Algebra & Number Theory

msp.org/ant

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

Explicit Chabauty over number fields

Samir Siksek

Let C be a smooth projective absolutely irreducible curve of genus $g \geq 2$ over a number field K of degree d, and let J denote its Jacobian. Let r denote the Mordell–Weil rank of $J(K)$. We give an explicit and practical Chabauty-style criterion for showing that a given subset $\mathcal{K} \subseteq C(K)$ is in fact equal to $C(K)$. This criterion is likely to be successful if $r \leq d(g-1)$. We also show that the only solution to the equation $x^2 + y^3 = z^{10}$ in coprime nonzero integers is $(x, y, z) = (\pm 3, -2, \pm 1)$. This is achieved by reducing the problem to the determination of K-rational points on several genus-2 curves where $K = \mathbb{Q}$ or $\mathbb{Q}(\sqrt[3]{2})$ and applying the method of this paper.

1. Introduction

Let C be a smooth projective absolutely irreducible curve of genus $g \geq 2$ defined over a number field K, and write J for the Jacobian of C. Suppose that the rank of the Mordell–Weil group $J(K)$ is at most $g - 1$. In a pioneering paper, Chabauty [1941] proved the finiteness of the set of K-rational points on C. This has since been superseded by the proof of Faltings [1983] of the Mordell conjecture, which gives the finiteness of $C(K)$ without any assumption on the rank of $J(K)$. Chabauty's approach, where applicable, does however have two considerable advantages:

(a) Chabauty's method can be refined to give explicit bounds for the cardinality of $C(K)$ as shown by Coleman [1985a]. Coleman's bounds are realistic and occasionally even sharp; see for example [Grant 1994; Flynn 1995b]. Coleman's approach has been adapted to give bounds (assuming some reasonable conditions) for the number of solutions of Thue equations [Lorenzini and Tucker 2002], the number of rational points on Fermat curves [McCallum 1992; 1994], the number of points on curves of the form $y^2 = x^5 + A$ [Stoll 2006b], and the number of rational points on twists of a given curve [Stoll 2006a].

(b) The Chabauty–Coleman strategy can often be adapted to compute $C(K)$ as in [Bruin 2002; 2003; Flynn 1997; Flynn and Wetherell 1999; 2001; McCallum

---

and Poonen 2010; Wetherell 1997] and even the K-rational points on the
symmetric powers of C [Siksek 2009].

This paper is inspired by a talk[1] given by Joseph Wetherell at the MSRI on December 11, 2000. In that talk, Wetherell suggested that it should be possible to adapt the Chabauty strategy to compute the set of K-rational points on C provided the rank r of the Mordell–Weil group $J(K)$ satisfies $r \leq d(g-1)$, where $d = [K : \mathbb{Q}]$. Wetherell has never published details of his method, which we believe is similar to ours.

In this paper, we give a practical Chabauty-style method for determining $C(K)$ that should succeed if the inequality $r \leq d(g-1)$ holds (but see the discussion at the end of Section 2). We suppose that we have been supplied with a basis $D_1, \ldots, D_r$ for a subgroup of $J(K)$ of full rank and hence finite index; the elements of this basis are represented as degree-0 divisors on C (modulo linear equivalence). Obtaining a basis for a subgroup of full rank is often the happy outcome of a successful descent calculation [Cassels and Flynn 1996; Flynn 1994; Poonen and Schaefer 1997; Schaefer 1995; Schaefer and Wetherell 2005; Stoll 1998; 2001; 2002a]. Obtaining a basis for the full Mordell–Weil group is often time-consuming for genus-2 curves [Flynn 1995a; Flynn and Smart 1997; Stoll 1999; 2002b] and simply not feasible in the present state of knowledge for curves of genus at least 3. We also assume the knowledge of at least one rational point $P_0 \in C(K)$. If a search for rational points on C does not reveal any points, then experience suggests that $C(K) = \varnothing$ and that some combination of descent and Mordell–Weil sieving [Bruin and Stoll 2008; 2009; 2010] is likely to prove this.

This paper is organized as follows. Section 2 gives a heuristic explanation of why Chabauty's approach should be applicable when the rank r of the Mordell–Weil group satisfies $r \leq g(d - 1)$. Section 3 gives a quick summary of basic facts regarding υ-adic integration on curves and Jacobians. In Section 4, for $Q \in C(K)$ and a rational prime p, we define a certain neighborhood of Q in $\prod_{\upsilon|p} C(K_\upsilon)$ that we call the p-unit ball around Q and give a Chabauty-style criterion for Q to be the unique K-rational point belonging to this neighborhood. In Section 5, we explain how to combine our Chabauty criterion with the Mordell–Weil sieve and deduce a practical criterion for a given set $\mathcal{H} \subseteq C(K)$ to be equal to $C(K)$. In Section 6, we use our method to prove the following theorem:

Theorem 1. *The only solutions to the equation*

$$x^2 + y^3 = z^{10} \tag{1}$$

in coprime integers x, y, and z are

$$(\pm 3, -2, \pm 1), \quad (\pm 1, 0, \pm 1), \quad (\pm 1, -1, 0), \quad and \quad (0, 1, \pm 1).$$

---

[1]http://msri.org/publications/ln/msri/2000/arithgeo/wetherell/1/banner/01.html

We note that Dahmen [2008, Chapter 3.3.2] has solved the equation $x^2 + z^{10} = y^3$ using Galois representations and level-lowering. We have been unable to solve (1) by using Galois representations; the difficulty arises from the additional "nontrivial" solution $(x, y, z) = (\pm 3, -2, \pm 1)$, which is not present for the equation $x^2 + z^{10} = y^3$. We solve (1) by reducing the problem to determining the K-rational points on several genus-2 curves where K is either $\mathbb{Q}$ or $\mathbb{Q}(\sqrt[3]{2})$. For all these genus-2 curves, the inequality $r \leq d(g-1)$ is satisfied and we are able to determine the K-rational points using the method of this paper.

Recently, David Brown [2012] has given an independent and entirely different proof of Theorem 1. Brown's method is rather intricate and makes use of elliptic-curve Chabauty, mod-5 level-lowering, and number-field enumeration.

2. A heuristic explanation of Wetherell's idea

In this section, we explain the heuristic idea behind Chabauty's method and then how the heuristic can be modified for curves over number fields. Let C be a smooth projective curve of genus $g \geq 2$ defined over K. Let J be the Jacobian of C and r the rank of the Mordell–Weil group $J(K)$. Fix a rational point $P_0 \in C(K)$, and let $\jmath : C \hookrightarrow J$ be the Abel–Jacobi map with base point P_0. We use $\jmath$ to identify C as a subvariety of J.

To explain the usual Chabauty method, it is convenient to assume that $K = \mathbb{Q}$. Choose a finite prime p. Inside $J(\mathbb{Q}_p)$, it is clear that

$$C(\mathbb{Q}) \subseteq C(\mathbb{Q}_p) \cap J(\mathbb{Q}) \subseteq C(\mathbb{Q}_p) \cap \overline{J(\mathbb{Q})},$$

where $\overline{J(\mathbb{Q})}$ is the closure of $J(\mathbb{Q})$ in the p-adic topology. Now $J(\mathbb{Q}_p)$ is a $\mathbb{Q}_p$-Lie group of dimension g, and $\overline{J(\mathbb{Q})}$ is a $\mathbb{Q}_p$-Lie subgroup of dimension at most r. Moreover, $C(\mathbb{Q}_p)$ is a one-dimensional submanifold of $J(\mathbb{Q}_p)$. If $r + 1 \leq g$, then we expect that the intersection $C(\mathbb{Q}_p) \cap \overline{J(\mathbb{Q})}$ is finite. It turns out that this intersection is indeed finite if $r \leq g - 1$, and Coleman [1985a] gives a bound for the cardinality of this intersection under some further (but mild) hypotheses. Moreover, in practice, this intersection can be computed to any desired p-adic accuracy.

Now we return to the general setting by letting K be a number field of degree d. Define the Weil restrictions

$$V = \operatorname{Res}_{K/\mathbb{Q}} C \quad \text{and} \quad A = \operatorname{Res}_{K/\mathbb{Q}} J. \qquad (2)$$

Then V is a variety of dimension d and A an abelian variety of dimension gd, both defined over $\mathbb{Q}$. The Weil restriction of the morphism $\jmath : C \hookrightarrow J$ is a morphism $V \hookrightarrow A$ defined over $\mathbb{Q}$ that we use to identify V as a subvariety of A. This Weil restriction defines a bijection between $C(K)$ and $V(\mathbb{Q})$, and

$$\operatorname{rank} A(\mathbb{Q}) = \operatorname{rank} J(K) = r.$$

Mimicking the previous argument,

$$V(\mathbb{Q}) \subseteq V(\mathbb{Q}_p) \cap \overline{A(\mathbb{Q})}.$$

Now $\overline{A(\mathbb{Q})}$ is at most r-dimensional, $V(\mathbb{Q}_p)$ is d-dimensional, and the intersection is taking place in the $\mathbb{Q}_p$-Lie group $A(\mathbb{Q}_p)$ of dimension gd. If $r + d \leq gd$, we expect that the intersection is finite.

Remark. As Wetherell points out, even if $r + d \leq gd$, it is possible for the intersection $V(\mathbb{Q}_p) \cap \overline{A(\mathbb{Q})}$ to be infinite. For example, let C be a curve defined over $\mathbb{Q}$ with Mordell–Weil rank at least g. One normally expects that $\overline{J(\mathbb{Q})}$ is g-dimensional. Assume that this is the case. Then the intersection

$$C(\mathbb{Q}_p) \cap \overline{J(\mathbb{Q})}$$

will contain a neighborhood in $C(\mathbb{Q}_p)$ of the base point P_0 and so will be infinite. Now let V and A be obtained from C and J by first base-extending to number field K and then taking Weil restriction back to $\mathbb{Q}$. One has a natural injection

$$C(\mathbb{Q}_p) \cap \overline{J(\mathbb{Q})} \hookrightarrow V(\mathbb{Q}_p) \cap \overline{A(\mathbb{Q})}$$

proving that the latter intersection is infinite. This is true regardless of whether the inequality $r \leq d(g-1)$ is satisfied. However, for a random curve C defined over a number field K, on the basis for the above heuristic argument, we expect the intersection $V(\mathbb{Q}_p) \cap \overline{A(\mathbb{Q})}$ to be finite when the inequality $r \leq d(g-1)$ is satisfied. We are led to the following open question:

Open question. *Let C be a smooth projective curve of genus $g \geq 2$ over a number field K of degree d. Suppose, for every smooth projective curve D defined over a subfield $L \subseteq K$ and satisfying $D \times_L K \cong_K C$, that the inequality*

$$\mathrm{rank}\, J_D(L) \leq [L : \mathbb{Q}](g - 1)$$

holds, where J_D denotes the Jacobian of D. Let V and A be given by (2). *Is $V(\mathbb{Q}_p) \cap \overline{A(\mathbb{Q})}$ necessarily finite?*

3. Preliminaries

In this section, we summarize various results on p-adic integration that we need. The definitions and proofs can be found in [Coleman 1985b; Colmez 1998]. For an introduction to the ideas involved in Chabauty's method, we warmly recommend the thesis [Wetherell 1997] and the survey paper [McCallum and Poonen 2010] as well as [Coleman 1985a].

Integration. Let p be a (finite) rational prime. Let K_υ be a finite extension of $\mathbb{Q}_p$ and $\mathbb{O}_\upsilon$ the ring of integers in K_υ. Let $\mathcal{W}$ be a smooth proper connected scheme of finite type over $\mathbb{O}_\upsilon$, and write W for the generic fiber. Coleman [1985b, Section II] describes how to integrate "differentials of the second kind" on W. We shall however only be concerned with global 1-forms (i.e., differentials of the first kind) and so shall restrict our attention to these. Among the properties of integration [Coleman 1985b, Section II] that we shall need are the following, where P, Q, R lie in $W(K_\upsilon)$, while ω and ω' are global 1-forms on W, and α is an element of K_υ:

(i)
$$\int_P^Q \omega = -\int_Q^P \omega.$$

(ii)
$$\int_Q^P \omega + \int_P^R \omega = \int_Q^R \omega.$$

(iii)
$$\int_Q^P (\omega + \omega') = \int_Q^P \omega + \int_Q^P \omega'.$$

(iv)
$$\int_Q^P \alpha\omega = \alpha \int_Q^P \omega.$$

We shall also need a change of variables formula [Coleman 1985b, Theorem 2.7]: if $\mathcal{W}_1$ and $\mathcal{W}_2$ are smooth proper connected schemes of finite type over $\mathbb{O}_\upsilon$ and $\varrho : W_1 \to W_2$ is a morphism of their generic fibers, then

$$\int_Q^P \varrho^*\omega = \int_{\varrho(Q)}^{\varrho(P)} \omega$$

for all global 1-forms ω on W_2 and P, $Q \in W_1(K_\upsilon)$.

Now let A be an abelian variety of dimension g over K_υ, and write Ω_A for the K_υ-space of global 1-forms on A. Consider the pairing

$$\Omega_A \times A(K_\upsilon) \to K_\upsilon, \quad (\omega, P) \mapsto \int_0^P \omega. \tag{3}$$

This pairing is bilinear. It is K_υ-linear on the left by (iii) and (iv). It is $\mathbb{Z}$-linear on the right; this is a straightforward consequence [Coleman 1985b, Theorem 2.8] of the "change of variables formula". The kernel on the left is 0, and on the right is the torsion subgroup of $A(K_\upsilon)$ [Bourbaki 1989, III.7.6].

Notation. Henceforth, we shall be concerned with curves over number fields and their Jacobians. We fix once and for all the following notation:

- K is a number field,
- C is a smooth projective absolutely irreducible curve defined over K of genus at least 2,
- J is the Jacobian of C,

υ is a non-Archimedean prime of K of good reduction for C,

K_υ is the completion of K at υ,

k_υ is the residue field of K at υ,

$\mathbb{O}_\upsilon$ is the ring of integers in K_υ,

$x \mapsto \tilde{x}$ is the natural map $\mathbb{O}_\upsilon \to k_\upsilon$,

$\mathscr{C}_\upsilon$ is a minimal regular proper model for C over $\mathbb{O}_\upsilon$,

$\tilde{C}_\upsilon$ is the special fiber of $\mathscr{C}_\upsilon$ at υ, and

Ω_{C/K_υ} is the K_υ-vector space of global 1-forms on C.

Integration on curves and Jacobians. For any field extension M/K (not necessarily finite), we shall write $\Omega_{C/M}$ and $\Omega_{J/M}$ for the M-vector spaces of global 1-forms on C/M and J/M, respectively. We shall assume the existence of some $P_0 \in C(K)$. The point P_0 gives rise to an Abel–Jacobi map

$$j : C \hookrightarrow J, \quad P \mapsto [P - P_0].$$

It is well known that the pull-back $j^* : \Omega_{J/K} \to \Omega_{C/K}$ is an isomorphism of K-vector spaces [Milne 1986, Proposition 2.2]. Clearly any two Abel–Jacobi maps differ by a translation on J. As 1-forms on J are translation invariant, the map j^* is independent of the choice of P_0 [Wetherell 1997, Section 1.4]. Let υ be a non-Archimedean place for K. The isomorphism j^* extends to an isomorphism $\Omega_{J/K_\upsilon} \to \Omega_{C/K_\upsilon}$, which we shall also denote j^*. For any global 1-form $\omega \in \Omega_{J/K_\upsilon}$ and any two points $P, Q \in C(K_\upsilon)$, we have

$$\int_Q^P j^* \omega = \int_{jQ}^{jP} \omega = \int_0^{[P-Q]} \omega$$

using the properties of integration above. We shall henceforth use j^* to identify Ω_{C/K_υ} with Ω_{J/K_υ}. With this identification, the pairing (3) with $J = A$ gives the bilinear pairing

$$\Omega_{C/K_\upsilon} \times J(K_\upsilon) \to K_\upsilon, \quad \left(\omega, \left[\sum P_i - Q_i\right]\right) \mapsto \sum \int_{Q_i}^{P_i} \omega, \qquad (4)$$

whose kernel on the right is 0 and on the left is the torsion subgroup of $J(K_\upsilon)$. We ease notation a little by defining, for divisor class $D = \sum P_i - Q_i$ of degree 0, the integral

$$\int_D \omega = \sum \int_{Q_i}^{P_i} \omega.$$

Note that this integral depends on the equivalence class of D and not on its decomposition as $D = \sum P_i - Q_i$.

Uniformizers. The usual Chabauty approach when studying rational points in a residue class is to work with a local coordinate (defined shortly) and create power-series equations in terms of the local coordinate whose solutions, roughly speaking, contain the rational points. In our case, we find it more convenient to shift the local coordinate so that it becomes a uniformizer at a rational point in the residue class.

Fix a non-Archimedean place v of good reduction for C and a minimal regular proper model $\mathscr{C}_v$ for C over v. Since our objective is explicit computation, we point out that in our case of good reduction, such a model is simply a system of equations for the nonsingular curve that reduces to a nonsingular system modulo v. Let $Q \in C(K_v)$, and let $\tilde{Q}$ be its reduction on the special fiber $\tilde{C}_v$; as we are considering a regular model, $\tilde{Q}$ is a smooth point. Choose a rational function $s_Q \in K_v(C)$ so that the maximal ideal in $\mathbb{O}_{\mathscr{C}_v, \tilde{Q}}$ is (s_Q, π), where π is a uniformizing element for K_v. The function s_Q is called [Lorenzini and Tucker 2002, Section 1] a *local coordinate* at Q. Let $t_Q = s_Q - s_Q(Q)$. We shall refer to t_Q, constructed as above, as a *well behaved uniformizer* at Q. A uniformizer at a smooth point Q means a local coordinate that vanishes with multiplicity 1 at Q. The reason for the adjective "well behaved" will be clear from Lemma 3.1 below.

