinguished triangles:



More precisely, we know that the vertical triangles as well as the middle horizontal one are all distinguished. The bottom one must be distinguished too. □

Back to our running example, we see that the long exact sequence obtained by applying $\mathrm{R}\Gamma(\mathbb{P}^1_{k/K}, -)$ to the distinguished triangle

$$\mathrm{R}\underline{\Gamma}_{\mathbb{A}^1_k}\mathcal{O}^\dagger_{\mathbb{P}^1_k/K} \to j^\dagger_{\mathbb{A}^1_k}\mathcal{O}^\dagger_{\mathbb{P}^1_k/K} \to j^\dagger_{\mathbb{A}^1_k}\mathrm{R}\beta_*\mathcal{O}^\dagger_{\infty/K} \to$$

reads

$$0 \to K[t]^\dagger \to \mathcal{R} \to K[1/t]^{\mathrm{an}}/K \to 0.$$

We can summarize the situation as follows:

(1) There exist two triples of adjoint functors (up means left):

$$\mathcal{O}^\dagger_{T_U}\text{-Mod} \;\underset{\alpha_*}{\overset{\alpha_!}{\underset{\longleftarrow}{\overset{\longrightarrow}{\alpha^{-1}}}}}\; \mathcal{O}^\dagger_{T}\text{-Mod} \;\underset{\beta*}{\overset{\beta!}{\underset{\longrightarrow}{\overset{\longleftarrow}{\beta^{-1}}}}}\; \mathcal{O}^\dagger_{T_Z}\text{-Mod}.$$

Moreover, α_* is *exact* and *preserves isocrystals* (and so do α^{-1} and β^{-1}).

(2) There exist two functors with support (that preserve injectives):

$$\underline{\Gamma}_U \circlearrowright \mathcal{O}^\dagger_T\text{-Mod} \circlearrowleft \underline{\Gamma}^\dagger_Z.$$

Moreover, $\Gamma^\dagger_Z$ *preserves isocrystals* and is *exact on isocrystals*.

(3) There exist two other functors:

$$j^\dagger_U \circlearrowright \mathcal{O}^\dagger_T\text{-Mod} \xleftarrow{\beta_\dagger} \mathcal{O}^\dagger_{T_Z}\text{-Mod}.$$

They are both *exact* and *preserve isocrystals* (but not injectives). If E is an isocrystal on T, we have

$$j^\dagger_U E = \alpha_* E_{|U} \quad \text{and} \quad \underline{\Gamma}^\dagger_Z E = \beta_\dagger E_{|Z}.$$

3. Constructibility

Recall that K denotes a complete ultrametric field with ring of integers $\mathcal{V}$ and residue field k. We let X be an algebraic variety over k and T a (overconvergent) presheaf over X/K. Roughly speaking, T is some family of varieties X' over X which embed into a formal $\mathcal{V}$-scheme P', together with a morphism of analytic K-varieties $V' \to P'_K$. A (*overconvergent*) *module* $\mathcal{F}$ on T is then a compatible family of $i_{X'}^{-1}\mathcal{O}_{V'}$-modules $\mathcal{F}_{V'}$, where $i_{X'} :]X'[_{V'} \hookrightarrow V'$ denotes the inclusion of the tube (the reader is redirected to Section 1 for the details).

Definition 3.1. A module $\mathcal{F}$ on T is said to be *constructible* (with respect to X) if there exists a locally finite covering of X by locally closed subvarieties Y such that $\mathcal{F}_{|Y}$ is locally finitely presented.

Recall that a locally finitely presented module is the same thing as an isocrystal with coherent realizations. It is important to notice however that a constructible module is *not* necessarily an isocrystal (the transition maps might not be bijective). We'll give an example later.

Proposition 3.2. (1) *Constructible modules on T form an additive category which is stable under cokernel, extension, tensor product and internal Hom.*

(2) *Constructible isocrystals on T form an additive category* $\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(T)$ *which is stable under cokernel, extension and tensor product.*

Proof. The analog to the first assertion for locally finitely presented modules is completely formal besides the internal Hom question that was proved in Proposition 3.3.12 of [Le Stum 2011]. The analog to the second assertion for all isocrystals was proved in Corollary 3.3.9 of [Le Stum 2011]. Since the restriction maps $\mathcal{F} \mapsto \mathcal{F}_{|Y}$ are exact and commute with tensor product and internal Hom, everything follows. □

Note however that $\mathcal{H}om(E_1, E_2)$ need *not* be an isocrystal (see example below) when E_1 and E_2 are two constructible isocrystals.

Proposition 3.3. *Let $\mathcal{F}$ be a module on T.*

(1) *The module $\mathcal{F}$ is constructible if and only if there exists a locally finite covering by locally closed subvarieties Y of X such that $\mathcal{F}_{|Y}$ is constructible.*

(2) *If $T' \to T$ is any morphism of overconvergent presheaves and $\mathcal{F}$ is constructible, then $\mathcal{F}_{|T'}$ is constructible. The converse also is true if $T' \to T$ is a covering.*

(3) *Assume that T is actually a presheaf on X'/K for some $f : X' \to X$. If $\mathcal{F}$ is constructible with respect to X, then it is also constructible with respect to X'.*

Proof. The first assertion is an immediate consequence of the transitivity of locally finite coverings by locally closed subsets: if $X = \bigcup X_i$ and $X_i = \bigcup X_{ij}$ are such coverings, so is the covering $X = \bigcup X_{ij}$.

In order to prove the second assertion, note first that it is formally satisfied by locally finitely presented modules. Moreover, if Y is a locally closed subvariety of X, we have $(\mathcal{F}_{|T'})_{|Y} = (\mathcal{F}_{|Y})_{|T'_Y}$. The result follows.

Finally, for the third assertion, if $X = \bigcup X_i$ is a locally finite covering by locally closed subvarieties, so is $X' = \bigcup f^{-1}(X_i)$. Moreover, by definition $\mathcal{F}_{|f^{-1}(X_i)} = \mathcal{F}_{|X_i}$ and there is nothing to do. □

Together with Corollary 2.18 above, the next proposition will allow us to move freely along a closed or open embedding when we consider constructible isocrystals (note that this is obviously wrong for overconvergent isocrystals with coherent realizations):

Proposition 3.4. (1) *If $\alpha : U \hookrightarrow X$ is an open immersion of algebraic varieties, then a module $\mathcal{F}''$ on T_U is constructible if and only if $\alpha_+\mathcal{F}''$ is constructible.*

(2) *If $\beta : Z \hookrightarrow X$ is a closed embedding of algebraic varieties, then a module $\mathcal{F}'$ on T_Z is constructible if and only if $\beta_\dagger\mathcal{F}'$ is constructible.*

Proof. We may assume that U and Z are open and closed complements. We saw in Corollary 2.3 that $(\alpha_+\mathcal{F}'')_{|Z} = \mathcal{F}''$ and $(\alpha_+\mathcal{F}')_{|U} = 0$. We also saw in Proposition 2.14 that $(\beta_\dagger\mathcal{F}')_{|Z} = \mathcal{F}'$ and $(\beta_\dagger\mathcal{F}')_{|U} = 0$. □

It is easy to see that the *usual* dual to a constructible isocrystal is *not* an isocrystal in general: if $\beta : Z \hookrightarrow X$ is a closed embedding of algebraic varieties and E is an overconvergent isocrystal on Z with coherent realizations, it follows from Proposition 2.14 that

$$(\beta_\dagger E)^\vee := \mathcal{H}om(\beta_\dagger E, \mathcal{O}^\dagger_T) = \beta_+ \mathcal{H}om(E, \mathcal{O}^\dagger_{T_Z}) = \beta_+ E^\vee,$$

which is constructible but is not an isocrystal in general (as we saw in Section 2).

The next property is also very important because it allows the use of noetherian induction to reduce some assertions about constructible isocrystals to analogous assertions about overconvergent isocrystals with coherent realizations.

Lemma 3.5. *A module $\mathcal{F}$ on T is constructible if and only if there exists a closed subvariety Z of X such that, if $U := X \setminus Z$, then both $\mathcal{F}_{|Z}$ and $\mathcal{F}_{|U}$ are constructible. We may even assume that U is dense in X and $\mathcal{F}_{|U}$ is locally finitely presented.*

Proof. The condition is sufficient thanks to assertion (1) of Proposition 3.3. Conversely, if ξ is a generic point of X, then there exists a locally closed subset Y of X such that $\xi \in Y$ and $\mathcal{F}_{|Y}$ is locally finitely presented. The subset Y contains necessarily an open neighborhood U_ξ of ξ in X. We may choose $U := \bigcup U_\xi$. □

Proposition 3.6. *An isocrystal E on T is constructible if and only if there exists an exact sequence*

$$0 \to \beta_\dagger E' \to E \to \alpha_+ E'' \to 0, \tag{9}$$

where E'' (resp. E') is a constructible isocrystal on a closed subvariety Z of X (resp. on $U := X \setminus Z$) and $\beta : Z \hookrightarrow X$ (resp. $\alpha : U \hookrightarrow X$) denotes the inclusion map. We may assume that U is dense in X and that E'' has coherent realizations.

Proof. If we are given such an exact sequence, we may pull back along α and β in order to obtain $E' \simeq E_{|Z}$ and $E'' \simeq E_{|U}$. Conversely, we may set $E' := E_{|Z}$ and $E'' := E_{|U}$ in order to get such an exact sequence by Proposition 2.16. □

Note that this property is specific to constructible *isocrystals* and that the analog for constructible modules is wrong.

It follows from Proposition 2.15 that any extension such as (9) comes from a unique morphism $\alpha_* E'' \to j_U^\dagger \beta_* E'$. This is a classical gluing method and the correspondence is given by the morphism of exact sequences

$$\begin{array}{ccccccccc} 0 & \longrightarrow & \beta_\dagger E' & \longrightarrow & \beta_* E' & \longrightarrow & j_U^\dagger \beta_* E' & \longrightarrow & 0 \\ & & \| & & \uparrow & & \uparrow & & \\ 0 & \longrightarrow & \beta_\dagger E' & \longrightarrow & E & \longrightarrow & \alpha_* E'' & \longrightarrow & 0 \end{array}$$

We can do the computations in the very special case of $\alpha : \mathbb{A}^1_k \hookrightarrow \mathbb{P}^1_k$ and $\beta : \infty \hookrightarrow \mathbb{P}^1_k$. We have $E' = \mathcal{O}^\dagger_{\infty/K} \otimes_K H$ for some finite-dimensional vector space H, and E'' is given by a finite free $K[t]^\dagger$-module M of finite rank endowed with a (overconvergent) connection. One can show that there exists a canonical isomorphism

$$\begin{aligned} \mathrm{Ext}(\alpha_* E'', \beta_\dagger E') &= \mathrm{Hom}(\alpha_* E'', j_U^\dagger \beta_* E') \\ &= \mathrm{Hom}_\nabla(M, \mathcal{R} \otimes_K H) \\ &= H^0_{\mathrm{dR}}(M^\vee \otimes_{K[t]^\dagger} \mathcal{R}) \otimes_K H \end{aligned}$$

(the second identity is not trivial). A slight generalization will give a classification of constructible isocrystals on smooth projective curves as in Theorem 6.15 of [Le Stum 2014].

4. Integrable connections and constructibility

In this section, we will give a more concrete description of constructible isocrystals in the case when T is representable by some overconvergent variety (X, V), in the case $T = X_V/O$, where (X, V) is a variety over some overconvergent variety (C, O), and finally when $T = X/O$, where X is a variety over C (see Section 1).

Definition 4.1. Let (X, V) be an overconvergent variety. An $i_X^{-1}\mathcal{O}_V$-module $\mathcal{F}$ is *constructible* if there exists a locally finite covering of X by locally closed subvarieties Y such that $i_Y^{-1} i_{X*} \mathcal{F}$ is a coherent $i_Y^{-1}\mathcal{O}_V$-module.

Of course, we have $i_Y^{-1} i_{X*}\mathcal{F} = i_{Y\subset X}^{-1}\mathcal{F}$ if we denote by $i_{Y\subset X} :]Y[_V \hookrightarrow]X[_V$ the inclusion of the tubes.

Proposition 4.2. *Let (X, V) be an overconvergent variety. Then*:

(1) $\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X, V)$ *is an abelian subcategory of* $\mathrm{Isoc}^\dagger(X, V)$.

(2) *The realization functor induces an equivalence between* $\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X, V)$ *and the category of all constructible $i_X^{-1}\mathcal{O}_V$-modules.*

Proof. It was shown in Proposition 3.3.8 of [Le Stum 2011] that the realization functor induces an equivalence between $\mathrm{Isoc}^\dagger(X, V)$ and the category of all $i_X^{-1}\mathcal{O}_V$-modules. Overconvergent isocrystals correspond to coherent modules. The second assertion is an immediate consequence of these observations. The first assertion then follows immediately from the analogous result about coherent modules. □

Recall that an $i_X^{-1}\mathcal{O}_V$-module may be endowed with an overconvergent stratification. Then, we have:

Proposition 4.3. *Let (X, V) be an overconvergent variety over another overconvergent variety (C, O).*

(1) *If V is universally flat over O in a neighborhood of* $]X[$, *then* $\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X_V/O)$ *is an abelian subcategory of* $\mathrm{Isoc}^\dagger(X_V/O)$.

(2) *The realization functor induces an equivalence between* $\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X_V/O)$ *and the category of constructible $i_X^{-1}\mathcal{O}_V$-modules $\mathcal{F}$ endowed with an overconvergent stratification.*

Proof. According to Proposition 3.5.3 of [Le Stum 2011], its corollary and Proposition 3.5.5 of the same paper, the proof goes exactly as in Proposition 4.2. □

The next corollary is valid if we work with *good* overconvergent varieties (which we may have assumed from the beginning).

Corollary 4.4. *If (C, O) is an (good) overconvergent variety and X is an algebraic variety over C, then* $\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X/O)$ *is an abelian subcategory of* $\mathrm{Isoc}^\dagger(X/O)$.

Proof. Using Proposition 4.6.3 of [Le Stum 2011], we may assume that X has a geometric realization over (C, O) and use the second part of Proposition 3.5.8 in the same paper. □

We could have included a description of constructible isocrystal as modules endowed with an overconvergent stratification on some geometric realization of X/O but we are heading towards a finer description (this is what the rest of this section is all about).

Recall that any overconvergent stratification will induce, by pull-back at each level, a usual stratification. This is a faithful construction and we want to show that

it is actually *fully* faithful when we work with constructible modules (in suitable geometric situations). Thus, we have the sequence of injective maps

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(\mathcal{F},\mathcal{G}) \hookrightarrow \mathrm{Hom}_{\mathrm{Strat}}(\mathcal{F},\mathcal{G}) \hookrightarrow \mathrm{Hom}(\mathcal{F},\mathcal{G})$$

and we wonder whether the first one is actually bijective. In order to do so, we will also have to study the injectivity of the maps in the sequence

$$\mathrm{Ext}_{\mathrm{Strat}^\dagger}(\mathcal{F},\mathcal{G}) \to \mathrm{Ext}_{\mathrm{Strat}}(\mathcal{F},\mathcal{G}) \to \mathrm{Ext}(\mathcal{F},\mathcal{G}).$$

We start with the following observation:

Proposition 4.5. *Let (X,V) be a variety over an overconvergent variety (C,O), $\alpha : U \hookrightarrow X$ the inclusion of an open subvariety of X and $\beta : Z \hookrightarrow X$ the inclusion of a closed complement. Let $\mathcal{F}'$ be an $i_Z^{-1}\mathcal{O}_V$-module and $\mathcal{F}''$ an $i_U^{-1}\mathcal{O}_V$-module. Then a usual (resp. an overconvergent) stratification on the direct sum $]\beta[_!\mathcal{F}' \oplus]\alpha[_*\mathcal{F}''$ is uniquely determined by its restrictions to $\mathcal{F}'$ and $\mathcal{F}''$.*

Proof. Let us denote by

$$\epsilon^{(n)} = \begin{pmatrix}]\beta[_!\epsilon'^{(n)} & \varphi_n \\ \psi_n &]\alpha[_*\epsilon''^{(n)} \end{pmatrix}$$

the stratification of $]\beta[_!\mathcal{F}' \oplus]\alpha[_*\mathcal{F}''$ (recall that the maps $]\beta[_!$ and $]\alpha[_*$ are fully faithful). Then the maps

$$\varphi_n : i_X^{-1}\mathcal{O}_{V^{(n)}} \otimes_{i_X^{-1}\mathcal{O}_V}]\alpha[_*\mathcal{F}'' \to]\beta[_!\mathcal{F}' \otimes_{i_X^{-1}\mathcal{O}_V} i_X^{-1}\mathcal{O}_{V^{(n)}}$$

and

$$\psi_n : i_X^{-1}\mathcal{O}_{V^{(n)}} \otimes_{i_X^{-1}\mathcal{O}_V}]\beta[_!\mathcal{F}' \to]\alpha[_*\mathcal{F}'' \otimes_{i_X^{-1}\mathcal{O}_V} i_X^{-1}\mathcal{O}_{V^{(n)}}$$

are necessarily zero, as one may see by considering the fibers.

On the other hand, denote by

$$\epsilon = \begin{pmatrix}]\beta[_!\epsilon' & \varphi \\ \psi &]\alpha[_*\epsilon'' \end{pmatrix}$$

the overconvergent stratification of $]\beta[_!\mathcal{F}' \oplus]\alpha[_*\mathcal{F}''$. Then the maps

$$\varphi :]p_2[^\dagger]\alpha[_*\mathcal{F}'' \simeq]p_1[^\dagger]\beta[_!\mathcal{F}' \quad \text{and} \quad \psi :]p_2[^\dagger]\beta[_!\mathcal{F}' \simeq]p_1[^\dagger]\alpha[_*\mathcal{F}''$$

are necessarily zero, as one may see by considering the fibers and using the fact that $p_i^\dagger$ commutes with $]\alpha[_*$ and $]\beta[_!$. □

We keep the assumptions and the notation of the proposition for a while and assume that $\mathcal{F}'$ and $\mathcal{F}''$ are both endowed with a usual (resp. an overconvergent) stratification. From the general fact that

$$\mathrm{Hom}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0 \quad \text{and} \quad \mathrm{Hom}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') = 0,$$

we can deduce that

$$\mathrm{Hom}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0 \quad \big(\text{resp. } \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0\big),$$
$$\mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') = 0 \quad \big(\text{resp. } \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') = 0\big).$$

Since we also know that

$$\mathrm{Ext}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0,$$

we can deduce the following result from the proposition:

Corollary 4.6. *If $\mathcal{F}'$ and $\mathcal{F}''$ are both endowed with a usual* (*resp. an overconvergent*) *stratification*, *then we have*

$$\mathrm{Ext}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0 \quad \big(\text{resp. } \mathrm{Ext}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{F}'') = 0\big).$$

Alternatively, it means that any short exact sequence of $i_Z^{-1}\mathcal{O}_V$-modules (resp. with a usual stratification, resp. with an overconvergent stratification)

$$0 \to]\alpha[_*\mathcal{F}'' \to \mathcal{F} \to]\beta[_!\mathcal{F}' \to 0$$

splits (and the splitting is compatible with the extra structure).

From the proposition, we may also deduce the following:

Corollary 4.7. *If $\mathcal{F}'$ and $\mathcal{F}''$ are both endowed with a usual stratification*, *then the map*

$$\mathrm{Ext}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') \hookrightarrow \mathrm{Ext}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}')$$

is injective. If $\mathcal{F}'$ and $\mathcal{F}''$ are both endowed with an overconvergent stratification, *then the maps*

$$\mathrm{Ext}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') \hookrightarrow \mathrm{Ext}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}') \hookrightarrow \mathrm{Ext}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{F}')$$

are injective.

Alternatively, it means that if $\mathcal{F}$ is an $i_X^{-1}\mathcal{O}_V$-module with a usual (resp. an overconvergent) stratification, *and* if the exact sequence of $i_X^{-1}\mathcal{O}_V$-modules

$$0 \to]\beta[_!\mathcal{F}_{|]Z[} \to \mathcal{F} \to]\alpha[_*\mathcal{F}_{|]U[} \to 0$$

splits, then the splitting is always compatible with the (resp. the overconvergent) stratifications.

We are now ready to prove our main result:

Proposition 4.8. *Let*

$$\begin{array}{ccccccc}
X & \hookrightarrow & P & \leftarrow & P_K & \leftarrow & V \\
\downarrow{\scriptstyle f} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle v_K} & & \downarrow{\scriptstyle u} \\
C & \hookrightarrow & S & \leftarrow & S_K & \leftarrow & O
\end{array}$$

be a formal morphism of overconvergent varieties with f quasicompact, v smooth at X, O locally separated and V a good neighborhood of X in $P_K \times_{S_K} O$. If $\mathcal{F}$ and $\mathcal{G}$ are two constructible $i_X^{-1}\mathcal{O}_V$-modules endowed with an overconvergent stratification, then

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(\mathcal{F},\mathcal{G}) \simeq \mathrm{Hom}_{\mathrm{Strat}}(\mathcal{F},\mathcal{G}).$$

Proof. Since we know that the map is injective, we may rephrase the assertion as follows: we are given a morphism $\varphi : \mathcal{F} \to \mathcal{G}$ of constructible $i_X^{-1}\mathcal{O}_V$-modules which is compatible with the usual stratifications and we have to show that φ is actually compatible with the overconvergent stratifications. This question is clearly local on O, which is locally compact. We may therefore assume that the image of O in S_K is contained in some S'_K with S' quasicompact. We may then pull back the diagram along $S' \to S$ and assume that X is finite-dimensional (use assertion (3) of Proposition 3.3). This will allow us to use noetherian induction.

We know (use, for example, Propositions 3.5 and 4.3) that there exists a dense open subset U of X such that the restrictions $\mathcal{F}''$ and $\mathcal{G}''$ to U of $\mathcal{F}$ and $\mathcal{G}$ are coherent. Moreover, it was shown in Corollary 3.4.10 of [Le Stum 2011] that the proposition is valid for $\mathcal{F}''$ and $\mathcal{G}''$ on U. Let us denote as usual by $\alpha : U \hookrightarrow X$ the inclusion map. Since $]\alpha[_*$ is fully faithful, we see that the proposition is valid for $]\alpha[_*\mathcal{F}''$ and $]\alpha[_*\mathcal{G}''$. In other words, we have a bijection

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'',]\alpha[_*\mathcal{G}'') \simeq \mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\alpha[_*\mathcal{G}''). \tag{10}$$

We denote now by $\beta : Z \hookrightarrow X$ the inclusion of a closed complement of U and let $\mathcal{F}'$ and $\mathcal{G}'$ be the restrictions of $\mathcal{F}$ and $\mathcal{G}$ to Z. We observe the following commutative diagram:

$$\begin{array}{ccc}
0 & & 0 \\
\downarrow & & \downarrow \\
\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'', \mathcal{G}) & \hookrightarrow & \mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'', \mathcal{G}) \\
\downarrow & & \downarrow \\
\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'',]\alpha[_*\mathcal{G}'') & \xrightarrow{\simeq} & \mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\alpha[_*\mathcal{G}'') \\
\downarrow & & \downarrow \\
\mathrm{Ext}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{G}') & \hookrightarrow & \mathrm{Ext}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{G}')
\end{array}$$

The columns are exact because $\mathrm{Hom}(]\alpha[_*\mathcal{F}'',]\beta[_!\mathcal{G}') = 0$, the bottom map is injective thanks to Corollary 4.7 and the middle map is the isomorphism (10). It follows from the five lemma (or an easy diagram chase) that the upper map is

necessarily bijective: we have

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'', \mathcal{G}) \simeq \mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'', \mathcal{G}''). \tag{11}$$

We turn now to the other side: by induction, the proposition is valid for $\mathcal{F}'$ and $\mathcal{G}'$ on Z, and since $]\beta[_!$ is fully faithful, it also holds for $]\beta[_!\mathcal{F}'$ and $]\beta[_!\mathcal{G}'$. Hence, we have

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}',]\beta[_!\mathcal{G}') \simeq \mathrm{Hom}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}',]\beta[_!\mathcal{G}'). \tag{12}$$

Now, we consider the commutative square

$$\begin{array}{ccc} \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}',]\beta[_!\mathcal{G}') & \xrightarrow{\simeq} & \mathrm{Hom}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}',]\beta[_!\mathcal{G}') \\ \downarrow \simeq & & \downarrow \simeq \\ \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}', \mathcal{G}) & \hookrightarrow & \mathrm{Hom}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}', \mathcal{G}) \end{array}$$

The vertical maps are bijective because $\mathrm{Hom}(]\beta[_!\mathcal{F}',]\alpha[_*\mathcal{G}'') = 0$ and the upper map is simply the isomorphism (12). It follows that we have an isomorphism

$$\mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_!\mathcal{F}', \mathcal{G}) \simeq \mathrm{Hom}_{\mathrm{Strat}}(]\beta[_!\mathcal{F}', \mathcal{G}). \tag{13}$$

In order to end the proof, we will need to kill another obstruction. Since the proposition holds for $]\alpha[_*\mathcal{F}''$ and *any* constructible $\mathcal{G}$, the following canonical map is necessarily injective:

$$\mathrm{Ext}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'', \mathcal{G}) \hookrightarrow \mathrm{Ext}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'', \mathcal{G}). \tag{14}$$

We consider now the commutative diagram with exact columns

$$\begin{array}{ccc} 0 & & 0 \\ \downarrow & & \downarrow \\ \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'', \mathcal{G}) & \xrightarrow{\simeq} & \mathrm{Hom}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'', \mathcal{G}) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_{\mathrm{Strat}^\dagger}(\mathcal{F}, \mathcal{G}) & \hookrightarrow & \mathrm{Hom}_{\mathrm{Strat}}(\mathcal{F}, \mathcal{G}) \\ \downarrow & & \downarrow \\ \mathrm{Hom}_{\mathrm{Strat}^\dagger}(]\beta[_\dagger\mathcal{F}', \mathcal{G}) & \xrightarrow{\simeq} & \mathrm{Hom}_{\mathrm{Strat}}(]\beta[_\dagger\mathcal{F}', \mathcal{G}) \\ \downarrow & & \downarrow \\ \mathrm{Ext}_{\mathrm{Strat}^\dagger}(]\alpha[_*\mathcal{F}'', \mathcal{G}) & \hookrightarrow & \mathrm{Ext}_{\mathrm{Strat}}(]\alpha[_*\mathcal{F}'', \mathcal{G}) \end{array}$$

The horizontal isomorphisms are just (11) and (13) and the bottom injection is (14). It is then sufficient to apply the five lemma again. □

We may reformulate the statement of the proposition as follows:

Corollary 4.9. *The forgetful functor from constructible $i_X^{-1}\mathcal{O}_V$-modules endowed with an overconvergent stratification to $i_X^{-1}\mathcal{O}_V$-modules endowed with a usual stratification is fully faithful.*

It is also worth mentioning the following immediate consequence:

Corollary 4.10. *If $\mathcal{F}$ and $\mathcal{G}$ are two constructible $i_X^{-1}\mathcal{O}_V$-modules endowed with an overconvergent stratification, then we have an injective map*

$$\mathrm{Ext}_{\mathrm{Strat}^\dagger}(\mathcal{F},\mathcal{G}) \hookrightarrow \mathrm{Ext}_{\mathrm{Strat}}(\mathcal{F},\mathcal{G}). \tag{15}$$

It means that if

$$0 \to \mathcal{F} \to \mathcal{G} \to \mathcal{H} \to 0 \tag{16}$$

is a short exact sequence of constructible $i_X^{-1}\mathcal{O}_V$-modules endowed with an overconvergent stratification, then any splitting for the usual stratifications will be compatible with the overconvergent stratifications. I strongly suspect that much more is actually true: if we are given an exact sequence (16) of constructible $i_X^{-1}\mathcal{O}_V$-modules endowed with usual stratifications and if the stratifications of $\mathcal{F}'$ and $\mathcal{F}''$ are overconvergent, then the stratification of $\mathcal{F}$ should also be overconvergent. In other words, the injective map (15) would be an isomorphism.

If (X, V) is a variety over an overconvergent variety (C, O), we will denote by

$$\mathrm{MIC}^\dagger_{\mathrm{cons}}(X, V/O)$$

the category of constructible $i_X^{-1}\mathcal{O}_V$-modules $\mathcal{F}$ endowed with an overconvergent connection (recall that it means that the connection extends to some overconvergent stratification). Then, we can also state the following corollary:

Corollary 4.11. *If* $\mathrm{Char}(K) = 0$, *then the realization functor induces an equivalence of categories*

$$\mathrm{Isoc}^\dagger_{\mathrm{cons}}(X_V/O) \simeq \mathrm{MIC}^\dagger_{\mathrm{cons}}(X, V/O).$$

As a consequence, we observe that we will have, for a constructible isocrystal E on X_V/O,

$$\Gamma(X_V/O, E) \simeq H^0_{\mathrm{dR}}(E_V),$$

and we expect the same to hold for higher cohomology spaces; we only know at this point that

$$H^1(X_V/O, E) \subset H^1_{\mathrm{dR}}(E_V).$$

Again, we need to work with good overconvergent varieties for the theorem to hold:

Theorem 4.12. *Assume that* $\mathrm{Char}(K) = 0$ *and that we are given a commutative diagram*

$$\begin{array}{ccccccc} X & \hookrightarrow & P & \leftarrow & P_K & \leftarrow & V \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle v} & & \downarrow{\scriptstyle v_K} & & \downarrow{\scriptstyle u} \\ C & \hookrightarrow & S & \leftarrow & S_K & \leftarrow & O \end{array} \tag{17}$$

in which P *is a formal scheme over* S *which is proper and smooth around* X, *and* V *is a neighborhood of the tube of* X *in* $P_K \times_{S_K} O$ *(and* O *is good in the neighborhood of* $]C[$*). Then the realization functor induces an equivalence of categories*

$$\mathrm{Isoc}^{\dagger}_{\mathrm{cons}}(X/O) \simeq \mathrm{MIC}^{\dagger}_{\mathrm{cons}}(X, V/O)$$

between constructible overconvergent isocrystals on X/O *and constructible* $i_X^{-1}\mathcal{O}_V$*-modules endowed with an overconvergent connection.*

Proof. Using the second assertion of Proposition 3.5.8 in [Le Stum 2011], this follows immediately from Corollary 4.11. □

As a consequence of the theorem, we see that the notion of a constructible module endowed with an overconvergent connection only depends on X and *not* on the choice of the geometric realization (17). It is likely that this could have been proven directly using Berthelot's technique of diagonal embedding. However, we believe that our method is much more natural because functoriality is built in.

Acknowledgments

Many thanks to Tomoyuki Abe, Pierre Berthelot, Florian Ivorra, Vincent Mineo-Kleiner, Laurent Moret-Bailly, Matthieu Romagny and Atsushi Shiho, with whom I had interesting conversations related to some questions discussed here.

References

[Abe 2013] T. Abe, "Langlands correspondence for isocrystals and existence of crystalline companion for curves", preprint, 2013. arXiv 1310.0528

[Abe and Caro 2013] T. Abe and D. Caro, "Theory of weights in p-adic cohomology", preprint, 2013. arXiv 1303.0662

[Berkovich 1990] V. G. Berkovich, *Spectral theory and analytic geometry over non-Archimedean fields*, Mathematical Surveys and Monographs **33**, American Mathematical Society, Providence, RI, 1990. MR 1070709 Zbl 0715.14013

[Berkovich 1993] V. G. Berkovich, "Étale cohomology for non-Archimedean analytic spaces", *Inst. Hautes Études Sci. Publ. Math.* **78** (1993), 5–161. MR 1259429 Zbl 0804.32019

[Berthelot 1996a] P. Berthelot, "Cohomologie rigide et cohomologie à support propre, I", preprint 96(03):89, Institut de Recherche Math. de Rennes, 1996, available at https://perso.univ-rennes1.fr/pierre.berthelot/publis/Cohomologie_Rigide_I.pdf.

[Berthelot 1996b] P. Berthelot, "$\mathscr{D}$-modules arithmétiques, I: Opérateurs différentiels de niveau fini", *Ann. Sci. École Norm. Sup.* (4) **29**:2 (1996), 185–272. MR 1373933 Zbl 0886.14004

[Berthelot 2002] P. Berthelot, "Introduction à la théorie arithmétique des $\mathscr{D}$-modules", pp. 1–80 in *Cohomologies p-adiques et applications arithmétiques, II*, edited by P. Berthelot et al., Astérisque **279**, Société Mathématique de France, Paris, 2002. MR 1922828 Zbl 1098.14010

[Borel et al. 1987] A. Borel, P.-P. Grivel, B. Kaup, A. Haefliger, B. Malgrange, and F. Ehlers, *Algebraic D-modules*, Perspectives in Mathematics **2**, Academic Press, Boston, 1987. MR 882000 Zbl 0642.32001

[Caro 2009] D. Caro, "$\mathscr{D}$-modules arithmétiques surholonomes", *Ann. Sci. Éc. Norm. Supér.* (4) **42**:1 (2009), 141–192. MR 2518895 Zbl 1168.14013

[Deligne 1970] P. Deligne, *Équations différentielles à points singuliers réguliers*, Lecture Notes in Mathematics **163**, Springer, 1970. MR 0417174 Zbl 0244.14004

[Iversen 1986] B. Iversen, *Cohomology of sheaves*, Springer, 1986. MR 842190 Zbl 1272.55001

[Kashiwara 1984] M. Kashiwara, "The Riemann–Hilbert problem for holonomic systems", *Publ. Res. Inst. Math. Sci.* **20**:2 (1984), 319–365. MR 743382 Zbl 0566.32023

[Kashiwara 2004] M. Kashiwara, "t-structures on the derived categories of holonomic $\mathscr{D}$-modules and coherent $\mathscr{O}$-modules", *Mosc. Math. J.* **4**:4 (2004), 847–868, 981. MR 2124169 Zbl 1073.14023

[Le Stum 2011] B. Le Stum, *The overconvergent site*, Mém. Soc. Math. Fr. (N.S.) **127**, 2011. MR 2952779 Zbl 1246.14028

[Le Stum 2014] B. Le Stum, "Constructible ∇-modules on curves", *Selecta Math.* (*N.S.*) **20**:2 (2014), 627–674. MR 3177929 Zbl 1320.14035

[Zureick-Brown 2010] D. Zureick-Brown, *Rigid cohomology for algebraic stacks*, Ph.D. thesis, University of California Berkeley, 2010, available at http://www.mathcs.emory.edu/~dzb/math/papers/davidBrownThesis.pdf.

[Zureick-Brown 2014] D. Zureick-Brown, "Cohomology with closed support on the overconvergent site", preprint, 2014. arXiv 1408.1970

Communicated by Kiran S. Kedlaya
Received 2016-01-26 Revised 2016-09-05 Accepted 2016-11-12

bernard.le-stum@univ-rennes1.fr *Institut de Recherche Mathematique (IRMAR), Universite de Rennes I, 35042 Rennes, France*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2153

Canonical heights on genus-2 Jacobians

Jan Steffen Müller and Michael Stoll

Scale New Heights!

Motto of International (now Jacobs) University Bremen,
where the first author started his Ph.D. under the supervision of the second.

Let K be a number field and let C/K be a curve of genus 2 with Jacobian variety J. We study the canonical height $\hat{h}\colon J(K) \to \mathbb{R}$. More specifically, we consider the following two problems, which are important in applications:

(1) for a given $P \in J(K)$, compute $\hat{h}(P)$ efficiently;
(2) for a given bound $B > 0$, find all $P \in J(K)$ with $\hat{h}(P) \leq B$.

We develop an algorithm running in polynomial time (and fast in practice) to deal with the first problem. For the second problem, we show how to tweak the naive height h to obtain significantly improved bounds for the difference $h - \hat{h}$, which allows a much faster enumeration of the desired set of points.

Our approach is to use the standard decomposition of $h(P) - \hat{h}(P)$ as a sum of local "height correction functions". We study these functions carefully, which leads to efficient ways of computing them and to essentially optimal bounds. To get our polynomial-time algorithm, we have to avoid the factorization step needed to find the finite set of places where the correction might be nonzero. The main innovation is to replace factorization into primes by factorization into coprimes.

Most of our results are valid for more general fields with a set of absolute values satisfying the product formula.

An errata was submitted on 30 Dec 2022 and posted online on 16 Feb 2023.

MSC2010: primary 11G50; secondary 11G30, 11G10, 14G40, 14Q05, 14G05.
Keywords: canonical height, hyperelliptic curve, curve of genus 2, Jacobian surface, Kummer surface.

1. Introduction

Let K be a global field and let C/K be a curve of genus 2 with Jacobian variety J. There is a map $\kappa\colon J \to \mathbb{P}^3$ that corresponds to the class of twice the theta divisor on J; it identifies a point on J with its negative, and its image is the Kummer surface KS of J. Explicit versions of κ can be found in the book [Cassels and Flynn 1996] for C given in the form $y^2 = f(x)$ and in the paper [Müller 2010] by the first author for general C (also in characteristic 2). Thus κ gives rise to a height function $h\colon J(K) \to \mathbb{R}$, which we call the *naive height* on J. It is defined by

$$h(P) = \sum_{v \in M_K} \log \max\{|\kappa_1(P)|_v, |\kappa_2(P)|_v, |\kappa_3(P)|_v, |\kappa_4(P)|_v\},$$

where $\kappa(P) = (\kappa_1(P) : \kappa_2(P) : \kappa_3(P) : \kappa_4(P))$, M_K is the set of places of K, and $|\cdot|_v$ is the v-adic absolute value, normalized so that the product formula holds:

$$\prod_{v \in M_K} |x|_v = 1 \quad \text{for all } x \in K^\times.$$

By general theory [Hindry and Silverman 2000, Chapter B] the limit

$$\hat{h}(P) = \lim_{n \to \infty} \frac{h(nP)}{n^2}$$

exists; it is called the *canonical height* (or *Néron–Tate height*) of $P \in J(K)$. The difference $h - \hat{h}$ is bounded. The canonical height induces a positive definite quadratic form on $J(K)/J(K)_{\mathrm{tors}}$ (and on the $\mathbb{R}$-vector space $J(K) \otimes_{\mathbb{Z}} \mathbb{R}$).

In this paper, we tackle the following two problems:

Problem 1.1. Find an efficient algorithm for the computation of $\hat{h}(P)$ for a given point $P \in J(K)$.

Problem 1.2. Find an efficient algorithm for the enumeration of all $P \in J(K)$ which satisfy $\hat{h}(P) \le B$, where B is a given real number.

These problems are important because such algorithms are needed if we want to saturate a given finite-index subgroup of $J(K)$ (see the discussion at the end of Section 18). This, in turn, is necessary for the computation of generators of $J(K)$. Such generators are required, for instance, to carry out the method described in [Bugeaud et al. 2008] for the computation of all integral points on a hyperelliptic curve over $\mathbb{Q}$. Furthermore, the regulator of $J(K)$ appearing in the conjecture of Birch and Swinnerton-Dyer is the Gram determinant of a set of generators of $J(K)/J(K)_{\mathrm{tors}}$ with respect to the canonical height. So Problems 1.1 and 1.2 are also important in the context of gathering numerical evidence for this conjecture as in [Flynn et al. 2001].

It is a classical fact, going back to work by Néron [1965], that $\hat{h}(P)$ and the difference $h(P) - \hat{h}(P)$ can be decomposed into a finite sum of local terms. In our situation, this can be done explicitly as follows. The duplication map $P \mapsto 2P$ on J induces a morphism $\delta\colon \mathrm{KS} \to \mathrm{KS}$, given by homogeneous polynomials $(\delta_1, \delta_2, \delta_3, \delta_4)$ of degree 4; explicit equations can again be found in [Cassels and Flynn 1996] and [Müller 2010]. For a point $Q \in J(K_v)$, where K_v is the completion of K at a place $v \in M_K$, such that $\kappa(Q) = (x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(K_v)$, we set

$$\tilde{\varepsilon}_v(Q) = -\log \max\{|\delta_j(x_1, x_2, x_3, x_4)|_v : 1 \le j \le 4\} + 4 \log \max\{|x_j|_v : 1 \le j \le 4\}.$$

Note that this does not depend on the scaling of the coordinates. We can then write $\hat{h}(P)$ in the following form (compare Lemma 2.4):

$$\hat{h}(P) = h(P) - \sum_{v \in M_K} \sum_{n=0}^{\infty} 4^{-(n+1)} \tilde{\varepsilon}_v(2^n P).$$

We set, for $Q \in J(K_v)$ as above,

$$\tilde{\mu}_v(Q) = \sum_{n=0}^{\infty} 4^{-(n+1)} \tilde{\varepsilon}_v(2^n Q), \tag{1-1}$$

and we deduce the decomposition

$$h(P) - \hat{h}(P) = \sum_{v \in M_K} \tilde{\mu}_v(P), \tag{1-2}$$

which is valid for all points $P \in J(K)$. In addition, $\tilde{\varepsilon}_v = \tilde{\mu}_v = 0$ for all but finitely many v (the exceptions are among the places of bad reduction, the places where the given equation of C is not integral and the archimedean places). The maps $\tilde{\varepsilon}_v \colon J(K_v) \to \mathbb{R}$ are continuous maps (with respect to the v-adic topology) with compact domains, so they are bounded. Therefore $\tilde{\mu}_v$ is also bounded.

Let us first discuss Problem 1.1. Because of (1-2), it suffices to compute $h(P)$ (which is easy) and $\sum_{v \in M_K} \tilde{\mu}_v(P)$ in order to compute $\hat{h}(P)$ for a point $P \in J(K)$. Building on earlier work of Flynn and Smart [1997], the second author introduced an algorithm for the computation of $\tilde{\mu}_v(P)$ in [Stoll 2002]. One of the main problems with this approach is that we need integer factorization to compute the sum $\tilde{\mu}^{\mathrm{f}}(P) := \sum_v \tilde{\mu}_v(P)$, where v runs through the finite primes v such that $\tilde{\mu}_v(P) \neq 0$, because we need to find these primes, or at least a finite set of primes containing them.

We use an idea which was already exploited in [Müller and Stoll 2016] to obtain a polynomial-time algorithm for the computation of the canonical height of a point on an elliptic curve (in fact, we first used this technique in genus 2 and only later realized that it also works, and is actually easier to implement, for elliptic curves). When v is nonarchimedean, there is a constant $c_v > 0$ such that the function

$$\mu_v := \tilde{\mu}_v / c_v$$

maps $J(K_v)$ to $\mathbb{Q}$. More precisely, $\tilde{\mu}^{\mathrm{f}}(P)$ is a sum of rational multiples of logarithms of positive integers. As in [Müller and Stoll 2016], we find a bound on the denominator of μ_v that depends only on the valuation of the discriminant; this allows us to devise an algorithm that computes $\tilde{\mu}^{\mathrm{f}}(P)$ in quasilinear time. We can compute $\tilde{\mu}_v(P)$ for archimedean v essentially from the definition of $\tilde{\mu}_v$. This leads to a factorization-free algorithm that computes $\hat{h}(P)$ in polynomial time:

Theorem 1.3. *Let J be the Jacobian of a curve of genus 2 defined over $\mathbb{Q}$, and let $P \in J(\mathbb{Q})$. There is an algorithm that computes $\hat{h}(P)$ in time quasilinear in the size of the coordinates of P and the coefficients of the given equation of C, and quasiquadratic in the desired number of digits of precision.*

See Theorem 14.5 for a precise statement. We expect a similar result to be true for any number field K in place of $\mathbb{Q}$.

We now move on to Problem 1.2. If we have an upper bound β for $h - \hat{h}$, then the set of all points $P \in J(K)$ such that $h(P) \leq B + \beta$ contains the set $\{P \in J(K) : \hat{h}(P) \leq B\}$. Since the naive height h is a logarithmic height, β contributes exponentially to the size of the box we need to search for the enumeration. Therefore it is crucial to keep β as small as possible.

We write $\tilde{\beta}_v = \max\{\tilde{\mu}_v(Q) : Q \in J(K_v)\}$, and we obtain the bound

$$h(P) - \hat{h}(P) \leq \sum_{v \in M_K} \tilde{\beta}_v$$

from (1-2). If we write

$$\tilde{\gamma}_v = \max\{\tilde{\varepsilon}_v(Q) : Q \in J(K_v)\},$$

then clearly $\frac{1}{4}\tilde{\gamma}_v \le \tilde{\beta}_v \le \frac{1}{3}\tilde{\gamma}_v$. In [Stoll 1999], it is shown that for curves given in the form $y^2 = f(x)$, where f has v-adically integral coefficients, we have

$$\tilde{\gamma}_v \le -\log|2^4 \operatorname{disc}(f)|_v = -\log|2^{-4}\Delta|_v,$$

with $\operatorname{disc}(f)$ denoting the discriminant of f considered as a polynomial of degree 6 and Δ denoting the discriminant of the given equation of C. When v is nonarchimedean and the normalized additive valuation of Δ is 1, we can take $\tilde{\gamma}_v = \tilde{\beta}_v = 0$ [Stoll 2002].

The results of the present paper improve on this; they are based on a careful study of the functions $\tilde{\mu}_v$. It turns out that when v is nonarchimedean, the set of points where μ_v (or equivalently, $\tilde{\mu}_v$) vanishes forms a group. Moreover, the function μ_v factors through the component group of the Néron model of J when the given model of C/K_v, which we assume to have v-integral coefficients in the following, has rational singularities; see Theorem 7.4. If the minimal proper regular model of C is semistable, then we can use results of Zhang [1993] and Heinz [2004] to give explicit formulas for μ_v in terms of the resistance function on the reduction graph of C (which is essentially the dual graph of the special fiber of the minimal proper regular model, suitably metrized). We use this to find simple explicit formulas for μ_v that apply in the most frequent cases of bad reduction, namely nodal or cuspidal reduction. These explicit formulas give us the optimal bounds for $\tilde{\mu}_v$ in these cases. By reducing to the semistable case and tracking how μ_v changes as we change the Weierstrass equation of C, we deduce the general upper bound

$$\tilde{\beta}_v \le -\tfrac{1}{4}\log|\Delta|_v \tag{1-3}$$

for nonarchimedean v; see Theorem 11.3.

When v is archimedean, we also get a new bound for $\tilde{\mu}_v$ by iterating the bound obtained in [Stoll 1999], leading to vast improvements for $\tilde{\beta}_v$. Combining the archimedean and nonarchimedean bounds, we find a nearly optimal bound β for $h - \hat{h}$.

To get even smaller search spaces for the enumeration, we make use of the observation that we can replace the naive height h by any function h' such that $|h' - h|$ is bounded. Using the results on nearly optimal bounds for μ_v and such a modified naive height h' (which is also better suited than h for the enumeration process itself) we get a much smaller bound on the difference $h' - \hat{h}$ than what was previously possible. This makes the enumeration feasible in many cases that were completely out of reach so far.

As an example, we compute explicit generators for the Mordell–Weil group of the Jacobian of the curve

$$C\colon\ y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600 \quad (1\text{-}4)$$

over $\mathbb{Q}$, conditional on the generalized Riemann hypothesis (which is needed to show that the rank is 22). See Proposition 19.1. This curve has at least 642 rational points, which is the current record for the largest number of known rational points on a curve of genus 2; see [Stoll 2008].

The paper is divided into four parts. In Part I, we first generalize the usual notion of the naive height on projective space and clarify the relation between these generalized naive heights and suitable canonical heights, all in Section 2. We then introduce local height correction functions ε and μ ($= \mu_v$ in the notation introduced above) on the Jacobian of a genus-2 curve over a nonarchimedean local field in Section 3. This is followed in Section 4 by a study of certain canonical local heights constructed in terms of μ. We close Part I by introducing and investigating the notion of stably minimal Weierstrass models of curves of genus 2 in Section 5 and recalling some well-known results on Igusa invariants in Section 6.

Part II is in some sense the central part of the present paper. Here we study the local height correction function μ over a nonarchimedean local field. Using Picard functors, we show in Section 7 that μ factors through the component group of the Néron model of the Jacobian when the given model of the curve has rational singularities. We then relate μ to the reduction graph of C in Section 8. Building on this, the following sections contain simple explicit formulas for μ when the reduction of the curve is nodal (Section 9), respectively cuspidal (Section 10). A simple argument then gives the improved general upper bound (1-3) for μ; see Section 11.

In Part III we describe our factorization-free algorithm for the computation of $\hat{h}(P)$ for $P \in J(K)$, where K is a global field. We start in Section 12 by showing how to compute $\mu_v(P)$ for nonarchimedean v, using a bound on its denominator. The following section deals with archimedean places, before we finally combine these results in Section 14 into an algorithm for the computation of $\hat{h}(P)$ that runs in polynomial time; this proves Theorem 1.3. Some examples are discussed in Section 15.

In Part IV we turn to Problem 1.2. Section 16 contains two methods for bounding $\tilde{\mu}_v$ for archimedean v. In Section 17 we describe a modified naive height h' such that the bound on the difference $h' - \hat{h}$ becomes small. We use this, the results of Section 16, and our nearly optimal bounds for the nonarchimedean height correction functions from Part II to give an efficient algorithm for the enumeration of the set of rational points with bounded canonical height in Section 18. In Section 19 we compute generators of the Mordell–Weil group of the record curve (1-4).

Part I. Generalities on heights and genus-2 Jacobians

2. Generalized naive heights

Let K be a field with a set M_K of places v and associated absolute values $|\cdot|_v$ satisfying the product formula

$$\prod_{v\in M_K} |x|_v = 1 \quad \text{for all } x \in K^\times.$$

We write K_v for the completion of K at v. For a tuple $x = (x_1, \dots, x_m) \in K_v^m$ we set $\|x\|_v = \max\{|x_1|_v, \dots, |x_m|_v\}$.

In the following we will introduce some flexibility into our notion of height on projective spaces. (This is similar to the framework of "admissible families" in [Zarkhin 1995].)

Definition 2.1. (1) Let $v \in M_K$. A *local height function* on $\mathbb{P}^m$ at v is a map $h_v\colon K_v^{m+1} \setminus \{0\} \to \mathbb{R}$ such that

(i) $h_v(\lambda x) = \log|\lambda|_v + h_v(x)$ for all $x \in K_v^{m+1} \setminus \{0\}$ and all $\lambda \in K_v^\times$, and

(ii) $\big|h_v(x) - \log\|x\|_v\big|$ is bounded.

(2) A function $h\colon \mathbb{P}^m(K) \to \mathbb{R}$ is a *height* on $\mathbb{P}^m$ over K if there are local height functions h_v such that for all $x \in \mathbb{P}^m(K)$ we have

$$h((x_1 : x_2 : \cdots : x_{m+1})) = \sum_{v\in M_K} h_v(x_1, x_2, \dots, x_{m+1})$$

and $h_v(x) = \log\|x\|_v$ for all but finitely many places v.

Note that property (i) of local height functions together with the product formula imply that h is invariant under scaling of the coordinates and hence is well-defined.

One example of such a height is the standard height h_{std}, which we obtain by setting $h_v(x) = \log\|x\|_v$ for all v. We then have the following simple fact.

Lemma 2.2. *Let h be any height on $\mathbb{P}^m$ over K and let h_{std} be the standard height. Then there is a constant $c = c(h)$ such that*

$$|h(P) - h_{\mathrm{std}}(P)| \le c \quad \textit{for all } P \in \mathbb{P}^m(K).$$

Proof. This follows from property (ii) of local height functions and the requirement that $h_v(x) = \log\|x\|_v$ for all but finitely many v. □

Example 2.3. Other examples of heights can be obtained in the following way. For each place v, fix a linear form $l_v(x_1, \dots, x_{m+1}) = a_{v,1}x_1 + \cdots + a_{v,m+1}x_{m+1}$, with

$a_{v,1}, \dots, a_{v,m+1} \in K_v$ and $a_{v,m+1} \neq 0$, such that $l_v(x) = x_{m+1}$ for all but finitely many v. Then

$$h((x_1 : \cdots : x_m : x_{m+1})) = \sum_{v \in M_K} \log \max\{|x_1|_v, \dots, |x_m|_v, |l_v(x_1, \dots, x_{m+1})|_v\}$$

is a height on $\mathbb{P}^m$.

More generally, we could consider a family of automorphisms A_v of K_v^{m+1} with A_v equal to the identity for all but finitely many v, and take

$$h(x) = \sum_{v \in M_K} \log \max \|A_v(x)\|_v.$$

Now consider a projective variety $V \subset \mathbb{P}^m_K$ and an endomorphism $\varphi\colon V \to V$ of degree d (i.e., given by homogeneous polynomials of degree d). Then by general theory (see, e.g., [Hindry and Silverman 2000, Theorem B.2.5]) $|h_{\mathrm{std}}(\varphi(P)) - d h_{\mathrm{std}}(P)|$ is bounded on $V(K)$. We write $\varphi^{\circ n}$ for the n-fold iteration of φ. Then the *canonical height*

$$\hat{h}(P) = \lim_{n \to \infty} d^{-n} h_{\mathrm{std}}(\varphi^{\circ n}(P))$$

exists (and satisfies $\hat{h}(\varphi(P)) = d\hat{h}(P)$) [Hindry and Silverman 2000, Theorem B.4.1]. Let h be any height on $\mathbb{P}^m$. Since $|h - h_{\mathrm{std}}|$ is bounded, we can replace h_{std} by h in the definition of $\hat{h}$ without changing the result. We can then play the usual telescoping series trick in our more general setting.

Lemma 2.4. *Let*

$$\varphi((x_1 : \cdots : x_{m+1})) = (\varphi_1(x) : \cdots : \varphi_{m+1}(x))$$

with homogeneous polynomials $\varphi_j \in K[x_1, \dots, x_{m+1}]$ of degree d. We have

$$\hat{h}(P) = h(P) - \sum_{v \in M_K} \tilde{\mu}_v(P),$$

where

$$\tilde{\mu}_v(P) = \sum_{n=0}^{\infty} d^{-(n+1)} \tilde{\varepsilon}_v(\varphi^{\circ n}(P))$$

and, when $P = (x_1 : \cdots : x_{m+1})$ and $x = (x_1, \dots, x_{m+1})$,

$$\tilde{\varepsilon}_v(P) = d h_v(x) - h_v(\varphi_1(x), \dots, \varphi_{m+1}(x)).$$

Proof. Note that $\tilde{\varepsilon}_v$ is well-defined: scaling x by λ adds $|\lambda|_v$ to $h_v(x)$ and $d|\lambda|_v$ to $h_v(\varphi_1(x), \dots, \varphi_{m+1}(x))$. Let x be projective coordinates for P and write $x^{(n)}$ for

the result of applying $(\varphi_1, \dots, \varphi_{m+1})$ n times to $x = x^{(0)}$. Then

$$\begin{aligned}
\hat{h}(P) &= \lim_{n\to\infty} d^{-n} h(\varphi^{\circ n}(P)) \\
&= h(P) + \sum_{n=0}^{\infty} d^{-(n+1)} \big(h(\varphi^{\circ(n+1)}(P)) - d h(\varphi^{\circ n}(P))\big) \\
&= h(P) + \sum_{n=0}^{\infty} d^{-(n+1)} \sum_{v\in M_K} \big(h_v(x^{(n+1)}) - d h_v(x^{(n)})\big) \\
&= h(P) - \sum_{v\in M_K} \sum_{n=0}^{\infty} d^{-(n+1)} \tilde{\varepsilon}_v(\varphi^{\circ n}(P)) \\
&= h(P) - \sum_{v\in M_K} \tilde{\mu}_v(P). \qquad \square
\end{aligned}$$

We call the functions $\tilde{\mu}_v \colon \mathbb{P}^m(K_v) \to \mathbb{R}$ *local height correction functions*.

Note that when K_v is a discretely valued field such that $|x|_v = \exp(-c_v v(x))$ for $x \in K^\times$ with a constant $c_v > 0$ (and where we abuse notation and write $v \colon K_v^\times \twoheadrightarrow \mathbb{Z}$ also for the normalized additive valuation associated to the place v) and $h = h_{\mathrm{std}}$, then we have

$$\tilde{\mu}_v(P) = c_v \mu_v(P) \quad \text{and} \quad \tilde{\varepsilon}_v(P) = c_v \varepsilon_v(P),$$

where

$$\mu_v(P) = \sum_{n=0}^{\infty} d^{-(n+1)} \varepsilon_v(P)$$

and

$$\varepsilon_v(P) = \min\{v(\varphi_1(x)), \dots, v(\varphi_{m+1}(x))\} - d \min\{v(x_1), \dots, v(x_{m+1})\},$$

if $x = (x_1, \dots, x_{m+1})$ are homogeneous coordinates for P. This is the situation that we will study in some detail in Part II of this paper, for the special case when $V \subset \mathbb{P}^3$ is the Kummer surface associated to a curve of genus 2 and its Jacobian J and φ is the duplication map (then $d = 4$).

To deal with Problem 1.1, we work with the standard height h_{std}. We use our detailed results on the local height correction functions to deduce a bound on the denominator of μ_v (its values are rational) in terms of the valuation of the discriminant of the curve. This is the key ingredient that leads to our new factorization-free and fast algorithm for computing $\hat{h}$; see Part III.

To deal with Problem 1.2, we use the flexibility in choosing the (naive) height h and modify the standard height in such a way that the sum $\sum_{v\in M_K} \sup \tilde{\mu}_v(J(K_v))$ that bounds the difference $h - \hat{h}$ is as small as we can make it. The local height functions we use are as in Example 2.3 above, with $l_v(x_1, x_2, x_3, x_4) = x_4/s_v$ for

certain $s_v \in K_v^\times$ in most cases. Every height function of this type has the property that for any point $P = (x_1 : x_2 : x_3 : x_4) \in \mathbb{P}^3(K)$ different from $(0:0:0:1)$ we have

$$0 \le h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le h(P).$$

This is relevant, since we can fairly easily enumerate all points P as above that are on the Kummer surface and satisfy $h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le B$; see Part IV. Refinements of the standard height constructed using Arakelov theory were also used by Holmes [2014] to give an "in principle" algorithm for the enumeration of points of bounded canonical height on Jacobians of hyperelliptic curves over global fields.

3. Local height correction functions for genus-2 Jacobians

Until further notice, we let k be a nonarchimedean local field with additive valuation v, normalized to be surjective onto $\mathbb{Z}$. Let $\mathcal{O}$ denote the valuation ring of k with residue class field $\mathfrak{k}$ and let π be a uniformizing element of $\mathcal{O}$. We consider a smooth projective curve C of genus 2 over k, given by a Weierstrass equation

$$Y^2 + H(X, Z)Y = F(X, Z) \tag{3-1}$$

in weighted projective space $\mathbb{P}_k(1, 3, 1)$, with weights 1, 3 and 1 assigned to the variables X, Y and Z, respectively. Here

$$F(X, Z) = f_0Z^6 + f_1XZ^5 + f_2X^2Z^4 + f_3X^3Z^3 + f_4X^4Z^2 + f_5X^5Z + f_6X^6$$

and

$$H(X, Z) = h_0Z^3 + h_1XZ^2 + h_2X^2Z + h_3X^3$$

are binary forms of degrees 6 and 3, respectively, such that the discriminant $\Delta(F, H)$ of the Weierstrass equation (3-1) is nonzero. In characteristic different from 2, this discriminant is defined as

$$\Delta(F, H) = 2^{-12} \operatorname{disc}(4F + H^2) \in \mathbb{Z}[h_0, \dots, h_3, f_0, \dots, f_6],$$

and in general, we define it by the generic polynomial given by this formula. The curve defined by the equation is smooth if and only if $\Delta(F, H) \neq 0$.

For the remainder of this section we assume that $F, H \in \mathcal{O}[X, Z]$, so that (3-1) defines an *integral Weierstrass model* $\mathcal{C}$ of the curve in the terminology of Section 5 below. The discriminant of this model is then defined to be $\Delta(\mathcal{C}) := \Delta(F, H)$. We may assume that C is given by such an integral equation if k is the completion at a nonarchimedean place of a number field K and C is obtained by base change from K, since we can choose a globally integral Weierstrass equation for the curve. But also in general, we can always assume that C is given by an integral equation after applying a transformation defined over k, since we know from Corollary 4.6 in

the next section how the local height correction function μ defined in Definition 3.1 below behaves under such transformations.

We now generalize the definition of ε given in [Stoll 2002] (where the author works with Weierstrass equations that have $H = 0$) to our more general setting. As in the Introduction, let J denote the Jacobian of C and let KS be its Kummer surface, constructed explicitly together with an explicit embedding into $\mathbb{P}^3$ in [Cassels and Flynn 1996] in the case $H = 0$ and in [Müller 2010] in the general case. Also let $\kappa : J \to \mathbb{P}^3$ denote the composition of the quotient map from J to KS with this embedding; it maps the origin $O \in J(k)$ to the point $(0 : 0 : 0 : 1)$. A quadruple $x = (x_1, x_2, x_3, x_4) \in k^4$ is called a set of *Kummer coordinates* on KS if x is a set of projective coordinates for a point in $\mathrm{KS}(k)$; we denote the set of sets of Kummer coordinates on KS by $\mathrm{KS}_{\mathbb{A}}$ (this is the set of k-rational points on the pointed affine cone over KS). For $x \in \mathrm{KS}_{\mathbb{A}}$ we write $v(x) = \min\{v(x_1), \dots, v(x_4)\}$, and we say that x is *normalized* if $v(x) = 0$. If $P \in J(k)$, we say that $x \in \mathrm{KS}_{\mathbb{A}}$ is a set of *Kummer coordinates for* P if $\kappa(P) = (x_1 : x_2 : x_3 : x_4)$.

We let δ denote the duplication map on KS, which is given by homogeneous polynomials $\delta_1, \dots, \delta_4 \in \mathcal{O}[x_1, \dots, x_4]$ of degree 4 such that $\delta(0, 0, 0, 1) = (0, 0, 0, 1)$. We recall that there is a symmetric matrix $B = (B_{ij})_{1 \le i,j \le 4}$ of polynomials that are bihomogeneous of degree 2 in $x_1, \dots, x_4$ and also in $y_1, \dots, y_4$ and have coefficients in $\mathcal{O}$. They have the following properties; see Chapter 3 of [Cassels and Flynn 1996] and [Müller 2010].

(i) Let $x, y \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates for $P, Q \in J(k)$. Then there are Kummer coordinates $w, z \in \mathrm{KS}_{\mathbb{A}}$ for $P + Q$ and $P - Q$, respectively, such that

$$w * z := (w_i z_j + n_{ij} w_j z_i)_{1 \le i,j \le 4} = B(x, y)$$

and hence $v(w) + v(z) = v(B(x, y))$; here $n_{ij} = 1$ if $i \ne j$ and $n_{ij} = 0$ if $i = j$.

(ii) If $x \in \mathrm{KS}_{\mathbb{A}}$, then $B(x, x) = \delta(x) * (0, 0, 0, 1)$.

We specialize the notions introduced in Section 2 to our situation: we consider the Kummer surface $\mathrm{KS} \subset \mathbb{P}^3$ with the duplication map δ of degree $d = 4$. We use the standard local height on $\mathbb{P}^3$.

Definition 3.1. Let $x \in \mathrm{KS}_{\mathbb{A}}$ be a set of Kummer coordinates on KS. Then we set

$$\varepsilon(x) = v(\delta(x)) - 4v(x) \in \mathbb{Z} \quad \text{and} \quad \mu(x) = \sum_{n=0}^{\infty} \frac{1}{4^{n+1}} \varepsilon(\delta^{\circ n}(x)),$$

where $\delta^{\circ n}$ denotes the n-fold composition $\delta \circ \cdots \circ \delta$.

Because δ is given by homogeneous polynomials of degree 4, $\varepsilon(x)$ does not depend on the scaling of x, so it makes sense to define $\varepsilon(P) = \varepsilon(x)$ for points $P \in \mathrm{KS}(k)$, where $x \in \mathrm{KS}_{\mathbb{A}}$ is any set of Kummer coordinates for P, and to define

$\varepsilon(P) = \varepsilon(\kappa(P))$ for points $P \in J(k)$. We likewise extend the definition of μ. Then we have

$$\mu(2P) - 4\mu(P) = -\varepsilon(P) \quad \text{for all } P \in J(k).$$

Note that our assumption $F, H \in \mathcal{O}[X, Z]$ implies that $\varepsilon \geq 0$. If k is a local field (as we assume here), then $\mathrm{KS}(k)$ is compact in the v-adic topology, and ε is continuous, so ε is bounded.

Remark 3.2. More generally, if k is a field with a discrete valuation and not of characteristic 2, then the arguments in [Stoll 1999] show that when $H = 0$, $\varepsilon \leq v(2^4 \operatorname{disc}(F))$, so ε is bounded also for these more general fields.

If k is any field with a discrete valuation, then one can still conclude that ε is bounded, by making use of the fact that the duplication map is well-defined on KS, which implies that the ideal generated by the δ_j and the polynomial δ_0 defining KS contains a power of the irrelevant ideal. So for some $N > 0$, one can express every x_j^N as a linear combination of $\delta_0(x), \ldots, \delta_4(x)$ with coefficients that are homogeneous polynomials of degree $N - 4$ with coefficients in k. The negative of the minimum of the valuations of these coefficients then gives a bound for ε.

Remark 3.3. If k is the completion of a global field at a place v, then for $\alpha \in k^\times$, $v(\alpha)/\log \|\alpha\|_v = -c_v$ is a negative constant. So for $P \in J(k)$ we have $\varepsilon(P) = c_v \tilde{\varepsilon}_v(P)$ and $\mu(P) = c_v \tilde{\mu}_v(P)$, where $\tilde{\varepsilon}_v$ and $\tilde{\mu}_v$ are as defined in the introduction.

We will also have occasion to use the following function. Let $x, y \in \mathrm{KS}_{\mathbb{A}}$ and define

$$\varepsilon(x, y) = v(B(x, y)) - 2v(x) - 2v(y). \tag{3-2}$$

In the same way as for $\varepsilon(x)$ above, we can extend this to points in $\mathrm{KS}(k)$ and $J(k)$.

Lemma 3.4. *Let $x, y, w, z \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates satisfying $w * z = B(x, y)$. Then we have*

$$\delta(w) * \delta(z) = B(\delta(x), \delta(y)).$$

Proof. The proof carries over verbatim from the proof of [Stoll 2002, Lemma 3.2]. □

We deduce the following:

Lemma 3.5. *Let $x, y, w, z \in \mathrm{KS}_{\mathbb{A}}$ be Kummer coordinates satisfying $w * z = B(x, y)$. Then we have*

$$\varepsilon(\delta(x), \delta(y)) + 2\varepsilon(x) + 2\varepsilon(y) = \varepsilon(w) + \varepsilon(z) + 4\varepsilon(x, y).$$

Proof. Using Lemma 3.4, relation (3-2), and property (i) above for $\delta(w)$, $\delta(z)$, $\delta(x)$ and $\delta(y)$, we obtain

$$v(\delta(w)) + v(\delta(z)) = v(B(\delta(x), \delta(y))) = \varepsilon(\delta(x), \delta(y)) + 2v(\delta(x)) + 2v(\delta(y)).$$

Subtracting four times the corresponding relation for w, z, x and y, we get

$$\varepsilon(w)+\varepsilon(z)=\varepsilon(\delta(x),\delta(y))-4\varepsilon(x,y)+2\varepsilon(x)+2\varepsilon(y),$$

which is the claim. □

We state a few general facts on the functions ε and μ.