Before stating the lemma, we define the *v-unit ball around Q* to be

$$\mathscr{B}_v(Q) = \{\, P \in C(K_v) : \tilde{P} = \tilde{Q} \,\}. \tag{5}$$

Lemma 3.1.

  (i) t_Q *is a uniformizer at* Q.

 (ii) $\tilde{t}_Q$ *is a uniformizer at* $\tilde{Q}$.

(iii) t_Q *defines a bijection*

$$\mathscr{B}_v(Q) \to \pi\mathbb{O}_v, \quad P \mapsto t_Q(P),$$

    where π is any uniformizing element for K_v. In particular, for $P \in \mathscr{B}_v(Q)$, we have $t_Q(P) = 0$ if and only if $P = Q$.

Proof. Parts (i) and (ii) are clear from the construction. Part (iii) is standard; see [Lorenzini and Tucker 2002, Section 1; Wetherell 1997, Sections 1.7 and 1.8] for example. $\qquad\square$

Estimating integrals on curves.

Lemma 3.2. *Let p be an odd rational prime that does not ramify in K. Let v be a place of K above p. Let $Q \in C(K_v)$, and let $t_Q \in K_v(C)$ be a well behaved uniformizer at Q. Let $\omega \in \Omega_{\mathscr{C}_v/\mathbb{O}_v}$. Then there is a power series*

$$\phi(x) = \alpha_1 x + \alpha_2 x^2 + \alpha_3 x^3 + \cdots \in K_v[[x]] \tag{6}$$

that converges for $x \in \pi\mathbb{O}_\upsilon$ such that

$$\int_Q^P \omega = \phi(z)$$

for all $P \in \mathcal{B}_\upsilon(Q)$, where $z = t_Q(P)$. Moreover, the coefficient α_1 is given by

$$\alpha_1 = \left(\frac{\omega}{dt_Q}\right)(Q) \in \mathbb{O}_\upsilon, \tag{7}$$

where we interpret ω/dt_Q as an element of $K_\upsilon(C)$, and

$$\phi(z) = \int_Q^P \omega \equiv \alpha_1 z \pmod{z^2\mathbb{O}_\upsilon}. \tag{8}$$

Proof. We can expand ω (after viewing it as an element in $\Omega_{\hat{\mathbb{O}}_Q}$) as a formal power series

$$\omega = (\gamma_0 + \gamma_1 t_Q + \gamma_2 t_Q^2 + \cdots)dt_Q,$$

where the coefficients γ_i belong to $\mathbb{O}_\upsilon$ (see [Lorenzini and Tucker 2002, Proposition 1.6; Wetherell 1997, Chapters 1.7 and 1.8] for example); here we have not used the assumption that $t_Q(Q) = 0$, merely that t_Q is a local coordinate at Q. We note that $(\omega/dt_Q)(Q) = \gamma_0$ and is hence integral.

Let $P \in \mathcal{B}_\upsilon(Q)$ and $z = t_Q(P)$. Then (see [Lorenzini and Tucker 2002, Proposition 1.3] for example)

$$\int_Q^P \omega = \sum_{j=0}^\infty \frac{\gamma_j}{j+1} z^{j+1}. \tag{9}$$

Thus, in (6), we take the coefficients $\alpha_i = \gamma_{i-1}/i$. The power series $\phi(x)$ converges for $x \in \pi\mathbb{O}_\upsilon$ as the γ_i are integral. In particular, $\phi(z)$ converges since $\mathrm{ord}_\upsilon(z) \geq 1$ by Lemma 3.1(iii). To complete the proof, observe that

$$\phi(z) - \alpha_1 z = z^2\left(\frac{\gamma_1}{2} + \frac{\gamma_2}{3}z + \frac{\gamma_3}{4}z^2 + \cdots\right).$$

We must show the sum in brackets belongs to $\mathbb{O}_\upsilon$. Thus, it is sufficient to show that

$$\mathrm{ord}_\upsilon(j+2) \leq j$$

for all $j \geq 0$. But $K_\upsilon/\mathbb{Q}_p$ is unramified, and so $\mathrm{ord}_\upsilon(j+2) = \mathrm{ord}_p(j+2)$. Hence, we need to show that $\mathrm{ord}_p(j+2) \leq j$ for all $j \geq 0$ and all odd primes p. This is now an easy exercise. $\square$

4. Chabauty in a single unit ball

Let C be a smooth projective curve over a number field K. Let J be the Jacobian of C, and write r for the rank of the Mordell–Weil group $J(K)$. Let $D_1, \ldots, D_r$ be a basis for a free subgroup of finite index in $J(K)$.

Let p a rational prime such that

(p1) p is odd,

(p2) p is unramified in K, and

(p3) every prime υ of K above p is a prime of good reduction for the curve C.

Let $Q \in C(K)$. For $\upsilon \mid p$, let $\mathscr{B}_\upsilon(Q)$ be as in (5), and define the *p-unit ball around Q* to be

$$\mathscr{B}_p(Q) = \prod_{\upsilon \mid p} \mathscr{B}_\upsilon(Q). \tag{10}$$

We will shortly give a criterion for a point $Q \in C(K)$ to be the unique K-rational point in its own p-unit ball. Our criterion and its proof are rather involved. As motivation, we first explain the case $K = \mathbb{Q}$.

Motivation. Suppose $K = \mathbb{Q}$. Let $P \in C(\mathbb{Q}) \cap \mathscr{B}_p(Q)$. We will write down equations that give information about P and which, with appropriate assumptions, allow us to show that $P = Q$.

Let m be the index

$$m := [J(\mathbb{Q}) : \langle D_1, \ldots, D_r \rangle].$$

There are integers $n'_1, \ldots, n'_r$ such that

$$m(P - Q) = n'_1 D_1 + \cdots + n'_r D_r, \tag{11}$$

where the equality takes place in $\mathrm{Pic}^0(C)$. Let $\omega_1, \ldots, \omega_g$ be a $\mathbb{Z}_p$ basis for $\Omega_{\mathscr{C}_p / \mathbb{Z}_p}$. By the properties of integration explained in Section 3,

$$m \int_Q^P \omega_i = n'_1 \tau_{i,1} + \cdots + n'_r \tau_{i,r}, \quad i = 1, \ldots, g,$$

where the $\tau_{i,j}$ are given by

$$\tau_{i,j} = \int_{D_j} \omega_i, \quad j = 1, \ldots, r.$$

Let $n_i = n'_i / m \in \mathbb{Q}$. Thus,

$$\int_Q^P \omega_i = n_1 \tau_{i,1} + \cdots + n_r \tau_{i,r}.$$

Let t_Q be a well behaved uniformizer at Q as defined on page 771, and write $z = t_Q(P)$. By Lemma 3.1, we know that $z \in p\mathbb{Z}_p$. By Lemma 3.2, there are power series ϕ_i with coefficients in $\mathbb{Q}_p$, converging on $p\mathbb{Z}_p$, such that

$$\int_Q^P \omega_i = \phi_i(z) = \alpha_i z + \alpha_i' z^2 + \alpha_i'' z^2 + \cdots.$$

Then

$$n_1 \tau_{i,1} + \cdots + n_r \tau_{i,r} = \phi_i(z), \quad i = 1, \ldots, g. \tag{12}$$

This is a system of g equations in $r + 1$ unknowns, $n_1, n_2, \ldots, n_r, z$. If $r \le g - 1$, we can use linear algebra to eliminate $n_1, n_2, \ldots, n_r$ to obtain $g - r$ equations

$$\theta_1(z) = \theta_2(z) = \cdots = \theta_{g-r}(z) = 0,$$

where the $\theta_i(z)$ are power series with coefficients in $\mathbb{Q}_p$, converging on $p\mathbb{Z}_p$. In practical computations, it is usual at this point to use Newton polygons and other techniques to bound the number of solutions to this system with $z \in p\mathbb{Z}_p$. By Lemma 3.1(iii), the map $\mathcal{B}_p(Q) \to p\mathbb{Z}_p$ given by $P \mapsto t_Q(P) = z$ is bijective; thus, we also obtain a bound on the number of $P \in C(\mathbb{Q}) \cap \mathcal{B}_p(Q)$.

Now we want a practical criterion for Q to be the unique rational point in its p-unit ball or equivalently that $z = 0$. The system of equations (12) is easier to analyze if we consider only the linear terms of the power series ϕ_i. By (8),

$$n_1 \tau_{i,1} + \cdots + n_r \tau_{i,r} \equiv \alpha_i z \pmod{z^2 \mathbb{Z}_p},$$

where $\alpha_i = (\omega_i / dt_Q)(Q) \in \mathbb{Z}_p$. Let T be the $g \times r$ matrix $(\tau_{i,j})$ — this has entries in $\mathbb{Q}_p$. Let A be the column vector (α_i). We can rewrite this linear system of congruences in matrix form

$$T\boldsymbol{n} \equiv Az \pmod{z^2 \mathbb{Z}_p},$$

where $\boldsymbol{n}$ is the column vector (n_i). Choose a non-negative integer a such that $p^a T$ has entries in $\mathbb{Z}_p$. Let U be a unimodular matrix with entries in $\mathbb{Z}_p$ so that $U \cdot p^a T$ is in Hermite normal form (HNF) (see [Cohen 2000, Section 1.4.2] for the theory of HNF). Let h be the number of zero rows of $U \cdot p^a T$; as $U \cdot p^a T$ is in HNF, these are the last h rows. Let $M_p(Q)$ be the vector in $\mathbb{Z}_p^h$ formed by the last h elements of UA.

Lemma 4.1. *With the above assumptions and notation, suppose $h > 0$ and let $\tilde{M}_p(Q) \in \mathbb{F}_p^h$ denote the vector obtained by reducing $M_p(Q)$ mod p. If $\tilde{M}_p(Q) \neq \boldsymbol{0}$, then $C(\mathbb{Q}) \cap \mathcal{B}_p(Q) = \{Q\}$.*

Proof. From the above, we know that

$$M_p(Q)z \equiv 0 \pmod{z^2 \mathbb{Z}_p}$$

and that $M_p(Q)$ has entries in $\mathbb{Z}_p$. Suppose β is some entry of $M_p(Q)$ such that $\beta \not\equiv 0 \pmod{p}$. Then $\beta z \equiv 0 \pmod{z^2 \mathbb{Z}_p}$. As $z \in p\mathbb{Z}_p$, we must have that $z = 0$. By the above discussion, this forces $P = Q$. □

We do not take any credit for this lemma; the ideas involved can found in [Coleman 1985a]. We have however expressed our lemma and the ideas leading up to it in a way that motivates our generalization to the case where $[K : \mathbb{Q}] > 1$.

The general case. We return to the general case where K is a number field. The p-unit ball $\mathcal{B}_p(Q)$ is defined in (10). We would like to give a criterion for Q to be the unique K-rational point in its p-unit ball. Again, let $P \in C(K) \cap \mathcal{B}_p(Q)$. We will write equations that give information about P and devise a Chabauty-style criterion that forces $P = Q$. In the case $K = \mathbb{Q}$, the variable $z = t_Q(P)$ measured the "distance from P to Q along $C(\mathbb{Q}_p)$". Now the field K has several embeddings K_υ with $\upsilon \mid p$. We will need to replace z by a vector whose entries "measure the distances from P to Q along $\prod_{\upsilon|p} C(K_\upsilon)$".

To state our criterion — Theorem 2 below — we need to define a pair of matrices T and A. The matrix T depends on the basis $D_1, \ldots, D_r$. The matrix A depends on the point $Q \in C(K)$. Let $\upsilon_1, \ldots, \upsilon_n$ be the places of K above p. For each place υ above p, we fix once and for all a $\mathbb{Z}_p$-basis $\theta_{\upsilon,1}, \ldots, \theta_{\upsilon,d_\upsilon}$ for $\mathbb{O}_\upsilon$, where $d_\upsilon = [K_\upsilon : \mathbb{Q}_p]$. Of course, $d_\upsilon = [\mathbb{O}_\upsilon : \mathbb{Z}_p] = [k_\upsilon : \mathbb{F}_p]$ as p is unramified in K. We also choose an $\mathbb{O}_\upsilon$-basis $\omega_{\upsilon,1}, \ldots, \omega_{\upsilon,g}$ for $\Omega_{\mathscr{C}_\upsilon/\mathbb{O}_\upsilon}$.

Now fix υ above p, and let $\omega \in \Omega_{\mathscr{C}_\upsilon/\mathbb{O}_\upsilon}$. Let

$$\tau_j = \int_{D_j} \omega, \quad j = 1, \ldots, r. \tag{13}$$

Write

$$\tau_j = \sum_{i=1}^{d_\upsilon} t_{ij} \theta_{\upsilon,i}, \quad t_{ij} \in \mathbb{Q}_p. \tag{14}$$

Let

$$T_{\upsilon,\omega} = (t_{ij})_{i=1,\ldots,d_\upsilon,\ j=1,\ldots,r}; \tag{15}$$

that is, $T_{\upsilon,\omega}$ is the $d_\upsilon \times r$ matrix with entries t_{ij}. Recall that $\omega_{\upsilon,1}, \ldots, \omega_{\upsilon,g}$ is a basis for $\Omega_{\mathscr{C}_\upsilon/\mathbb{O}_\upsilon}$. Let

$$T_\upsilon = \begin{pmatrix} T_{\upsilon,\omega_{\upsilon,1}} \\ T_{\upsilon,\omega_{\upsilon,2}} \\ \vdots \\ T_{\upsilon,\omega_{\upsilon,g}} \end{pmatrix}; \tag{16}$$

this is a $gd_v \times r$ matrix with entries in $\mathbb{Q}_p$. We now define the matrix T needed for our criterion below:

$$T = \begin{pmatrix} T_{v_1} \\ T_{v_2} \\ \vdots \\ T_{v_n} \end{pmatrix}. \tag{17}$$

Note that T is a $gd \times r$ matrix with entries in $\mathbb{Q}_p$, where $d = [K : \mathbb{Q}] = d_{v_1} + \cdots + d_{v_n}$.

Let $Q \in C(K)$. We now define the second matrix A (depending on Q) needed to state our criterion for $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$. For each place v of K above p, we have chosen a minimal proper regular model $\mathscr{C}_v$. Let t_Q be a well behaved uniformizer at Q as defined in Section 3. Let $\omega \in \Omega_{\mathscr{C}_v/\mathbb{O}_v}$, and let α be given by (7). By Lemma 3.2, $\alpha \in \mathbb{O}_v$. Recall we have fixed a basis $\theta_{v,1}, \ldots, \theta_{v,d_v}$ for $\mathbb{O}_v/\mathbb{Z}_p$. Write

$$\alpha \cdot \theta_{v,j} = \sum_{i=1}^{d_v} a_{ij} \theta_{v,i}, \quad j = 1, \ldots, d_v, \tag{18}$$

with $a_{ij} \in \mathbb{Z}_p$. Let

$$A_{v,\omega} = (a_{ij})_{i,j=1,\ldots,d_v}. \tag{19}$$

Let

$$A_v = \begin{pmatrix} A_{v,\omega_1} \\ A_{v,\omega_2} \\ \vdots \\ A_{v,\omega_g} \end{pmatrix}; \tag{20}$$

this is a $d_v g \times d_v$ matrix with entries in $\mathbb{Z}_p$. Let

$$A = \begin{pmatrix} A_{v_1} & 0 & \cdots & 0 \\ 0 & A_{v_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{v_n} \end{pmatrix}. \tag{21}$$

Then A is a $dg \times d$ matrix with entries in $\mathbb{Z}_p$.

Choose a non-negative integer a such that $p^a T$ has entries in $\mathbb{Z}_p$. Let U be a unimodular matrix with entries in $\mathbb{Z}_p$ such that $U \cdot p^a T$ is in HNF. Let h be the number of zero rows of $U \cdot p^a T$; these are the last h rows. Let $M_p(Q)$ be the $h \times d$ matrix (with entries in $\mathbb{Z}_p$) formed by the last h rows of UA.

Theorem 2. *With the assumptions and notation above, let $\tilde{M}_p(Q)$ denote the matrix with entries in $\mathbb{F}_p$ obtained by reducing $M_p(Q)$ modulo p. If $\tilde{M}_p(Q)$ has rank d, then $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$.*

Remarks. (i) Let $u_1, \ldots, u_h$ be a $\mathbb{Z}_p$-basis for the kernel of the homomorphism of $\mathbb{Z}_p$-modules $\mathbb{Z}_p^{gd} \to \mathbb{Z}_p^r$ given by $p^a T$. Then $u_1 A, \ldots, u_h A$ span the same $\mathbb{Z}_p$-module as the rows of $M_p(Q)$, showing that the rank of $\tilde{M}_p(Q)$ is independent of the choice of U.

(ii) Since the matrix T is $gd \times r$, it is evident that $h \geq \max(gd - r, 0)$ and, very likely, $h = \max(gd - r, 0)$. Now the matrix $\tilde{M}_p(Q)$ is $h \times d$, and so a necessary condition for the criterion to hold is that $h \geq d$. Thus, it is sensible to apply the theorem when $gd - r \geq d$ or equivalently when $r \leq d(g - 1)$.

(iii) In practice, we do not compute the matrix T exactly, merely an approximation to it. Thus, we won't be able to provably determine h unless $h = \max(gd - r, 0)$.

Proof of Theorem 2. Suppose that $P \in C(K) \cap \mathcal{B}_p(Q)$. We need to show $P = Q$. Let m be the index

$$m := [J(K) : \langle D_1, \ldots, D_r \rangle].$$

There are integers $n_1', \ldots, n_r'$ such that (11) holds, where the equality takes place in $\mathrm{Pic}^0(C)$.

Let v be one of the places $v_1, \ldots, v_n$ above p. Recall that we have chosen a well behaved uniformizer t_Q at Q. Write $z = t_Q(P)$. By Lemma 3.1(iii), $\mathrm{ord}_v(z) \geq 1$. We will show that $z = 0$, and so again by Lemma 3.1(iii), $P = Q$, which is what we want to prove.

We write

$$z = z_{v,1} \theta_{v,1} + \cdots + z_{v,d_v} \theta_{v,d_v},$$

where $z_{v,i} \in \mathbb{Z}_p$. As $\tilde{\theta}_{v,1}, \ldots, \tilde{\theta}_{v,d_v}$ is a basis for $k_v/\mathbb{F}_p$ and $\mathrm{ord}_v(z) \geq 1$, we see that $\mathrm{ord}_v(z_{v,i}) \geq 1$ for $i = 1, \ldots, d_v$. Let

$$s_v = \min_{1 \leq i \leq d_v} \mathrm{ord}_p(z_{v,i}). \tag{22}$$

We will show that $s_v = \infty$, which implies that $z_{v,i} = 0$ for $i = 1, \ldots, d_v$, and so $z = 0$ as required. For now, we note that $s_v \geq 1$.

Now fix an $\omega \in \Omega_{\mathscr{C}_v/\mathbb{O}_v}$, and let $\alpha \in \mathbb{O}_v$ be as in Lemma 3.2; by that lemma,

$$\int_Q^P \omega = \alpha z + \beta z^2$$

for some $\beta \in \mathbb{O}_v$. However, by (11) and the properties of integration explained in Section 3,

$$m \int_Q^P \omega = n_1' \tau_1 + \cdots + n_r' \tau_r,$$

where the τ_j are given in (13). Let $n_i = n_i'/m \in \mathbb{Q}$. Thus,

$$\int_Q^P \omega = n_1\tau_1 + \cdots + n_r\tau_r.$$

Hence,

$$n_1\tau_1 + \cdots + n_r\tau_r = \alpha(z_{v,1}\theta_{v,1} + \cdots + z_{v,d_v}\theta_{v,d_v}) + \beta(z_{v,1}\theta_{v,1} + \cdots + z_{v,d_v}\theta_{v,d_v})^2.$$

From this and (22), we obtain

$$n_1\tau_1 + \cdots + n_r\tau_r \equiv z_{v,1}(\alpha\theta_{v,1}) + \cdots + z_{v,d_v}(\alpha\theta_{v,d_v}) \pmod{p^{2s_v}\mathbb{O}_v}. \qquad (23)$$

Write

$$\boldsymbol{n} = \begin{pmatrix} n_1 \\ n_2 \\ \vdots \\ n_r \end{pmatrix} \quad \text{and} \quad \boldsymbol{z_v} = \begin{pmatrix} z_{v,1} \\ z_{v,2} \\ \vdots \\ z_{v,d_v} \end{pmatrix}, \qquad (24)$$

and note that the entries of $\boldsymbol{n}$ are in $\mathbb{Q}$ and the entries of $\boldsymbol{z_v}$ are in $p^{s_v}\mathbb{Z}_p$. Recall that we have expressed $\tau_j = \sum t_{ij}\theta_{v,i}$ in (14) and $\alpha \cdot \theta_{v,j} = \sum a_{ij}\theta_{v,i}$ in (18), where t_{ij} are in $\mathbb{Q}_p$ and the a_{ij} are in $\mathbb{Z}_p$. Substituting in (23) and comparing the coefficients for $\theta_{v,i}$, we obtain

$$T_{v,\omega}\boldsymbol{n} \equiv A_{v,\omega}\boldsymbol{z_v} \pmod{p^{2s_v}},$$

where $T_{v,\omega}$ and $A_{v,\omega}$ are respectively given in (15) and (19).

Let T_v and A_v be as given in (16) and (20), respectively. Then

$$T_v\boldsymbol{n} \equiv A_v\boldsymbol{z_v} \pmod{p^{2s_v}}.$$

Now let

$$\boldsymbol{z} = \begin{pmatrix} \boldsymbol{z_{v_1}} \\ \boldsymbol{z_{v_2}} \\ \vdots \\ \boldsymbol{z_{v_n}} \end{pmatrix}.$$

Then $\boldsymbol{z}$ is of length $d = [K : \mathbb{Q}]$ with entries in $p\mathbb{Z}_p$. Write

$$s = \min_{v=v_1,\dots,v_n} s_v = \min_{i,j} \operatorname{ord}_{v_j}(z_{i,v_j}), \qquad (25)$$

where the s_v are defined in (22). Clearly $s \geq 1$. It is sufficient to show that $s = \infty$ since then all of the $z_{i,v_j} = 0$, implying that $P = Q$.

Let T and A be as given in (17) and (21). Then

$$T\boldsymbol{n} \equiv A\boldsymbol{z} \pmod{p^{2s}}, \qquad (26)$$

where we note once again that T is $dg \times r$ with entries in $\mathbb{Q}_p$ and A is $dg \times d$ with entries in $\mathbb{Z}_p$.

Let U, $M_p(Q)$, and h be as in the paragraph preceding the statement of the theorem. Suppose that $\tilde{M}_p(Q)$ has rank d. Suppose $s < \infty$, and we will derive a contradiction. Recall that the last h rows of UT are zero. From (26), we have that $M_p(Q)\mathbf{z} \equiv 0 \pmod{p^{2s}}$ since, by definition, $M_p(Q)$ is the matrix formed by the last h rows of UA. In particular, $M_p(Q)$ has entries in $\mathbb{Z}_p$ since both U and A have entries in $\mathbb{Z}_p$. From the definition of s in (25), we can write $\mathbf{z} = p^s \mathbf{w}$, where the entries of $\mathbf{w}$ are in $\mathbb{Z}_p$ and $\mathbf{w} \not\equiv \mathbf{0} \pmod{p}$. However, $M_p(Q)\mathbf{w} \equiv 0 \pmod{p^s}$, and as $s \geq 1$, we have that $M_p(Q)\mathbf{w} \equiv 0 \pmod{p}$. Since $\mathbf{w} \in \mathbb{Z}_p^d$, if $\tilde{M}_p(Q)$ has rank d, then $\mathbf{w} \equiv \mathbf{0} \pmod{p}$, giving the desired contradiction. □

Remark. In the above, we are only considering the linear terms of the power series $\phi(z)$, and this is enough for our criterion for $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$. Of course, if there is another rational point sharing the same p-unit ball as Q, then our criterion will fail. It may then be possible to obtain an upper bound for the number of rational points in the p-unit ball by writing out higher terms of the power series and eliminating the n_i. This is likely to be technical, and we have not attempted it in practice. However, we note that we can always choose a large enough prime p so that no two known rational points share the same p-unit ball. If our necessary condition $r \leq d(g-1)$ is satisfied, then we expect to be able to find a prime p so that our Chabauty criterion succeeds in showing $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$ for all known rational points Q. We then expect to be able to complete our determination of the rational points using the Mordell–Weil sieve as explained below.

5. Chabauty and the Mordell–Weil sieve

For the complete determination of the set of rational points on a curve of genus at least 2, it is often necessary to combine Chabauty with the Mordell–Weil sieve. Before giving details of how this works in our case, we sketch the idea behind the Mordell–Weil sieve.

We continue with the notation of the previous sections. In particular, C is a smooth curve defined over a number field K and P_0 is a fixed K-rational point on C. Let $\mathcal{K}$ be the subset of known K-rational points on C, and suppose that we would like to prove that $C(K) = \mathcal{K}$. Using our Theorem 2, it may be possible to show, for some prime p, that $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$ for every $Q \in \mathcal{K}$. In this situation, Chabauty tells us that to show that $C(K) = \mathcal{K}$, all you have to do is show that every $P \in C(K)$ belongs to the p-unit ball $\mathcal{B}_p(Q)$ for some $Q \in \mathcal{K}$. This is where we turn to the Mordell–Weil sieve.

Let υ be a place of K of good reduction for C. Let red denote the natural maps

$$\text{red} : C(K) \to C(k_\upsilon) \quad \text{and} \quad \text{red} : J(K) \to J(k_\upsilon).$$

Let $_J$ denote the Abel–Jacobi maps

$$C(K) \hookrightarrow J(K) \quad \text{and} \quad C(k_\upsilon) \hookrightarrow J(k_\upsilon)$$

respectively associated to P_0 and $\tilde{P}_0$. A glance at the commutative diagram

$$
\begin{array}{ccc}
C(K) & \xrightarrow{\;J\;} & J(K) \\
\downarrow{\scriptstyle \text{red}} & & \downarrow{\scriptstyle \text{red}} \\
C(k_\upsilon) & \xrightarrow{\;J\;} & J(k_\upsilon)
\end{array}
$$

shows that $_J(C(K)) \subseteq W_\upsilon + L_\upsilon$, where $L_\upsilon := \ker(J(K) \to J(k_\upsilon))$ is a subgroup of finite index in $J(K)$ and W_υ is a set of coset representatives for $\mathrm{red}^{-1}{}_J(C(k_\upsilon))$. In practice, if one knows a Mordell–Weil basis for $J(K)$, then W_υ and L_υ are straightforward to compute. Now let S be a finite set of places υ of K, all of good reduction for C. We can write

$$\bigcap_{\upsilon \in S}(W_\upsilon + L_\upsilon) = W_S + L_S, \tag{27}$$

where W_S is a finite subset of $J(K)$ and $L_S = \cap_{\upsilon \in S} L_\upsilon$ (of course, the elements of W_S are unique up to translation by elements of L_S). Clearly $_J(C(K)) \subseteq W_S + L_S$. With a judicious choice of places S, it is sometimes possible to show that

$$_J(\mathcal{H}) + L_S = W_S + L_S \supseteq {}_J(C(K)),$$

where the index $[J(K) : L_S]$ is large. In other words, any rational point is "close" to a known rational point in the profinite topology on the Mordell–Weil group. Suppose next that there is some rational prime p satisfying assumptions (p1)–(p3) on page 773 and that L_S is contained in the kernel of the diagonal map

$$J(K) \to \prod_{\upsilon \mid p} J(k_\upsilon).$$

It then follows that, for every $P \in C(K)$, there is some $Q \in \mathcal{H}$ (one of the known rational points) such that $P \in \mathcal{B}_p(Q)$. We may then attempt to apply Theorem 2 to show that $C(K) \cap \mathcal{B}_p(Q) = \{Q\}$ for all $Q \in \mathcal{H}$; if we can show this, then we will have shown that $C(K) = \mathcal{H}$.

The standard references for the Mordell–Weil sieve, e.g., [Bruin and Elkies 2002; Bruin and Stoll 2008; 2010; Bugeaud et al. 2008], as well as the above sketch assume full knowledge of the Mordell–Weil group. For reasons that we now explain, we need to adapt the Mordell–Weil sieve to work with a subgroup of the Mordell–Weil group of finite (but unknown) index. Let L_0 be a subgroup of $J(K)$ of finite index containing the free subgroup L generated by $D_1, \ldots, D_r$ of the previous section. We can take $L_0 = L$, but for our purpose, it is preferable to

include the torsion subgroup of $J(K)$ in L_0. The usual p-saturation method [Siksek 1995b; 1995a; Flynn and Smart 1997] shows how to enlarge L_0 so that its index in $J(K)$ is not divisible by any given small prime p. One expects, after checking p-saturation for all small primes p up to some large bound, that L_0 is in fact equal to $J(K)$. However, proving that $J(K) = L_0$ requires an explicit theory of heights on the Jacobian J. This is not yet available for Jacobians of curves of genus at least 3. For Jacobians of curves of genus 2, there is an explicit theory of heights [Flynn 1995a; Flynn and Smart 1997; Stoll 1999; 2002b] though the bounds over number fields other than the rationals are likely to be impractically large.

Before we give the details, we point out that substantial improvements can be made to the version of the Mordell–Weil sieve outlined below. It has certainly been sufficient for the examples we have computed so far (including the ones detailed in the next section). But we expect that for some other examples it will be necessary (though not difficult) to incorporate the improvements to the Mordell–Weil sieve found in [Bruin and Stoll 2008; 2010].

Lemma 5.1 (Mordell–Weil sieve). *Let L_0 be a subgroup of $J(K)$ of finite index $n = [J(K) : L_0]$. Let $P_0 \in C(K)$, and let $\jmath$ denote the Abel–Jacobi maps associated to P_0 as above. Let $\upsilon_1, \ldots, \upsilon_s$ be places of K such that each $\upsilon = \upsilon_i$ satisfies the following two conditions:*

(υ1) υ is a place of good reduction for C and

(υ2) the index n is coprime to $\#J(k_\upsilon)$.

To ease notation, write k_i for the residue field k_{υ_i}. Define inductively a sequence of subgroups

$$L_0 \supseteq L_1 \supseteq L_2 \supseteq L_3 \supseteq \cdots \supseteq L_s$$

and finite subsets $W_0, W_1, \ldots, W_s \subseteq L_0$ as follows. Let $W_0 = \{\mathbf{0}\}$. Suppose we have defined L_i and W_i, where $i \leq s - 1$. Let L_{i+1} be the kernel of the composition

$$L_i \hookrightarrow J(K) \to J(k_{i+1}).$$

To define W_{i+1}, choose a complete set $\mathfrak{Q}$ of coset representatives for L_i / L_{i+1} and let

$$W'_{i+1} = \{\, \mathbf{w} + \mathbf{q} : \mathbf{w} \in W_i \text{ and } \mathbf{q} \in \mathfrak{Q} \,\}.$$

Let

$$W_{i+1} = \{\, \mathbf{w} \in W'_{i+1} : \mathrm{red}(\mathbf{w}) \in \jmath(C(k_{i+1})) \,\}.$$

Then, for every $i = 0, \ldots, s$ and every $Q \in C(K)$, there is some $\mathbf{w} \in W_i$ such that

$$n(\jmath(Q) - \mathbf{w}) \in L_i. \tag{28}$$

Remark. If $L_0 = J(K)$, there is no difference between the usual Mordell–Weil sieve sketched at the beginning of this section and the Mordell–Weil sieve of the lemma. However, we have expressed the Mordell–Weil sieve in the lemma iteratively as this reflects how it is used in practice; we compute the intersection (27) gradually rather than all at once.

Proof of Lemma 5.1. The proof is by induction on i. Since L_0 has index n in $J(K)$, (28) is true with $\boldsymbol{w} = 0$. Let $i \le s - 1$. Suppose $Q \in C(K)$, $\boldsymbol{w}' \in W_i$, and $\boldsymbol{l}' \in L_i$ satisfy

$$n(J(Q) - \boldsymbol{w}') = \boldsymbol{l}'. \tag{29}$$

By definition of L_{i+1}, the quotient group L_i/L_{i+1} is isomorphic to a subgroup of $J(k_{i+1})$. It follows from assumption ($\upsilon 2$) that n is coprime to the order of L_i/L_{i+1}. Recall that $\mathfrak{Q}$ was defined as a complete set of coset representatives for L_i/L_{i+1}. Thus, $n\mathfrak{Q}$ is also a set of coset representatives. Hence, we may express $\boldsymbol{l}' \in L_i$ as

$$\boldsymbol{l}' = n\boldsymbol{q} + \boldsymbol{l},$$

where $\boldsymbol{q} \in \mathfrak{Q}$ and $\boldsymbol{l} \in L_{i+1}$. Let $\boldsymbol{w} = \boldsymbol{w}' + \boldsymbol{q}$. Then $\boldsymbol{w} \in W'_{i+1}$. By (29), we see that

$$n(J(Q) - \boldsymbol{w}) = \boldsymbol{l}' - n\boldsymbol{q} = \boldsymbol{l} \in L_{i+1}.$$

To complete the inductive argument, all we need to show is that $\boldsymbol{w} \in W_{i+1}$ or equivalently that $\text{red}(\boldsymbol{w}) \in J(C(k_{i+1}))$. However, since L_{i+1} is contained in the kernel of $\text{red} : J(K) \to J(k_{i+1})$, we see that

$$n(J(\tilde{Q}) - \text{red}(\boldsymbol{w})) = 0 \quad \text{in } J(k_{i+1}).$$

Using the fact that n is coprime to $\#J(k_{i+1})$ once again gives $\text{red}(\boldsymbol{w}) = J(\tilde{Q})$ as required. $\square$

The following theorem puts together the Mordell–Weil sieve with Theorem 2 to give a criterion for $C(K) = \mathcal{K}$. It is precisely the argument sketched before Lemma 5.1 but adapted to take account of the possibility that the index n may not be 1.

Theorem 3 (Chabauty with the Mordell–Weil sieve). *We continue with the above notation and assumptions. Let $L_0 \supseteq L_1 \supseteq \cdots \supseteq L_s$ and $W_0, \ldots, W_s$ be the sequences constructed in Lemma 5.1. Let $\mathcal{K}$ be a subset of $C(K)$. Let $P_0 \in \mathcal{K}$, and let J denote the maps associated to P_0 as above. Suppose that for every $\boldsymbol{w} \in W_s$, there is a point $Q \in \mathcal{K}$ and a prime p such that the following conditions hold*:

(a) *p satisfies conditions (p1)–(p3) on page 773.*

(b) *In the notation of the previous section, the matrix $\tilde{M}_p(Q)$ has rank d.*

(c) *The kernel of the homomorphism*

$$J(K) \longrightarrow \prod_{\upsilon|p} J(k_\upsilon) \tag{30}$$

contains both the group L_s and the difference $\jmath(Q) - \mathbf{w}$.

(d) *The index $n = [J(K) : L_0]$ is coprime to the orders of the groups $J(k_\upsilon)$ for $\upsilon \mid p$.*

Then $C(K) = \mathcal{H}$.