Lemma 3.6. *For points $P, Q \in J(k)$, we have the relation*

$$\mu(P+Q)+\mu(P-Q)-2\mu(P)-2\mu(Q)=-\varepsilon(P,Q).$$

Proof. Let x and y be Kummer coordinates for P and Q, respectively; then w and z as in Lemma 3.5 are Kummer coordinates for $P+Q$ and $P-Q$ (in some order). The claim now follows from the formula in Lemma 3.5:

$$\begin{aligned}
&\mu(P+Q)+\mu(P-Q)-2\mu(P)-2\mu(Q)\\
&\quad=\sum_{n=0}^{\infty}4^{-n-1}\bigl(\varepsilon(2^nP+2^nQ)+\varepsilon(2^nP-2^nQ)-2\varepsilon(2^nP)-2\varepsilon(2^nQ)\bigr)\\
&\quad=\sum_{n=0}^{\infty}4^{-n-1}\bigl(\varepsilon(\delta^{\circ n}(w))+\varepsilon(\delta^{\circ n}(z))-2\varepsilon(\delta^{\circ n}(x))-2\varepsilon(\delta^{\circ n}(y))\bigr)\\
&\quad=\sum_{n=0}^{\infty}4^{-n-1}\bigl(\varepsilon(\delta^{\circ(n+1)}(x),\,\delta^{\circ(n+1)}(y))-4\varepsilon(\delta^{\circ n}(x),\,\delta^{\circ n}(y))\bigr)\\
&\quad=-\varepsilon(x,y)=-\varepsilon(P,Q).
\end{aligned}$$

□

Lemma 3.7. *If $P \in J(k)$ satisfies $\mu(P)=0$, then $\mu(P+Q)=\mu(Q)$ for all $Q \in J(k)$.*

Proof. We apply Lemma 3.6 with P and Q replaced by $Q+nP$ and P, respectively, where $n \in \mathbb{Z}$. Taking into account that $\mu(P)=0$ and writing a_n for $\mu(Q+nP)$, this gives

$$a_{n+1}-2a_n+a_{n-1}=-\varepsilon(P,Q+nP).$$

As k is a nonarchimedean local field, the multiples of P accumulate at the origin O in $J(k)$. Recall that ε is locally constant. This implies that every value $\varepsilon(P,Q+nP)$ occurs for infinitely many $n \in \mathbb{Z}$, since $Q+(n+N)P$ will be close to $Q+nP$ for suitably chosen N. We have for any $m>0$

$$a_{m+1}-a_m-a_{-m}+a_{-m-1}=\sum_{n=-m}^{m}(a_{n+1}-2a_n+a_{n-1})=-\sum_{n=-m}^{m}\varepsilon(P,Q+nP).$$

Since μ is bounded, the left-hand side is bounded independently of m. We also know that $\varepsilon(P,Q+nP)\geq 0$. But if $\varepsilon(P,Q+nP)$ were nonzero for some n, then by the discussion above, the right-hand side would be unbounded as $m\to\infty$.

Therefore it follows that $\varepsilon(P, Q+nP) = 0$ for all $n \in \mathbb{Z}$. This in turn implies $a_{n+1} - 2a_n + a_{n-1} = 0$ for all $n \in \mathbb{Z}$. The only bounded solutions of this recurrence are constant sequences. In particular, we have

$$\mu(P+Q) = a_1 = a_0 = \mu(Q). \qquad \square$$

Proposition 3.8. *The subset $U = \{P \in J(k) : \mu(P) = 0\}$ is a subgroup of finite index in $J(k)$. The functions ε and μ factor through the quotient $J(k)/U$.*

Proof. Lemma 3.7 shows that U is a subgroup. We have $\varepsilon(P) = 0$ for $P \in J(k)$ sufficiently close to the origin. So taking a sufficiently small subgroup neighborhood U' of the origin in $J(k)$, we see that $\varepsilon(2^n P) = 0$ for all $P \in U'$ and all $n \geq 0$. This implies that $\mu = 0$ on U', so $U \supset U'$. Because k is a local field, U' and therefore also U have finite index in $J(k)$. By Lemma 3.7 again, μ factors through $J(k)/U$, and since $\varepsilon(P) = 4\mu(P) - \mu(2P)$, the same is true for ε. $\square$

We will now show that we actually have

$$U = \{P \in J(k) : \varepsilon(P) = 0\}$$

(the inclusion "$\subset$" is clear from the definition and Proposition 3.8). This is equivalent to the implication $\varepsilon(x) = 0 \Rightarrow \varepsilon(\delta(x)) = 0$ and generalizes [Stoll 2002, Theorem 4.1]. For this we first provide a characteristic-2 analogue of Proposition 3.1(1) of the same paper.

We temporarily let k denote an arbitrary field. Let $C_{F,H}$ be a (not necessarily smooth) curve in the weighted projective plane with respective weights 1, 3, 1 assigned to the variables X, Y, Z that is given by an equation

$$Y^2 + H(X, Z)Y = F(X, Z), \tag{3-3}$$

where $F, H \in k[X, Z]$ are binary forms of respective degrees 6 and 3. Let $\mathrm{KS}_{F,H}$ denote the subscheme of $\mathbb{P}^3$ given by the vanishing of the equation defining the Kummer surface of $C_{F,H}$ if $C_{F,H}$ is nonsingular. Then the construction of $\delta = (\delta_1, \delta_2, \delta_3, \delta_4)$ still makes sense in this context, but we may now have $\delta_i(x) = 0$ for all $1 \leq i \leq 4$ (which we abbreviate by $\delta(x) = 0$) for a set x of Kummer coordinates on $\mathrm{KS}_{F,H}$. We generalize Proposition 3.1 in [Stoll 2002] (which assumes $H = 0$) to the case considered here.

Note that two equations (3-3) for $C_{F,H}$ are related by a transformation τ acting on an affine point (ξ, η) by

$$\tau(\xi, \eta) = \left(\frac{a\xi + b}{c\xi + d}, \frac{e\eta + U(\xi, 1)}{(c\xi + d)^3} \right), \tag{3-4}$$

type	H	F	conditions
1	0	0	
2	Z^3	0	
3	Z^3	aXZ^5	$a \neq 0$
4	XZ^2	aXZ^5	$a \neq 0$
5	XZ^2	bX^3Z^3	$b \neq 0$
6	Z^3	$aXZ^5 + bX^3Z^3$	$ab \neq 0$
7	XZ^2	0	
8	$XZ(X+Z)$	0	
9	$XZ(X+Z)$	bX^3Z^3	$b(b+1) \neq 0$
10	$XZ(X+Z)$	$aXZ^5 + bX^3Z^3$	$a(a+b)(a+b+1) \neq 0$
11	XZ^2	$aXZ^5 + bX^3Z^3$	$ab \neq 0$
12	0	XZ^5	
13	0	X^3Z^3	

Table 1. Representatives in characteristic 2.

where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(k)$, $e \in k^\times$ and $U \in k[X, Z]$ is homogeneous of degree 3. The transformation τ also acts on the forms F and H by

$$\tau^* F(X, Z) = (ad - bc)^{-6}(e^2 F^A + (eH^A - U^A)U^A),$$
$$\tau^* H(X, Z) = (ad - bc)^{-3}(eH^A - 2U^A),$$

where we write

$$S^A = S(dX - bZ, -cX + aZ)$$

for a binary form $S \in k[X, Z]$.

Lemma 3.9. *Let $x \in \mathrm{KS}_{F,H}(k)$. If $\delta(\delta(x)) = 0$, then we already have $\delta(x) = 0$.*

Proof. If k has characteristic different from 2, we can apply a transformation so that the new Weierstrass equation will have $H = 0$; the statement is then [Stoll 2002, Proposition 3.1(1)]. So from now on, k has characteristic 2. We may assume without loss of generality that k is algebraically closed. If the given curve is smooth, then the result is obvious, because the situation described in the statement can never occur. If it is not smooth, we can act on F and H using transformations of the form (3-4), so it is enough to consider only one representative of each orbit under such transformations. This is analogous to the strategy in the proof of [Stoll 2002, Proposition 3.1]. We can, for example, pick the representatives listed in Table 1.

For these representatives, elementary methods as in that proof can be used to check that $\delta(x) = 0$ indeed follows from $\delta(\delta(x)) = 0$. □

We can use the above to analyze the group U.

Theorem 3.10. *Suppose that k is a nonarchimedean local field and that J is the Jacobian of a smooth projective curve of genus 2, given by a Weierstrass equation (3-1) with integral coefficients. Then the set $\{P \in J(k) : \varepsilon(P) = 0\}$ equals the subgroup U in Proposition 3.8. In particular, U is a subgroup of finite index in $J(k)$ and ε and μ factor through the quotient $J(k)/U$. Moreover we have that $\varepsilon(-P) = \varepsilon(P)$ and U contains the kernel of reduction $J(k)^1$ with respect to the given model of J, i.e., the subgroup of points whose image in $\mathrm{KS}(\mathfrak{k})$ equals that of O.*

Proof. The statement in Lemma 3.9 implies $\varepsilon(P) = 0 \Rightarrow \varepsilon(2P) = 0$ for points $P \in J(k)$, since $\varepsilon(P) = 0$ is equivalent to $\delta(\tilde{x}) \neq 0$ if x are normalized Kummer coordinates for P, with reduction $\tilde{x}$. This shows that $\varepsilon(P) = 0$ implies $\mu(P) = 0$ (and conversely), so $\{P \in J(k) : \varepsilon(P) = 0\} = \{P \in J(k) : \mu(P) = 0\} = U$. The remaining statements now are immediate from Proposition 3.8, taking into account that, for P in the kernel of reduction, we trivially have $\varepsilon(P) = 0$. □

An algorithm for the computation of $\mu(P)$ which is based on Theorem 3.10 (for $H = 0$) is given in [Stoll 2002, §6]. Using the relation in Lemma 3.6, we obtain the following alternative procedure for computing $\mu(P)$.

1. Let x be normalized Kummer coordinates for P. Set $y_0 = (0, 0, 0, 1)$ and $y_1 = x$.
2. For $n = 1, 2, \ldots$, do the following.
   a. Using pseudoaddition (see [Flynn and Smart 1997, §4]), compute normalized Kummer coordinates y_{n+1} for nP from x, y_{n-1} and y_n; record $\varepsilon(P, nP)$, which is the shift in valuation occurring when normalizing y_{n+1}.
   b. If $\varepsilon(P, nP) = 0$, check whether $v(\delta(y_n)) = 0$ (by Theorem 3.10, this is equivalent to $nP \in U$). If yes, let $N = n$ and exit the loop.
3. Return
$$\mu(P) = \frac{1}{2N} \sum_{n=1}^{N-1} \varepsilon(P, nP).$$

To see that this works, note that by Lemma 3.6 we have

$$\mu((n+1)P) - 2\mu(nP) + \mu((n-1)P) = 2\mu(P) - \varepsilon(P, nP).$$

The sequence $(\mu(nP))_{n \in \mathbb{Z}}$ is periodic with period N, where N is the smallest positive integer n such that $nP \in U$ (which exists according to Theorem 3.10). Taking the sum over one period gives

$$2N\mu(P) = \sum_{n=0}^{N-1} \varepsilon(P, nP) = \sum_{n=1}^{N-1} \varepsilon(P, nP).$$

From the periodicity we can also deduce the possible denominators of $\mu(P)$. As ε has integral values, we see that $\mu(P) \in \frac{1}{2N}\mathbb{Z}$ if N is a period of $(\mu(nP))_{n\in\mathbb{Z}}$. In fact, we can show a little bit more.

Corollary 3.11. *Let $P \in J(k)$ and $N = \min\{n \in \mathbb{Z}_{>0} : \mu(nP) = 0\}$. Then*

$$\mu(P) \in \begin{cases} \frac{1}{N}\mathbb{Z} & \text{if } N \text{ is odd,} \\ \frac{1}{2N}\mathbb{Z} & \text{if } N \text{ is even.} \end{cases}$$

Proof. The sequence $(\varepsilon(P, nP))_{n\in\mathbb{Z}}$ has period N and is symmetric. So if N is odd, we actually have

$$\mu(P) = \frac{1}{2N}\sum_{n=1}^{N-1} \varepsilon(P, nP) = \frac{1}{N}\sum_{n=1}^{\frac{1}{2}(N-1)} \varepsilon(P, nP) \in \frac{1}{N}\mathbb{Z}. \qquad \square$$

Analyzing the possible denominators of $\mu(P)$ will play a key role in Section 12, where we discuss another algorithm for the computation of $\mu(P)$.

4. Canonical local heights on Kummer coordinates

We now define a notion of canonical local height for Kummer coordinates. We keep the notation of the previous section.

Definition 4.1. Let $x \in \mathrm{KS}_{\mathbb{A}}$ be a set of Kummer coordinates on KS. The *canonical local height of x* is given by

$$\hat{\lambda}(x) = -v(x) - \mu(x).$$

Remark 4.2. We can also define the canonical local height on an archimedean local field in an analogous way. Then, if K is a global field and x is a set of Kummer coordinates for a point $J(K)$, we have

$$\hat{h}(P) = \sum_{v\in M_K} \frac{1}{c_v}\hat{\lambda}_v(x),$$

where c_v is the constant introduced in Remark 3.3 for a nonarchimedean place v and $c_v = [K_v : \mathbb{R}]^{-1}$ if v is archimedean.

The canonical local height $\hat{\lambda}$ on Kummer coordinates has somewhat nicer properties than the canonical local height defined (for instance, in [Flynn and Smart 1997] or, more generally, in [Hindry and Silverman 2000, §B.9]) with respect to a divisor on J.

Proposition 4.3. *Let $x, y, z, w \in \mathrm{KS}_{\mathbb{A}}$. Then the following hold*:

(i) $\hat{\lambda}(\delta(x)) = 4\hat{\lambda}(x)$.

(ii) *If* $w * z = B(x, y)$, *then* $\hat{\lambda}(z) + \hat{\lambda}(w) = 2\hat{\lambda}(x) + 2\hat{\lambda}(y)$.

(iii) $\hat{\lambda}(x) = -\lim_{n\to\infty} 4^{-n} v(\delta^{\circ n}(x))$.

(iv) *If* k'/k *is a finite extension of ramification index* e *and* $\widehat{\lambda'}$ *is the canonical local height over* k', *then we have* $\widehat{\lambda'}(x) = e \cdot \hat{\lambda}(x)$.

Proof. (i) This follows easily from the two relations

$$v(\delta(x)) = 4v(x) + \varepsilon(x) \quad \text{and} \quad \mu(\delta(x)) = 4\mu(x) - \varepsilon(x).$$

(ii) This is similar, using Lemma 3.6 and $\varepsilon(x, y) = v(w) + v(z) - 2v(x) - 2v(y)$.

(iii) This follows from (i) and the fact that $\mu(x)$ is a bounded function, implying

$$\hat{\lambda}(x) = 4^{-n}\hat{\lambda}(\delta^{\circ n}(x)) = -4^{-n} v(\delta^{\circ n}(x)) + O(4^{-n}).$$

(iv) This is obvious from the definition of $\hat{\lambda}$. □

The canonical local height on Kummer coordinates also behaves well under isogenies.

Proposition 4.4. *Let* C *and* C' *be two curves of genus* 2 *over* k *given by Weierstrass equations, with associated Jacobians* J *and* J', *Kummer surfaces* KS *and* KS′ *and sets of sets of Kummer coordinates* $\mathrm{KS}_{\mathbb{A}}$ *and* $\mathrm{KS}'_{\mathbb{A}}$, *respectively. Let* $\alpha\colon J \to J'$ *be an isogeny defined over* k. *Then* α *induces a map* $\alpha\colon \mathrm{KS} \to \mathrm{KS}'$; *let* d *denote its degree. We also get a well-defined induced map* $\alpha\colon \mathrm{KS}_{\mathbb{A}} \to \mathrm{KS}'_{\mathbb{A}}$ *if we fix* $a \in k^\times$ *and require* $\alpha(0, 0, 0, 1) = (0, 0, 0, a)$. *Then we have*

$$\hat{\lambda}(\alpha(x)) = d\hat{\lambda}(x) - v(a) \quad \textit{for all } x \in \mathrm{KS}_{\mathbb{A}}.$$

Proof. All assertions except for the last one are obvious. By the definition of $\hat{\lambda}$, we can reduce to the case $a = 1$. Using part (iii) of Proposition 4.3 it is then enough to show that

$$v(\delta^{\circ n}(\alpha(x))) = d\, v(\delta^{\circ n}(x)) + O(1).$$

However, we have $v(\alpha(x)) - d\,v(x) = O(1)$ by assumption, so it suffices to show that

$$v(\delta^{\circ n}(\alpha(x))) = v(\alpha(\delta^{\circ n}(x))). \tag{4-1}$$

But since $\alpha\colon J \to J'$ is an isogeny, $\delta^{\circ n}(\alpha(x))$ and $\alpha(\delta^{\circ n}(x))$ represent the same point on KS′, hence they are projectively equal. Because they also have the same degree, the factor of proportionality is independent of x. It therefore suffices to check (4-1) for a single x; we take $x = (0, 0, 0, 1) \in \mathrm{KS}_{\mathbb{A}}$. Because we have $\delta(x) = x$ and, by assumption, $\alpha(x) = x'$, where $x' = (0, 0, 0, 1) \in \mathrm{KS}'_{\mathbb{A}}(k)$, we find

$$\delta^{\circ n}(\alpha(x)) = x' \quad \text{and} \quad \alpha(\delta^{\circ n}(x)) = x',$$

thereby proving (4-1) and hence the proposition. □

Remark 4.5. Canonical local heights with similar functorial properties were constructed by Zarhin [1995] on total spaces of line bundles (without the zero section). See also [Bombieri and Gubler 2006] for an approach to canonical local heights using rigidified metrized line bundles.

The preceding proposition is particularly useful for analyzing the behavior of the canonical local height under a change of Weierstrass equation of the curve.

Recall that two Weierstrass equations for C are related by a transformation τ as in (3-4), specified by a triple (A, e, U), where $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(k)$, $e \in k^\times$ and

$$U = u_0 Z^3 + u_1 X Z^2 + u_2 X^2 Z + u_3 X^3 \in k[X, Z]$$

is homogeneous of degree 3. Such a transformation induces a map on $\mathrm{KS}_{\mathbb{A}}$ as follows: Let $x = (x_1, x_2, x_3, x_4) \in \mathrm{KS}_{\mathbb{A}}$. Then $\tau(x)$ is given by the quadruple

$$(ad-bc)^{-1}\big(d^2x_1 + cdx_2 + c^2x_3,\ 2bdx_1 + (ad+bc)x_2 + 2acx_3,\\ b^2x_1 + abx_2 + a^2x_3,\ (ad-bc)^{-2}(e^2x_4 + l_1x_1 + l_2x_2 + l_3x_3)\big),$$

where l_1, l_2, l_3 do not depend on x. More precisely, we can write

$$l_i = l_{i,1} + l_{i,2} + l_{i,3},$$

where

$$l_{i,1} = \frac{e^2}{(ad-bc)^4}\, l'_{i,1} \quad \text{with } l'_{i,1} \in \mathbb{Z}[f_0, \ldots, f_6, a, b, c, d],$$
$$l_{i,2} = \frac{e}{(ad-bc)^4}\, l'_{i,2} \quad \text{with } l'_{i,2} \in \mathbb{Z}[h_0, \ldots, h_3, u_0, \ldots, u_3, a, b, c, d],$$
$$l_{i,3} = \frac{1}{(ad-bc)^4}\, l'_{i,3} \quad \text{with } l'_{i,3} \in \mathbb{Z}[u_0, \ldots, u_3, a, b, c, d]$$

for $i = 1, 2, 3$. All of the $l'_{i,j}$ are homogeneous of degree 8 in a, b, c, d and homogeneous in the other variables.

So we see that τ acts on k^4 as a linear map τ' whose determinant has valuation

$$v(\tau) := v(\det(\tau')) = 2v(e) - 3v(ad-bc).$$

In this situation, Proposition 4.4 implies:

Corollary 4.6. *Let $\tau = ([a, b, c, d], e, U)$ be a transformation (3-4) between two Weierstrass equations W and W' of a smooth projective curve C/k of genus 2 and let* KS *be the model of the Kummer surface associated to W. Then we have*

$$\hat{\lambda}(\tau(x)) = \hat{\lambda}(x) - v(\tau) \quad \textit{for all } x \in \mathrm{KS}_{\mathbb{A}}.$$

In particular,

$$\mu(x) = \mu(\tau(x)) + v(\tau(x)) - v(x) - v(\tau).$$

This can be used to construct a canonical local height which does not depend on the choice of Weierstrass equation.

Definition 4.7. Let C/k be a smooth projective curve of genus 2 given by a Weierstrass equation (3-1) with discriminant Δ and let KS be the associated Kummer surface. We call the function

$$\tilde{\lambda}\colon \mathrm{KS}_{\mathbb{A}} \to \mathbb{R}, \quad x \mapsto \hat{\lambda}(x) + \tfrac{1}{10} v(\Delta),$$

the *normalized canonical local height on* $\mathrm{KS}_{\mathbb{A}}$.

Corollary 4.8. *The normalized canonical local height is independent of the given Weierstrass equation of C, in the following sense: if W and W' are two Weierstrass equations for C, with associated sets of sets of Kummer coordinates $\mathrm{KS}_{\mathbb{A}}$ and $\mathrm{KS}'_{\mathbb{A}}$ and canonical local heights $\tilde{\lambda}$ and $\tilde{\lambda}'$, respectively, and τ is a transformation (3-4) between them, then for all $x \in \mathrm{KS}_{\mathbb{A}}$ we have $\tilde{\lambda}'(\tau(x)) = \tilde{\lambda}(x)$.*

Proof. Let Δ and Δ' be the respective discriminants of W and W'. By [Liu 1996, §2], we have

$$v(\Delta') = v(\Delta) + 10v(\tau), \tag{4-2}$$

so, using Corollary 4.6,

$$\begin{aligned}\tilde{\lambda}'(\tau(x)) &= \hat{\lambda}'(\tau(x)) + \tfrac{1}{10} v(\Delta') \\ &= \hat{\lambda}(x) - v(\tau) + \tfrac{1}{10} v(\Delta') \\ &= \hat{\lambda}(x) + \tfrac{1}{10} v(\Delta) = \tilde{\lambda}(x).\end{aligned}$$

□

We will not need the normalized canonical local height in the remainder of this paper.

5. Stably minimal Weierstrass models

In this section, k continues to denote a nonarchimedean local field with valuation ring $\mathcal{O}$ and residue field $\mathfrak{k}$. We build on results established by Liu [1996] in the more general context of hyperelliptic curves of arbitrary genus.

Recall that an equation of the form (3-1) defining a curve C over k of genus 2 is an *integral Weierstrass model* of C if the polynomials F and H have coefficients in $\mathcal{O}$. (Note that this is slightly different from the notion of an "integral equation" as defined in [Liu 1996, Définition 2], but the difference is irrelevant for our purposes, since any minimal Weierstrass model is actually given by an integral equation; see [Liu 1996, Remarque 4].) It is a *minimal Weierstrass model* of C if it is integral and the valuation of its discriminant is minimal among all integral Weierstrass models of C [Liu 1996, Définition 3]. We introduce the following variant of this notion.

Definition 5.1. An integral Weierstrass model of a smooth projective curve C over k

of genus 2 is *stably minimal* if it is a minimal Weierstrass model for C over k' for every finite field extension k' of k.

Stably minimal Weierstrass models can be characterized in terms of the multiplicities of the points on the special fiber.

Definition 5.2. Only for this definition let k be an arbitrary field, and let $C_{F,H}$ be a curve in $\mathbb{P}_k(1,3,1)$ given by an equation of the form (3-1) over k; we assume that $C_{F,H}$ is reduced. The *multiplicity* $m(P, C_{F,H})$ of a geometric point $P \in C_{F,H}(\bar{k})$ is defined as follows:

- If P is a singular point of type A_n (relative to the embedding of $C_{F,H}$ into $\mathbb{P}_k(1,3,1)$), then $m(P, C_{F,H}) = n+1$.
- If P is fixed by the involution $\iota(X:Y:Z) = (X:-Y-H(X,Z):Z)$ and is nonsingular, then $m(P, C_{F,H}) = 1$.
- Otherwise $m(P, C_{F,H}) = 0$.

Singularities of type A_n were defined by Arnold over the complex numbers, and hence for arbitrary fields of characteristic zero; see for instance [Barth et al. 1984, §II.8]. For the case of positive characteristic, see [Greuel and Kröning 1990]. Note that if the characteristic of k is not 2, then $\pi(P)$ is a root of multiplicity $m(P, C_{F,H})$ of F^2+4H, where $\pi : C_{F,H} \to \mathbb{P}^1$ sends $(X:Y:Z)$ to $(X:Z)$.

We will use this notion in the context of points on the special fiber of a Weierstrass model of a curve of genus 2 over a complete local field. In this context, Definition 5.2 is equivalent to [Liu 1996, Définition 9] when the curve is reduced; see [Liu 1996, Remarque 8].

An algorithm that computes the multiplicity was given by Liu [1996, §6.1]. Liu defines [1996, Définition 10] further multiplicities $\lambda_r(P)$ for points on the special fiber of an integral Weierstrass model (and $r \geq 1$) that allow us to characterize when such a model is minimal. We note here that $\lambda_r(P)$ gives the value of $\lambda(P) = \lambda_1(P)$ after making a field extension of ramification index r. Also, Lemme 7(e) of [Liu 1996] states for r sufficiently large that $\lambda_r(P) = m(P)$ if the special fiber is reduced and implies that $\lambda_r(P) \geq r$ if the special fiber is nonreduced. In the reduced case, we also have $\lambda(P) \leq m(P)$.

Setting $\lambda = \lambda_1$, Corollaire 2 in [Liu 1996] states (for $g = 2$) that the model is minimal if and only if $\lambda(P) \leq 3$ and $\lambda'(P) \leq 4$ (and is the unique minimal Weierstrass model up to $\mathcal{O}$-isomorphism, if and only if in addition $\lambda'(P) \leq 3$) for all $\mathfrak{k}$-points P on the special fiber, where $\lambda'(P)$ is a number satisfying $\lambda'(P) \leq 2\lceil \lambda(P)/2 \rceil$; see [Liu 1996, Lemme 9(c)].

Lemma 5.3. *An integral Weierstrass model of a smooth projective curve C over k of genus 2 is stably minimal if and only if its special fiber is reduced and the multiplicity of every geometric point on the special fiber is at most 3.*

If the special fiber is reduced and all multiplicities are at most 2, *then the model is the unique minimal Weierstrass model of* C *over any finite extension* k' *of* k, *up to isomorphism over the valuation ring of* k'.

Proof. First note that the multiplicity of a point is a geometric property; it does not change when we replace k by a finite extension. If the special fiber of an integral Weierstrass model has the given properties, then it follows from Liu's results mentioned above that $\lambda(P) \le m(P) \le 3$ and therefore $\lambda'(P) \le 4$ for all points P on the special fiber, even after replacing k by a finite extension. It follows that the model is stably minimal.

If $m(P) \le 2$ for all P, then $\lambda(P) \le 2$ and $\lambda'(P) \le 2$, so by Liu's results, the model is the unique minimal Weierstrass model of C over k'.

Conversely, assume that the special fiber does not have the given properties. Then either the special fiber is nonreduced, or else there is a point P on the special fiber of multiplicity $m(P) \ge 4$. If the special fiber is nonreduced, then after replacing k by a sufficiently ramified extension k', there is a point P on the special fiber such that $\lambda(P) > 3$ over k' (ramification index 4 is sufficient). If the special fiber is reduced and there is a (geometric) point P on the special fiber with $m(P) > 3$, then again after replacing k by a sufficiently large finite extension k' (such that P is defined over the residue field and the ramification index is at least $m(P)$), we have $\lambda(P) = m(P) > 3$ over k'. Liu's results then show that the model is not minimal over k'. □

Lemma 5.4. *If* C *is a smooth projective curve over* k *of genus* 2, *then there is a finite extension* k' *of* k *such that*

(i) *the minimal proper regular model of* C *over the valuation ring of* k' *has semistable reduction, and*

(ii) *each minimal Weierstrass model of* C *over* k' *is already stably minimal.*

Proof. That there is a finite extension with the first property is a special case of the semistable reduction theorem [Deligne and Mumford 1969]. After a further unramified extension, we can assume that all geometric components of the special fiber of the minimal proper regular model (which all have multiplicity 1) are defined over the residue field and that at least one component has a smooth point defined over the residue field. This implies by Hensel's lemma that $C(k') \ne \varnothing$. It then follows from [Liu 1996, Corollaire 5] that every minimal Weierstrass model of C over k' is dominated by the minimal proper regular model. Since the latter has reduced special fiber, the same is true for each minimal Weierstrass model.

Now assume that there exists a stably minimal Weierstrass model of C over k'. Then every minimal Weierstrass model of C over k' must already be stably minimal, since both models must have the same valuation of the discriminant, and the discriminant of the stably minimal model remains minimal over any finite field extension of k'. So it is enough to show that a stably minimal model exists.

We now consider the various possibilities for the special fiber of the minimal proper regular model. The possible configurations are shown in Figures 1, 2, 3 and 5 (on pages 2187, 2188, 2189 and 2196). If the reduction type is $[\mathrm{I}_{m_1-m_2-m_3}]$ in the notation of [Namikawa and Ueno 1973], then the Weierstrass model whose special fiber contains the component(s) that are not (-2)-curves has the property that all points on the special fiber have multiplicity at most 2; this is then the unique minimal Weierstrass model, and it is stably minimal by Lemma 5.3. It remains to consider reduction type $[\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$. We see that the Weierstrass models that correspond to components in the chain linking the two polygons and also those coming from the component of one of the polygons that is connected to the chain satisfy the conditions of Lemma 5.3 and are thus stably minimal. On the other hand, Weierstrass models whose special fiber does not correspond to a component in the chain or to one of its neighbors have a point in the special fiber whose multiplicity is at least 4 and so cannot be stably minimal. □

6. Igusa invariants

In this section we describe how we can easily distinguish between different types of reduction using certain invariants of genus-2 curves introduced by Igusa [1960]. The results of this section are essentially due to Liu [1993]; see also [Mestre 1991].

Let k be an arbitrary field of characteristic not equal to 2 and consider the invariants J_2, J_4, J_6, J_8, J_{10} defined in [Igusa 1960], commonly called *Igusa invariants*. Then $J_{2i}(F)$ is an invariant of degree $2i$ of binary sextics, and if

$$F(X, Z) = f_0Z^6 + f_1XZ^5 + f_2X^2Z^4 + f_3X^3Z^3 + f_4X^4Z^2 + f_5X^5Z + f_6X^6$$

is a binary sextic, then

$$J_{2i}(F) \in \mathbb{Z}\big[\tfrac{1}{2}, f_0, \ldots, f_6\big].$$

For example, $J_{10}(F) = 2^{-12}\operatorname{disc}(F)$. It is shown in [Igusa 1960] that the invariants J_2, J_4, J_6, J_{10} generate the even-degree part of the ring of invariants of binary sextics.

Now let F and H be the generic binary forms over $\mathbb{Z}$ of degrees 6 and 3, respectively, with coefficients $f_0, \ldots, f_6$ and $h_0, \ldots, h_3$ as before. It turns out that $J_{2i}(4F + H^2)$ is an element of $\mathbb{Z}[f_0, \ldots, f_6, h_0, \ldots, h_3]$.

Definition 6.1. Let k be an arbitrary field and let $H, F \in k[X, Z]$ be binary forms of respective degrees 3 and 6 over k. Let $C_{F,H}$ be the curve given by the equation $Y^2 + H(X, Z)Y = F(X, Z)$ in the weighted projective plane $\mathbb{P}_k(1, 3, 1)$. For $1 \le i \le 5$ we define the *Igusa invariant* $J_{2i}(C_{F,H})$ *of* $C_{F,H}$ as

$$J_{2i}(C_{F,H}) = J_{2i}(4F + H^2).$$

Following [Liu 1993], we also define two additional invariants, namely

$$I_4(C_{F,H}) = J_2(C_{F,H})^2 - 24J_4(C_{F,H})$$

and

$$I_{12}(C_{F,H}) = -8J_4(C_{F,H})^3 + 9J_2(C_{F,H})J_4(C_{F,H})J_6(C_{F,H}) - 27J_6(C_{F,H})^2 - J_2(C_{F,H})^2 J_8(C_{F,H}).$$

The following is a consequence of [Liu 1993, Théorème 1].

Proposition 6.2. *Let k be a field and let $C_{F,H}/k$ be the curve given by the equation*

$$Y^2 + H(X, Z)Y = F(X, Z)$$

in $\mathbb{P}_k(1, 3, 1)$, where $H, F \in k[X, Z]$ are binary forms of degree 3 and 6, respectively. For $1 \le i \le 5$ and $j \in \{4, 12\}$ we set $J_{2i} = J_{2i}(C_{F,H})$ and $I_j = I_j(C_{F,H})$.

(i) *$C_{F,H}$ is smooth $\iff J_{10} \neq 0$.*

(ii) *$C_{F,H}$ has a unique node and no point of higher multiplicity $\iff J_{10} = 0$ and $I_{12} \neq 0$.*

(iii) *$C_{F,H}$ has exactly two nodes $\iff J_{10} = I_{12} = 0$, $I_4 \neq 0$, and $J_4 \neq 0$ or $J_6 \neq 0$.*

(iv) *$C_{F,H}$ has three nodes $\iff J_{10} = I_{12} = J_4 = J_6 = 0$ and $I_4 \neq 0$.*

(v) *$C_{F,H}$ has a cusp $\iff J_{10} = I_{12} = I_4 = 0$ and $J_{2i} \neq 0$ for some $i \le 4$.*

(vi) *$C_{F,H}$ is nonreduced or has a point of multiplicity at least 4 $\iff J_{2i} = 0$ for all i.*

When C is a curve of genus 2 over a nonarchimedean local field, then Igusa invariants can also be used to obtain information on the reduction type of C; see [Liu 1993, Théorème 1, Proposition 2].

Proposition 6.3. *Let k be a nonarchimedean local field with normalized additive valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ and valuation ring $\mathcal{O}$, and let C/k be a smooth projective genus-2 curve, given by a minimal Weierstrass model with reduced special fiber. Suppose that the minimal proper regular model $\mathcal{C}^{\min}$ of C over $\operatorname{Spec}\mathcal{O}$ is semistable and has reduction type $\mathcal{K}$ in the notation of [Namikawa and Ueno 1973]. We set $J_{2i} = J_{2i}(C)$ for $i \in \{1, \dots, 5\}$ and $I_4 = I_4(C)$, $I_{12} = I_{12}(C)$.*

(i) *If $\mathcal{K} = [\mathrm{I}_{m-0-0}]$, where $m > 0$, then $m = v(J_{10})$.*

(ii) *If $\mathcal{K} = [\mathrm{I}_{m_1-m_2-0}]$, where $0 < m_1 \le m_2$, then*

$$m_1 = \min\left\{v(I_{12}), \tfrac{1}{2}v(J_{10})\right\} \quad \text{and} \quad m_2 = v(J_{10}) - m_1.$$

(iii) *If* $\mathcal{K} = [\mathrm{I}_{m_1-m_2-m_3}]$, *where* $0 < m_1 \le m_2 \le m_3$, *then*

$$m_1 = \min\{v(J_4), \tfrac{1}{3}v(J_{10}), \tfrac{1}{2}v(I_{12})\},$$
$$m_2 = \min\{v(I_{12}) - m_1, \tfrac{1}{2}(v(J_{10}) - m_1)\}, \text{ and}$$
$$m_3 = v(J_{10}) - m_1 - m_2.$$

(iv) *If* $\mathcal{K} = [\mathrm{I}_0 - \mathrm{I}_0 - l]$, *then* $l = \frac{1}{12}v(J_{10})$.

(v) *If* $\mathcal{K} = [\mathrm{I}_{m_1} - \mathrm{I}_0 - l]$, *where* $m_1 > 0$, *then*

$$l = \tfrac{1}{12}v(I_{12}) \quad \text{and} \quad m_1 = v(J_{10}) - v(I_{12}).$$

(vi) *If* $\mathcal{K} = [\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$, *where* $m_2 \ge m_1 > 0$ *and* $l > 0$, *then*

$$l = \tfrac{1}{4}v(I_4),$$
$$m_1 = \min\{v(I_{12}) - 3v(I_4), \tfrac{1}{2}(v(J_{10}) - 3v(I_4))\}, \text{ and}$$
$$m_2 = v(J_{10}) - 3v(I_4) - m_1.$$

Part II. Study of local height correction functions

In Part II of the paper, k will always denote a nonarchimedean local field with residue field $\mathfrak{k}$, valuation ring $\mathcal{O}$ and normalized additive valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$. We let C be a curve of genus 2 over k, given by an integral Weierstrass model $\mathcal{C}$, which we consider as a subscheme of the weighted projective plane $\mathbb{P}_S(1, 3, 1)$, where $S = \mathrm{Spec}(\mathcal{O})$. In the following five sections we find explicit formulas and bounds for the local height correction function μ for the most frequent cases of bad reduction and use these to deduce a general bound on μ. We denote the minimal proper regular model of C over S by $\mathcal{C}^{\min}$. Let J be the Jacobian of C; we denote its Néron model over S by $\mathcal{J}$. We write $\mathcal{C}_v$, $\mathcal{C}_v^{\min}$ and $\mathcal{J}_v$ for the respective special fibers of $\mathcal{C}$, $\mathcal{C}^{\min}$ and $\mathcal{J}$.

7. The "kernel" of μ

By Theorem 3.10, the set

$$U = \{P \in J(k) : \varepsilon(P) = 0\}$$

is a group and the local height correction function μ factors through the quotient $J(k)/U$. In this section we relate U to the Néron model of J when $\mathcal{C}$ has rational singularities. See [Artin 1986] for a brief account of the theory of rational singularities on arithmetic surfaces.

For the remainder of this section we assume that $\mathcal{C}/S$ is normal and reduced. We let $\mathcal{J}^0$ denote the fiberwise-connected component of the identity of $\mathcal{J}$. Then $\mathcal{J}^0$ has generic fiber $\mathcal{J}_k \cong J$ and special fiber $\mathcal{J}_v^0$, the connected component of the

identity of $\mathcal{J}_v$. If $\mathcal{C}' \to \mathcal{C}$ is a desingularization of $\mathcal{C}$, then the identity components $\mathrm{Pic}^0_{\mathcal{C}'/S}$ and $\mathrm{Pic}^0_{\mathcal{C}/S}$ of the respective relative Picard functors of $\mathcal{C}'$ and $\mathcal{C}$ can both be represented by separated schemes; see [Bosch et al. 1990, Theorem 9.7.1]. There are canonical S-group scheme morphisms

$$\mathrm{Pic}^0_{\mathcal{C}/S} \to \mathrm{Pic}^0_{\mathcal{C}'/S} \xrightarrow{\sim} \mathcal{J}^0; \tag{7-1}$$

the latter map is an isomorphism by [Bosch et al. 1990, Theorem 9.4.2]. Let $\alpha\colon \mathrm{Pic}^0_{\mathcal{C}/S} \to \mathcal{J}^0$ denote the composition of the morphisms from (7-1); note that α does not depend on the choice of the desingularization $\mathcal{C}'$. We will show that if $P \in J(k)$ has reduction on $\mathcal{J}$ in the image of α, then $\varepsilon(P) = \mu(P) = 0$. The idea is to first show that this is true for points in the image of a certain open subscheme; we then prove that this suffices for the general case.

Let $\mathcal{C}_{\mathrm{sm}}$ be the smooth locus of $\mathcal{C}$. Following [Bosch et al. 1990, §9.3], we define an S-subscheme W of the symmetric square $\mathcal{C}^{(2)}_{\mathrm{sm}}$ of $\mathcal{C}_{\mathrm{sm}}$ consisting of the points $w \in \mathcal{C}^{(2)}_{\mathrm{sm}}$ that satisfy the following conditions:

- $H^1(\mathcal{C}, \mathcal{O}_{\mathcal{C}}(D_w)) = 0$, where D is the universal Cartier divisor $D \subset \mathcal{C} \times_S \mathcal{C}^{(2)}_{\mathrm{sm}}$ induced by the canonical map $\mathcal{C}^{(2)}_{\mathrm{sm}} \to \mathrm{Div}^2_{\mathcal{C}/S}$.
- If $w = \{w_1, w_2\}$ with w_1, w_2 geometric points on the special fiber of $\mathcal{C}$, then the hyperelliptic involution ι maps the component containing w_1 to the component containing w_2.

Then W has the following properties:

(i) W is an open subscheme of $\mathcal{C}^{(2)}_{\mathrm{sm}}$.

(ii) There is a strict S-birational group law on W, induced by the group law on $\mathrm{Pic}_{\mathcal{C}/S}$.

(iii) $\mathrm{Pic}^0_{\mathcal{C}/S}$ is the S-group scheme associated with this strict S-birational group law.

For (ii) and (iii) see the discussion preceding [Bosch et al. 1990, Theorem 9.3.7].

Let $\mathrm{Pic}^{[2]}_{\mathcal{C}/S}$ be the open subfunctor of $\mathrm{Pic}_{\mathcal{C}/S}$ whose elements have total degree 2. Let $\rho\colon W \to \mathrm{Pic}^{[2]}_{\mathcal{C}/S}$ be the canonical map induced by D; by [Bosch et al. 1990, Lemma 9.3.5] it is an open immersion. Replacing S by the spectrum of the valuation ring of a finite unramified extension of k, if necessary, we can find a section $x_0 \in \mathbb{P}^1_S(S)$ such that its pullback D_0 under the covering map $\mathcal{C} \to \mathbb{P}^1_S$ is horizontal and does not intersect the singular locus of $\mathcal{C}$. We denote by c_0 the class of D_0 in $\mathrm{Pic}^{[2]}_{\mathcal{C}/S}$. Let $w = \{P_1, P_2\} \in W$; using the condition on the action of ι on the components P_1 and P_2 lie on, we find that

$$\rho_0(w) := \rho(w) - c_0 \in \mathrm{Pic}^0_{\mathcal{C}/S}\,.$$

In fact, ρ_0 defines an open immersion $\rho_0\colon W \to \mathrm{Pic}^0_{\mathcal{C}/S}$; see [Bosch et al. 1990, Lemma 9.3.6].

Lemma 7.1. *Suppose that the residue characteristic of k is not 2. Let $P \in J(k)$ such that the reduction of P on $\mathcal{J}_v$ is in $\alpha(\rho_0(W))$. Then $\varepsilon(P) = 0$.*

Proof. We may assume that $\mathcal{C}\colon Y^2 = F(X, Z)$. Let J_F denote the model of J in $\mathbb{P}^{15}$ constructed in [Cassels and Flynn 1996, Chapter 2] and let $\mathcal{J}_F/S$ denote the model it defines over S. Following [Bruin and Stoll 2010, §5], we denote by $\mathcal{J}_F^0$ the fiberwise-connected component of the identity of the smooth locus of $\mathcal{J}_F$, so that the generic fiber is $\mathcal{J}_F$ and the special fiber $\mathcal{J}_{F,v}^0$ is the connected component of the identity of the smooth locus of the special fiber $\mathcal{J}_{F,v}$. We have a morphism $\psi : \mathcal{C}_{\mathrm{sm}}^{(2)} \to \mathcal{J}_F^0$, defined using the expressions for the coordinates on J_F in [Cassels and Flynn 1996, Chapter 2]; see the proof of [Bruin and Stoll 2010, Lemma 5.7]. We also denote the restriction of this morphism to W by ψ.

The Néron mapping property yields a natural map $\varphi\colon \mathcal{J}_F^0 \to \mathcal{J}$. In general, its image can be a proper subset of $\mathcal{J}^0$. Nevertheless, the following diagram of S-scheme morphisms is commutative by [Liu 2002, Proposition 3.3.11], since W is reduced, $\mathcal{J}^0$ is separated and the diagram is commutative when restricted to generic fibers:

$$\begin{array}{ccc} W & \xrightarrow{\psi} & \mathcal{J}_F^0 \\ {\scriptstyle\rho_0}\downarrow & & \downarrow{\scriptstyle\varphi} \\ \mathrm{Pic}^0_{\mathcal{C}/S} & \xrightarrow{\alpha} & \mathcal{J}^0 \end{array} \tag{7-2}$$

It follows from [Bruin and Stoll 2010, Proposition 5.10] that a point $P \in J(k)$ satisfies $\varepsilon(P) = 0$ if and only if P reduces to $\mathcal{J}_{F,v}^0(\mathfrak{k})$. So if P has reduction in $\alpha(\rho_0(W))$, then the commutativity of the diagram (7-2) shows that $\varepsilon(P) = 0$. □

If the residue characteristic is 2, then no explicit analogue of the group scheme $\mathcal{J}_F$ is known. Instead, we have to work with explicit expressions to prove a result analogous to Lemma 7.1.

Let $\tilde{F}$ and $\tilde{H}$ be the reductions of F and H, respectively. In analogy with [Bruin and Stoll 2010, Definition 5.1], we define the subscheme $\tilde{\mathcal{D}}$ of $\mathbb{A}^3_{\mathfrak{k}} \times \mathbb{A}^4_{\mathfrak{k}} \times \mathbb{A}^5_{\mathfrak{k}}$ consisting of all triples

$$(A, B, C) = \big((a_0, a_1, a_2), (b_0, b_1, b_2, b_3), (c_0, c_1, c_2, c_3, c_4)\big) \in \mathbb{A}^3_{\mathfrak{k}} \times \mathbb{A}^4_{\mathfrak{k}} \times \mathbb{A}^5_{\mathfrak{k}}$$

such that

$$AC = \tilde{F} - B^2 - B\tilde{H},$$

where

$$\begin{aligned} A &= a_0Z^2 + a_1XZ + a_2X^2, \\ B &= b_0Z^3 + b_1XZ^2 + b_2X^2Z + b_3X^3, \\ C &= c_0Z^4 + c_1XZ^2 + c_2X^2Z^2 + c_3X^3Z + c_4X^4. \end{aligned}$$

Moreover, we set $\mathcal{D} := (\pi_2 \times \mathrm{id})(\mathrm{pr}_{12}(\tilde{\mathcal{D}}))$, where pr_{12} is the projection onto the first two factors and π_2 is the canonical map $\mathbb{A}^3_{\mathfrak{k}} \setminus \{(0,0,0)\} \to \mathbb{P}^2_{\mathfrak{k}}$.

Note that if the curve $\mathcal{C}_v$ defined by $Y^2 + \tilde{H}(X,Z)Y = \tilde{F}(X,Z)$ in $\mathbb{P}_{\mathfrak{k}}(1,3,1)$ is nonsingular, then $\mathcal{D}(\mathfrak{k})$ is in bijective correspondence with the possible Mumford representations of effective divisors of degree 2 on $\mathcal{C}_v$.

In general, this correspondence still holds for the subset $\mathcal{D}'$ of all $(A,B) \in \mathcal{D}$ such that A does not vanish at the image in $\mathbb{P}^1$ of a singular point of $\mathcal{C}_v$, and those effective divisors with support in the smooth locus of $\mathcal{C}_v$. More precisely, we get a map $\zeta : \mathcal{D}' \to C_v^{(2)}$ such that if $\zeta((A,B)) = \{\tilde{P}_1, \tilde{P}_2\}$, then there are representatives (X_i, Y_i, Z_i) of $\tilde{P}_i$ $(i = 1,2)$ satisfying

(i) $A(X,Z) = (Z_1X - X_1Z)(Z_2X - X_2Z)$,

(ii) $Y_i = B(X_i, Z_i)$ for $i = 1,2$.

If $\mathcal{C}_v$ is nonsingular, and $(A,B) \in \mathcal{D}$, then we can compose the natural surjection $\mathcal{D} \to \mathrm{Jac}(\mathcal{C}_v) \setminus \{O\}$ with the quotient map $\mathrm{Jac}(\mathcal{C}_v) \to \mathrm{KS}_{\tilde{F},\tilde{H}}$. In the general case one can also define a surjection $\omega\colon \mathcal{D} \to \mathrm{KS}_{\tilde{F},\tilde{H}} \setminus \{(0:0:0:1)\}$ with the following property: if $P = [(P_1) - (P_2)] \in J(k)$ is such that the reductions $\tilde{P}_1$ and $\tilde{P}_2$ are both smooth points on $\mathcal{C}_v$, and if $(A,B) \in \mathcal{D}'$ is such that $\zeta((A,B)) = \{\tilde{P}_1, \iota(\tilde{P}_2)\}$, then the reduction of $\kappa(P)$ on $\mathrm{KS}_{\tilde{F},\tilde{H}}$ is $\omega((A,B))$. The image of a pair $(A,B) \in \mathcal{D}$ under ω is of the form $(a_0 : -a_1 : a_2 : x_4)$.

Lemma 7.2. *Suppose that the residue characteristic of k is 2. Let $P \in J(k)$ such that the reduction of P on $\mathcal{J}$ is in $\alpha(\rho_0(W))$. Then $\varepsilon(P) = 0$.*

Proof. Let $(A,B) \in D'_{\tilde{F},\tilde{H}}$ such that $\zeta((A,B)) = \{\tilde{P}_1, \tilde{P}_2\} \in W$. By the discussion preceding the lemma, it suffices to show that we have $\delta(x) \neq 0$ for $x = \omega((A,B))$.

Changing the given model, if necessary, we can assume that $\tilde{H}$ and $\tilde{F}$ are as in the list of representatives 1–13 in Table 1. Table 2 contains conditions on x which are equivalent to the vanishing of $\delta(x)$ for each representative and additional conditions which a point $x = (x_1 : x_2 : x_3 : x_4) \in \mathbb{P}^3$ satisfying $\delta(x) = 0$ must satisfy in order to lie on $\mathrm{KS}_{\tilde{F},\tilde{H}}$. Finally, we have listed the multiplicities $m(\infty)$, $m(0)$, $m(1)$ that $\mathcal{C}_v$ has at the points with $(X:Z) = (1:0)$, $(X:Z) = (0:1)$ and $(X:Z) = (1:1)$, respectively, in case the multiplicities there are greater than 1. Note that we do not have to treat type 1, as $\mathcal{C}_v$ is assumed to be reduced.

Since $A(X,Z)$ does not vanish at the image in $\mathbb{P}^1$ of a singular point, we get $x_1 \neq 0$ and, if $(0,0)$ is a singular point, also $x_3 \neq 0$. Using Table 2, this already implies that $\delta(x) \neq 0$ whenever $\mathcal{C}_v$ is irreducible. In the reducible cases 2, 7 and 8, $\mathcal{C}_v$ has two irreducible components, and one checks easily that x_4 does not vanish because, by definition of W, ι maps the component containing $\tilde{P}_1$ to the component containing $\tilde{P}_2$. Hence $\delta(x) \neq 0$ by Table 2. □

The next proposition follows from Lemmas 7.1 and 7.2.

type	condition	additional	$m(\infty)$	$m(0)$	$m(1)$
1	$x_4 = 0$				
2	$x_4 = 0$		6		
3	$x_4 = 0$	$x_1 = 0$	5		
4	$x_4 = 0$	$x_1 = 0$	4		
5	$x_1x_3 = x_4 = 0$		3	2	
6	$x_1 = x_4 = 0$		3		
7	$x_4 = 0$		4	2	
8	$x_4 = 0$		2	2	2
9	$x_1x_3 = x_4 = 0$		2	2	
10	$x_1 = x_4 = 0$		2		
11	$x_1 = x_4 = 0$		3		
12	$x_4 = 0$	$x_1 = 0$	5		
13	$x_4 = 0$	$x_1x_3 = 0$	3	3	

Table 2. Conditions for the vanishing of $\delta(x)$.

Proposition 7.3. *Let $\alpha\colon \mathrm{Pic}^0_{\mathcal{C}/S} \to \mathcal{J}^0$ be the canonical homomorphism. If the reduction of $P \in J(k)$ on $\mathcal{J}_v$ is in the image of α, then $\varepsilon(P) = \mu(P) = 0$.*

Proof. If T is an S-scheme and $x \in \mathrm{Pic}^0_{\mathcal{C}/S}(T)$, then by properties (ii) and (iii) of W, there is an étale cover T'/T and $w_1, \dots, w_n \in W(T')$ such that

$$x = \rho_0(w_1) + \cdots + \rho_0(w_n),$$

where the sum is taken with respect to the group law on $\mathrm{Pic}^0_{\mathcal{C}/S}$. In fact we can take $n = 2$; this follows from [Bosch et al. 1990, Lemma 5.1.4] and the discussion following Lemma 5.2.4 of the same paper. Using this and Theorem 3.10, it suffices to show that $\varepsilon(P) = 0$ when the reduction of P on $\mathcal{J}_v$ is in $\alpha(\rho_0(W))$. Hence the result follows from Lemmas 7.1 and 7.2. □

Let $J_0(k)$ denote the subgroup of $J(k)$ consisting of points whose image on the special fiber of $\mathcal{J}$ is in $\mathcal{J}^0(\mathfrak{k})$. By [Bosch and Liu 1999, Lemma 2.1] the group $\Phi(\mathfrak{k})$ of $\mathfrak{k}$-rational points in the component group Φ of J satisfies

$$\Phi(\mathfrak{k}) \cong J(k)/J_0(k).$$

We can now give a criterion for when ε and μ factor through $\Phi(\mathfrak{k})$.

Theorem 7.4. *Let C be a smooth projective curve of genus 2 defined over a nonarchimedean local field k, given by an integral Weierstrass model $\mathcal{C}$ with rational singularities. Then ε and μ factor through $\Phi(\mathfrak{k})$.*

Proof. First note that if $\mathcal{C}$ has rational singularities, then $\mathcal{C}$ is normal and reduced. Moreover, according to [Bosch et al. 1990, Theorem 9.7.1], the homomorphism α is an isomorphism if and only if $\mathcal{C}$ has rational singularities. This implies that the

image of α, restricted to the generic fiber, is $J_0(k)$. By Proposition 7.3, we have $\varepsilon(P) = \mu(P) = 0$ for P in the image of α. Theorem 3.10 implies that μ and ε factor through $\Phi(\mathfrak{k})$. □

Remark 7.5. A nonminimal Weierstrass model cannot have rational singularities. Moreover, there are minimal (even stably minimal) Weierstrass models of curves of genus 2 that have nonrational singularities. See Example 10.4 for a stably minimal Weierstrass model having $\mu(P) \neq 0$ for some points $P \in J_0(k)$.

This behavior cannot occur for elliptic curves; here μ always factors through $\Phi(\mathfrak{k})$, provided the given Weierstrass model is minimal; see [Silverman 1988]. This is crucial for the usual algorithms to compute canonical heights on elliptic curves. Note that a Weierstrass model of an elliptic curve is minimal if and only if it has rational singularities by [Conrad 2005, Corollary 8.4].

8. Néron functions and reduction graphs

Our next goal is to derive a formula for $\mu(P)$ in the case when the minimal proper regular model of C is semistable and μ factors through $\Phi(\mathfrak{k})$. To this end, we need the notion of Néron functions. The following result is due to Néron; see [Lang 1983, §11.1].

Proposition 8.1. *Let A be an abelian variety defined over a local field k. Then we can associate to any divisor $D \in \mathrm{Div}_A(\bar{k})$ a function $\lambda_D \colon A(\bar{k}) \setminus \mathrm{supp}(D) \to \mathbb{R}$ such that the following conditions are satisfied, where we write $\lambda \equiv \lambda'$ mod const. to indicate that the functions λ and λ' differ by a constant.*

(1) *If $D, E \in \mathrm{Div}_A(\bar{k})$, then $\lambda_{D+E} \equiv \lambda_D + \lambda_E$ mod const.*

(2) *If $D = \mathrm{div}(f) \in \mathrm{Div}_A(\bar{k})$ is principal, then $\lambda_D \equiv \bar{v} \circ f$ mod const., where $\bar{v}$ is the extension of v to $\bar{k}$.*

(3) *If $D \in \mathrm{Div}_A(\bar{k})$ and $T_P \colon A \to A$ is the translation map by a point $P \in A(\bar{k})$, then we have $\lambda_{T_P^* D} \equiv \lambda_D \circ T_P$ mod const.*

Also, λ_D is uniquely determined up to adding a constant.

We call a function λ_D as in Proposition 8.1 a *Néron function associated with D*.

We can use local heights on Kummer coordinates to construct Néron functions on the Jacobian J of our genus-2 curve C. If $P_0 \in C(\bar{k})$, then we have an embedding $C_{\bar{k}} \to J_{\bar{k}}$ (defined over $\bar{k}$) that maps $P \in C(\bar{k})$ to the divisor class $[(P) - (P_0)] \in \mathrm{Pic}^0_C(\bar{k}) = J(\bar{k})$. Its image is the theta divisor Θ_{P_0}. We set $\Theta^{\pm}_{P_0} = \Theta_{P_0} + \Theta_{\iota(P_0)}$; then $\Theta^{\pm}_{P_0}$ is symmetric and in the linear equivalence class of 2Θ (where Θ is a theta divisor coming from taking a Weierstrass point as base-point). For the following, fix a point $\infty \in C(\bar{k})$ at infinity. For $i \in \{1, \dots, 4\}$, we set

$$D_i = \Theta^{\pm}_{\infty} + \mathrm{div}\left(\frac{\kappa_i}{\kappa_1}\right)$$

and we define a function $\hat{\lambda}_i : J(k) \setminus \mathrm{supp}(D_i) \to \mathbb{R}$ by

$$\hat{\lambda}_i(P) = \hat{\lambda}\left(\frac{\kappa(P)}{\kappa_i(P)}\right).$$

Lemma 8.2. *Let $\infty \in C(\bar{k})$ be a point at infinity as above and let $i \in \{1, \dots, 4\}$. Then D_i is defined over k and the function $\hat{\lambda}_i$ is a Néron function associated with D_i.*

Proof. If $\infty \notin C(k)$, then we have $\infty \in C(k')$ for some quadratic extension k' of k and the nontrivial element of the Galois group $\mathrm{Gal}(k'/k)$ maps ∞ to $\iota(\infty)$, proving the first assertion. For a proof of the second assertion, see [Uchida 2011, Theorem 5.3]. □

Definition 8.3. Assume that C has semistable reduction over k. Let $C' = \mathcal{C}^{\min}_{v,\bar{\mathfrak{k}}}$ denote the special fiber of the minimal proper regular model $\mathcal{C}^{\min}$ of C, considered over the algebraic closure of the residue field $\mathfrak{k}$. The *reduction graph* $R(C)$ of C is a graph with vertex set the set of irreducible components of C'; two vertices Γ_1 and Γ_2 are connected by n edges, where n is the number of intersection points of Γ_1 and Γ_2 if $\Gamma_1 \neq \Gamma_2$, and n is the number of nodes of Γ_1 if $\Gamma_1 = \Gamma_2$. The Galois group of $\mathfrak{k}$ acts on $R(C)$ in a natural way.

We consider $R(C)$ as a metric graph by giving each edge length 1. For two vertices Γ_1 and Γ_2, we define $r(\Gamma_1, \Gamma_2)$ as the resistance between the vertices, when $R(C)$ is considered as an electric network with unit resistance along every edge.

Remark 8.4. We can compute $r(\Gamma_1, \Gamma_2)$ as follows. Order the vertices of $R(C)$ in some way and let M be the intersection matrix with respect to this ordering. Since all components of the special fiber have multiplicity one, the kernel of M is spanned by the "all-ones" vector and the image of M consists of the vectors whose entries sum to zero. Let v be the vector with entries zero except that the entry corresponding to Γ_1 is 1 and the entry corresponding to Γ_2 is -1. Then there is a vector g with rational entries such that $Mg = v$, and

$$r(\Gamma_1, \Gamma_2) = -g \cdot v$$

is, up to sign, the standard inner product of the two vectors. (Note that g is not unique, but adding a vector in the kernel of M to it will not change the result.) See for instance [Cinkir 2011, Lemma 6.1].

Note that the linear map given by M on the space of functions on the vertices can be interpreted as the discrete Laplace operator on the graph $R(C)$. It is then easy to see that g, viewed as a function on the vertices, is piecewise linear along sequences of edges not containing Γ_1, Γ_2 or a vertex of degree at least 3. This makes it quite easy to find g and to compute $r(\Gamma_1, \Gamma_2)$.

The reduction graph is unchanged when we replace k by an unramified extension. If we base-change to a ramified extension k' of k with ramification index e, then

the new reduction graph is obtained by subdividing the edges of $R(C)$ into e new edges. We can give these new edges length $1/e$; then the underlying metric space remains the same. In particular, $r(\Gamma_1, \Gamma_2)$ does not depend on k'. This allows us to replace k by a finite extension if necessary. The scaling of the length corresponds to extending the valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ to $\bar{k}^\times \to \mathbb{Q}$ instead of considering the normalized valuation on k'. All notions defined in terms of the valuation (for example, intersection numbers) are then scaled accordingly.

Proposition 8.5. *We assume that $\mathcal{C}^{\min}$ is semistable. Let $P = [(P_1) - (P_2)] \in J(k)$, with $P_1, P_2 \in C(k)$ mapping to components Γ_1 and Γ_2, respectively, of the special fiber of $\mathcal{C}^{\min}$. We make the following further assumptions.*

(i) *If $Q_1, Q_2 \in C(k)$ map to Γ_1 and Γ_2, respectively, then $\mu(P) = \mu([(Q_1)-(Q_2)])$.*

(ii) *There is a constant $\mu_1 \in \mathbb{Q}$ such that $\mu([(Q_1) - (Q_1')]) = \mu_1$ for all $Q_1, Q_1' \in C(k)$ mapping to Γ_1 such that the images of Q_1 and Q_1' on the special fiber of $\mathcal{C}^{\min}$ are distinct.*

(iii) *There is a constant $\mu_2 \in \mathbb{Q}$ such that $\mu([(Q_2) - (Q_2')]) = \mu_2$ for all $Q_2, Q_2' \in C(k)$ mapping to Γ_2 such that the images of Q_2 and Q_2' on the special fiber of $\mathcal{C}^{\min}$ are distinct.*

Then we have

$$\mu(P) = r(\Gamma_1, \Gamma_2) + \tfrac{1}{2}(\mu_1 + \mu_2).$$

Proof. By the discussion preceding the statement of the theorem, we can assume that k is sufficiently large for $C(k)$ to contain all points we might be interested in.

Let $P_0 \in C(k)$. The embedding with respect to P_0 is obtained from the "difference map" $\psi\colon C \times C \to J$ that sends a pair of points (P_1, P_2) to $[(P_1) - (P_2)]$ by specializing the second argument to P_0. One easily checks that

$$\psi^*\Theta_{P_0} = \Delta_C + (\{\iota(P_0)\} \times C) + (C \times \{P_0\}),$$

where Δ_C denotes the diagonal and ι is the hyperelliptic involution on C. We then have

$$\psi^*\Theta^{\pm}_{P_0} = 2\Delta_C + \mathrm{pr}_1^* D_0 + \mathrm{pr}_2^* D_0,$$

where $D_0 = (P_0) + (\iota(P_0))$. By the results in [Heinz 2004] this implies that, taking λ_0 to be a Néron function associated to $\Theta^{\pm}_{P_0}$,

$$\lambda_0([(P_1) - (P_2)]) = 2\langle P_1, P_2\rangle + \langle P_1 + P_2, P_0 + \iota(P_0)\rangle + c$$

for all points $P_1, P_2 \in C(k^{\mathrm{nr}})$ with $P_1 \neq P_2$ and $\{P_1, P_2\} \cap \{P_0, \iota(P_0)\} = \varnothing$, where $\langle \cdot, \cdot \rangle$ is the pairing in [Heinz 2004, Theorem 4.4] and $c \in \mathbb{R}$ is a constant.

If $\mathcal{C}^{\min}$ has semistable reduction, then, by [Heinz 2004, Remark 4.6], the pairing $\langle \cdot, \cdot \rangle$ coincides with Zhang's admissible pairing $(\cdot, \cdot)_a$ [1993] in terms of harmonic

analysis on the reduction graph $R(C)$. In these terms, we have, for $Q, Q' \in C(k^{\mathrm{nr}})$,

$$\langle Q, Q' \rangle = (Q, Q')_a = i(\bar{Q}, \bar{Q}') + g_\nu(\Gamma, \Gamma'),$$

where $i(\bar{Q}, \bar{Q}')$ is the intersection multiplicity of the sections $\bar{Q}, \bar{Q}' \in \mathcal{C}^{\min}(\mathcal{O}^{\mathrm{nr}})$ induced by Q and Q', respectively, and $g_\nu(\Gamma, \Gamma')$ is the Green's function associated to a certain measure ν on $R(C)$, with Γ and Γ' being the respective components of the special fiber of $\mathcal{C}^{\min}$ that Q and Q' reduce to. See [Zhang 1993, §4]. We extend g_ν to a bilinear map on the free abelian group generated by the vertices of $R(C)$.

Lemma 8.2 gives, for $P_0 = \infty$ and $P = [(P_1) - (P_2)]$ with normalized Kummer coordinates $x(P) = (x_1(P), \ldots, x_4(P))$,

$$\begin{aligned}\mu(P) &= v(x_1(P)) - \hat{\lambda}_1(P) \\ &= v(x_1(P)) - 2i(\bar{P}_1, \bar{P}_2) - i(\bar{P}_1 + \bar{P}_2, \bar{P}_0 + \overline{\iota(P_0)}) \\ &\qquad - 2g_\nu(\Gamma_1, \Gamma_2) - g_\nu(\Gamma_1 + \Gamma_2, \Gamma_0 + \Gamma_0') - c,\end{aligned}$$

where Γ_1 and Γ_2 are the respective components that P_1 and P_2 reduce to, and Γ_0 and Γ_0' are the respective components that P_0 and $\iota(P_0)$ reduce to. We assume for a moment that the images of P_1 and P_2 on the special fiber of the original model $\mathcal{C}$ are distinct from the images of the points at infinity. By assumption (i), $\mu(P)$ is unchanged when we replace the points P_1 and P_2 by other points still mapping to Γ_1 and Γ_2, respectively. We can therefore assume that the images of P_1 and P_2 on the special fiber of $\mathcal{C}^{\min}$ are distinct from each other and also from the images of P_0 and $\iota(P_0)$. This implies that $v(x_1(P)) = 0$ and that the intersection numbers in the formula above are zero. We can choose further points Q_1 and Q_2 that also reduce to Γ_1 and Γ_2 with reductions on the special fiber of $\mathcal{C}$ distinct from those of P_0 and $\iota(P_0)$ and such that P_1, P_2, Q_1 and Q_2 all reduce to distinct points on the special fiber of $\mathcal{C}^{\min}$. Using assumptions (ii) and (iii), we obtain the relations

$$\begin{aligned}-\tfrac{1}{2}\mu_1 &= -\tfrac{1}{2}\mu([(P_1) - (Q_1)]) = g_\nu(\Gamma_1, \Gamma_1) + g_\nu(\Gamma_1, \Gamma_0 + \Gamma_0') + \tfrac{1}{2}c, \\ \mu(P) &= \mu([(P_1) - (P_2)]) = -2g_\nu(\Gamma_1, \Gamma_2) - g_\nu(\Gamma_1 + \Gamma_2, \Gamma_0 + \Gamma_0') - c, \\ -\tfrac{1}{2}\mu_2 &= -\tfrac{1}{2}\mu([(P_2) - (Q_2)]) = g_\nu(\Gamma_2, \Gamma_2) + g_\nu(\Gamma_2, \Gamma_0 + \Gamma_0') + \tfrac{1}{2}c.\end{aligned}$$

Adding them together gives

$$\mu(P) - \tfrac{1}{2}(\mu_1 + \mu_2) = g_\nu(\Gamma_1 - \Gamma_2, \Gamma_1 - \Gamma_2) = r(\Gamma_1, \Gamma_2),$$

as desired. See [Zhang 1993, §3] for the last equality.