Proof. Suppose that $P \in C(K)$. We would like to show that $P \in \mathcal{H}$. By Lemma 5.1, there is some $\mathbf{w} \in W_s$ such that $n(\jmath(P) - \mathbf{w}) \in L_s$. Let $Q \in \mathcal{H}$ and prime p satisfy conditions (a)–(d) of the theorem. By (c), L_s is contained in the kernel of (30), and hence,

$$n(\jmath(\tilde{P}) - \mathrm{red}(\mathbf{w})) = 0$$

in $J(k_\upsilon)$ for all $\upsilon \mid p$. Since p satisfies assumption (d), it follows that

$$\jmath(\tilde{P}) - \mathrm{red}(\mathbf{w}) = 0$$

in $J(k_\upsilon)$ for all $\upsilon \mid p$. But by assumption (c) again,

$$\jmath(\tilde{Q}) - \mathrm{red}(\mathbf{w}) = 0$$

in $J(k_\upsilon)$ for all $\upsilon \mid p$. It follows that $\tilde{P} = \tilde{Q}$ in $C(k_\upsilon)$ for all $\upsilon \mid p$. Hence, $P \in \mathcal{B}_p(Q)$, where $\mathcal{B}_p(Q)$ is the p-unit ball around Q defined in (10). By assumption (b) and Theorem 2, we see that $P = Q \in \mathcal{H}$, completing the proof. $\qquad\square$

6. The generalized Fermat equation with signature (2, 3, 10)

Let $p, q, r \in \mathbb{Z}_{\geq 2}$. The equation

$$x^p + y^q = z^r \tag{31}$$

is known as the generalized Fermat equation (or the Fermat–Catalan equation) with signature (p, q, r). As in Fermat's last theorem, one is interested in integer solutions x, y, and z. Such a solution is called *nontrivial* if $xyz \neq 0$ and *primitive* if x, y, and z are coprime. Let $\chi = p^{-1} + q^{-1} + r^{-1}$. The parametrization of nontrivial primitive solutions for (p, q, r) with $\chi \geq 1$ has now been completed [Edwards 2004]. The generalized Fermat conjecture [Darmon 1997; Darmon and Granville 1995] is concerned with the case $\chi < 1$. It states that the only nontrivial primitive solutions to (31) with $\chi < 1$ are those shows in Table 1.

The generalized Fermat conjecture has been established for many signatures (p, q, r) including for several infinite families of signatures: Fermat's last theorem (p, p, p) by Wiles and Taylor [Wiles 1995; Taylor and Wiles 1995], $(p, p, 2)$ and $(p, p, 3)$ by Darmon and Merel [1997], $(2, 4, p)$ by Ellenberg [2004] and Bennett

$$1 + 2^3 = 3^2, \qquad 17^7 + 76271^3 = 21063928^2,$$
$$2^5 + 7^2 = 3^4, \qquad 43^8 + 96222^3 = 30042907^2,$$
$$7^3 + 13^2 = 2^9, \qquad 33^8 + 1549034^2 = 15613^3,$$
$$2^7 + 17^3 = 71^2, \qquad 1414^3 + 2213459^2 = 65^7,$$
$$3^5 + 11^4 = 122^2, \qquad 9262^3 + 15312283^2 = 113^7.$$

Table 1. Known (and conjecturally only) primitive solutions to $x^p + y^q = z^r$ with $p^{-1} + q^{-1} + r^{-1} < 1$.

et al. [2010], and $(2p, 2p, 5)$ by Bennett [Bennett 2006]. Recently, Chen and Siksek [2009] have solved the generalized Fermat equation with signatures $(3, 3, p)$ for a set of prime exponents p having Dirichlet density 28219/44928. For an exhaustive survey, see the book of Cohen [2007, Chapter 14]. An older but still very useful survey is [Kraus 1999].

There is an abundance of solutions for generalized Fermat equations with signatures $(2, 3, n)$ [Edwards 2004; Cohen 2007, Chapter 14], and so this subfamily is particularly interesting. The condition $\chi > 1$ within this subfamily coincides with the condition $n \geq 7$. The cases $n = 7, 8, 9$ are solved respectively in [Poonen et al. 2007; Bruin 2003; 2005]. The case $n = 10$ appears to be the first hitherto unresolved case within this subfamily, and this of course corresponds to Equation (1).

In this section, we solve Equation (1) in coprime integers x, y, and z, thereby proving Theorem 1. Equation (1) does not define a curve in $\mathbb{P}^3$; however, using standard factorization arguments, we will reduce its resolution to the determination of K-rational points on a family of genus-2 curves where $K = \mathbb{Q}(\sqrt[3]{2})$. One of these curves has Jacobian Mordell–Weil rank 2, and two others have Jacobian Mordell–Weil rank 3. Classical Chabauty is inapplicable as these curves defy the bound $r \leq g - 1$. However, they do satisfy the weaker bound $r \leq d(g - 1)$, which is a necessary condition for the applicability of our method. In what follows, we sketch how we successfully applied the method of this paper to determine the K-rational points on these curves. We used Magma [Bosma et al. 1997] for all our calculations. It includes implementations by Nils Bruin and Michael Stoll of 2-descent on Jacobians of hyperelliptic curves over number fields; the algorithm is detailed in [Stoll 2001]. Magma also includes an implementation of Chabauty for genus-2 curves over $\mathbb{Q}$.

Case I (y is odd). From (1), we immediately see that

$$x + z^5 = u^3 \quad \text{and} \quad x - z^5 = v^3,$$

where u and v are coprime and odd. Hence, $2z^5 = u^3 - v^3$.

Case I.1 ($3 \nmid z$). Then

$$u - v = 2a^5 \quad \text{and} \quad u^2 + uv + v^2 = b^5,$$

where a and b are coprime integers with $z = ab$. We now use the identity

$$(u - v)^2 + 3(u + v)^2 = 4(u^2 + uv + v^2) \tag{32}$$

to obtain $4a^{10} + 3c^2 = 4b^5$, where $c = u + v$. Dividing by $4a^{10}$, we obtain a rational point $(X, Y) = (b/a^2, 3c/2a^5)$ on the genus-2 curve

$$C : Y^2 = 3(X^5 - 1).$$

Using Magma, we are able to show that the Jacobian of this genus-2 curve C has Mordell–Weil rank 0 and torsion subgroup isomorphic to $\mathbb{Z}/2\mathbb{Z}$. It is immediate that $C(\mathbb{Q}) = \{\infty, (1, 0)\}$.

Working backwards, we obtain the solutions $(x, y, z) = (0, 1, \pm 1)$ to (1).

Case I.2 ($3 \mid z$). Recall that $2z^5 = u^3 - v^3$ and u and v are odd and coprime. Thus,

$$u - v = 2 \cdot 3^4 a^5 \quad \text{and} \quad u^2 + uv + v^2 = 3b^5,$$

where $z = 3ab$. Now we use identity (32) to obtain $4 \cdot 3^8 a^{10} + 3c^2 = 12b^5$, where $c = u + v$. Hence, we obtain a rational point $(X, Y) = (b/a^2, c/2a^5)$ on the genus-2 curve

$$C : Y^2 = X^5 - 3^7.$$

Let J be the Jacobian of C. Using Magma, we can show that $J(\mathbb{Q})$ is free of rank 1 with generator

$$\left(\frac{-9 + 3\sqrt{-3}}{2}, \frac{81 + 27\sqrt{-3}}{2} \right) + \left(\frac{-9 - 3\sqrt{-3}}{2}, \frac{81 - 27\sqrt{-3}}{2} \right) - 2\infty.$$

Using Magma's built-in Chabauty command, we find that $C(\mathbb{Q}) = \{\infty\}$. Working backwards, we obtain $(x, y, z) = (\pm 1, -1, 0)$.

Case II (y is even). We would now like to solve (1) with y even and x and y coprime. Replacing x by $-x$ if necessary, we obtain $x \equiv z^5 \pmod 4$. Thus,

$$x + z^5 = 2u^3 \quad \text{and} \quad x - z^5 = 4v^3,$$

where $y = -2uv$. Hence,

$$u^3 - 2v^3 = z^5 \quad \text{with } u \text{ and } v \text{ coprime and } u \text{ and } z \text{ odd.} \tag{33}$$

If $3 \mid z$, then this equation is impossible modulo 9. Hence, $3 \nmid z$.

Let $\theta = \sqrt[3]{2}$. We shall work in the number field $K = \mathbb{Q}(\theta)$. This has ring of integers $\mathbb{O}_K = \mathbb{Z}[\theta]$ with class number 1. The unit group is isomorphic to $\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ with $\epsilon = 1 - \theta$ a fundamental unit.

Observe that
$$(u - v\theta)(u^2 + uv\theta + v^2\theta^2) = z^5,$$

where the two factors on the left-hand side are coprime as u and v are coprime and z is neither divisible by 2 nor 3. Hence,

$$u - v\theta = \epsilon^s \alpha^5 \quad \text{and} \quad u^2 + uv\theta + v^2\theta^2 = \epsilon^{-s}\beta^5, \tag{34}$$

where $-2 \le s \le 2$ and $\alpha, \beta \in \mathbb{Z}[\theta]$ satisfy $z = \alpha\beta$. We now use the identity

$$(u - v\theta)^2 + 3(u + v\theta)^2 = 4(u^2 + uv\theta + v^2\theta^2)$$

to obtain

$$\epsilon^{2s}\alpha^{10} + 3(u + v\theta)^2 = 4\epsilon^{-s}\beta^5.$$

Let C_s be the genus-2 curve defined over K given by

$$C_s : Y^2 = 3(4\epsilon^{-s}X^5 - \epsilon^{2s}).$$

We see that

$$(X, Y) = \left(\frac{\beta}{\alpha^2}, \frac{3(u + v\theta)}{\alpha^5}\right) \tag{35}$$

is a K-rational point on C_s. To complete our proof of Theorem 1, we need to determine $C_s(K)$ for $-2 \le s \le 2$. Let J_s be the Jacobian of C_s. Using reduction at various places of K, we easily showed that the torsion subgroup of $J_s(K)$ is trivial in all cases. The 2-Selmer ranks of $J_s(K)$ are respectively 1, 3, 2, 3, and 0 for $s = -2, -1, 0, 2, 1$. We searched for K-rational points on each J_s by first searching for points on the associated Kummer surface. We are fortunate to have found enough independent points in $J_s(K)$ in each case to show that the Mordell–Weil rank is equal to the 2-Selmer rank. In other words, we have determined a basis for a subgroup of $J_s(K)$ of finite index, and this is given in Table 2.

In each case, the rank r is at most $3 = d(g - 1)$, where $d = [K : \mathbb{Q}] = 3$ and $g = 2$ is the genus. We note that the bound $r \le g - 1$ needed to apply classical Chabauty fails for $s = -1, 0, 1$.

We implemented our method in Magma. Our program succeeded in determining $C_s(K)$ for all $-2 \le s \le 2$, and the results are given in Table 2. The entire computation took approximately 2.5 hours on a 2.8 GHz dual-core AMD Opteron; this includes the time taken for computing Selmer groups and searching for points on the Kummer surfaces. It is appropriate to give more details, and we do this for the case $s = 1$. Let $C = C_1$, and write J for its Jacobian. Let

$$\mathcal{H} = \{\infty, P_0, P_0', P_1, P_1'\},$$

s	basis for subgroup of $J_s(K)$ of finite index	$C_s(K)$
-2	$(\theta^2+\theta+1, \theta^2+2\theta+1)-\infty$	$\infty, (\theta^2+\theta+1, \pm(\theta^2+2\theta+1))$
-1	$(-\theta^2-\theta-1, 11\theta^2+13\theta+17)-\infty,$ $\sum_{i=1,2}(\Phi_i, (2\theta^2+2\theta+3)\Phi_i+2\theta^2+3\theta+4)-2\infty,$ $\sum_{i=3,4}(\Phi_i, (4\theta^2+6\theta+10)\Phi_i+9\theta^2+11\theta+13)-2\infty$ $\infty, (\frac{-\theta^2-2\theta-1}{3}, \frac{\pm(\theta^2-\theta+1)}{3}),(-\theta^2-\theta-1, \pm(11\theta^2+13\theta+17))$	
0	$(1, 3)-\infty, (\frac{\theta^2+2\theta+1}{3}, \frac{10\theta^2+8\theta+13}{3})-\infty$ $\infty, (\frac{\theta^2+2\theta+1}{3}, \frac{\pm(10\theta^2+8\theta+13)}{3}), (1, \pm 3)$	
1	$D_1 = (-\theta^2-\theta-1, -40\theta^2-53\theta-67)-\infty,$ $D_2 = (-1, 3\theta+3)-\infty,$ $D_3 = \sum_{i=5,6}(\Phi_i, (2\theta-2)\Phi_i-\theta+1)-2\infty$ $\infty, (-\theta^2-\theta-1, \pm(40\theta^2+53\theta+67)), (-1, \pm(3\theta+3))$	
2	$\varnothing$	∞

Table 2. Notation: Φ_1 and Φ_2 are the roots of $2\Phi^2+(\theta^2+\theta+2)\Phi+(\theta^2+\theta+2) = 0$; Φ_3 and Φ_4 are the roots of $3\Phi^2+(4\theta^2+5\theta+4)\Phi+(4\theta^2+5\theta+7) = 0$; Φ_5 and Φ_6 are the roots of $3\Phi^2+(\theta^2-\theta-2)\Phi+(-2\theta^2+2\theta+1) = 0$.

where

$$P_0 = (-\theta^2 - \theta - 1, 40\theta^2 + 53\theta + 67), \quad P_1 = (-1, 3\theta + 3),$$

and P_0' and P_1' are respectively the images of P_0 and P_1 under the hyperelliptic involution. Let D_1, D_2, D_3 be the basis given in Table 2 for a subgroup of $J(K)$ of finite index. Let $L_0 = \langle D_1, D_2, D_3 \rangle$. Our program verified that the index of L_0 in $J(K)$ is not divisible by any prime less than 75. Our program used the point

$$P_0 = (-\theta^2 - \theta - 1, 40\theta^2 + 53\theta + 67)$$

as the base point for the Abel–Jacobi map $\jmath$. The image of $\mathcal{H}$ under $\jmath$ is

$$\jmath(\mathcal{H}) = \{D_1, 0, 2D_1, D_1 + D_2, D_1 - D_2\},$$

where we have listed the elements of $\jmath(\mathcal{H})$ so that they correspond to the above list of points of $\mathcal{H}$. Next our program applied the Mordell–Weil sieve as in Lemma 5.1. The program chose twenty-two places υ that are places of good reduction for C

with $\#J(k_\upsilon)$ divisible only by primes less than 75. In the notation of Lemma 5.1,

$$L_{22} = \langle 1386000D_1 + 16632000D_2 + 18018000D_3, 24948000D_2, 24948000D_3 \rangle$$

and

$$W_{22} = \{0, D_1 - D_2, D_1, D_1 + D_2, 2D_1, D_1 + 12474000D_2 + 87318000D_3,$$
$$277201D_1 + 5821200D_2 + 51004800D_3, 277201D_1 - 6652800D_2 - 36313200D_3,$$
$$- 277199D_1 + 6652800D_2 + 36313200D_3,$$
$$- 277199D_1 - 5821200D_2 - 51004800D_3\}.$$

Next we would like to apply Theorem 3, and so we need primes p satisfying conditions (a)–(d) of that theorem. In particular, our program searches for odd primes p, unramified in K, so that every place $\upsilon \mid p$ is a place of good reduction for C, $\#J(k_\upsilon)$ is divisible only by primes less than 75, and L_{22} is contained in the kernel of the homomorphism (30). The smallest prime satisfying these conditions is $p = 109$, which splits completely in K, and so there are three degree-1 places υ_1, υ_2, and υ_3 above 109. It turns out that

$$J(k_\upsilon) \cong (\mathbb{Z}/110)^2$$

for $\upsilon = \upsilon_1, \upsilon_2, \upsilon_3$. The reader can easily see that

$$L_{22} \subset 110L_0 \subseteq 110J(K),$$

and so clearly L_{22} is in the kernel of (30) with $p = 109$. Moreover, the reader will easily see that every $\boldsymbol{w} \in W_{22}$ is equivalent modulo $110L_0$ to some element of $J(\mathcal{H})$. Hence, conditions (a), (c), and (d) of Theorem 3 are satisfied for each $\boldsymbol{w} \in W_{22}$ with $p = 109$. To show that $C(K) = \mathcal{H}$, it is enough to show that $\tilde{M}_{109}(Q)$ has rank 3 for all $Q \in \mathcal{H}$.

It is convenient to take

$$\omega_1 = \frac{dx}{y} \quad \text{and} \quad \omega_2 = \frac{x\,dx}{y}$$

as basis for the 1-forms on C. With this choice, we computed the matrices $\tilde{M}_{109}(Q)$ for $Q \in \mathcal{H}$. For example, we obtained

$$\tilde{M}_{109}(\infty) = \begin{pmatrix} 79 & 64 & 0 \\ 31 & 0 & 0 \\ 104 & 0 & 82 \end{pmatrix} \pmod{109};$$

this matrix of course depends on our choice of U used to compute the HNF on page 776 though, as observed in the remarks after Theorem 2, its rank is independent of this choice of U. The matrix $\tilde{M}_{109}(\infty)$ clearly has nonzero determinant and so

rank 3. It turns out that the four other $\tilde{M}_{109}(Q)$ also have rank 3. This completes the proof that $C(K) = \mathcal{H}$.

We now return to the general case where $-2 \leq s \leq 2$ and would like to recover the coprime integer solutions u and v to Equation (33) from the K-rational points on C_s and hence the solutions (x, y, z) to (1) with y even and $x \equiv z^5 \pmod 4$. From (35) and (34), we see that

$$Y = \frac{3(u + v\theta)}{\alpha^5} = 3\epsilon^s \left(\frac{u + v\theta}{u - v\theta} \right).$$

Thus,

$$\frac{u}{v} = \theta \cdot \left(\frac{Y + 3\epsilon^s}{Y - 3\epsilon^s} \right).$$

Substituting in here the values of Y and s from the K-rational points on the curves C_s, the only $\mathbb{Q}$-*rational* values for u/v we obtain are $-1, 2, 0, 5/4$, and 1; these come from the points $(\theta^2 + \theta + 1, -\theta^2 - 2\theta - 1)$, $(-\theta^2 - \theta - 1, -11\theta^2 - 13\theta - 17)$, $(1, -3)$, $(-\theta^2 - \theta - 1, 40\theta^2 + 53\theta + 67)$, and $(-1, 3\theta + 3)$, respectively. This immediately allows us to complete the proof of Theorem 1.

The reader can find the Magma code for verifying the above computations at http://www.warwick.ac.uk/staff/S.Siksek/progs/chabnf/.

Remarks. (i) Although our approach solves Equation (1) completely, we point out that it is possible to eliminate some cases by using Galois representations and level-lowering as Dahmen [2008] does for the equation $x^2 + z^{10} = y^3$. Indeed, by mimicking Dahmen's approach and making use of the work of Darmon and Merel [1997] and the so called "method for predicting the exponents of constants" [Cohen 2007, Section 15.7], we were able to reduce to the case $s = 1$, and it is this case that corresponds to our nontrivial solution $(x, y, z) = (\pm 3, -2, \pm 1)$. It seems however that the approach via Galois representations cannot in the current state of knowledge deal with case $s = 1$.

(ii) Note that to solve our original problem (1), we did not need all K-rational points on the curves C_s, merely those $(X, Y) \in C_s(K)$ with

$$\theta \cdot \left(\frac{Y + 3\epsilon^s}{Y - 3\epsilon^s} \right) \in \mathbb{Q}.$$

Mourao [2013] has recently developed a higher-dimensional analogue of elliptic curve Chabauty that is applicable in such situations, and this may provide an alternative approach to (1).

Acknowledgments

References

[Bennett 2006] M. A. Bennett, "The equation $x^{2n} + y^{2n} = z^5$", *J. Théor. Nombres Bordeaux* **18**:2 (2006), 315–321. MR 2007i:11048 Zbl 1138.11009

[Bennett et al. 2010] M. A. Bennett, J. S. Ellenberg, and N. C. Ng, "The Diophantine equation $A^4 + 2^\delta B^2 = C^n$", *Int. J. Number Theory* **6**:2 (2010), 311–338. MR 2011k:11045 Zbl 1218.11035

[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 1484478 Zbl 0898.68039

[Bourbaki 1989] N. Bourbaki, *Elements of Mathematics: Lie groups and Lie algebras, Chapters 1–3*, Springer, Berlin, 1989. MR 89k:17001 Zbl 0672.22001

[Brown 2012] D. Brown, "Primitive integral solutions to $x^2 + y^3 = z^{10}$", *Int. Math. Res. Not.* **2012**:2 (2012), 423–436. MR 2012k:11036 Zbl 06013326

[Bruin 2002] N. R. Bruin, *Chabauty methods and covering techniques applied to generalized Fermat equations*, Ph.D. thesis, University of Leiden, Amsterdam, 2002. MR 2003i:11042 Zbl 1043.11029

[Bruin 2003] N. Bruin, "Chabauty methods using elliptic curves", *J. Reine Angew. Math.* **562** (2003), 27–49. MR 2004j:11051 Zbl 1135.11320

[Bruin 2005] N. Bruin, "The primitive solutions to $x^3 + y^9 = z^2$", *J. Number Theory* **111**:1 (2005), 179–189. MR 2006e:11040 Zbl 1081.11019

[Bruin and Elkies 2002] N. Bruin and N. D. Elkies, "Trinomials $ax^7 + bx + c$ and $ax^8 + bx + c$ with Galois groups of order 168 and $8 \cdot 168$", pp. 172–188 in *Algorithmic number theory* (Sydney, 2002), edited by C. Fieker and D. R. Kohel, Lecture Notes in Computer Science **2369**, Springer, Berlin, 2002. MR 2005d:11094 Zbl 1058.11044

[Bruin and Stoll 2008] N. Bruin and M. Stoll, "Deciding existence of rational points on curves: an experiment", *Experiment. Math.* **17**:2 (2008), 181–189. MR 2009d:11100 Zbl 1218.11065

[Bruin and Stoll 2009] N. Bruin and M. Stoll, "Two-cover descent on hyperelliptic curves", *Math. Comp.* **78**:268 (2009), 2347–2370. MR 2010e:11059 Zbl 1208.11078

[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 2011j:11118 Zbl 05947723

[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 2010b:11066 Zbl 1168.11026

[Cassels and Flynn 1996] J. W. S. Cassels and E. V. Flynn, *Prolegomena to a middlebrow arithmetic of curves of genus 2*, London Math. Soc. Lecture Note Ser. **230**, Cambridge University Press, 1996. MR 97i:11071 Zbl 0857.14018

[Chabauty 1941] C. Chabauty, "Sur les points rationnels des variétés algébriques dont i'irrégularité est supérieure à la dimension", *C. R. Acad. Sci. Paris* **212** (1941), 1022–1024. MR 6,102e Zbl 0025.24903

[Chen and Siksek 2009] I. Chen and S. Siksek, "Perfect powers expressible as sums of two cubes", *J. Algebra* **322**:3 (2009), 638–656. MR 2011d:11070 Zbl 1215.11026

[Cohen 2000] H. Cohen, *Advanced topics in computational number theory*, Graduate Texts in Mathematics **193**, Springer, New York, 2000. MR 2000k:11144 Zbl 0977.11056

[Cohen 2007] H. Cohen, *Number theory, II: Analytic and modern tools*, Graduate Texts in Mathematics **240**, Springer, New York, 2007. MR 2008e:11002 Zbl 1119.11002

[Coleman 1985a] R. F. Coleman, "Effective Chabauty", *Duke Math. J.* **52**:3 (1985), 765–770. MR 87f:11043 Zbl 0588.14015

[Coleman 1985b] R. F. Coleman, "Torsion points on curves and p-adic abelian integrals", *Ann. of Math.* (2) **121**:1 (1985), 111–168. MR 86j:14014 Zbl 0578.14038

[Colmez 1998] P. Colmez, *Intégration sur les variétés p-adiques*, Astérisque **248**, 1998. MR 2000e: 14026 Zbl 0930.14013

[Dahmen 2008] S. R. Dahmen, *Classical and modular methods applied to Diophantine equations*, Ph.D. thesis, Universiteit Utrecht, 2008, Available at http://igitur-archive.library.uu.nl/dissertations/ 2008-0820-200949/dahmen.p% df.

[Darmon 1997] H. Darmon, "Faltings plus epsilon, Wiles plus epsilon, and the generalized Fermat equation", *C. R. Math. Rep. Acad. Sci. Canada* **19**:1 (1997), 3–14. MR 98h:11034a Zbl 0932.11022

[Darmon and Granville 1995] H. Darmon and A. Granville, "On the equations $z^m = F(x, y)$ and $Ax^p + By^q = Cz^r$", *Bull. London Math. Soc.* **27**:6 (1995), 513–543. MR 96e:11042 Zbl 0838.11023

[Darmon and Merel 1997] H. Darmon and L. Merel, "Winding quotients and some variants of Fermat's last theorem", *J. Reine Angew. Math.* **490** (1997), 81–100. MR 98h:11076 Zbl 0976.11017

[Edwards 2004] J. Edwards, "A complete solution to $X^2 + Y^3 + Z^5 = 0$", *J. Reine Angew. Math.* **571** (2004), 213–236. MR 2005e:11035 Zbl 1208.11045

[Ellenberg 2004] J. S. Ellenberg, "Galois representations attached to $\mathbb{Q}$-curves and the generalized Fermat equation $A^4 + B^2 = C^p$", *Amer. J. Math.* **126**:4 (2004), 763–787. MR 2005g:11089 Zbl 1059.11041

[Faltings 1983] G. Faltings, "Endlichkeitssätze für abelsche Varietäten über Zahlkörpern", *Invent. Math.* **73**:3 (1983), 349–366. MR 85g:11026a Zbl 0588.14026

[Flynn 1994] E. V. Flynn, "Descent via isogeny in dimension 2", *Acta Arith.* **66**:1 (1994), 23–43. MR 95g:11057 Zbl 0835.14009

[Flynn 1995a] E. V. Flynn, "An explicit theory of heights", *Trans. Amer. Math. Soc.* **347**:8 (1995), 3003–3015. MR 95j:11052 Zbl 0864.11033

[Flynn 1995b] E. V. Flynn, "On a theorem of Coleman", *Manuscripta Math.* **88**:4 (1995), 447–456. MR 97b:11082 Zbl 0865.14012

[Flynn 1997] E. V. Flynn, "A flexible method for applying Chabauty's theorem", *Compositio Math.* **105**:1 (1997), 79–94. MR 97m:11083 Zbl 0882.14009

[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 98f:11066 Zbl 0895.11026

[Flynn and Wetherell 1999] E. V. Flynn and J. L. Wetherell, "Finding rational points on bielliptic genus 2 curves", *Manuscripta Math.* **100**:4 (1999), 519–533. MR 2001g:11098 Zbl 1029.11024

[Flynn and Wetherell 2001] E. V. Flynn and J. L. Wetherell, "Covering collections and a challenge problem of Serre", *Acta Arith.* **98**:2 (2001), 197–205. MR 2002b:11088 Zbl 1049.11066

[Grant 1994] D. Grant, "A curve for which Coleman's effective Chabauty bound is sharp", *Proc. Amer. Math. Soc.* **122**:1 (1994), 317–319. MR 94k:14019 Zbl 0834.14015

[Kraus 1999] A. Kraus, "On the equation $x^p + y^q = z^r$: a survey", *Ramanujan J.* **3**:3 (1999), 315–333. MR 2001f:11046 Zbl 0939.11016

[Lorenzini and Tucker 2002] D. Lorenzini and T. J. Tucker, "Thue equations and the method of Chabauty–Coleman", *Invent. Math.* **148**:1 (2002), 47–77. MR 2003d:11088 Zbl 1048.11023

[McCallum 1992] W. G. McCallum, "The arithmetic of Fermat curves", *Math. Ann.* **294**:3 (1992), 503–511. MR 93j:11037 Zbl 0766.14013

[McCallum 1994] W. G. McCallum, "On the method of Coleman and Chabauty", *Math. Ann.* **299**:3 (1994), 565–596. MR 95c:11079 Zbl 0824.14017

[McCallum and Poonen 2010] W. McCallum and B. Poonen, "The method of Chabauty and Coleman", preprint, 2010, Available at http://www-math.mit.edu/~poonen/papers/chabauty.pdf.

[Milne 1986] J. S. Milne, "Jacobian varieties", pp. 167–212 in *Arithmetic geometry* (Storrs, CT, 1984), edited by G. Cornell and J. H. Silverman, Springer, New York, 1986. MR 861976 Zbl 0604.14018

[Mourao 2013] M. Mourao, "Extending elliptic curve Chabauty to higher genus curves", *Manusc. Math.* (2013). arXiv 1111.5506

[Poonen and Schaefer 1997] B. Poonen and E. F. Schaefer, "Explicit descent for Jacobians of cyclic covers of the projective line", *J. Reine Angew. Math.* **488** (1997), 141–188. MR 98k:11087 Zbl 0888.11023

[Poonen et al. 2007] B. Poonen, E. F. Schaefer, and M. Stoll, "Twists of $X(7)$ and primitive solutions to $x^2 + y^3 = z^7$", *Duke Math. J.* **137**:1 (2007), 103–158. MR 2008i:11085 Zbl 1124.11019

[Schaefer 1995] E. F. Schaefer, "2-descent on the Jacobians of hyperelliptic curves", *J. Number Theory* **51**:2 (1995), 219–232. MR 96c:11066 Zbl 0832.14016

[Schaefer and Wetherell 2005] E. F. Schaefer and J. L. Wetherell, "Computing the Selmer group of an isogeny between abelian varieties using a further isogeny to a Jacobian", *J. Number Theory* **115**:1 (2005), 158–175. MR 2006g:11116 Zbl 1095.11033

[Siksek 1995a] S. Siksek, *Descents on curves of genus* 1, Ph.D. thesis, University of Exeter, 1995, Available at http://homepages.warwick.ac.uk/~maseap/papers/phdnew.pdf.

[Siksek 1995b] S. Siksek, "Infinite descent on elliptic curves", *Rocky Mountain J. Math.* **25**:4 (1995), 1501–1538. MR 97g:11053 Zbl 0852.11028

[Siksek 2009] S. Siksek, "Chabauty for symmetric powers of curves", *Algebra Number Theory* **3**:2 (2009), 209–236. MR 2010b:11069 Zbl 05566607

[Stoll 1998] M. Stoll, "On the arithmetic of the curves $y^2 = x^l + A$ and their Jacobians", *J. Reine Angew. Math.* **501** (1998), 171–189. MR 99h:11069 Zbl 0902.11024

[Stoll 1999] M. Stoll, "On the height constant for curves of genus two", *Acta Arith.* **90**:2 (1999), 183–201. MR 2000h:11069 Zbl 0932.11043

[Stoll 2001] M. Stoll, "Implementing 2-descent for Jacobians of hyperelliptic curves", *Acta Arith.* **98**:3 (2001), 245–277. MR 2002b:11089 Zbl 0972.11058

[Stoll 2002a] M. Stoll, "On the arithmetic of the curves $y^2 = x^l + A$, II", *J. Number Theory* **93**:2 (2002), 183–206. MR 2003d:11090 Zbl 1004.11038

[Stoll 2002b] M. Stoll, "On the height constant for curves of genus two, II", *Acta Arith.* **104**:2 (2002), 165–182. MR 2003f:11093 Zbl 1139.11318

[Stoll 2006a] M. Stoll, "Independence of rational points on twists of a given curve", *Compos. Math.* **142**:5 (2006), 1201–1214. MR 2007m:14025 Zbl 1128.11033

[Stoll 2006b] M. Stoll, "On the number of rational squares at fixed distance from a fifth power", *Acta Arith.* **125**:1 (2006), 79–88. MR 2007g:11039 Zbl 1162.11326

[Taylor and Wiles 1995] R. Taylor and A. Wiles, "Ring-theoretic properties of certain Hecke algebras", *Ann. of Math.* (2) **141**:3 (1995), 553–572. MR 96d:11072 Zbl 0823.11030

[Wetherell 1997] J. L. Wetherell, *Bounding the number of rational points on certain curves of high rank*, Ph.D. thesis, University of California, Berkeley, 1997, Available at www.williamstein.org/swc/notes/files/99WetherellThesis.pdf. MR 2696280

[Wiles 1995] A. Wiles, "Modular elliptic curves and Fermat's last theorem", *Ann. of Math.* (2) **141**:3 (1995), 443–551. MR 96d:11071 Zbl 0823.11029

s.siksek@warwick.ac.uk   *Department of Mathematics, University of Warwick,
                                Coventry, CV4 7AL, United Kingdom*
                                http://www.warwick.ac.uk/~maseap/

Moduli spaces for point modules on naïve blowups

Thomas A. Nevins and Susan J. Sierra

The *naïve blowup algebras* developed by Keeler, Rogalski, and Stafford, after examples of Rogalski, are the first known class of connected graded algebras that are noetherian but not strongly noetherian. This failure of the strong noetherian property is intimately related to the failure of the point modules over such algebras to behave well in families: puzzlingly, there is no fine moduli scheme for such modules although point modules correspond bijectively with the points of a projective variety X. We give a geometric structure to this bijection and prove that the variety X is a coarse moduli space for point modules. We also describe the natural moduli stack X_∞ for *embedded point modules* — an analog of a "Hilbert scheme of one point" — as an infinite blowup of X and establish good properties of X_∞. The natural map $X_\infty \to X$ is thus a kind of "Hilbert–Chow morphism of one point" for the naïve blowup algebra.

1. Introduction

One of the important achievements of noncommutative projective geometry is the classification of noncommutative projective planes, such as the three-dimensional Sklyanin algebra Skl_3, by Artin, Tate, and Van den Bergh [Artin et al. 1990]. More formally, these are *Artin–Schelter regular algebras* of dimension 3, noncommutative graded rings that are close analogs of a commutative polynomial ring in three variables; see [Stafford and Van den Bergh 2001] for a discussion. The key method of [Artin et al. 1990] is to study *point modules*, that is, cyclic graded modules with the Hilbert series of a point in projective space. Given a noncommutative projective plane R, the authors describe a moduli scheme for its point modules. This allows them to construct a homomorphism from R to a well understood ring, providing a first step in describing the structure of the noncommutative plane itself.

The techniques described above work in a more general context. Let $\Bbbk$ be an algebraically closed field; we assume $\Bbbk$ is uncountable although for some of the results quoted this hypothesis is unnecessary. A $\Bbbk$-algebra R is said to be *strongly noetherian* if, for any commutative noetherian $\Bbbk$-algebra C, the tensor

product $R \otimes_{\Bbbk} C$ is again noetherian. By a general result of Artin and Zhang [2001, Theorem E4.3], if R is a strongly noetherian $\mathbb{N}$-graded $\Bbbk$-algebra, then its point modules are parametrized by a projective scheme. Rogalski and Zhang [2008] used this result to extend the method of [Artin et al. 1990] to strongly noetherian connected graded $\Bbbk$-algebras that are generated in degree 1. (An $\mathbb{N}$-graded $\Bbbk$-algebra R is *connected graded* if $R_0 = \Bbbk$.) Their method constructs a map from the algebra to a twisted homogeneous coordinate ring (see Section 2 for definitions) on the scheme X parametrizing point modules. For example, Sklyanin algebras are strongly noetherian, and here X is an elliptic curve. The homomorphism here gives the well known embedding of an elliptic curve in a noncommutative $\mathbb{P}^2$.

Although it was believed for a time that all connected graded noetherian algebras would be strongly noetherian, Rogalski [2004] showed this was not the case. His example was generalized in joint work with Keeler and Stafford [Keeler et al. 2005; Rogalski and Stafford 2007] to give a geometric construction of a beautiful class of noncommutative graded algebras, known as *naïve blowups*, that are noetherian but not strongly noetherian. Along the way, they showed that point modules for naïve blowups — viewed as objects of noncommutative projective geometry in a way we make precise below — cannot behave well in families: there is no fine moduli scheme of finite type for such modules.