If our assumption that the images of P_1 and P_2 on the special fiber of the original model $\mathcal{C}$ are distinct from the images of the points at infinity is not satisfied, then we choose another point P_0 for which the assumption is satisfied. We can then perform a change of coordinates τ over $\mathcal{O}$ that moves P_0 to infinity and apply the

result above. By Corollary 4.6 (note that $v(\tau) = 0$ in this case) and the fact that $v(\tau(x)) = v(x)$, $\mu(P)$ is unchanged by τ. □

Remark 8.6. We see from the proof that for two points Q, Q' both having image on a component Γ, but with distinct reductions that are also distinct from those of P_0 and $\iota(P_0)$, we always have

$$\mu([(Q) - (Q')]) = -2g_v(\Gamma, \Gamma) - 2g_v(\Gamma, \Gamma_0 + \Gamma_0') - c.$$

So the assumption that this value does not depend on the choice of Q and Q' is not really necessary.

Theorem 8.7. *Let C be a smooth projective curve of genus 2 defined over a nonarchimedean local field k, given by an integral Weierstrass model. Let J be the Jacobian of C and $\mathcal{J}$ its Néron model over $S = \operatorname{Spec} \mathcal{O}$. Assume that the minimal proper regular model $\mathcal{C}^{\min}$ of C over S is semistable and that μ factors through the component group $\Phi(\mathfrak{k})$ of $\mathcal{J}$. Let $P \in J(k)$ be such that its image in $\Phi(\mathfrak{k})$ is $[\Gamma_1 - \Gamma_2]$, where Γ_1 and Γ_2 are components of the special fiber of $\mathcal{C}^{\min}$. Then we have*

$$\mu(P) = r(\Gamma_1, \Gamma_2).$$

Proof. Since μ factors through $\Phi(\mathfrak{k})$, it follows that $\mu([(P_1) - (P_2)])$ vanishes when P_1 and P_2 map to the same component on the special fiber of $\mathcal{C}^{\min}$ and in general depends only on the components P_1 and P_2 map to. This shows that assumptions (i)–(iii) in Proposition 8.5 are satisfied with $\mu_1 = \mu_2 = 0$. The claim follows. □

9. Formulas and bounds for $\mu(P)$ in the nodal reduction case

In this section and the next, we will deduce explicit formulas for $\mu(P)$ when we have a stably minimal Weierstrass model $\mathcal{C}$. Recall that $\mathcal{C}^{\min}$ denotes the minimal proper regular model of $\mathcal{C}$. In the following, when we speak of components, points, and so on, of the special fiber of $\mathcal{C}$ or $\mathcal{C}^{\min}$, we always mean *geometric* components, points, and so on.

In this section we shall use Theorem 8.7 and Remark 8.4 to find explicit formulas for $\mu(P)$ whenever C/k has nodal reduction, i.e., the special fiber $\mathcal{C}_v$ of $\mathcal{C}$ is reduced and all multiplicities are at most 2. In this case $\mathcal{C}$ is semistable and therefore it has rational singularities. Let $\Delta = \Delta(\mathcal{C})$ denote the discriminant of $\mathcal{C}$; we assume that there is at least one node, so that $v(\Delta) > 0$.

Since there are at most three nodes in the special fiber of $\mathcal{C}$, we have to consider three different cases.

First suppose that there is a unique node in the special fiber of $\mathcal{C}$ and set $m = v(\Delta)$. In the notation of [Namikawa and Ueno 1973] this is reduction type $[\mathrm{I}_{m-0-0}]$. If $m = 1$, then $\mathcal{C}$ is regular over S. In general, there is a unique component, which we denote by A, of genus 1 in the special fiber of $\mathcal{C}^{\min}$. As in the case of multiplicative

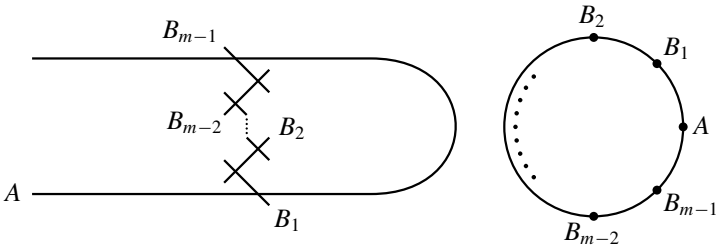


Figure 1. The special fiber of reduction type $[\mathrm{I}_{m-0-0}]$ and its reduction graph.

reduction of elliptic curves (see, for example, [Silverman 1994]), the singular point on the special fiber is replaced by a string of $m-1$ components of $\mathcal{C}^{\min}$, all of genus 0 and multiplicity 1. We choose one of the two components intersecting A and call it B_1 and number the other components $B_2, \ldots, B_{m-1}$ consecutively as in Figure 1.

Using [Bosch et al. 1990, Theorem 9.6.1], it is easy to see that the geometric component group $\Phi(\bar{\mathfrak{k}})$ of the Néron model is generated by $[B_1 - A]$ and is isomorphic to $\mathbb{Z}/m\mathbb{Z}$. We have $[B_j - A] = j \cdot [B_1 - A]$ in $\Phi(\bar{\mathfrak{k}})$.

We set $B_0 := B_m := A$. Then we have the following result.

Proposition 9.1. *Suppose that there is a unique node in the special fiber of $\mathcal{C}$; let m and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If $P \in J(k)$ maps to $[B_i - A]$ in the component group, then we have*

$$\mu(P) = \frac{i(m-i)}{m}.$$

Proof. Since the given model is semistable, we can use Theorem 8.7 and Remark 8.4. One choice of g as in Remark 8.4 is given by

$$g(B_j) = \begin{cases} -\dfrac{j(m-i)}{m} & \text{if } 0 \le j \le i, \\ -\dfrac{i(m-j)}{m} & \text{if } i \le j \le m. \end{cases}$$

Then

$$\mu(P) = r(B_i, A) = -(g(B_i) - g(A)) = \frac{i(m-i)}{m}. \qquad \square$$

Remark 9.2. Proposition 9.1 resembles the formula for the canonical local height on an elliptic curve with split multiplicative reduction given, for instance, in [Silverman 1988].

Now suppose that there are precisely two nodes in the special fiber of $\mathcal{C}$. The reduction type is $[\mathrm{I}_{m_1-m_2-0}]$ in the notation of [Namikawa and Ueno 1973], where $m_1, m_2 \ge 1$ and $m_1 + m_2 = v(\Delta)$. The special fiber of $\mathcal{C}^{\min}$ is obtained by blowing

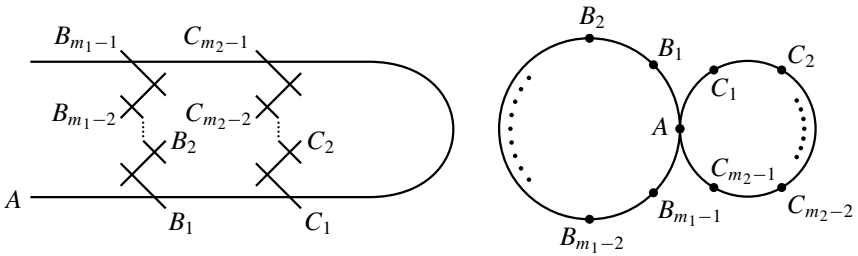


Figure 2. The special fiber of reduction type $[\mathrm{I}_{m_1-m_2-0}]$ and its reduction graph.

up the two singular points of the special fiber of $\mathcal{C}$ repeatedly and replacing them with a chain of $m_1 - 1$ and $m_2 - 1$ curves of genus 0, respectively. We call these components $B_1, \ldots, B_{m_1-1}, C_1, \ldots, C_{m_2-1}$, numbered as in Figure 2, where A contains all images of points reducing to a nonsingular point and we pick components B_1 and C_1 intersecting A as in the case of a unique node. The component group $\Phi(\bar{\mathfrak{k}})$ is isomorphic to $\mathbb{Z}/m_1\mathbb{Z} \times \mathbb{Z}/m_2\mathbb{Z}$ and is generated by $[B_1 - A]$ and $[C_1 - A]$; this follows again using [Bosch et al. 1990, Theorem 9.6.1]. If we have $m_1 = 1$ or $m_2 = 1$, then the corresponding singular point on the special fiber of $\mathcal{C}$ is regular and is therefore not blown up.

We set $B_0 := B_{m_1} := C_0 := C_{m_2} := A$. Then every element of the component group has a representative of the form $[B_i - C_j]$ with $0 \le i \le m_1$ and $0 \le j \le m_2$. The following result expresses $\mu(P)$ in terms of this representative.

Proposition 9.3. *Suppose that there are exactly two nodes in the special fiber of $\mathcal{C}$; let m_1 and m_2 and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If $P \in J(k)$ maps to $[B_i - C_j]$ in the component group, then we have*

$$\mu(P) = \frac{i(m_1 - i)}{m_1} + \frac{j(m_2 - j)}{m_2}.$$

Proof. This is an easy computation along the same lines as in the proof of Proposition 9.1. □

The final case that we have to consider is the case of three nodes in the special fiber of $\mathcal{C}$, which then has two components. We call these components A and E. The special fiber of the minimal proper regular model is obtained using a sequence of blowups of the singular points; they are replaced by a chain of $m_i - 1$ curves of genus 0 and multiplicity 1, respectively, where $v(\Delta) = m_1 + m_2 + m_3$. Hence the special fiber of $\mathcal{C}^{\min}$ contains the two components A and E, connected by three chains of curves of genus 0 that we call $B_1, \ldots, B_{m_1-1}$, $C_1, \ldots, C_{m_2-1}$ and $D_1, \ldots, D_{m_3-1}$, respectively, where B_1, C_1 and D_1 intersect A, as shown in Figure 3. The reduction type is $[\mathrm{I}_{m_1-m_2-m_3}]$.

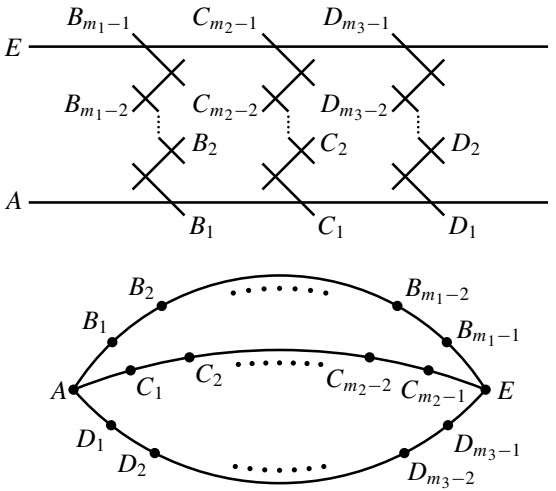


Figure 3. The special fiber of reduction type $[\mathrm{I}_{m_1-m_2-m_3}]$ and its reduction graph.

By [Bosch et al. 1990, Proposition 9.6.10], the group $\Phi(\bar{\mathfrak{k}})$ is isomorphic to $\mathbb{Z}/d\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$, where

$$d = \gcd(m_1, m_2, m_3) \quad \text{and} \quad n = \frac{m_1m_2 + m_1m_3 + m_2m_3}{d}.$$

We set $B_0 := C_0 := D_0 := A$ and $B_{m_1} := C_{m_2} := D_{m_3} := E$. Then it is not hard to see that each element of $\Phi(\bar{\mathfrak{k}})$ can be written in one of the forms

$$[B_i - C_j], \quad [C_j - D_l] \quad \text{or} \quad [D_l - B_i]$$

with $0 \le i \le m_1$, $0 \le j \le m_2$, $0 \le l \le m_3$. The following result allows us to express $\mu(P)$ for any $P \in J(k)$ in terms of the component P maps to.

Proposition 9.4. *Suppose that there are three nodes in the special fiber of $\mathcal{C}$; let m_1, m_2, m_3 and the notation for the components of the special fiber of $\mathcal{C}^{\min}$ be as above. If P maps to $[B_i - C_j]$ in the component group for some $0 \le i \le m_1$ and $0 \le j \le m_2$, then we have*

$$\mu(P) = \frac{m_2 i(m_1 - i) + m_3(i + j)(m_1 - i + m_2 - j) + m_1 j(m_2 - j)}{m_1m_2 + m_1m_3 + m_2m_3}.$$

The formulas for $[C_j - D_l]$ and $[D_l - B_i]$ are analogous.

Proof. The proof is analogous to those of Propositions 9.1 and 9.3. To find g, use that it is piecewise linear on the segments $AB_1 \cdots B_i$, $B_i \cdots B_{m_1-1}E$, $AC_1 \cdots C_j$, $C_j \cdots C_{m_2-1}E$, $AD_1 \cdots D_{m_3-1}E$ and the relations at the vertices A, E, B_i, C_j. □

Remark 9.5. Using the relation $\varepsilon(P) = 4\mu(P) - \mu(2P)$, one can show by a somewhat tedious computation involving a number of different cases that if the image of P in $\Phi(\mathfrak{k})$ is $[\Gamma_1 - \Gamma_2]$, where Γ_1 and Γ_2 are components of the special fiber of $\mathcal{C}^{\min}$, then $\varepsilon(P)$ is the "distance" between Γ_1 and Γ_2 in the reduction graph, where the "length" of the path between B_i and B_j (say, analogously for C_i, C_j and D_i, D_j) is $\min\{2|i-j|, m_1\}$, and otherwise "lengths" are additive. In particular, if $\Phi(\mathfrak{k}) = \Phi(\bar{\mathfrak{k}})$, then

$$\gamma = \max\{\varepsilon(P) : P \in J(k)\} = \max\{m_i + m_j - \delta_{ij} : 1 \le i < j \le 3\},$$

where $\delta_{ij} = 0$ if both m_i and m_j are even, and $\delta_{ij} = 1$ otherwise.

Remark 9.6. In order to use the results of this section to actually compute $\mu(P)$ for a given point $P \in J(k)$, we need to be able to find the component of $\mathcal{J}_v$ that P reduces to. One approach is to find $P_1, P_2 \in C$ such that $P = [(P_1) - (P_2)]$ and find the reductions of P_1 and P_2 to $\mathcal{C}_v^{\min}$. Another approach is to use a transformation (possibly defined over an unramified extension of k) to move the singular points to ∞, $(0, 0)$ and $(1, 0)$, respectively. Then we can (possibly after applying another transformation) read off the component that P maps to directly from the Kummer coordinates of P.

The discussion of this section shows that we get the following results on the local height constant $\beta = \max\{\mu(P) : P \in J(k)\}$. Recall that $\gamma = \max\{\varepsilon(P) : P \in J(k)\}$ and that $\frac{1}{4}\gamma \le \beta \le \frac{1}{3}\gamma$. We will see that in many cases the lower bound is attained.

Let P be a node on $\mathcal{C}_v$; it is defined over a finite extension of $\mathfrak{k}$. We say that the node P is *split* if the two tangent directions of the branches at P are defined over every extension that P is defined over, otherwise P is *nonsplit*. We say that P is *even* if its contribution m_i to the valuation of the discriminant is even, and *odd* otherwise.

Corollary 9.7. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there is a unique node in the special fiber of $\mathcal{C}$ and let $m = v(\Delta)$. Then we have*

$$\beta = \frac{1}{2m}\left\lfloor \frac{m^2}{2} \right\rfloor \le \frac{v(\Delta)}{4}$$

if the node is split or even, and $\beta = 0$ otherwise.

Proof. This follows from Proposition 9.1, taking into account that if m is odd and the node is nonsplit, then the group $\Phi(\mathfrak{k})$ is trivial. □

Remark 9.8. Using the relation $\varepsilon(P) = 4\mu(P) - \mu(2P)$, one can check that

$$\varepsilon(P) = 2\min\{i, m-i\} \quad \text{if } P \text{ maps to } [B_i - A] \text{ in } \Phi(\mathfrak{k}).$$

If m is even (and $\beta > 0$), then $\beta = \frac{1}{4}m = \frac{1}{4}\gamma$. If m is odd, then $\beta = \frac{1}{4}\left(m - \frac{1}{m}\right)$ and $\gamma = m - 1$, so $\beta/\gamma = \frac{1}{4}\left(1 + \frac{1}{m}\right)$ approaches $\frac{1}{4}$ as $m \to \infty$, but for $m = 3$ (the worst case), we have $\beta = \frac{1}{3}\gamma$.

Corollary 9.9. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there are exactly two nodes in the special fiber of $\mathcal{C}$. Let $v(\Delta) = m_1 + m_2$ as above. Then we have*

$$\beta = \frac{1}{2m_1}\left\lfloor \frac{m_1^2}{2} \right\rfloor + \frac{1}{2m_2}\left\lfloor \frac{m_2^2}{2} \right\rfloor \le \frac{v(\Delta)}{4}$$

if each of the nodes is split or even,

$$\beta = \frac{1}{2m_i}\left\lfloor \frac{m_i^2}{2} \right\rfloor$$

if the node corresponding to m_i is split or even and the other node is nonsplit and odd, and $\beta = 0$ if both nodes are nonsplit and odd.

Proof. This follows from Proposition 9.3, taking into account the action of Frobenius on $\Phi(\bar{\mathfrak{k}})$. □

If we have three nodes, then it helps to take the field of definition of the nodes into account.

Corollary 9.10. *Suppose that C/k is a smooth projective curve of genus 2 given by an integral Weierstrass model $\mathcal{C}$ such that there are three nodes in the special fiber of $\mathcal{C}$. We say that $\mathcal{C}$ is split if the two components A and E of the special fiber of $\mathcal{C}^{\min}$ are defined over $\mathfrak{k}$; otherwise $\mathcal{C}$ is nonsplit. Let $v(\Delta) = m_1 + m_2 + m_3$ as above and set $M = m_1m_2 + m_1m_3 + m_2m_3$.*

(a) *If all nodes are $\mathfrak{k}$-rational, $\mathcal{C}$ is split, and we have $m_1 \ge m_3$ and $m_2 \ge m_3$, then*

$$\beta = \frac{1}{2M}\left(m_2\left\lfloor \frac{m_1^2}{2} \right\rfloor + m_3\left\lfloor \frac{(m_1+m_2)^2}{2} \right\rfloor + m_1\left\lfloor \frac{m_2^2}{2} \right\rfloor\right) \le \frac{m_1+m_2}{4} < \frac{v(\Delta)}{4}.$$

(b) *If all nodes are $\mathfrak{k}$-rational, but $\mathcal{C}$ is nonsplit, then*

$$\beta = \max\{0\} \cup \left\{\tfrac{1}{4}(m_i + m_j) : 1 \le i < j \le 3,\ m_i \text{ and } m_j \text{ even}\right\}.$$

(c) *If two of the nodes lie in a quadratic extension of $\mathfrak{k}$ and are conjugate over $\mathfrak{k}$ and one is $\mathfrak{k}$-rational, then*

$$\beta = \begin{cases} \dfrac{m_1}{M}\max\left\{\left\lfloor \dfrac{m_1^2}{2} \right\rfloor + m_1m_3, \left\lfloor \dfrac{m_3^2}{2} \right\rfloor + m_1\left\lfloor \dfrac{m_3}{2} \right\rfloor\right\} & \text{if } \mathcal{C} \text{ is split,} \\ \dfrac{m_1}{2} & \text{if } \mathcal{C} \text{ is nonsplit and } m_1 \text{ is even,} \\ 0 & \text{otherwise,} \end{cases}$$

where m_3 corresponds to the rational node (and $m_1 = m_2$).

(d) *If all nodes are defined over a cubic extension of* $\mathfrak{k}$ *and are conjugate over* $\mathfrak{k}$, *then* $m_1 = m_2 = m_3 = \frac{1}{3}v(\Delta)$ *and*

$$\beta = \begin{cases} \frac{1}{9}v(\Delta) & \text{if } \mathcal{C} \text{ is split,} \\ 0 & \text{otherwise.} \end{cases}$$

Proof. The proof of (a) follows easily from Proposition 9.4.

For the other cases, note that in the nonsplit case some power of Frobenius acts as negation on the component group $\Phi(\bar{\mathfrak{k}})$, so the only elements of $\Phi(\mathfrak{k})$ are elements of order 2 in $\Phi(\bar{\mathfrak{k}})$, which correspond to $[B_{m_1/2} - C_{m_2/2}]$ if m_1 and m_2 are even $\big($where μ takes the value $\frac{1}{4}(m_1 + m_2)\big)$, and similarly with the obvious cyclic permutations.

In the situation of (c), we must have $m_1 = m_2$. If $P = [(P_1) - (P_2)] \in J(k)$ and $P_1 \in C(\bar{k})$ maps to one of the conjugate nodes, then P_2 must map to the other, so all $P \in J(k)$ must map to a component of the form $[B_i - C_j]$ or $[D_i - D_j]$. Now the result in the split case follows from a case distinction depending on whether $m_1 \le m_3$ or not. In the nonsplit case, the only element of order 2 that is defined over $\mathfrak{k}$ is $[B_{m_1/2} - C_{m_1/2}]$ if it exists.

In the situation of (d), the group $\Phi(\mathfrak{k})$ is of order 3 (generated by $[E - A]$) in the split case and trivial in the nonsplit case. □

Extending the valuation $v\colon k^\times \twoheadrightarrow \mathbb{Z}$ to $\bar{v}\colon \bar{k}^\times \to \mathbb{Q}$, we get extensions of ε and μ to $J(\bar{k})$. Denote $\max\{\mu(P) : P \in J(\bar{k})\}$ by $\bar{\beta}$ and $\max\{\varepsilon(P) : P \in J(\bar{k})\}$ by $\bar{\gamma}$. Then by the discussion at the beginning of Section 8 and the results above, we find that

$$\bar{\beta} = \tfrac{1}{4}\bar{\gamma} = \tfrac{1}{4}v(\Delta),$$

when there are one or two nodes, and

$$\tfrac{1}{6}v(\Delta) \le \bar{\beta} = \tfrac{1}{4}\bar{\gamma} = \tfrac{1}{4}(v(\Delta) - \min\{m_1, m_2, m_3\}) < \tfrac{1}{4}v(\Delta),$$

when there are three nodes. (Equality is achieved as soon as the Galois action on $R(C)$ is trivial and the ramification index is even.)

10. Formulas and bounds for $\mu(P)$ in the cuspidal reduction case

In this section we consider the case of a stably minimal Weierstrass model $\mathcal{C}$ such that there are (one or two) points of multiplicity 3 on the special fiber. These points are either both $\mathfrak{k}$-rational or they are defined over a quadratic extension of $\mathfrak{k}$ and are conjugate over $\mathfrak{k}$.

In the notation of [Namikawa and Ueno 1973], the reduction type is of the form $[\mathcal{K}_1 - \mathcal{K}_2 - l]$, where $l \ge 0$ and $\mathcal{K}_j$ is an elliptic Kodaira type for $j \in \{1, 2\}$. We can

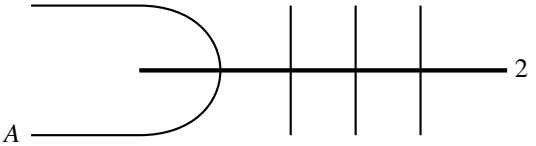


Figure 4. The special fiber of reduction type $[\mathrm{I}_0 - \mathrm{I}_0^* - 0]$.

compute $\mathcal{K}_1$, $\mathcal{K}_2$ and l as in [Liu 1994, §6.1]. By [Liu 1994, §7], we have

$$\Phi(\bar{\mathfrak{k}}) \cong \Phi_1(\bar{\mathfrak{k}}) \times \Phi_2(\bar{\mathfrak{k}}),$$

where Φ_j is the component group of an elliptic curve with Kodaira type $\mathcal{K}_j$. As in the previous section, we write $\Delta = \Delta(\mathcal{C})$ for the discriminant of the model $\mathcal{C}$.

If $\mathcal{C}$ is not regular, then we can compute the minimal proper regular model $\mathcal{C}^{\min}$ of C from $\mathcal{C}$ by a sequence of blowups in the singular point(s) of $\mathcal{C}$, so the corresponding morphism $\zeta : \mathcal{C}^{\min} \to \mathcal{C}$ is the minimal desingularization of $\mathcal{C}$.

Suppose that $l > 0$. Then the special fiber of $\mathcal{C}^{\min}$ consists of Kodaira types $\mathcal{K}_1$ and $\mathcal{K}_2$, connected by a chain of $l-1$ rational curves. See for example Figure 5. The desingularization ζ contracts $\mathcal{K}_2$ to one of the singular points; in this case we say that this point *corresponds to* $\mathcal{K}_2$. If there is another singular point in $\mathcal{C}_v(\bar{\mathfrak{k}})$, then it corresponds to $\mathcal{K}_1$; otherwise we must have $\mathcal{K}_1 = \mathrm{I}_0$.

Suppose now that $l = 0$. If both $\mathcal{K}_1$ and $\mathcal{K}_2$ are good or multiplicative, then we are in the situation $[\mathrm{I}_{m_1 - m_2 - 0}]$ for some $m_1, m_2 \geq 0$, which we have discussed in the previous section. So we may assume that at least one of the $\mathcal{K}_j$ is additive, say $\mathcal{K}_2$. Then $\mathcal{C}_v^{\min}$ looks like Kodaira type $\mathcal{K}_2$, but with one of the rational curves replaced by (see [Namikawa and Ueno 1973]):

- a curve A of genus 1 if $\mathcal{K}_1 = \mathrm{I}_0$ (see Figure 4 for the case $\mathcal{K}_2 = \mathrm{I}_0^*$);
- one of the rational components of $\mathcal{K}_1$, otherwise; the remainder of $\mathcal{K}_1$ is then attached to this component.

We say that a singularity corresponds to one of the Kodaira types $\mathcal{K}_1$ or $\mathcal{K}_2$ similarly to the case $l > 0$.

Lemma 10.1. *Suppose that the residue characteristic of k is not 2. Let C be given by a stably minimal Weierstrass model with reduction type $[\mathcal{K}_1 - \mathcal{K}_2 - l]$. Then after at most a quadratic unramified extension of k there is a stably minimal Weierstrass model*

$$\mathcal{C}\colon Y^2 = F(X, Z) = f_6X^6 + f_5X^5Z + f_4X^4Z^2 + X^3Z^3 + f_2X^2Z^4 + f_1XZ^5 + f_0Z^6$$

of C, isomorphic to the given model of C, such that the elliptic curve with Weierstrass model

$$\mathcal{E}_1\colon Y^2Z = X^3 + f_2X^2Z + f_1XZ^2 + f_0Z^3$$

has Kodaira type $\mathcal{K}_1$ and the elliptic curve with Weierstrass model

$$\mathcal{E}_2 \colon Y^2Z = X^3 + f_4X^2Z + f_5XZ^2 + f_6Z^3$$

has Kodaira type $\mathcal{K}_2$.

Proof. After possibly making a quadratic unramified extension and applying a transformation, we can assume that there is a unique point $\infty \in \mathcal{C}_v(\mathfrak{k})$ at infinity on the special fiber and that it is a cusp, corresponding to $\mathcal{K}_2$; see the discussion preceding the lemma. Moreover, we can assume that if there is another singular point in $\mathcal{C}_v(\bar{\mathfrak{k}})$, then this point is $P = (0, 0) \in \mathcal{C}_v(\mathfrak{k})$ (in which case it must correspond to $\mathcal{K}_1$).

Because the residue characteristic is not 2, we may assume that $\mathcal{C}$ has $H = 0$ and that f_3 is a unit. By Hensel's lemma there is a factorization $F = F_1F_2$, where F_2 is a cubic form reducing to Z^3. Similarly, we may assume that F_1 reduces to X^3 if there is a cusp at P and to $X^2(X + aZ)$ with $a \neq 0$ if there is a node at P; otherwise F_1 is squarefree. Consider the elliptic curves given by the Weierstrass models

$$\mathcal{D}_1 \colon Y^2Z = F_1(X, Z) \quad \text{and} \quad \mathcal{D}_2 \colon Y^2Z = F_2(Z, X).$$

We first show that $\mathcal{D}_1$ has Kodaira type $\mathcal{K}_1$ and $\mathcal{D}_2$ has Kodaira type $\mathcal{K}_2$.

If $\mathcal{D}_1$ is not minimal, then we can apply a transformation to $\mathcal{C}$ which makes $\mathcal{D}_1$ minimal. This decreases the valuation of the discriminant $\Delta(\mathcal{D}_1)$, but increases the valuation of $\Delta(\mathcal{D}_2)$ by the same amount. The resulting model is still stably minimal and the resulting F_2 still reduces to Z^3. Hence we may assume that $\mathcal{D}_1$ is minimal.

Let $Q = (0, 0) \in \mathcal{D}_{1,v}(\mathfrak{k})$; then $\mathcal{D}_1$ is smooth outside Q. Note that F_2 is a unit in $\mathcal{O}_{\mathcal{C},P}$, so that P is a smooth point if and only if Q is a smooth point, in which case $\mathcal{D}_1$ has reduction type $\mathrm{I}_0 = \mathcal{K}_1$. More generally, $\mathcal{C}$ is regular at P if and only if $\mathcal{D}_1$ is regular at Q, and P is a node (resp. a cusp) if and only if Q is a node (resp. a cusp). Recall that P corresponds to $\mathcal{K}_1$, so that $\mathcal{D}_1$ has reduction type I_1 (resp. II) if and only if $\mathcal{K}_1 = \mathrm{I}_1$ (resp. $\mathcal{K}_1 = \mathrm{II}$).

Now suppose that $\mathcal{C}$ is not regular at P and $\mathcal{D}_1$ is not regular at Q. The minimal desingularization $\xi \colon \mathcal{C}' \to \mathcal{C}$ in P can be computed by a sequence of blowups, starting with the blowup of $\mathcal{C}$ in P. The preimage of P under the latter map is contained in the chart $\mathcal{C}^1$ obtained by dividing the x- and y-coordinates by the uniformizing element π. Similarly, in order to compute the minimal desingularization $\xi_1 \colon \mathcal{D}_1' \to \mathcal{D}_1$ in Q, we first blow up $\mathcal{D}_1$ in Q; then the chart $\mathcal{D}_1^1$ obtained by dividing the x- and y-coordinates by π contains the preimage of Q. But because F_2 reduces to Z^3, the special fibers of $\mathcal{C}^1$ and $\mathcal{D}_1^1$ are identical. This continues to hold after further blowups (if any are necessary), so we have $\xi^{-1}(P) = \xi_1^{-1}(Q)$. There are no exceptional components in these preimages, since we assumed that $\mathcal{D}_1$ is minimal. Therefore $\mathcal{D}_1'$ is in fact the minimal proper regular model of the elliptic curve defined by $\mathcal{D}_1$. Since the minimal desingularization of $\mathcal{C}'$ in the point $\infty \in \mathcal{C}_v'(\mathfrak{k})$ leads to $\mathcal{C}^{\min}$, and since P corresponds to $\mathcal{K}_1$, we deduce that $\mathcal{D}_1$ has Kodaira type $\mathcal{K}_1$.

A similar argument (for which we first apply a transformation to make $\mathcal{D}_2$ minimal) shows that $\mathcal{D}_2$ has Kodaira type $\mathcal{K}_2$. To complete the proof of the lemma, we therefore only need to make sure that $\mathcal{E}_i$ has the same reduction type as $\mathcal{D}_i$ for $i = 1, 2$. This is certainly satisfied if the coefficients of $\mathcal{E}_i$ and $\mathcal{D}_i$ agree modulo π^{N_i+1}, where N_i is the number of blowups needed to construct the minimal desingularization of $\mathcal{D}_i$. Suppose that $F_1 = a_0Z^3 + a_1XZ^2 + a_2X^2Z + a_3X^3$ and $F_2 = b_3Z^3 + b_2XZ^2 + b_1X^2Z + b_0X^3$. Writing out the coefficients of F in terms of the coefficients of F_1 and F_2, we see that it suffices to have

$$v(a_0b_2) > v(a_1), \quad v(a_0b_1 + a_2b_2) > v(a_2),$$
$$v(b_0a_2) > v(b_1), \quad v(a_1b_0 + a_2b_1) > v(b_2).$$

If this is not satisfied, it can be achieved by acting on the given stably minimal Weierstrass model via a suitable element of $\mathrm{GL}_2(\mathcal{O})$ as in Section 4. Finally, we scale the variables to get $f_3 = 1$. □

Remark 10.2. If the residue characteristic is 2, then it is not hard to see that one can also construct a stably minimal Weierstrass model $\mathcal{C}$ and corresponding elliptic Weierstrass models $\mathcal{E}_1$ and $\mathcal{E}_2$ as in the lemma in a similar way. The construction is more cumbersome, since we cannot assume $H = 0$.

In view of Theorem 7.4 we want a condition for $\mathcal{C}$ to have rational singularities.

Lemma 10.3. *The model $\mathcal{C}$ has rational singularities if and only if $l = 0$.*

Proof. We may assume that $\mathcal{C}$ is as in Lemma 10.1 or Remark 10.2. Then all points in $\mathcal{C}_v(\bar{\mathfrak{k}}) \setminus \{\infty, P\}$ are nonsingular, where $\infty \in \mathcal{C}_v(\mathfrak{k})$ is the unique point at infinity, and $P = (0, 0) \in \mathcal{C}_v(\mathfrak{k})$. If $\mathcal{C}$ is regular in P, then P is a rational singularity. If not, then, by [Artin 1966, Theorem 3], P is a rational singularity if and only if the fundamental cycle of $\xi^{-1}(P)$ has arithmetic genus 0, where ξ is any desingularization of P. In particular, the assertion that P is a rational singularity depends only on the configuration of $\xi^{-1}(P)$, where $\xi : \mathcal{C}' \to \mathcal{C}$ is the minimal desingularization of P. Now let $\mathcal{E}_1$ be as in Lemma 10.1 or Remark 10.2, and let $\xi_1 : \mathcal{E}_1' \to \mathcal{E}_1$ denote the minimal desingularization of the singular point $Q = (0, 0) \in \mathcal{E}_{1,v}(\mathfrak{k})$; then the assertion that Q is a rational singularity depends only on the configuration of $\xi_1^{-1}(Q)$. We have $\xi^{-1}(P) = \xi_1^{-1}(Q)$ as in the proof of Lemma 10.1 (this also works when $\mathrm{char}\,\mathfrak{k} = 2$ and does not require minimality of $\mathcal{E}_1$). In particular, P is a rational singularity if and only if Q is a rational singularity.

A similar argument proves the corresponding statement for $\mathcal{E}_2$. Hence $\mathcal{C}$ has rational singularities if and only if both $\mathcal{E}_1$ and $\mathcal{E}_2$ have rational singularities. By [Conrad 2005, Corollary 8.4] a Weierstrass model of an elliptic curve has rational singularities if and only if it is minimal. But it is easy to see that $\mathcal{E}_1$ and $\mathcal{E}_2$ are both minimal if and only if $l = 0$. □

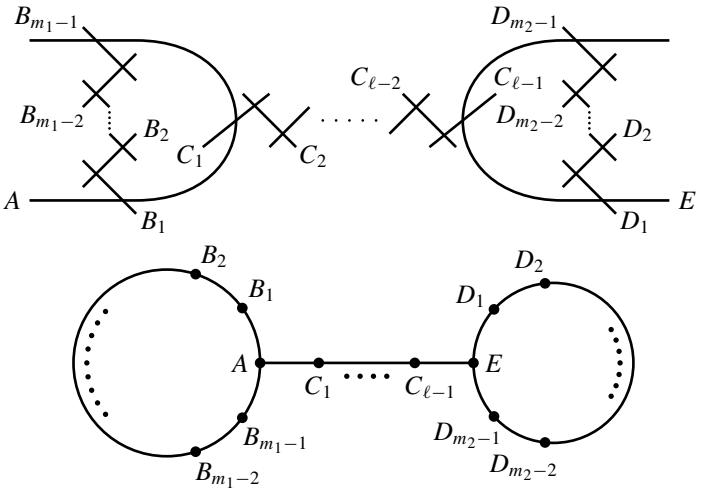


Figure 5. The special fiber of reduction type $[\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$ and its reduction graph.

According to Lemma 10.3, not all singularities of the given stably minimal Weierstrass model $\mathcal{C}$ are rational when $l > 0$. The following example shows that in this situation $\varepsilon(P) \neq 0$, and hence $\mu(P) \neq 0$, can indeed occur for $P \in J_0(k)$.

Example 10.4. Let p be an odd prime and let $C/\mathbb{Q}_p$ be given by

$$Y^2 = Z(X^2 + Z^2)(X^3 + p^5 X Z^2 + p^8 Z^3).$$

Let $P_1 = (0, p^4) \in C(\mathbb{Q}_p)$ and $P_2 = \iota(P_1)$. The reduction type is $[\mathrm{I}_0 - \mathrm{III} - 1]$ and hence $\#\Phi(\bar{\mathfrak{k}}) = 2$. It turns out that both P_1 and P_2 map to the same component and so we have $P = [(P_1) - (P_2)] \in J_0(k)$. The image of P on the Kummer surface is of the form $(x_1 : 0 : 0 : x_4)$, where $v(x_4) - v(x_1) = 2$. We get $\varepsilon(P) = \varepsilon(2P) = 6$ and $\mu(P) = \mu(2P) = 2$.

The case of semistable reduction, corresponding to reduction type $[\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$ (see Figure 5) deserves special attention. Here $l \geq 1$, by the discussion above. Note that $m_1 = 0$ (or $m_2 = 0$) is possible; in that case A (or E) is a curve of genus 1 and there are no components B_i (or D_i). If $m_1 = 1$ (or $m_2 = 1$), then A (or E) is a nodal curve (again there are no B_i or D_i). After perhaps an unramified quadratic extension, we can assume that all components in the "chain" that connects the two polygons in the special fiber of $\mathcal{C}^{\min}$ are defined over $\mathfrak{k}$. There are then $l + 1$ different (meaning pairwise nonisomorphic over $\mathcal{O}$) minimal Weierstrass models of the curve; compare the proof of Lemma 5.4. Explicitly, these models can be taken to have the form

$$\begin{aligned} \mathcal{C}_j \colon Y^2 + (h_0 \pi^{3j} Z^3 + h_1 \pi^j Z^2 X + h_2 \pi^{l-j} Z X^2 + h_3 \pi^{3(l-j)} X^3) Y \\ = f_0 \pi^{6j} Z^6 + f_1 \pi^{4j} X Z^5 + f_2 \pi^{2j} X^2 Z^4 + X^3 Z^3 \\ + f_4 \pi^{2(l-j)} X^4 Z^2 + f_5 \pi^{4(l-j)} X^5 Z + f_6 \pi^{6(l-j)} X^6 \end{aligned} \quad (10\text{-}1)$$

for $j = 0, 1, \ldots, l$, where

$$y^2 + h_1xy + h_0y = x^3 + f_2x^2 + f_1x + f_0$$

and

$$y^2 + h_2xy + h_3y = x^3 + f_4x^2 + f_5x + f_6$$

are minimal Weierstrass equations of elliptic curves of reduction types I_{m_1} and I_{m_2}, respectively. Such a model corresponds to the vertex C_j of the reduction graph (where we set $C_0 = A$ and $C_l = E$); the corresponding component of the special fiber of $\mathcal{C}^{\min}$ is the one that is visible in the special fiber of $\mathcal{C}_j$. The valuation of the discriminant of $\mathcal{C}_j$ is $m_1 + m_2 + 12l$ and does not depend on j.

A *simple path* in $R(C)$ is a subgraph that is a tree without vertices of valency ≥ 3. Let $P_1, P_2 \in C(k)$ reduce to components Γ_1 and Γ_2 of the special fiber of $\mathcal{C}^{\min}$, respectively. Consider the model $\mathcal{C}_j$ of C. If there is a simple path from Γ_1 to Γ_2 in the reduction graph that passes through C_j, then we say that $\mathcal{C}_j$ *lies between* P_1 and P_2. We denote the μ-function computed with respect to $\mathcal{C}_j$ by μ_j.

Proposition 10.5. *Assume that C has semistable reduction of type $[\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$. Let $P_1, P_2 \in C(k)$ be points reducing to components Γ_1 and Γ_2 of the special fiber of $\mathcal{C}^{\min}$ and let $j \in \{0, 1, \ldots, l\}$. Define $j_{\min}$ and $j_{\max}$ to be the smallest and largest $j' \in \{0, 1, \ldots, l\}$ such that $\mathcal{C}_{j'}$ lies between P_1 and P_2. Let $P = [(P_1) - (P_2)] \in J(k)$. Then*

$$r(\Gamma_1, \Gamma_2) + j_{\max} - j_{\min} \leq \mu_j(P) \leq r(\Gamma_1, \Gamma_2) + |j - j_{\max}| + |j - j_{\min}|.$$

If $\mathcal{C}_j$ lies between P_1 and P_2, then the inequalities are equalities.

Proof. First note that the last statement follows from the first, since $j_{\min} \leq j \leq j_{\max}$ implies $j_{\max} - j_{\min} = |j - j_{\max}| + |j - j_{\min}|$.

Let $B_0 = B_{m_1} = A$ and $D_0 = D_{m_2} = E$. We prove a number of lemmas.

Lemma 10.6. *If $j = j_{\max} = j_{\min} \in \{0, l\}$, then $\mu_j(P) = r(\Gamma_1, \Gamma_2)$.*

Proof. We assume that $j = j_{\max} = j_{\min} = l$; the other case is analogous. Then Γ_1 and Γ_2 are both of the form D_i, and we consider the model $\mathcal{C}_l$. We first claim that $\mu(P) = 0$ if $\Gamma_1 = \Gamma_2$, but the images of P_1 and P_2 on Γ_1 are distinct. This is clear if $\Gamma_1 = D_0 = E$, since in this case P is in the image of α; compare Lemmas 7.1 and 7.2. Otherwise, we note that the points with nonzero multiplicity on the special fiber of $\mathcal{C}_l$ have multiplicities 1, 2 and 3. Transforming the equation over $\mathcal{O}$ if necessary, we can assume that its reduction is case 7 in Table 1 of [Stoll 2002] or (if the residue characteristic is 2) case 5 in Table 2 here.

Recall that $\Gamma_1 = \Gamma_2 = D_i$, where we can assume $0 < i \leq \frac{1}{2}m_2$. Applying a transformation, we may assume that the points $P_1 = (\xi_1 : \eta_1 : 1)$ and $P_2 = (\xi_2 : \eta_2 : 1)$ both reduce to $(0 : 0 : 1)$ modulo π and that $m_2 = \min\{v(f_0), 2v(f_1)\}$. First suppose

that $i < \frac{1}{2}m_2$. We then have $v(\xi_1) = v(\xi_2) = v(\xi_1 - \xi_2) = i$. Normalizing the Kummer coordinates x of P so that $x_1 = 1$, we can check that $v(x_2)$ and $v(x_3)$ are positive, but that $v(x_4) = 0$. This follows because $\Gamma_1 = D_i = \Gamma_2$ implies that $v(f_2\xi_1\xi_2 + 2\eta_1\eta_2) = 2i$ if $\mathrm{char}(\mathfrak{k}) \neq 2$ and $H = 0$ and that $v(\xi_1\eta_2 + \xi_2\eta_1) = 2i$ if $\mathrm{char}(\mathfrak{k}) = 2$. By a similar argument, the reduction of the image of P on the Kummer surface has nonvanishing last coordinate if m_2 is even and $i = \frac{1}{2}m_2$. According to the tables, this implies that $\varepsilon(P) = 0$ and therefore also $\mu(P) = 0$.

Now consider the case that Γ_1 and Γ_2 do not necessarily coincide. The considerations above imply that the assumptions of Proposition 8.5 are satisfied with $\mu_1 = \mu_2 = 0$ (where we use Lemma 3.7 for the first assumption); the proposition then establishes the claim. □

Lemma 10.7. *Assume that $\Gamma_1 = \Gamma_2 = C_j$ with $0 < j < l$. Then $\mu_j(P) = 0$.*

Proof. In this case, P is in the image of α, so the claim follows by Proposition 7.3. □

Note that Lemmas 10.6 and 10.7 establish the claim of Proposition 10.5 in all cases such that $j = j_{\min} = j_{\max}$.

Lemma 10.8. *Assume that both $\mathcal{C}_j$ and $\mathcal{C}_{j+1}$ lie between P_1 and P_2, where $0 \le j < l$. Then $\mu_j(P) = \mu_{j+1}(P)$.*

Proof. Let $\tau\colon (\xi : \eta : \zeta) \mapsto (\pi\xi : \eta : \pi^{-1}\zeta)$; then τ gives an isomorphism from the generic fiber of $\mathcal{C}_j$ to that of $\mathcal{C}_{j+1}$. The induced map on Kummer coordinates is

$$(x_1, x_2, x_3, x_4) \mapsto (\pi^{-2}x_1, x_2, \pi^2 x_3, x_4);$$

we have $v(\tau) = 0$. Since both $\mathcal{C}_j$ and $\mathcal{C}_{j+1}$ lie between P_1 and P_2, assuming that Γ_1 is to the left and Γ_2 to the right of C_j and C_{j+1}, we must have that the x-coordinate of P_1 on $\mathcal{C}_j$ does not reduce to infinity, whereas that of P_2 does. For normalized Kummer coordinates $x = (x_1, x_2, x_3, x_4)$ of P on the Kummer surface associated to $\mathcal{C}_j$, this implies $v(x_2) = 0$ (the point is not in the kernel of reduction, so $v(x_4) \ge \min\{v(x_1), v(x_2), v(x_3)\}$) and $v(x_1) > 0$. Comparing valuations in the equation of $\mathcal{C}_j$, we see that $P_2 = (1 : \eta : \zeta)$ must have $v(\zeta) \ge 2$, which implies $v(x_1) \ge 2$. It follows that $v(\tau(x)) = 0 = v(x)$. By Corollary 4.6 we also have $\hat{\lambda}(\tau(x)) = \hat{\lambda}(x)$ (recall that $v(\tau) = 0$). Since

$$-v(x) - \mu_j(P) = \hat{\lambda}(x) = \hat{\lambda}(\tau(x)) = -v(\tau(x)) - \mu_{j+1}(P),$$

the claim follows. □

Lemma 10.9. *If $\mathcal{C}_j$ lies between P_1 and P_2, then $\mu_j(P)$ depends only on Γ_1 and Γ_2.*

Proof. Let $P_1', P_2' \in C(k)$ be points also mapping to Γ_1 and Γ_2, respectively. We assume without loss of generality that Γ_1 is to the left of Γ_2. By Lemma 10.6 or

Lemma 10.7, we have that $\mu_{j_{\min}}([(P_1)-(P_1')])=0$ and $\mu_{j_{\max}}([(P_2)-(P_2')])=0$. Using Lemmas 10.8 and 3.7, we obtain

$$\begin{aligned}\mu_j([(P_1')-(P_2')]) &= \mu_{j_{\min}}([(P_1')-(P_2')]) = \mu_{j_{\min}}([(P_1)-(P_2')])\\ &= \mu_{j_{\max}}([(P_1)-(P_2')]) = \mu_{j_{\max}}([(P_1)-(P_2)]) = \mu_j(P). \quad \square\end{aligned}$$

Lemma 10.10. *Let $P_1', P_2' \in C(k)$ be points mapping to distinct points on the same component of the special fiber of $\mathcal{C}^{\min}$ and let $P' = [(P_1')-(P_2')] \in J(k)$. Let j_0 be the unique index such that $\mathcal{C}_{j_0}$ lies between P_1' and P_2'. Then $\mu_j(P') = 2|j-j_0|$.*

Proof. By Lemmas 10.6 and 10.7, we have $\mu_{j_0}(P')=0$. Since the images of P_1' and P_2' on the special fiber of $\mathcal{C}^{\min}$ are distinct, P' is not in the kernel of reduction with respect to $\mathcal{C}_{j_0}$. If

$$x^{(j_0)} = \big(x_1^{(j_0)}, x_2^{(j_0)}, x_3^{(j_0)}, x_4^{(j_0)}\big)$$

are normalized Kummer coordinates for P' on the Kummer surface associated to $\mathcal{C}_{j_0}$, we therefore have

$$0 = v(x^{(j_0)}) = \min\big\{v(x_1^{(j_0)}), v(x_2^{(j_0)}), v(x_3^{(j_0)})\big\}.$$

Applying a suitable power of τ (see the proof of Lemma 10.8), we find that

$$x^{(j)} = \big(\pi^{2(j_0-j)}x_1^{(j_0)}, x_2^{(j_0)}, \pi^{2(j-j_0)}x_3^{(j_0)}, x_4^{(j_0)}\big)$$

are (not necessarily normalized) Kummer coordinates for P' on the Kummer surface associated to $\mathcal{C}_j$. For definiteness, assume that $j > j_0$, the case $j = j_0$ being clear. Similarly to the proof of Lemma 10.8, we find that $0 = v(x^{(j_0)}) = v(x_1^{(j_0)})$, which implies that $v(x^{(j)}) = -2(j-j_0)$. In the same way as in the proof of Lemma 10.8, we deduce $\mu_j(P') = 2(j-j_0) = 2|j-j_0|$. $\square$

To continue the proof of the proposition, we now first consider the case that $\mathcal{C}_j$ lies between P_1 and P_2. In this case, Lemmas 10.9 and 10.10 show that the assumptions in Proposition 8.5 hold with $\mu_1 = 2|j-j_{\min}|$ and $\mu_2 = 2|j-j_{\max}|$ or conversely. So the statement follows from Proposition 8.5 and $|j-j_{\max}|+|j-j_{\min}| = j_{\max}-j_{\min}$.

Now assume that $\mathcal{C}_j$ does not lie between P_1 and P_2. We assume for definiteness that $j > j_{\max}$. For normalized Kummer coordinates $x^{(j_{\max})}$ for $P = [(P_1)-(P_2)]$ on the Kummer surface associated to $\mathcal{C}_{j_{\max}}$, we have

$$v(x_2^{(j_{\max})}) \le \min\big\{v(x_1^{(j_{\max})}), v(x_3^{(j_{\max})})\big\};$$

compare the proof of Lemma 10.8 above. Then $x^{(j)} = \tau^{j-j_{\max}}(x^{(j_{\max})})$ are Kummer coordinates for $[(P_1)-(P_2)]$ on the Kummer surface associated to $\mathcal{C}_j$, and we have

$$v(x^{(j_{\max})}) - 2(j-j_{\max}) \le v(x^{(j)}) \le v(x^{(j_{\max})}).$$

It follows that

$$\begin{aligned}\mu_j(P) - \mu_{j_{\max}}(P) &= \bigl(-\hat{\lambda}(x^{(j)}) - v(x^{(j)})\bigr) - \bigl(-\hat{\lambda}(x^{(j_{\max})}) - v(x^{(j_{\max})})\bigr)\\ &= v(x^{(j_{\max})}) - v(x^{(j)}) \in \{0, 1, \dots, 2(j - j_{\max})\}.\end{aligned}$$

As $\mu_{j_{\max}}(P) = r(\Gamma_1, \Gamma_2) + j_{\max} - j_{\min}$ by the case already discussed, the result follows, and the proof of Proposition 10.5 is finished. □

Corollary 10.11. *Let $\mathcal{C}$ be a stably minimal Weierstrass model of C with discriminant Δ; assume that C has reduction type $[\mathrm{I}_{m_1} - \mathrm{I}_{m_2} - l]$ with $l > 0$. As usual, let*

$$\beta(\mathcal{C}) = \max\{\mu(P) : P \in J(k)\} \quad \text{and} \quad \bar{\beta}(\mathcal{C}) = \max\{\mu(P) : P \in J(\bar{k})\},$$

where μ is computed with respect to $\mathcal{C}$. Then we have

$$\beta(\mathcal{C}) \le \bar{\beta}(\mathcal{C}) = \tfrac{1}{4}(m_1 + m_2) + 2l < \tfrac{1}{4}v(\Delta) \quad \text{and} \quad \bar{\beta} \ge \tfrac{1}{6}v(\Delta).$$

Proof. The assumption on the reduction type implies that the model is equivalent to one of the form (10-1). Proposition 10.5 then gives upper bounds for $\mu([(P_1) - (P_2)])$, with $P_1, P_2 \in C(\bar{k})$, depending on the images Γ_1 and Γ_2 of P_1 and P_2 in the reduction graph. The maximizing case occurs for $\Gamma_1 = B_{m_1/2}$ and $\Gamma_2 = D_{m_2/2}$, giving

$$\mu([(P_1) - (P_2)]) = r(B_{m_1/2}, D_{m_2/2}) + l = \tfrac{1}{4}m_1 + l + \tfrac{1}{4}m_2 + l.$$

For the remaining inequalities, recall that $v(\Delta) = m_1 + m_2 + 12l$ and that $l > 0$. □

We state a technical lemma which will be needed for the proof of Theorem 10.13.

Lemma 10.12. *Suppose that the residue characteristic of k is not 2. Consider a degenerate Weierstrass equation of the form*

$$\mathcal{C}\colon Y^2 = f_0 Z^6 + f_1 X Z^5 + f_2 X^2 Z^4 + X^3 Z^3$$

and let

$$\mathcal{E}\colon y^2 = f_0 + f_1 x + f_2 x^2 + x^3$$

be an elliptic Weierstrass equation. If $Q_1 = (x_1, y_1)$ and $Q_2 = (x_2, y_2)$ are points in $\mathcal{E}(k)$, then $P_1 = (x_1 : y_1 : 1)$ and $P_2 = (x_2 : y_2 : 1)$ are points in $\mathcal{C}(k)$, and if $x_1, x_2 \in \mathcal{O}$, then $\mu_{\mathcal{C}}([(P_1) - (P_2)]) \le \mu_{\mathcal{E}}(Q_1 - Q_2)$.

Here $\mu_{\mathcal{E}}$ is the height correction function for the elliptic curve $\mathcal{E}$, and $\mu_{\mathcal{C}}$ denotes the height correction function defined in the same way as μ in the smooth case in terms of the equation $\mathcal{C}$.

Proof. Let $\underline{\delta}_{\mathcal{C}} = (\delta_{\mathcal{C},1}, \delta_{\mathcal{C},2}, \delta_{\mathcal{C},3}, \delta_{\mathcal{C},4})$ be the duplication polynomials on the Kummer surface associated to $\mathcal{C}$, and let $\underline{\delta}_{\mathcal{E}} = (\delta_{\mathcal{E},1}, \delta_{\mathcal{E}_2})$ be the duplication polynomials for the numerator and denominator of the x-coordinate associated to $\mathcal{E}$. Then a generic computation shows that, if $(\xi_1 : \xi_2 : \xi_3 : \xi_4)$ is the image of $[(P_1) - (P_2)]$ on

the Kummer surface, we have $(\xi_4 : \xi_1) = x(Q_1 - Q_2)$. In addition, we find that $\delta_{\mathcal{C},1}(\xi_1, \xi_2, \xi_3, \xi_4) = \delta_{\mathcal{E},2}(\xi_4, \xi_1)$ and $\delta_{\mathcal{C},4}(\xi_1, \xi_2, \xi_3, \xi_4) = \delta_{\mathcal{E},1}(\xi_4, \xi_1)$ (as polynomials in the ξ_j).

That $P_1, P_2 \in \mathcal{C}(k)$ is obvious from the equations. For the last statement, we observe that $\min\{v(\xi_1), v(\xi_2), v(\xi_3), v(\xi_4)\} = \min\{v(\xi_1), v(\xi_4)\}$ (this is where we use that x_1 and x_2 are integral), which implies

$$\begin{aligned}\mu_{\mathcal{C}}([(P_1) - (P_2)]) &= \lim_{n\to\infty} 4^{-n} v(\underline{\delta}_{\mathcal{C}}^{\circ n}(\underline{\xi})) - v(\underline{\xi}) \\ &\le \lim_{n\to\infty} 4^{-n} v(\underline{\delta}_{\mathcal{E}}^{\circ n}(\xi_4, \xi_1)) - \min\{v(\xi_1), v(\xi_4)\} \\ &= \mu_{\mathcal{E}}(Q_1 - Q_2). \qquad \square\end{aligned}$$

The following consequence is useful for practical purposes. For simplicity, we state it for the case of residue characteristic $\neq 2$, but we expect that the statement remains true for residue characteristic 2.

Theorem 10.13. *Suppose that the residue characteristic of k is not 2. Let $\mathcal{C}$ be a stably minimal Weierstrass model of C such that C has reduction type $[\mathcal{K}_1 - \mathcal{K}_2 - l]$. Then*

$$\beta(\mathcal{C}) \le \beta(\mathcal{K}_1) + \beta(\mathcal{K}_2) + 2l,$$

where $\beta(\mathcal{K})$ denotes the maximum of μ for an elliptic curve of reduction type $\mathcal{K}$, taking the action of Frobenius into account. (See Table 1 in [Cremona et al. 2006] for the values of $\beta(\mathcal{K})$.)

Proof. We may assume that the point(s) of multiplicity 3 on the special fiber are defined over $\mathfrak{k}$, at the cost of an at most quadratic unramified extension of k. Then we can move these points to have x-coordinates 0 and ∞, respectively, and so we can assume that our model $\mathcal{C}$ is as in Lemma 10.1. Let $P \in J(k)$; we write $P = [(P_1) - (P_2)]$ with points $P_1, P_2 \in C(k')$ for a finite extension k' of k such that the reduction of C over k' is semistable. We can find $\mathcal{C}_0$, $\mathcal{C} = \mathcal{C}_j$ and $\mathcal{C}_l$ as vertices in the reduction graph of the minimal proper regular model of C over k'. Then the part of the graph to the left of $\mathcal{C}_0$ corresponds to the reduction graph of $\mathcal{E}_1$ over k', in the sense that we consider a semistable model that dominates $\mathcal{E}_1$ (and is minimal with that property); the graph then is either a line segment (potentially good reduction) or a line segment joined to a circle (potentially multiplicative reduction), with $\mathcal{E}_1$ corresponding to the end of the line segment joined to the remaining graph of $\mathcal{C}$. Similarly, the part of the graph to the right of $\mathcal{C}_l$ corresponds to the reduction graph of $\mathcal{E}_2$ over k'.

Now assume that both P_1 and P_2 map (strictly) to the left of $\mathcal{C}_0$ in the reduction graph. This means that the x-coordinates of the points have positive valuation. We can then find points P_1' and P_2' in $\mathcal{E}_1(k')$ with the same x-coordinates as P_1 and P_2 and nearby y-coordinates. Then $P_1' - P_2'$ is in $\mathcal{E}_1(k)$ and P_1' and P_2' have the same images as P_1 and P_2 in the reduction graph. By our previous results for the semistable

case, the value of (or at least the upper bound given in Proposition 10.5 for) $\mu_0(P)$ depends only on the part of the graph to the left of $\mathcal{C}_0$. We can therefore let l tend to infinity; then Lemma 10.12 and the discussion preceding Lemma 10.3 show that $\mu_0(P)$ is bounded by the value of $\mu_{\mathcal{E}_1}$ on the difference $P_1' - P_2'$. By the arguments in the proof of Proposition 10.5, we have

$$\mu_{\mathcal{C}}(P) = \mu_j(P) \le \mu_0(P) + 2j \le \beta(\mathcal{K}_1) + 2l.$$

The case that P_1 and P_2 both map to the right of $\mathcal{C}_l$ is similar.

If (say) P_1 maps to the left of $\mathcal{C}_0$ and P_2 maps to the right of $\mathcal{C}_0$, but not to the right of $\mathcal{C}_l$, then by the formula of Proposition 10.5, we can bound $\mu_{\mathcal{C}}(P)$ by $\mu_1 + 2l$, where μ_1 comes from the part of the graph between P_1 and $\mathcal{C}_0$. By an argument similar to the one used in the previous paragraph, μ_1 can be bounded by $\mu_{\mathcal{E}_1}(P_1')$, where P_1' is the point on $\mathcal{E}_1$ corresponding to P_1 and we take the second point to be on the component visible in $\mathcal{C}_0$. If P_2 maps to the right of $\mathcal{C}_l$, then we similarly obtain a bound of the form $\mu_1 + \mu_2 + 2l \le \beta(\mathcal{K}_1) + \beta(\mathcal{K}_2) + 2l$. The remaining cases are similar or follow directly from Proposition 10.5. □

The example in Section 19 demonstrates the effect of the improved bounds on β as given in the preceding section. For other examples the bounds established in this section will be similarly useful.

11. General upper and lower bounds for $\bar{\beta}$

In this section we derive an upper bound for the geometric height constant $\bar{\beta}(\mathcal{C})$ in the general case by reducing to the semistable situation. We also give a lower bound of the same order of magnitude. We note the following consequence of the results obtained so far; see the discussion at the end of Section 9 and Corollary 10.11.

Corollary 11.1. *Assume that $\mathcal{C}$ is a stably minimal Weierstrass model of C over k and that the minimal proper regular model $\mathcal{C}^{\min}$ of C over k has semistable reduction. Denoting the discriminant of $\mathcal{C}$ by Δ and writing $\bar{\beta}(\mathcal{C}) = \max\{\mu_{\mathcal{C}}(P) : P \in J(\bar{k})\}$, where $\mu_{\mathcal{C}}$ denotes μ with respect to the model $\mathcal{C}$ and J is the Jacobian of C, we have*

$$\tfrac{1}{6}v(\Delta) \le \bar{\beta}(\mathcal{C}) \le \tfrac{1}{4}v(\Delta).$$

When $\mathcal{C}^{\min}$ does not have semistable reduction, the idea is to pass to a suitable field extension k'/k and apply Corollary 11.1 over k'. In order to compare the corresponding geometric height constants $\bar{\beta}$, we need to analyze how μ changes under minimization. We first prove the following key lemma:

Lemma 11.2. *There exists a transformation $\tau : \mathcal{C} \to \mathcal{C}'$, defined over k, such that $\mathcal{C}'$ is a minimal Weierstrass model and*

$$v(\tau(x)) + v(\tau) \le v(x) \quad \text{for all } x \in \mathrm{KS}_{\mathbb{A}}.$$

Proof. If $\mathcal{C}$ is already minimal, then there is nothing to prove. Otherwise, [Liu 1996, Remarque 11] implies that we can compute a minimal Weierstrass model by going through the following steps for finitely many points P on the special fiber of $\mathcal{C}$.

(a) Move P to $(0, 0)$.

(b) Scale x by $1/\pi$.

(c) Replace $\mathcal{C}$ by the normalization of the resulting model.

As transformations of the form (a) do not change $v(x)$ and have determinant of valuation 0, it suffices to prove

$$v(\tau(x)) + v(\tau) \le v(x) \quad \text{for all } x \in \mathrm{KS}_{\mathbb{A}}$$

for a transformation $\tau = \sigma \circ \rho$, where ρ is as in (b) and σ is as in (c). Note that such a transformation decreases the valuation of the discriminant; cf. [Liu 1996, Lemme 9, Corollaire 2]. By the discussion following Proposition 4.4, the transformation ρ maps $x \in \mathrm{KS}_{\mathbb{A}}$ to $(\pi x_1, x_2, \pi^{-1}x_3, \pi^3 x_4)$.

Suppose $v(2) = 0$ and, without loss of generality, $H = 0$. According to [Liu 1996, Remarque 2], the normalization can be computed using the transformation σ mapping an affine point (ξ, η) to $\sigma(\xi, \eta) = (\xi, \eta\pi^{-s})$ for some nonnegative integer s. As $v(\tau) = 3-2s$, we must have $s \ge 2$, since otherwise τ would increase the valuation of the discriminant. Because $\tau(x) = (\pi x_1, x_2, \pi^{-1}x_3, \pi^{3-2s}x_4)$ for $x \in \mathrm{KS}_{\mathbb{A}}$, we find that $v(\tau(x)) \le v(x) + 1$, implying

$$v(\tau(x)) + v(\tau) - v(x) \le -2s + 4 \le 0.$$

The case $v(2) > 0$ is slightly more complicated. Here one computes the normalization by repeatedly applying transformations

$$(\xi, \eta) \mapsto \left(\xi, \frac{\eta + R(\xi, 1)}{\pi}\right), \tag{11-1}$$

where $R \in \mathcal{O}[X, Z]$ is a certain cubic form, until the minimum of the valuations of the coefficients of $F + RH - R^2$ is equal to 1. See [Liu 1996, Remarque 2]. Such a transformation maps Kummer coordinates $x = (x_1, x_2, x_3, x_4)$ to

$$(x_1, x_2, x_3, \pi^{-2}x_4 + l_1x_1 + l_2x_2 + l_3x_3)$$

and the expressions for the l_i given in Section 4 show that $v(l_i) \ge -2$ for all i. As the determinant of a transformation (11-1) has valuation -2, we need to apply at least two such transformations, because otherwise the valuation of the discriminant would increase. In other words, $\sigma = \sigma_s \circ \cdots \circ \sigma_1$, where $s \ge 2$ and every σ_i is of the form (11-1).

By the properties of the transformations (11-1), it suffices to show the desired inequality for the case $s = 2$, since further applications of transformations σ_i will

only make the left-hand side of the desired inequality smaller and will not change the right-hand side. So suppose that $\sigma = \sigma_2 \circ \sigma_1$; then $\tau = \sigma \circ \rho$ maps $x \in \mathrm{KS}_{\mathbb{A}}$ to

$$\tau(x) = \big(\pi x_1,\ x_2,\ \pi^{-1}x_3,\ \pi^{-1}x_4 + \pi l_1 x_1 + \pi l_2 x_2 + \pi l_3 x_3 + \pi l_1' x_1 + l_2' x_2 + \pi^{-1} l_3' x_3\big),$$

where the l_i arise from σ_1 and the l_i' arise from σ_2. As $v(\tau) = -1$, it clearly suffices to prove that

$$v(\tau(x)) \le v(x) + 1. \tag{11-2}$$

But if (11-2) is false, then $v(x) = v(x_4) < \min\{v(x_1), v(x_2)+1, v(x_3)+2\}$. In this situation it follows from the lower bounds $v(l_i) \ge -2$ and $v(l_i') \ge -2$ that we get

$$v\big(\pi l_1 x_1 + \pi l_2 x_2 + \pi l_3 x_3 + \pi l_1' x_1 + l_2' x_2 + \pi^{-1} l_3' x_3\big) > v(x_4) - 1.$$

This implies (11-2) and therefore finishes the proof of the lemma. □

Theorem 11.3. *Let C be a smooth projective curve of genus 2 defined over a nonarchimedean local field k, given by an integral Weierstrass model $\mathcal{C}$. Then we have*

$$\bar{\beta}(\mathcal{C}) \le \tfrac{1}{4} v(\Delta(\mathcal{C})).$$

Proof. By Lemma 5.4 there is a finite extension k'/k such that the minimal proper regular model of C over k' is semistable and such that all minimal Weierstrass models of C over k' are stably minimal. By Corollary 11.1, the claim therefore holds for any minimal Weierstrass model of C over k'.

It follows from Lemma 11.2 that there is a transformation $\tau : \mathcal{C} \to \mathcal{C}'$ defined over k' such that $\mathcal{C}'$ is a minimal (and hence stably minimal) Weierstrass model over k' and such that

$$v(\tau(x)) + v(\tau) \le v(x) \tag{11-3}$$

for all $x \in \mathrm{KS}_{\mathbb{A}}$.