In the present paper, we systematically develop the moduli theory of point modules for the naïve blowups S of [Keeler et al. 2005; Rogalski and Stafford 2007]. Roughly speaking, we show that there is an analog of a "Hilbert scheme of one point on Proj(S)" that is an infinite blowup of a projective variety. This infinite blowup is quasicompact and noetherian as an *fpqc-algebraic* stack (a notion we make precise in Section 4). Furthermore, we show there is a *coarse* "moduli space for one point on Proj(S)" — it is, in fact, the projective variety from which the naïve blowup was constructed. These are the first descriptions in the literature of moduli structures for point modules on a naïve blowup.

More precisely, let X be a projective $\Bbbk$-variety of dimension at least 2, let σ be an automorphism of X, and let $\mathcal{L}$ be a σ-ample (see Section 2) invertible sheaf on X. We follow the standard convention that $\mathcal{L}^\sigma := \sigma^* \mathcal{L}$. Let $P \in X$ (in the body of the paper, we let P be any zero-dimensional subscheme of X), and assume that the σ-orbit of P is *critically dense*; that is, it is infinite and every infinite subset is Zariski dense. For $n \geq 0$, let

$$\mathcal{I}_n := \mathcal{I}_P \mathcal{I}_P^\sigma \cdots \mathcal{I}_P^{\sigma^{n-1}} \quad \text{and} \quad \mathcal{L}_n := \mathcal{L} \otimes \mathcal{L}^\sigma \otimes \cdots \otimes \mathcal{L}^{\sigma^{n-1}}.$$

Define $\mathcal{S}_n := \mathcal{I}_n \otimes \mathcal{L}_n$, and let

$$S := S(X, \mathcal{L}, \sigma, P) := \bigoplus_{n \geq 0} H^0(X, \mathcal{S}_n).$$

The algebra S is the naïve blowup associated to the data $(X, \mathcal{L}, \sigma, P)$.

If $\mathcal{L}$ is sufficiently ample, then S is generated in degree 1; alternatively, a sufficiently large Veronese of S is always generated in degree 1. We will assume throughout that S is generated in degree 1.

A *point module* is a graded cyclic S-module M with Hilbert series $1+t+t^2+\cdots$. We say M is an *embedded point module* if we are given, in addition, a surjection $S \to M$ of graded modules. Two embedded point modules M and M' are *isomorphic* if there is an S-module isomorphism from M to M' that intertwines the maps from S.

We begin by constructing a moduli stack for embedded point modules. Recall that X_∞ is a *fine moduli space* (or stack) for embedded point modules if there is an S-module quotient $S \otimes_\Bbbk \mathcal{O}_{X_\infty} \to M$ that is a universal family for point modules; that is, M is an X_∞-flat family of embedded S-point modules with the property that if $S \otimes_\Bbbk C \to M'$ is any C-flat family of embedded point modules for a commutative $\Bbbk$-algebra C, then there is a morphism $\mathrm{Spec}(C) \xrightarrow{f} X_\infty$ and an isomorphism $f^* M \cong M'$ of families of embedded S-point modules. Let X_n be the blowup of X at $\mathcal{I}_n$; there is an inverse system $\cdots \to X_n \to X_{n-1} \to \cdots \to X$ of schemes. Let $X_\infty := \varprojlim X_n$. This inverse limit exists as a stack. More precisely, in [Definition 4.1](), we introduce the notion of an *fpqc-algebraic* stack. We then have:

Theorem 1.1. *The inverse limit X_∞ is a noetherian fpqc-algebraic stack. The morphism $X_\infty \to X$ is quasicompact. Moreover, X_∞ is a fine moduli space for embedded S-point modules.*

We have been told that similar results were known long ago to M. Artin; however, they seem not to have been very widely known even among experts, nor do they seem to have appeared in the literature.

Note that the stack X_∞ is discrete: its points have no stabilizers. Thus, X_∞ is actually a $\Bbbk$-*space* in the terminology of [Laumon and Moret-Bailly 2000]; in particular, this justifies our use of the phrase "fine moduli space" in the statement of the theorem. However, X_∞ does not seem to have an étale cover by a scheme and hence does not have the right to be called an algebraic space.

We recall that, by definition, the noncommutative projective scheme associated to S is the quotient category Qgr-S = Gr-S/Tors-S of graded right S-modules by the full subcategory of locally bounded modules. A *point object* in Qgr-S is the image of (a shift of) a point module. If S is a commutative graded algebra generated in degree 1, Qgr-S is equivalent to the category of quasicoherent sheaves on $\mathrm{Proj}(S)$; this justifies thinking of Qgr-S as the noncommutative analog of a projective scheme.

If R is strongly noetherian and generated in degree 1, then a result of Artin and Stafford [Keeler et al. 2005, Theorem 10.2] shows that point objects of Qgr-R are parametrized by the same projective scheme X that parametrizes embedded point

modules. On the other hand, for naïve blowups $S = S(X, \mathcal{L}, \sigma, P)$ as above, we have:

Theorem 1.2 [Keeler et al. 2005, Theorem 1.1]. *The algebra S is noetherian but not strongly noetherian. Moreover*, *there is no fine moduli scheme of finite type over $\Bbbk$ parametrizing point objects of* Qgr-S.

By contrast, [Keeler et al. 2005] gives a simple classification (that fails in families), namely that point objects are in bijective correspondence with points of X: to a point $x \in X$, we associate the S-module $\bigoplus H^0(X, \Bbbk_x \otimes \mathcal{L}_n)$. In the present paper, we explain how these two facts about point objects of Qgr-S naturally fit together.

Assume that $\mathcal{L}$ is sufficiently ample (in the body of the paper, we work with any σ-ample $\mathcal{L}$ by considering shifts of point modules). Let F be the moduli functor of embedded point modules over S. Define an equivalence relation $\sim$ on $F(C)$ by saying that $M \sim N$ if their images are isomorphic in Qgr-$S \otimes_{\Bbbk} C$. We obtain a functor $G :$ Affine schemes $\to$ Sets by sheafifying (in the fpqc topology) the presheaf G^{pre} of sets defined by Spec $C \mapsto F(C)/\sim$.

A scheme Y is a *coarse moduli scheme* for point objects if it corepresents the functor G; that is, there is a natural transformation $G \to \mathrm{Hom}_{\Bbbk}(\,\cdot\,, Y)$ that is universal for natural transformations from G to schemes.

Our main result is the following:

Theorem 1.3. *The variety X is a coarse moduli scheme for point objects in* Qgr-S.

This gives a geometric structure to the bijection discovered by Keeler, Rogalski, and Stafford.

Corollary 1.4. *There is a fine moduli space X_∞ for embedded S-point modules but only a coarse moduli scheme X for point objects of* Qgr-S.

It may be helpful to compare the phenomenon described by Corollary 1.4 to a related, though quite different, commutative phenomenon. Namely, let Y be a smooth projective (commutative) surface. Fix $n \geq 1$. Let $R = \mathbb{C}[Y]$ denote a homogeneous coordinate ring of Y (associated to a sufficiently ample invertible sheaf on Y), and consider graded quotient modules $R \to M$ such that $\dim M_l = n$ for $l \gg 0$. By a general theorem of Serre, the moduli space for such quotients is the *Hilbert scheme of n points on Y*, denoted $\mathrm{Hilb}^n(Y)$. This is a smooth projective variety of dimension $2n$. Alternatively, remembering only the corresponding objects $[M]$ of Qgr-$R \simeq \mathrm{Qcoh}(Y)$ and imposing the further S-equivalence relation [Huybrechts and Lehn 1997, Example 4.3.6], we get the moduli space $\mathrm{Sym}^n(Y)$ for semistable length-n sheaves on Y, which equals the n-th symmetric product of Y. The latter moduli space is only a coarse moduli space for semistable sheaves. One has the Hilbert–Chow morphism $\mathrm{Hilb}^n(Y) \to \mathrm{Sym}^n(Y)$, which is defined by taking a quotient $R \to M$ to the equivalence class of M. It is perhaps helpful to view the

moduli spaces and map $X_\infty \to X$ associated to the algebra S in light of the theorems stated above, that is, as a kind of "noncommutative Hilbert–Chow morphism of one point" for a naïve blowup algebra $S(X, \mathcal{L}, \sigma, P)$.

In work in preparation, we generalize the results in [Rogalski and Zhang 2008] by proving a converse, of sorts, to Theorem 1.3. Namely, suppose R is a connected graded noetherian algebra generated in degree 1, that R has a fine moduli space X_∞ for embedded point modules, that R has a projective coarse moduli scheme X for point objects of Qgr-R, and that the spaces X_∞ and X and the morphism $X_\infty \to X$ between them have geometric properties similar to those of the spaces we encounter in the theorems above. Then, we show, there exist an automorphism σ of X, a zero-dimensional subscheme $P \subset X$ supported on points with critically dense orbits, an ample and σ-ample invertible sheaf $\mathcal{L}$ on X, and a homomorphism $\phi : R \to S(X, \mathcal{L}, \sigma, P)$ from R to the naïve blowup associated to this data; furthermore, ϕ is surjective in large degree. This construction gives a new tool for analyzing the structure of rings that are noetherian but not strongly noetherian. Details will appear in [Nevins and Sierra 2012].

2. Background

In this section, we give needed definitions and background. We begin by discussing *bimodule algebras*: this is the correct way to think of the sheaves $\mathcal{S}_n$ defined above. Most of the material in this section was developed in [Van den Bergh 1996; Artin and Van den Bergh 1990], and we refer the reader there for references. Our presentation follows that in [Keeler et al. 2005; Sierra 2011].

Convention 2.1. Throughout the paper, by *variety* (over $\Bbbk$), we mean an integral separated scheme of finite type over $\Bbbk$.

Throughout this section, let $\Bbbk$ be an algebraically closed field and let A denote an affine noetherian $\Bbbk$-scheme, which we think of as a base scheme.

Definition 2.2. Let X be a scheme of finite type over A. An $\mathcal{O}_X$-*bimodule* is a quasicoherent $\mathcal{O}_{X \times X}$-module $\mathcal{F}$ such that, for every coherent submodule $\mathcal{F}' \subseteq \mathcal{F}$, the projection maps $p_1, p_2 : \mathrm{Supp}\,\mathcal{F}' \to X$ are both finite morphisms. The left and right $\mathcal{O}_X$-module structures associated to an $\mathcal{O}_X$-bimodule $\mathcal{F}$ are defined respectively as $(p_1)_*\mathcal{F}$ and $(p_2)_*\mathcal{F}$. We make the notational convention that when we refer to an $\mathcal{O}_X$-bimodule simply as an $\mathcal{O}_X$-module, we are using the left-handed structure (for example, when we refer to the global sections or higher cohomology of an $\mathcal{O}_X$-bimodule). All $\mathcal{O}_X$-bimodules are assumed to be $\mathcal{O}_A$-symmetric.

There is a tensor product operation on the category of bimodules that has the expected properties [Van den Bergh 1996, Section 2].

All the bimodules that we consider will be constructed from bimodules of the following form:

Definition 2.3. Let X be a projective scheme over A, and let $\sigma, \tau \in \text{Aut}_A(X)$. Let (σ, τ) denote the map

$$X \to X \times_A X \quad \text{defined by} \quad x \mapsto (\sigma(x), \tau(x)).$$

If $\mathscr{F}$ is a quasicoherent sheaf on X, we define the $\mathbb{O}_X$-bimodule ${}_\sigma\mathscr{F}_\tau := (\sigma, \tau)_*\mathscr{F}$. If $\sigma = 1$ is the identity, we will often omit it; thus, we write $\mathscr{F}_\tau$ for ${}_1\mathscr{F}_\tau$ and $\mathscr{F}$ for the $\mathbb{O}_X$-bimodule ${}_1\mathscr{F}_1 = \Delta_*\mathscr{F}$, where $\Delta : X \to X \times_A X$ is the diagonal.

Definition 2.4. Let X be a projective scheme over A. An $\mathbb{O}_X$-*bimodule algebra*, or simply a *bimodule algebra*, $\mathscr{B}$ is an algebra object in the category of bimodules. That is, there are a unit map $1 : \mathbb{O}_X \to \mathscr{B}$ and a product map $\mu : \mathscr{B} \otimes \mathscr{B} \to \mathscr{B}$ that have the usual properties.

We follow [Keeler et al. 2005] and define the following:

Definition 2.5. Let X be a projective scheme over A, and let $\sigma \in \text{Aut}_A(X)$. A bimodule algebra $\mathscr{B}$ is a *graded $(\mathbb{O}_X, \sigma)$-bimodule algebra* if

(1) there are coherent sheaves $\mathscr{B}_n$ on X such that $\mathscr{B} = \bigoplus_{n \in \mathbb{Z}} {}_1(\mathscr{B}_n)_{\sigma^n}$,

(2) $\mathscr{B}_0 = \mathbb{O}_X$, and

(3) the multiplication map μ is given by $\mathbb{O}_X$-module maps $\mathscr{B}_n \otimes \mathscr{B}_m^{\sigma^n} \to \mathscr{B}_{n+m}$, satisfying the obvious associativity conditions.

Definition 2.6. Let X be a projective scheme over A, and let $\sigma \in \text{Aut}_A(X)$. Let $\mathscr{R} = \bigoplus_{n \in \mathbb{Z}} (\mathscr{R}_n)_{\sigma^n}$ be a graded $(\mathbb{O}_X, \sigma)$-bimodule algebra. A *right $\mathscr{R}$-module* $\mathscr{M}$ is a quasicoherent $\mathbb{O}_X$-module $\mathscr{M}$ together with a right $\mathbb{O}_X$-module map $\mu : \mathscr{M} \otimes \mathscr{R} \to \mathscr{M}$ satisfying the usual axioms. We say that $\mathscr{M}$ is *graded* if there is a direct sum decomposition $\mathscr{M} = \bigoplus_{n \in \mathbb{Z}} (\mathscr{M}_n)_{\sigma^n}$ with multiplication giving a family of $\mathbb{O}_X$-module maps $\mathscr{M}_n \otimes \mathscr{R}_m^{\sigma^n} \to \mathscr{M}_{n+m}$ obeying the appropriate axioms.

We say that $\mathscr{M}$ is *coherent* if there are a coherent $\mathbb{O}_X$-module $\mathscr{M}'$ and a surjective map $\mathscr{M}' \otimes \mathscr{R} \to \mathscr{M}$ of ungraded $\mathscr{R}$-modules. We make similar definitions for left $\mathscr{R}$-modules. The bimodule algebra $\mathscr{R}$ is *right (left) noetherian* if every right (left) ideal of $\mathscr{R}$ is coherent. A graded $(\mathbb{O}_X, \sigma)$-bimodule algebra is right (left) noetherian if and only if every graded right (left) ideal is coherent.

We recall here some standard notation for module categories over rings and bimodule algebras. Let C be a commutative ring, and let R be an $\mathbb{N}$-graded C-algebra. We define Gr-R to be the category of $\mathbb{Z}$-graded right R-modules; morphisms in Gr-R preserve degree. Let Tors-R be the full subcategory of modules that are direct limits of right bounded modules. This is a Serre subcategory of Gr-R, so we may form the *quotient category*

$$\text{Qgr-}R := \text{Gr-}R/\text{Tors-}R.$$

(We refer the reader to [Gabriel 1962] as a reference for the category theory used here.) There is a canonical quotient functor from Gr-R to Qgr-R.

We make similar definitions on the left. Further, throughout this paper, we adopt the convention that if Xyz is a category, then xyz is the full subcategory of noetherian objects. Thus, we have gr-R and qgr-R, R-qgr, etc. If X is a scheme, $\mathcal{O}_X$-Mod and $\mathcal{O}_X$-mod will denote the categories of quasicoherent and coherent sheaves on X, respectively.

Given a module $M \in$ gr-R, we define $M[n] := \bigoplus_{i \in \mathbb{Z}} M[n]_i$, where $M[n]_i = M_{n+i}$.

For a graded $(\mathcal{O}_X, \sigma)$-bimodule algebra $\mathcal{R}$, we likewise define Gr-$\mathcal{R}$ and gr-$\mathcal{R}$. The full subcategory Tors-$\mathcal{R}$ of Gr-$\mathcal{R}$ consists of direct limits of modules that are coherent as $\mathcal{O}_X$-modules, and we similarly define Qgr-$\mathcal{R}$:= Gr-$\mathcal{R}$/Tors-$\mathcal{R}$. We define qgr-$\mathcal{R}$ in the obvious way.

If $\mathcal{R}$ is an $\mathcal{O}_X$-bimodule algebra, its global sections $H^0(X, \mathcal{R})$ inherit an $\mathcal{O}_A$-algebra structure. We call $H^0(X, \mathcal{R})$ the *section algebra* of $\mathcal{R}$. If $\mathcal{R} = \bigoplus (\mathcal{R}_n)_{\sigma^n}$ is a graded $(\mathcal{O}_X, \sigma)$-bimodule algebra, then multiplication on $H^0(X, \mathcal{R})$ is induced from the maps

$$H^0(X, \mathcal{R}_n) \otimes_A H^0(X, \mathcal{R}_m) \xrightarrow{1 \otimes \sigma^n} H^0(X, \mathcal{R}_n) \otimes_A H^0(X, \mathcal{R}_m^{\sigma^n}) \xrightarrow{\mu} H^0(X, \mathcal{R}_{n+m}).$$

If $\mathcal{M}$ is a graded right $\mathcal{R}$-module, then $H^0(X, \mathcal{M}) = \bigoplus_{n \in \mathbb{Z}} H^0(X, \mathcal{M}_n)$ is a right $H^0(X, \mathcal{R})$-module in the obvious way; thus, $H^0(X, \cdot)$ is a functor from Gr-$\mathcal{R}$ to Gr-$H^0(X, \mathcal{R})$.

If $R = H^0(X, \mathcal{R})$ and M is a graded right R-module, define $M \otimes_R \mathcal{R}$ to be the sheaf associated to the presheaf $V \mapsto M \otimes_R \mathcal{R}(V)$. This is a graded right $\mathcal{R}$-module, and the functor $\cdot \otimes_R \mathcal{R} :$ Gr-$R \to$ Gr-$\mathcal{R}$ is a right adjoint to $H^0(X, \cdot)$.

The following is a relative version of a standard definition:

Definition 2.7. Let A be an affine $\Bbbk$-scheme, and let $q : X \to A$ be a projective morphism. Let $\sigma \in \mathrm{Aut}_A(X)$, and let $\{\mathcal{R}_n\}_{n \in \mathbb{N}}$ be a sequence of coherent sheaves on X. The sequence of bimodules $\{(\mathcal{R}_n)_{\sigma^n}\}_{n \in \mathbb{N}}$ is *right ample* if, for any coherent $\mathcal{O}_X$-module $\mathcal{F}$, the following properties hold:

(1) $\mathcal{F} \otimes \mathcal{R}_n$ is globally generated for $n \gg 0$ (the natural map $q^* q_*(\mathcal{F} \otimes \mathcal{R}_n) \to \mathcal{F} \otimes \mathcal{R}_n$ is surjective for $n \gg 0$) and

(2) $R^i q_*(\mathcal{F} \otimes \mathcal{R}_n) = 0$ for $n \gg 0$ and $i \geq 1$.

The sequence $\{(\mathcal{R}_n)_{\sigma^n}\}_{n \in \mathbb{N}}$ is *left ample* if, for any coherent $\mathcal{O}_X$-module $\mathcal{F}$, the following properties hold:

(1) the natural map $q^* q_*(\mathcal{R}_n \otimes \mathcal{F}^{\sigma^n}) \to \mathcal{R}_n \otimes \mathcal{F}^{\sigma^n}$ is surjective for $n \gg 0$ and

(2) $R^i q_*(\mathcal{R}_n \otimes \mathcal{F}^{\sigma^n}) = 0$ for $n \gg 0$ and $i \geq 1$.

If $A = \Bbbk$, we say that an invertible sheaf $\mathscr{L}$ on X is σ-*ample* if the $\mathbb{O}_X$-bimodules

$$\{(\mathscr{L}_n)_{\sigma^n}\}_{n \in \mathbb{N}} = \{\mathscr{L}_\sigma^{\otimes n}\}_{n \in \mathbb{N}}$$

form a right ample sequence. By [Keeler 2000, Theorem 1.2], this is true if and only if the $\mathbb{O}_X$-bimodules $\{(\mathscr{L}_n)_{\sigma^n}\}_{n \in \mathbb{N}}$ form a left ample sequence.

The following result is a special case of a result due to Van den Bergh [1996, Theorem 5.2] although we follow the presentation of [Keeler et al. 2005, Theorem 2.12]:

Theorem 2.8 (Van den Bergh). *Let X be a projective $\Bbbk$-scheme, and let σ be an automorphism of X. Let $\mathscr{R} = \bigoplus (\mathscr{R}_n)_{\sigma^n}$ be a right noetherian graded $(\mathbb{O}_X, \sigma)$-bimodule algebra such that the bimodules $\{(\mathscr{R}_n)_{\sigma^n}\}$ form a right ample sequence. Then $R = H^0(X, \mathscr{R})$ is also right noetherian, and the functors $H^0(X, \cdot)$ and $\cdot \otimes_R \mathscr{R}$ induce an equivalence of categories* qgr-$\mathscr{R} \simeq$ qgr-R.

Castelnuovo–Mumford regularity is a useful tool for measuring ampleness and studying ample sequences. We will need to use relative Castelnuovo–Mumford regularity; we review the relevant background here. In the next three results, let X be a projective $\Bbbk$-scheme, and let A be a noetherian $\Bbbk$-scheme. Let $X_A := X \times A$, and let $p : X_A \to X$ and $q : X_A \to A$ be the projection maps.

Fix a very ample invertible sheaf $\mathbb{O}_X(1)$ on X. Let $\mathbb{O}_{X_A}(1) := p^* \mathbb{O}_X(1)$; note $\mathbb{O}_{X_A}(1)$ is relatively ample for $q : X_A \to A$. If $\mathscr{F}$ is a coherent sheaf on X_A and $n \in \mathbb{Z}$, let $\mathscr{F}(n) := \mathscr{F} \otimes_{X_A} \mathbb{O}_{X_A}(1)^{\otimes n}$. We say $\mathscr{F}$ is *m-regular with respect to $\mathbb{O}_{X_A}(1)$*, or just *m-regular*, if $R^i q_* \mathscr{F}(m - i) = 0$ for all $i > 0$. Since $\mathbb{O}_{X_A}(1)$ is relatively ample, $\mathscr{F}$ is m-regular for some m. The *regularity* of $\mathscr{F}$ is the minimal m for which $\mathscr{F}$ is m-regular; we write it reg$(\mathscr{F})$.

Castelnuovo–Mumford regularity is usually defined only for $\Bbbk$-schemes, so we will spend a bit of space on the technicalities of working over a more general base. First note:

Lemma 2.9. *Let $\mathscr{F}$ be a coherent sheaf on X. Then* reg$(\mathscr{F}) =$ reg$(p^* \mathscr{F})$. $\qquad\square$

The fundamental result on Castelnuovo–Mumford regularity is due to Mumford.

Theorem 2.10 [Lazarsfeld 2004, Example 1.8.24]. *Let $\mathscr{F}$ be an m-regular coherent sheaf on X_A. Then for every $n \geq 0$,*

(1) *$\mathscr{F}$ is $(m + n)$-regular;*

(2) *$\mathscr{F}(m + n)$ is generated by its global sections; that is, the natural map*

$$q^* q_* \mathscr{F}(m + n) \to \mathscr{F}(m + n)$$

is surjective;

(3) *the natural map $q_* \mathscr{F}(m) \otimes_A q_* \mathbb{O}_{X_A}(n) \to q_* \mathscr{F}(m + n)$ is surjective.*

Lemma 2.11. *For any* 0*-regular invertible sheaf* $\mathscr{H}$ *on* X_A *and any* A*-point* y *of* X, *the natural map* $q_*\mathscr{H} \xrightarrow{\alpha} q_*(\mathbb{O}_y \otimes_{X_A} \mathscr{H})$ *is surjective.*

Proof. This is standard, but we check the details. Since cohomology commutes with flat base change, it suffices to consider the case that $A = \operatorname{Spec} C$, where C is a local ring. Then for any $n \in \mathbb{Z}$, we may consider $\mathbb{O}_y \otimes_{X_A} \mathscr{H}(n)$ as an invertible sheaf on A. Since C is local, as a C-module, this is isomorphic to C.

We thus have $q_*(\mathbb{O}_y \otimes_{X_A} \mathscr{H}) \cong C$. Let $I := \operatorname{Im}(\alpha)$; this is an ideal of C.

Let $n \geq 0$, and consider the natural maps

$$
\begin{array}{ccc}
q_*\mathscr{H} \otimes_C q_*\mathbb{O}_{X_A}(n) & \xrightarrow{\quad\quad\quad\mu\quad\quad\quad} & q_*\mathscr{H}(n) \\
\Big\downarrow & \searrow^{\displaystyle f} & \Big\downarrow \\
I \otimes_C q_*\mathbb{O}_{X_A}(n) \xrightarrow{\ \alpha\otimes 1\ } q_*(\mathbb{O}_y \otimes_{X_A} \mathscr{H}) \otimes_C q_*\mathbb{O}_{X_A}(n) \longrightarrow & q_*(\mathbb{O}_y \otimes_{X_A} \mathscr{H}(n)) \cong C
\end{array}
\tag{2.12}
$$

This diagram clearly commutes, and $\alpha \otimes 1$ factors through $I \otimes_C q_*\mathbb{O}_{X_A}(n)$ by construction. Thus, $\operatorname{Im} f \subseteq I$ for all n.

On the other hand, by Theorem 2.10(3), μ is surjective. As $\mathbb{O}_{X_A}(1)$ is relatively ample, for $n \gg 0$, the right-hand vertical map is surjective. Thus, f is surjective for $n \gg 0$, and so $I = C$. $\qquad\square$

Let Z be a closed subscheme of X_A. We say that Z has *relative dimension* $\leq d$ if, for all $x \in A$, the fiber $q^{-1}(x)$ has dimension $\leq d$ as a $\Bbbk(x)$-scheme.

The following is a relative version of Proposition 2.7 of [Keeler 2010]:

Proposition 2.13. *Let* X *be a projective* $\Bbbk$*-scheme. There exists a constant* D, *depending only on* X *and on* $\mathbb{O}_X(1)$, *so that the following holds: for any noetherian* $\Bbbk$*-scheme* A *and for any coherent sheaves* $\mathscr{F}, \mathscr{G}$ *on* X_A *such that the closed subscheme of* X_A *where* $\mathscr{F}$ *and* $\mathscr{G}$ *both fail to be locally free has relative dimension* ≤ 2, *we have*

$$\operatorname{reg}(\mathscr{F} \otimes_{X_A} \mathscr{G}) \leq \operatorname{reg}(\mathscr{F}) + \operatorname{reg}(\mathscr{G}) + D.$$

Proof. The statement is local on the base, so we may assume without loss of generality that $A = \operatorname{Spec} C$ is affine. Since standard results such as Theorem 2.10 and Lemma 2.11 hold in this relative context, we may repeat the proof of [Keeler 2010, Proposition 2.7]. The relative dimension assumption ensures the vanishing of Rq_* that is needed in the proof. $\qquad\square$

To end the introduction, we define *naïve blowups*: these are the algebras and bimodule algebras that we will work with throughout the paper. Let X be a projective $\Bbbk$-variety. Let $\sigma \in \operatorname{Aut}_\Bbbk(X)$, and let $\mathscr{L}$ be a σ-ample invertible sheaf on X. Let P be a zero-dimensional subscheme of X. We define ideal sheaves

$$\mathscr{I}_n := \mathscr{I}_P \mathscr{I}_P^\sigma \cdots \mathscr{I}_P^{\sigma^{n-1}}$$

for $n \geq 0$. Then we define a bimodule algebra $\mathcal{S}(X, \mathcal{L}, \sigma, P) := \bigoplus_{n \geq 0}(\mathcal{S}_n)_{\sigma^n}$, where $\mathcal{S}_n := \mathcal{I}_n \mathcal{L}_n$. Define $S(X, \mathcal{L}, \sigma, P) := H^0(X, \mathcal{S}(X, \mathcal{L}, \sigma, P))$.

Theorem 2.14 [Rogalski and Stafford 2007, Theorems 1.2 and 3.1]. *Let X be a projective $\Bbbk$-variety with $\dim X \geq 2$. Let $\sigma \in \mathrm{Aut}_{\Bbbk}(X)$, and let $\mathcal{L}$ be a σ-ample invertible sheaf on X. Let P be a zero-dimensional subscheme of X, and let $\mathcal{S} := \mathcal{S}(X, \mathcal{L}, \sigma, P)$ and $S := S(X, \mathcal{L}, \sigma, P)$.*

If all points in P have critically dense σ-orbits, then the following hold:

(1) *The sequence of bimodules $\{(\mathcal{S}_n)_{\sigma^n}\}$ is a left and right ample sequence.*

(2) *S and $\mathcal{S}$ are left and right noetherian, the categories qgr-S and qgr-$\mathcal{S}$ are equivalent via the global sections functor. Likewise, S-qgr and $\mathcal{S}$-qgr are equivalent.*

(3) *The isomorphism classes of simple objects in qgr-$S \simeq$ qgr-$\mathcal{S}$ are in one-to-one correspondence with the closed points of X, where $x \in X$ corresponds to the $\mathcal{S}$-module $\bigoplus \Bbbk_x \otimes \mathcal{L}_n$. However, the simple objects in qgr-S are not parametrized by any scheme of finite type over $\Bbbk$.*

For technical reasons, we will want to assume that our naïve blowup algebra S is generated in degree 1. By [Rogalski and Stafford 2007, Propositions 3.18 and 3.19], this will always be true if we either replace S by a sufficiently large Veronese or replace $\mathcal{L}$ by a sufficiently ample line bundle (for example, if $\mathcal{L}$ is ample, by a sufficiently high tensor power of $\mathcal{L}$). If S is generated in degree 1, then by [Rogalski and Stafford 2007, Corollary 4.11], the simple objects in qgr-S are the images of shifts of point modules.

3. Blowing up arbitrary zero-dimensional schemes

For the rest of the paper, let $\Bbbk$ be an uncountable algebraically closed field. Let X be a projective variety over $\Bbbk$, let $\sigma \in \mathrm{Aut}_{\Bbbk}(X)$, and let $\mathcal{L}$ be a σ-ample invertible sheaf on X. Let P be a zero-dimensional subscheme of X supported at points with dense (later, critically dense) orbits. Let $\mathcal{S} := \mathcal{S}(X, \mathcal{L}, \sigma, P)$, and let $S := S(X, \mathcal{L}, \sigma, P)$. In this paper, we compare three objects: the scheme parametrizing length-n truncated point modules over S, the scheme parametrizing length-n truncated point modules over $\mathcal{S}$, and the blowup of X at the ideal sheaf $\mathcal{I}_n = \mathcal{I}_P \cdots \mathcal{I}_P^{\sigma^{n-1}}$. In this section, we focus on the blowup of X. We first give some general lemmas on blowing up the defining ideals of zero-dimensional schemes. These are elementary, but we give proofs for completeness.

Suppose that X is a variety and that $f : Y \to X$ is a surjective, projective morphism of schemes. Let η be the generic point of X. We define

$$Y^o := \overline{f^{-1}(\eta)}$$

and refer to Y^o, by abuse of terminology, as the *relevant component* of Y. In our situation, f will always be generically one-to-one and Y^o will be irreducible with $f|_{Y^o}$ birational onto its image.

Lemma 3.1. *Let A be a variety of dimension ≥ 2. Let $\mathcal{I}$ be the ideal sheaf of a zero-dimensional subscheme of A, and let $\pi : X \to A$ be the blowup of A at $\mathcal{I}$. Let W be the scheme parametrizing colength-1 ideals inside $\mathcal{I}$. Let $\phi : W \to A$ be the canonical morphism that sends an ideal $\mathcal{J}$ to the support of $\mathcal{I}/\mathcal{J}$. Then there is a closed immersion $c : X \to W$ that gives an isomorphism between X and W^o. Further, the following diagram commutes:*

$$
\begin{array}{ccc}
X & \xrightarrow{c} & W \\
& \pi \searrow \quad \swarrow \phi & \\
& A &
\end{array}
$$

Proof. Without loss of generality, $A = \operatorname{Spec} C$ is affine; let $I := \mathcal{I}(A)$. We may identify W with $\operatorname{Proj} \operatorname{Sym}_C(I)$ [Kleiman 1990, Proposition 2.2]; under this identification, $\phi : W \to A$ is induced by the inclusion $C \hookrightarrow \operatorname{Sym}_C(I)$. There is a canonical surjective map of graded C-algebras $\operatorname{Sym}_C(I) \to C \oplus \bigoplus_{n \geq 1} I^n$, which is the identity on C. This induces a closed immersion $c : X \to W$ with $\phi c = \pi$ as claimed. Further, both $\pi : X \to A$ and $\phi : W \to A$ are isomorphisms away from $\operatorname{Cosupp} \mathcal{I}$. Thus, c gives a birational closed immersion (and therefore an isomorphism) onto W^o. $\square$

Lemma 3.2. *Let A be a variety of dimension ≥ 2, and let $\mathcal{I}$ and $\mathcal{J}$ be ideal sheaves on A. Let $\mathcal{K} := \mathcal{I}\mathcal{J}$. Define $i : X \to A$ to be the blowup of A at $\mathcal{I}$, $j : Y \to A$ to be the blowup of A at $\mathcal{J}$, and $k : Z \to A$ to be the blowup of A at $\mathcal{K}$.*

(a) *There are morphisms $\xi : Z \to X$ and $\omega : Z \to Y$ so that the diagram*

$$
\begin{array}{ccc}
Z & \xrightarrow{\xi} & X \\
\omega \downarrow & \searrow k & \downarrow i \\
Y & \xrightarrow{j} & A
\end{array}
$$

commutes.

(b) *We have $Z \cong (X \times_A Y)^o$.*

(c) *Let W be the moduli scheme of subsheaves of $\mathcal{K}$ of colength 1, and let V be the moduli scheme of subsheaves of $\mathcal{I}$ of colength 1. Let $c : Z \to W$ and $d : X \to V$ be the maps from Lemma 3.1, and let $Z' := c(Z)$ and $X' := d(X)$. Then the map $\xi' : Z' \to X'$ induced from ξ sends $\mathcal{K}' \subset \mathcal{K}$ to $(\mathcal{K}' : \mathcal{J}) \cap \mathcal{I}$.*

Proof. (a) Since $\xi^{-1}(\mathcal{K})\mathcal{O}_Z = \xi^{-1}(\mathcal{I})\xi^{-1}(\mathcal{J})\mathcal{O}_Z$ is invertible, the inverse images of both $\mathcal{I}$ and $\mathcal{J}$ on Z are invertible. By the universal property of blowing up

[Hartshorne 1977, Proposition 7.14], the morphisms $\xi : Z \to X$ and $\omega : Z \to Y$ exist and commute as claimed.

For (b) and (c), we may without loss of generality assume $A = \operatorname{Spec} C$ is affine.

(b) Let $U := (X \times_A Y)^o$. Let $A' := A \smallsetminus \operatorname{Cosupp} \mathcal{K}$. Then U is the closure of A' in $\mathbb{P}_A^n \times_A \mathbb{P}_A^m$ for appropriate n and m.