Then, by the above, we have

$$\mu(\tau(x)) \le \tfrac{1}{4} v(\Delta(\mathcal{C}')).$$

Now using Corollary 4.6 and the relation (4-2), we find

$$\begin{aligned}
\mu(x) &= \mu(\tau(x)) - v(x) + v(\tau(x)) - v(\tau) \\
&\le \tfrac{1}{4} v(\Delta(\mathcal{C}')) - v(x) + v(\tau(x)) - v(\tau) \\
&= \tfrac{1}{4} v(\Delta(\mathcal{C})) - v(x) + v(\tau(x)) + \tfrac{3}{2} v(\tau) \\
&\le \tfrac{1}{4} v(\Delta(\mathcal{C})),
\end{aligned}$$

where we have used (11-3) and $v(\tau) \le 0$. □

Remark 11.4. When the residue characteristic is not 2, then we can easily show that $\bar{\beta}(\mathcal{C})$ is indeed always comparable to $v(\Delta(\mathcal{C}))$. We can assume that $H = 0$ and write $F = cF_0$ with F_0 primitive. We consider the points of order 2 on J. Such a point P is given by a factorization $F_0 = G_1G_2$ with G_1 and G_2 primitive of degrees 2 and 4, respectively. An explicit computation shows that

$$\varepsilon(P) = 4v(c) + 2v(R(P)),$$

where $R(P)$ denotes the resultant of G_1 and G_2, and we have $4\mu(P) = \varepsilon(P)$. Since $v(\Delta(\mathcal{C})) = v(\operatorname{disc}(F)) = 10v(c) + v(\operatorname{disc}(F_0))$ and $4v(\operatorname{disc}(F_0))$ is the sum of the valuations of the 15 resultants $R(P)$, we find that

$$\begin{aligned}\bar{\beta}(\mathcal{C}) &\geq \tfrac{1}{4} \max_{O \neq P \in J[2]} (4v(c) + 2v(R(P))) \geq v(c) + \tfrac{1}{30} \sum_{O \neq P \in J[2]} v(R(P)) \\ &= v(c) + \tfrac{2}{15} v(\operatorname{disc}(F_0)) \geq \tfrac{1}{10} v(\Delta(\mathcal{C})).\end{aligned}$$

A similar statement should be true when the residue characteristic is 2.

Recall that we denote $\max\{\varepsilon(P) : P \in J(\bar{k})\}$ by $\bar{\gamma}(\mathcal{C})$.

Corollary 11.5. *Let C be a smooth projective curve of genus 2 defined over a nonarchimedean local field k, given by an integral Weierstrass model $\mathcal{C}$. Then we have*

$$\bar{\gamma}(\mathcal{C}) \leq v(\Delta(\mathcal{C})).$$

If $H = 0$ and $\operatorname{char}(k) \neq 2$, then this can be improved to

$$\bar{\gamma}(\mathcal{C}) \leq v(2^{-4}\Delta(\mathcal{C})).$$

Proof. The first inequality follows from $\varepsilon(P) = 4\mu(P) - \mu(2P)$ and Theorem 11.3. The second inequality is Theorem 6.1 of [Stoll 1999]. □

Question 11.6. If $\mathcal{C}$ is a minimal Weierstrass model, does $\bar{\beta}(\mathcal{C})$ only depend on the special fiber of $\mathcal{C}^{\min}$?

Note that the corresponding statement holds for elliptic curves [Cremona et al. 2006]. In our situation, however, there may be several nonisomorphic minimal Weierstrass models, which complicates the picture.

Part III. Efficient computation of canonical heights

In this part we show how to compute the canonical height $\hat{h}(P)$ efficiently for a point P over a number field, global function field or more general field with a system of absolute values as in Section 2. We first explain how to compute the local height correction functions. We use $\mathsf{M}(d)$ to denote the time needed to multiply two d-bit integers.

12. Computing μ at nonarchimedean places

In this section, k is a nonarchimedean local field again, with valuation ring $\mathcal{O}$, uniformizer π, normalized valuation v and residue class field $\mathfrak{k}$. Let $\mathcal{C}$ be an integral Weierstrass model for a genus-2 curve C over k. We make no assumptions on the reduction type of C. We already discussed a method for the computation of $\mu(P)$ for a given point $P \in J(k)$ in Section 3. In this section, we provide an alternative fast algorithm and show that its running time is

$$\ll (\log v(\Delta))\, \mathsf{M}\big((\log v(\Delta))v(\Delta)(\log \#\mathfrak{k})\big),$$

where $\Delta = \Delta(\mathcal{C})$.

Lemma 12.1. *Assume that M is a positive integer such that $M\mu(P) \in \mathbb{Z}$. Further assume that* $\max\{\varepsilon(P) : P \in J(k)\} \le B$. *Then*

$$\mu(P) = \frac{1}{M}\left\lceil M \sum_{n=0}^{\lfloor \log(\frac{1}{3}BM)/\log 4 \rfloor} 4^{-n-1}\varepsilon(2^n P)\right\rceil.$$

Proof. This follows from $M\mu(P) \in \mathbb{Z}$ and from

$$0 \le M \sum_{n \ge m} 4^{-n-1}\varepsilon(2^n P) \le \frac{BM}{3 \cdot 4^m}. \qquad \square$$

If we know that the reduction is nodal, then we get an upper bound B for $\varepsilon(P)$ and all possible denominators of $\mu(P)$ from the results of Section 9. More generally, if we know the smallest positive period N of the sequence $(\mu(nP))_n$, then we can take $M = N$ (respectively, $M = 2N$) if N is odd (respectively, even) by Corollary 3.11. Also note that we can always take $B = v(\Delta)$ (or even $B = v(2^{-4}\Delta)$ if $\mathrm{char}(k) \ne 2$ and the equation of the curve has $H = 0$); see Corollary 11.5.

If we only know an upper bound for the denominator of $\mu(P)$, then the following alternative approach can be used. This is analogous to [Müller and Stoll 2016, Lemma 4.2].

Lemma 12.2. *Assume that $M \ge 2$ is an integer such that $M'\mu(P) \in \mathbb{Z}$ for some $0 < M' \le M$. Assume in addition that* $\max\{\varepsilon(P) : P \in J(k)\} \le B$, *and set*

$$m = \left\lfloor \frac{\log(\frac{1}{3}BM^2)}{\log 4} \right\rfloor.$$

Then $\mu(P)$ is the unique fraction with denominator less than or equal to M in the interval $[\mu_0, \mu_0 + 1/M^2]$, where

$$\mu_0 = \sum_{n=0}^{m} 4^{-n-1}\varepsilon(2^n P).$$

Proof. Note that

$$\mu_0 \le \mu(P) \le \mu_0 + \sum_{n>m} 4^{-n-1} B < \mu_0 + \frac{1}{M^2}.$$

But since $M \ge 2$, the interval $[\mu_0, \mu_0 + 1/M^2]$ contains at most one fraction with denominator bounded by M; by assumption, $\mu(P)$ is such a fraction. □

In order to apply Lemma 12.2, we now find a general upper bound M on the possible denominators of μ. Let $\mathcal{J}$ denote the Néron model of J over $S = \mathrm{Spec}(\mathcal{O})$ and write Φ for the component group of $\mathcal{J}$.

Proposition 12.3. *Let N denote the exponent of $\Phi(\bar{\mathfrak{k}})$ and let $P \in J(k)$. Then we have*

$$\mu(P) \in \frac{1}{2N}\mathbb{Z}.$$

If N is odd or if C has a k^{nr}-rational Weierstrass point, then we have

$$\mu(P) \in \frac{1}{N}\mathbb{Z}.$$

Proof. Let $i \in \{1, \dots, 4\}$ be such that $\kappa_i(P) \ne 0$. Recall from Lemma 8.2 that the function $\hat{\lambda}_i = \hat{\lambda} \circ (\kappa/\kappa_i)$ is a Néron function with respect to the divisor D_i. As $P \notin \mathrm{supp}\, D_i$, we find

$$\mu(P) \equiv \hat{\lambda}(x) \equiv \hat{\lambda}_i(P) \pmod{\mathbb{Z}}$$

for any set of Kummer coordinates x for P. It follows from the results of [Néron 1965] and [Lang 1983, §11.5] that

$$\hat{\lambda}_i(P) \equiv j(D_i, (P) - (O)) \pmod{\mathbb{Z}},$$

where $j(\cdot, \cdot)$ denotes Néron's bilinear j-pairing, defined in [Néron 1965, §III.3].

By [Néron 1965, Proposition III.2], the values of the j-pairing lie in $\frac{1}{2N'}\mathbb{Z}$, where $N' = \#\Phi(\bar{\mathfrak{k}})$. It is easy to see that we can replace N' by the exponent N in the proof of [Néron 1965, Proposition III.2], so the first statement of the proposition follows.

For the second statement, note that the j-pairing takes values in $\frac{1}{N}\mathbb{Z}$ if N is odd, again by [Néron 1965, Proposition III.2] and its proof. If C has a k^{nr}-rational Weierstrass point P_0, then the divisor D_i is linearly equivalent over k^{nr} to $2\Theta_{P_0}$, where Θ_{P_0} is the theta divisor with respect to P_0. The Néron model does not change under unramified extensions, and $\mu(P) \bmod \mathbb{Z}$ does not depend on the Weierstrass model of C by Corollary 4.6. Hence we can assume that $i = 1$ and $D_1 = 2\Theta_{P_0}$, so the linearity of the j-pairing in the first variable proves the claim. □

Remark 12.4. In the notation of [Namikawa and Ueno 1973], the only reduction types for which Proposition 12.3 does not show that $\mu(P) \in \frac{1}{N}\mathbb{Z}$ (where N is the

exponent of $\Phi(\bar{\mathfrak{k}})$), are $[2\mathrm{III} - l]$ and $[2\mathrm{III}^* - l]$ for $l \geq 0$; $[2\mathrm{I}_n^* - l]$ for $n, l \geq 0$; and $[2\mathrm{I}_n - l]$ for $n > 0$ even and $l \geq 0$. We have not found an example where $\mu(P) \notin \frac{1}{N}\mathbb{Z}$.

We can compute the group $\Phi(\bar{\mathfrak{k}})$ in practice using [Bosch et al. 1990, §9.6]. For this we need to know the intersection matrix of the special fiber of a regular model of C over S. This is implemented in Magma, but can be rather slow. If the residue characteristic is not 2, then we can apply Liu's algorithm [1994] to compute the reduction type and read off $\Phi(\bar{\mathfrak{k}})$.

In general, an upper bound for the exponent of $\Phi(\bar{\mathfrak{k}})$ suffices to apply Lemma 12.2. We give a bound which only depends on the valuation of the discriminant $\Delta = \Delta(\mathcal{C})$.

Lemma 12.5. *The exponent of $\Phi(\bar{\mathfrak{k}})$ is bounded from above by*

$$M := \max\left\{2, \left\lfloor \tfrac{1}{3} v(\Delta)^2 \right\rfloor\right\}.$$

Moreover, the denominator of $\mu(P)$ is bounded from above by M for all $P \in J(k)$.

Proof. This follows from a case-by-case analysis, using the list of groups $\Phi(\bar{\mathfrak{k}})$ from [Liu 1994, §8] for all reduction types in [Namikawa and Ueno 1973], and Proposition 12.3. □

Remark 12.6. By going through all reduction types, it is possible to obtain better upper bounds for the denominator M' of $\mu(P)$ from the Igusa invariants discussed in Section 6. First note that if the special fiber of $\mathcal{C}$ is nonreduced, then we have

(i) $M' \leq 4$ if $v(\Delta) \leq 12$,

(ii) $M' \leq \max\{12, v(\Delta) - 15\}$ otherwise.

Suppose that $\mathcal{C}$ is reduced; then, by Proposition 6.2, we can use the Igusa invariants of the special fiber to distinguish between the multiplicities of its singularities.

(i) If all points on the special fiber of $\mathcal{C}$ have multiplicity at most 2, then we can bound M' using Proposition 6.3(i)–(iii) and Propositions 9.1, 9.3, and 9.4.

(ii) If there is a point of multiplicity 3 on the special fiber, then we have
- $M' \leq \min\{6, v(\Delta) + 1\}$ if $v(\Delta) \leq 10$,
- $M' \leq 12$ if $v(\Delta) \leq 20$,
- $M' \leq \left\lfloor \frac{1}{4}(v(\Delta) - 12)^2 \right\rfloor$ otherwise.

(iii) If there is a point of multiplicity ≥ 4 on the special fiber, then we have
- $M' \leq 3v(\Delta) - 10$ if $v(\Delta) \leq 10$,
- $M' \leq 4v(\Delta) - 20$ if $v(\Delta) > 10$ and the model is minimal,
- $M' \leq \left\lfloor \frac{1}{3}(v(\Delta) - 10)^2 \right\rfloor$ if the model is not minimal.

The results of this section lead to an efficient algorithm for the computation of $\mu(P)$, which is analogous to Algorithm 4.4 of [Müller and Stoll 2016]. We

assume that the coefficients of F and H and the coordinates of P are given to sufficient v-adic precision (in practice, they will be given exactly as elements of a number field or function field).

1. If $\mathrm{char}(k) \neq 2$ and $H = 0$, set $B := v(2^{-4}\Delta)$. Otherwise, set $B := v(\Delta)$.
2. Set $M := \max\{2, \lfloor \frac{1}{3} v(\Delta)^2 \rfloor\}$.
3. Set $m := \lfloor \log(\frac{1}{3} B M^2) / \log 4 \rfloor$.
4. Set $\mu_0 := 0$. Let x be normalized Kummer coordinates for P with $(m+1)B+1$ v-adic digits of precision.
5. For $n := 0$ to m do:
   a. Compute $x' := \delta(x)$ (to $(m+1)B+1$ v-adic digits of precision).
   b. If $v(x') = 0$, then return μ_0.
   c. Set $\mu_0 := \mu_0 + 4^{-n-1} v(x')$.
   d. Set $x := \pi^{-v(x')} x'$.
6. Return the unique fraction with denominator at most M in the interval between μ_0 and $\mu_0 + 1/M^2$.

The fraction in the final step can be computed easily, for instance, using continued fractions.

For the complexity analysis in the following proposition, we assume that elements of $\mathcal{O}$ are represented as truncated power series in π, whose coefficients are taken from a complete set of representatives for the residue classes. Operations on these coefficients can be performed in time $\ll \mathsf{M}(\log \#\mathfrak{k})$.

Proposition 12.7. *The algorithm above computes $\mu(P)$. Its running time is*

$$\ll (\log v(\Delta))\, \mathsf{M}\big((\log v(\Delta)) v(\Delta) (\log \#\mathfrak{k})\big)$$

as $v(\Delta) \to \infty$, with an absolute implied constant.

Proof. The following proof is analogous to the proof of [Müller and Stoll 2016, Proposition 4.5]. Corollary 11.5 shows that B is a suitable upper bound for ε and Lemma 12.5 shows that M is an upper bound for the denominator of μ. Because $M \geq 2$, the loop in step 5 computes the sum in Lemma 12.2. Note that when $v(x') = 0$ in step 5b, we have $\mu(P) = \mu_0$ by Theorem 3.10. At each duplication step, the precision loss is $\varepsilon(2^n P) \leq B$, so that with our choice of starting precision, after the $m+1$ steps in the loop the resulting x still has at least one digit of precision. This proves the correctness of the algorithm.

Clearly the running time of the algorithm is dominated by the running time of the loop in step 5. Step 5a consists of a fixed number of additions and multiplications of elements of $\mathcal{O}$ which are given to a precision of $(m+1)B+1$ digits.

Because steps 5b–5d take negligible time compared to step 5a, each pass through the loop takes

$$\ll \mathsf{M}\big(((m+1)B+1)(\log \#\mathfrak{k})\big)$$

operations, leading to a total running time that is

$$\begin{aligned}&\ll (m+1)\,\mathsf{M}\big(((m+1)B+1)(\log \#\mathfrak{k})\big)\\ &\ll m\,\mathsf{M}(mB(\log \#\mathfrak{k}))\\ &\ll (\log v(\Delta))\,\mathsf{M}\big((\log v(\Delta))v(\Delta)(\log \#\mathfrak{k})\big)\end{aligned}$$

as $v(\Delta)\to\infty$. Here we use that $B \ll v(\Delta)$ and $M \ll v(\Delta)^2$, so that $m \ll \log v(\Delta)$. □

Remark 12.8. In step 2, we can use Remark 12.6 to compute a sharper upper bound for the denominator of μ. See also the discussion following Remark 12.4. Of course, if we want to find $\mu(P)$ for several points P, the quantities M, B and m only have to be computed once.

Remark 12.9. We can compute $\mu(P)$ using the algorithm above in more general situations. Suppose that k is any discretely valued field with valuation ring $\mathcal{O}$ and uniformizer π. In that case, the sequence $(\mu(nP))_n$ might not have a finite period, so the method for the computation of $\mu(P)$ discussed in Section 3 might not be applicable. However, Lemmas 12.1, 12.2 and 12.5 and Proposition 12.3 remain valid. If $\mathrm{char}(k)\neq 2$ and if $H=0$, then we have the upper bound $\varepsilon(P)\le v(2^{-4}\Delta)$ (cf. Remark 3.2), so the algorithm above can be used and Proposition 12.7 remains valid as well, in the sense that the computation can be done using $\ll \log v(\Delta)$ operations with elements of $\mathcal{O}/\pi^n\mathcal{O}$, where $n \ll v(\Delta)\log v(\Delta)$. In the remaining cases, we can compute an upper bound B on ε as in Remark 3.2, and we can apply the algorithm with this choice of B.

13. Computing μ at archimedean places

In this section, k is an archimedean local field, so $k=\mathbb{R}$ or $k=\mathbb{C}$. We assume that the curve C is given by a Weierstrass equation $\mathcal{C}$ with $H=0$. In the following, $\log_+ x = \max\{0,\log x\}$.

Let $x\in k^4$ be a set of Kummer coordinates. Recall that

$$\tilde{\varepsilon}(x) = -[k:\mathbb{R}](\log\|\delta(x)\|_\infty - 4\log\|x\|_\infty)$$

and

$$\tilde{\mu}(x) = \sum_{n=0}^{\infty} 4^{-n-1}\tilde{\varepsilon}(\delta^{\circ n}(x)).$$

We easily obtain a lower bound for $\tilde{\varepsilon}$ using the standard estimate for $\|\delta(x)\|_\infty$. Since the coefficients of the duplication polynomials δ_j are universal polynomials

of degree at most 4 in the coefficients of F, this gives

$$-\tilde{\varepsilon} \ll 1 + \log_+ \|F\|_\infty,$$

where $\|F\|_\infty$ is the maximum norm of the coefficient vector of F. We recall that the method described in Section 7 of [Stoll 1999], leading to equation (7.1) there, provides an upper bound $\tilde{\gamma}$ for $\tilde{\varepsilon}$ that can be explicitly computed for any given Weierstrass equation $\mathcal{C}$ of the curve (provided $H = 0$). It is given by

$$\begin{aligned}\tilde{\gamma} &= \log \max_i \left(\sum\nolimits_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum\nolimits_{1\le j\le 4} |b_{\{S,S'\},j}|} \right)^2 \\ &\le \log 400 + 2 \log \max_{i,\{S,S'\}} |a_{i,\{S,S'\}}| + \log \max_{\{S,S'\},j} |b_{\{S,S'\},j}|\end{aligned}$$

with certain numbers $a_{i,\{S,S'\}}$, $b_{\{S,S'\},j}$, where $i, j \in \{1, 2, 3, 4\}$ and $\{S, S'\}$ runs through the ten partitions of the set of roots of F into two sets of three. Using the formulas in [Stoll 1999, §10] and Mignotte's bound (see, for example, [von zur Gathen and Gerhard 1999, Corollary 6.33]), we see that

$$\log \max_{\{S,S'\},j} |b_{\{S,S'\},j}| \ll 1 + \log_+ \|F\|_\infty$$

and

$$\log \max_{i,\{S,S'\}} |a_{i,\{S,S'\}}| \ll 1 + \log_+ \|F\|_\infty + \log_+ \max_{\{S,S'\}} |R(S, S')|^{-1},$$

where $R(S, S')$ is the resultant of the two factors G, G' of F corresponding to the partition of the roots. Using Mignotte's bound again, we find that

$$|R(S, S')|^{-1} = \frac{\sqrt{|\operatorname{disc} G|\,|\operatorname{disc} G'|}}{\sqrt{|\operatorname{disc} F|}} \ll \|F\|_\infty^2 |\Delta(\mathcal{C})|^{-1/2},$$

leading finally to the estimate

$$|\tilde{\varepsilon}| \ll 1 + \log_+ \|F\|_\infty + \log_+ |\Delta(\mathcal{C})|^{-1} =: s(F).$$

If $|\tilde{\varepsilon}(x)| \le \tilde{\eta}$ for all $x \in \mathrm{KS}_{\mathbb{A}}$, then we have

$$\left| \sum_{n\ge N} 4^{-n-1} \tilde{\varepsilon}(\delta^{\circ n}(x)) \right| \le \tfrac{1}{3}\tilde{\eta} 4^{-N},$$

so we need to sum the first

$$N = \left\lceil \frac{d}{2} + \frac{\log\left(\frac{1}{3}\tilde{\eta}\right)}{\log 4} \right\rceil \ll d + \log s(F)$$

terms to obtain an accuracy of 2^{-d}. Comparing the largest term in any of the δ_j and the lower bound on $\|\delta(x)\|_\infty$, we obtain a bound $\tilde{\theta}$ on the loss of relative precision

(in terms of bits) in the computation of $\delta(x)$; we have $\tilde{\theta} \ll s(F)$. To achieve the desired precision at the end, we therefore need to compute with an initial precision of

$$d + 1 + N\tilde{\theta} \ll (d + \log s(F))s(F)$$

bits. The time needed for each duplication is then

$$\ll \mathsf{M}((d + \log s(F))s(F)).$$

A logarithm can be computed to d bits of precision in time $\ll (\log d)\,\mathsf{M}(d)$ by one of several quadratically converging algorithms (see, for example, [Borwein and Borwein 1987, Chapter 7]), so we obtain the following result.

Proposition 13.1. *Given Kummer coordinates x of a point P in $J(k)$ (or $\mathrm{KS}(k)$) to sufficient precision, we can compute $\tilde{\mu}(P)$ to an accuracy of d bits in time*

$$\ll (d + \log s(F))(\log d)\,\mathsf{M}((d + \log s(F))s(F)),$$

where

$$s(F) = 1 + \log_+ \|F\|_\infty + \log_+ |\Delta(\mathcal{C})|^{-1}.$$

In the applications, k will be the completion of a number field at a real or complex place. If the number field is $\mathbb{Q}$ and the given equation $\mathcal{C}$ of C is integral, then $|\Delta(\mathcal{C})| \geq 1$ and we have $s(F) = 1 + \log \|F\|_\infty = 1 + h(F)$, where $h(F)$ denotes the (logarithmic) height of the coefficient vector of F as a point in affine space. In general, we have the estimate (denoting the value of $s(F)$ for a place v by $s_v(F)$)

$$\begin{aligned}\sum_{v|\infty} s_v(F) &\leq [K : \mathbb{Q}] + \sum_{v|\infty} \log_+ \|F\|_v + \sum_{v|\infty} \log_+ |\Delta(\mathcal{C})|_v^{-1} \\ &\leq [K : \mathbb{Q}] + h(F) + h(\Delta(\mathcal{C})) \ll h(F)\end{aligned}$$

for $h(F)$ large. This implies that we can compute the infinite part of the height correction function in time

$$\ll (d + \log h(F))(\log d)\,\mathsf{M}((d + \log h(F))h(F)),$$

which is polynomial in d and $h(F)$.

14. Computing the canonical height of rational points

The first algorithm for computing the canonical height on a genus-2 Jacobian over $\mathbb{Q}$ was introduced by Flynn and Smart [1997]. It does not require any integer factorization, but can be impractical even for simple examples; see the discussion in [Stoll 2002, §1]. A more practical algorithm was introduced in [Stoll 2002]; here the local height correction functions are computed separately, so some integer factorization is required. Uchida [2011] later introduced a similar algorithm. De Jong and Müller [2014] used division polynomials for a different approach.

Building on the Arakelov-theoretic Hodge index theorem for arithmetic surfaces due to Faltings and Hriljac, Holmes [2012] and Müller [2014] independently developed algorithms for the computation of canonical heights of points on Jacobians of hyperelliptic curves of arbitrary genus over global fields. While these algorithms can be used to compute canonical heights for genus as large as 10 (see [Müller 2014, Example 6.2]), they are much slower than the algorithm from [Stoll 2002] when the genus is 2.

In this section we now combine the results of Sections 12 and 13 into an efficient algorithm for computing the canonical height of a point on the Jacobian of a curve of genus 2 over a global field K.

If K is a function field, then there are no archimedean places and factorization is reasonably cheap. So in this case, the best approach seems to be to first find the places v of K such that $\mu_v(P)$ is possibly nonzero (this includes the places at which the given equation of the curve is nonintegral) and then compute the corrections $\mu_v(P)$ for each place separately as in the algorithm of Proposition 12.7, if necessary changing first to an integral model and correcting for the transformation afterwards. In fact this approach can be used whenever K is a field with a set of absolute values that satisfy the product formula, because the algorithm before Proposition 12.7 is applicable over any discretely valued field; see Remark 12.9. This includes function fields such as $\mathbb{Q}(t)$ and $\mathbb{C}(t)$.

If K is a number field, then we compute the contribution from the archimedean places as described in Section 13. The finite part of our algorithm is analogous to our quasilinear algorithm for the computation of the finite part of the canonical height of a point on an elliptic curve in [Müller and Stoll 2016]; see Proposition 14.3 below. For simplicity, we take K to be $\mathbb{Q}$ in the following. We write ε_p and μ_p for the local height correction functions over $\mathbb{Q}_p$ as given by Definition 3.1 and $\tilde{\mu}_\infty$ for the local height correction function over $\mathbb{R}$ as defined in equation (1-1).

We assume that our curve is given by a model $\mathcal{C}\colon Y^2 = F(X, Z)$ with $F \in \mathbb{Z}[X, Z]$, and we set $\Delta = \Delta(\mathcal{C})$. Our goal is to devise an algorithm for the computation of $\hat{h}(P)$ that runs in time polynomial in $\log \|F\|_\infty$, $h(P)$ and the required precision d (measured in bits after the binary dot). We note that $h(P)$ can be computed in time

$$\ll \log(h(P) + d)\, \mathsf{M}(h(P) + d),$$

since it is just a logarithm. By Proposition 13.1, the height correction function $\tilde{\mu}_\infty(P)$ can be computed in polynomial time. So we only have to find an efficient algorithm for the computation of the "finite part" $\tilde{\mu}^{\mathrm{f}}(P) := \sum_p \mu_p(P) \log p$ of the height correction.

Fix $P \in J(\mathbb{Q})$. We call a set x of Kummer coordinates for P *primitive* if $x \in \mathbb{Z}^4$ and $\gcd(x) = 1$. We set $g_n = \gcd(\delta(x^{(n)}))$, where $x^{(n)}$ is a primitive set of Kummer

coordinates for $2^n P$. Then

$$\tilde{\mu}^{\mathrm{f}}(P) = \sum_{n=0}^{\infty} 4^{-n-1} \log g_n.$$

We also know by [Stoll 1999] that g_n divides $D = \frac{1}{2^4}|\Delta| = 2^4|\mathrm{disc}(F)|$, which implies that $\log g_n \le \log D$ for all n. To achieve a precision of 2^{-d}, it is therefore enough to take the sum up to

$$n = m := \left\lfloor \tfrac{1}{2}d + \log\left(\tfrac{1}{3}\log D\right)\right\rfloor \ll d + \log\log D \ll d + \log\log \|F\|_\infty.$$

Since at each duplication step we have to divide by g_n to obtain primitive coordinates again, it suffices to do the computation modulo D^{m+2}. This leads to the following algorithm.

1. Let $D = \frac{1}{16}|\Delta|$ and set $m := \left\lfloor \frac{1}{2}d + \log\log D - \log 3 \right\rfloor$.
2. Let x be primitive Kummer coordinates for P.
3. Set $\mu := 0$.
4. For $n := 0$ to m do:
   a. Compute $x' := \delta(x) \bmod D^{m+2}$.
   b. Set $g_n := \gcd(D, \gcd(x'))$ and $x := x'/g_n$.
   c. Set $\mu := \mu + 4^{-n-1}\log g_n$ (to d bits of precision).
5. Return $\tilde{\mu}^{\mathrm{f}}(P) \approx \mu$.

Proposition 14.1. *This algorithm computes $\tilde{\mu}^{\mathrm{f}}(P)$ to d bits of precision in time*

$$\ll (d + \log\log D)\log(d + \log\log D)\,\mathsf{M}((d + \log\log D)\log D) + h(P).$$

Proof. The discussion preceding the algorithm shows that it is correct. The duplication in step 4a can be computed in time

$$\ll \mathsf{M}((m+2)\log D) \ll \mathsf{M}((d + \log\log D)\log D),$$

while the gcd in step 4b can be computed in time

$$\begin{aligned} &\ll \mathsf{M}((m+2)\log D)\log((m+2)\log D) \\ &\ll \log(d + \log\log D)\,\mathsf{M}((d + \log\log D)\log D); \end{aligned}$$

the division is even faster, since g_n is small. The computation of the logarithm takes time $\ll \log(d + \log D)\,\mathsf{M}(d + \log D)$; this is dominated by the time for computing the gcd. This gives a time complexity of

$$\ll (d + \log\log D)\log(d + \log\log D)\,\mathsf{M}((d + \log\log D)\log D) + h(P),$$

where the last term comes from processing the input x. □

Note that $\log D \ll \log \|F\|_\infty$, so this bound is similar to (and even better by a factor of $\log d$ than) the complexity for computing $\tilde{\mu}_\infty(P)$.

Remark 14.2. An alternative way to proceed is to compute

$$x' = \delta^{\circ(m+1)}(x) \bmod D^{m+2}$$

(without dividing out gcds in between) and then use $\mu = 4^{-m-1} \log \gcd(x')$. The advantage of the algorithm above is that we can actually work mod D^{m+2-n}, which makes the computation more efficient. The advantage of the alternative is that it can also be used when working over a number field with nontrivial class group (replacing $\log \gcd(x')$ by the logarithm of the ideal norm of the ideal generated by x'). The resulting complexity is similar, with the implied constant depending on the base field.

We now show that we can in fact do quite a bit better than this, by using the strategy already employed in [Müller and Stoll 2016]. Note that $\tilde{\mu}^{\mathrm{f}}(P)$ is a rational linear combination of logarithms of positive integers. We can compute such a representation exactly and efficiently by the following algorithm. We again assume that x is a set of primitive Kummer coordinates for P.

1. Set $x' := \delta(x)$, $g_0 := \gcd(x')$ and $x := x'/g_0$.
2. Set $D := \gcd(2^4 \operatorname{disc}(F), g_0^\infty)$ and $B := \lfloor \log D / \log 2 \rfloor$.
3. If $B \le 1$, return 0. Otherwise, set $M := \max\{2, \lfloor \frac{1}{3}(B+4)^2 \rfloor\}$ and $m := \lfloor \log(\frac{1}{3}B^3M^2)/\log 4 \rfloor$.
4. For $n := 1$ to m do:
   a. Compute $x' := \delta(x) \bmod D^{m+1}g_0$.
   b. Set $g_n := \gcd(D, \gcd(x'))$ and $x := x'/g_n$.
5. Using the algorithm in [Bernstein 2004] (or in [Bernstein 2005]), compute a sequence $(q_1, \dots, q_r)$ of pairwise coprime positive integers such that each g_n (for $n = 0, \dots, m$) is a product of powers of the q_i: $g_n = \prod_{i=1}^r q_i^{e_{i,n}}$.
6. For $i := 1$ to r do:
   a. Compute $a := \sum_{n=0}^m 4^{-n-1} e_{i,n}$.
   b. Let μ_i be the simplest fraction between a and $a + 1/(B^2M^2)$.
7. Return $\sum_{i=1}^r \mu_i \log q_i$ (a formal linear combination of logarithms).

Proposition 14.3. *The preceding algorithm computes $\tilde{\mu}^{\mathrm{f}}(P)$ in time*

$$\ll (\log\log D)^2\, \mathsf{M}((\log\log D)(\log D)) + \mathsf{M}(h(P))(\log h(P)).$$

Note that $D \le \frac{1}{16}|\Delta|$ and $\log D \ll \log \|F\|_\infty$.

Proof. If $B \le 1$ in step 3, then we either have $g_0 = 1$ and $\tilde{\mu}^{\mathrm{f}}(P) = 0$, or we have $D \in \{2, 3\}$. In the latter case, g_0 is a power of $p = 2$ or 3 and $v_p(\Delta) = 1$, which would imply that $\varepsilon_p(P) = 0$ by [Stoll 2002, Proposition 5.2], so $g_0 = 1$, and we get a contradiction.

If a prime p does not divide g_0, then $\varepsilon_p(P) = 0$, implying $\mu_p(P) = 0$. Suppose now that p divides g_0; then we have $v_p(D) \le B$ and $v_p(\Delta) \le B+4$, so B, M and m are suitable values for Lemma 12.2. We have $v_p(g_n) = \varepsilon_p(2^n P)$ for all $n \le m$, because $p^{(m+1)v_p(D)+1} \mid D^{m+1} g_0$ (compare the proof of Proposition 12.7). All the g_n are power products of the q_i, so there will be exactly one $i = i(p) \in \{1, \dots, r\}$ such that $p \mid q_{i(p)}$. Setting $b_p = v_p(q_{i(p)})$ and $a = \sum_{n=0}^{m} 4^{-n-1} e_{i(p),n}$, we have

$$\sum_{n=0}^{m} 4^{-n-1} \varepsilon_p(2^n P) = \sum_{n=0}^{m} 4^{-n-1} v_p(g_n) = b_p a,$$

implying

$$\mu_p(P) = \sum_{n=0}^{\infty} 4^{-n-1} \varepsilon_p(2^n P) = b_p a + \sum_{n=m+1}^{\infty} 4^{-n-1} \varepsilon_p(2^n P).$$

Here the last sum is in $[0, 1/(B^2 M^2)]$ by the definition of m (compare the proof of Lemma 12.2). Therefore

$$a \le \frac{\mu_p(P)}{b_p} \le a + \frac{1}{b_p B^2 M^2} \le a + \frac{1}{B^2 M^2}.$$

Since the denominator of $\mu_p(P)$ is at most M and since we have $b_p \le v_p(D) \le B$, the denominator of $\mu_p(P)/b_p$ is at most BM. Hence $\mu_p(P)/b_p$ is the unique fraction in $[a, a + 1/(B^2 M^2)]$ with denominator bounded by BM, so $\mu_p(P)/b_p = \mu_{i(p)}$ by step 6b. Now

$$\sum_p \mu_p(P) \log p = \sum_p \mu_{i(p)} b_p \log p = \sum_{i=1}^{r} \mu_i \sum_{p \mid q_i} b_p \log p = \sum_{i=1}^{r} \mu_i \log q_i,$$

so the algorithm is correct.

The complexity analysis is as in the proof of Proposition 6.1 in [Müller and Stoll 2016]. Namely, the computations in step 1 can be done in time $\ll \mathsf{M}(h(P)) \log h(P)$. The computations in steps 2 and 3 take negligible time. Each pass through the loop in step 4 takes time $\ll \log((m+2)\log D)\, \mathsf{M}((m+2)\log D)$, so the total time for step 4 is

$$\ll m\, \mathsf{M}(m \log D) \log(m \log D) \ll (\log\log D)^2\, \mathsf{M}((\log\log D)(\log D)),$$

because $m \ll \log\log D$. The coprime factorization algorithm in [Bernstein 2004] (or in [Bernstein 2005]) computes suitable q_i for a pair (a, b) of positive integers in time

$\ll (\log ab)(\log\log ab)^2$. We iterate this algorithm, applying it first to g_0 and g_1, then to each of the resulting q_i and g_2, and so on. There are always $\ll \log D$ terms in the sequence of the q_i and we have $g_n \le D$ for all n. Hence step 5 takes time $\ll \log D(\log\log D)^3$. Because this is dominated by the time for the loop and because the remaining steps take negligible time, the result follows. □

Note that the complexity of the algorithm above is quasilinear in $\log D$ and $h(P)$. In practice, the efficiency of this approach can be improved somewhat:

- We can split off the contributions of all sufficiently small primes p by choosing a suitable bound T and trial factoring Δ up to T; the corresponding μ_p can then be computed using the algorithm of Proposition 12.7; see also Remark 12.8. In step 3, we can then set $B := \lfloor \log D'/\log T \rfloor$, where D' is the unfactored part of D, and replace $B+4$ by B in the definition of M. If the coefficients of F are sufficiently large, then this trial division can become quite expensive (even for small values of T). So when $h(F)$ is large, it is usually preferable to avoid trial division altogether.

- We can update the q_i after each pass through the loop in step 4 using the new g_n; we can also do the computation in step 4a modulo suitable powers of the q_i instead of modulo $D^{m+1}g_0$. Moreover, it is possible to use separate values of B, M and m for each q_i; these will usually be smaller than those computed in steps 2 and 3. In this way, we can integrate steps 4, 5 and 6 into one loop.

Remark 14.4. Over a more general number field K in place of $\mathbb{Q}$, the algorithm as stated does not quite work, since we cannot always divide out greatest common divisors. In this case we first compute $x^{(1)} = \delta(x)$ and the ideal g_0 generated by D and the entries of $x^{(1)}$. Then we compute $x^{(2)} = \delta(x^{(1)}), \ldots, x^{(m+1)} = \delta(x^{(m)})$ modulo the ideal $D^{m+1}g_0$. Let G_j be the ideal generated by the entries of $x^{(j)}$ and D^{m+1} and set

$$g_1 = g_0^{-4}G_2, \quad g_2 = G_2^{-4}G_3, \quad g_3 = G_3^{-4}G_4, \quad \ldots \quad g_m = G_m^{-4}G_{m+1}.$$

The coprime factorization algorithms in [Bernstein 2004; 2005] also work for ideals. In the final result, $\log q_i$ has to be replaced by $\log N(q_i)$, where $N(q_i)$ is the norm of the ideal q_i. This should result in a complexity similar to that over $\mathbb{Q}$ (with the implied constant depending on K), or at least one that is dominated by the complexity of computing the naive height and the contributions from the archimedean places. Unfortunately, no complexity analysis for standard operations with ideals in number fields seems to be available in the literature; this prevents us from making a precise statement. Alternatively, we can take the approach described in Remark 14.2.

Combining this with the results for archimedean places, we obtain an efficient algorithm for computing the canonical height $\hat{h}(P)$ of a point $P \in J(\mathbb{Q})$. As

mentioned above, we expect a similar result to hold for any number field K in place of $\mathbb{Q}$, with the implied constant depending on K.

Theorem 14.5. *Let C be given by the model $Y^2 = F(X, Z)$ with $F \in \mathbb{Z}[X, Z]$ and let $P \in J(\mathbb{Q})$ be given by primitive Kummer coordinates x (i.e., the coordinates are coprime integers). We can compute $\hat{h}(P)$ to d bits of precision in time*

$$\ll \log(d + h(P))\, \mathsf{M}(d + h(P))$$
$$+ (d + \log\log\|F\|_\infty)(\log d + \log\log\|F\|_\infty)\, \mathsf{M}((d + \log\log\|F\|_\infty)\log\|F\|_\infty).$$

Proof. The first term comes from computing $h(P)$. The second term dominates both the complexity bound for $\tilde{\mu}_\infty(P)$ from Proposition 13.1 and the complexity of computing $\tilde{\mu}^{\mathrm{f}}(P)$ using the algorithm of Proposition 14.3, since we have $D \le \frac{1}{16}|\Delta|$ and $\log D \ll \log\|F\|_\infty$. The time for the numerical evaluation of the logarithms $\log q_i$ to d bits of precision is also dominated by this term. □

Note that the complexity is quasilinear in $\log\|F\|_\infty$ and in $h(P)$, and quasi-quadratic in d. The latter is caused by the (only) linear convergence of the computation of $\tilde{\mu}_\infty(P)$. For elliptic curves one can use a quadratically convergent algorithm due to Bost and Mestre [1993] (see also [Müller and Stoll 2016]); such an algorithm in the genus-2 case would lead to a complexity that is quasilinear in d as well.

In Section 15 below we illustrate the efficiency of our algorithm by applying it to a family of curves and points with the property that the number g_0 above is large, so that the previously known algorithms have problems factoring it.

15. Examples

We have implemented our algorithm using the computer algebra system Magma [Bosma et al. 1997]. For the factorization into coprimes we have implemented a simple quadratic algorithm due to Buchmann and Lenstra [1994, Proposition 6.5] instead of the quasilinear, but more complicated, algorithms of [Bernstein 2004] or [Bernstein 2005].

Since the estimates for the required precision in the computation of the archimedean contribution as given in Section 13 are too wasteful in practice, we instead compute this contribution repeatedly using a geometrically increasing sequence of digits of precision until the results agree up to the desired number of bits.

We now compare our implementation with Magma's built-in `CanonicalHeight` (version 2.21-2), which is based on [Flynn and Smart 1997] and [Stoll 2002], for a family of genus-2 curves. In `CanonicalHeight`, the duplication on the Kummer surface is done using arithmetic over $\mathbb{Q}$, making the implementation slow when points with large coordinates show up during the computation. No factorization of the discriminant is required. However, to find a set of primes such that $\mu_p(P) \neq 0$

for every prime p not in the set, `CanonicalHeight` factors the integer $\gcd(\delta(x))$, where x are primitive Kummer coordinates for P.

Example 15.1. For an integer $a \neq 0$, consider the curve C_a of genus 2 defined by the integral Weierstrass model $y^2 = x^5 + a^2x + a^2$. Let J_a denote the Jacobian of C_a. Then the point $P = [((0, a)) - (\infty)] \in J_a(\mathbb{Q})$ is nontorsion. A set of primitive Kummer coordinates is given by $x = (0, 1, 0, 0)$ and we have $\delta(x) = (4a^2, 0, 0, a^4)$. Hence `CanonicalHeight` needs to factor a^2.

We choose this family of curves because (a) there is an obvious rational point P on the Jacobian that is generically nontorsion and (b) $\gcd(\delta(x))$ involves a large integer, where x is a set of primitive integral Kummer coordinates for P. For a random sextic polynomial in $\mathbb{Z}[x]$, very likely the discriminant will have a large squarefree part, and so $\gcd(\delta(x))$ will be fairly small. Of course, the advantages of our algorithm show most clearly when $\gcd(\delta(x))$ is too large to be factored quickly.

Consider

$$a = 58076586049885709421603671222868245057879201906396781960722099044468153398453014079361023706360328 2,$$

with partial factorization $2 \cdot 7 \cdot 643 \cdot 804743 \cdot a'$, where a' has 89 decimal digits, and its smallest prime factor has 34 decimal digits. Our implementation computes $\hat{h}(P)$ in 0.51 seconds, whereas Magma's `CanonicalHeight` needs about 15 minutes.

Next, we look at

$$a = 200403772956059488950289789507853617719701760528626768445669337185652379002740222523854354057543152846830555620006935999906608809182174662282078076286357255031457727185777958196892 0.$$

This factors as $a = 2^3 \cdot 5 \cdot 17 \cdot a'$, where a' has 178 decimal digits and no prime divisor with less than 50 decimal digits. Here, our implementation took 1.04 seconds to compute $\hat{h}(P)$, whereas Magma did not terminate in 8 weeks.

For $a = p \cdot q$, where p and q are the smallest primes larger than 10^{200} and 10^{250}, respectively, the canonical height of P was computed in 5.87 seconds using our implementation.

For the computations in these examples, we used a single-core Xeon CPU E7-8837 having 2.67GHz. All heights were computed to 30 decimal digits of precision.

We conclude this part with an example over the rational function field $\mathbb{Q}(t)$.

Example 15.2. Consider the curve $C/\mathbb{Q}(t)$ given by the equation

$$y^2 = x^6 - 2t(t+1)x^5 + (t+1)(t^3 - 5t^2 + 4t - 2)x^4 + 2t(t+1)^2(3t^2+1)x^3 - (t+1)(3t^4 - 2t^2 + 4t - 1)x^2 - 4t^2(t+1)^3(t^2 + 2t - 1)x + 4t^4(t+1)^4.$$

It has the points

$$P_1 = (1:1:0), \quad P_2 = (0, -2t^2(t+1)^2), \quad P_3 = (t+1, 2t(t-1)(t+1))$$

(and also points with x-coordinate $t(t+1)$ and a Weierstrass point $(-t-1, 0)$). Let $Q = [(P_1) - 2(P_2) + (P_3)] \in J(\mathbb{Q}(t))$. Its image on the Kummer surface has coordinates

$$(1 : -t+1 : -2t^2(t+1) : 0).$$

Applying the duplication polynomials and looking at the gcd of the result, we see that we have to compute the height correction functions at the places given by $t = 0$, $t = 1$ and $t = -1$. We also have to consider the place at infinity, since our model of C is not integral there. We use the algorithms of Section 12. Consider the place $t = 0$. From the valuations of the Igusa invariants (see Section 6) we can deduce that the reduction type is $[\mathrm{I}_{7-3-2}]$, which gives us $M = 41$ for the exponent of the component group and a bound $B = 10$ for ε. We follow Lemma 12.1 and compute

$$\mu_0(Q) = \frac{1}{41}\left\lceil 41 \sum_{n=0}^{3} 4^{-n-1}\varepsilon_0(2^n Q)\right\rceil = \frac{1}{41}\left\lceil 41\left(\frac{8}{4} + \frac{4}{4^2} + \frac{7}{4^3} + \frac{6}{4^4}\right)\right\rceil = \frac{98}{41}.$$

At $t = 1$, the model is not stably minimal. We can deduce from the Igusa invariants that there is a stably minimal model over an extension of ramification index 4, which has reduction type $[\mathrm{I}_{12-2-2}]$. This shows that the denominator of μ_1 is divisible by $4 \cdot 26 = 104$. With $M = 104$ and $B = 9$ we get $m = 4$ in Lemma 12.1; we obtain

$$\mu_1(Q) = \frac{1}{104}\left\lceil 104 \sum_{n=0}^{4} 4^{-n-1}\varepsilon_1(2^n Q)\right\rceil = \frac{1}{104}\left\lceil 104\left(\frac{4}{4} + \frac{4}{4^2} + \frac{3}{4^3} + \frac{2}{4^4} + \frac{2}{4^5}\right)\right\rceil = \frac{17}{13}.$$

At $t = -1$, the situation is similar. There is a stably minimal model over an extension with ramification index 4 again, which has reduction type $[\mathrm{I}_{20-0-0}]$. This leads to $M = 4 \cdot 20 = 80$ and $B = 20$, so $m = 4$, and

$$\mu_{-1}(Q) = \frac{1}{80}\left\lceil 80 \sum_{n=0}^{4} 4^{-n-1}\varepsilon_{-1}(2^n Q)\right\rceil = \frac{1}{80}\left\lceil 80\left(\frac{7}{4} + \frac{10}{4^2} + \frac{8}{4^3} + \frac{10}{4^4} + \frac{8}{4^5}\right)\right\rceil = \frac{51}{20}.$$

Finally, at the infinite place, there is a stably minimal integral model over an extension with ramification degree 2, which has reduction type $[\mathrm{I}_{8-0-0}]$. In a similar way as for $t = -1$ and taking into account a shift of -8 coming from making the model integral, we obtain $\mu_\infty(Q) = \frac{19}{4} - 8 = -\frac{13}{4}$. This results in

$$\begin{aligned}\hat{h}(Q) &= h(Q) - \mu_0(Q) - \mu_1(Q) - \mu_{-1}(Q) - \mu_\infty(Q) \\ &= 3 - \frac{98}{41} - \frac{17}{13} - \frac{51}{20} + \frac{13}{4} = \frac{11}{5330}.\end{aligned}$$

To our best knowledge, the point Q is the point of smallest known nonzero canonical height on the Jacobian of a curve of genus 2 over $\mathbb{Q}(t)$. The curve was found by Andreas Kühn (a student of the second author) in the course of a systematic search for curves with many points mapping into a subgroup of rank 1 in the Jacobian.

Part IV. Efficient search for points with bounded canonical height

16. Bounding the height difference at archimedean places

We now describe two approaches for getting a better upper bound $\tilde{\beta}$ on $\tilde{\mu}$ than the one coming from the bound on $\tilde{\varepsilon}$ given in [Stoll 1999, Equation (7.1)], when k is an archimedean local field and C/k is a smooth projective curve of genus 2, given by a Weierstrass equation $Y^2 = F(X, Z)$ in $\mathbb{P}_K(1, 3, 1)$.

We write $\|x\|_\infty = \max\{|x_1|, |x_2|, |x_3|, |x_4|\}$ for the maximum norm.

16A. ***Bounding $\tilde{\varepsilon}$ closely.*** For the first approach we assume that $k = \mathbb{R}$. We describe how to approximate $\max\{\tilde{\varepsilon}(P) : P \in J(\mathbb{R})\}$ to any desired accuracy, which gives us an essentially optimal bound $\tilde{\gamma}$. Recall that

$$\tilde{\varepsilon}(P) = -\log \frac{\max\{|\delta_1(x_1, x_2, x_3, x_4)|, \dots, |\delta_4(x_1, x_2, x_3, x_4)|\}}{\max\{|x_1|, |x_2|, |x_3|, |x_4|\}^4},$$

where $(x_1 : x_2 : x_3 : x_4)$ is the image of $P \in J(\mathbb{R})$ on the Kummer surface. We can normalize the Kummer coordinates in such a way that $\|x\|_\infty = 1$ and one of the coordinates is 1. We then have to minimize $\max\{|\delta_1|, \dots, |\delta_4|\}$ over four three-dimensional unit cubes, restricted to the points on the Kummer surface that are in the image of $J(\mathbb{R})$. This means that the relevant points satisfy the equation defining the Kummer surface and in addition the value of (at least) one of four further auxiliary polynomials is positive. (In general, the values of these polynomials are squares if the point comes from the Jacobian, and the converse holds for any one of the polynomials when its value is nonzero. One can choose four such polynomials in such a way that they do not vanish simultaneously on the Kummer surface.)

The idea is now to successively subdivide the given cubes. For each small cube, we check if it may contain points in the image of $J(\mathbb{R})$, by evaluating the various polynomials at the center of the cube and bounding the gradient on the cube. If it can be shown that the defining equation cannot vanish on the cube or that one of the auxiliary polynomials takes only negative values on the cube, then the cube can be discarded. Otherwise, we find upper and lower estimates for $\max\{|\delta_1|, \dots, |\delta_4|\}$ in a similar way. If the lower bound is larger than our current best upper bound for the minimum, the cube can also be discarded. (At the beginning, we have a trivial upper bound of 1 for the minimum, coming from the origin.) Otherwise, we keep it and subdivide it further. We continue until the difference of the upper and lower

bounds for $\tilde{\varepsilon}$ on the cube with the smallest lower bound for $\max\{|\delta_1|, \dots, |\delta_4|\}$ becomes smaller than a specified tolerance. The upper bound for $\tilde{\varepsilon}$ on that cube is then our bound $\tilde{\gamma}$, and we take (as before) $\tilde{\beta} = \frac{1}{3}\tilde{\gamma}$.

We have implemented this approach in Magma [Bosma et al. 1997]. After a considerable amount of fine-tuning, our implementation usually takes a few seconds to produce the required bound. In many cases the new bound, which is essentially optimal as a bound on $\tilde{\varepsilon}$, is considerably better than the bound of [Stoll 1999, Equation (7.1)], but there are also cases for which it turns out that the old bound is actually pretty good.

We used the following tricks to get the implementation reasonably fast.

- We keep the polynomials shifted and rescaled so that the cube under consideration is $[-1, 1]^3$.
- The shifting and scaling is done using linear algebra (working with vectors of coefficients and matrices) and not using polynomial arithmetic.
- The coordinates of the centers and vertices of all cubes are dyadic fractions. We scale everything (by $2^4 = 16$ at each subdivision step — note that the polynomials involved are of degree 4) so that we can compute with integers instead.

16B. ***Iterating Stoll's bound.*** We now describe a different approach that also works for complex places. Instead of trying to get an optimal bound on $\tilde{\varepsilon}$, we aim at a bound on $\tilde{\mu}$ by iterating the bound obtained from equation (7.1) in [Stoll 1999]. We recall how this bound was obtained. There is an elementary abelian group scheme G of order 32 that maps onto $J[2]$ and acts on the space of quadratic forms in the coordinates of the $\mathbb{P}^3$ containing the Kummer surface. This representation splits into a direct sum of ten one-dimensional representations that correspond to the ten partitions $\{S, S'\}$ of the set of ramification points of the double cover $C \to \mathbb{P}^1$ into two sets of three. We write $y_{\{S,S'\}}$ for suitably normalized generators of these eigenspaces ([Stoll 1999] gives explicit formulas in the case $H = 0$). We can then express the squares x_i^2 as linear combinations of these quadratic forms,

$$x_i^2 = \sum_{\{S,S'\}} a_{i,\{S,S'\}} y_{\{S,S'\}}(x),$$

for certain complex numbers $a_{i,\{S,S'\}}$ that can be explicitly determined. On the other hand, $y_{\{S,S'\}}^2$ is a quartic form invariant under the action of $J[2]$ (the representation of G on quartic forms descends to a representation of $J[2]$) and is therefore a linear combination of the duplication polynomials δ_j and the quartic defining the Kummer surface. So there are complex numbers $b_{\{S,S'\},j}$ that can also be explicitly determined such that

$$y_{\{S,S'\}}(x)^2 = \sum_{1 \le j \le 4} b_{\{S,S'\},j} \delta_j(x)$$

if x is a set of Kummer coordinates. Taking absolute values and using the triangle inequality, we obtain

$$\begin{aligned}|x_i|^4 &\le \left(\sum\nolimits_{\{S,S'\}} |a_{i,\{S,S'\}}||y_{\{S,S'\}}(x)|\right)^2 \\ &\le \left(\sum\nolimits_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum\nolimits_{1\le j\le 4} |b_{\{S,S'\},j}||\delta_j(x)|}\right)^2\end{aligned}$$

for all $(x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(\mathbb{C})$. This gives a bound for $\tilde{\varepsilon}$ in terms of the $a_{i,\{S,S'\}}$ and $b_{\{S,S'\},j}$ as in equation (7.1) of [Stoll 1999].

We refine this as follows. Define a function $\varphi\colon \mathbb{R}^4_{\ge 0} \to \mathbb{R}^4_{\ge 0}$ by

$$(d_1, d_2, d_3, d_4) \mapsto \left(\sqrt{\sum\nolimits_{\{S,S'\}} |a_{i,\{S,S'\}}| \sqrt{\sum\nolimits_{1\le j\le 4} |b_{\{S,S'\},j}| d_j}}\right)_{1\le i\le 4}.$$

Lemma 16.1. *Define a sequence $(b_n)_n$ in $\mathbb{R}^4_{\ge 0}$ by*

$$b_0 = (1, 1, 1, 1) \quad \text{and} \quad b_{n+1} = \varphi(b_n).$$

Then (b_n) converges to a limit b and we have

$$\tilde{\mu}(P) \le \frac{4^N}{4^N - 1} \log \|b_N\|_\infty$$

for all $N \ge 1$ and all $P \in J(\mathbb{C})$. In particular, $\sup \tilde{\mu}(J(\mathbb{C})) \le \log \|b\|_\infty$.

Proof. By our previous considerations, it is clear that $|\delta_j(x)| \le d_j$ for all j implies $|x_i| \le \varphi_i(d_1, d_2, d_3, d_4)$ for all i. We deduce by induction on N that

$$\log \|x\|_\infty \le \log \|b_N\|_\infty + 4^{-N} \log \|\delta^{\circ N}(x)\|_\infty$$

for all $N \ge 1$. Writing

$$\tilde{\mu}(P) = -\sum_{m=0}^{\infty} 4^{-mN} \left(\log \|\kappa(2^{mN} P)\|_\infty - 4^{-N} \log \|\delta^{\circ N}(\kappa(2^{mN} P))\|_\infty\right),$$

we obtain an upper bound of $\log \|b_N\|_\infty$ for each of the terms in parentheses, which gives the desired bound.

To see that (b_n) converges, we consider

$$\Phi(x) = \left(\log \varphi_i(\exp(x_1), \ldots, \exp(x_4))\right)_{1\le i\le 4}.$$

It is easy to see that the partial derivatives $\partial\Phi_i/\partial x_j$ are positive and that, for each i, summing them over j gives $\frac{1}{4}$. (This comes from the fact that φ_i is homogeneous of degree $\frac{1}{4}$.) This implies that $\|\Phi(x') - \Phi(x)\|_\infty \le \frac{1}{4}\|x' - x\|_\infty$, so that Φ is contracting with contraction factor $\le \frac{1}{4}$. The Banach fixed point theorem then guarantees the

existence of a unique fixed point of Φ, which every iteration sequence converges to. This implies the corresponding statement for φ. □

If we are dealing with a real place, then we may gain a little bit more by making use of the fact that the $\delta_j(x)$ are real, while some of the coefficients $b_{\{S,'S\},j}$ may be genuinely complex. This can lead to a better bound on $|y_{\{S,S'\}}|$.

For example, considering the curve with the record number of known rational points, we get an improvement from 7.726 to 0.973 for the upper bound on $-\tilde{\mu}$ using Lemma 16.1. See Section 19 for more details. In practice it appears that this second approach is at the same time more efficient and leads to better bounds than the approach described in Section 16A above.

The approach described here can also be applied in the context of heights on genus-3 hyperelliptic Jacobians; see [Stoll 2014].

17. Optimizing the naive height

We now consider an arbitrary local field k, with absolute value $|\cdot|$. Let C be given by an equation

$$Y^2 = F(X, Z),$$

and let W be the canonical class on C. The first three coordinates of the image of a point $P = [(X_1 : Y_1 : Z_1) + (X_2 : Y_2 : Z_2)] - W \in J$ on the Kummer surface are given by Z_1Z_2, $X_1Z_2 + Z_1X_2$, X_1X_2, whereas the fourth coordinate is homogeneous of degree 1 in the coefficients f_j of F $\left(\text{if we consider } Y_1 \text{ and } Y_2 \text{ to be of degree } \frac{1}{2}\right)$. This has the effect that the fourth coordinate usually differs by a factor of about $\|F\| := \max\{|f_0|, |f_1|, \dots, |f_6|\}$ from the other three, which gives this last coordinate a much larger (when $\|F\|$ is large; this is usually the case when k is archimedean) or smaller (this may occur when k is nonarchimedean) influence on the local contribution to the naive height when $k = K_v$ and K is a global field. This imbalance tends to increase the difference $h_{\mathrm{std}} - \hat{h}$ between naive and canonical height. This observation suggests to modify the naive height in the following way, so as to give all coordinates roughly the same weight. Compare Section 2 for the general setup. Let x be a set of Kummer coordinates over a global field K and set

$$h'(x) := \sum_{v \in M_K} \log \max\{|x_1|_v, |x_2|_v, |x_3|_v, |x_4|_v / \|F\|_v\}.$$

This is a height as in Example 2.3.

We state the following simple result, which will help us use this modified height.

Lemma 17.1. *Let $F_0 \in k[X, Z]$ be squarefree and homogeneous of degree 6. For $c \in k^\times$, let $C^{(c)}$ denote the curve $Y^2 = cF_0(X, Z)$. The Kummer surfaces* $\mathrm{KS}^{(1)}$

of $C^{(1)}$ *and* $\mathrm{KS}^{(c)}$ *of* $C^{(c)}$ *are isomorphic via*

$$\iota\colon \mathrm{KS}^{(1)} \to \mathrm{KS}^{(c)}, \quad (x_1 : x_2 : x_3 : x_4) \mapsto (x_1 : x_2 : x_3 : cx_4).$$

We abuse notation and write ι *also for the linear map*

$$(x_1, x_2, x_3, x_4) \mapsto (x_1, x_2, x_3, cx_4).$$

Write $\delta^{(c)}$ *for the duplication polynomials on* $\mathrm{KS}^{(c)}$. *Then*

$$\delta^{(c)}(\iota(x)) = c^3\iota(\delta^{(1)}(x)) \quad \textit{for each } x \in \mathrm{KS}^{(1)}_{\mathbb{A}}.$$

Proof. This can be checked by an easy calculation. □

If k is nonarchimedean and we use the modified local height given by

$$h'_v(x) = \log \max\{|x_1|_v, |x_2|_v, |x_3|_v, |x_4|_v/\|F\|_v\},$$

then we need to change the definition of ε accordingly to (compare Lemma 2.4)

$$\varepsilon(x) = \min\{v(\delta_1(x)), v(\delta_2(x)), v(\delta_3(x)), v(\delta_4(x)) - v(F)\} \\ - 4\min\{v(x_1), v(x_2), v(x_3), v(x_4) - v(F)\},$$

where $v(F) = v(\{f_0, \dots, f_6\})$. By Lemma 17.1 with $c = \pi^{v(F)}$, where π is a uniformizer of k, and $F_0 = c^{-1}F$, we then have, denoting the objects associated to F_0 by δ_0, ε_0 and μ_0,

$$\begin{aligned}\varepsilon(x) &= v\big(\iota^{-1}(\delta(x))\big) - 4v(\iota^{-1}(x)) \\ &= v\big(c^3\delta_0(\iota^{-1}(x))\big) - 4v(\iota^{-1}(x)) = 3v(F) + \varepsilon_0(\iota^{-1}(x)).\end{aligned}$$

This implies $\mu(x) = v(F) + \mu_0(\iota^{-1}(x))$. Let C_0 be the curve given by $Y^2 = F_0(X, Z)$. We then get that

$$\beta(C) \le v(F) + \bar{\beta}(C_0).$$

Note that the Jacobians of C and C_0 are in general only isomorphic over the ramified quadratic extension $k(\sqrt{\pi})$, so we cannot necessarily use $\beta(C_0)$ here. If $v(F)$ is even, however, then the isomorphism is defined over k, and we have $\beta(C) = v(F) + \beta(C_0)$.

So, except for the correction term $v(F)$, the effect is that we use the Kummer surface associated to the quadratic twist C_0 of C, which has a primitive polynomial on the right-hand side of its equation. Note in addition that this also allows us to deal with nonintegral equations; in this case, we again implicitly scale to make the polynomial on the right integral and primitive.

When $k = K_v \cong \mathbb{Q}_2$ (say) and we can write $F = 4F_1 + H^2$ with binary forms F_1 and H with integral coefficients, then C is isomorphic to the curve C' given by

the Weierstrass equation

$$Y^2 + H(X, Z)Y = F_1(X, Z),$$

and we can use the Kummer surface of the latter to define the local contribution to the naive height. The isomorphism between the Kummer surfaces is given by (see [Müller 2010, p. 53]; note that this is the inverse of the map given there)

$$(x_1 : x_2 : x_3 : x_4) \mapsto \big(x_1 : x_2 : x_3 : \tfrac{1}{4}x_4 + \tfrac{1}{2}(h_0h_2x_1 + h_0h_3x_2 + h_1h_2x_3)\big).$$

The scaling factor this induces for the δ polynomials is 2^6 in this case. So defining the local component at v of $h'(x)$ to be

$$\log \max\big\{|x_1|_v, |x_2|_v, |x_3|_v, \big|\tfrac{1}{4}x_4 + \tfrac{1}{2}(h_0h_2x_1 + h_0h_3x_2 + h_1h_2x_3)\big|_v\big\},$$

we can replace the bound for μ_v by the bound we get on C' plus 2. If we use this at the places above 2 where it applies (instead of, or combined with, the scaling described above), we still obtain a height as in Example 2.3.

If v is an archimedean place, then the approach described in Section 16B above can easily be adapted to the modified naive height. We just have to replace $b_{\{S,S'\},4} = 1$ by $\|F\|_v$ and $a_{4,\{S,S'\}}$ by $a_{4,\{S,S'\}}/\|F\|_v^2$. This will usually lead to a *negative* upper bound for $\tilde{\mu}_v$, which is fairly close to $-\log \|F\|_v$, at least when F is reduced in the sense of [Stoll and Cremona 2003] and its roots are not too close together. This is because the scaled $a_{i,\{S,S'\}}$ are now all of size $\approx \|F\|_\infty^{-2}$ and the scaled $b_{\{S,S'\},j}$ are all of size $\approx \|F\|_\infty$, so Φ as in the proof of Lemma 16.1 roughly satisfies $\|\Phi(x)\|_\infty \approx -\frac{3}{4}\log \|F\|_\infty + \frac{1}{4}\|x\|_\infty$, which has $-\log \|F\|_\infty$ as its fixed point.

Note that for a point $(0 : 0 : 0 : 1) \neq P = (x_1 : x_2 : x_3 : x_4) \in \mathrm{KS}(K)$ we have, for all versions h' of the modified height,

$$h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le h'(P).$$

We will therefore find all points P with $h'(P) \le B$, if we can enumerate all P with $h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le B$. This can be done (over $\mathbb{Q}$) by using the `-a` option of the second author's program `j-points`, which is available at [Stoll 2006]. (This option is also available in Magma version 2.22 or later.) In this way, enumerating all points as above with B up to roughly $\log 50\,000$ is feasible. See the discussion in Section 18.

Note that it is quite possible that we end up with a bound

$$h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le h'(P) \le \hat{h}(P) + \tilde{\beta} \quad \text{for all } P \in J(\mathbb{Q}) \setminus \{O\}$$

with $\tilde{\beta} < 0$. In this case $-\tilde{\beta}$ is a lower bound on the canonical height of any nontrivial point in $J(\mathbb{Q})$; in particular, the torsion subgroup of $J(\mathbb{Q})$ must be trivial. To give an indication of when we can expect $\tilde{\beta}$ to be close to zero or negative, write $|2^4 \operatorname{disc}(F)| = DD'$ with D and D' coprime and D' squarefree and odd. Then the contribution of the finite places to $\tilde{\beta}$ can be bounded by $\frac{1}{4}\log D$, and we get

$\tilde{\beta} \approx -\log \|F\|_\infty + \frac{1}{4}\log D$. So if $D \ll \|F\|_\infty^4$, we are in good shape. Note that $|\mathrm{disc}(F)| \ll \|F\|_\infty^{10}$, so this means that 60% or more of $\log|\mathrm{disc}(F)|$ comes from primes p dividing the discriminant exactly once. For curves that are not very special this is very likely to be the case.

In Section 19 we show how this approach can be used to get a very small bound for the height difference even for a curve with ten-digit coefficients.

18. Efficient enumeration of points of bounded canonical height

Let $C\colon y^2 = f(x)$ be a curve of genus 2 over $\mathbb{Q}$ with Jacobian J. In this section we describe the algorithm for enumerating all points $P \in J(\mathbb{Q})$ with $\hat{h}(P) \le B$ that follows from the considerations above. We assume that $f \in \mathbb{Z}[x]$ and proceed as follows.

1. Compute the complex roots of f numerically.
2. Compute the coefficients $a_{i,\{S,S'\}}$ and $b_{\{S,S'\},j}$ from the roots and the leading coefficient of f according to the formulas given in [Stoll 1999, §10].
3. Multiply all $a_{4,\{S,S'\}}$ by $\|f\|_\infty^{-2}$ and multiply all $b_{\{S,S'\},4}$ by $\|f\|_\infty$.
4. Iterate the function φ from Section 17 (but using the modified coefficients) a number of times, starting at $(1, 1, 1, 1)$, until there is little change; let $\tilde{\beta}_\infty$ be the upper bound for $\tilde{\mu}_\infty$ as in Lemma 16.1.
5. Factor the discriminant of f. Let g be the gcd of the coefficients of f.
6. For each prime divisor p of $2\,\mathrm{disc}(f)$, do the following.
   a. Let e_p be the p-adic valuation of g and set $f_1 = p^{-e_p} f$.
   b. If $p = 2$ and $f_1 = h^2 + 4f_2$ for polynomials $f_2, h \in \mathbb{Z}[x]$, then set $C_1\colon y^2 + h(x)y = f_2(x)$ and replace g by $4g$; otherwise set $C_1\colon y^2 = f_1(x)$. Let J_1 be the Jacobian of C_1.
   c. If e_p is even, let β_p be the bound for μ_p on $J_1(\mathbb{Q}_p)$ as obtained in Part II. Otherwise, let β_p be the bound for μ_p on $J_1(\overline{\mathbb{Q}}_p)$.
7. Set $\tilde{\beta} = \tilde{\beta}_\infty + \sum_p \beta_p \log p + \log g$.
8. Use `j-points` with the `-a` option to enumerate all points $O \ne P \in J(\mathbb{Q})$ such that $h_{\mathrm{std}}((\kappa_1(P) : \kappa_2(P) : \kappa_3(P))) \le B + \tilde{\beta}$.
9. Add O to this set and return it.

Note that $\log g$ is the sum of the correction terms $v_p(f)\log p$.

It follows from the discussion in the previous sections that the set returned by this algorithm contains all points with canonical height at most B. If necessary, one can compute the actual canonical heights using the algorithm from Part III and discard the points whose height is too large.

The actual enumeration is done by running through all points $(x_1 : x_2 : x_3) \in \mathbb{P}^2$ of (standard) height at most $B + \tilde{\beta}$ and checking whether there are rational numbers x_4

such that $(x_1 : x_2 : x_3 : x_4)$ is on the Kummer surface. For each of these points on the Kummer surface, we then check if it lifts to the Jacobian. Both these conditions are equivalent to some expression in the coordinates (and the coefficients of f) being a square. The `j-points` program tries to do this efficiently by using information modulo a number of primes to filter out triples that do not lift to rational points on J. Let $N = \lfloor \exp(B + \tilde{\beta}) \rfloor$. Then `j-points` usually takes a couple of seconds when $N = 1000$, a few minutes when $N = 5000$ and a few days when $N = 50\,000$. The running time scales with N^3, but the scaling factor depends on how effective the sieving mod p is. For Jacobians of high rank, the program tends to take longer than for "random" Jacobians.

Since the running time depends exponentially on $B + \tilde{\beta}$, it is very important to obtain a small bound $\tilde{\beta}$ for the difference between naive and canonical height. The improvement at the infinite place that we can achieve by considering a modified naive height is crucial for making the enumeration feasible also in cases when the defining polynomial has large coefficients. This is demonstrated by the example in Section 19.

If the discriminant of f is too large to be factored, then one can use

$$\tilde{\beta} = \tilde{\beta}_\infty + \tfrac{1}{4} \log |\mathrm{disc}(f_1)| + \log g$$

(or use information from small prime divisors as in the algorithm above and $\frac{1}{4} \log D$ for the remaining primes, where D is the unfactored part of the discriminant). But note that it is usually a great advantage to know the bad primes, since we can take $\beta_p = 0$ for primes p such that $v_p(\mathrm{disc}(f)) = 1$. In most cases, this leads to a much smaller bound $\tilde{\beta}$.

One of the most important applications of this enumeration algorithm is its use in saturating a given finite-index subgroup of $J(\mathbb{Q})$, which gives (generators of) the full group $J(\mathbb{Q})$. This is a necessary ingredient of the method for obtaining all integral points on C developed in [Bugeaud et al. 2008], for example, and for computing the regulator of $J(\mathbb{Q})$.

There are essentially two ways of performing the saturation. Let $G \subset J(\mathbb{Q})$ denote the known subgroup.

(i) Let ρ be (an upper bound for) the covering radius of the lattice $\Lambda = (G/G_{\mathrm{tors}}, \hat{h})$. Then $J(\mathbb{Q})$ is generated by G together with all points $P \in J(\mathbb{Q})$ that satisfy $\hat{h}(P) \le \rho^2$; see [Stoll 2002, Proposition 7.1]. This approach is feasible when $\tilde{\beta} + \rho^2$ is sufficiently small.

(ii) Let $I = (J(\mathbb{Q}) : G)$ denote the index; we assume $J(\mathbb{Q})_{\mathrm{tors}} \subset G$. If $m_1, \dots, m_r$ are the successive minima of Λ and there are no points $P \in J(\mathbb{Q}) \setminus G$ with $\hat{h}(P) < B$, then

$$I \le \sqrt{\frac{R \cdot \gamma_r^r}{\prod_{j=1}^r \min\{m_j, B\}}};$$

see [Flynn and Smart 1997, §7]. Here γ_r is (an upper bound for) the Hermite constant for lattices of rank r, and R is the regulator of G (i.e., the determinant of the Gram matrix of any basis of Λ). This can be used to get a bound on I whenever B is strictly positive, so for the enumeration we only need $\tilde{\beta}$ to be sufficiently small. (If $\tilde{\beta} < 0$, then we can do entirely without enumeration to get an index bound.) In a second step, one then has to check that G is p-saturated in $J(\mathbb{Q})$ (or find the largest group $G \subset G' \subset J(\mathbb{Q})$ with $(G' : G)$ a power of p) for all primes p up to the index bound. This can be done by considering the intersection of the kernels of the maps $J(\mathbb{Q})/pJ(\mathbb{Q}) \to J(\mathbb{F}_q)/pJ(\mathbb{F}_q)$ for a set of good primes q (such that the group on the right is nontrivial). If this intersection is trivial, then G is p-saturated; otherwise it tells us where to look for points that are potentially divisible by p. Since the index bound gets smaller with increasing B (as long as $B < m_r$), it makes sense to pick B in such a way as to balance the time spent in the two steps of this approach.

19. Example

As an example that demonstrates the use of our nearly optimal upper bound for the difference $h - \hat{h}$ between naive and canonical height (which is based on the optimal bounds for the μ_p obtained in Sections 9, 10 and 11 and the variation of the naive height discussed in Section 17), we consider the curve

$$C\colon\ y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600.$$

This curve is of interest, since it holds the current record for the largest number of known rational points (which is 642 for this curve); see [Stoll 2008]. A 2-descent on its Jacobian J (assuming GRH) as described in [Stoll 2001] and implemented in Magma gives an upper bound of 22 for the rank of $J(\mathbb{Q})$, and the differences of the known rational points generate a group of rank 22. The latter statement can be checked by computing the determinant R of the height pairing matrix of the 22 points in $J(\mathbb{Q})$ listed in Table 3, which is fairly fast using the algorithm for computing canonical heights described in Section 14. The points are given in Mumford representation $(a(x), b(x))$, which stands for $[(\theta_1, b(\theta_1)) + (\theta_2, b(\theta_2))] - W$, where θ_1, θ_2 are the two roots of $a(x)$ and W is the canonical class. Not all of these points are differences of rational points, but they are linear combinations of such differences.

We can easily check that $J(\mathbb{Q})$ has trivial torsion subgroup by computing the order of $J(\mathbb{F}_p)$ for a few good primes p.

The discriminant of C factors as

$$\Delta = 2^{47} \cdot 3^5 \cdot 5^9 \cdot 11^2 \cdot 13^2 \cdot 17^6 \cdot 19^4 \cdot 23^2 \cdot 41^4 \cdot 73^3 \cdot 2707 \cdot 43579 \cdot 108217976921 \cdot 8723283517315751077.$$

$$
\begin{array}{rr}
(x^2+x, 18868x+15740), & \left(x^2-\frac{1}{3}x, \frac{216800}{3}x-15740\right),\\
\left(x^2+\frac{2}{3}x-\frac{1}{3}, \frac{11747}{3}x+\frac{21131}{3}\right), & (x^2+5x+4, 276256x+273128),\\
\left(x^2+\frac{4}{3}x-\frac{5}{9}, 16315x+\frac{26195}{9}\right), & \left(x^2+\frac{53}{12}x+\frac{5}{3}, \frac{1433669}{6}x+\frac{371650}{3}\right),\\
(x^2-3x-4, 34104x+30976), & (x^2-4x-5, 65987x+69115),\\
\left(x^2+\frac{8}{5}x+\frac{3}{5}, 67671x+64543\right), & \left(x^2-5x-6, \frac{883626}{7}x+\frac{905522}{7}\right),\\
\left(x^2-\frac{3}{4}x-\frac{7}{4}, 31875x+35003\right), & \left(x^2+\frac{5}{7}x-\frac{2}{7}, \frac{432898}{49}x+\frac{279626}{49}\right),\\
\left(x^2+\frac{29}{6}x-\frac{178}{9}, \frac{3014179}{6}x-\frac{10824742}{9}\right), & \left(x^2+\frac{19}{84}x-\frac{65}{84}, \frac{4287373}{294}x+\frac{5207005}{294}\right),\\
\left(x^2+\frac{97}{42}x-\frac{37}{42}, \frac{23742013}{294}x-\frac{5459431}{294}\right), & \left(x^2-\frac{5}{11}x, \frac{1089388}{121}x-15740\right),\\
\left(x^2+\frac{325}{84}x-\frac{11}{21}, \frac{30014567}{147}x-\frac{2230444}{147}\right), & \left(x^2-\frac{683}{140}x-\frac{279}{140}, \frac{45519013}{490}x+\frac{5478709}{490}\right),\\
\left(x^2-\frac{91}{769}x-\frac{584}{769}, \frac{6911886712}{591361}x+\frac{16665656516}{591361}\right), & \left(x^2-\frac{259}{96}x+\frac{163}{72}, \frac{52305719}{768}x-\frac{13101271}{576}\right),\\
\left(x^2-\frac{3073}{2307}x-\frac{1252}{769}, \frac{54505985456}{1774083}x+\frac{25990632928}{591361}\right), & \left(x^2-\frac{137}{51}x+\frac{40}{51}, \frac{47131040}{867}x-\frac{8471860}{867}\right).
\end{array}
$$

Table 3. Generators of the known part of $J(\mathbb{Q})$.

The results of [Stoll 1999; Stoll 2002] lead to a bound of

$$\tfrac{1}{3}(43\log 2+3\log 3+9\log 5+2\log 11+2\log 13 \\ +6\log 17+4\log 19+2\log 23+4\log 41+3\log 73)\approx 40.1$$

for the contribution of the finite places to the height difference bound. When trying to get a better bound (for γ_p) by essentially doing an exhaustive search over the p-adic points of the Kummer surface, Magma gets stuck at $p=2$ for a long while, but eventually finishes with a contribution of 26.434 from the finite places and a total bound of 34.163. This contribution turns out to be $\frac{1}{3}\gamma_p \log p$ in all cases except for $p=73$, where it is $\frac{2}{3}\log 73$ instead of $\frac{1}{3}\log 73$. Our new results from this paper give bounds on the local contributions as shown in Table 4. Φ_p is the component group (ε and μ factor through it in all cases) and "gain" gives the gain in the bound on the height difference obtained by using the optimal bound on μ versus the bound $\frac{1}{3}\gamma$, where γ is the maximum of the values of ε.

This now gives a bound of ≈ 20.429 for the contribution of the finite places. The optimization of the naive height does not give any improvement at the odd finite places, since the polynomial f defining the curve is primitive. On the other hand, we note that f is congruent to a square mod 4, so we could use the Kummer surface of the curve $y^2+(x^2+x)y=f_1(x)$ (where $f(x)=4f_1(x)+(x^2+x)^2$) for the local height at 2, but this results in no improvement, since we have already used a minimal model to get our bound.

Now we consider the contribution of the infinite place. The bound obtained from [Stoll 1999, Equation (7.1)] is 7.726. Using Lemma 16.1 with $N=10$ improves

p	reduction type	Φ_p	β_p	$\frac{1}{3}\gamma_p$	gain
2	$[\mathrm{I}_{10-9-8}]$	$\mathbb{Z}/242\mathbb{Z}$	$2+{}^{1145}/_{242}$	${}^{26}/_{3}$	1.341
3	$[\mathrm{I}_0-\mathrm{IV}-0]$	$\mathbb{Z}/3\mathbb{Z}$	${}^{2}/_{3}$	${}^{2}/_{3}$	0.000
5	$[\mathrm{I}_{4-3-2}]$	$\mathbb{Z}/26\mathbb{Z}$	${}^{22}/_{13}$	2	0.495
11	$[\mathrm{I}_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	${}^{1}/_{2}$	${}^{2}/_{3}$	0.400
13	$[\mathrm{I}_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	${}^{1}/_{2}$	${}^{2}/_{3}$	0.427
17	$[\mathrm{I}_{2-2-2}]$	$\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/6\mathbb{Z}$	1	${}^{4}/_{3}$	0.944
19	$[\mathrm{I}_{2-1-1}]$	$\mathbb{Z}/5\mathbb{Z}$	${}^{3}/_{5}$	${}^{2}/_{3}$	0.196
23	$[\mathrm{I}_{2-0-0}]$	$\mathbb{Z}/2\mathbb{Z}$	${}^{1}/_{2}$	${}^{2}/_{3}$	0.523
41	$[\mathrm{I}_{2-1-1}]$	$\mathbb{Z}/5\mathbb{Z}$	${}^{3}/_{5}$	${}^{2}/_{3}$	0.248
73	$[\mathrm{I}_{1-1-1}]$	$\mathbb{Z}/3\mathbb{Z}$	${}^{1}/_{3}$	${}^{1}/_{3}$	0.000

Table 4. Bounds for β_p.

this to 0.973; increasing N further gives no significant improvement. However, modifying the local height at the infinite place by scaling the contribution of the fourth coordinate by $\|f\|_\infty^{-1}$ reduces this bound drastically to $\tilde{\mu}_\infty \le -19.25654$ (compare this to $-\log\|f\|_\infty \approx -21.59708$). This finally gives

$$h'(P) \le \hat{h}(P) + 1.17273$$

for our modified naive height h'.

So if we enumerate all points $P \in J(\mathbb{Q})$ with $h'(P) \le \log N$ and do not find points that are not in the known subgroup G, then we obtain a bound for the index $I = (J(\mathbb{Q}) : G)$ as follows (see the discussion at the end of Section 18):

$$I \le \sqrt{\frac{R\cdot\gamma_{22}^{22}}{\prod_{j=1}^{22}\min\{m_j, \log N - 1.17273\}}}.$$

Here R is the regulator of G and $m_1, m_2, \ldots, m_{22}$ are the successive minima of the lattice $(G, \hat{h})$, which are

8.5276, 8.5668, 8.5956, 8.8594, 9.0256, 9.0776, 9.1426, 9.1753,
9.4456, 9.7428, 9.7747, 9.9047, 9.9465, 9.9611, 9.9704, 10.1408,
10.3472, 10.3784, 10.5284, 10.5356, 10.6318, 10.9287.

With $N = 10\,000$ we obtain $I \le 6842$, with $N = 20\,000$ we get $I \le 2835$ and with $N \ge 178\,245$ we obtain the best possible bound $I \le 900$. We checked that there are no unknown points P with $\kappa(P) = (x_1 : x_2 : x_3 : x_4)$ such that $h_{\mathrm{std}}((x_1 : x_2 : x_3)) \le \log 20\,000$ and verified that the index is not divisible by any prime $p \le 2835$. The first computation took about two days on a single core, the second less than half a day. This implies the following.

Proposition 19.1. *Assume the generalized Riemann hypothesis. Let*

$$C\colon y^2 = 82342800x^6 - 470135160x^5 + 52485681x^4 + 2396040466x^3 + 567207969x^2 - 985905640x + 247747600$$

and denote by J the Jacobian of C. Then $J(\mathbb{Q})$ is a free abelian group of rank 22, freely generated by the points listed in Table 3. In particular, $J(\mathbb{Q})$ is generated by the differences of rational points on C.

Acknowledgments

We would like to thank David Holmes for suggesting the strategy of the proof of Proposition 7.3, Elliot Wells for pointing out an inaccuracy in the complexity analysis in Propositions 14.1 and 14.3, and the anonymous referee for some useful remarks and suggestions.

References

[Artin 1966] M. Artin, "On isolated rational singularities of surfaces", *Amer. J. Math.* **88** (1966), 129–136. MR Zbl 10

[Artin 1986] M. Artin, "Lipman's proof of resolution of singularities for surfaces", pp. 267–287 in *Arithmetic geometry* (Storrs, CT, 1984), edited by G. Cornell and J. H. Silverman, Springer, 1986. MR Zbl 7

[Barth et al. 1984] W. Barth, C. Peters, and A. Van de Ven, *Compact complex surfaces*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **4**, Springer, 1984. MR Zbl 5

[Bernstein 2004] D. J. Bernstein, "Research announcement: faster factorization into coprimes", preprint, 2004, https://cr.yp.to/lineartime/dcba2-20041009.ps. 5., 14, 14.4, 15

[Bernstein 2005] D. J. Bernstein, "Factoring into coprimes in essentially linear time", *J. Algorithms* **54**:1 (2005), 1–30. MR Zbl 5., 14, 14.4, 15

[Bombieri and Gubler 2006] E. Bombieri and W. Gubler, *Heights in Diophantine geometry*, New Mathematical Monographs **4**, Cambridge University Press, 2006. MR Zbl 4.5

[Borwein and Borwein 1987] J. M. Borwein and P. B. Borwein, *Pi and the AGM: a study in analytic number theory and computational complexity*, John Wiley & Sons, New York, 1987. MR Zbl 13

[Bosch and Liu 1999] S. Bosch and Q. Liu, "Rational points of the group of components of a Néron model", *Manuscripta Math.* **98**:3 (1999), 275–293. MR Zbl 7

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **21**, Springer, 1990. MR Zbl 7, 7, 7, 7, 7, 9, 9, 9, 12

[Bosma et al. 1997] W. Bosma, J. Cannon, and C. Playoust, "The Magma algebra system, I: The user language", *J. Symbolic Comput.* **24**:3-4 (1997), 235–265. MR Zbl 15, 16A

[Bost and Mestre 1993] J.-B. Bost and J.-F. Mestre, "Calcul de la hauteur archimédienne des points d'une courbe elliptique par un algorithme quadratiquement convergent et application au calcul de la capacité de l'union de deux intervalles", unpublished manuscript, 1993. 14

[Bruin and Stoll 2010] N. Bruin and M. Stoll, "The Mordell–Weil sieve: proving non-existence of rational points on curves", *LMS J. Comput. Math.* **13** (2010), 272–306. MR Zbl 7, 7

[Buchmann and Lenstra 1994] J. A. Buchmann and H. W. Lenstra, Jr., "Approximating rings of integers in number fields", *J. Théor. Nombres Bordeaux* **6**:2 (1994), 221–260. MR Zbl 15

[Bugeaud et al. 2008] Y. Bugeaud, M. Mignotte, S. Siksek, M. Stoll, and S. Tengely, "Integral points on hyperelliptic curves", *Algebra Number Theory* **2**:8 (2008), 859–885. MR Zbl 1, 18

[Cassels and Flynn 1996] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus* 2, London Mathematical Society Lecture Note Series **230**, Cambridge University Press, 1996. MR Zbl 1, 1, 3, 7

[Cinkir 2011] Z. Cinkir, "Zhang's conjecture and the effective Bogomolov conjecture over function fields", *Invent. Math.* **183**:3 (2011), 517–562. MR Zbl 8.4

[Conrad 2005] B. Conrad, "Minimal models for elliptic curves", unpublished manuscript, 2005, http://math.stanford.edu/~conrad/papers/minimalmodel.pdf. 7.5, 10

[Cremona et al. 2006] J. E. Cremona, M. Prickett, and S. Siksek, "Height difference bounds for elliptic curves over number fields", *J. Number Theory* **116**:1 (2006), 42–68. MR Zbl 10.13, 11

[Deligne and Mumford 1969] P. Deligne and D. Mumford, "The irreducibility of the space of curves of given genus", *Inst. Hautes Études Sci. Publ. Math.* **36** (1969), 75–109. MR Zbl 5

[Flynn and Smart 1997] E. V. Flynn and N. P. Smart, "Canonical heights on the Jacobians of curves of genus 2 and the infinite descent", *Acta Arith.* **79**:4 (1997), 333–352. MR Zbl 1, a., 4, 14, 15, 18

[Flynn et al. 2001] E. V. Flynn, F. Leprévost, E. F. Schaefer, W. A. Stein, M. Stoll, and J. L. Wetherell, "Empirical evidence for the Birch and Swinnerton-Dyer conjectures for modular Jacobians of genus 2 curves", *Math. Comp.* **70**:236 (2001), 1675–1697. MR Zbl 1

[von zur Gathen and Gerhard 1999] J. von zur Gathen and J. Gerhard, *Modern computer algebra*, Cambridge University Press, New York, 1999. MR Zbl 13

[Greuel and Kröning 1990] G.-M. Greuel and H. Kröning, "Simple singularities in positive characteristic", *Math. Z.* **203**:2 (1990), 339–354. MR Zbl 5

[Heinz 2004] N. Heinz, "Admissible metrics for line bundles on curves and abelian varieties over non-Archimedean local fields", *Arch. Math. (Basel)* **82**:2 (2004), 128–139. MR Zbl 1, 8

[Hindry and Silverman 2000] M. Hindry and J. H. Silverman, *Diophantine geometry: an introduction*, Graduate Texts in Mathematics **201**, Springer, 2000. MR Zbl 1, 2, 4

[Holmes 2012] D. Holmes, "Computing Néron–Tate heights of points on hyperelliptic Jacobians", *J. Number Theory* **132**:6 (2012), 1295–1305. MR Zbl 14

[Holmes 2014] D. Holmes, "An Arakelov-theoretic approach to naïve heights on hyperelliptic Jacobians", *New York J. Math.* **20** (2014), 927–957. MR Zbl 2

[Igusa 1960] J.-I. Igusa, "Arithmetic variety of moduli for genus two", *Ann. of Math.* (2) **72** (1960), 612–649. MR Zbl 6

[de Jong and Müller 2014] R. de Jong and J. S. Müller, "Canonical heights and division polynomials", *Math. Proc. Cambridge Philos. Soc.* **157**:2 (2014), 357–373. MR Zbl 14

[Lang 1983] S. Lang, *Fundamentals of Diophantine geometry*, Springer, 1983. MR Zbl 8, 12

[Liu 1993] Q. Liu, "Courbes stables de genre 2 et leur schéma de modules", *Math. Ann.* **295**:2 (1993), 201–222. MR Zbl 6, 6.1, 6, 6

[Liu 1994] Q. Liu, "Modèles minimaux des courbes de genre deux", *J. Reine Angew. Math.* **453** (1994), 137–164. MR Zbl 10, 12, 12

[Liu 1996] Q. Liu, "Modèles entiers des courbes hyperelliptiques sur un corps de valuation discrète", *Trans. Amer. Math. Soc.* **348**:11 (1996), 4577–4610. MR Zbl 4, 5, 5, 5, 11, 11, 11

[Liu 2002] Q. Liu, *Algebraic geometry and arithmetic curves*, Oxford Graduate Texts in Mathematics **6**, Oxford University Press, 2002. MR Zbl 7

[Mestre 1991] J.-F. Mestre, "Construction de courbes de genre 2 à partir de leurs modules", pp. 313–334 in *Effective methods in algebraic geometry* (Castiglioncello, 1990), edited by T. Mora and C. Traverso, Progr. Math. **94**, Birkhäuser, Boston, 1991. MR Zbl 6

[Müller 2010] J. S. Müller, "Explicit Kummer surface formulas for arbitrary characteristic", *LMS J. Comput. Math.* **13** (2010), 47–64. MR Zbl 1, 1, 3, 17

[Müller 2014] J. S. Müller, "Computing canonical heights using arithmetic intersection theory", *Math. Comp.* **83**:285 (2014), 311–336. MR Zbl 14

[Müller and Stoll 2016] J. S. Müller and M. Stoll, "Computing canonical heights on elliptic curves in quasi-linear time", *LMS J. Comput. Math.* **19**:suppl. A (2016), 391–405. MR 1, 12, 12, 12, 14, 14, 14, 14

[Namikawa and Ueno 1973] Y. Namikawa and K. Ueno, "The complete classification of fibres in pencils of curves of genus two", *Manuscripta Math.* **9** (1973), 143–186. MR Zbl 5, 6.3, 9, 9, 10, 10, 12.4, 12

[Néron 1965] A. Néron, "Quasi-fonctions et hauteurs sur les variétés abéliennes", *Ann. of Math.* (2) **82** (1965), 249–331. MR Zbl 1, 12

[Silverman 1988] J. H. Silverman, "Computing heights on elliptic curves", *Math. Comp.* **51**:183 (1988), 339–358. MR Zbl 7.5, 9.2

[Silverman 1994] J. H. Silverman, *Advanced topics in the arithmetic of elliptic curves*, Graduate Texts in Mathematics **151**, Springer, 1994. MR Zbl 9

[Stoll 1999] M. Stoll, "On the height constant for curves of genus two", *Acta Arith.* **90**:2 (1999), 183–201. MR Zbl 1, 1, 3.2, 11, 13, 14, 16, 16A, 16B, 2., 19, 19

[Stoll 2001] M. Stoll, "Implementing 2-descent for Jacobians of hyperelliptic curves", *Acta Arith.* **98**:3 (2001), 245–277. MR Zbl 19

[Stoll 2002] M. Stoll, "On the height constant for curves of genus two, II", *Acta Arith.* **104**:2 (2002), 165–182. MR Zbl 1, 1, 3, 3, 3, 3, 3, 3, 10, 14, 14, 15, 18, 19

[Stoll 2006] M. Stoll, "`j-points`, a program for searching rational points on genus 2 Jacobians", computer program, 2006, http://www.mathe2.uni-bayreuth.de/stoll/programs/index.html. 17

[Stoll 2008] M. Stoll, "A genus 2 curve with at least 642 rational points", electronic reference, 2008, http://www.mathe2.uni-bayreuth.de/stoll/recordcurve.html. 1, 19

[Stoll 2014] M. Stoll, "An explicit theory of heights for hyperelliptic Jacobians of genus three", preprint, 2014, http://www.mathe2.uni-bayreuth.de/stoll/papers/Kummer-g3-hyp-2014-05-15.pdf. 16B

[Stoll and Cremona 2003] M. Stoll and J. E. Cremona, "On the reduction theory of binary forms", *J. Reine Angew. Math.* **565** (2003), 79–99. MR Zbl 17

[Uchida 2011] Y. Uchida, "Canonical local heights and multiplication formulas for the Jacobians of curves of genus 2", *Acta Arith.* **149**:2 (2011), 111–130. MR Zbl 8, 14

[Zarkhin 1995] Y. G. Zarkhin, "Local heights and Néron pairings", *Trudy Mat. Inst. Steklov.* **208** (1995), 111–127. In Russian; translated in *Proc. Steklov Inst. Math.* **208** (1995), 100–114. MR Zbl 2, 4.5

[Zhang 1993] S. Zhang, "Admissible pairing on a curve", *Invent. Math.* **112**:1 (1993), 171–193. MR Zbl 1, 8

Communicated by Joseph H. Silverman
Received 2016-03-31 Revised 2016-08-02 Accepted 2016-09-05

jan.steffen.mueller@uni-oldenburg.de *Institut für Mathematik, Carl von Ossietzky Universität Oldenburg, D-26111 Oldenburg, Germany*

michael.stoll@uni-bayreuth.de *Mathematisches Institut, Universität Bayreuth, D-95440 Bayreuth, Germany*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2235

Combinatorial degenerations of surfaces and Calabi–Yau threefolds

Bruno Chiarellotto and Christopher Lazda

In this article we study combinatorial degenerations of minimal surfaces of Kodaira dimension 0 over local fields, and in particular show that the "type" of the degeneration can be read off from the monodromy operator acting on a suitable cohomology group. This can be viewed as an arithmetic analogue of results of Persson and Kulikov on degenerations of complex surfaces, and extends various particular cases studied by Matsumoto, Liedtke and Matsumoto, and Hernández Mada. We also study "maximally unipotent" degenerations of Calabi–Yau threefolds, following Kollár and Xu, showing in this case that the dual intersection graph is a 3-sphere.

1. Introduction

Fix a complete discrete valuation ring R with perfect residue field k of characteristic $p > 3$ and fraction field F. Let π be a uniformiser for R, and let X be a smooth and projective scheme over F. Let $\overline{F}$ be a separable closure of F.

MSC2010: primary 14J28; secondary 14G20, 11G25.
Keywords: monodromy, surfaces, good reduction.

Definition 1.1. A *model* of X over R is a regular algebraic space $\mathscr{X}$, proper and flat over $\mathscr{X}$ over R, whose generic fibre is isomorphic to X, and whose special fibre is a scheme. We say that a model is semistable if it is étale locally smooth over $R[x_1, \dots, x_d](x_1 \cdots x_r - \pi)$, and strictly semistable if furthermore the irreducible components of the special fibre Y are smooth over k.

A major question in arithmetic geometry is that of determining criteria under which X has good or semistable reduction over F, i.e., admits a model $\mathscr{X}$ which is smooth and proper over R, or semistable over R. In general the question of determining good reduction criteria comes in two flavours:

(1) Does there exists a model $\mathscr{X}$ of X which is smooth over R?

(2) Given a semistable model $\mathscr{X}$ of X, can we tell whether or not $\mathscr{X}$ is smooth?

We will refer to the first of these as the problem of "abstract" good reduction, and the second as the problem of "concrete" good reduction. The sorts of criteria we expect are those that can be expressed in certain homological or homotopical invariants of the variety in question. In this article we will mainly concentrate on these problems for minimal smooth projective surfaces over F of Kodaira dimension 0. These naturally fall into four classes:

- K3 surfaces;
- Enriques surfaces;
- abelian surfaces;
- bielliptic surfaces,

and in each case we have both the abstract and concrete good reduction problem. Note that for this article we will generally use "abelian surface" to mean a surface over F that is geometrically an abelian surface, i.e., we do not necessarily assume the existence of an F-rational point (or thus of a group law).

In the analogous complex analytic situation (i.e., that of a semistable, projective degeneration $X \to \Delta$ over the open unit disc with general fibre X_t a minimal complex algebraic surface with $\kappa = 0$) it was shown by Persson [1977] and Kulikov [1977] that, under a certain (reasonably strong) hypothesis on the total space X one could quite explicitly describe the "shape" of the special fibre, and that these shapes naturally fall into three "types" depending on the nilpotency index of the logarithm of the monodromy on a suitable cohomology group. Our main result here is an analogue of this result in an "arithmetic" context, namely classifying the special fibre of a strictly semistable scheme over R whose generic fibre is a surface of one of the above types, in terms of the monodromy operator on a suitable cohomology group. The exact form of the theorem is somewhat tricky to state simply, so here we content ourselves with providing a rough outline and refer to the body of the article for more detailed statements.

Theorem 1.2 (Theorems 6.4, 7.5, 8.3 and 9.3). *Let X/F be a minimal surface with $\kappa = 0$, and let ℓ be a prime (possibly equal to p). Let $\mathscr{X}/R$ be a "minimal" model of X in the sense of Definition 5.1. Then the special fibre Y of $\mathscr{X}$ is "combinatorial", and moreover there exists an "ℓ-adic local system" V_ℓ on X such that Y is of Type I, II or III as the nilpotency index of a certain monodromy operator on $H^i(X, V_\ell)$ is* 1, 2 *or* 3 *respectively.*

Remark 1.3. (1) We will not give the definition of "combinatorial" surfaces here, see Definitions 5.4, 5.5, 5.6 and 5.7.

(2) When $\mathrm{char}(F) = 0$ or $\mathrm{char}(F) = p \neq \ell$ then the local system V_ℓ is a $\mathbb{Q}_\ell$-étale sheaf on X, and the corresponding cohomology group is $H^i_{\text{ét}}(X_{\bar{F}}, V_\ell)$. This is an ℓ-adic representation of G_F, de Rham when $\ell = p$ and $\mathrm{char}(F) = 0$, and hence has a monodromy operator attached to it.

(3) When $\mathrm{char}(F) = p = \ell$ then the local system $V_\ell = V_p$ is an overconvergent F-isocrystal, and the corresponding cohomology group is a certain form of rigid cohomology $H^i_{\text{rig}}(X/\mathscr{R}_K, V_p)$. This is a (φ, ∇)-module over the Robba ring $\mathscr{R}_K$ and hence has a monodromy operator by the p-adic local monodromy theorem. For more details on p-adic cohomology in equicharacteristic p case see Section 2.

Certain types of results of this sort have been studied before, for example by Matsumoto [2015] (for $\mathrm{char}(F) \neq \ell$ and X a K3 surface), Liedtke and Matsumoto [2016] ($\mathrm{char}(F) = 0$, $\ell \neq p$ and X K3 or Enriques), Hernández Mada [2015] ($\mathrm{char}(F) = 0$, $\ell = p$ and X K3 or Enriques), and Pérez Buendía [2014] ($\mathrm{char}(F) = 0$, $\ell = p$ and X K3), and our purpose here is partly to unify these existing results into a broader picture, and partly to fill in various gaps, for example allowing $\ell = p = \mathrm{char}(F)$ in the case of K3 surfaces. It is perhaps worth noting that even treating the case of abelian surfaces is not quite as irrelevant as it may seem (given the rather well-known results on good reduction criteria for abelian varieties) since our result describes the possible shape of the special fibre of a *proper*, but not necessarily smooth model. We also relate these shapes to the more classical description of the special fibre of the Néron model, at least after a finite base change (Proposition 10.5).

In each case (K3, Enriques, abelian, bielliptic) the proof of the theorem is in two parts. The first consists of showing that the special fibre Y is combinatorial; this uses coherent cohomology and some basic (logarithmic) algebraic geometry. The second then divides the possible shapes into types depending on the nilpotency index of a certain monodromy operator N; this uses the weight spectral sequence and the weight monodromy conjecture (which in all cases is known for dimensions ≤ 2). Although we do not use it explicitly, constantly lurking in the background here is a Clemens–Schmid type exact sequence of the sort considered in [Chiarellotto and

Tsuzuki 2014]. Unfortunately, while the structure of the argument in all 4 cases is similar, we were not able to provide a single argument to cover all of them, hence parts of this article may seem somewhat repetitive.

The major hypothesis in the theorem is "minimality" of the model $\mathscr{X}$, which is more or less the assumption that the canonical divisor $K_{\mathscr{X}}$ of $\mathscr{X}$ is numerically trivial. For K3 surfaces one expects that such models exist (at least after a finite base change), and Matsumoto [2015] showed that this is true if the semistable reduction conjecture is true for K3 surfaces. For abelian surfaces, this argument adapts to show that one does always have such a model after a finite base change (Theorem 10.3), however, for Enriques surfaces there are counterexamples to the existence of such models (see [Liedtke and Matsumoto 2016]) and it seems likely that the same true for bielliptic surfaces. Unfortunately, the methods used by Persson, Kulikov et al. to describe the special fibre when one does not necessarily have these "minimal models" do not seem to be at all adaptable to the arithmetic situation.

Finally, we turn towards addressing similar questions in higher dimensions by looking at certain "maximally unipotent" degenerations of Calabi–Yau threefolds. The inspiration here is the recent work of Kollár and Xu [2016] on log Calabi–Yau pairs, using recently proved results on the minimal model program for threefolds in positive characteristic (in particular the existence of Mori fibre spaces from [Birkar and Waldron 2016]). The main result we obtain (Theorem 11.5) is only part of the story, unfortunately, proceeding any further (at least using the methods of this article) will require knowing that the weight monodromy conjecture holds in the given situation, so is only likely to be currently possible in equicharacteristic. A key part of the proof uses a certain description of the homotopy type (in particular the fundamental group) of Berkovich spaces, which forces us to restrict to models $\mathscr{X}/R$ which are schemes, rather than algebraic spaces. As the example of K3 surfaces shows, however, any result concerning the "abstract" good reduction problem is likely to involve algebraic spaces, and will therefore require methods to handle this case.

Notation and conventions. Throughout k will be a perfect field of characteristic $p>3$, R will be a complete DVR with residue field k and fraction field F, which may be of characteristic 0 or p. We will choose a uniformiser π for F, and let $\bar{F}$ denote a separable closure. We will denote by q some fixed power of p such that $\mathbb{F}_q \subset k$.

A variety over a field will be a separated scheme of finite type, and when X is proper and $\mathscr{F}$ is a coherent sheaf on X we will write

$$h^i(X,\mathscr{F}) = \dim H^i(X,\mathscr{F}) \quad \text{and} \quad \chi(X,\mathscr{F}) = \sum_i (-1)^i h^i(X,\mathscr{F}).$$

We will also write $\chi(X) = \chi(X,\mathscr{O}_X)$; since we always mean coherent Euler–Poincaré characteristics (rather than topological ones) this should not cause confusion.

Unless otherwise mentioned, a surface over any field will always mean a smooth, projective and geometrically connected surface. A ruled surface of genus g is a surface X together with a morphism $f : X \to C$ to a smooth projective surface C of genus g, whose generic fibre is isomorphic to $\mathbb{P}^1$. If we let F denote a smooth fibre of f then an *n-ruling* of f (for some $n \geq 1$) will be a smooth curve $D \subset X$ such $D \cdot F = n$, a 1-ruling will be referred to simply as a ruling.

2. Review of p-adic cohomology in equicharacteristic

In this section we will briefly review some of the material from [Lazda and Pál 2016] on p-adic cohomology when $\mathrm{char}(F) = p$, and explain some of the facts alluded to in the introduction, in particular the existence of monodromy operators. We will therefore let $W = W(k)$ denote the ring of Witt vectors of k, K its fraction field, and σ the q-power Frobenius on W and K. In this situation, we have an isomorphism $F \cong k((\pi))$, where π is our choice of uniformiser. We will let $\mathscr{R}_K$ denote the Robba ring over K, that is the ring of series $\sum_i a_i t^i$ with $a_i \in K$ such that

- for all $\rho < 1$, $|a_i|\rho^i \to 0$ as $i \to \infty$;
- for some $\eta < 1$, $|a_i|\eta^i \to 0$ as $i \to -\infty$.

In other words, it is the ring of functions convergent on some semiopen annulus $\eta \leq |t| < 1$. The ring of integral elements $\mathscr{R}_K^{\mathrm{int}}$ (i.e., those with $a_i \in W$) is therefore a lift of F to characteristic 0, in the sense that mapping $t \mapsto \pi$ induces $\mathscr{R}_K^{\mathrm{int}}/(p) \cong F$. We will denote by σ a Frobenius on $\mathscr{R}_K$, i.e., a continuous σ-linear endomorphism preserving $\mathscr{R}_K^{\mathrm{int}}$ and lifting the absolute q-power Frobenius on F, we will moreover assume that $\sigma(t) = ut^q$ for some $u \in (W[\![t]\!] \otimes_W K)^\times$. The reader is welcome to assume that $\sigma\left(\sum_i a_i t^i\right) = \sum_i \sigma(a_i) t^{iq}$. Let $\partial_t : \mathscr{R}_K \to \mathscr{R}_K$ denote the derivation given by differentiation with respect to t.

Definition 2.1. A (φ, ∇)-module over $\mathscr{R}_K$ is a finite free $\mathscr{R}_K$-module M together with

- a connection, that is a K-linear map $\nabla : M \to M$ such that
$$\nabla(rm) = \partial_t(r)m + r\nabla(m) \quad \text{for all } r \in \mathscr{R}_K \text{ and } m \in M;$$
- a horizontal Frobenius $\varphi : \sigma^* M := M \otimes_{\mathscr{R}_K, \sigma} \mathscr{R}_K \xrightarrow{\sim} M$.

Then (φ, ∇)-modules over $\mathscr{R}_K$ should be considered as p-adic analogues of Galois representations, for example, they satisfy a local monodromy theorem (see [Kedlaya 2004]) and hence have a canonical monodromy operator N attached to them (see [Marmora 2008]). More specifically, the connection ∇ should be viewed as an analogue of the action of the inertia subgroup I_F and the Frobenius φ the action of some Frobenius lift in G_F. The analogue for (φ, ∇)-modules

of inertia acting unipotently (on an ℓ-adic representation for $\ell \neq p$) or of a p-adic Galois representation being semistable (when $\mathrm{char}(F) = 0$) is therefore the connection acting unipotently, i.e., there being a basis $m_1, \dots, m_n$ such that $\nabla(m_i) \in \mathscr{R}_K m_1 + \cdots + \mathscr{R}_K m_{i-1}$ for all i. The analogue of being unramified or crystalline for a (φ, ∇)-module M is therefore the connection acting trivially, or in other words M admitting a basis of horizontal sections. We call such (φ, ∇)-modules M solvable.

Let $\mathscr{E}_K^\dagger \subset \mathscr{R}_K$ denote the bounded Robba ring, that is the subring consisting of series $\sum_i a_i t^i$ such that $|a_i|$ is bounded; we therefore have the notion of a (φ, ∇)-module over $\mathscr{E}_K^\dagger$, as in Definition 2.1. The main purpose of the book [Lazda and Pál 2016] was to define cohomology groups

$$X \mapsto H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger)$$

for $i \geq 0$ associated to any $k((\pi))$-variety X (i.e., separated $k((\pi))$-scheme of finite type), as well as versions with compact support $H^i_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger)$ or support in a closed subscheme $Z \subset X$, $H^i{}_{Z,\mathrm{rig}}(X/\mathscr{E}_K^\dagger)$. These are (φ, ∇)-modules over $\mathscr{E}_K^\dagger$ and enjoy all the same formal properties as ℓ-adic étale cohomology for $\ell \neq p$. Here we list a few of them:

(1) If X is of dimension d then $H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger) = H^i_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) = H^i_{Z,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) = 0$ for i outside the range $0 \leq i \leq 2d$.

(2) (Künneth formula) For any X, Y over $k((\pi))$ we have

$$H^n_{c,\mathrm{rig}}(X \times Y/\mathscr{E}_K^\dagger) \cong \bigoplus_{i+j=n} H^i_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \otimes_{\mathscr{E}_K^\dagger} H^j_{c,\mathrm{rig}}(Y/\mathscr{E}_K^\dagger)$$

and if X and Y are smooth over $k((\pi))$ we also have

$$H^n_{\mathrm{rig}}(X \times Y/\mathscr{E}_K^\dagger) \cong \bigoplus_{i+j=n} H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \otimes_{\mathscr{E}_K^\dagger} H^j_{\mathrm{rig}}(Y/\mathscr{E}_K^\dagger).$$

(3) (Poincaré duality) For any X smooth over $k((\pi))$ of equidimension d we have a perfect pairing

$$H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \times H^{2d-i}_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \to H^{2d}_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \cong \mathscr{E}_K^\dagger(-d)$$

where $(-d)$ is the Tate twist which multiplies the Frobenius structure on the constant (φ, ∇)-module $\mathscr{E}_K^\dagger$ by q^d.

(4) (Excision) For any closed $Z \subset X$ with complement $U \subset X$ we have long exact sequences

$$\cdots \to H^i_{Z,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \to H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \to H^i_{\mathrm{rig}}(U/\mathscr{E}_K^\dagger) \to \cdots$$

and

$$\cdots \to H^i_{c,\mathrm{rig}}(U/\mathscr{E}_K^\dagger) \to H^i_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \to H^i_{c,\mathrm{rig}}(Z/\mathscr{E}_K^\dagger) \to \cdots .$$

(5) (Gysin) For any closed immersion $Z \hookrightarrow X$ of smooth schemes over $k((\pi))$, of constant codimension c there is a Gysin isomorphism

$$H^i_{Z,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \cong H^{i-2c}_{\mathrm{rig}}(Z/\mathscr{E}_K^\dagger)(-c).$$

(6) There is a "forget supports" map $H^i_{c,\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \to H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger)$ which is an isomorphism whenever X is proper over $k((\pi))$.

(7) Let $U \subset C$ be an open subcurve of a smooth projective curve C of genus g, with complementary divisor D of degree d. Then

$$\dim_{\mathscr{E}_K^\dagger} H^1_{\mathrm{rig}}(U/\mathscr{E}_K^\dagger) = \begin{cases} 2g-1+d & \text{if } d \geq 1, \\ 2g & \text{if } d = 0. \end{cases}$$

(8) Let A be an abelian variety over $k((\pi))$ of dimension g. Then $H^1_{\mathrm{rig}}(A/\mathscr{E}_K^\dagger)$ is (more or less) isomorphic to the contravariant Dieudonné module of the p-divisible group $A[p^\infty]$ of A, has dimension $2g$, and

$$H^i_{\mathrm{rig}}(A/\mathscr{E}_K^\dagger) \cong \textstyle\bigwedge^i H^1_{\mathrm{rig}}(A/\mathscr{E}_K^\dagger).$$

All of these properties were proved in [Lazda and Pál 2016]. We may therefore define, for any variety $X/k((\pi))$

$$H^i_{\mathrm{rig}}(X/\mathscr{R}_K) := H^i_{\mathrm{rig}}(X/\mathscr{E}_K^\dagger) \otimes_{\mathscr{E}_K^\dagger} \mathscr{R}_K$$

as (φ,∇)-modules over $\mathscr{R}_K$. That the property of a (φ,∇)-module being solvable (resp. unipotent) really is the correct analogue of a Galois representation being unramified or crystalline (resp. unipotent or semistable) is suggested by the following result.

Theorem 2.2 [Lazda and Pál 2016, §5]. *Let $X/k((\pi))$ be smooth and proper. If X has good (resp. semistable reduction) then $H^i_{\mathrm{rig}}(X/\mathscr{R}_K)$ is solvable (resp. unipotent) for all $i \geq 0$. If moreover X is an abelian variety, then the converse also holds.*

In [Lazda and Pál 2016] was also shown an equicharacteristic analogue of the C_{st}-conjecture, namely that when $\mathscr{X}/R$ is proper and semistable, the cohomology $H^i_{\mathrm{rig}}(X/\mathscr{R}_K)$ of the generic fibre can be recovered from the log-crystalline cohomology $H^i_{\log\text{-cris}}(Y^{\log}/W^{\log}) \otimes_W K$ of the special fibre. Our task for the remainder of this section is to generalise this result to algebraic spaces (with fibres that are schemes).

So fix a smooth and proper variety X/F and a semistable model $\mathscr{X}/R$ (see Definition 1.1) for X. Let $Y^{\log}$ denote the special fibre of $\mathscr{X}$ with its induced log structure, and let $W^{\log}$ denote W with the log structure defined by $1 \mapsto 0$. Then the log-crystalline cohomology $H^i_{\log\text{-cris}}(Y^{\log}/W^{\log}) \otimes_W K$ is a (φ, N)-module over K, i.e., a vector space with semilinear Frobenius φ and nilpotent monodromy operator

N satisfying $N\varphi = q\varphi N$, and the rigid cohomology $H^i_{\mathrm{rig}}(X/\mathscr{R}_K)$ is a (φ,∇)-module over $\mathscr{R}_K$. There is a fully faithful functor

$$(-)\otimes_K \mathscr{R}_K : \underline{\mathbf{M}\Phi}^N_K \to \underline{\mathbf{M}\Phi}^\nabla_{\mathscr{R}_K}$$

from the category $\underline{\mathbf{M}\Phi}^N_K$ of (φ, N)-modules over K to that of (φ,∇)-modules over $\mathscr{R}_K$, whose essential image consists exactly of the unipotent (φ,∇)-modules, i.e., those which are iterated extensions of constant ones. The analogue of Fontaine's C_{st} conjecture in the equicharacteristic world is then the following.

Proposition 2.3. *There is an isomorphism*

$$\left(H^i_{\log\text{-cris}}(Y^{\log}/W^{\log})\otimes_W K\right)\otimes_K \mathscr{R}_K \cong H^i_{\mathrm{rig}}(X/\mathscr{R}_K)$$

in $\underline{\mathbf{M}\Phi}^\nabla_{\mathscr{R}_K}$.

Proof. Thanks to the extension of logarithmic crystalline cohomology and Hyodo–Kato cohomology to algebraic stacks by Olsson [2007], in particular base change [Olsson 2007, Theorem 2.6.2] and the construction of the monodromy operator [loc. cit., §6.5], the same proof as given in the scheme case (see Chapter 5 of [Lazda and Pál 2016]) works for algebraic spaces as well. □

In [Lazda and Pál 2016] was defined the notion of an overconvergent F-isocrystal on X, relative to K. These play the role in the p-adic theory of lisse ℓ-adic sheaves in ℓ-adic cohomology. Classically, i.e., over k, one can associate these objects to p-adic representations of the fundamental group, and we will need to do this also over Laurent series fields. We only need this for representations ρ with finite image, and in this case the construction is simple. So let $\rho : \pi_1^{\text{ét}}(X,\bar{x}) \to G$ be a finite quotient of the étale fundamental group of a smooth and proper variety over F, then this corresponds to a finite, étale, Galois cover $f : X' \to X$, and hence from results of [Lazda and Pál 2016] we have a pushforward functor

$$f_* : F\text{-Isoc}^\dagger(X'/K) \to F\text{-Isoc}^\dagger(X/K)$$

from overconvergent F-isocrystals on X' to those on X. We may therefore define $V_\rho \in F\text{-Isoc}^\dagger(X/K)$ to be the pushforward $f_*\mathscr{O}^\dagger_{X'/K}$ of the constant isocrystal on X'.

3. SNCL varieties

In this section, following F. Kato [1996, §11], we will introduce the key notion of a simple normal crossings log variety over k, or SNCL variety for short.

Definition 3.1. We say a geometrically connected variety Y/k is a normal crossings variety over k if it is étale locally étale over $k[x_0,\dots,x_d]/(x_0\cdots x_r)$.

Definition 3.2. Let Y denote a normal crossings variety over k, and let M_Y be a log structure on Y. Then we say that M_Y is of embedding type if étale locally on Y it is (isomorphic to) the log structure associated to the homomorphism of monoids

$$\mathbb{N}^{r+1} \to \frac{k[x_0, \ldots, x_d]}{(x_0 \cdots x_r)}$$

sending the i-th basis element of $\mathbb{N}^{r+1}$ to x_i.

Note that the existence of such a log structure imposes conditions on Y, and the log structure M_Y is *not* determined by the geometry of the underlying scheme Y. In fact, one can show that such a log structure exists if and only if, denoting by D the singular locus of Y, there exists a line bundle $\mathscr{L}$ on Y such that $\mathscr{E}xt^1(\Omega^1_{Y/k}, \mathscr{O}_Y) \cong \mathscr{L} \otimes \mathscr{O}_D$ (see for example Theorem 11.7 of [Kato 1996]).

Definition 3.3. We say that a log scheme $Y^{\log}$ of embedding type is of semistable type if there exists a log smooth morphism $Y^{\log} \to \mathrm{Spec}(k)^{\log}$ where the latter is endowed with the log structure of the punctured point.

Again, the existence of such a morphism implies conditions on Y, namely that $\mathscr{E}xt^1(\Omega^1_{Y/k}, \mathscr{O}_Y) \cong \mathscr{O}_D$ (where again D is the singular locus).

Definition 3.4. A SNCL variety over k is a smooth log scheme $Y^{\log}$ over $k^{\log}$ of semistable type, such that the irreducible components of Y are all smooth.

Any SNCL variety $Y^{\log}$ is log smooth over $k^{\log}$ by definition, and for all $p \geq 0$ we will let $\Lambda^p_{Y^{\log}/k^{\log}}$ denote the locally free sheaf of logarithmic p-forms on Y. We will also let $\omega_Y = \Lambda^{\dim Y}_{Y^{\log}/k^{\log}}$ denote the line bundle of top degree differential forms.

Proposition 3.5. *The sheaf ω_Y is a dualising sheaf for Y.*

Proof. This follows immediately from [Tsuji 1999, Proposition 2.14 and Theorem 2.21]. □

We will also need a spectral sequence for the cohomology of semistable varieties. This should be well-known, but we could not find a suitable reference.

Lemma 3.6. *Let $Y^{\log}$ be a SNCL variety over k of dimension n, with smooth components $Y_1, \ldots, Y_N$. For each $0 \leq s \leq n$ write*

$$Y^{(s)} = \coprod_{\substack{I \subset \{1,\ldots,N\} \\ |I| = s+1}} \bigcap_{i \in I} Y_i,$$

and let $i_s : Y^{(s)} \to Y$ denote the natural map. For $1 \leq t \leq s+1$ let

$$\partial^s_t : Y^{(s+1)} \to Y^{(s)}$$

be the canonical map induced by the natural inclusion $Y_{\{i_1,\dots,i_{s+1}\}} \to Y_{\{i_1,\dots,\hat{i_t},\dots,i_{s+1}\}}$. *Then the there exists an exact sequence*

$$0 \to \mathscr{O}_Y \xrightarrow{d^{-1}} i_{0*}\mathscr{O}_{Y^{(0)}} \xrightarrow{d^0} \cdots \xrightarrow{d^{n-1}} i_{n*}\mathscr{O}_Y^{(n)} \to 0$$

of sheaves on Y, *where* $d^{-1} = i_0^*$ *and*

$$d^s = \sum_{t=1}^{s+1} (-1)^t \partial_t^{s*} \quad \textit{for } s \geq 0.$$

Proof. We define a complex

$$0 \to \mathscr{O}_Y \to i_{0*}\mathscr{O}_{Y^{(0)}} \to \cdots \to i_{n*}\mathscr{O}_Y^{(n)} \to 0$$

using the formulae in the statement of the lemma; to check it is in fact exact (or indeed, to check that it is even a complex) we may work locally, and hence assume that Y is smooth over $\mathrm{Spec}(k[x_1,\dots,x_d]/(x_1\dots x_r))$. But now we can just use flat base change to reduce to the case where $Y = \mathrm{Spec}(k[x_1,\dots,x_d]/(x_1\dots x_r))$, which follows from a straightforward computation. □

Corollary 3.7. *In the above situation, there exists a spectral sequence*

$$E_1^{s,t} := H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) \Rightarrow H^{s+t}(Y, \mathscr{O}_Y).$$

4. Some useful results

In this section we prove three lemmas that will come in handy later on. The first characterises surfaces with effective anticanonical divisor of a certain form, analogous to Lemma 3.3.7 of [Persson 1977] in the complex case.

Lemma 4.1. *Let* k *be an algebraically closed field, and* V *a surface with canonical divisor* K_V. *Let* $\{C_i\}$ *be a nonempty family of smooth curves* C_i *on* V, *such that the divisor* $D = \sum_i C_i$ *is a simple normal crossings divisor, and we have* $K_V + D = 0$ *in* $\mathrm{Pic}(V)$. *Then one of the following must happen*:

(1) V *is an elliptic ruled surface, and* $D = E_1 + E_2$ *is a sum of disjoint elliptic curves, which are rulings on* V.

(2) V *is an elliptic ruled surface, and* $D = E$ *is a single elliptic curve, which is a* 2*-ruling on* V.

(3) V *is rational, and* $D = E$ *is an elliptic curve.*

(4) V *is rational, and* $D = \sum_{i=1}^d C_i$ *is a cycle of rational curves on* V, *i.e., either* $d = 2$ *and* $C_1 \cdot C_2 = 2$, *or* $d > 2$ *and* $C_1 \cdot C_2 = C_2 \cdot C_3 = \cdots = C_d \cdot C_1 = 1$, *with all other intersection numbers* 0.

Proof. The point is that since the classification of surfaces is essentially the same in characteristic p as characteristic 0, Persson's original proof carries over verbatim. We reproduce it here for the reader's benefit.

The hypotheses imply that V is of Kodaira dimension $-\infty$, and hence is either rational or ruled. For each curve C_i, let T_{C_i} denote the number of double points on C_i, that is $\sum_{j\neq i} C_i \cdot C_j$. By the genus formula we have

$$2g(C_i) - 2 = C_i \cdot (C_i + K_V) = -T_{C_i}$$

(here K_V is the canonical divisor) and hence either $T_{C_i} = 0$ and $g(C_i) = 1$ or $T_{C_i} = 2$ and $g(C_i) = 0$. Hence D is a disjoint sum of elliptic curves and cycles of rational curves.

Let $\pi : V \to V_0$ be a map onto a minimal model. For any i such that π does not contract C_i, let $C_{0i} = \pi(C_i)$, and let $D_0 := \pi(D)$. Any exceptional curve E has to either be a component of a rational cycle or meet exactly one component of D in exactly one point (because $D \cdot E = -K_V \cdot E = 1$). It then follows that D_0 has the same form as D (i.e., is a disjoint union of elliptic curves and cycles of rational curves) except that it might also contain nodal rational curves, not meeting any other components. If $V_0 \cong \mathbb{P}^2$, then the only possibilities for D_0 are a triangle of lines, a conic plus a line, a single elliptic curve or a nodal cubic. Therefore (V, D) has the form claimed.

Otherwise, V_0 is a $\mathbb{P}^1$ bundle over a smooth projective curve, let $F \subset V_0$ be a fibre intersecting all C_{0i} properly. Applying the genus formula again gives $K_{V_0} \cdot F = -2$, hence $D_0 \cdot F = 2 = \sum_i C_{0i} \cdot F$. Each connected component of D_0 is either a rational cycle, a nodal rational curve or an elliptic curve, and the first two kinds of components have to intersect F with multiplicity ≥ 2 (in the second case this is because it cannot be either a fibre or a degree 1 cover of the base). Hence if some C_{0i} is an elliptic curve E_1, then either $E_1 \cdot F = 2$, in which case $D_0 = E_1$, or $E_1 \cdot F = 1$, in which case we must have $D_0 = E_1 + E_2$ for some other elliptic curve E_2. In the first case V_0 can be elliptic ruled, in which case E_1 is a 2-ruling, or rational. In the second case V_0 must be elliptic ruled, and both E_1 and E_2 are rulings. Otherwise, each C_{0i} is a rational curve, V_0 must be rational and D_0 is either a single cycle of smooth rational curves or a single nodal rational curve. Again, this implies that (V, D) has the form claimed. □

We will also need the following cohomological computation.

Lemma 4.2. (1) *Let V be an elliptic ruled surface over k, and let ℓ be a prime number $\neq p$. Then* $\dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(V_{\bar{k}}, \mathbb{Q}_\ell) = \dim_K H^1_{\text{rig}}(V/K) = 2$.

(2) *Let V be a rational surface over k, and let ℓ be a prime number $\neq p$. Then* $\dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(V_{\bar{k}}, \mathbb{Q}_\ell) = \dim_K H^1_{\text{rig}}(V/K) = 0$.

Proof. One may use the excision exact sequence in either rigid or ℓ-adic étale cohomology to see that the first Betti number of a smooth projective surface is unchanged under monoidal transformations, and is hence a birational invariant. We may therefore reduce to the case of $E \times \mathbb{P}^1$ or $\mathbb{P}^1 \times \mathbb{P}^1$, which follows from the Künneth formula. □

Finally, we have the following (well known) result.

Lemma 4.3. *Let $\mathscr{X}/R$ be proper and flat. Assume that the generic fibre X is geometrically connected. Then so is the special fibre Y.*

Proof. Since $\mathscr{X}$ is proper and flat over R, the zeroth cohomology $H^0(\mathscr{X}, \mathscr{O}_{\mathscr{X}})$ is torsion free and finitely generated over R, hence it is free. Since the generic fibre is geometrically connected, it is of rank 1, and the natural map $R \to H^0(\mathscr{X}, \mathscr{O}_{\mathscr{X}})$ is an isomorphism. Since this also holds after any finite flat base change $R \to R'$, it follows from Zariski's Main Theorem [Belmans et al. 2005–, Tag 0A1C] that Y must in fact be geometrically connected. □

5. Minimal models, logarithmic surfaces and combinatorial reduction

The purpose of this section is to introduce the notion of a minimal model of a surface of Kodaira dimension 0, as well as the corresponding logarithmic and combinatorial versions of these surfaces. The basic idea in all cases is that we have

$$\text{minimal} \Rightarrow \text{logarithmic} \Rightarrow \text{combinatorial}$$

and although the general form that the picture takes is the same in all 4 cases, there are enough differences to merit describing how it works separately in each case. This unfortunately means that the next few sections are somewhat repetitive.

Let X/F be a smooth, projective, geometrically connected minimal surface of Kodaira dimension 0, and denote the canonical sheaf by ω_X. Then X falls into one of the following four cases:

(1) $\omega_X \cong \mathscr{O}_X$ and $h^1(X, \mathscr{O}_X) = 0$. Then X is a K3 surface.

(2) $h^0(X, \omega_X) = 0$ and $h^1(X, \mathscr{O}_X) = 0$. Then X is an Enriques surface.

(3) $\omega_X \cong \mathscr{O}_X$ and $h^1(X, \mathscr{O}_X) = 2$. Then X is an abelian surface.

(4) $h^0(X, \omega_X) = 0$ and $h^1(X, \mathscr{O}_X) = 1$. Then X is a bielliptic surface.

Note that if X is an Enriques surface we have $\omega_X^{\otimes 2} \cong \mathscr{O}_X$ and if X is a bielliptic surface we have $\omega_X^{\otimes m} \cong \mathscr{O}_X$ for $m = 2, 3, 4$ or 6. Also note that since $p > 3$ the classification of such surfaces is the same over k as over F (i.e., we do not have to consider the "extraordinary" Enriques or bielliptic surfaces). In all cases we may therefore define an integer m as the smallest positive integer such that $\omega_X^{\otimes m} \cong \mathscr{O}_X$. If $\mathscr{X}/R$ is a semistable model for X then we will let $\mathscr{X}^{\log}$ denote the log scheme with

log structure induced by the special fibre; this is log smooth over $R^{\log}$, where the log structure is again induced by the special fibre $\pi = 0$. We will let $\omega_{\mathscr{X}} = \Lambda^2_{\mathscr{X}^{\log}/R^{\log}}$ denote the line bundle of logarithmic 2-forms on $\mathscr{X}$. We will also let Y denote the special fibre, and $Y^{\log}/k^{\log}$ the smooth log scheme whose log structure is the one pulled back from that on $\mathscr{X}$.

Definition 5.1. Let $\mathscr{X}/R$ be a semistable model for X. Then we say that $\mathscr{X}$ is minimal if it is strictly semistable and $\omega_{\mathscr{X}}^{\otimes m} \cong \mathscr{O}_{\mathscr{X}}$.

Warning. When X is an Enriques surface, there are counter-examples to the existence of such minimal models, even allowing for finite extensions of R.

The first stage is in passing from minimal models to logarithmic surfaces of Kodaira dimension 0, the latter being defined by logarithmic analogues of the above criteria.

Definition 5.2. Let $Y^{\log}/k^{\log}$ be a proper SNCL scheme over k, of dimension 2, and let $\omega_Y = \Lambda^2_{Y^{\log}/k^{\log}}$ be its canonical sheaf. Then we say that $Y^{\log}$ is a

(1) logarithmic K3 surface if $\omega_Y \cong \mathscr{O}_Y$ and $h^1(Y, \mathscr{O}_Y) = 0$;

(2) logarithmic Enriques surface if ω_Y is torsion in $\mathrm{Pic}(Y)$, $h^0(Y, \omega_Y) = 0$ and $h^1(Y, \mathscr{O}_Y) = 0$;

(3) logarithmic abelian surface if $\omega_Y \cong \mathscr{O}_Y$ and $h^1(Y, \mathscr{O}_Y) = 2$;

(4) logarithmic bielliptic surface if ω_Y is torsion in $\mathrm{Pic}(Y)$, $h^0(Y, \omega_Y) = 0$ and $h^1(Y, \mathscr{O}_Y) = 1$.

Proposition 5.3. *Let X/F be a minimal surface of Kodaira dimension* 0, *and $\mathscr{X}/R$ a minimal model. Then $Y^{\log}$ is a logarithmic K*3 (*resp. Enriques*, *abelian*, *bielliptic*) *surface if X is K*3 (*resp. Enriques*, *abelian*, *bielliptic*).

Proof. Note that the only obstruction to $Y^{\log}/k^{\log}$ being an SNCL variety is geometric connectedness, which follows from Lemma 4.3. The conditions on the canonical sheaf ω_Y in Definition 5.2 follow from the definition of minimality, it therefore suffices to verify the required dimensions of the coherent cohomology groups on Y. We divide into the four cases.

First assume that X is a K3 surface. Then we have $\chi(X, \mathscr{O}_X) = 2$, and hence by local constancy of χ under a flat map (see [Hartshorne 1977, Chapter III, Theorem 9.9]) we must also have that $\chi(Y, \mathscr{O}_Y) = 2$. Since Y is geometrically connected by Lemma 4.3, we have $h^0(Y, \mathscr{O}_Y) = 1$, and therefore $h^2(Y, \mathscr{O}_Y) - h^1(Y, \mathscr{O}_Y) = 1$. But by Proposition 3.5 we must have $h^2(Y, \mathscr{O}_Y) = h^0(Y, \omega_Y)$, and by definition of minimality we know that $\omega_Y \cong \mathscr{O}_Y$. Hence $h^2(Y, \mathscr{O}_Y) = 1$ and therefore $h^1(Y, \mathscr{O}_Y) = 0$. Hence $Y^{\log}$ is a logarithmic K3 surface.

Next assume that X is Enriques. Then as above, we have that $h^0(Y, \mathscr{O}_Y) = 1$ and hence by local constancy of χ, that $h^1(Y, \mathscr{O}_Y) = h^2(Y, \mathscr{O}_Y)$. Let $\pi : \widetilde{\mathscr{X}} \to \mathscr{X}$ denote

the canonical double cover coming from the 2-torsion element $\omega_{\mathscr{X}} \in \mathrm{Pic}(\mathscr{X})$, with generic fibre $\widetilde{X} \to X$ and special fibre $\widetilde{Y} \to Y$. Then $\widetilde{\mathscr{X}}$ is a minimal model of the K3 surface $\widetilde{X}$, and hence $\widetilde{Y}^{\log}$ is a logarithmic K3 surface. Hence $h^1(\widetilde{Y}, \mathscr{O}_{\widetilde{Y}}) = 0$, and since $\mathscr{O}_Y \subset \pi_* \mathscr{O}_{\widetilde{Y}}$ is a direct summand, we must have $h^1(Y, \mathscr{O}_Y) = 0$, and therefore $h^0(Y, \omega_Y) = h^2(Y, \mathscr{O}_Y) = 0$. Thus $Y^{\log}$ is a logarithmic Enriques surface.

The case of abelian surfaces is handled entirely similarly to that of K3 surfaces, and the case of bielliptic surfaces is then deduced as Enriques is deduced from K3. □

The next notion is that of combinatorial versions of the above four cases.

Definition 5.4. Let Y be a proper surface over k (not necessarily smooth). We say that Y is a combinatorial K3 surface if, geometrically (i.e., over $\bar{k}$), one of the following situations occurs:

- (Type I) Y is a smooth K3 surface.
- (Type II) $Y = Y_1 \cup \cdots \cup Y_N$ is a chain with Y_1, Y_N smooth rational surfaces and all other Y_i elliptic ruled surfaces, with each double curve on each "inner" component a ruling. The dual graph of $Y_{\bar{k}}$ is a straight line with endpoints Y_1 and Y_N.
- (Type III) Y is a union of smooth rational surfaces, the double curves on each component form a cycle of rational curves, and the dual graph of $Y_{\bar{k}}$ is a triangulation of S^2.

Definition 5.5. Let Y be a proper surface over k (not necessarily smooth). We say that Y is a combinatorial Enriques surface if, geometrically, one of the following situations occurs:

- (Type I) Y is a smooth Enriques surface.
- (Type II) $Y = Y_1 \cup \cdots \cup Y_N$ is a chain of surfaces, with Y_1 rational and all others elliptic ruled, with each double curve on each "inner" component a ruling and the double curve on Y_N a 2-ruling. The dual graph of $Y_{\bar{k}}$ is a straight line with endpoints Y_1 and Y_N.
- (Type III) Y is a union of smooth rational surfaces, the double curves on each component form a cycle of rational curves, and the dual graph of $Y_{\bar{k}}$ is a triangulation of $\mathbb{P}^2(\mathbb{R})$.

Definition 5.6. Let Y be a proper surface over k (not necessarily smooth). We say that Y is a combinatorial abelian surface if, geometrically, one of the following situations occurs:

- (Type I) Y is a smooth abelian surface.
- (Type II) $Y = Y_1 \cup \cdots \cup Y_N$ is a cycle of elliptic ruled surfaces, with each double curve a ruling. The dual graph of $Y_{\bar{k}}$ is a circle.

- (Type III) Y is a union of smooth rational surfaces, the double curves on each component form a cycle of rational curves, and the dual graph of $Y_{\bar{k}}$ is a triangulation of the torus $S^1 \times S^1$.

Definition 5.7. Let Y be a proper surface over k (not necessarily smooth). We say that Y is a combinatorial bielliptic surface if, geometrically, one of the following situations occurs:

- (Type I) Y is a smooth bielliptic surface.
- (Type II) $Y = Y_1 \cup \cdots \cup Y_N$ is either a cycle or chain of elliptic ruled surfaces, with each double curve either a ruling (cycles or "inner" components of a chain) or a 2-ruling ("end" components of a chain). The dual graph of $Y_{\bar{k}}$ is either a circle or a line segment.
- (Type III) Y is a union of smooth rational surfaces, the double curves on each component form a cycle of rational curves, and the dual graph of $Y_{\bar{k}}$ is a triangulation of the Klein bottle.

Of course, in each case logarithmic surfaces will turn out to be combinatorial; this has been proved by Nakkajima for K3 and Enriques surfaces, and we will show it during the course of this article for abelian (Theorem 8.1) and bielliptic (Theorem 9.1) surfaces.

6. K3 surfaces

In this section, we will properly state and prove Theorem 1.2 for K3 surfaces. The case when $\mathrm{char}(F) = 0$ and $\ell = p$ is due to Hernández Mada [2015], and Perez Buendía [2014] and the case $\ell \neq p$ should be well-known (and at least part of it is implicitly proved in [Matsumoto 2015]), however, we could not find a reference in the literature so we include a proof here for completeness. We begin with a result of Nakkajima.

Theorem 6.1 [Nakkajima 2000, §3]. *Let $Y^{\log}$ be a logarithmic K3 surface over k. Then the underlying scheme Y is a combinatorial K3 surface.*

Remark 6.2. A proof of this result given entirely in terms of coherent cohomology can be given as in Theorem 8.1 below.

Corollary 6.3. *Let $\mathscr{X}/R$ be a minimal semistable model of a K3 surface X/F. Then the special fibre Y is a combinatorial K3 surface.*

For a K3 surface X/K, and for all $\ell \neq p$, the second cohomology group $H^2_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$ is a finite dimensional $\mathbb{Q}_\ell$ vector space with a continuous Galois action, which is quasiunipotent. If $\ell = p$ and $\mathrm{char}(F) = 0$ then $H^2_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_p)$ is a de Rham representation of G_F, and if $\mathrm{char}(F) = p$ then $H^2_{\text{rig}}(X/\mathscr{R}_K)$ is a (φ, ∇)-module over $\mathscr{R}_K$.

If we therefore let $H^2(X)$ stand for

- $H^2_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$ if $\ell \neq p$;
- $H^2_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_p)$ if $\ell = p$ and $\text{char}(F) = 0$;
- $H^2_{\text{rig}}(X/\mathscr{R}_K)$ if $\ell = p$ and $\text{char}(F) = p$,

then in all cases we get a monodromy operator N on $H^2(X)$.

Theorem 6.4. *Let $\mathscr{X}/R$ be a minimal semistable model of a K3 surface X, and Y its special fibre, which is a combinatorial K3 surface. Then Y is of Type I, II or III respectively as the nilpotency index of N on $H^2(X)$ is 1, 2 or 3.*

Proof.

When $\mathscr{X}$ is a scheme, the case $\ell = p$ and $\text{char}(F) = 0$ is due to Hernández Mada, and in fact the case $\ell = \text{char}(F) = p$ also follows from his result by applying the results in Chapter 5 of [Lazda and Pál 2016].

To deal with the case $\ell \neq \text{char}(k)$ (and $\mathscr{X}$ an algebraic space), we use the weight spectral sequence (for algebraic spaces this is Proposition 2.3 of [Matsumoto 2015]). Let $Y = Y_1 \cup \cdots \cup Y_N$ be the components of Y, $C_{ij} = Y_i \cap Y_j$ the double curves and

$$Y^{(0)} = \coprod_i Y_i, \quad Y^{(1)} = \coprod_{i<j} C_{ij}, \quad Y^{(2)} = \coprod_{i<j<k} Y_i \cap Y_j \cap Y_k.$$

We consider the weight spectral sequence

$$E_1^{s,t} = \bigoplus_{j \geq \max\{0,-s\}} H^{t-2j}_{\text{ét}}(Y_{\bar{k}}^{(s+2j)}, \mathbb{Q}_\ell)(-j) \Rightarrow H^{s+t}_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$$

which degenerates at E_2 and is compatible with monodromy in the sense that there exists a morphism $N : E_r^{s,t} \to E_r^{s+2,t-2}$ of spectral sequences abutting to the monodromy operator on $H^{s+t}_{\text{ét}}(X_{\bar{K}}, \mathbb{Q}_\ell)$. Moreover, by the weight-monodromy conjecture (see [Nakkajima 2006, Remark 6.8(1)]) we know that N^r induces an isomorphism $E_2^{-r,w+r} \xrightarrow{\sim} E_2^{r,w-r}$. Hence we can characterise the three cases where N has nilpotency index 1, 2 or 3 in terms of the weight spectral sequence as follows:

(1) $N = 0$ if and only if $E_2^{1,1} = E_2^{2,0} = 0$.

(2) $N \neq 0$, $N^2 = 0$ if and only if $E_2^{1,1} \neq 0$ and $E_2^{2,0} = 0$.

(3) $N^2 \neq 0$, $N^3 = 0$ if and only if $E_2^{1,1}, E_2^{2,0} \neq 0$.

Hence it suffices to show the following:

(1) If Y is of Type I, then $E_2^{1,1} = 0$.

(2) If Y is of Type II, then $E_2^{1,1} \neq 0$ and $E_2^{2,0} = 0$.

(3) If Y is of Type III, then $E_2^{2,0} \neq 0$.

The first of these is clear, and in both the Type II and III cases the term

$$E_2^{2,0} = \operatorname{coker}\big(H^0(Y_{\bar{k}}^{(1)}, \mathbb{Q}_\ell) \to H^0(Y_{\bar{k}}^{(2)}, \mathbb{Q}_\ell)\big)$$

is simply the second singular cohomology $H^2_{\mathrm{sing}}(\Gamma, \mathbb{Q}_\ell)$ of the dual graph Γ. For Type II this is 0, and for Type III this is 1-dimensional over $\mathbb{Q}_\ell$, hence it suffices to show that if Y is of Type II, then $E_2^{1,1} \neq 0$.

But we know that

$$\dim_{\mathbb{Q}_\ell} E_2^{-1,2} + \dim_{\mathbb{Q}_\ell} E_2^{0,1} + \dim_{\mathbb{Q}_\ell} E_2^{1,0} = \dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(X, \mathbb{Q}_\ell) = 0$$

and hence $\dim_{\mathbb{Q}_\ell} E_2^{0,1} = 0$. Therefore we have

$$\dim_{\mathbb{Q}_\ell} E_2^{1,1} = \dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(Y^{(1)}, \mathbb{Q}_\ell) - \dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(Y^{(0)}, \mathbb{Q}_\ell)$$

which using Lemma 4.2 we can check to be equal to $2(N-1) - 2(N-2) = 2$. Hence $E_2^{1,1} \neq 0$ as required.

To deal with the case $\ell = \mathrm{char}(k)$ and $\mathscr{X}$ an algebraic space, we argue entirely similarly, using the p-adic weight spectral sequence and Proposition 2.2(4) of [Matsumoto 2015] (F mixed characteristic) or Proposition 2.3 (F equicharacteristic). □

7. Enriques surfaces

To deal with the case of Enriques surfaces, we again start with a result of Nakkajima, analogous to the one quoted above.

Theorem 7.1 [Nakkajima 2000, §7]. *Let $Y^{\log}$ be a logarithmic Enriques surface over k. Then the underlying scheme Y is a combinatorial Enriques surface.*

Remark 7.2. Again, it is possible to prove this only using coherent cohomology as in Theorem 9.1 below.

Corollary 7.3. *Let $\mathscr{X}/R$ be a minimal semistable model of an Enriques surface X/F. Then the special fibre Y is a combinatorial Enriques surface.*

If X/F is an Enriques surface, then for all $\ell \neq p$ the second homotopy group $\pi_2^{\text{ét}}(X_{\bar{F}})_{\mathbb{Q}_\ell}$ (for the definition of the higher homotopy groups of algebraic varieties, see [Artin and Mazur 1969]) is a finite dimensional $\mathbb{Q}_\ell$ vector space with a continuous Galois action, which is quasiunipotent. If $\ell = p$ and $\mathrm{char}(F) = 0$ then $\pi_2^{\text{ét}}(X_{\bar{F}})_{\mathbb{Q}_p}$ is a de Rham representation of G_F. If $\mathrm{char}(F) = p$ there is (currently!) no general theory of higher homotopy groups, so instead we cheat somewhat and use the known properties of the higher étale homotopy groups to justify making the following definition.

Definition 7.4. We define $\pi_2^{\mathrm{rig}}(X/\mathscr{R}_K) := H^2_{\mathrm{rig}}(\widetilde{X}/\mathscr{R}_K)^\vee$, where $\widetilde{X} \to X$ is the canonical double cover of X.

Thus $\pi_2^{\mathrm{rig}}(X/\mathscr{R}_K)$ is a (φ,∇)-module over $\mathscr{R}_K$. Again, if we let $\pi_2(X)$ stand for any of $\pi_2^{\text{ét}}(X_{\bar{F}})_{\mathbb{Q}_\ell}$, $\pi_2^{\text{ét}}(X_{\bar{F}})_{\mathbb{Q}_p}$ or $\pi_2^{\mathrm{rig}}(X/\mathscr{R}_K)$, then in all cases we have a monodromy operator N associated to $\pi_2(X)$.

Theorem 7.5. *Let $\mathscr{X}/R$ be a minimal semistable model of an Enriques surface X, and Y its special fibre, which is a combinatorial Enriques surface. Then Y is of Type I, II or III respectively as the nilpotency index of N on $\pi_2(X)$ is* 1, 2 *or* 3.

Remark 7.6. (1) As noted in the introduction, a result very similar to this was proved in [Hernandez Mada 2015].

(2) The result as stated here is slightly different to Theorem 1.2. There are in fact two ways of stating it, one using the second homotopy group π_2 and one using the cohomology of a certain rank 2 local system V on X, given by pushing forward the constant sheaf on the K3 double cover of X.

Proof. If we let $\widetilde{\mathscr{X}}$ denote the canonical double cover of $\mathscr{X}$, with special fibre $\widetilde{Y}$ and generic fibre $\widetilde{X}$, then as remarked above, $\widetilde{X}$ is a smooth K3 surface over K, and $\widetilde{\mathscr{X}}$ is a minimal semistable model for $\widetilde{X}$. Hence $\widetilde{Y}$ is a combinatorial K3 surface, whose type can be deduced from the nilpotency index of the monodromy operator N on $H^2_{\text{ét}}(\widetilde{X}_{\bar{F}},\mathbb{Q}_\ell)$.

Now note that since $\widetilde{X}$ is simply connected, we have

$$\pi_2^{\text{ét}}(X_{\bar{F}})_{\mathbb{Q}_\ell} \cong \pi_2^{\text{ét}}(\widetilde{X}_{\bar{F}})_{\mathbb{Q}_\ell} \cong H_2^{\text{ét}}(\widetilde{X}_{\bar{F}},\mathbb{Q}_\ell) \cong H^2_{\text{ét}}(\widetilde{X}_{\bar{F}},\mathbb{Q}_\ell)^\vee$$

for all ℓ (including $\ell = p$ when $\mathrm{char}(F) = 0$), and the corresponding isomorphism holds by definition for $\pi_2^{\mathrm{rig}}(X/\mathscr{R}_K)$. Hence $\widetilde{Y}$ is of Type I, II or III respectively as the nilpotency index of N on $\pi_2(X)$ is 1, 2 or 3. It therefore suffices to show that the type of $\widetilde{Y}$ is the same as that of Y.

Note that we have a finite étale map $f : \widetilde{Y} \to Y$, therefore if $\widetilde{Y}$ is of Type I, that is a smooth K3 surface, then we must also have that Y is smooth, hence of Type I. If Y is not smooth, then let the components of Y be $Y_1,\dots,Y_N$, and the components of $\widetilde{Y}$ be $\widetilde{Y}_1,\dots,\widetilde{Y}_M$. After pulling back f to each component Y_i, one of two things can occur:

(1) $f^{-1}(Y_i)$ is irreducible, and we get a nontrivial 2-cover $\widetilde{Y}_j \to Y_i$;

(2) $f^{-1}(Y_i)$ splits into 2 disjoint components $\widetilde{Y}_j, \widetilde{Y}_{j'}$, each mapping isomorphically onto Y.

If $\widetilde{Y}$ is of Type III, then each component $\widetilde{Y}_j$ is rational, hence, since rational varieties are simply connected each component of Y is also rational, and Y is of Type III. If $\widetilde{Y}$ is of Type II, then one of two things can happen:

(1) $M > 2$ and there exists a component of $\widetilde{Y}$ which is an elliptic ruled surface.

(2) $M = 2$ and $\widetilde{Y} = \widetilde{Y}_1 \cup \widetilde{Y}_2$ consists of 2 rational surfaces meeting along an elliptic curve.

In the first case, one verifies that Y must also have a component isomorphic to an elliptic ruled surface (since a rational surface cannot be an unramified cover of an elliptic ruled surface), and is therefore of Type II. In the second case, Y must also have 2 components, (since otherwise Y, and therefore $\widetilde{Y}$, would be smooth), and each component of $\widetilde{Y}$ would be a nontrivial double cover of a component of Y. But since the components of Y are either rational or elliptic ruled, this cannot happen. □

8. Abelian surfaces

In order to deal with abelian surfaces, we need the following analogue of Nakkajima's result,

Theorem 8.1. *Let $Y^{\log}$ be a logarithmic abelian surface over k. Then the underlying scheme Y is a combinatorial abelian surface.*

Proof. We may assume that $k = \bar{k}$. We adapt the proof of Theorem II of [Kulikov 1977]. Let $Y_1, \ldots, Y_N$ denote the components of Y, $C_{ij} = Y_i \cap Y_j$ for $i \neq j$ the double curves, and $T_{C_{ij}}$ the number of triple points on each curve C_{ij}. We may assume that $N > 1$.

Note that $\omega_Y|_{Y_i} \cong \Omega^2_{Y_i/k}(\log \sum_{j\neq i} C_{ij}) \cong \mathscr{O}_{Y_i}$ and hence the divisor $K_{Y_i} + \sum_{j\neq i} C_{ij}$ on Y_i is principal, where K_{Y_i} is a canonical divisor on Y_i. Write $D_i = \sum_{j\neq i} C_{ij}$. Now applying Lemma 4.1 gives us the following possibilities for each (Y_i, D_i):

(1) Y_i is an elliptic ruled surface, and either:
   (a) $D_i = E_1 + E_2$ where E_1, E_2 are 2 nonintersecting rulings;
   (b) a $D_i = E$ is a single 2-ruling.

(2) Y_i is a rational surface, and either:
   (a) $D_i = E$ is an elliptic curve inside Y_i;
   (b) $D_i = C_1 + \cdots + C_d$ is a cycle of rational curves on Y_i.

First suppose that there is some i such that case (2)(b) happens. Then this must also occur on each neighbour of Y_i, and since Y is connected, it follows that this occurs on each component. The dual graph Γ is therefore a triangulation of a compact surface without border.

Write

$$Y^{(0)} = \coprod_i Y_i, \quad Y^{(1)} = \coprod_{i<j} C_{ij}, \quad Y^{(2)} = \coprod_{i<j<k} Y_i \cap Y_j \cap Y_k,$$

and consider the spectral sequence $H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) \Rightarrow H^{s+t}(Y, \mathscr{O}_Y)$ constructed in Section 3. Since the components Y_i and the curves C_{ij} are rational, it follows

that $H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) = 0$ for $t > 0$ (see for example, [Chatzistamatiou and Rülling 2011, Theorem 1]), and therefore that the coherent cohomology $H^i(Y, \mathscr{O}_Y)$ of Y is the same as the k-valued singular cohomology $H^i_{\text{sing}}(\Gamma, k)$ of Γ. But since $p \neq 2$, the k-Betti numbers $\dim_k H^i_{\text{sing}}(\Gamma, k)$ are the same as the $\mathbb{Q}$-Betti numbers $\dim_{\mathbb{Q}} H^i_{\text{sing}}(\Gamma, \mathbb{Q})$, the latter must therefore be $1, 2, 1$ and by the classification of closed 2-manifolds we can deduce that Γ is a torus.

Finally let us suppose that all the double curves C_{ij} are elliptic curves, so that each $T_{C_{ij}} = 0$ (see the proof of Lemma 4.1). Again examining the spectral sequence $H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) \Rightarrow H^{s+t}(Y, \mathscr{O}_Y)$ and using the fact that $\chi(E) = 0$ for an elliptic curve, we can see that $0 = \chi(Y) = \chi(Y^{(0)}) = \bigoplus_i \chi(Y_i)$. Since each Y_i is either rational ($\chi = 1$) or elliptic ruled ($\chi = 0$), it follows that each Y_i must be elliptic ruled, and we are in the case (1) above. The dual graph Γ is one dimensional, and since each component has on it at most two double curves, Γ is either a line segment or a circle.

If Γ were a line segment, then $Y = Y_1 \cup_{E_1} \cdots \cup_{E_{N-1}} Y_N$ would be a chain. Then birational invariance of coherent cohomology would imply that the maps

$$H^0(Y_i, \mathscr{O}_{Y_i}) \to H^0(E_i, \mathscr{O}_{E_i}), \quad H^0(Y_{i+1}, \mathscr{O}_{Y_{i+1}}) \to H^0(E_i, \mathscr{O}_{E_i}),$$
$$H^1(Y_i, \mathscr{O}_{Y_i}) \to H^1(E_i, \mathscr{O}_{E_i}), \quad H^1(Y_{i+1}, \mathscr{O}_{Y_{i+1}}) \to H^1(E_i, \mathscr{O}_{E_i}),$$

would be isomorphisms, and hence some basic linear algebra would imply surjectivity of the maps

$$H^0(Y^{(0)}, \mathscr{O}_{Y^{(0)}}) \to H^0(Y^{(1)}, \mathscr{O}_{Y^{(1)}}), \quad H^1(Y^{(0)}, \mathscr{O}_{Y^{(0)}}) \to H^1(Y^{(1)}, \mathscr{O}_{Y^{(1)}}).$$

Also, we would have $\dim_k H^1(Y^{(0)}, \mathscr{O}_{Y^{(0)}}) = N$ and $\dim_k H^1(Y^{(1)}, \mathscr{O}_{Y^{(1)}}) = N - 1$, so again examining the spectral sequence $H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) \Rightarrow H^{s+t}(Y, \mathscr{O}_Y)$ would imply that $\dim_k H^1(Y, \mathscr{O}_Y) = 1$. Since we know that in fact $\dim_k H^1(Y, \mathscr{O}_Y) = 2$ (by the definition of a logarithmic abelian surface), this cannot happen. Hence Γ must be a circle and Y is of Type II. □

Corollary 8.2. *Let $\mathscr{X}/R$ be a minimal semistable model of an abelian surface X/F. Then the special fibre Y is a combinatorial abelian surface.*

If X/F is an abelian surface, then for any prime $\ell \neq p$ we consider the quasiunipotent G_F-representation $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_\ell)$. For $\ell = p$ and $\text{char}(F) = 0$ we may also consider the de Rham representation $H^2_{\text{ét}}(X_{\overline{K}}, \mathbb{Q}_p)$, and when $\text{char}(F) = p = \ell$ the (φ, ∇)-module $H^2_{\text{rig}}(X/\mathscr{R}_K)$. Again letting $H^2(X)$ stand for any of the above second cohomology groups then, in each case, we have a nilpotent monodromy operator N associated to $H^2(X)$.

Theorem 8.3. *Let $\mathscr{X}/R$ be a minimal semistable model for X, with special fibre Y. Then Y is combinatorial of Type I, II or III respectively as the nilpotency index of N on $H^2(X)$ is $1, 2$ or 3.*

Proof. We will treat the case $\ell \neq p$ and $\mathrm{char}(F) = 0$; the other cases are handled entirely similarly. Let $Y_1, \dots, Y_N$ be the smooth components of the special fibre Y. For any $I = \{i_1, \dots, i_n\}$ write $Y_I = \bigcap_{i \in I} Y_i$ and for any $s \geq 0$ write $Y^{(s)} = \coprod_{|I|=s+1} Y_I$; these are all smooth over k and empty if $s > 2$.

As in the proof of Theorem 6.4, we consider the weight spectral sequence

$$E_1^{s,t} = \bigoplus_{j \geq \max\{0,-s\}} H^{t-2j}_{\text{ét}}(Y_{\bar{k}}^{(s+2j)}, \mathbb{Q}_\ell)(-j) \Rightarrow H^{s+t}_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell).$$

As before it suffices to show the following:

(1) If Y is of Type I, then $E_2^{1,1} = 0$.

(2) If Y is of Type II, then $E_2^{1,1} \neq 0$ and $E_2^{2,0} = 0$.

(3) If Y is of Type III, then $E_2^{2,0} \neq 0$.

Again, the first of these is trivial, and in both the Type II and III cases the term $E_2^{2,0}$ is the second singular cohomology $H^2_{\text{sing}}(\Gamma, \mathbb{Q}_\ell)$ of the dual graph Γ. It therefore suffices to show that if Y is of Type II, then $E_2^{1,1} \neq 0$.

To show this, note that we have $\dim_{\mathbb{Q}_\ell} E_2^{i,0} = \dim_{\mathbb{Q}_\ell} H^i_{\text{sing}}(\Gamma, \mathbb{Q}_\ell)$, which is 1 for $i = 0, 1$ and zero otherwise. Hence we may deduce that $\dim_{\mathbb{Q}_\ell} E_2^{-1,2} = 1$, from the fact that $E_2^{-r,w+r} \xrightarrow{\sim} E_2^{r,w-r}$, and that $\dim_{\mathbb{Q}_\ell} E_2^{0,1} = 2$, from the fact that

$$\dim_{\mathbb{Q}_\ell} E_2^{-1,2} + \dim_{\mathbb{Q}_\ell} E_2^{0,1} + \dim_{\mathbb{Q}_\ell} E_2^{1,0} = \dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell) = 4.$$

If we write $Y = Y_1 \cup \cdots \cup Y_N$ as a union of N elliptic ruled surfaces, then $Y^{(1)}$ is a disjoint union of N elliptic curves. Hence by Lemma 4.2 we must have

$$\dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(Y_{\bar{k}}^{(0)}, \mathbb{Q}_\ell) = \dim_{\mathbb{Q}_\ell} H^1_{\text{ét}}(Y_{\bar{k}}^{(1)}, \mathbb{Q}_\ell) = 2N.$$

Hence $\dim_{\mathbb{Q}_\ell} E_2^{1,1} = \dim_{\mathbb{Q}_\ell} E_2^{0,1} = 2$ and therefore $E_2^{1,1} \neq 0$.

When $\ell = p$, the ℓ-adic weight spectral should be replaced by the p-adic one constructed by Mokrane [1993]. That this abuts to the p-adic étale cohomology when $\mathrm{char}(F) = 0$ follows from Matsumoto's [2015] extension of Fontaine's C_{st} conjecture to algebraic spaces, and that it abuts to the $\mathscr{R}_K$-valued rigid cohomology when $\mathrm{char}(F) = p$ follows from Proposition 2.3. □

9. Bielliptic surfaces

We can now complete our treatment of minimal models of surfaces of Kodaira dimension 0 by investigating what happens for bielliptic surfaces.

Theorem 9.1. *Let $Y^{\log}$ be a logarithmic bielliptic surface over k. Then the underlying scheme Y is a combinatorial bielliptic surface.*

Proof. We may assume $k = \bar{k}$. Let $\pi : \widetilde{Y}^{\log} \to Y^{\log}$ be the canonical m-cover associated to $\omega_{Y^{\log}}$. Then one easily checks that $\widetilde{Y}^{\log}$ is a logarithmic abelian surface

over k, and hence is combinatorial of Type I, II or III. If $\widetilde{Y}$ is of Type I, then $\widetilde{Y}$, and therefore Y, must be smooth over k, and hence Y is a smooth bielliptic surface over k, i.e., of Type I.

So assume that $\widetilde{Y}$ is of Type II or III. Let $\widetilde{Y}_1, \dots, \widetilde{Y}_M$ denote the components of $\widetilde{Y}$ and $Y_1, \dots, Y_N$ those of Y. Note that as in the proof of Theorem 8.1 we have

$$m(K_{Y_i} + \sum_{j \neq i} C_{ij}) = 0$$

in $\mathrm{Pic}(Y_i)$, where C_{ij} are the double curves.

Suppose that $\widetilde{Y}$ is of Type II. Note that each component of $\widetilde{Y}$ is finite étale over some component of Y, and hence each component of Y is an elliptic ruled surface. For each Y_i choose some $\widetilde{Y}_l \to Y_i$ finite étale, and let $\widetilde{C}_{lj}$ be the inverse image of the double curves. Then we have

$$K_{\widetilde{Y}_l} + \sum_j \widetilde{C}_{lj} = 0$$

in $\mathrm{Pic}(\widetilde{Y}_l)$. Applying Lemma 4.1 we can see that $\sum_j \widetilde{C}_{il}$ is either a single elliptic curve E, which is a 2-ruling on $\widetilde{Y}_l$, or two disjoint rulings E_1, E_2. Hence the same is true for $\sum_j C_{ij}$ on Y_i, and therefore Y is of Type II.

Finally, suppose that $\widetilde{Y}$ is of Type III. Then again, each component of $\widetilde{Y}$ is finite étale over some component of Y, hence all of the latter are rational. Since the Picard group of a rational surface is torsion free, it follows that we must have

$$K_{Y_i} + \sum_{j \neq i} C_{ij} = 0$$

on each Y_i. Hence applying Lemma 4.1 as in the proof of Theorem 8.1 it suffices to show that the dual graph Γ of Y is a triangulation of the Klein bottle. But now examining the spectral sequence

$$E_1^{s,t} := H^t(Y^{(s)}, \mathscr{O}_{Y^{(s)}}) \Rightarrow H^{s+t}(Y, \mathscr{O}_Y)$$

(where $Y^{(s)}$ is defined similarly to before), and using the fact that $\mathrm{char}(k) > 2$, we can see that the Betti numbers of Γ are the same as the dimensions of the coherent cohomology of Y, and therefore Y is of Type III. □

To formulate the analogue of Theorem 8.3 for bielliptic surfaces, we will need to construct a family of canonical local systems on our bielliptic surface X. Note that the torsion element $\omega_X \in \mathrm{Pic}(X)[m] \in H^1(X, \mu_m)$ gives rise to a μ_m-torsor over X, and hence a canonical $\mathbb{Q}$-valued permutation representation ρ of the fundamental group $\pi_1^{\text{ét}}(X, \bar{x})$, and we can use this to construct canonical ℓ- or p-adic local systems on X. When $\ell \neq p$ we obtain a continuous representation $\rho \otimes_{\mathbb{Q}} \mathbb{Q}_\ell$ of $\pi_1^{\text{ét}}(X, \bar{x})$ and hence a lisse ℓ-adic sheaf V_ℓ on X, and when $\ell = p$ and $\mathrm{char}(F) = 0$ we may

do the same to obtain a lisse p-adic sheaf V_p on X, and when $\ell = \mathrm{char}(F) = p$ we obtain an overconvergent F-isocrystal V_p on X/K using the construction of Section 2.

Then the local systems V_ℓ, V_p do not depend on the choice of point $\bar{x}$, and the G_F-representations $H^2_{\text{ét}}(X_{\bar{F}}, V_\ell)$ and $H^2_{\text{ét}}(X_{\bar{F}}, V_p)$ when $\mathrm{char}(F) = 0$ are quasiunipotent and de Rham respectively; we may also consider the (φ, ∇)-module

$$H^2_{\text{rig}}(X/\mathscr{R}_K, V_p) := H^2_{\text{rig}}(X/\mathscr{E}_K^\dagger, V_p) \otimes_{\mathscr{E}_K^\dagger} \mathscr{R}_K$$

over $\mathscr{R}_K$. Representing any of $H^2_{\text{ét}}(X_{\bar{F}}, V_\ell)$, $H^2_{\text{ét}}(X_{\bar{F}}, V_p)$ or $H^2_{\text{ét}}(X/\mathscr{R}_K, V_p)$ by $H^2(X, V)$, in all cases we obtain monodromy operators N associated to $H^2(X, V)$.

Remark 9.2. This construction might seem a little laboured, since what we are really constructing is simply the pushforward of the constant sheaf via the canonical abelian cover of X. The point of describing it in the above way is to emphasise the fact that the local systems V_ℓ, V_p are entirely intrinsic to X.

Theorem 9.3. *Let $\mathscr{X}/R$ be a minimal semistable model for X, with special fibre Y. Then Y is combinatorial of Type I, II or III respectively as the nilpotency index of N on $H^2(X, V)$ is* 1, 2 *or* 3.

Proof. The local systems V_ℓ, V_p are by construction such that there exists a finite étale cover $\widetilde{\mathscr{X}} \to \mathscr{X}$ Galois with group G, such that $\widetilde{\mathscr{X}}$ is a minimal model of an abelian surface $\widetilde{X}$ and $H^2(X, V) \cong H^2(\widetilde{X})$. The special fibre $\widetilde{Y}$ is therefore a finite étale cover of Y, also Galois with group G, and is a combinatorial abelian surface of Type I, II or III according to the nilpotency index of N on $H^2(X, V)$. Hence we must show that $\widetilde{Y}$ and Y have the same type; this was shown during the course of the proof of Theorem 9.1. □

10. Existence of models and abstract good reduction

As explained in the introduction, our results so far are essentially "one half" of the good reduction problem for surfaces with $\kappa = 0$, the other half consists of trying to actually find models nice enough to be able to apply the above methods.

Definition 10.1. Let X/F be a minimal surface of Kodaira dimension 0. Then we say that X admits potentially combinatorial reduction if after replacing F by a finite separable extension, there exists a minimal model $\mathscr{X}/R$ of X.

Then thanks to the results of the previous sections, for surfaces with potentially combinatorial reduction, we can describe the "type" of the reduction in terms of the nilpotency index of the monodromy operator on a suitable cohomology or homotopy group of X (either ℓ-adic or p-adic). We can therefore answer questions of "abstract reduction" type by establishing whether or not surfaces have potentially

combinatorial reduction. The strongest result one might hope for is that every such surface has potentially combinatorial reduction. Unfortunately, this is not the case.

Example 10.2 [Liedtke and Matsumoto 2016, Theorem 2.8]. There exist Enriques surfaces over $\mathbb{Q}_p$ which do not admit potentially combinatorial reduction.

This can in fact be seen already in the complex analytic case of a degenerating family $\mathscr{X} \to \Delta$ of Kähler manifolds over a disc (see [Persson 1977, Appendix 2]). In Proposition 2.1 of [Liedtke and Matsumoto 2016] it is shown that if a K3 surface over F admits potentially strictly semistable reduction, then it admits potentially combinatorial reduction. Again, while the former is always conjectured, it can only be proved under certain conditions, see Corollary 2.2 of [loc. cit.]. Since we know that abelian surfaces admit potentially strictly semistable reduction, we can use their argument to prove the following.

Theorem 10.3. *Abelian surfaces X/F admit potentially combinatorial reduction.*

Proof. By Theorem 4.6 of [Künnemann 1998], after replacing F by a finite separable extension, we may assume that there exists a strictly semistable scheme model $\mathscr{X}/R$ of X. By applying the minimal model program of [Kawamata 1994] there exists another scheme model $\mathscr{X}'$ for X such that:

(1) the components of the special fibre of $\mathscr{X}'$ are geometrically normal and integral $\mathbb{Q}$-Cartier divisors on $\mathscr{X}'$;

(2) $\mathscr{X}'$ is regular away from a finite set Σ of closed points on its special fibre, and $\mathscr{X}'$ has only terminal singularities at these points;

(3) the special fibre is a normal crossings divisor on $\mathscr{X} \setminus \Sigma$;

(4) the relative canonical Weil divisor $K_{\mathscr{X}'/R}$ is $\mathbb{Q}$-Cartier and n.e.f. relative to R.

Now, since the canonical divisor K_X on the generic fibre is trivial, it follows that we may write $K_{\mathscr{X}'/R}$ as a linear combination $\sum_i a_i V_i$ of the components of the special fibre Y' of $\mathscr{X}'$. Moreover since $\sum_i V_i = 0$ we may in fact assume that $a_i \leq 0$ for all i and $a_i = 0$ for some i. Since $K_{\mathscr{X}'/\mathbb{Q}}$ is n.e.f. relative to R, arguing as in Lemma 4.7 of [Maulik 2014] shows that in fact we must have $a_i = 0$ for all i, and hence $K_{\mathscr{X}'/R} = 0$. In particular it is Cartier (not just $\mathbb{Q}$-Cartier) and therefore applying Theorem 4.4 of [Kawamata 1994] we can see that in fact $\mathscr{X}'$ is strictly semistable away from a finite set of isolated rational double points on components of Y'.

Finally, applying Theorem 2.9.2 of [Saito 2004] and Theorem 2 of [Artin 1974] we may, after replacing F by a finite separable extension, find a strictly semistable algebraic space model $\mathscr{X}''/R$ for X and a birational morphism $\mathscr{X}'' \to \mathscr{X}'$ which is an isomorphism outside a closed subset of each special fibre, of codimension ≥ 2 in the total space. Since we know that $K_{\mathscr{X}'/R} = 0$, it follows that $K_{\mathscr{X}''/R} = 0$, and therefore $\mathscr{X}''$ is a minimal model in the sense of Definition 5.1. □

Remark 10.4. Of course, this begs the question as to whether or not bielliptic surfaces admit potentially combinatorial reduction; we are not sure whether to expect this or not.

Finally, we would like to relate the "type" of combinatorial reduction for abelian (and hence bielliptic) surfaces to the more traditional invariants associated to abelian varieties with semiabelian reduction. So suppose that we have an abelian surface X/F. Then after a finite separable extension, we may assume that X admits the structure of an abelian variety over F; let us therefore call it A instead. After making a further extension, we may assume that A has semiabelian reduction, i.e., there exist a semiabelian scheme over R whose generic fibre is A. In this situation we have a "uniformisation cross" for A (see for example [Coleman and Iovita 1999, §2]), which is a diagram

$$\begin{array}{ccccc} & & T & & \\ & & \downarrow & & \\ \Gamma & \longrightarrow & G & \xrightarrow{\pi} & A \\ & & \downarrow & & \\ & & B & & \end{array}$$

where T is a torus over F, B is an abelian variety with good reduction, G is an extension of B by T and Γ is a discrete group. Fixing a prime $\ell \neq p$, the monodromy operator on $H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$ can be defined as follows. We have an exact sequence

$$0 \to \mathrm{Hom}(\Gamma, \mathbb{Q}_\ell) \to H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell) \to H^1_{\text{ét}}(G_{\bar{F}}, \mathbb{Q}_\ell) \to 0$$

and a nondegenerate pairing

$$\Gamma \times \mathrm{Hom}(T, \mathbb{G}_m) \to \mathbb{Q}$$

and the monodromy operator on $H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$ is then the composition

$$H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell) \to H^1_{\text{ét}}(T_{\bar{F}}, \mathbb{Q}_\ell) \to \mathrm{Hom}(T, \mathbb{G}_m) \otimes_{\mathbb{Z}} \mathbb{Q}_\ell \to \mathrm{Hom}(\Gamma, \mathbb{Q}_\ell) \to H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$$

(see for example [Coleman and Iovita 1999]). Since the first map is surjective, the last injective, and all others are isomorphisms, we have that the dimension of the image of monodromy on $H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$ is equal to the dimension of $H^1_{\text{ét}}(T_{\bar{F}}, \mathbb{Q}_\ell)$, and therefore to the rank of T. Using some simple linear algebra, one can therefore give the nilpotency index of N on $H^2_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell) = \bigwedge^2 H^1_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$ as follows:

(1) $\mathrm{rank}(T) = 0 \Longrightarrow N = 0$ on $H^2_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$;

(2) $\mathrm{rank}(T) = 1 \Longrightarrow N \neq 0$, $N^2 = 0$ on $H^2_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$;

(3) $\mathrm{rank}(T) = 2 \Longrightarrow N^2 \neq 0$, $N^3 = 0$ on $H^2_{\text{ét}}(A_{\bar{F}}, \mathbb{Q}_\ell)$.

Hence we have the following.

Proposition 10.5. *A has potentially combinatorial reduction of Type I, II or III as* $\operatorname{rank}(T)$ *is* 0, 1 *or* 2 *respectively.*

11. Towards higher dimensions

In this final section of the article, we begin to investigate the shape of degenerations in higher dimensions, in particular looking at Calabi–Yau threefolds and concentrating on the "maximal intersection case", analogous to the Type III degeneration of K3 surfaces. In characteristic 0 some fairly general results in this direction are proved in [Kollár and Xu 2016], and the approach there provides much of the inspiration for the main result of this section, Theorem 11.5, as well as some of the key ingredients of its proof. Many of the proofs there rely on results from the log minimal model program (LMMP), which happily has recently been solved for threefolds in characteristics > 5 [Hacon and Xu 2015; Birkar 2016; Birkar and Waldron 2016]. Given these results, many of our proofs consist of working through special low dimensional cases of [Kollár and Xu 2016] explicitly (and gaining slightly more information than given there), although there are certain places where specifically characteristic p arguments are needed.

Since we will need to use the LMMP for threefolds, we will assume throughout that $p > 5$. Unfortunately, since we will also need to know results on the homotopy type of Berkovich spaces, we will also need to assume that our models are in fact schemes, rather than algebraic spaces.

Definition 11.1. A Calabi–Yau variety over F is a smooth, projective, geometrically connected variety X/F such that:

- the canonical sheaf $\omega_X = \Omega^{\dim X}_{X/F}$ is trivial, i.e., $\omega_X \cong \mathscr{O}_X$;
- X is geometrically simply connected, i.e., $\pi_1^{\text{ét}}(X_{\bar{F}}, x) = \{1\}$ for any $x \in X(\bar{F})$;
- $H^i(X, \mathscr{O}_X) = 0$ for all $0 < i < \dim X$.

In dimension 2 these are exactly the K3 surfaces, and we will be interested in what we can say about degenerations of Calabi–Yau varieties in dimension 3. Here one expects to be able to divide "suitably nice" semistable degenerations into 4 "types" depending on the nilpotency index of N acting on $H^3(X)$ (for some suitable Weil cohomology theory). In this section we will treat the "Type IV" situation.

Definition 11.2. We say that a morphism $f : X \to S$ of algebraic varieties (over an algebraically closed field) is a Mori fibre space if it is projective with connected fibres, and the anticanonical divisor $-K_X$ is f-ample, i.e., ample on all fibres of f.

Definition 11.3. Let $Y = \bigcup_i V_i$ be a simple normal crossings variety over k of dimension 3. We say that Y is a combinatorial Calabi–Yau of Type IV if geometrically (i.e., over $\bar{k}$) we have:

- each component V_i is birational to a Mori fibre space over a unirational base;
- each connected component of every double surface S_{ij} is rational;
- each connected component of every triple curve C_{ijk} is rational;
- the dual graph Γ of Y is a triangulation of the 3-sphere S^3.

Remark 11.4. (1) It is worth noting that in characteristic 0 these conditions imply that V_i is rationally connected, and the analogue of the condition in dimension 2 implies rationality, even in characteristic p.

(2) We may in fact assume that we have the above shape after a finite extension of k.

Let $H^3(X)$ stand for either $H^3_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$ if $\text{char}(F) = 0$ or $\ell \neq p$, or $H^3_{\text{rig}}(X/\mathscr{R}_K)$ if $\text{char}(F) = p$. In all cases, we have a natural monodromy operator N acting on $H^3(X)$, such that $N^4 = 0$. As a first step in the study of Calabi–Yau degenerations in dimension 3, the main result of this section is the following.

Theorem 11.5. *Let $\mathscr{X}$ be a strictly semistable R-scheme with generic fibre X a Calabi–Yau threefold. Assume moreover that the sheaf of logarithmic 3-forms $\omega_{\mathscr{X}}$ on $\mathscr{X}$ relative to R is trivial, and that $N^3 \neq 0$ on $H^3(X)$. Then the special fibre Y of $\mathscr{X}$ is a combinatorial Calabi–Yau of Type IV.*

As before, we will only treat the case $\text{char}(F) = 0$ and $\ell \neq p$; the others are handled identically. We may also assume that $k = \bar{k}$. Let V_i denote the components of Y, S_{ij} the double surfaces, C_{ijk} the triple curves and P_{ijkl} the quadruple points. Write $Y^{(0)} = \coprod_i V_i$, $Y^{(1)} = \coprod_{ij} S_{ij}$ et cetera. The only point where the hypothesis on the nilpotency index of N is used is to prove the following lemma.

Lemma 11.6. *Suppose that $N^3 \neq 0$. Then Y has "maximal intersection", i.e., there exists a quadruple point P_{ijkl}.*

Proof. If there is no quadruple point P_{ijkl} then $Y^{(3)} = \varnothing$. Let W_n denote the weight filtration on $H^3_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$, so that $W_{-1} = 0$ and $W_6 = H^3_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)$. The monodromy operator N^3 sends W_i into W_{i-6}, in particular $N^3(H^3_{\text{ét}}(X_{\bar{F}}, \mathbb{Q}_\ell)) \subset W_0$. But $Y^{(3)} = \varnothing$ implies that $W_0 = 0$ and hence $N^3 = 0$. □

Note that we do not need to know the weight-monodromy conjecture in order for the lemma to hold, we simply need to know compatibility of N with the weight filtration.

For each i we will let $D_i = \sum_{j \neq i} S_{ij}$, so that by the assumption $\omega_{\mathscr{X}} \cong \mathscr{O}_{\mathscr{X}}$ and the adjunction formula we have $-K_{V_i} = D_i$ for all i. Similarly setting $E_{ij} = \sum_{k \neq i,j} C_{ijk}$ we obtain $-K_{S_{ij}} = E_{ij}$ and setting $F_{ijk} = \sum_{l \neq ijk} P_{ijkl}$ we can see that $-K_{C_{ijk}} = F_{ijk}$. The lemma shows that there exists some V_i containing a quadruple point, and the first key step in proving Theorem 11.5 is showing that this is actually true for every i. The main ingredient in this is the following.

Proposition 11.7. *Let (V, D) be a pair consisting of a smooth projective threefold V over $\bar{k}$ and a nonempty strict normal crossings divisor $D \subset V$. Assume that $K_V + D = 0$, and that D is disconnected. Then D consists of two disjoint irreducible components D_1 and D_2.*

Remark 11.8. The corresponding result for surfaces follows from Lemma 4.1.

Proof. The characteristic 0 version of this result is Proposition 4.37 of [Kollár 2013]. However, thanks to the proof of the minimal model program for threefolds in characteristic $p > 5$, in particular the connectedness principle and the existence of Mori fibre spaces in [Birkar 2016; Birkar and Waldron 2016], the same proof works here. So we will run the MMP on the smooth 3-fold V. It follows from Theorem 1.7 of [Birkar and Waldron 2016] that this terminates in a Mori fibre space $p : V^* \to S$, and by the connectedness principle ([Birkar 2016, Theorem 1.8]) it suffices to prove that the strict transform $D^* \subset V^*$ consists of 2 irreducible components. Now we simply follow the proof of Proposition 4.37 of [Kollár 2013], which goes as follows.

We know that there exists some component $D_1^* \subset D^*$ which positively intersects the ray contracted by p. Choose another component $D_2^* \subset D^*$ disjoint from D_1^*, and choose some fibre F_s of p meeting D_2^*. Since D_2^* is disjoint from D_1^*, it follows that it cannot contain F_s, and hence intersects F_s positively. Hence both D_1^* and D_2^* are p-ample, intersecting the contracted ray positively. Hence the generic fibre of p is of dimension 1, and is a regular (not necessarily smooth) Fano curve. It then follows that if we choose a general fibre F_g of p, then $D_i^* \cdot F_g = 1$ for $i = 1, 2$ and all other components of D^* are p-vertical, hence trivial as claimed. □

Corollary 11.9. *Every component of Y contains a quadruple point.*

Proof. By connectedness of Y it suffices to show that each neighbour of V_i also contains a quadruple point. Note that by Proposition 11.7 the divisor D_i is connected, by hypothesis there exists a double surface S_{ij} in D_i containing a quadruple point, and hence it suffices to show that each double surface S_{ik} meeting S_{ij} contains a quadruple point. But if not, then C_{ijk} would form a connected component of E_{ij} and hence again applying Lemma 4.1 we would see that S_{ij} could not contain a quadruple point. Therefore S_{ik} must contain a quadruple point, and we are done. □

Of course this also shows that each double surface S_{ij} contains a quadruple point, hence by repeatedly applying Lemma 4.1 we can conclude that each surface S_{ij} and each curve C_{ijk} is rational. We may therefore see as in the proof of Theorem 8.1 that the dual graph of each D_i is a closed 2-manifold. Moreover, applying the MMP to each V_i produces a Mori fibre space $W_i \to Z_i$, such that the divisor $D_i = \sum_{j \neq i} S_{ij}$ dominates Z_i. Therefore V_i has the form described in Definition 11.3.

Finally, to show that the dual graph Γ is a 3-sphere, we consider, for every vertex γ corresponding to a component V_i of Y, the "star" of γ, i.e., the subcomplex of Γ

consisting of those cells meeting γ. This is a cone over the dual graph of D_i, hence Γ is a closed 3-manifold.

Proposition 11.10. *The dual graph Γ is simply connected.*

Proof. Let $\mathbb{C}_p$ denote the completion of the algebraic closure of F, and $\mathscr{O}_{\mathbb{C}_p}$ its ring of integers. Let $\mathfrak{X}$ denote the base change to $\mathscr{O}_{\mathbb{C}_p}$ of the π-adic completion of $\mathscr{X}$, this $\mathfrak{X}$ is polystable over $\mathscr{O}_{\mathbb{C}_p}$ in the sense of Definition 1.2 of [Berkovich 1999]. Let $X^{\mathrm{an}}_{\mathbb{C}_p}$ denote the generic fibre of $\mathfrak{X}$, considered as a Berkovich space, or in other words the analytification of the base change of X to $\mathbb{C}_p$.

Let $\pi_1^{\text{ét}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ denote the étale fundamental group of $X^{\mathrm{an}}_{\mathbb{C}_p}$ in the sense of [de Jong 1995], and by $\pi_1^{\mathrm{top}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ the fundamental group of the underlying topological space of $X^{\mathrm{an}}_{\mathbb{C}_p}$. Theorem 2.10(iii) of [de Jong 1995] together with rigid analytic GAGA shows that the profinite completion of $\pi_1^{\text{ét}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ is trivial, since it is isomorphic to the algebraic étale fundamental group $\pi_1^{\text{ét}}(X_{\mathbb{C}_p})$ of $X_{\mathbb{C}_p}$, and X is Calabi–Yau. Next, by Remark 2.11 of [de Jong 1995] together with Theorem 9.1 of [Berkovich 1999] we have a surjection $\pi_1^{\text{ét}}(X^{\mathrm{an}}_{\mathbb{C}_p}) \to \pi_1^{\mathrm{top}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ and hence the profinite completion of $\pi_1^{\mathrm{top}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ is trivial.

Now by Theorem 8.2 of [Berkovich 1999] we have $\pi_1(\Gamma) \cong \pi_1^{\mathrm{top}}(X^{\mathrm{an}}_{\mathbb{C}_p})$ and hence the profinite completion of $\pi_1(\Gamma)$ is trivial. Since Γ is a 3-manifold, we may finally apply [Hempel 1987] to conclude that $\pi_1(\Gamma)$ is trivial as claimed. □

We may now conclude the proof of Theorem 11.5 using the Poincaré conjecture. In fact, if we know that the weight monodromy conjecture holds, then we have the following converse.

Proposition 11.11. *Let $\mathscr{X}$ be a strictly semistable R-scheme whose generic fibre is a Calabi–Yau threefold X, such that $\omega_{\mathscr{X}} \cong \mathscr{O}_{\mathscr{X}}$. Assume that the special fibre Y is a combinatorial Calabi–Yau of Type IV. If the weight monodromy conjecture holds for $H^3(X)$, then $N^3 \neq 0$.*

Proof. Again, we assume that $\ell \neq p$; the other cases are handled similarly. Consider the weight spectral sequence $E_r^{p,q}$ for $\mathscr{X}$. The hypotheses imply that N^3 induces an isomorphism

$$N^3 : E_2^{-3,6} \to E_2^{3,0}$$

and to show that $N^3 \neq 0$ it therefore suffices to show that $E_2^{3,0} \neq 0$. Writing out the weight spectral sequence explicitly we see that we have an isomorphism

$$E_2^{3,0} \cong H^3_{\mathrm{sing}}(\Gamma, \mathbb{Q}_\ell),$$

where $\Gamma \simeq S^3$ is the dual graph of Y, and hence the claim follows. □

This is in particular the case if $\mathrm{char}(F) = p$ (when $\ell \neq p$ this is [Ito 2005], when $\ell = p$ it is [Lazda and Pál 2016, Chapter 5]) or $\mathrm{char}(F) = 0$, $\ell \neq p$ and X

is a complete intersection in some projective space (which follows from [Scholze 2012]).

Acknowledgements

B. Chiarellotto was supported by the grant MIUR-PRIN 2010-11 "Arithmetic Algebraic Geometry and Number Theory". C. Lazda was supported by a Marie Curie fellowship of the Istituto Nazionale di Alta Matematica. Both authors would like to thank the anonymous referee for a careful reading of the paper, and for suggesting several important improvements.

References

[Artin 1974] M. Artin, "Algebraic construction of Brieskorn's resolutions", *J. Algebra* **29** (1974), 330–348. MR Zbl

[Artin and Mazur 1969] M. Artin and B. Mazur, *Etale homotopy*, Lecture Notes in Mathematics **100**, Springer, Berlin, 1969. MR Zbl

[Belmans et al. 2005–] P. Belmans, A. J. de Jong, et al., "The Stacks project", electronic reference, 2005–, Available at http://stacks.math.columbia.edu.

[Berkovich 1999] V. G. Berkovich, "Smooth p-adic analytic spaces are locally contractible", *Invent. Math.* **137**:1 (1999), 1–84. MR Zbl

[Birkar 2016] C. Birkar, "Existence of flips and minimal models for 3-folds in char p", *Ann. Sci. Éc. Norm. Supér.* (4) **49**:1 (2016), 169–212. MR Zbl

[Birkar and Waldron 2016] C. Birkar and J. Waldron, "Existence of Mori fibre spaces for 3-folds in char p", preprint, 2016. arXiv

[Chatzistamatiou and Rülling 2011] A. Chatzistamatiou and K. Rülling, "Higher direct images of the structure sheaf in positive characteristic", *Algebra Number Theory* **5**:6 (2011), 693–775. MR Zbl

[Chiarellotto and Tsuzuki 2014] B. Chiarellotto and N. Tsuzuki, "Clemens–Schmid exact sequence in characteristic p", *Math. Ann.* **358**:3-4 (2014), 971–1004. MR Zbl

[Coleman and Iovita 1999] R. Coleman and A. Iovita, "The Frobenius and monodromy operators for curves and abelian varieties", *Duke Math. J.* **97**:1 (1999), 171–215. MR Zbl

[Hacon and Xu 2015] C. D. Hacon and C. Xu, "On the three dimensional minimal model program in positive characteristic", *J. Amer. Math. Soc.* **28**:3 (2015), 711–744. MR Zbl

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR Zbl

[Hempel 1987] J. Hempel, "Residual finiteness for 3-manifolds", pp. 379–396 in *Combinatorial group theory and topology* (Alta, UT, 1984), edited by S. M. Gersten and J. R. Stallings, Ann. of Math. Stud. **111**, Princeton Univ. Press, 1987. MR Zbl

[Hernandez Mada 2015] G. Hernandez Mada, *Monodromy criterion for the good reduction of surfaces*, Ph.D. thesis, Università Degli Studi di Padova, Concordia University, Université de Bordeaux, 2015, Available at http://paduaresearch.cab.unipd.it/8812/.

[Ito 2005] T. Ito, "Weight-monodromy conjecture over equal characteristic local fields", *Amer. J. Math.* **127**:3 (2005), 647–658. MR Zbl

[de Jong 1995] A. J. de Jong, "Étale fundamental groups of non-Archimedean analytic spaces", *Compositio Math.* **97**:1-2 (1995), 89–118. MR Zbl

[Kato 1996] F. Kato, "Log smooth deformation theory", *Tohoku Math. J.* (2) **48**:3 (1996), 317–354. MR Zbl

[Kawamata 1994] Y. Kawamata, "Semistable minimal models of threefolds in positive or mixed characteristic", *J. Algebraic Geom.* **3**:3 (1994), 463–491. MR Zbl

[Kedlaya 2004] K. S. Kedlaya, "A p-adic local monodromy theorem", *Ann. of Math.* (2) **160**:1 (2004), 93–184. MR Zbl

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics **200**, Cambridge University Press, 2013. MR Zbl

[Kollár and Xu 2016] J. Kollár and C. Xu, "The dual complex of Calabi–Yau pairs", *Invent. Math.* **205**:3 (2016), 527–557. MR Zbl

[Kulikov 1977] V. S. Kulikov, "Degenerations of K3 surfaces and Enriques surfaces", *Izv. Akad. Nauk SSSR Ser. Mat.* **41**:5 (1977), 1008–1042, 1199. In Russian; translated in *Math. USSR-Izv.* **11**:5 (1977), 957–989. MR Zbl

[Künnemann 1998] K. Künnemann, "Projective regular models for abelian varieties, semistable reduction, and the height pairing", *Duke Math. J.* **95**:1 (1998), 161–212. MR Zbl

[Lazda and Pál 2016] C. Lazda and A. Pál, *Rigid cohomology over Laurent series fields*, Algebra and Applications **21**, Springer, Cham, 2016. MR Zbl

[Liedtke and Matsumoto 2016] C. Liedtke and Y. Matsumoto, "Good reduction of K3 surfaces", preprint, 2016. arXiv

[Marmora 2008] A. Marmora, "Facteurs epsilon p-adiques", *Compos. Math.* **144**:2 (2008), 439–483. MR Zbl

[Matsumoto 2015] Y. Matsumoto, "Good reduction criterion for K3 surfaces", *Math. Z.* **279**:1-2 (2015), 241–266. MR Zbl

[Maulik 2014] D. Maulik, "Supersingular K3 surfaces for large primes", *Duke Math. J.* **163**:13 (2014), 2357–2425. MR Zbl

[Mokrane 1993] A. Mokrane, "La suite spectrale des poids en cohomologie de Hyodo–Kato", *Duke Math. J.* **72**:2 (1993), 301–337. MR Zbl

[Nakkajima 2000] Y. Nakkajima, "Liftings of simple normal crossing log K3 and log Enriques surfaces in mixed characteristics", *J. Algebraic Geom.* **9**:2 (2000), 355–393. MR Zbl

[Nakkajima 2006] Y. Nakkajima, "Signs in weight spectral sequences, monodromy-weight conjectures, log Hodge symmetry and degenerations of surfaces", *Rend. Sem. Mat. Univ. Padova* **116** (2006), 71–185. MR Zbl

[Olsson 2007] M. C. Olsson, *Crystalline cohomology of algebraic stacks and Hyodo–Kato cohomology*, Astérisque **316**, Société Mathématique de France, Paris, 2007. MR Zbl

[Pérez Buendía 2014] J. R. Pérez Buendía, *A crystalline criterion for good reduction on semi-stable K3-surfaces over a p-adic field*, Ph.D. thesis, Concordia University, 2014, Available at http://spectrum.library.concordia.ca/978195/.

[Persson 1977] U. Persson, *On degenerations of algebraic surfaces*, Mem. Amer. Math. Soc. **189**, 1977. MR Zbl

[Saito 2004] T. Saito, "Log smooth extension of a family of curves and semi-stable reduction", *J. Algebraic Geom.* **13**:2 (2004), 287–321. MR Zbl

[Scholze 2012] P. Scholze, "Perfectoid spaces", *Publ. Math. Inst. Hautes Études Sci.* **116** (2012), 245–313. MR Zbl

[Tsuji 1999] T. Tsuji, "Poincaré duality for logarithmic crystalline cohomology", *Compositio Math.* **118**:1 (1999), 11–41. MR Zbl

Communicated by Kiran S. Kedlaya
Received 2016-04-26 Revised 2016-07-28 Accepted 2016-09-05

chiarbru@math.unipd.it *Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova, Via Trieste, 63, I-35121 Padova, Italy*

lazda@math.unipd.it *Dipartimento di Matematica "Tullio Levi-Civita", Università degli Studi di Padova, Via Trieste, 63, I-35121 Padova, Italy*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2267

The Voronoi formula and double Dirichlet series

Eren Mehmet Kıral and Fan Zhou

We prove a Voronoi formula for coefficients of a large class of L-functions including Maass cusp forms, Rankin–Selberg convolutions, and certain noncuspidal forms. Our proof is based on the functional equations of L-functions twisted by Dirichlet characters and does not directly depend on automorphy. Hence it has wider application than previous proofs. The key ingredient is the construction of a double Dirichlet series.

1. Introduction

A Voronoi formula is an identity involving Fourier coefficients of automorphic forms, with the coefficients twisted by additive characters on either side. A history of the Voronoi formula can be found in [Miller and Schmid 2004]. Since its introduction in [loc. cit.], the Voronoi formula on GL(3) of Miller and Schmid has become a standard tool in the study of L-functions arising from GL(3), and has found important applications such as those in [Blomer 2012; Blomer et al. 2013; Khan 2012; Li 2009; 2011; Li and Young 2012; Miller 2006; Munshi 2013; 2015]. As of yet the general GL(N) formula has had fewer applications, a notable one being found in [Kowalski and Ricotta 2014].

The first proof of a Voronoi formula on GL(3) was found by Miller and Schmid [2006] using the theory of automorphic distributions. Later, a Voronoi formula was established for GL(N) with $N \geq 4$ in [Goldfeld and Li 2006; 2008; Miller and Schmid 2011], with [Miller and Schmid 2011] being more general and earlier than [Goldfeld and Li 2008] (see the addendum there). Goldfeld and Li's proof [2008] is more akin to the classical proof in GL(2) [Good 1981], obtaining the associated Dirichlet series through a shifted "vertical" period integral and making use of automorphy. An adelic version was established by Ichino and Templier [2013], allowing ramifications and applications to number fields. Another direction

MSC2010: primary 11F30; secondary 11F68, 11L05.
Keywords: Voronoi formula, automorphic form, Maass form, multiple Dirichlet series, Gauss sum, Kloosterman sum, Rankin–Selberg L-function.

of generalization with more complicated additive twists on either side has been considered in an unpublished work of Li and Miller and in [Zhou 2016].

In this article, we prove a Voronoi formula for a large class of automorphic objects or L-functions, including cusp forms for $\mathrm{SL}(N,\mathbb{Z})$, Rankin–Selberg convolutions, and certain noncuspidal forms. Previous works [Miller and Schmid 2011; Goldfeld and Li 2008; Ichino and Templier 2013] do not offer a Voronoi formula for Rankin–Selberg convolutions or noncuspidal forms. Even for Maass cusp forms, our new proof is shorter than any previous one, and uses a completely different set of techniques.

Let us briefly summarize our method of proof. We first reduce the statement of a Voronoi formula to a formula involving Gauss sums of Dirichlet characters. We construct a complex function of two variables and write it as double Dirichlet series in two different ways by applying a functional equation. Using the uniqueness theorem of Dirichlet series, we get an identity between coefficients of these two double Dirichlet series. This leads us to the Voronoi formula with Gauss sums.

One of our key steps in obtaining the Voronoi formula is the use of functional equations of L-functions twisted by Dirichlet characters. The relationship between the Dirichlet twists and the additive twists was expected, but not fully understood, such as in [Duke and Iwaniec 1990; Goldfeld and Li 2006, Section 4; Buttcane and Khan 2015; Zhou 2016]. In these works, only prime modulus is dealt with, which is a significant restriction. Miller and Schmid [2006, Section 6] derived the functional equation of L-functions twisted by a Dirichlet character of prime conductor from the Voronoi formula. However there is a combinatorial difficulty in reversing this process, i.e., obtaining additive twists of general nonprime conductors from multiplicative ones, which was acknowledged in both [Miller and Schmid 2006, p. 430] and [Ichino and Templier 2013, p. 68]. The method presented here is able to overcome this difficulty by discovering an interlocking structure among a family of Voronoi formulas with different conductors.

Our proof of the Voronoi formula is complete for additive twists of all conductors, prime or not, and unlike [Ichino and Templier 2013], [Miller and Schmid 2006], or [Miller and Schmid 2011], does not depend directly on automorphy of the cusp forms. This fact allows us to apply our theorem to many conjectural Langlands functorial transfers. For example, the Rankin–Selberg convolutions (also called functorial products) for $\mathrm{GL}(m)\times\mathrm{GL}(n)$ are not yet known to be automorphic on $\mathrm{GL}(m\times n)$ in general. Yet we know the functional equations of $\mathrm{GL}(m)\times\mathrm{GL}(n)$ L-functions twisted by Dirichlet characters. Thus, our proof provides a Voronoi formula for the Rankin–Selberg convolutions on $\mathrm{GL}(m)\times\mathrm{GL}(n)$ (see Example 1.7). Voronoi formulas for these functorial cases are unavailable from [Goldfeld and Li 2008], [Miller and Schmid 2011] or [Ichino and Templier 2013]. In Theorem 1.3 we reformulate our Voronoi formula like the classical converse theorem of Weil, i.e.,

assuming every L-function twisted by a Dirichlet character is entire, has an Euler product (or satisfies Hecke relations), and satisfies the precise functional equations, then the Voronoi formula as in Theorem 1.1 is valid. We do not have to assume it is a standard L-function coming from a cusp form.

Furthermore, by Theorem 1.3, we obtain a Voronoi formula for certain noncuspidal forms, such as isobaric sums (see Example 1.8). This is not readily available from any previous work but it is believed (see [Miller and Schmid 2011, p. 176]) that one may derive a formula by using formulas on smaller groups through a possibly complicated procedure. Such complication does not occur in our method because we work directly with L-functions.

We first state the main results for Maass cusp forms. Denote

$$e(x) := \exp(2\pi i x)$$

for $x \in \mathbb{R}$. Let $N \geq 3$ be an integer. Let $a, n \in \mathbb{Z}$, $c \in \mathbb{N}$ and let

$$\boldsymbol{q} = (q_1, q_2, \dots, q_{N-2}) \quad \text{and} \quad \boldsymbol{d} = (d_1, d_2, \dots, d_{N-2})$$

be tuples of positive integers satisfying the divisibility conditions

$$d_1 | q_1 c, \quad d_2 \,\Big|\, \frac{q_1 q_2 c}{d_1}, \quad \dots, \quad d_{N-2} \,\Big|\, \frac{q_1 \cdots q_{N-2} c}{d_1 \cdots d_{N-3}}. \tag{1}$$

In this case, to simplify notation we set

$$\xi_i := \frac{q_1 \cdots q_i c}{d_1 \cdots d_i}.$$

Define the hyper-Kloosterman sum as

$$\mathrm{Kl}(a, n, c; \boldsymbol{q}, \boldsymbol{d}) = \sideset{}{^*}\sum_{x_1 \bmod \xi_1} \sideset{}{^*}\sum_{x_2 \bmod \xi_2} \cdots \sideset{}{^*}\sum_{x_{N-2} \bmod \xi_{N-2}} e\left(\frac{d_1 x_1 a}{c} + \frac{d_2 x_2 \overline{x_1}}{\xi_1} + \cdots + \frac{d_{N-2} x_{N-2} \overline{x_{N-3}}}{\xi_{N-3}} + \frac{n \overline{x_{N-2}}}{\xi_{N-2}}\right),$$

where $\sum^*$ indicates that the summation is over reduced residue classes, and $\overline{x_i}$ denotes the multiplicative inverse of x_i modulo ξ_i. When $N = 3$, $\mathrm{Kl}(a, n, c; q_1, d_1)$ becomes the classical Kloosterman sum $S(aq_1, n; \xi_1)$. For the degenerate case $N = 2$, we define $\mathrm{Kl}(a, n, c; \ , \ ) := e(an/c)$.

Let F be a Hecke–Maass cusp form for $\mathrm{SL}(N, \mathbb{Z})$ with the spectral parameters $(\lambda_1, \dots, \lambda_N) \in \mathbb{C}^N$. Let $A(m_1, \dots, m_{N-1})$, with $(m_1, \dots, m_{N-1}) \in \mathbb{N}^{N-1}$, be the Fourier–Whittaker coefficients of F normalized as $A(1, \dots, 1) = 1$. We refer to [Goldfeld 2006] for the definitions and the basic results of Maass forms for $\mathrm{SL}(N, \mathbb{Z})$.

The Fourier coefficients satisfy the Hecke relations

$$A(m_1 m_1', \dots, m_{N-1} m_{N-1}') = A(m_1, \dots, m_{N-1}) A(m_1', \dots, m_{N-1}') \tag{2}$$

if $(m_1 \cdots m_{N-1}, m_1' \cdots m_{N-1}') = 1$ is satisfied,

$$A(1, \dots, 1, n) A(m_{N-1}, \dots, m_1) = \sum_{\substack{d_0 \cdots d_{N-1} = n \\ d_1 | m_1, \dots, d_{N-1} | m_{N-1}}} A\left(\frac{m_{N-1} d_{N-2}}{d_{N-1}}, \dots, \frac{m_2 d_1}{d_2}, \frac{m_1 d_0}{d_1} \right), \tag{3}$$

and

$$A(n, 1, \dots, 1) A(m_1, \dots, m_{N-1}) = \sum_{\substack{d_0 \cdots d_{N-1} = n \\ d_1 | m_1, \dots, d_{N-1} | m_{N-1}}} A\left(\frac{m_1 d_0}{d_1}, \frac{m_2 d_1}{d_2}, \dots, \frac{m_{N-1} d_{N-2}}{d_{N-1}} \right). \tag{4}$$

The dual Maass form of F is denoted by $\widetilde{F}$. Let $B(*, \dots, *)$ be the Fourier–Whittaker coefficients of $\widetilde{F}$. These coefficients satisfy

$$B(m_1, \dots, m_{N-1}) = A(m_{N-1}, \dots, m_1). \tag{5}$$

Define the ratio of Gamma factors

$$G_\pm(s) := i^{-N\delta} \pi^{-N(1/2-s)} \prod_{j=1}^{N} \Gamma\left(\frac{\delta + 1 - s - \overline{\lambda_j}}{2} \right) \Gamma\left(\frac{\delta + s - \lambda_j}{2} \right)^{-1}, \tag{6}$$

where for even Maass forms, we define $\delta = 0$ in G_+ and $\delta = 1$ in G_-, and for odd Maass forms, we define $\delta = 1$ in G_+ and $\delta = 0$ in G_-. We refer to [Goldfeld 2006, Section 9.2] for the definition of even and odd Maass forms.

Theorem 1.1 (Voronoi formula on GL(N) of Miller and Schmid [2011]). *Let F be a Hecke–Maass cusp form with coefficients $A(*, \dots, *)$, and $G_\pm$ a ratio of Gamma factors as in (6). Let $c > 0$ be an integer and let a be any integer with $(a, c) = 1$. Denote by $\bar{a}$ the multiplicative inverse of a modulo c. Let the additively twisted Dirichlet series be given as*

$$L_q\left(s, F, \frac{a}{c}\right) = \sum_{n=1}^{\infty} \frac{A(q_{N-2}, \dots, q_1, n)}{n^s} e\left(\frac{\bar{a} n}{c}\right) \tag{7}$$

for $\Re(s) > 1$. This Dirichlet series has an analytic continuation to all $s \in \mathbb{C}$ and satisfies the functional equation

$$
\begin{aligned}
&L_{\boldsymbol{q}}(s, F, a/c)\\
&= \frac{G_+(s) - G_-(s)}{2} \sum_{d_1 \mid q_1 c} \sum_{d_2 \mid \frac{q_1 q_2 c}{d_1}} \cdots \sum_{d_{N-2} \mid \frac{q_1 \cdots q_{N-2} c}{d_1 \cdots d_{N-3}}}\\
&\qquad \sum_{n=1}^{\infty} \frac{A(n, d_{N-2}, \ldots, d_2, d_1)\, \mathrm{Kl}(a, n, c; \boldsymbol{q}, \boldsymbol{d})}{n^{1-s} c^{Ns-1} d_1 d_2 \cdots d_{N-2}} \frac{d_1^{(N-1)s} d_2^{(N-2)s} \cdots d_{N-2}^{2s}}{q_1^{(N-2)s} q_2^{(N-3)s} \cdots q_{N-2}^{s}}\\
&+ \frac{G_+(s) + G_-(s)}{2} \sum_{d_1 \mid q_1 c} \sum_{d_2 \mid \frac{q_1 q_2 c}{d_1}} \cdots \sum_{d_{N-2} \mid \frac{q_1 \cdots q_{N-2} c}{d_1 \cdots d_{N-3}}}\\
&\qquad \sum_{n=1}^{\infty} \frac{A(n, d_{N-2}, \ldots, d_2, d_1)\, \mathrm{Kl}(a, -n, c; \boldsymbol{q}, \boldsymbol{d})}{n^{1-s} c^{Ns-1} d_1 d_2 \cdots d_{N-2}} \frac{d_1^{(N-1)s} d_2^{(N-2)s} \cdots d_{N-2}^{2s}}{q_1^{(N-2)s} q_2^{(N-3)s} \cdots q_{N-2}^{s}}, \qquad (8)
\end{aligned}
$$

in the region of convergence of the expression on the right-hand side ($\Re(s) < 0$).

The traditional Voronoi formula, involving weight functions instead of Dirichlet series, is obtained after taking an inverse Mellin transform against a suitable test function.

Choose a Dirichlet character χ modulo c, which is not necessarily primitive, multiply both sides of (8) by $\chi(a)$, and sum this equality over the reduced residue system modulo c. We obtain the following Voronoi formula with Gauss sums. In Section 3B we show through elementary finite arithmetic that the formulas (8) and (11) are equivalent.

Theorem 1.2 (Voronoi formula with Gauss sums). *Let χ be a Dirichlet character modulo c, induced from the primitive character χ^* modulo c^* with $c^* \mid c$. Define for $\boldsymbol{q} = (q_1, \ldots, q_{N-2})$ a tuple of positive integers*

$$
H(\boldsymbol{q}; c, \chi^*, s) = \sum_{n=1}^{\infty} \frac{A(q_{N-2}, \ldots, q_1, n)\, g(\overline{\chi^*}, c, n)}{n^s (c/c^*)^{1-2s}} \qquad (9)
$$

for $\Re(s) > 1$, and

$$
\begin{aligned}
G(\boldsymbol{q}; c, \chi^*, s) &= \frac{G(s)\chi^*(-1)}{c^{Ns-1}(c/c^*)^{1-2s}} \sum_{d_1 c^* \mid q_1 c} \sum_{d_2 c^* \mid \frac{q_1 q_2 c}{d_1}} \cdots \sum_{d_{N-2} c^* \mid \frac{q_1 \cdots q_{N-2} c}{d_1 \cdots d_{N-3}}}\\
&\sum_{n=1}^{\infty} \frac{A(n, d_{N-2}, \ldots, d_1)}{n^{1-s} d_1 d_2 \cdots d_{N-2}} \frac{d_1^{(N-1)s} d_2^{(N-2)s} \cdots d_{N-2}^{2s}}{q_1^{(N-2)s} q_2^{(N-3)s} \cdots q_{N-2}^{s}}\\
&\qquad \times g(\chi^*, c, d_1) g(\chi^*, \xi_1, d_2) \cdots g(\chi^*, \xi_{N-3}, d_{N-2}) g(\chi^*, \xi_{N-2}, n) \qquad (10)
\end{aligned}
$$

for $\Re(s) < 0$, where G equals G_+ or G_- depending on whether $\chi^*(-1)$ is 1 or -1, and $g(\chi^*, \ell c^*, *)$ is the Gauss sum of the induced character modulo ℓc^* from χ^*,

which is defined in Definition 2.1. Both functions have analytic continuation to all $s \in \mathbb{C}$, *and the equality*

$$H(\boldsymbol{q}; c, \chi^*, s) = G(\boldsymbol{q}; c, \chi^*, s) \tag{11}$$

is satisfied.

In proving (11), we define

$$Z(s, w) = \frac{L_{\boldsymbol{q}}(2w - s, F)L(s, F \times \chi^*)}{L(2w - 2s + 1, \overline{\chi^*})}, \tag{12}$$

where $\boldsymbol{q} = (q_1, \ldots, q_{N-2})$ is a tuple of positive integers, and the function $L_{\boldsymbol{q}}(s, F)$ is given as the Dirichlet series

$$L_{\boldsymbol{q}}(s, F) = \sum_{n=1}^{\infty} \frac{A(q_{N-2}, \ldots, q_1, n)}{n^s}$$

for $\Re(s) \gg 1$. We express $Z(s, w)$ as a double Dirichlet series in two different ways. In one region of convergence we express the L-functions as Dirichlet series and obtain

$$Z(s, w) = \sum_{n=1}^{\infty} \frac{a_n(s)}{n^{2w}}.$$

On the other hand, we apply the functional equation of $L(s, F \times \chi^*)$, replacing s with $1 - s$, and write $Z(s, w)$ as the Dirichlet series

$$Z(s, w) = \sum_{n} \frac{b_n(s)}{n^{2w}}.$$

By the uniqueness of Dirichlet series, we must have $a_n(s) = b_n(s)$. This equality leads us to the Voronoi formula with Gauss sums.

Our proof only uses the Hecke relations about the Fourier coefficients of F and the exact form of the functional equations. The expression of Gamma factors, or the automorphy of F, plays no role. Hence we can formulate our theorem in a style similar to the classical converse theorem of Weil. First, let us list the properties of Fourier coefficients that we use in order to state the following theorem.

The Fourier coefficients of F grow moderately, i.e.,

$$A(m_1, \ldots, m_{N-1}) \ll (m_1 \cdots m_{N-1})^{\sigma} \tag{13}$$

for some $\sigma > 0$. Given a primitive Dirichlet character χ^* modulo c^*, define the twisted L-function

$$L(s, F \times \chi^*) = \sum_{n=1}^{\infty} \frac{A(1, \ldots, 1, n)\chi^*(n)}{n^s} \tag{14}$$

for $\Re(s) > \sigma + 1$. It has analytic continuation to the whole complex plane, and satisfies the functional equation

$$L(s, F \times \chi^*) = \tau(\chi^*)^N c^{*-Ns} G(s) L(1-s, \widetilde{F} \times \overline{\chi^*}), \tag{15}$$

where $G(s) = G_+(s)$ or $G_-(s)$ depending on whether $\chi^*(-1) = 1$ or -1.

Theorem 1.3. *Let F be a symbol and assume that with F come numbers*

$$A(m_1, \dots, m_{N-1}) \in \mathbb{C}$$

attached to every $(N-1)$-tuple $(m_1, \dots, m_{N-1})$ of natural numbers. Assume $A(1, \dots, 1) = 1$.

Assume that these "coefficients" $A(*, \dots, *)$ satisfy the aforementioned Hecke relations (2), (3) and (4). Further assume that they grow moderately as in (13).

Let $\widetilde{F}$ be another symbol whose associated coefficients $B(*, \dots, *) \in \mathbb{C}$ are given as in (5) and assume that they also satisfy the same properties. Further, assume that there are two meromorphic functions $G_+(s)$ and $G_-(s)$ associated to the pair $(F, \widetilde{F})$, so that for a given primitive character χ^*, the function $L(s, F \times \chi^*)$ as defined in (14) satisfies the functional equation (15).

Under all these assumptions, $L_{\boldsymbol{q}}(s, F, a/c)$, defined as in (7) for $\Re(s) > 1 + \sigma$, has analytic continuation to all $s \in \mathbb{C}$, and satisfies the Voronoi formula (8). (The Dirichlet series on the right side of (8) is absolutely convergent for $\Re(s) < -\sigma$.)

Equivalently the functions $H(\boldsymbol{q}; c, \chi^*, s)$ and $G(\boldsymbol{q}; c, \chi^*, s)$ as defined by the formulas (9) and (10) have analytic continuations to all s and equal each other as in (11).

Remark 1.4 (the structure of this article). Theorem 1.3 is our main result. For the most part our focus is on the case $N \geq 3$, and we deal with the case $N = 2$ in Remark 3.2. The Voronoi formula (8) is proved to be equivalent to a formula (11) involving Gauss sums. The equivalence is shown in Proposition 3.5. A convolved version of (11) is obtained in Theorem 3.1 by comparing Dirichlet coefficients of two different expressions of a double Dirichlet series. We later show in Proposition 3.3 that this convolved version yields (11).

Remark 1.5. If we start with an L-series $L(s, F)$ with an Euler product

$$L(s, F) = \sum_{n=1}^{\infty} \frac{A(1, \dots, 1, n)}{n^s} = \prod_p \prod_{i=1}^{N} \left(1 - \frac{\alpha_i(p)}{p^s}\right)^{-1}$$

and with $\prod_i \alpha_i(p) = 1$ for any p, then we can define $A(p^{k_1}, \dots, p^{k_{N-1}})$ by the Casselman–Shalika formula [Zhou 2014, Proposition 5.1] and they are compatible with the Hecke relations. More explicitly, for a prime number p, we define $A(p^{k_1}, \dots, p^{k_{N-1}}) = S_{k_1, \dots, k_{N-1}}(\alpha_1(p), \dots, \alpha_N(p))$ by the work of Shintani, where

$S_{k_1,\dots,k_{N-1}}(x_1,\dots,x_N)$ is the Schur polynomial, which can be found in [Goldfeld 2006, p. 233].

We extend the definition to all $A(*,\dots,*)$ multiplicatively by (2). One can prove that $A(*,\dots,*)$ satisfies the Hecke relations (2)–(4). In summary, the "coefficients" $A(*,\dots,*)$ along with the Hecke relations can be generated by an L-function with an Euler product.

The following examples satisfy the conditions in Theorem 1.3, and hence we have a Voronoi formula for each of them.

Example 1.6 (automorphic form for $\mathrm{SL}(N,\mathbb{Z})$). Any cuspidal automorphic form for $\mathrm{SL}(N,\mathbb{Z})$ satisfies the conditions in Theorem 1.3. It can have an unramified or ramified component at the archimedean place, because only the exact form of the $G_\pm$ function would change; see [Godement and Jacquet 1972]. The Hecke–Maass cusp forms considered in Theorem 1.1 are included in this category, and therefore, we prove Theorem 1.3 instead of Theorem 1.1.

Example 1.7 (Rankin–Selberg convolution). Let F_1 and F_2 be even Hecke–Maass cusp forms for $\mathrm{SL}(N_1,\mathbb{Z})$ and $\mathrm{SL}(N_2,\mathbb{Z})$ with the spectral parameters

$$(\lambda_1,\dots,\lambda_{N_1})\in\mathbb{C}^{N_1} \quad\text{and}\quad (\mu_1,\dots,\mu_{N_2})\in\mathbb{C}^{N_2},$$

respectively. Assume $F_1\neq\widetilde{F}_2$ if $N_1=N_2$. The automorphic forms F_1 and F_2 have the standard L-functions

$$L(s,F_1)=\prod_p\prod_{i=1}^{N_1}\Big(1-\frac{\alpha_i(p)}{p^s}\Big)^{-1} \quad\text{and}\quad L(s,F_2)=\prod_p\prod_{i=1}^{N_2}\Big(1-\frac{\beta_i(p)}{p^s}\Big)^{-1}.$$

Let $L(s,F_1\times F_2)$ be the Rankin–Selberg L-function of F_1 and F_2 defined by

$$L(s,F_1\times F_2)=\prod_p\prod_{i_1=1}^{N_1}\prod_{i_2=1}^{N_2}\Big(1-\frac{\alpha_{i_1}(p)\beta_{i_2}(p)}{p^s}\Big)^{-1}.$$

The L-function is of degree $N:=N_1N_2$. The work of Jacquet, Piatetskii-Shapiro, and Shalika [Jacquet et al. 1983] shows that $L(s,F\times\chi^*)=L(s,(F_1\times\chi^*)\times F_2)$ is holomorphic and satisfies the functional equation (15) for $F:=F_1\times F_2$.

Define $A(p^{k_1},\dots,p^{k_{N-1}})$ by the Schur polynomials as in Remark 1.5:

$$A(p^{k_1},\dots,p^{k_{N-1}}):= S_{k_1,\dots,k_{N-1}}\big(\alpha_1(p)\beta_1(p),\dots,\alpha_{i_1}(p)\beta_{i_2}(p),\dots,\alpha_{N_1}(p)\beta_{N_2}(p)\big).$$

Extend the definition to all $A(*,\dots,*)$ multiplicatively by (2). Define

$$G_\pm(s) := i^{-N\delta}\pi^{-N(1/2-s)}\prod_{i_1=1}^{N_1}\prod_{i_2=1}^{N_2}\Gamma\left(\frac{\delta+1-s-\overline{\lambda_{i_1}}-\overline{\mu_{i_2}}}{2}\right)\Gamma\left(\frac{\delta+s-\lambda_{i_1}-\mu_{i_2}}{2}\right)^{-1},$$

where one takes $\delta=0$ and $\delta=1$ for G_+ and G_-, respectively. Theorem 1.3 gives us a Voronoi formula for the Rankin–Selberg convolution $F=F_1\times F_2$ with the $A(*,\dots,*)$ and $G_\pm$ defined above.

Example 1.8 (isobaric sum, Eisenstein series). For $i=1,\dots,k$, let F_i be a Hecke–Maass cusp form for $\mathrm{SL}(N_i,\mathbb{Z})$. Let s_i be complex numbers with $\sum_i N_i s_i=0$. Define the isobaric sum $F=\big(F_1\times|\cdot|_{\mathbb{A}}^{s_1}\big)\boxplus\big(F_2\times|\cdot|_{\mathbb{A}}^{s_2}\big)\boxplus\cdots\boxplus\big(F_k\times|\cdot|_{\mathbb{A}}^{s_k}\big)$, whose L-function is $L(s,F)=\prod_i L(s+s_i,F_i)$. This isobaric sum F is associated with a noncuspidal automorphic form on $\mathrm{GL}(N)$, an Eisenstein series twisted by Maass forms, where $N=\sum_i N_i$; see [Goldfeld 2006, Section 10.5]. The L-function twisted by a character is simply given by $L(s,F\times\chi^*)=\prod_i L(s+s_i,F_i\times\chi^*)$, which satisfies the conditions of Theorem 1.3.

Example 1.9 (symmetric powers on GL(2)). Let f be a modular form of weight k for $\mathrm{SL}(2,\mathbb{Z})$, and define $F:=\mathrm{Sym}^2 f$. The symmetric square F satisfies the conditions in Theorem 1.3 by the work of Shimura [1975]. Here we do not need to involve automorphy using Gelbart–Jacquet lifting. One may have similar results for higher symmetric powers depending on the recent progress in the theory of Galois representations.

As a last remark, let us explain the construction of the double Dirichlet series $Z(s,w)$ given by (12). This construction originates from the Rankin–Selberg convolution of a cusp form F and an Eisenstein series on GL(2). The Fourier coefficients of the Eisenstein series $E(z,s,\chi^*)$ can be written in terms of the divisor function $\sigma_{2s-1}(n,\chi^*)$ defined in Definition 2.1:

$$\frac{1}{n^{2s-1}}\frac{\sigma_{2s-1}(n,\chi^*)}{L(2s,\overline{\chi^*})}\quad\text{or}\quad\sum_{\ell=1}^{\infty}\frac{g(\overline{\chi^*},\ell c^*,n)}{(\ell c^*)^{2s}}.$$

Therefore, in the case of F on GL(2), the Rankin–Selberg integral of F and $E\big(*,w-s+\frac12,\chi^*\big)$ produces the double Dirichlet series

$$\sum_{n=1}^{\infty}\sum_{\ell=1}^{\infty}\frac{A(n)g(\overline{\chi^*},\ell c^*,n)}{n^s(\ell c^*)^{2w+1-2s}}.$$

A similar expression appears on the left-hand side of the Voronoi formula with Gauss sums (9). The Rankin–Selberg convolution of the cusp form F and an

Eisenstein series can be written as a product of two copies of a standard L-function of F, namely

$$\frac{L(2w-s,F)L(s,F\times\chi^*)}{L(2w-2s+1,\overline{\chi^*})}.$$

Applying the functional equation to only $L(s,F\times\chi^*)$ gives us another expression, which is similar to the right-hand side (10) of the Voronoi formula with Gauss sums. Since $L(2w-s,F)$ was not used in this process, we have the freedom to replace $L(2w-s,F)$ by $L_{\boldsymbol{q}}(2w-s,F)$ in the case of GL(N), and it gives us enough generality to prove the Voronoi formula (11) with Gauss sums. In the case of GL(3), this construction is similar to Bump's double Dirichlet series; see [Goldfeld 2006, Chapter 6.6] or [Bump 1984, Chapter X].

2. Background on Gauss sums

Here we collect information about the Gauss sums of Dirichlet characters which are not necessarily primitive.

Definition 2.1. Let χ be a Dirichlet character modulo c induced from a primitive Dirichlet character χ^* modulo c^*. Define the divisor function

$$\sigma_s(m,\chi)=\sum_{d|m}\chi(d)d^s.$$

Define the Gauss sum of χ to be

$$g(\chi^*,c,m)=\sum_{\substack{(u,c)=1\\ u \bmod c}}\chi(u)e\Big(\frac{mu}{c}\Big).$$

The standard Gauss sum for χ^* is given as $\tau(\chi^*)=g(\chi^*,c^*,1)$.

The Gauss sum $g(\chi^*,c,m)$ is the same as the Gauss sum $\tau_m(\chi)$ in other literature. However we prefer our notation because we come upon numerous Gauss sums of characters χ induced from a single primitive character χ^*.

Lemma 2.2 (Gauss sum of nonprimitive characters [Miyake 1989, Lemma 3.1.3(2)]). *Let χ be a character modulo c induced from a primitive character χ^* modulo c^*. Then the Gauss sum of χ is given by*

$$g(\chi^*,c,a)=\tau(\chi^*)\sum_{d|(a,c/c^*)}d\chi^*\Big(\frac{c}{c^*d}\Big)\overline{\chi^*}\Big(\frac{a}{d}\Big)\mu\Big(\frac{c}{c^*d}\Big).$$

Lemma 2.3 [Montgomery and Vaughan 2007, Theorem 9.12]. *Let χ^* be a primitive character modulo c^* and assume $c^*\mid c$. Then we have*

$$g(\chi^*,c,a)=\tau(\chi^*)\frac{\phi(c)}{\phi(c/(c,a))}\mu\Big(\frac{c}{c^*(c,a)}\Big)\chi^*\Big(\frac{c}{c^*(c,a)}\Big)\overline{\chi^*}\Big(\frac{a}{(c,a)}\Big)$$

if $c^*\mid c/(a,c)$. Otherwise, $g(\chi^*,c,a)$ is zero.

The next lemma is a generalization of a famous formula of Ramanujan:

$$\frac{\sigma_{s-1}(n)}{n^{s-1}} = \zeta(s) \sum_{\ell=1}^{\infty} \frac{c_\ell(n)}{\ell^s},$$

where $c_\ell(n)$ is the Ramanujan sum.

Lemma 2.4. *Let $\Re(s) > 1$. Define a Dirichlet series*

$$I(s, \chi^*, c^*, m) = \sum_{\ell=1}^{\infty} \frac{g(\chi^*, \ell c^*, m)}{\ell^s}$$

as a generating function for the nonprimitive Gauss sums induced from χ^*. It satisfies the identity

$$\tau(\chi^*)\sigma_{s-1}(m, \overline{\chi^*}) = m^{s-1} I(s, \chi^*, c^*, m) L(s, \chi^*).$$

Proof. We prove the equivalent formula

$$\tau(\chi^*) m^{1-s} \sigma_{s-1}(m, \overline{\chi^*}) L(s, \chi^*)^{-1} = I(s, \chi^*, c^*, m).$$

For $\Re(s) > 1$, the function $\tau(\chi^*)m^{1-s}\sigma_{s-1}(m, \overline{\chi^*})L(s, \chi^*)^{-1}$ can be written as a Dirichlet series

$$\tau(\chi^*) \sum_{d|m} \frac{d\overline{\chi^*}(m/d)}{d^s} \sum_{n=1}^{\infty} \frac{\chi^*(n)\mu(n)}{n^s} = \tau(\chi^*) \sum_{\ell=1}^{\infty} \frac{\sum_{d|(m,\ell)} d\overline{\chi^*}(m/d)\mu(\ell/d)\chi^*(\ell/d)}{\ell^s},$$

and this equals $I(s, \chi^*, c^*, m)$ by Lemma 2.2. □

Lemma 2.5. *For any two positive integers n and m, and a primitive Dirichlet character χ^* modulo c^*, we have*

$$\sum_{\ell d = n} \chi^*(d) g(\chi^*, \ell c^*, m) = \begin{cases} \tau(\chi^*)\overline{\chi^*(m/n)}n & \text{if } n \mid m, \\ 0 & \text{otherwise.} \end{cases}$$

Proof. We start with the formula,

$$\frac{\tau(\chi^*)\sigma_{s-1}(m, \overline{\chi^*})}{m^{s-1}} = I(s, \chi^*, c^*, m) L(s, \chi^*).$$

Both sides are Dirichlet series and we equate coefficients. The left-hand side is given as

$$\tau(\chi^*) \sum_{e|m} \frac{\overline{\chi^*(m/e)}e}{e^s},$$

whereas the right-hand side is

$$\sum_{\ell=1}^{\infty} \frac{g(\chi^*, \ell c^*, m)}{\ell^s} \sum_{d=1}^{\infty} \frac{\chi^*(d)}{d^s} = \sum_{n=1}^{\infty} \frac{\sum_{d\ell=n} \chi^*(d) g(\chi^*, \ell c^*, m)}{n^s}. \qquad \square$$

3. The Voronoi formula

3A. ***Double Dirichlet series.*** We begin by proving a convolved version of (11).

Theorem 3.1. *For $N \geq 3$, $\boldsymbol{q} = (q_1, \ldots, q_{N-2}) \in \mathbb{N}^{N-2}$, and $n \in \mathbb{N}$, define*

$$\mathcal{H}(\boldsymbol{q}; n, s) := \sum_{d_1 | q_1, \ldots, d_{N-2} | q_{N-2}} \frac{\chi^*(d_1 \cdots d_{N-2})}{(d_1 \cdots d_{N-2})^s} \sum_{d\ell=n} \chi^*(d) H(\boldsymbol{q}'; \ell c^*, \chi^*, s)$$

for $\Re(s) \gg 1$, and

$$\mathcal{G}(\boldsymbol{q}; n, s) := \sum_{d_1 | q_1, \ldots, d_{N-2} | q_{N-2}} \frac{\chi^*(d_1 \cdots d_{N-2})}{(d_1 \cdots d_{N-2})^s} \sum_{d\ell=n} \chi^*(d) G(\boldsymbol{q}'; \ell c^*, \chi^*, s)$$

for $\Re(1-s) \gg 1$, where we abbreviate

$$\boldsymbol{q}' = \Big(\frac{q_1 d}{d_1}, \frac{q_2 d_1}{d_2}, \ldots, \frac{q_{N-2} d_{N-3}}{d_{N-2}} \Big). \tag{16}$$

The functions $\mathcal{H}(\boldsymbol{q}; n, s)$ and $\mathcal{G}(\boldsymbol{q}; n, s)$ have analytic continuation to all $s \in \mathbb{C}$ and these analytic continuations satisfy

$$\mathcal{H}(\boldsymbol{q}; n, s) = \mathcal{G}(\boldsymbol{q}; n, s). \tag{17}$$

Proof. The region of absolute convergence for $\mathcal{H}(\boldsymbol{q}; n, s)$ is a right half plane $\Re(s) \gg 1$, and the region of absolute convergence of $\mathcal{G}(\boldsymbol{q}; n; s)$ is a left half plane $\Re(1-s) \gg 1$. Let $Z(s, w)$ be defined as in (12). For any $s \in \mathbb{C}$ and w with $\Re(w)$ large enough so that $\Re(2w-s) \gg 1$ and $\Re(w-s) > 0$, writing $L_{\boldsymbol{q}}(2w-s, F)$ and $L(2w-2s+1, \overline{\chi^*})^{-1}$ as Dirichlet series, we derive

$$Z(s, w) = L(s, F \times \chi^*) \sum_{n=1}^{\infty} \frac{\sum_{d|n} A(q_{N-2}, \ldots, q_1, d) d^s \overline{\chi^*}(n/d) \mu(n/d) (n/d)^{2s-1}}{n^{2w}}.$$

Hence, we have

$$Z(s, w) = \sum_{n=1}^{\infty} \frac{a_n(s)}{n^{2w}},$$

where

$$a_n(s) = L(s, F \times \chi^*) \sum_{d|n} A(q_{N-2}, \ldots, q_1, d) d^s \overline{\chi^*}(n/d) \mu(n/d) (n/d)^{2s-1}.$$

Here $a_n(s)$ is an analytic function of $s \in \mathbb{C}$, because $L(s, F \times \chi^*)$ is entire. The computation below shows that $a_n(s)$ equals either side of (17) in their respective

regions of absolute convergence, up to scaling by a constant $\tau(\overline{\chi^*})$. This proves the analytic continuation of $\mathcal{H}$ and $\mathcal{G}$ as well as their equality.

For $\Re(s) \gg 1$, $\Re(w-s) > 0$, we expand the two L-functions in the numerator of $Z(s,w)$ as Dirichlet series, obtaining

$$\begin{aligned} Z(s,w) &= \frac{1}{L(2w-2s+1,\overline{\chi^*})} \sum_{n,m=1}^{\infty} \frac{A(q_{N-2},\dots,q_1,n)A(1,\dots,1,m)\chi^*(m)}{n^{2w-s}m^s} \\ &= \frac{1}{L(2w-2s+1,\overline{\chi^*})} \sum_{n,m=1}^{\infty} \Bigg(\frac{\chi^*(m)}{n^{2w-s}m^s} \\ &\qquad \times \sum_{\substack{d_0d_1\cdots d_{N-1}=m \\ d_0\mid n, d_1\mid q_1,\dots,d_{N-2}\mid q_{N-2}}} A\Big(\frac{q_{N-2}d_{N-3}}{d_{N-2}},\dots,\frac{q_1d_0}{d_1},\frac{nd_{N-1}}{d_0}\Big)\Bigg), \end{aligned}$$

where we have used the Hecke relation (3). We change the variable $n/d_0 \to n$ and combine $h = nd_{N-1}$, giving

$$\begin{aligned} Z(s,w) &= \frac{1}{L(2w-2s+1,\overline{\chi^*})} \sum_{n,d_0,d_{N-1}=1}^{\infty} \sum_{\substack{d_i\mid q_i \\ i=1,\dots,N-2}} \frac{\chi^*(d_0\cdots d_{N-1})}{n^{2w-s}d_0^{2w-s}(d_0\cdots d_{N-1})^s} \\ &\qquad \times A\Big(\frac{q_{N-2}d_{N-3}}{d_{N-2}},\dots,\frac{q_1d_0}{d_1},nd_{N-1}\Big) \\ &= \frac{1}{L(2w-2s+1,\overline{\chi^*})} \sum_{d_0,h=1}^{\infty} \sum_{\substack{d_i\mid q_i \\ i=1,\dots,N-2}} \frac{\chi^*(d_0\cdots d_{N-2})}{d_0^{2w-s}(d_0\cdots d_{N-2})^s} \\ &\qquad \times A\Big(\frac{q_{N-2}d_{N-3}}{d_{N-2}},\dots,\frac{q_1d_0}{d_1},h\Big)\frac{\sigma_{2w-2s}(h,\chi^*)}{h^{2w-s}}. \end{aligned}$$

Applying Lemma 2.4, we get

$$\begin{aligned} Z(s,w) = \tau(\overline{\chi^*})^{-1} \sum_{d_0=1}^{\infty} \sum_{\substack{d_i\mid q_i \\ i=1,\dots,N-2}} \Bigg(&\frac{\chi^*(d_0\cdots d_{N-2})}{d_0^{2w}(d_1\cdots d_{N-2})^s} \\ &\times \sum_{h=1}^{\infty} \frac{1}{h^s} A\Big(\frac{q_{N-2}d_{N-3}}{d_{N-2}},\dots,\frac{q_1d_0}{d_1},h\Big) \sum_{\ell=1}^{\infty} \frac{g(\overline{\chi^*},\ell c^*,h)}{\ell^{2w-2s+1}}\Bigg). \end{aligned}$$

Therefore, defining $\boldsymbol{q}'$ as in (16), we reach

$$\begin{aligned} Z(s,w) = \tau(\overline{\chi^*})^{-1} \sum_{n=1}^{\infty} \frac{1}{n^{2w}} \sum_{d_1\mid q_1,\dots,d_{N-2}\mid q_{N-2}} \Bigg(&\frac{\chi^*(d_1\cdots d_{N-2})}{(d_1\cdots d_{N-2})^s} \\ &\times \sum_{d\ell=n} \chi^*(d) H(\boldsymbol{q}';\ell c^*,\chi^*,s)\Bigg). \end{aligned} \tag{18}$$

On the other hand, let us apply the functional (15) to $L(s, F \times \chi^*)$ in $Z(s, w)$, giving

$$Z(s, w) = \frac{G(s)\tau(\chi^*)^N}{c^{*Ns}} \frac{L_q(2w-s, F)L(1-s, \widetilde{F} \times \overline{\chi^*})}{L(2w-2s+1, \overline{\chi^*})}.$$

Given $\Re(1-s) \gg 1$ and $\Re(2w-s) \gg 1$, we open the expression as a Dirichlet series:

$$\begin{aligned}
&Z(s, w)\\
&= \frac{G(s)\tau(\chi^*)^N c^{*-Ns}}{L(2w-2s+1, \overline{\chi^*})} \sum_{n,m=1}^{\infty} \frac{A(q_{N-2}, \dots, q_1, n)A(m, 1, \dots, 1)\overline{\chi^*}(m)}{n^{2w-s}m^{1-s}}\\
&= \frac{G(s)\tau(\chi^*)^N c^{*-Ns}}{L(2w-2s+1, \overline{\chi^*})}\\
&\quad \times \sum_{n,m=1}^{\infty} \frac{\overline{\chi^*}(m)}{n^{2w-s}m^{1-s}} \sum_{\substack{d_0 d_1 \cdots d_{N-1}=m\\ d_0|n, d_1|q_1, \dots, d_{N-2}|q_{N-2}}} A\Big(\frac{q_{N-2}d_{N-1}}{d_{N-2}}, \dots, \frac{q_1 d_2}{d_1}, \frac{nd_1}{d_0}\Big)\\
&= \frac{G(s)\tau(\chi^*)^N c^{*-Ns}}{L(2w-2s+1, \overline{\chi^*})}\\
&\quad \times \sum_{n,m=1}^{\infty} \sum_{\substack{d_0 d_1 \cdots d_{N-1}=m\\ d_0|n, d_1|q_1, \dots, d_{N-2}|q_{N-2}}} \frac{\overline{\chi^*}(d_0 d_1 \cdots d_{N-1})A\Big(\frac{q_{N-2}d_{N-1}}{d_{N-2}}, \dots, \frac{q_1 d_2}{d_1}, \frac{nd_1}{d_0}\Big)}{(n/d_0)^{2w-s} d_0^{1+2w-2s} (d_1 \cdots d_{N-1})^{1-s}},
\end{aligned}$$

where we have combined the Fourier coefficients by the Hecke relation (4). We change the variable $n/d_0 \to n$. Then the sum over d_0 cancels with $L(2w-2s+1, \overline{\chi^*})$ in the denominator, giving

$$\begin{aligned}
&Z(s, w)\\
&= \frac{G(s)\tau(\chi^*)^N c^{*-Ns}}{L(2w-2s+1, \overline{\chi^*})} \sum_{n,d_0,d_{N-1}=1}^{\infty} \sum_{\substack{d_i|q_i\\ i=1,\dots,N-2}} A\Big(\frac{q_{N-2}d_{N-1}}{d_{N-2}}, \dots, \frac{q_1 d_2}{d_1}, d_1 n\Big)\\
&\qquad \times \frac{\overline{\chi^*}(d_0 d_1 \cdots d_{N-1})}{n^{2w-s} d_0^{1+2w-2s} (d_1 \cdots d_{N-1})^{1-s}}\\
&= \frac{G(s)\tau(\chi^*)^N}{c^{*Ns}} \sum_{n,d_{N-1}=1}^{\infty} \sum_{\substack{d_i|q_i\\ i=1,\dots,N-2}} \frac{\overline{\chi^*}(d_1 \cdots d_{N-1})}{n^{2w-s}(d_1 \cdots d_{N-1})^{1-s}}. \qquad (19)
\end{aligned}$$

If we denote the right-hand side of (17) by $\tau(\overline{\chi^*})b_n(s)$, our goal is to transform (19) into $R := \sum_{n=1}^{\infty} b_n(s)n^{-2w}$. But at this point it is easier to start from R. More explicitly, we have

$$R = \tau(\overline{\chi^*})^{-1} \sum_{h=1}^{\infty} \frac{1}{h^{2w}} \sum_{d_1|q_1,\dots,d_{N-2}|q_{N-2}} \Bigg(\frac{\chi^*(d_1\cdots d_{N-2})}{(d_1\cdots d_{N-2})^s} \times \sum_{d\ell=h} \chi^*(d) G(\boldsymbol{q}'; \ell c^*, \chi^*, s) \Bigg). \quad (20)$$

Here $\boldsymbol{q}'$ has been defined in (16). We plug in the definition of $G(\boldsymbol{q}'; \ell c^*, \chi^*, s)$ from (10) for $\boldsymbol{q}'$, giving

$$\begin{aligned} &G(\boldsymbol{q}'; \ell c^*, \chi^*, s) \\ &= \frac{G(s)\chi^*(-1)}{c^{*Ns-1}\ell^{(N-2)s}} \sum_{f_1 \mid \frac{q_1 d\ell}{d_1}} \sum_{f_2 \mid \frac{q_1 q_2 d\ell}{f_1 d_2}} \cdots \sum_{f_{N-2} \mid \frac{q_1\cdots q_{N-2} d\ell}{f_1\cdots f_{N-3} d_{N-2}}} \\ &\quad \sum_{n=1}^{\infty} \frac{A(n, f_{N-2}, \dots, f_1)}{n^{1-s} f_1 f_2 \cdots f_{N-2}} \frac{f_1^{(N-1)s} f_2^{(N-2)s} \cdots f_{N-2}^{2s}}{q_1^{(N-2)s} q_2^{(N-3)s} \cdots q_{N-2}^{s}} \frac{(d_1\cdots d_{N-2})^s}{d^{(N-2)s}} \\ &\qquad \times g(\chi^*, \ell c^*, f_1) g\Big(\chi^*, \frac{q_1 d\ell c^*}{f_1 d_1}, f_2\Big) \\ &\qquad \cdots \times g\Big(\chi^*, \frac{q_1\cdots q_{N-3} d\ell c^*}{f_1\cdots f_{N-3} d_{N-3}}, f_{N-2}\Big) g\Big(\chi^*, \frac{q_1\cdots q_{N-2} d\ell c^*}{f_1\cdots f_{N-2} d_{N-2}}, n\Big). \end{aligned}$$

We substitute $G(\boldsymbol{q}'; \ell c^*, \chi^*, s)$ with this expression in (20) and change the orders of summation between f_i and d_i. The summations over d and d_i collapse with the repeated use of Lemma 2.5, giving

$$\begin{aligned} R = \tau(\overline{\chi^*})^{-1} \frac{G(s)\chi^*(-1)}{c^{*Ns-1}} \sum_{h=1}^{\infty}\sum_{n=1}^{\infty} \sum_{\substack{h \mid f_1 \\ f_1 \mid q_1 h}} \sum_{\substack{\frac{q_1 h}{f_1} \mid f_2 \\ f_2 \mid \frac{q_1 q_2 h}{f_1}}} \cdots \sum_{\substack{\frac{q_1\cdots q_{N-3} h}{f_1\cdots f_{N-3}} \mid f_{N-2} \\ f_{N-2} \mid \frac{q_1\cdots q_{N-2} h}{f_1\cdots f_{N-3}}}} \sum_{\frac{q_1\cdots q_{N-2} h}{f_1\cdots f_{N-2}} \mid n} \frac{\tau(\chi^*)^{N-1}}{h^{2w}} \\ \times \overline{\chi^*}\Big(\frac{f_1}{h}\Big)\overline{\chi^*}\Big(\frac{f_1 f_2}{h q_1}\Big)\cdots\overline{\chi^*}\Big(\frac{f_1 f_2\cdots f_{N-2}}{h q_1\cdots q_{N-3}}\Big)\overline{\chi^*}\Big(\frac{f_1 f_2\cdots f_{N-2} n}{h q_1\cdots q_{N-2}}\Big) \\ \times \Big(\frac{q_1}{f_1}\Big)^{N-2}\Big(\frac{q_2}{f_2}\Big)^{N-3}\cdots\Big(\frac{q_{N-2}}{f_{N-2}}\Big) h^{N-1-Ns+2s} \\ \times \frac{A(n, f_{N-2}, \dots, f_1)}{n^{1-s} f_1\cdots f_{N-2}} \frac{f_1^{(N-1)s}\cdots f_{N-2}^{2s}}{q_1^{(N-2)s}\cdots q_{N-2}^{s}}. \end{aligned}$$

Define $e_1 = f_1/h$ and $e_i = (f_1\cdots f_i)/(q_1\cdots q_{i-1}h)$ for $i = 2, \dots, N-2$, so that the double conditions under the sums simplify to $e_i|q_i$. Extend this to all positive integers by setting $e_{N-1} = (f_1\cdots f_{N-2}n)/(hq_1\cdots q_{N-2})$. Finally, noting $\tau(\overline{\chi^*})^{-1} = \chi^*(-1)\tau(\chi^*)/c^*$, we get

$$R = \frac{G(s)\tau(\chi^*)^N}{c^{*Ns}} \sum_{h,e_{N-1}=1}^{\infty} \frac{1}{h^{2w-s}} \sum_{\substack{e_i|q_i \\ i=1,\dots,N-2}} \frac{\overline{\chi^*}(e_1\cdots e_{N-2}e_{N-1})}{(e_1\cdots e_{N-1})^{1-s}} \times A\left(\frac{e_{N-1}q_{N-2}}{e_{N-2}}, \dots, \frac{e_2q_1}{e_1}, e_1h\right),$$

which in turn, by (19), equals $Z(s, w)$ as well as (18). We complete the proof by applying the uniqueness theorem for Dirichlet series [Apostol 1976, Theorem 11.3] to the equality between (18) and (20). □

Remark 3.2. The above proof works for $N \geq 3$ but not for $N = 2$. We can prove the Voronoi formula for $\mathrm{SL}(2, \mathbb{Z})$ similarly and easily by considering

$$Z(s, w) = \frac{L(2w - s, F)L(s, F \times \chi^*)}{L(2w - 2s + 1, \overline{\chi^*})L(2w, \chi^*)}.$$

We have, from the Hecke relations on GL(2),

$$Z(s, w) = \tau(\overline{\chi^*})^{-1} \sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} \frac{A(n)}{n^s} \frac{g(\overline{\chi^*}, \ell c^*, n)}{\ell^{1+2w-2s}},$$

and applying the functional equation for $L(s, F \times \chi^*)$ we have

$$Z(s, w) = \tau(\chi^*)c^{*-2s}G(s) \sum_{\ell=1}^{\infty}\sum_{n=1}^{\infty} \frac{A(n)}{n^{1-s}} \frac{g(\chi^*, \ell c^*, n)}{\ell^{2w}}.$$

Applying the uniqueness theorem for Dirichlet series to the variable w, we get the Voronoi formula with Gauss sums on GL(2).

Proposition 3.3. *Equation* (11) *is equivalent to Theorem 3.1.*

Proof. Construct the following summation:

$$T := \sum_{e_0|n} \sum_{e_1|q_1e_0} \cdots \sum_{e_{N-2}|q_{N-2}e_{N-3}} \frac{\mu(e_0\cdots e_{N-2})\chi^*(e_0\cdots e_{N-2})}{(e_1\cdots e_{N-2})^s} \times \mathcal{H}\left(\frac{q_1e_0}{e_1}, \dots, \frac{q_{N-2}e_{N-3}}{e_{N-2}}; \frac{n}{e_0}, s\right)$$

$$= \sum_{e_0|n} \sum_{e_1|q_1e_0} \cdots \sum_{e_{N-2}|q_{N-2}e_{N-3}} \left(\frac{\mu(e_0\cdots e_{N-2})\chi^*(e_0\cdots e_{N-2})}{(e_1\cdots e_{N-2})^s} \times \sum_{\substack{d_i|q_ie_{i-1}/e_i \\ i=1,\dots,N-2}} \frac{\chi^*(d_1\cdots d_{N-2})}{(d_1\cdots d_{N-2})^s} \sum_{d_0|n/e_0} \chi^*(d_0) \times H\left(\frac{q_1e_0d_0}{e_1d_1}, \dots, \frac{q_{N-2}e_{N-3}d_{N-3}}{e_{N-2}d_{N-2}}; \frac{n}{e_0d_0}c^*, \chi^*, s\right)\right).$$

Change variables $e_i d_i \to a_i$ for $i = 0, \dots, N-2$, and change orders of summation, getting

$$T = \sum_{a_0|n} \sum_{e_0|a_0} \sum_{a_1|q_1 e_0} \sum_{e_1|a_1} \cdots \sum_{a_{N-2}|q_{N-2} e_{N-3}} \sum_{e_{N-2}|a_{N-2}} \frac{\chi^*(a_0 \cdots a_{N-2})}{(a_1 \cdots a_{N-2})^s} \times H\left(\frac{q_1 a_0}{a_1}, \frac{q_2 a_1}{a_2}, \dots, \frac{q_{N-2} a_{N-3}}{a_{N-2}}; \frac{nc^*}{a_0}, \chi^*, s\right) \mu(e_0) \cdots \mu(e_{N-2}).$$

One by one, the Möbius summation over e_i forces $a_i = 1$, and thus we obtain $T = H(\boldsymbol{q}; nc^*, \chi^*, s)$. By Theorem 3.1, we have $\mathscr{H} = \mathscr{G}$, and the same calculations yield $T = G(\boldsymbol{q}; nc^*, \chi^*, s)$. This proves the theorem. □

3B. ***Equivalence between equations (8) and (11).*** First we prove a lemma showing that the hyper-Kloosterman sum on the right-hand side of (8) becomes a product of $(N-2)$ Gauss sums after averaging against a Dirichlet character.

Lemma 3.4. *Let χ be a Dirichlet character modulo c which is induced from the primitive character χ^* modulo c^*. Let $\boldsymbol{q} = (q_1, \dots, q_{N-2})$ and $\boldsymbol{d} = (d_1, \dots, d_{N-2})$ be two tuples of positive integers, and assume that all the divisibility conditions in (1) are met. Consider the summation*

$$S := \sum_{\substack{a \bmod c \\ (a,c)=1}} \chi(a)\, \mathrm{Kl}(a, n, c; \boldsymbol{q}, \boldsymbol{d}).$$

The quantity S is zero unless the divisibility conditions

$$d_1 c^* | q_1 c, \quad d_2 c^* \,\Big|\, \frac{q_1 q_2 c}{d_1}, \quad d_3 c^* \,\Big|\, \frac{q_1 q_2 q_3 c}{d_1 d_2}, \quad \dots, \quad d_{N-2} c^* \,\Big|\, \frac{q_1 \cdots q_{N-2} c}{d_1 \cdots d_{N-3}} \tag{21}$$

are satisfied. Under such divisibility conditions, setting $\xi_i := (q_1 \cdots q_i c)/(d_1 \cdots d_i)$, S can be written as a product of Gauss sums:

$$S = g(\chi^*, c, d_1) g(\chi^*, \xi_1, d_2) \cdots g(\chi^*, \xi_{N-3}, d_{N-2}) g(\chi^*, \xi_{N-2}, n).$$

Proof. The divisibility conditions (1) imply

$$d_1 \mid q_1(c, d_1), \quad d_2 \mid q_2(\xi_1, d_2), \quad \dots, \quad d_{N-2} \mid q_{N-2}(\xi_{N-3}, d_{N-2}). \tag{22}$$

We open up the hyper-Kloosterman sum in S. The forthcoming computation is an iterative process. The summation over a yields a Gauss sum, which in turn produces the term $\overline{\chi^*}(x_1)$. Then the summation over x_1 yields another Gauss sum, which produces the term $\overline{\chi^*}(x_2)$, and so on.

First, we sum over a modulo c:

$$S = \sum_{a \bmod c} \chi(a) \sideset{}{^*}\sum_{x_1 \bmod \xi_1} e\Big(\frac{d_1 x_1 a}{c}\Big)\Big(\sideset{}{^*}\sum_{x_2 \bmod \xi_2} e\Big(\frac{d_2 x_2 \overline{x_1}}{\xi_1}\Big)\cdots\Big)$$
$$= \sideset{}{^*}\sum_{x_1 \bmod \xi_1} g(\chi^*, c, x_1 d_1)\Big(\sideset{}{^*}\sum_{x_2 \bmod \xi_2} e\Big(\frac{d_2 x_2 \overline{x_1}}{\xi_1}\Big)\cdots\Big).$$

Now, because $(c, x_1 d_1) = ((c, q_1 c), x_1 d_1) = (c, (q_1 c, x_1 d_1)) = (c, d_1)$, we deduce from Lemma 2.3 that

$$g(\chi^*, c, x_1 d_1) = \overline{\chi^*(x_1)} g(\chi^*, c, d_1).$$

By Lemma 2.3, this Gauss sum is zero unless $c^* \mid c/(c, d_1)$, which implies the first divisibility condition of (21) because, by (22),

$$c^* \,\Big|\, \frac{c}{(c, d_1)} = \frac{d_1}{(c, d_1)} \frac{c}{d_1} \,\Big|\, \frac{q_1 c}{d_1}.$$

Next we sum over x_1. Notice that $\overline{x_1}$ is its multiplicative inverse modulo $q_1 c/d_1$, and hence modulo c^*. This means that $\chi^*(\overline{x_1}) = \overline{\chi^*(x_1)}$. We change variables in the x_1 summation $x_1 \to \overline{x_1}$, and change orders of summation to obtain

$$S = g(\chi^*, c, d_1) \sideset{}{^*}\sum_{x_1 \bmod \xi_1} \overline{\chi^*(x_1)}\Big(\sideset{}{^*}\sum_{x_2 \bmod \xi_2} e\Big(\frac{d_2 x_2 \overline{x_1}}{\xi_1}\Big)\cdots\Big)$$
$$= g(\chi^*, c, d_1) \sideset{}{^*}\sum_{x_2 \bmod \xi_2} \sideset{}{^*}\sum_{x_1 \bmod \xi_1} \chi^*(x_1) e\Big(\frac{d_2 x_2 x_1}{\xi_1}\Big)\Big(\sideset{}{^*}\sum_{x_3 \bmod \xi_3} e\Big(\frac{d_3 x_3 \overline{x_2}}{\xi_2}\Big)\cdots\Big)$$
$$= g(\chi^*, c, d_1) \sideset{}{^*}\sum_{x_2 \bmod \xi_2} g(\chi^*, \xi_1, d_2 x_2)\Big(\sideset{}{^*}\sum_{x_3 \bmod \xi_3} e\Big(\frac{d_3 x_3 \overline{x_2}}{\xi_2}\Big)\cdots\Big).$$

Once again, the equalities $(\xi_1, d_2 x_2) = ((\xi_1, d_2 \xi_2), d_2 x_2) = (\xi_1, (d_2 \xi_2, d_2 x_2)) = (\xi_1, d_2)$ imply that we can pull out $\overline{\chi^*}(x_2)$ from the Gauss sum. Then we have

$$S = g(\chi^*, c, d_1) g(\chi^*, \xi_1, d_2) \sideset{}{^*}\sum_{x_2 \bmod \xi_2} \overline{\chi^*(x_2)}\Big(\sideset{}{^*}\sum_{x_3 \bmod \xi_3} e\Big(\frac{d_3 x_3 \overline{x_2}}{\xi_2}\Big)\cdots\Big).$$

The second Gauss sum $g(\chi^*, \xi_1, d_2)$ vanishes unless $c^* \mid \xi_1/(\xi_1, d_2)$ by Lemma 2.3. This in turn implies $c^* \mid \xi_1/(\xi_1, d_2) \mid \xi_2$ by (22), which is the second divisibility condition of (21). We complete the proof after repeating this process $(N-2)$ times. □

Proposition 3.5. *The equations (8) and (11) are equivalent.*

Proof. Let χ be a Dirichlet character modulo c induced from the primitive Dirichlet character χ^* modulo c^*. Multiply both sides of (8) by $\chi(a)$ and sum over reduced

residue classes modulo c. On the left-hand side of (8), one gets

$$\sum_{\substack{a \bmod c \\ (a,c)=1}} \chi(a) L_{\boldsymbol{q}}(s, F, a/c) = (c/c^*)^{1-2s} H(\boldsymbol{q}; c, \chi^*, s),$$

whereas on the right-hand side of (8), one obtains $(c/c^*)^{1-2s} G(\boldsymbol{q}; c, \chi^*, s)$ by making use of Lemma 3.4 and the fact that

$$g(\chi^*, \xi_{N-2}, -n) = \pm g(\chi^*, \xi_{N-2}, n),$$

depending on whether $\chi(-1)$ is 1 or -1. This shows that (8) implies (11).

Conversely, if we multiply both sides of (11) by $\overline{\chi(a)}/\phi(c)$ and sum over all Dirichlet characters (both primitive and nonprimitive) modulo c, we obtain (8) by using the orthogonality relation for Dirichlet characters. Since both of the aforementioned summations that shuttle between (8) and (11) are finite, the properties of absolute convergence and analytic continuation are preserved. □

Acknowledgements

The authors thank Dorian Goldfeld and Wenzhi Luo for helpful suggestions, Matthew Young for his extensive help in organization of the manuscript and his encouragement, and Jeffrey Hoffstein, in whose class the seed for this work originated.

References

[Apostol 1976] T. M. Apostol, *Introduction to analytic number theory*, Springer, New York, 1976. MR Zbl

[Blomer 2012] V. Blomer, "Subconvexity for twisted L-functions on GL(3)", *Amer. J. Math.* **134**:5 (2012), 1385–1421. MR Zbl

[Blomer et al. 2013] V. Blomer, R. Khan, and M. Young, "Distribution of mass of holomorphic cusp forms", *Duke Math. J.* **162**:14 (2013), 2609–2644. MR Zbl

[Bump 1984] D. Bump, *Automorphic forms on* GL$(3, \mathbb{R})$, Lecture Notes in Mathematics **1083**, Springer, Berlin, 1984. MR Zbl

[Buttcane and Khan 2015] J. Buttcane and R. Khan, "L^4-norms of Hecke newforms of large level", *Math. Ann.* **362**:3-4 (2015), 699–715. MR Zbl

[Duke and Iwaniec 1990] W. Duke and H. Iwaniec, "Estimates for coefficients of L-functions, I", pp. 43–47 in *Automorphic forms and analytic number theory* (Montreal, 1989), edited by R. Murty, Univ. Montréal, 1990. MR Zbl

[Godement and Jacquet 1972] R. Godement and H. Jacquet, *Zeta functions of simple algebras*, Lecture Notes in Mathematics **260**, Springer, Berlin, 1972. MR Zbl

[Goldfeld 2006] D. Goldfeld, *Automorphic forms and L-functions for the group* GL$(n, \mathbb{R})$, Cambridge Studies in Advanced Mathematics **99**, Cambridge University Press, 2006. MR Zbl

[Goldfeld and Li 2006] D. Goldfeld and X. Li, "Voronoi formulas on GL(n)", *Int. Math. Res. Not.* **2006** (2006), art. id 86295. MR Zbl

[Goldfeld and Li 2008] D. Goldfeld and X. Li, "The Voronoi formula for GL$(n, \mathbb{R})$", *Int. Math. Res. Not.* **2008**:2 (2008), art. id rnm144. MR Zbl

[Good 1981] A. Good, "Cusp forms and eigenfunctions of the Laplacian", *Math. Ann.* **255**:4 (1981), 523–548. MR Zbl

[Ichino and Templier 2013] A. Ichino and N. Templier, "On the Voronoĭ formula for GL(n)", *Amer. J. Math.* **135**:1 (2013), 65–101. MR Zbl

[Jacquet et al. 1983] H. Jacquet, I. I. Piatetskii-Shapiro, and J. A. Shalika, "Rankin–Selberg convolutions", *Amer. J. Math.* **105**:2 (1983), 367–464. MR Zbl

[Khan 2012] R. Khan, "Simultaneous non-vanishing of GL(3) × GL(2) and GL(2) L-functions", *Math. Proc. Cambridge Philos. Soc.* **152**:3 (2012), 535–553. MR Zbl

[Kowalski and Ricotta 2014] E. Kowalski and G. Ricotta, "Fourier coefficients of GL(N) automorphic forms in arithmetic progressions", *Geom. Funct. Anal.* **24**:4 (2014), 1229–1297. MR Zbl

[Li 2009] X. Li, "The central value of the Rankin–Selberg L-functions", *Geom. Funct. Anal.* **18**:5 (2009), 1660–1695. MR Zbl

[Li 2011] X. Li, "Bounds for GL(3) × GL(2) L-functions and GL(3) L-functions", *Ann. of Math.* (2) **173**:1 (2011), 301–336. MR Zbl

[Li and Young 2012] X. Li and M. P. Young, "The L^2 restriction norm of a GL_3 Maass form", *Compos. Math.* **148**:3 (2012), 675–717. MR Zbl

[Miller 2006] S. D. Miller, "Cancellation in additively twisted sums on GL(n)", *Amer. J. Math.* **128**:3 (2006), 699–729. MR Zbl

[Miller and Schmid 2004] S. D. Miller and W. Schmid, "Summation formulas, from Poisson and Voronoi to the present", pp. 419–440 in *Noncommutative harmonic analysis*, edited by P. Delorme and M. Vergne, Progress in Mathematics **220**, Birkhäuser, Boston, 2004. MR Zbl

[Miller and Schmid 2006] S. D. Miller and W. Schmid, "Automorphic distributions, L-functions, and Voronoi summation for GL(3)", *Ann. of Math.* (2) **164**:2 (2006), 423–488. MR Zbl

[Miller and Schmid 2011] S. D. Miller and W. Schmid, "A general Voronoi summation formula for GL($n, \mathbb{Z}$)", pp. 173–224 in *Geometry and analysis, No. 2*, edited by L. Ji, Advanced Lectures in Mathematics **18**, Int. Press, Somerville, MA, 2011. MR Zbl

[Miyake 1989] T. Miyake, *Modular forms*, Springer, Berlin, 1989. MR Zbl

[Montgomery and Vaughan 2007] H. L. Montgomery and R. C. Vaughan, *Multiplicative number theory, I: Classical theory*, Cambridge Studies in Advanced Mathematics **97**, Cambridge University Press, 2007. MR Zbl

[Munshi 2013] R. Munshi, "Shifted convolution sums for GL(3) × GL(2)", *Duke Math. J.* **162**:13 (2013), 2345–2362. MR Zbl

[Munshi 2015] R. Munshi, "The circle method and bounds for L-functions, IV: Subconvexity for twists of GL(3) L-functions", *Ann. of Math.* (2) **182**:2 (2015), 617–672. MR Zbl

[Shimura 1975] G. Shimura, "On the holomorphy of certain Dirichlet series", *Proc. London Math. Soc.* (3) **31**:1 (1975), 79–98. MR Zbl

[Zhou 2014] F. Zhou, "Weighted Sato–Tate vertical distribution of the Satake parameter of Maass forms on PGL(N)", *Ramanujan J.* **35**:3 (2014), 405–425. MR Zbl

[Zhou 2016] F. Zhou, "Voronoi summation formulae on GL(n)", *J. Number Theory* **162** (2016), 483–495. MR Zbl

Communicated by Philippe Michel
Received 2016-05-08 Revised 2016-07-19 Accepted 2016-09-23

ekiral@math.tamu.edu *Department of Mathematics, Texas A&M University, College Station, TX 77843, United States*

zhou.1406@math.osu.edu *Department of Mathematics, The Ohio State University, 231 West 18th Avenue, Columbus, OH 43210, United States*

ALGEBRA AND NUMBER THEORY 10:10 (2016)
dx.doi.org/10.2140/ant.2016.10.2287

Finite dimensional Hopf actions on algebraic quantizations

Pavel Etingof and Chelsea Walton

Let k be an algebraically closed field of characteristic zero. In joint work with J. Cuadra, we showed that a semisimple Hopf action on a Weyl algebra over a polynomial algebra $k[z_1, \ldots, z_s]$ factors through a group action, and this in fact holds for any finite dimensional Hopf action if $s = 0$. We also generalized these results to finite dimensional Hopf actions on algebras of differential operators. In this work we establish similar results for Hopf actions on other algebraic quantizations of commutative domains. This includes universal enveloping algebras of finite dimensional Lie algebras, spherical symplectic reflection algebras, quantum Hamiltonian reductions of Weyl algebras (in particular, quantized quiver varieties), finite W-algebras and their central reductions, quantum polynomial algebras, twisted homogeneous coordinate rings of abelian varieties, and Sklyanin algebras. The generalization in the last three cases uses a result from algebraic number theory due to A. Perucca.

1. Introduction

Throughout this paper, k will denote an algebraically closed field of characteristic zero. In [Etingof and Walton 2014, Theorem 1.3], we showed that any semisimple Hopf action on a commutative domain over k factors through a group action. Likewise, it was established in our joint work with Juan Cuadra that the same conclusion holds for semisimple Hopf actions on Weyl algebras $\boldsymbol{A}_n(k[z_1, \ldots, z_s])$ [Cuadra et al. 2015, Proposition 4.3]. Moreover, we showed that it holds for any (not necessarily semisimple) finite dimensional Hopf action on $\boldsymbol{A}_n(k)$ [Cuadra et al. 2016, Theorem 1.1], and, more generally, on algebras of differential operators of smooth affine varieties [Cuadra et al. 2016, Theorem 1.2]. Finally, in [Etingof and Walton 2016] we extended these results to certain finite dimensional Hopf actions on deformation quantizations (i.e., formal quantum deformations) of commutative domains. We say that there is *no finite quantum symmetry* in the settings above.

MSC2010: primary 16T05; secondary 16S80, 13A35, 16S38.
Keywords: algebraic quantization, filtered deformation, Hopf algebra action, quantum polynomial algebra, Sklyanin algebra, twisted coordinate ring.

The goal of this paper is to establish no finite quantum symmetry results for finite dimensional Hopf actions on other algebraic quantizations of commutative domains, i.e., quantizations whose parameters are elements of k (rather than formal variables). We now summarize our main results for various classes of algebraic quantizations.

1A. ***Semisimple Hopf actions on filtered quantizations.*** Our first main result concerns Hopf actions on *filtered deformations* (or *filtered quantizations*) of commutative domains, that is, on filtered k-algebras B where the associated graded algebra $\mathrm{gr}(B)$ is a commutative finitely generated domain.

Let B be a $\mathbb{Z}_+$-filtered algebra over k such that $\mathrm{gr}(B)$ is a commutative finitely generated domain. We will see that for sufficiently large primes p, the algebra B admits a reduction B_p modulo p, which is a domain over $\bar{\mathbb{F}}_p$. Namely, there exists an R-order $B_R \subset B$ over some finitely generated subring $R \subset k$, and

$$B_p = B_{\psi,p} := B_R \otimes_R \bar{\mathbb{F}}_p$$

for a homomorphism $\psi : R \to \bar{\mathbb{F}}_p$. (For details on R-orders in B, see Section 2A below).

Recall that a ring A is *PI* if it satisfies a polynomial identity over $\mathbb{Z}$. By Posner's and Ore's theorems [Posner 1960; Ore 1931; McConnell and Robson 2001, Theorem 13.6.5 and Corollary 1.14], a domain A is PI if and only if it is an Ore domain and its division ring of fractions $\mathrm{Frac}(A)$ is a central division algebra. In this case, $\mathrm{Frac}(A)$ is a division ring that is dimension d^2 over its center, where d is the *PI degree* of A [McConnell and Robson 2001, Definition 13.6.7].

Definition 1.1. Given B as above, we say that B is an *algebra with PI reductions* if it admits an order B_R such that B_p is PI for sufficiently large p (with any choice of ψ).[1]

Theorem 2.4. *If B is an algebra with PI reductions, then any semisimple Hopf action on B factors through a group action.*

Note that when the Hopf action preserves the filtration of B, Theorem 2.4 (even without the PI reduction assumption) is proved in [Etingof and Walton 2014, Proposition 5.4]; our main achievement here is that we eliminate this requirement.

A basic example of an algebra with PI reductions is the Weyl algebra $B = \boldsymbol{A}_n(k)$, and, in fact, the proof of Theorem 2.4 is analogous to the proof of [Cuadra et al. 2015, Theorem 4.1], which addresses this case. Moreover, a wide range of filtered quantizations (each defined in Section 2B below) are algebras with PI reductions, resulting in the following corollary.

[1] It follows from Lemma 2.1(ii) below that if this condition holds for one pair (R, B_R), then the condition holds for all such pairs.

Corollary 2.6. *Let B be one of the following filtered k-algebras*:

(i) *any filtered quantization B generated in filtered degree one*; *in particular*, *the enveloping algebra $U(\mathfrak{g})$ of a finite dimensional Lie algebra $\mathfrak{g}$, or the algebra $D_\omega(X)$ of twisted differential operators on a smooth affine irreducible variety X*;

(ii) *a finite W-algebra or its quotient by a central character*;

(iii) *a quantum Hamiltonian reduction of a Weyl algebra by a reductive group action*; *in particular*, *the coordinate ring of a quantized quiver variety*;

(iv) *a spherical symplectic reflection algebra*; *or*

(v) *the tensor product of any of the algebras above with any commutative finitely generated domain over k.*

Then any semisimple Hopf action on B factors through a group action.

Other applications of Theorem 2.4 have been investigated recently by Lomp and Pansera [2015]; for instance, they establish no finite semisimple quantum symmetry on certain iterated differential operator rings.

Remark 1.2. We do not know if a filtered quantization of a finitely generated commutative domain over k must be an algebra with PI reductions (i.e., if the PI reduction assumption is, in fact, vacuous); see the question in [Cuadra et al. 2015, Introduction] and [Etingof 2016, Question 1.1]. This is of independent interest in noncommutative ring theory.

1B. ***Finite dimensional Hopf actions on filtered quantizations.*** Like [Cuadra et al. 2015, Theorem 4.2], Theorem 2.4 and hence Corollary 2.6 hold for Hopf–Galois actions of any (not necessarily semisimple) finite dimensional Hopf algebra (see Theorem 3.1). The proof is parallel to the proofs of Theorem 2.4 and [Cuadra et al. 2015, Theorem 4.2].

Moreover, it turns out that even without the Hopf–Galois assumption, Theorem 2.4 extends to nonsemisimple Hopf actions for a somewhat more restrictive class of quantizations. To see this, let us recall some algebras introduced in [Cuadra et al. 2016].

Notation 1.3 (B, B_{p^m}, C_m, D_{p^m}, Z, $Z(m)$). Let B be a quantization with PI reductions, and let B_{p^m} be the reduction of B modulo p^m. Let C_m be the center of B_{p^m}, $\mathrm{Frac}(C_m)$ be its ring of fractions, and $D_{p^m} := B_{p^m} \otimes_{C_m} \mathrm{Frac}(C_m)$. The PI reduction condition implies that D_{p^m} is the full localization (i.e., ring of fractions) of B_{p^m}. These algebras are defined over the truncated Witt ring $W_{m,p}$ of $\overline{\mathbb{F}}_p$; cf. [Cuadra et al. 2016, Sections 2.1, 2.3, 2.4]. Let Z be the center of the central division algebra D_p. (Here and below, to lighten the notation, we often suppress dependence on p.) Let Z_m be the center of D_{p^m}, and let $Z(m)$ be its image in D_p under the map $D_{p^m} \twoheadrightarrow D_p$ (so $Z(1) = Z$). It is easy to see that $Z(m) \subset Z(m-1)$.

Definition 1.4. We say that an algebra B with PI reductions is *nondegenerate* if for almost all p one has $\bigcap_{m\geq 1} Z(m) = \bar{\mathbb{F}}_p$.

Theorem 3.5. *If B is a nondegenerate algebra with PI reductions, then any finite dimensional Hopf action over B factors through a group action* (i.e., the condition that H is semisimple in Theorem 2.4 can be dropped).

The proof of Theorem 3.5 is similar to the proof of [Cuadra et al. 2016, Theorem 1.2].

To illustrate when the nondegeneracy condition holds, recall that $\mathrm{gr}(B)$ carries a natural Poisson bracket. Namely, if B is commutative, this bracket is zero; otherwise, if d is the largest integer such that $[F_i B, F_j B] \subset F_{i+j-d}B$, then for $a_0 \in \mathrm{gr}_i(B)$ and $b_0 \in \mathrm{gr}_j(B)$, the Poisson bracket $\{a_0, b_0\}$ is the projection of $[a, b]$ to $\mathrm{gr}_{i+j-d}(B)$, where $a \in F_i B$ and $b \in F_j B$ are any lifts of a_0 and b_0, respectively. Thus, $\mathrm{gr}(B) = O(X)$, where X is an irreducible Poisson algebraic variety.

The nondegeneracy assumption is satisfied, in particular, when X is a *generically symplectic* Poisson variety, i.e., one having a symplectic dense open subset; see Theorem 3.6. Therefore, Theorem 3.5 holds for many of the examples of Corollary 2.6 — quantum Hamiltonian reductions of Weyl algebras, central reductions of finite W-algebras, spherical symplectic reflection algebras, and tensor products thereof (see Corollary 3.7).

1C. ***Quantum polynomial algebras.*** For our next main result, we consider finite dimensional Hopf actions on *quantum polynomial algebras* (or *quantized coordinate rings of affine n-space*):

$$k_{\boldsymbol{q}}[x_1, \ldots, x_n] := k\langle x_1, \ldots, x_n\rangle/(x_i x_j - q_{ij} x_j x_i),$$

where $\boldsymbol{q} = (q_{ij})$, $q_{ij} \in k^\times$ with $q_{ii} = 1$ and $q_{ij}q_{ji} = 1$. Thus we can view $\boldsymbol{q}$ as a point of the algebraic torus $(k^\times)^{n(n-1)/2}$ with coordinates q_{ij} for $i < j$.

There are many examples of semisimple Hopf actions on $k_{\boldsymbol{q}}[x_1, \ldots, x_n]$ that do not factor through group actions; the parameters q_{ij} are roots of unity in these examples. See, for instance, [Chan et al. 2016, Theorem 0.4; Etingof and Walton 2014, Example 5.10; Kirkman et al. 2009, Examples 7.4–7.6]. Still, we establish the following result.

Let $\langle \boldsymbol{q}\rangle$ be the subgroup in $(k^\times)^{n(n-1)/2}$ generated by $\boldsymbol{q}$, and let $G_{\boldsymbol{q}}$ be its Zariski closure. Let $G_{\boldsymbol{q}}^0$ be the connected component of the identity in $G_{\boldsymbol{q}}$.

Theorem 1.5 (Theorem 4.1). *Let H be a semisimple Hopf algebra of dimension d. If the order of $G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0$ is coprime to $d!$, then any H-action on $B := k_{\boldsymbol{q}}[x_1, \ldots, x_n]$ factors through a group action.*

If each q_{ij} is a root of unity of order r_{ij}, then $|G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0| = \mathrm{lcm}\{r_{ij}\}_{i<j}$. In particular, if $n = 2$, i.e., if $B = k\langle x, y\rangle/(xy - qyx)$, then the condition on $\boldsymbol{q} = q \in k^\times$ in

Theorem 1.5 is that the order of q is coprime to $d!$ or infinite. On the other hand, the condition on $\boldsymbol{q}$ in Theorem 1.5 is also satisfied if each q_{ij} is not a root of unity and the set of the q_{ij} is multiplicatively independent; here, $|G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0| = 1$. See Example A.3 for a discussion of how to compute $|G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0|$ in general.

One may compare Theorem 1.5 to a similar result, Theorem 4.3 of [Chan et al. 2014], in the case where the Hopf action preserves the grading of $k_{\boldsymbol{q}}[x_1, \dots, x_n]$. But note that without the degree-preserving assumption, semisimplicity is still needed in Theorem 1.5; see [Etingof and Walton 2015; 2016, Example 3.6] for counterexamples for $n = 1, 3$, respectively.

Moreover, Theorem 1.5 is valid for finite dimensional Hopf algebras in the Hopf–Galois case, where we can replace the condition "coprime to $d!$" with "coprime to d" (Proposition 5.1). Also, Theorem 1.5 has a straightforward generalization (with the same proof) to actions on the quantum torus $k_{\boldsymbol{q}}[x_1^{\pm 1}, \dots, x_n^{\pm 1}]$.

Another generalization of Theorem 1.5 to the nonsemisimple case can be made under a nondegeneracy assumption. Recall that $\boldsymbol{q}$ may be viewed as a skew-symmetric bicharacter on $\mathbb{Z}^n$ with values in $k^\times$, with $\boldsymbol{q}(e_i, e_j) = q_{ij}$ for the standard basis $\{e_i\}$. A bicharacter $\boldsymbol{q}$ is called *nondegenerate* if the character $\boldsymbol{q}(a, \cdot) : \mathbb{Z}^n \to k^\times$ is nontrivial whenever $a \neq 0$. Note that unlike skew-symmetric bilinear forms (which are always degenerate in odd dimensions), a skew-symmetric bicharacter can be nondegenerate for any $n \geq 2$.

Theorem 5.2. *Let H be a finite dimensional Hopf algebra of dimension d acting on $B := k_{\boldsymbol{q}}[x_1, \dots, x_n]$. Assume that the order of $G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0$ is coprime to $d!$, and $\boldsymbol{q}$ is a nondegenerate bicharacter. Then, the action of H on B factors through a group algebra.*

It is shown in Example 5.5 that Theorems 1.5 and 5.2 fail when hypotheses are removed; these examples involve actions of the nonsemisimple 4-dimensional Sweedler Hopf algebra.

1D. ***Twisted homogeneous coordinate rings of abelian varieties and Sklyanin algebras.*** Let X be an abelian variety over k, let $\mathcal{L}$ be an ample line bundle on X, and let $\sigma : X \to X$ be an automorphism given by translation by a point $s \in X$. Then we can define the twisted homogeneous coordinate ring

$$B(X, \sigma, \mathcal{L}) := \bigoplus_{n=0}^{\infty} H^0\left(X, \bigotimes_{i=0}^{n-1} (\sigma^i)^* \mathcal{L}\right),$$

with twisted multiplication $f * g := f(\sigma^n)^*(g)$, where f is of degree n [Artin and Van den Bergh 1990]. It is well-known that $B(X, \sigma, \mathcal{L})$ is a domain, and if $|\sigma| < \infty$, then $B(X, \sigma, \mathcal{L})$ is a PI domain of PI degree $|\sigma|$.

Let G_σ be the Zariski closure of the subgroup $\{s^i\}_{i \in \mathbb{Z}}$, and let G_σ^0 be the connected component of the identity in G_σ.

Theorem 1.6 (Theorem 6.1). *If H is a semisimple Hopf algebra of dimension d, and if the order of G_σ/G_σ^0 is coprime to $d!$, then any H-action on $B(X, \sigma, \mathcal{L})$ factors through a group action.*

In particular, if the subgroup $\{s^i\}_{i\in\mathbb{Z}}$ is Zariski-dense in X, then any semisimple Hopf action on $B(X, \sigma, \mathcal{L})$ factors through a group action. Moreover, if $X =: E$ is an elliptic curve, the condition on σ in Theorem 6.1 is that the order of σ is coprime to $d!$ or infinite.

Lastly, we study semisimple Hopf actions on another class of quantizations: the *3-dimensional Sklyanin algebras* $S(a, b, c)$ (Definition 6.3). To $S(a, b, c)$, one can naturally associate an elliptic curve $E_{abc} \subset \mathbb{P}^2_k$ and an automorphism σ_{abc} given by translation by a point; see [Artin et al. 1990, Introduction].

Theorem 1.7. *If H is a semisimple Hopf algebra of dimension d, and if the order of σ_{abc} is coprime to $d!$ or infinite, then any H-action on $S(a, b, c)$ factors through a group action.*

Remark 1.8. We believe that by adapting the techniques in this work, one could establish a version of Theorem 1.7 for semisimple Hopf actions on other elliptic algebras, such as in [Sklyanin 1982] (or, see [Smith and Stafford 1992]) and in [Etingof and Ginzburg 2010; Odesskiĭ 2002; Stephenson 1997]. Further, we believe that under an appropriate nondegeneracy condition, there are no finite quantum symmetry results for such elliptic algebras and for twisted homogeneous coordinate rings $B(X, \sigma, \mathcal{L})$ as above; compare to Theorem 5.2.

Our paper is organized as follows. We discuss semisimple Hopf actions on filtered quantizations in Section 2, nonsemisimple Hopf actions on filtered quantizations in Section 3, semisimple Hopf actions on quantum polynomial algebras in Section 4, nonsemisimple Hopf actions on quantum polynomial algebras in Section 5, and Hopf actions on twisted homogeneous coordinate rings of abelian varieties and Sklyanin algebras in Section 6. The results of Sections 4–6 rely on a number-theoretic result of Antonella Perucca discussed in the Appendix.

The notation and terminology of the introduction is used throughout this work, often without mention.

2. Semisimple Hopf actions on filtered quantizations

2A. ***The result on semisimple Hopf actions on quantizations with PI reductions.*** In this section, we study actions of semisimple Hopf algebras H on filtered quantizations B. Throughout this section, we let B denote a $\mathbb{Z}_+$-filtered algebra over k such that $\mathrm{gr}(B)$ is a commutative finitely generated domain; such B will be referred to as a filtered quantization.

Our goal is to prove Theorem 2.4. This result was established in [Cuadra et al. 2015] for B a Weyl algebra, and we generalize those techniques for our purpose here.

Let R be a finitely generated subring of k. By an *R-order* in a filtered quantization B we mean an R-subalgebra B_R of B such that $\mathrm{gr}(B_R)$ is a finitely generated R-algebra which is projective as an R-module, and such that the natural map $\mathrm{gr}(B_R) \otimes_R k \to \mathrm{gr}(B)$ is an isomorphism of graded k-algebras.

Lemma 2.1. (i) *Any filtered quantization B admits an R-order B_R for a suitable ring R.*

(ii) *For any two orders B_R over R and $B_{R'}$ over R', there exists a finitely generated ring $R'' \subset k$ containing R and R', and admitting an R''-algebra isomorphism $\phi : B_R \otimes_R R'' \to B_{R'} \otimes_{R'} R''$ such that $\mathrm{gr}\,\phi$ is an isomorphism.*

Proof. (i) By the Hilbert basis theorem, the algebra $\mathrm{gr}(B)$ is finitely presented. This implies that so is B, as we can lift the generators and defining relations of $\mathrm{gr}(B)$ to those of B.

More specifically, pick homogeneous generators $\bar{b}_1, \ldots, \bar{b}_n$ of $\mathrm{gr}(B)$ of degrees $m_1, \ldots, m_n$. Let

$$p_s(\bar{b}_1, \ldots, \bar{b}_n) = 0, \quad s = 1, \ldots, r$$

be a set of defining relations for $\mathrm{gr}(B)$, with $p_s \in k[X_1, \ldots, X_n]$ being homogeneous of degree d_s (this set may be chosen to be finite by the Hilbert basis theorem). Let b_j be lifts of $\bar{b}_j$ to B, and $\tilde{p}_s$ be homogeneous lifts of p_s to $k\langle X_1, \ldots, X_n\rangle$. Then $[b_i, b_j] = f_{ij}(b_1, \ldots, b_n)$, where $f_{ij} \in k\langle X_1, \ldots, X_n\rangle$ is a noncommutative polynomial of filtration degree at most $m_i + m_j - 1$, and $\tilde{p}_s(b_1, \ldots, b_n) = p'_s(b_1, \ldots, b_n)$, where $p'_s \in k\langle X_1, \ldots, X_n\rangle$ is a noncommutative polynomial of filtration degree at most $d_s - 1$.

Let $g_s := \tilde{p}_s - p'_s \in k\langle X_1, \ldots, X_n\rangle$. Thus, we have relations

$$[b_i, b_j] = f_{ij}(b_1, \ldots, b_n) \quad \text{and} \quad g_s(b_1, \ldots, b_n) = 0 \tag{1}$$

in B. It is easy to see that these relations are defining, since they are already defining at the graded level.

Using relations (1), we can find a suitable finitely generated subring $R \subset k$ and define $B_R \subset B$ as follows. We take $\widetilde{R}$ to be the ring generated by all the coefficients of the polynomials f_{ij}, g_s, and set $R = \widetilde{R}[1/f]$ for a suitable $f \in \widetilde{R}$. Now let B_R be the subalgebra of B generated over R by $b_1, \ldots, b_n$. We can choose f so that (1) are defining relations for B_R, and so that B_R is an R-order on B, since for a suitable choice of f, $\mathrm{gr}(B_R)$ is a projective (in fact, free) R-module by Grothendieck's generic freeness lemma [Eisenbud 1995, Theorem 14.4]. This proves (i).

(ii) Note that we have a natural isomorphism of filtered algebras

$$\widetilde{\phi} : B_R \otimes_R k \to B_{R'} \otimes_{R'} k$$

(as both are equal to B). This isomorphism is defined over some finitely generated ring $R'' \subset k$ containing R and R', which implies (ii). □

Lemma 2.2. *Suppose that B is a filtered quantization that carries an action of a finite dimensional Hopf algebra H. Let S be a finitely generated subring of k, and B_S be an S-order in B. Then one can find a finitely generated subring $R \subset k$ containing S and a Hopf order $H_R \subset H$* (cf. [Cuadra et al. 2015, Definition 2.1]), *so that there is an induced action of H_R on $B_R := B_S \otimes_S R$ which gives the action of H on B upon tensoring over R with k.*

Proof. We use the method of [Cuadra et al. 2015, Section 2]. Pick homogeneous generators $\bar{b}_1, \dots, \bar{b}_n$ of $\mathrm{gr}(B_S)$, and let b_j be lifts of $\bar{b}_j$ to B_S. Choose a basis $\{h_m\}$ of H. We have

$$h_m \cdot b_j = q_{mj}(b_1, \dots, b_n) \tag{2}$$

for some noncommutative polynomials $q_{mj} \in k\langle X_1, \dots, X_n\rangle$. Let R be generated over S by the structure constants of H in the basis $\{h_m\}$ and the coefficients of q_{mj}. Let $H_R \subset H$ be the span of h_m over R. Then, $H_R \subset H$ is a Hopf order, and H_R acts on B_R by formula (2). The lemma is proved. □

Thus, any action of H on B admits an R-order for some finitely generated ring $R \subset k$. Moreover, it is easy to see from Lemma 2.1(ii) that any two such orders over rings R and R' can be identified after tensoring with some finitely generated ring $R'' \subset k$ containing R and R', so an order is essentially unique.

Now fix a ring R and an R-order $B_R \subset B$ with an action of H_R. Let p be a sufficiently large prime, and $\psi : R \to \bar{\mathbb{F}}_p$ be a character. Following [Cuadra et al. 2015, Section 2], let $H_p := H_R \otimes_R \bar{\mathbb{F}}_p$, $B_p := B_R \otimes_R \bar{\mathbb{F}}_p$ be the corresponding reductions of H, B modulo p, where $\bar{\mathbb{F}}_p$ is an R-module via ψ. Then, H_p acts on B_p (by applying ψ to the action of H_R on B_R).

Lemma 2.3. *For a sufficiently large prime p, $\mathrm{gr}(B_p)$, and hence B_p, is a domain.*

Proof. We have $\mathrm{gr}(B_p) = \mathrm{gr}(B_R) \otimes_R \bar{\mathbb{F}}_p$. Therefore, the statement follows from [Grothendieck 1966, 9.7.7(i)] ("geometric irreducibility"). □

Theorem 2.4. *If B is an algebra with PI reductions, then any semisimple Hopf action on B factors through a group action.*

Proof. We may assume without loss of generality that the action of H on B is inner faithful (otherwise we can pass to an action of a quotient Hopf algebra).

Take $p \gg 0$. Then by [Cuadra et al. 2015, Proposition 2.4] (which applies with the same proof in our more general situation), H_p acts inner faithfully on B_p.

Moreover, as in [Cuadra et al. 2015, Lemma 2.5], H_p is a semisimple cosemisimple Hopf algebra over $\overline{\mathbb{F}}_p$.

Since B is an algebra with PI reductions, by Lemma 2.3, the algebra B_p is a PI domain. Let D_p be the division algebra of quotients of B_p. Then by [Etingof 2016, Corollary 3.2(ii)], D_p is a central division algebra of degree p^n for some $n \geq 0$ (which may depend on p). Moreover, H_p acts inner faithfully on D_p by [Skryabin and Van Oystaeyen 2006, Theorem 2.2].

Since $\deg D_p = p^n$ is coprime to $(\dim H)!$, [Cuadra et al. 2015, Proposition 3.3(ii)] implies that H_p is cocommutative. Thus, H is cocommutative (as in the proof of [Cuadra et al. 2015, Theorem 4.1]), and thus is a group algebra. □

2B. ***Some examples of filtered quantizations.*** As a consequence, Theorem 2.4 applies to semisimple Hopf actions on many classes of filtered quantizations. Namely, we will consider the following examples.

Twisted differential operators. Let X be a smooth affine irreducible algebraic variety over k, and ω a closed 2-form on X. Then we define the *algebra of twisted differential operators* $D_\omega(X)$ to be the algebra generated by $O(X)$ and elements L_v attached k-linearly to vector fields $v \in \mathrm{Der}O(X)$ on X, with defining relations

$$L_{fv} = fL_v, \quad [L_v, f] = v(f), \quad [L_v, L_w] = L_{[v,w]} + \omega(v, w)$$

for $f \in O(X)$, $v, w \in \mathrm{Der}\, O(X)$. Then $D_\omega(X)$ carries a filtration defined by $\deg O(X) = 0$ and $\deg L_v = 1$ for $v \in \mathrm{Der}\, O(X)$, and $\mathrm{gr}(D_\omega(X)) = O(T^*X)$, the algebra of functions on the symplectic variety T^*X.

The filtered algebra $D_\omega(X)$ depends only on the cohomology class $[\omega]$ of ω, and if $[\omega] = 0$, then $D_\omega(X) = D(X)$, the algebra of usual differential operators on X. For more on twisted differential operators, see, e.g., [Beĭlinson and Bernstein 1993, Section 2].

Quantum Hamiltonian reductions. Let G be a reductive algebraic group over k with Lie algebra $\mathfrak{g}$, and let $(V, (\cdot, \cdot))$ be a faithful finite dimensional symplectic representation of G. Let $\boldsymbol{A}(V)$ be the Weyl algebra of V, generated by $v \in V$ with relations $[v, w] = (v, w)$ for $v, w \in V$. We have a natural action of G on $\boldsymbol{A}(V)$ which preserves its filtration. In this case, we have a natural G-equivariant Lie algebra map $\hat{\mu} : \mathfrak{g} \to \boldsymbol{A}(V)$ called the *quantum moment map*, which quantizes the classical moment map $\mu : V \to \mathfrak{g}^*$, where $\mu(v)(a) = \frac{1}{2}(v, av)$ for $v \in V$, $a \in \mathfrak{g}$. Now, given a character $\chi : \mathfrak{g} \to k$, we can define the algebra

$$B(\chi) := [\boldsymbol{A}(V)/\boldsymbol{A}(V)(\hat{\mu}(a) - \chi(a),\ a \in \mathfrak{g})]^G,$$

called the *quantum Hamiltonian reduction of* $\boldsymbol{A}(V)$ *by* G *using* χ. It inherits a filtration from the Weyl algebra. See [Etingof 2007, Chapter 4] for further details.

Assume that the moment map μ is flat, and that the scheme $\mu^{-1}(0)$ is reduced and irreducible (i.e., $\mu^{-1}(0)$ is a reduced irreducible complete intersection). In this case, $X := \mu^{-1}(0)/G$ is an irreducible generically symplectic Poisson variety, and $B(\chi)$ is a filtered quantization of $O(X)$. See [Losev and Etingof 2015, Section 2.3; Nakajima 2015, Section 2(i)], and references therein for more details.

An interesting special case of this is when

$$G = \left(\prod_{i=1}^{r} \mathrm{GL}(V_i)\right)/k^\times \quad \text{and} \quad V = \bigoplus_{i,j} (V_i \otimes V_j^*)^{\oplus m_{ij}},$$

where V_i are finite dimensional vector spaces and $m_{ij} = m_{ji}$ are positive integers with m_{ii} even; i.e., V is the space of representations of a doubled quiver, and G is the group of linear transformations for this quiver. In this case, $B(\chi)$ is the *quantized quiver variety*; see, e.g., [Braden et al. 2012, Section 3.4]. The conditions under which μ is flat and $\mu^{-1}(0)$ is reduced and irreducible are given in [Crawley-Boevey 2001, Theorems 1.1 and 1.2].

Finite W-algebras. Let $\mathfrak{g}$ be a simple Lie algebra over k, and $e \in \mathfrak{g}$ a nilpotent element. To this data one can attach a Lie subalgebra $\mathfrak{m} \subset \mathfrak{g}$ with a character χ, and a *finite W-algebra* is

$$U(\mathfrak{g}, e) := \big(U(\mathfrak{g})/U(\mathfrak{g})(a - \chi(a),\ a \in \mathfrak{m})\big)^{\mathrm{ad}\,\mathfrak{m}},$$

a quantum Hamiltonian reduction of $U(\mathfrak{g})$. The algebra $U(\mathfrak{g}, e)$ has a filtration induced by the filtration in $U(\mathfrak{g})$, and its associated graded algebra is a polynomial algebra (of functions on the corresponding Slodowy slice). We refer the reader to [Losev 2010, Sections 2.3 and 2.4] and the references therein for details.

Also, the center $U(\mathfrak{g})^{\mathfrak{g}}$ of $U(\mathfrak{g})$ embeds into $U(\mathfrak{g}, e)$, so for any central character $\theta : U(\mathfrak{g})^{\mathfrak{g}} \to k$, one can consider the central reduction

$$U_\theta(\mathfrak{g}, e) := U(\mathfrak{g}, e)/(a - \theta(a),\ a \in U(\mathfrak{g})^{\mathfrak{g}}).$$

Then $\mathrm{gr}(U_\theta(\mathfrak{g}, e)) = O(X)$, where X is the nilpotent Slodowy slice, a generically symplectic Poisson variety.

Symplectic reflection algebras. Let G be a finite group and V a faithful finite dimensional symplectic representation of G, and assume that V is not a direct sum of two nonzero symplectic representations. The *symplectic reflection algebra* $H_{t,c}(G, V)$ is the most general filtered deformation of $kG \ltimes SV$, where $[F_i, F_j] \subset F_{i+j-2}$; here $t \in k$, and c is a conjugation invariant function on the set of symplectic reflections in G; see [Etingof 2007, Chapter 8].

Let $\boldsymbol{e} = |G|^{-1} \sum_{g \in G} g$ be the symmetrizing idempotent for G. Then, the algebra $\boldsymbol{e} H_{t,c}(G, V) \boldsymbol{e}$ is called the *spherical symplectic reflection algebra*. For $t = 1$, it is a

filtered quantization of $O(X)$, where $X = V/G$, a generically symplectic Poisson variety.

Remark 2.5. There are many other interesting examples of filtered quantizations, and our results given below can be extended to such examples. Since this extension is rather routine, we leave it outside the scope of this paper.

2C. ***Results on semisimple Hopf actions on specific filtered quantizations.*** Here are some concrete applications of Theorem 2.4.

Corollary 2.6. *Let B be one of the following filtered k-algebras*:

(i) *any filtered quantization B generated in filtered degree one*; *in particular*, *the enveloping algebra $U(\mathfrak{g})$ of a finite dimensional Lie algebra $\mathfrak{g}$, or the algebra $D_\omega(X)$ of twisted differential operators on a smooth affine irreducible variety X*;

(ii) *a finite W-algebra or its quotient by a central character*;

(iii) *a quantum Hamiltonian reduction of a Weyl algebra by a reductive group action*; *in particular*, *the coordinate ring of a quantized quiver variety*;

(iv) *a spherical symplectic reflection algebra*; *or*

(v) *the tensor product of any of the algebras above with any commutative finitely generated domain over k.*

Then any semisimple Hopf action on B factors through a group action.

Note that in some of these cases, a stronger statement is true: any finite dimensional (not necessarily semisimple) Hopf action on B factors through a group action; see Corollary 3.7 below. However, we still prefer to prove the weaker version here, since the proof is simpler (e.g., it does not require reduction modulo prime powers).

Proof. By Theorem 2.4, our job is to show that B is an algebra with PI reductions. In other words, we need to show that the division algebra D_p of fractions of B_p is central (i.e., there is a "p-center") for $p \gg 0$. We do so below in each case.

(i) We show that if a filtered quantization A of a commutative finitely generated domain A_0 over a field F of characteristic $p > 0$ is generated in degree one, then it is module-finite over its center after localization; this implies the required statement.

Let $A_0[i]$ be the degree i part of A_0. Then $A_0[0] = A[0]$ is a finitely generated commutative domain. Let $\bar{a}_1, \dots, \bar{a}_n$ be generators of $A_0[1]$ as an $A_0[0]$-module. Let a_i be lifts of $\bar{a}_i$ to A. Then, a_i and $A[0]$ generate A as an algebra. Also, the operators $[a_i, \cdot\,]$ are derivations of $A[0]$, and hence vanish on $A[0]^p$. Thus, $A[0]^p$ is central in A. Let K be the field of quotients of $A[0]^p$, and let $A' := A \otimes_{A[0]^p} K$. The K-algebra A' is generated in filtration degree 1, and $L := F_1 A'$ is a finite dimensional vector space over K (as it is spanned by $1, a_1, \dots, a_n$ over $A[0] \otimes_{A[0]^p} K$, and

$A[0]$ is module-finite over $A[0]^p$ as $A[0]$ is a finitely generated algebra). Also, L is closed under commutator. Thus, L is a finite dimensional Lie algebra over K, and A' is a quotient of the enveloping algebra $U(L)$. But the enveloping algebra of a finite dimensional Lie algebra in characteristic p is module-finite over its center (i.e., there is a p-center; see [Jacobson 1952; 1962, Chapter 6, Lemma 5]). This implies that A' is module-finite over its center, as desired.

(ii) Since a W-algebra is a quantum Hamiltonian reduction of the enveloping algebra $U(\mathfrak{g})$ of a semisimple Lie algebra $\mathfrak{g}$ [Losev 2010], the statement follows from (i).

(iii) This also follows from (i) and the definition of the quantum Hamiltonian reduction.

(iv) This holds by [Etingof 2006, Theorem 9.1.1 (in the appendix)].

(v) This follows easily from the previous cases. □

3. Finite dimensional Hopf actions on filtered quantizations

3A. ***Hopf–Galois actions.*** Theorem 2.4 does not hold for nonsemisimple Hopf actions, as there are many inner faithful actions of nonsemisimple finite dimensional Hopf algebras on commutative domains; see [Etingof and Walton 2015]. However, Theorem 2.4 is valid in the Hopf–Galois case.

Theorem 3.1. *Let B be a filtered quantization of a commutative finitely generated domain with PI reductions, and let H be a finite dimensional Hopf algebra over k which acts on B. Assume that this action gives rise to an H^*-Hopf–Galois extension $B^H \subset B$. Then H is a group algebra.*

Proof. The result follows from the arguments in the proofs of Theorem 2.4 and [Cuadra et al. 2015, Theorem 4.2]. Namely, recall Notation 1.3. We show, similarly to the proof of Theorem 2.4, that Z is H_p-stable, and then proceed as in the proof of [Cuadra et al. 2015, Theorem 4.2]. Specifically, by [Etingof 2016, Corollary 3.2(ii)], D_p has degree p^n over its center $Z = Z(D_p)$ for some n, so by [Cuadra et al. 2015, Proposition 3.3(i)], Z is H_p-invariant. Now, since the action of H on B gives rise to a Hopf–Galois extension, so does the action of H_p on Z, i.e., the algebra map $Z \otimes_{Z^{H_p}} Z \to Z \otimes H_p^*$ is an isomorphism. Thus, H_p^* is commutative and H_p is cocommutative, so H is cocommutative [Cuadra et al. 2015, Lemma 2.3(ii)], i.e., a group algebra by the Cartier–Gabriel–Kostant theorem [Montgomery 1993, Corollary 5.6.4(3) and Theorem 5.6.5]. □

3B. ***Preparatory results on nondegenerate quantizations.*** Another generalization of Theorem 2.4 concerns nondegenerate quantizations, defined in Definition 1.4. To obtain it, we first need to generalize [Cuadra et al. 2016, Theorem 3.2]. Let $\mathcal{H}$ be a finite dimensional Hopf algebra over an algebraically closed field F of

characteristic $p > 0$, and let $\mathcal{Z}$ be a finitely generated field extension of F. Let $\mathcal{Z}(m)$, for $m \geq 1$, be a collection of subfields of $\mathcal{Z}$ such that $\mathcal{Z}(m) \supset \mathcal{Z}(m+1)$ for all $m \geq 1$.

Theorem 3.2. *Suppose that* $\bigcap_{m\geq 1} \mathcal{Z}(m) = F$, *and that* $[\mathcal{Z} : \mathcal{Z}(m)]$ *is a power of* p *for all* $m \geq 1$. *Assume that* $\mathcal{H}$ *acts* F*-linearly and inner faithfully on* $\mathcal{Z}$. *If* $p > \dim \mathcal{H}$ *and* $\mathcal{H}$ *preserves* $\mathcal{Z}(m)$ *for all* m, *then* $\mathcal{H}$ *is a group algebra.*

Proof. The proof is the same as that of [Cuadra et al. 2016, Theorem 3.2]. Indeed, the only properties of the fields Z^{p^m} used in that proof are that their intersection is F and that the degree of Z over Z^{p^m} is a power of p. □

We will also need the lemma below from commutative algebra. We first introduce the following notation. Let $W_N = W_N(F) := W(F)/(p^N)$ be the N-th truncated Witt ring of F (W_N is an algebra over $\mathbb{Z}/p^N\mathbb{Z}$; cf. [Cuadra et al. 2016, Subsection 2.1]). Let Y be an irreducible smooth affine algebraic variety over F with structure algebra $A := O(Y)$, and $\widetilde{Y}$ be a flat deformation of Y over W_N. Let $1 \leq m \leq N$, and let $A_m := O(\widetilde{Y})/(p^m)$ (a free $\mathbb{Z}/p^m\mathbb{Z}$-module); thus $A_1 = A$ and $A_{m-1} = A_m/(p^{m-1})$ for $m \geq 2$. Let

$$d_m : A_m \to \Omega_{A_m/W_m}$$

be the differential.

Lemma 3.3. *For* $1 \leq m \leq N$, *the image of* $\ker(d_m)$ *in* A *is* A^{p^m}.

Proof. It is clear that the image of $\ker(d_m)$ contains A^{p^m}, so it remains to establish the opposite inclusion. We do so by induction in m.

The base of induction is the equality $\ker(d|_A) = A^p$, which is the Cartier isomorphism in degree zero [Katz 1970, Section 7]. Alternatively, here is a direct proof. Since A is integrally closed in its quotient field $L := \mathrm{Frac}(A)$, we may replace A with L. Note that L can be represented as a finite separable extension of $F(y_1, \ldots, y_n)$, where $n = \dim Y$. Given $f \in L$ such that $df = 0$, consider the minimal polynomial $P(t) = t^r + a_{r-1}t^{r-1} + \cdots + a_0$ of f over $E := F(y_1, \ldots, y_n)$. Applying the differential to the equation $P(f) = 0$, we get $\sum_{j=0}^{r-1} f^j da_j = 0$. Since P is the minimal polynomial, this implies that $da_j = 0$ for all j. Thus $a_j \in E^p$ (as the statement in question is easy for purely transcendental fields). Thus, $E^p(f)$ is a finite separable extension of E^p (as P is a separable polynomial). But $E^p(f)$ is a purely inseparable extension of $E^p(f^p)$. Hence, $E^p(f) = E^p(f^p)$, that is, $f \in E^p(f^p) \subset L^p$, as desired.

To perform the induction step, suppose $f \in \ker(d_m)$. Our job is to show that the image $\bar{f}$ of f in A is contained in A^{p^m}. By the induction assumption we know that $\bar{f} = b^{p^{m-1}}$ for some $b \in A$, so it remains to show that $b = c^p$ for some $c \in A$.

For this, we expand f in a power series in some local coordinate system $y_1, \ldots, y_n$ on $\widetilde{Y}$. It is easy to see by looking at monomials that if $g \in W_m[\![y_1, \ldots, y_n]\!]$

and $dg = 0$, then the reduction $\bar{g}$ of g modulo p lies in $F[\![y_1^{p^m}, \ldots, y_n^{p^m}]\!]$. In particular, $F[\![y_1^{p^m}, \ldots, y_n^{p^m}]\!]$ contains the power series expansion of $\bar{f}$ in y_i. This means that the power series expansion of b is in $F[\![y_1^p, \ldots, y_n^p]\!]$. Thus, $db = 0$. By the base of induction we conclude that $b = c^p$ for some $c \in A$, which completes the induction step. □

Moreover, we will need the result below.

Lemma 3.4. *Let B be an algebra with PI reductions, and let D_p denote the full localization (i.e., the ring of fractions) of the reduction B_p of B, for $p \gg 0$. Then the center Z of D_p is a finitely generated field extension of $\bar{\mathbb{F}}_p$.*

Proof. Let $v_1, \ldots, v_N$ be a basis of D_p over Z, and let $b_1, \ldots, b_n$ be generators of B_p. Then $b_s v_i = \sum_{j=1}^N \beta_{si}^j v_j$ for $\beta_{si}^j \in Z$. Let K denote the field $\bar{\mathbb{F}}_p(\beta_{si}^j)$.

Now take $z \in Z$. Then $z \in D_p$, so $z = c^{-1}b$, and hence $cz = b$ for some $b, c \in B_p$ with $c \neq 0$. Since $b, c \in B_p$, they are noncommutative polynomials in $b_1, \ldots, b_n$ over $\bar{\mathbb{F}}_p$. So, $bv_i = \sum \beta_i^j v_j$, $cv_i = \sum \gamma_i^j v_j$, with $\beta_i^j, \gamma_i^j \in K$. But $\gamma_i^j z = \beta_i^j$ and γ_i^j are not all zero. So, $z \in K$ and hence $Z = K$. Thus, Z is a finitely generated extension of $\bar{\mathbb{F}}_p$. □

3C. ***Hopf actions on nondegenerate quantizations.*** Now let B be a filtered quantization with PI reductions.

Theorem 3.5. *If B is a nondegenerate algebra with PI reductions, then any finite dimensional Hopf action on B factors through a group action* (i.e., the condition that H is semisimple in Theorem 2.4 can be dropped).

Proof. The proof is obtained by combining the proofs of Theorem 2.4 and [Cuadra et al. 2016, Theorem 1.1] with the following modifications.

1. In [Cuadra et al. 2016, Lemma 2.5] and below, x_i, y_i should be replaced by any finite set of generators $L_1, \ldots, L_r$ of B, and the number $2n$ in the proof of [Cuadra et al. 2016, Lemma 4.3] should be replaced by r (cf. [Cuadra et al. 2016, proof of Theorem 1.2]).

2. The discussion in [Cuadra et al. 2016, Subsection 2.4, Lemma 4.7, Proposition 4.8] (needed to justify the assumptions of [Cuadra et al. 2016, Theorem 3.2]) becomes unnecessary. Instead, note that if $a \in D_{p^m}$ is central modulo p^{m-1} for some $m \geq 2$, then a^p is central. Hence $Z(m) \supset Z(m-1)^p$, implying that $Z(m) \supset Z^{p^{m-1}}$ and therefore $[Z : Z(m)]$ is finite (by Lemma 3.4) and is a power of p. Now the proof proceeds by invoking Theorem 3.2, whose assumptions are satisfied by the nondegeneracy property of B and using a straightforward generalization of [Cuadra et al. 2016, Lemma 4.6].

3. The rest of the proof of [Cuadra et al. 2016, Theorem 1.1] is modified as in the proof of Theorem 2.4. Namely, we use the PI reduction condition and [Etingof 2016, Corollary 3.2], which says that the PI degree of B_p is a power of p. □

The next theorem shows that the nondegeneracy assumption is satisfied, in particular, when $\mathrm{gr}(B) = O(X)$, where X is generically symplectic.

Theorem 3.6. *Let B be a quantization with PI reductions. Assume $\mathrm{gr}(B) = O(X)$, where X is a generically symplectic Poisson variety. Then any action of a finite dimensional Hopf algebra H on B factors through a group action.*

Proof. By Theorem 3.5, it suffices to show that B is a nondegenerate quantization, i.e., that $\bigcap_{m\geq 1} Z(m) = \bar{\mathbb{F}}_p$ for $p \gg 0$.

Recall Notation 1.3. Let C be the center of B_p; thus, by Posner's theorem [McConnell and Robson 2001, Theorem 13.6.5], the field $\mathrm{Frac}(C)$ of fractions of C is Z. Let C_m be the center of B_{p^m}, and $C(m)$ be its image in B_p.

Let $a \in B_{p^m}$ be central modulo p (i.e., the image $\bar{a}$ of a in B_p lies in C). Then a^p is central modulo p^2, a^{p^2} is central modulo p^3, and so on. Hence, $C^{p^{m-1}} \subset C(m)$. Let C'_m be the preimage of $C^{p^{m-1}}$ in C_m. Then the image of C'_m in B_p is $C^{p^{m-1}}$.

We claim that

$$Z(m) = \mathrm{Frac}(C(m)). \tag{3}$$

Indeed, it is clear that $\mathrm{Frac}(C(m)) \subseteq Z(m)$. On the other hand, observe that any element $a \in D_{p^m}$ can be written as $a = c^{-1}b$, where $c \in C'_m$ is nonzero modulo p, and $b \in B_{p^m}$ (as this can be done modulo p, since $D_p = Z^{p^{m-1}} B_p$). Now given $z \in Z(m)$, let $\tilde{z}$ be its lift to Z_m. Writing $\tilde{z} = c^{-1}b$ as above, we see that $b := c\tilde{z} \in C_m$. Let $\bar{b} \in C(m)$ and $\bar{c} \in C^{p^{m-1}} \subset C(m)$ be the reductions of b and c modulo p, respectively. We have $\bar{b} = \bar{c}z$, hence $z = \bar{c}^{-1}\bar{b} \in \mathrm{Frac}(C(m))$, as claimed.

Now let $B_{0p^m} := \mathrm{gr}(B_{p^m})$. This is a Poisson algebra over the truncated Witt ring $W_{m,p}$. Let C_{0m} be the Poisson center of B_{0p^m}, and $C_0(m)$ be the image of C_{0m} in B_{0p}. Then $\mathrm{gr}(C_m) \subset C_{0m}$ and hence

$$\mathrm{gr}(C(m)) \subset C_0(m). \tag{4}$$

Let $Z_0 := \mathrm{Frac}(B_{0p})$. Since X is generically symplectic, C_{0m} coincides with the set of all $f \in B_{0p^m}$ such that $df = 0$. By Lemma 3.3 (taking $\widetilde{Y}$ to be the reduction modulo p^m of a symplectic dense affine open subset $U \subset X$), this implies that

$$\mathrm{Frac}(C_0(m)) \subset Z_0^{p^m}. \tag{5}$$

Now suppose that $z \in \bigcap_{m\geq 1} Z(m)$ with $z \neq 0$. Then by (3), for each m, we have $z = f_m/g_m$ for $f_m, g_m \in C(m)$. Let $f_m^0, g_m^0 \in \mathrm{gr}(C(m))$ be the leading terms of f_m, g_m. By (4), $f_m^0, g_m^0 \in C_0(m)$. Then for any m, n we have $f_m^0 g_n^0 = f_n^0 g_m^0$ since $f_m g_n = f_n g_m$. So $z_0 := f_m^0/g_m^0$ is independent of m and by (5) belongs to $Z_0^{p^m}$ for all $m \geq 0$. As $\bigcap_{m\geq 1} Z_0^{p^m}$ is a perfect field that is finitely generated over $\bar{\mathbb{F}}_p$, we get

that $\bigcap_{m\geq 1} Z_0^{p^m} = \bar{\mathbb{F}}_p$. So, $z_0 \in \bar{\mathbb{F}}_p$ is a nonzero constant, and $f_m^0 = z_0 g_m^0$ for all m; in particular,

$$\deg(f_m) = \deg(g_m). \tag{6}$$

Now $z - z_0 = (f_m - z_0 g_m)/g_m$, and the numerator has degree strictly less than $\deg g_m$. This violates (6), so $z - z_0 = 0$, i.e., $z \in \bar{\mathbb{F}}_p$. This proves the theorem. □

Corollary 3.7. *Let B be one of the following algebras*:

(i) *a quotient of a finite W-algebra by a central character*;

(ii) *a Hamiltonian reduction of a Weyl algebra by a reductive group action*; *in particular*, *the coordinate ring of a quantized quiver variety*;

(iii) *a spherical symplectic reflection algebra $H_{1,c}(G, V)$*; *or*

(iv) *the tensor product of any of the algebras in (i)–(iii).*

Then any action of a finite dimensional Hopf algebra H on B factors through a group action.

Proof. It is explained in Section 2B that in examples (i)–(iv), we have $\mathrm{gr}(B) = O(X)$, where X is generically symplectic. This implies the corollary. □

Proposition 3.8. *Theorems 2.4, 3.1, 3.5, and 3.6 remain valid if B is replaced by its quotient division algebra* $\mathrm{Frac}(B)$.

Proof. The proofs are obtained by combining the proofs of Theorems 2.4, 3.1, 3.5, and 3.6 with the proof of [Cuadra et al. 2015, Proposition 4.4]. (The exact form of the generators of B used in the proof of [Cuadra et al. 2015, Proposition 4.4] is irrelevant for the argument.) □

4. Semisimple Hopf actions on quantum polynomial algebras

We now turn to finite dimensional Hopf actions on quantum polynomial algebras

$$k_{\boldsymbol{q}}[x_1, \ldots, x_n] := k\langle x_1, \ldots, x_n\rangle/(x_i x_j - q_{ij} x_j x_i),$$

where $\boldsymbol{q} = (q_{ij})$, $q_{ij} \in k^\times$ with $q_{ii} = 1$ and $q_{ij} q_{ji} = 1$. We view $\boldsymbol{q}$ as a point of the algebraic torus $(k^\times)^{n(n-1)/2}$ with coordinates q_{ij}, $i < j$. Let $\langle \boldsymbol{q} \rangle$ be the subgroup in $(k^\times)^{n(n-1)/2}$ generated by $\boldsymbol{q}$, and let $G_{\boldsymbol{q}}$ be its Zariski closure. Let $G_{\boldsymbol{q}}^0$ be the connected component of the identity in $G_{\boldsymbol{q}}$.

Theorem 4.1. *Let H be a semisimple Hopf algebra of dimension d. If the order of $G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0$ is coprime to $d!$, then any H-action on $B := k_{\boldsymbol{q}}[x_1, \ldots, x_n]$ factors through a group action.*

Proof. We may assume that H acts on $B := k_{\boldsymbol{q}}[x_1, \ldots, x_n]$ inner faithfully. Let $R \subset k$ be a finitely generated subring containing q_{ij}, let $B_R := R_{\boldsymbol{q}}[x_1, \ldots, x_n]$ be

the quantum polynomial algebra defined over R, and let H_R be a Hopf R-order with an action on B_R which becomes the action of H on B upon tensoring with k.

Similarly to the proof of Theorem 2.4, we need to control the PI degree of B_R after reducing modulo p; we employ a version of the number-theoretic result of A. Perucca (as presented in the Appendix) to do so.

Given a number field K and a ring homomorphism $\xi : R \to K$, let $R' := \xi(R)$, $H_{R'} := H_R \otimes_R R'$, $B_{R'} := B_R \otimes_R R' = R'_{\xi(\boldsymbol{q})}[x_1, \dots, x_n]$. Then $H_{R'}$ acts on $B_{R'}$ inner faithfully. For a generic choice of ξ, any multiplicative relation satisfied by $\xi(q_{ij})$ is already satisfied by q_{ij}, so by Example A.3, we have $|G_{\boldsymbol{q}}/G^0_{\boldsymbol{q}}| = |G_{\xi(\boldsymbol{q})}/G^0_{\xi(\boldsymbol{q})}|$. By Corollary A.2, there exist infinitely many primes p with prime ideals $\mathfrak{p} \subset R'$ lying over them such that, for a generic homomorphism $\psi : R' \to \bar{\mathbb{F}}_p$ annihilating $\mathfrak{p}$, the order $N := N_{\mathfrak{p}}$ of $\psi \circ \xi(\boldsymbol{q})$ is finite and relatively prime to $d!$. Let $H_p := H_{R'} \otimes_{R'} \bar{\mathbb{F}}_p$ and $B_p := B_{R'} \otimes_{R'} \bar{\mathbb{F}}_p$ be the corresponding reductions of H and B modulo p. For large enough p, the Hopf algebra H_p is semisimple and cosemisimple by [Cuadra et al. 2015, Lemma 2.5], and B_p is a PI domain with PI degree dividing N^n (as x_i^N are central elements in B_p). Moreover, H_p acts on B_p inner faithfully by a version of [Cuadra et al. 2015, Proposition 2.4] adapted to the algebra B (with the same proof).

Let D_p be the quotient division algebra of B_p. Then the PI degree of D_p divides N^n, and is therefore coprime to $d!$. Further, H_p acts inner faithfully on D_p. Hence, [Cuadra et al. 2015, Proposition 3.3(ii)] implies that H_p is cocommutative. Since this happens for infinitely many primes, we conclude that $H_{R'}$ is cocommutative. Since this happens for generic maps ξ, this implies that H_R is cocommutative. Thus H is cocommutative, i.e., H is a group algebra. □

Corollary 4.2. *The conclusion of Theorem 4.1 holds when $q_{ij} = q^{m_{ij}}$, where $m_{ij} = -m_{ji}$ are integers, and the order of $q \in k^\times$ is infinite or is coprime to $d!$.*

Proof. This is a special case of Theorem 4.1. □

Example 4.3. The assumption in Theorem 4.1 and Corollary 4.2 that the order of $G_{\boldsymbol{q}}/G^0_{\boldsymbol{q}}$ is coprime to $d!$ cannot be removed. For instance, there exists an inner faithful action of the 8-dimensional noncommutative noncocommutative semisimple Hopf algebra on the quantum polynomial algebra $k_{-1}[x, y]$; see [Kirkman et al. 2009, Example 7.6]. In this case, $|G_{\boldsymbol{q}}/G^0_{\boldsymbol{q}}| = 2$.

5. Finite dimensional Hopf actions on quantum polynomial algebras

Let us now extend the results of the previous section to not necessarily semisimple Hopf algebras, under some additional assumptions.

First of all, when the action of H on B is Hopf–Galois, we can remove in Theorem 4.1 the assumption that H is semisimple, and also weaken the coprimeness assumption, replacing $d!$ with d.

Proposition 5.1. *Suppose that a finite dimensional Hopf algebra H acts on $B := k_{\boldsymbol{q}}[x_1, \ldots, x_n]$, and the order of $G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0$ is coprime to d. If this action gives rise to an H^*-Hopf–Galois extension $B^H \subset B$, then the action of H on B factors through a group algebra.*

Proof. The proof is parallel to the proof of Theorem 3.1. The weaker coprimeness assumption suffices since by the Hopf–Galois condition, $[D_p : D_p^H] = d$ (not just $\leq d$). Here, $p \gg 0$ and D_p is the full localization of B reduced modulo p via the method in the proof of Theorem 4.1. □

Let us now give a generalization of Theorem 4.1 to the nonsemisimple case under a nondegeneracy assumption.

Theorem 5.2. *Let H be a finite dimensional Hopf algebra of dimension d acting on $B := k_{\boldsymbol{q}}[x_1, \ldots, x_n]$. Assume that the order of $G_{\boldsymbol{q}}/G_{\boldsymbol{q}}^0$ is coprime to $d!$, and $\boldsymbol{q}$ is nondegenerate. Then the action of H on B factors through a group action.*

Proof. The proof is obtained by combining the proofs of Theorems 3.5 and 4.1. Let us describe the necessary changes.

We argue as in the proof of Theorem 4.1. Fix a generic character $\xi : R \to K$ from R to a number field K, and set $R' = \xi(R)$. By Corollary A.2, there exist infinitely many primes p with prime ideals $\mathfrak{p} \subset R'$ lying over them such that, for a generic homomorphism $\psi : R' \to \overline{\mathbb{F}}_p$ annihilating $\mathfrak{p}$, the order $N := N_{\mathfrak{p}}$ of $\psi \circ \xi(\boldsymbol{q})$ is finite and coprime to $d!$.

Consider the image $Z(m)$ of the center Z_m of D_{p^m} in D_p (thus, $Z(1) = Z$). By a straightforward generalization of [Cuadra et al. 2016, Lemma 4.6], $Z(m)$ is preserved by the action of H_p. It is clear that $Z(m)$ is generated by the monomials $x_1^{m_1} \cdots x_n^{m_n}$ such that $\prod_j q_{ij}^{m_j} = 1$ in the truncated ring of Witt vectors $W_{m,p}$ (see [Cuadra et al. 2016, Section 2.1]). Let $W'_{m,p}$ be the kernel of the natural map of multiplicative groups $W_{m,p}^\times \to \overline{\mathbb{F}}_p^\times$. Then every element of $W'_{m,p}$ has order a power of p. Hence, $[Z : Z(m)]$ is a power of p. Also it is clear from the nondegeneracy condition for $\boldsymbol{q}$ that $\bigcap_m Z(m) = \overline{\mathbb{F}}_p$. Thus, Theorem 3.2 applies, and yields that H_p is cocommutative. Hence $H_{R'}$ is cocommutative, implying that H_R is cocommutative and ultimately that H is cocommutative, i.e., a group algebra. □

Remark 5.3. If $q_{ij} = q^{m_{ij}}$, where q is not a root of unity, then $\boldsymbol{q}$ is nondegenerate if and only if $\det(m_{ij}) \neq 0$. Theorem 5.2 applies in this case. This gives a generalization of [Chan et al. 2014, Theorem 0.4] to nonhomogeneous Hopf actions for even n.

Proposition 5.4. *Theorem 4.1, Corollary 4.2, and Theorem 5.2 remain valid if the quantum polynomial algebra B is replaced by the quantum torus $k_{\boldsymbol{q}}[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ or the division algebra of quotients* $\mathrm{Frac}(B)$.

Proof. In the case of the quantum torus, the proof is analogous to the proof of Theorem 4.1. The case of the division algebra of quotients is obtained using the same argument as in the proof of [Cuadra et al. 2015, Proposition 4.4]. □

Example 5.5. The condition that H is semisimple cannot be dropped in Theorem 4.1, and the condition that $\boldsymbol{q}$ is nondegenerate cannot be dropped in Theorem 5.2.

Namely, let $A = A_0 \oplus A_1$ be a $\mathbb{Z}/2\mathbb{Z}$-graded domain with a nonzero central element $z \in A_1$, and take H to be the 4-dimensional Sweedler Hopf algebra generated by a group-like element g and a $(g, 1)$-skew-primitive element u with $g^2 = 1$, $u^2 = 0$ and $gu + ug = 0$.

(1) Then there is an action of H on A (not preserving the grading of A) given by $g \cdot a = (-1)^{\deg a} a$, and $u \cdot a = 0$ if $a \in A_0$ and $u \cdot a = za$ if $a \in A_1$. It is easy to check that this action is well-defined, and it is inner faithful since u acts by a nonzero operator.

(2) In particular, we have an inner faithful action of H on the quantum polynomial algebra $k_q[x, y]$, for q a root of unity of any odd order $2m - 1$, $m > 0$; namely, we can take $z = x^{2m-1}$.

(3) This gives an inner faithful action of H on the quantum torus $k_q[x_1^{\pm 1}, \ldots, x_n^{\pm 1}]$ if n is odd: we can take the central element

$$z = x_1 x_2^{-1} x_3 \cdots x_{n-1}^{-1} x_n.$$

For even n, such an action is impossible if q is not a root of unity by Theorem 5.2. Indeed, the matrix $m_{ij} := \operatorname{sign}(j - i)$ is nondegenerate if and only if n is even (see Remark 5.3).

(4) Finally, this gives an inner faithful Sweedler Hopf algebra action on the Weyl algebra $\boldsymbol{A}_n(F)$ when $\operatorname{char}(F) = p \geq 3$; the $\mathbb{Z}/2\mathbb{Z}$ grading is defined by giving the generators degree 1, and we can take, for instance, $z = x_1^p$. (Note that by [Cuadra et al. 2016, Theorem 1.1], this is impossible in characteristic zero; indeed, the center of $\boldsymbol{A}_n(k)$ is k.)

6. Semisimple Hopf actions on twisted homogeneous coordinate rings and 3-dimensional Sklyanin algebras

Now let us consider semisimple Hopf actions on twisted homogeneous coordinate rings of abelian varieties. We keep the notation of Section 1D.

Let H be a Hopf algebra over k of dimension d.

Theorem 6.1. *We have the following statements.*

(i) *If H is semisimple, and if the order of G_σ / G_σ^0 is coprime to $d!$, then any H-action on $B := B(X, \sigma, \mathcal{L})$ factors through a group action.*

(ii) *Moreover, part (i) holds for the H-action on the division algebra of quotients* $\mathrm{Frac}(B)$ *of* B.

(iii) *Part (i) also holds for not necessarily semisimple H if the order of G_σ/G_σ^0 is coprime to d and the H-action gives rise to a Hopf–Galois extension.*

Proof. The proofs of the statements (i)–(iii) are parallel to the proofs of Theorem 4.1, Proposition 5.4, and Proposition 5.1, respectively, where we use that the PI degree of B equals the order of σ. The only difference is that Corollary A.2 is applied to the abelian variety X with subgroup $\{s^i\}_{i\in\mathbb{Z}}$ rather than the torus $(k^\times)^{n(n-1)/2}$ with subgroup $\langle \boldsymbol{q}\rangle$. □

In particular, if $X =: E$ is an elliptic curve, Theorem 6.1 holds if the order of σ is coprime to $d!$ or infinite. Moreover, if σ has infinite order, the assumption that H is semisimple can be dropped.

Theorem 6.2. *Let E be an elliptic curve, and take $\sigma \in Aut(E)$ given by translation by a point of infinite order. Then any finite dimensional Hopf action on $B(E,\sigma,\mathcal{L})$ factors through a group action.*

Proof. The proof repeats the proofs of Theorems 5.2 and 3.5 without significant changes. □

Finally, let us consider semisimple Hopf actions on 3-dimensional Sklyanin algebras [Artin et al. 1990; Odesskiĭ and Feĭgin 1989]. Let F be an algebraically closed field of characteristic not equal to 2 or 3.

Definition 6.3. Let $a, b, c \in F^\times$ be such that

$$(3abc)^3 \neq (a^3+b^3+c^3)^3.$$

The *3-dimensional Sklyanin algebra*, denoted by $S(a,b,c)$ is generated over F by x, y, z with defining relations

$$ayz+bzy+cx^2 = azx+bxz+cy^2 = axy+byx+cz^2 = 0.$$

It is known that $S(a,b,c)$ is Koszul with Hilbert series $(1-t)^{-3}$ (see [Artin et al. 1990, Theorems 6.6(ii) and 6.8(i)] and a result of J. Zhang [Smith 1996, Theorem 5.11]), so that $S(a,b,c)$ is a flat deformation of the algebra of polynomials in three variables (see, e.g., [Tate and Van den Bergh 1996, Theorem 1.1]). Moreover, the center of $S(a,b,c)$ contains an element T of degree 3, and $S(a,b,c)/(T) = B(E,\sigma,\mathcal{L})$, where E is the elliptic curve in $\mathbb{P}^2$ given by the equation

$$(a^3+b^3+c^3)xyz = abc(x^3+y^3+z^3),$$

σ is given by translation by the point $(a:b:c) \in E$, and $\mathcal{L}$ is a line bundle of degree 3 on E.

Theorem 6.4. *Let $S(a, b, c)$ be a 3-dimensional Sklyanin algebra over k and let H be a semisimple Hopf algebra over k of dimension d. If the order of $\sigma \in Aut(E)$ is coprime to $d!$ or infinite, then any H-action on $S(a, b, c)$ factors through a group action.*

Proof. It is known from the theory of Sklyanin algebras that if σ has order N, then $S(a, b, c)$ is PI with PI degree N (see [Artin et al. 1994, part 5 of theorem on page 7]). Therefore, Theorem 6.4 is proved similarly to Theorem 4.1, using Corollary A.2 for elliptic curves, as in Theorem 5.2. □

Remark 6.5. The semisimplicity condition on H in Theorem 6.1 and the infinite order condition in Theorem 6.2 cannot be dropped, as there exists a Sweedler Hopf algebra action on $B := B(X, \sigma, \mathcal{L})$ if σ has odd order N. Namely, we take a sufficiently large odd number m such that the line bundle $\mathcal{L}^{\otimes m}$ is very ample (it exists since $\mathcal{L}$ is ample). Now $B[mN] \neq 0$ and there exists an eigenvector f of σ in $B[mN]$. We then take $z = f^N$, a nonzero central element of odd degree mN^2, so that a desired action is given by Example 5.5.

Also, the semisimplicity assumption in Theorem 6.4 cannot be dropped, as there exists a Sweedler Hopf algebra action on $S(a, b, c)$ for any a, b, c, given by Example 5.5 where we use the central element T in place of the element z.

Appendix

The goal of this Appendix is to provide number-theoretic results needed in Section 4. We start by quoting a result from [Perucca 2009] (in which we take F to be the number field K itself).

Theorem A.1 [Perucca 2009, Theorem 7]. *Let G be the product of an abelian variety and a torus defined over a number field K. Let $g \in G(K)$ be a K-rational point on G such that the Zariski closure G_g of the subgroup $\langle g \rangle \subset G(K)$ generated by g is connected. Fix a positive integer r. Then there exists a positive Dirichlet density of primes $\mathfrak{p}$ of K such that the order of g modulo $\mathfrak{p}$ is coprime to r.* □

Corollary A.2. *Let K, G be as in Theorem A.1, let $g \in G(K)$, and let $\ell := |G_g/G_g^0|$, where G_g^0 is the connected component of the identity in G_g (i.e., $G_g/G_g^0 = \mathbb{Z}/\ell\mathbb{Z}$). Fix a positive integer r coprime to ℓ. Then there exists a positive Dirichlet density of primes $\mathfrak{p}$ of K such that the order of g modulo $\mathfrak{p}$ is coprime to r.*

The corollary above is used in the proof of Theorem 4.1, where $d!$ is r and $N_\mathfrak{p}$ is the order of g modulo $\mathfrak{p}$.

Proof. The order of g in G_g/G_g^0 is ℓ, so $G_{g^\ell} = G_g^0$. Now the statement follows by applying Theorem A.1 to g^ℓ. □

Example A.3. Let G be a split m-dimensional torus, and consider an element $g := (q_1, \ldots, q_m) \in G$. We have the following statements.

(1) The group G_g is connected if and only if the group Γ generated by $q_1, \dots, q_m$ in $K^\times$ is free, i.e., does not contain nontrivial roots of unity. Indeed both conditions are equivalent to the condition that any character χ of G which maps g to an ℓ-th root of unity satisfies $\chi(g) = 1$.

(2) More generally, $|G_g/G_g^0| = \ell$ if and only if the group of roots of unity generated by $\chi(g)$, where χ runs through characters of G such that $\chi(g)$ is a root of unity, is μ_ℓ. In other words, ℓ is the order of the torsion subgroup in $\mathbb{Z}^m/g^\perp$, where $g^\perp$ is the subgroup of characters χ such that $\chi(g) = 1$. In particular, ℓ depends only on the multiplicative relations satisfied by q_{ij}.

(3) If $\dim G = 1$ (i.e., $G = \mathbb{G}_m$ or an elliptic curve), then G_g is connected if and only if g has infinite order or $g = 1$. More generally, $|G_g/G_g^0| = \ell > 1$ if and only if g has order ℓ.

Acknowledgments

We thank Bjorn Poonen for many useful discussions and for the number-theoretic reference [Perucca 2009], which is crucial for our arguments. We are also grateful to R. Bezrukavnikov, I. Losev, and H. Nakajima for useful discussions and explanations. We thank the referee for many useful comments that improved greatly the quality of this manuscript. The authors were supported by the National Science Foundation: NSF-grants DMS-1502244 and DMS-1550306.

References

[Artin and Van den Bergh 1990] M. Artin and M. Van den Bergh, "Twisted homogeneous coordinate rings", *J. Algebra* **133**:2 (1990), 249–271. MR 1067406 Zbl 0717.14001

[Artin et al. 1990] M. Artin, J. Tate, and M. Van den Bergh, "Some algebras associated to automorphisms of elliptic curves", pp. 33–85 in *The Grothendieck Festschrift, Vol. I*, edited by P. Cartier et al., Progress in Mathematics **86**, Birkhäuser, Boston, 1990. MR 1086882 Zbl 0744.14024

[Artin et al. 1994] M. Artin, W. Schelter, and J. Tate, "The centers of 3-dimensional Sklyanin algebras", pp. 1–10 in *Barsotti Symposium in Algebraic Geometry* (Abano Terme, 1991), edited by V. Cristante and W. Messing, Perspectives in Mathematics **15**, Academic Press, San Diego, 1994. MR 1307390 Zbl 0823.17019

[Beĭlinson and Bernstein 1993] A. Beĭlinson and J. Bernstein, "A proof of Jantzen conjectures", pp. 1–50 in *I. M. Gel'fand Seminar*, edited by S. Gel'fand and S. Gindikin, Advances in Soviet Mathematics **16**, American Mathematical Society, Providence, RI, 1993. MR 1237825 Zbl 0790.22007

[Braden et al. 2012] T. Braden, N. Proudfoot, and B. Webster, "Quantizations of conical symplectic resolutions, I: local and global structure", preprint, 2012. arXiv 1208.3863

[Chan et al. 2014] K. Chan, C. Walton, and J. Zhang, "Hopf actions and Nakayama automorphisms", *J. Algebra* **409** (2014), 26–53. MR 3198834 Zbl 1315.16027

[Chan et al. 2016] K. Chan, E. Kirkman, C. Walton, and J. J. Zhang, "Quantum binary polyhedral groups and their actions on quantum planes", *J. Reine Angew. Math.* **719** (2016), 211–252. MR 3552496 Zbl 06636678

[Crawley-Boevey 2001] W. Crawley-Boevey, "Geometry of the moment map for representations of quivers", *Compositio Math.* **126**:3 (2001), 257–293. MR 1834739 Zbl 1037.16007

[Cuadra et al. 2015] J. Cuadra, P. Etingof, and C. Walton, "Semisimple Hopf actions on Weyl algebras", *Adv. Math.* **282** (2015), 47–55. MR 3374522 Zbl 06473182

[Cuadra et al. 2016] J. Cuadra, P. Etingof, and C. Walton, "Finite dimensional Hopf actions on Weyl algebras", *Adv. Math.* **302** (2016), 25–39. MR 3545923 Zbl 06631651

[Eisenbud 1995] D. Eisenbud, *Commutative algebra, with a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, 1995. MR 1322960 Zbl 0819.13001

[Etingof 2006] P. Etingof, "The p-center of symplectic reflection algebras", (2006). Appendix to R. Bezrukavnikov, M. Finkelberg, and V. Ginzburg, "Cherednik algebras and Hilbert schemes in characteristic p", *Represent. Theory* **10** (2006), 254–298. MR 2219114 Zbl 1130.14005

[Etingof 2007] P. Etingof, *Calogero–Moser systems and representation theory*, European Mathematical Society, Zürich, 2007. MR 2296754 Zbl 1331.53002

[Etingof 2016] P. Etingof, "A PI degree theorem for quantum deformations", *J. Algebra* **466** (2016), 308–313. MR 3541688 Zbl 06623498

[Etingof and Ginzburg 2010] P. Etingof and V. Ginzburg, "Noncommutative del Pezzo surfaces and Calabi–Yau algebras", *J. Eur. Math. Soc.* (*JEMS*) **12**:6 (2010), 1371–1416. MR 2734346 Zbl 1204.14004

[Etingof and Walton 2014] P. Etingof and C. Walton, "Semisimple Hopf actions on commutative domains", *Adv. Math.* **251** (2014), 47–61. MR 3130334 Zbl 1297.16029

[Etingof and Walton 2015] P. Etingof and C. Walton, "Pointed Hopf actions on fields, I", *Transform. Groups* **20**:4 (2015), 985–1013. MR 3416436 Zbl 1338.16035

[Etingof and Walton 2016] P. Etingof and C. Walton, "Finite dimensional Hopf actions on deformation quantizations", *Proc. Amer. Math. Soc.* (online publication October 2016).

[Grothendieck 1966] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, III", *Inst. Hautes Études Sci. Publ. Math.* **28** (1966), 5–255. MR 0217086 Zbl 0144.19904

[Jacobson 1952] N. Jacobson, "A note on Lie algebras of characteristic p", *Amer. J. Math.* **74** (1952), 357–359. MR 0047026 Zbl 0046.03402

[Jacobson 1962] N. Jacobson, *Lie algebras*, Interscience Tracts in Pure and Applied Mathematics **10**, Interscience Publishers, New York, 1962. MR 0143793 Zbl 0121.27504

[Katz 1970] N. M. Katz, "Nilpotent connections and the monodromy theorem: applications of a result of Turrittin", *Inst. Hautes Études Sci. Publ. Math.* **39** (1970), 175–232. MR 0291177 Zbl 0221.14007

[Kirkman et al. 2009] E. Kirkman, J. Kuzmanovich, and J. J. Zhang, "Gorenstein subrings of invariants under Hopf algebra actions", *J. Algebra* **322**:10 (2009), 3640–3669. MR 2568355 Zbl 1225.16015

[Lomp and Pansera 2015] C. Lomp and D. Pansera, "A note on a paper by Cuadra, Etingof and Walton", preprint, 2015. To appear in *Comm. Alg.* arXiv 1506.07766

[Losev 2010] I. Losev, "Finite W-algebras", pp. 1281–1307 in *Proceedings of the International Congress of Mathematicians* (Hyderabad, 2010), vol. 3, edited by R. Bhatia et al., Hindustan Book Agency, New Delhi, 2010. MR 2827841 Zbl 1246.17015

[Losev and Etingof 2015] I. Losev and P. Etingof, "Bernstein inequality and holonomic modules", preprint, 2015. arXiv 1501.01260

[McConnell and Robson 2001] J. C. McConnell and J. C. Robson, *Noncommutative Noetherian rings*, revised ed., Graduate Studies in Mathematics **30**, American Mathematical Society, Providence, RI, 2001. MR 1811901 Zbl 0980.16019

[Montgomery 1993] S. Montgomery, *Hopf algebras and their actions on rings*, CBMS Regional Conference Series in Mathematics **82**, American Mathematical Society, Providence, RI, 1993. MR 1243637 Zbl 0793.16029

[Nakajima 2015] H. Nakajima, "Towards a mathematical definition of Coulomb branches of 3-dimensional $\mathcal{N} = 4$ gauge theories, I", preprint, 2015. arXiv 1503.03676

[Odesskiĭ 2002] A. V. Odesskiĭ, "Elliptic algebras", *Uspekhi Mat. Nauk* **57**:6 (2002), 87–122. In Russian; translated in *Russian Math. Surveys* **57**:6 (2002), 1127–1162. MR 1991863

[Odesskiĭ and Feĭgin 1989] A. V. Odesskiĭ and B. L. Feĭgin, "Sklyanin's elliptic algebras", *Funktsional. Anal. i Prilozhen.* **23**:3 (1989), 45–54. In Russian; translated in *Funct. Anal. Appl.* **23**:3 (1989), 207–214. MR 1026987

[Ore 1931] O. Ore, "Linear equations in non-commutative fields", *Ann. of Math.* (2) **32**:3 (1931), 463–477. MR 1503010 Zbl 0001.26601

[Perucca 2009] A. Perucca, "Prescribing valuations of the order of a point in the reductions of abelian varieties and tori", *J. Number Theory* **129**:2 (2009), 469–476. MR 2473894 Zbl 1166.14028

[Posner 1960] E. C. Posner, "Prime rings satisfying a polynomial identity", *Proc. Amer. Math. Soc.* **11** (1960), 180–183. MR 0111765 Zbl 0215.38101

[Sklyanin 1982] E. K. Sklyanin, "Some algebraic structures connected with the Yang–Baxter equation", *Funktsional. Anal. i Prilozhen.* **16**:4 (1982), 27–34, 96. In Russian; translated in *Funct. Anal. Appl.* **16**:4 (1983), 263–270. MR 684124 Zbl 0513.58028

[Skryabin and Van Oystaeyen 2006] S. Skryabin and F. Van Oystaeyen, "The Goldie Theorem for H-semiprime algebras", *J. Algebra* **305**:1 (2006), 292–320. MR 2264132 Zbl 1109.16033

[Smith 1996] S. P. Smith, "Some finite-dimensional algebras related to elliptic curves", pp. 315–348 in *Representation theory of algebras and related topics* (Mexico City, 1994), edited by R. Bautista et al., CMS Conf. Proc. **19**, American Mathematical Society, Providence, RI, 1996. MR 1388568 Zbl 0856.16009

[Smith and Stafford 1992] S. P. Smith and J. T. Stafford, "Regularity of the four-dimensional Sklyanin algebra", *Compositio Math.* **83**:3 (1992), 259–289. MR 1175941 Zbl 0758.16001

[Stephenson 1997] D. R. Stephenson, "Algebras associated to elliptic curves", *Trans. Amer. Math. Soc.* **349**:6 (1997), 2317–2340. MR 1390046 Zbl 0868.16028

[Tate and Van den Bergh 1996] J. Tate and M. Van den Bergh, "Homological properties of Sklyanin algebras", *Invent. Math.* **124**:1 (1996), 619–647. MR 1369430 Zbl 0876.17010

Communicated by Susan Montgomery
Received 2016-05-19 Revised 2016-08-01 Accepted 2016-10-22

etingof@math.mit.edu *Department of Mathematics, Massachusetts Institute of Technology, 77 Massachusetts Avenue, Cambridge, MA 02139-4307, United States*

notlaw@temple.edu *Department of Mathematics, Temple University, 1805 N. Broad Street, Philadelphia, PA 19122-6094, United States*

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 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.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

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 LaTeX but submissions in other varieties of TeX, 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 BibTeX 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@msp.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 10 No. 10 2016