Let $\phi : \mathbb{P}_A^n \times_A \mathbb{P}_A^m \to \Sigma_{n,m} \subset \mathbb{P}_A^{nm+m+n}$ be the Segre embedding. Note that the canonical embeddings $Z \subseteq W \subseteq \mathbb{P}_A^{nm+m+n}$ actually have $W \subseteq \Sigma_{n,m}$. Since $\phi' := \phi|_U$ is the identity over A' and $Z \subseteq \Sigma_{n,m}$ is the closure of A' in $\mathbb{P}_A^{nm+n+m}$, we have $\phi'(U) = Z$.
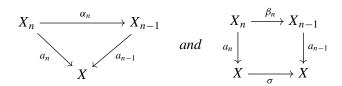
Let $p : X \times_A Y \to X$ and $q : X \times_A Y \to Y$ be the projection maps. From the commutative diagram in (a), we obtain a morphism $r : Z \to X \times_A Y$ with $qr = \omega$ and $pr = \xi$. Further, r restricts to $(\phi')^{-1}$ over A'. Thus, $r(Z) = U$, and $\phi' : U \to Z$ is an isomorphism.

(c) A point $(x, y) \in \mathbb{P}_A^n \times_A \mathbb{P}_A^m$ corresponds to a pair of linear ideals $\mathfrak{n} \subset C[x_0, \ldots, x_n]$ and $\mathfrak{m} \subset C[y_0, \ldots, y_m]$. Let $C[(x_i y_j)_{i,j}] \subset C[(x_i)_i][(y_j)_j]$ be the homogeneous coordinate ring of $\Sigma_{n,m}$. It is clear the ideal defining $\phi(x, y) = \{x\} \times \mathbb{P}^m \cap \mathbb{P}^n \times \{y\}$ in $C[x_i y_j]$ is generated by $\mathfrak{n}_1 \cdot (y_0, \ldots, y_m) + (x_0, \ldots, x_n) \cdot \mathfrak{m}_1$.

Let $(x, y) \in (X \times_A Y)^o$, where x corresponds to the colength-1 ideal $\mathscr{I}' \subseteq \mathscr{I}$ and y corresponds to $\mathscr{J}' \subseteq \mathscr{J}$. That $\mathscr{I}'\mathscr{J} + \mathscr{I}\mathscr{J}'$ gives the ideal $\mathcal{K}' \subset \mathcal{K}$ corresponding to $\phi(x, y)$ follows from the previous paragraph together with the fact that the isomorphism ϕ' between $(X \times_A Y)^o$ and Z is given by the Segre embedding.

Since ϕ' is an isomorphism, any ideal $\mathcal{K}'$ corresponding to a point $z \in Z$ may be written $\mathcal{K}' = \mathscr{I}'\mathscr{J} + \mathscr{I}\mathscr{J}'$ for appropriate $\mathscr{I}'$ and $\mathscr{J}'$. We thus have $\mathscr{I}' \subseteq (\mathcal{K}' : \mathscr{J}) \cap \mathscr{I} \subsetneqq \mathscr{I}$. Since $\mathscr{I}'$ is colength-1, this implies that $\mathscr{I}' = (\mathcal{K}' : \mathscr{J}) \cap \mathscr{I}$ as claimed. $\qquad \square$

Corollary 3.3. *Let X be a projective variety of dimension ≥ 2, let $\sigma \in \operatorname{Aut}_{\Bbbk}(X)$, and let $\mathscr{I}$ be an ideal sheaf on X. Let $\mathscr{I}_n := \mathscr{I}\mathscr{I}^\sigma \cdots \mathscr{I}^{\sigma^{n-1}}$. For all $n \geq 0$, let $a_n : X_n \to X$ be the blowup of X at $\mathscr{I}_n$. Then there are birational morphisms $\alpha_n : X_n \to X_{n-1}$ (for $n \geq 1$) and $\beta_n : X_n \to X_{n-1}$ (for $n \geq 2$) so that the diagrams*

$$
\begin{array}{ccc}
X_n & \xrightarrow{\ \alpha_n\ } & X_{n-1} \\
& \searrow{\scriptstyle a_n} \quad \swarrow{\scriptstyle a_{n-1}} & \\
& X &
\end{array}
\qquad and \qquad
\begin{array}{ccc}
X_n & \xrightarrow{\ \beta_n\ } & X_{n-1} \\
{\scriptstyle a_n}\downarrow & & \downarrow{\scriptstyle a_{n-1}} \\
X & \xrightarrow{\ \sigma\ } & X
\end{array}
$$

commute.

Proof. Let $\mathcal{K} := \mathscr{I}_n$, and let $\zeta : X'_{n-1} \to X$ be the blowup of X at $\mathscr{I}^\sigma_{n-1}$. Since $(\mathscr{I}_p)^\sigma \cong \mathscr{I}_{\sigma^{-1}(p)}$, there is an isomorphism $\theta : X'_{n-1} \to X_{n-1}$ so that the following
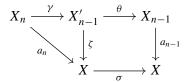
diagram commutes:

$$
\begin{array}{ccc}
X'_{n-1} & \xrightarrow{\theta} & X_{n-1} \\
\zeta \downarrow & & \downarrow a_{n-1} \\
X & \xrightarrow{\sigma} & X
\end{array}
$$

Apply Lemma 3.2(a) with $\mathscr{I} = \mathscr{I}^{\sigma}_{n-1}$ and $\mathscr{J} = \mathscr{I}_1$. We obtain a morphism $\gamma : X_n \to X'_{n-1}$ so that

$$
\begin{array}{ccccc}
X_n & \xrightarrow{\gamma} & X'_{n-1} & \xrightarrow{\theta} & X_{n-1} \\
& a_n \searrow & \downarrow \zeta & & \downarrow a_{n-1} \\
& & X & \xrightarrow{\sigma} & X
\end{array}
$$

commutes. Let $\beta_n := \theta\gamma : X_n \to X_{n-1}$.

Let α_n be the morphism $X_n \to X_{n-1}$ given by Lemma 3.2(a) with $\mathscr{I} = \mathscr{I}_{n-1}$ and $\mathscr{J} = \mathscr{I}^{\sigma^{n-1}}_1$. The diagram

$$
\begin{array}{ccc}
X_n & \xrightarrow{\alpha_n} & X_{n-1} \\
& a_n \searrow & \downarrow a_{n-1} \\
& & X
\end{array}
$$

commutes as required. $\qquad\square$

We will frequently suppress the subscripts on the maps α_n, a_n, etc., when the source and target are indicated. Note that the equation $a_n = \alpha_1 \circ \cdots \circ \alpha_n$ that follows from Corollary 3.3 may be written more compactly as $a = \alpha^n : X_n \to X$.

4. Infinite blowups

In this section, we prove some general properties of infinite blowups that will be useful when we consider moduli spaces of embedded point modules. Such infinite blowups can be handled in two ways: either as pro-objects in the category of schemes or as stacks via the (inverse) limits of such pro-objects in the category of spaces or of stacks. We've chosen to treat infinite blowups as the limits rather than as pro-objects. This is formally the correct choice in the sense that the limit formally contains less information than the pro-object. We note that in our setting, we could also work with the pro-object with no difficulties; however, we have found the language of stacks more natural.

We begin with some technical preliminaries on schemes and stacks. We will work with stacks in the fpqc (fidèlement plat et quasicompact) topology; the fpqc topology of schemes is discussed in [Vistoli 2005, Section 2.3.2]. We are interested

in a class of stacks that are apparently not algebraic but for which a certain amount of algebraic geometry is still possible. More precisely, recall that a stack $\mathscr{X}$ is called *algebraic* if the diagonal morphism of $\mathscr{X}$ is representable, separated, and quasicompact; and it has an fppf atlas $f : Z \to \mathscr{X}$ that is a scheme; that is, f is representable, faithfully flat, and finitely presented. By Artin's theorem [Laumon and Moret-Bailly 2000, Théorème 10.1], the second condition is equivalent to requiring the existence of a smooth, surjective, and representable f.

Our stacks are very similar to algebraic stacks, but it seems not to be possible to find a finite-type f for which Z is a scheme. On the other hand, we can find f for which Z is a scheme and f is fpqc — and even formally étale — so in some sense our stacks are the fpqc analogs of algebraic stacks.

Definition 4.1. We will refer to a stack $\mathscr{X}$ for which the diagonal $\Delta : \mathscr{X} \to \mathscr{X} \times_{\Bbbk} \mathscr{X}$ is representable, separated, and quasicompact, and which admits a representable fpqc morphism $Z \to \mathscr{X}$ from a scheme Z, as *fpqc-algebraic*.

Note that "separated" and "quasicompact" make sense for fpqc stacks by [EGA IV.2 1965, Proposition 2.7.1 and Corollaire 2.6.4]. Unfortunately, in this weaker setting, there are fewer notions of algebraic geometry that one can check fpqc-locally and hence fewer adjectives that one can sensibly apply to fpqc-algebraic stacks. Still, one can make sense, for example, of representable morphisms being separated, quasiseparated, locally of finite type or of finite presentation, proper, closed immersions, affine, etc., by [EGA IV.2 1965, Proposition 2.7.1].

Recall [EGA IV.4 1967, Définition 17.1.1] that a morphism of schemes $f : X \to Y$ is *formally étale* if, for every affine scheme Y', closed subscheme $Y'_0 \subset Y'$ defined by a nilpotent ideal, and morphism $Y' \to Y$, the map $\operatorname{Hom}_Y(Y', X) \to \operatorname{Hom}_Y(Y'_0, X)$ is bijective. By faithfully flat descent [Vistoli 2005], the definition extends immediately to stacks in the étale, fppf, and fpqc topologies of schemes.

We will say that an fpqc-algebraic stack $\mathscr{X}$ is *noetherian* if it admits an fpqc atlas $Z \to \mathscr{X}$ by a noetherian scheme Z. Unfortunately, since fpqc morphisms need not be of finite type even locally, it does not seem to be possible to check this property on an arbitrary atlas $Y \to \mathscr{X}$.

Suppose we have a sequence of schemes $\{X_n \mid n \in \mathbb{N}\}$ and projective morphisms $\pi_n : X_n \to X_{n-1}$. We define the infinite blowup X_∞ to be the presheaf of sets $X_\infty = \varprojlim X_n$: more precisely, let $h_{X_\infty} : \text{Schemes}^{\text{op}} \to \text{Sets}$ be the functor of points whose value on a scheme A is

$$h_{X_\infty}(A) = \{(\zeta_n : A \to X_n)_{n \in \mathbb{N}} \mid \pi_n \zeta_n = \zeta_{n-1}\}.$$

For each n, there is an induced map $\pi : X_\infty \to X_n$, where the target space X_n is indicated explicitly.

Proposition 4.2. *Suppose that $X := X_0$ is a variety of dimension ≥ 2 and there are maps $\pi_n : X_n \to X_{n-1}$ as above. Then the stack X_∞ is a sheaf in the fpqc topology. Further, suppose that the maps π_n satisfy the following conditions:*

(i) *For all n, π_n^{-1} is defined at all but finitely many points of X_{n-1}^o. That is, the set of exceptional points of $\pi^{-1} : X \dashrightarrow X_\infty$ is countable; let $\{z_m\}_{m \in \mathbb{N}}$ be an enumeration of this set.*

(ii) *The set $\{z_m\}$ is critically dense.*

(iii) *For all m, there is some $n(m)$ so that, for $n \geq n(m)$, the map π_n is a local isomorphism at all points in the preimage of z_m.*

(iv) *For all $m \in \mathbb{N}$, there is an ideal sheaf $\mathcal{J}_m$ on X, cosupported at z_m, so that $X_{n(m)}$ is a closed subscheme of $\operatorname{Proj} \mathcal{S}ym_X \mathcal{J}_m$ above a neighborhood of z_m. That is, $X_{n(m)} \to X$ factors as*

$$X_{n(m)} \xrightarrow{c_m} \operatorname{Proj} \mathcal{S}ym_X \mathcal{J}_m \xrightarrow{p_m} X,$$

where $p_m : \operatorname{Proj} \mathcal{S}ym_X \mathcal{J}_m \to X$ is the natural map and c_m is a closed immersion over a neighborhood of z_m.

(v) *There is some $D \in \mathbb{N}$ so that $\mathfrak{m}_{z_m}^D \mathcal{O}_{X,z_m} \subseteq \mathcal{J}_m \subseteq \mathcal{O}_{X,z_m}$ for every m.*

Then:

(1) *The stack X_∞ is fpqc-algebraic.*

   (1a) *X_∞ has a representable, formally étale, fpqc cover by an affine scheme $U \to X_\infty$.*

   (1b) *The diagonal morphism $\Delta : X_\infty \to X_\infty \times_\mathbb{k} X_\infty$ is representable, separated, and quasicompact.*

(2) *The morphism $\pi : X_\infty \to X$ is quasicompact.*

(3) *X_∞ is noetherian as an fpqc-algebraic stack.*

Proof. Because any limit of an inverse system of sheaves taken in the category of presheaves is already a sheaf [Hartshorne 1977, Exercise II.1.12], X_∞ is a sheaf in the fpqc topology.

Now assume that (i)–(v) hold. For $n \in \mathbb{N}$, let W_n be the scheme-theoretic image of c_n. Let

$$X_n' := W_0 \times_X W_1 \times_X \cdots \times_X W_{n-1}.$$

Let $\pi_n' : X_n' \to X_{n-1}'$ be projection on the first $n-1$ factors. We show we can assume without loss of generality that $X_n = X_n'$; that is, we claim that

$$\varprojlim_{\pi'} X_n' \cong \varprojlim_{\pi} X_n.$$

Let $k \in \mathbb{N}$. Let $K(k) := \max\{k, n(0), \ldots, n(k-1)\}$. For each $0 \le m \le k-1$, there is a morphism

$$X_{K(k)} \xrightarrow{\pi^{K(k)-n(m)}} X_{n(m)} \xrightarrow{c_m} W_m.$$

Since these agree on the base, we obtain an induced $\phi_k : X_{K(k)} \to X'_k$. The ϕ_k are clearly compatible with the inverse systems π and π'. Taking the limit, we obtain

$$\phi : \varprojlim X_{K(k)} \to \varprojlim X'_k.$$

Now let $N(k) := k + \max\{m \mid z_m \in F_k\}$, where F_k is the set of fundamental points of $X \dashrightarrow X_k$. We claim there is a morphism $\psi_k : X'_{N(k)} \to X_k$. There is certainly a rational map defined over $X \smallsetminus \{F_k\}$ since there X_k is locally isomorphic to X. Let $z_m \in F_k$, and let $n'(m) := \max\{k, n(m)\}$. The rational map

$$X'_{N(k)} \to W_m \dashrightarrow X_{n'(m)} \to X_k$$

is then defined over a neighborhood of z_m. These maps clearly agree on overlaps, so we may glue to define ψ_k as claimed. Let

$$\psi : \varprojlim X'_{N(k)} \to \varprojlim X_k$$

be the limit of the ψ_k. It is clear that $\psi = \phi^{-1}$; note that by construction both $N(k)$ and $K(k)$ go to infinity as $k \to \infty$.

Going forward, we replace X_n by X'_n. Thus, let $Y_n := \operatorname{Proj} \mathcal{S}ym_X \mathcal{J}_n$, and assume that there are closed immersions $i_n : X_n \to Y_0 \times_X \cdots \times_X Y_{n-1}$ so that the π_n are given by restricting the projection maps.

It suffices to prove the proposition in the case that $X = \operatorname{Spec} C$ is affine; note that we can choose an affine subset of X that contains all z_n. Let $J_n \subseteq C$ be the ideal cosupported at z_n so that $(J_n)_{z_n} = \mathcal{J}_n$. Let $\mathfrak{m}_p$ denote the maximal ideal of C corresponding to p.

Claim 4.3. *There is an N such that every ideal J_m, $m \in \mathbb{N}$, is generated by at most N elements.*

Proof. Embed X in an affine space; i.e., choose a closed immersion $X \subseteq \mathbb{A}^l$. Then each point of $\mathbb{A}^l$, hence a fortiori each point of X, is cut out scheme-theoretically by l elements of C, and the power of the maximal ideal $\mathfrak{m}^D_{z_m}$ appearing in hypothesis (v) of the proposition is generated by $N_0 := \binom{D+l-1}{l-1}$ elements of C. Now J_m contains $\mathfrak{m}^D_{z_m}$, and

$$\dim(J_m/\mathfrak{m}^D_{z_m}) \le \dim C/\mathfrak{m}^D_{z_m} \le \dim \Bbbk[u_1, \ldots, u_l]/(u_1, \ldots, u_l)^D =: N_1.$$

Thus, J_m is generated by at most $N := N_0 + N_1$ elements. $\qquad\square$

We continue with the proof of the proposition.

(1a) To construct an affine scheme U with a representable, formally étale morphism $U \to X_\infty$, we proceed as follows. For each n and for each $1 \le i \le N$, we choose hypersurfaces $D_{n,i} = V(d_{n,i}) \subset X$ with the following properties:

(A) For all n, the elements $d_{n,1}, \ldots, d_{n,N}$ generate J_n.

(B) For all n and i, the hypersurface $D_{n,i}$ does not contain any irreducible component Z of a hypersurface $D_{m,j}$ with $m < n$ or $m = n$ and $j < i$.

(C) For all n and each $m \ne n$, $z_m \notin D_{n,i}$ for any i.

We can make such choices because $\Bbbk$ is uncountable and X is affine (note that in order to satisfy (B), the choice of each $D_{n,i}$ will depend on finitely many earlier choices).

Let $Z_{(n_1,\ldots,n_N)} := D_{n_1,1} \cap \cdots \cap D_{n_N,N}$ for each N-tuple of positive integers $(n_1, \ldots, n_N)$. Note that $z_m \notin Z_{(n_1,\ldots,n_N)}$ unless $(n_1, \ldots, n_N) = (m, m, \ldots, m)$ by property (C). Note also that, since $Z_{(n_1,\ldots,n_N)}$ is a union of intersections of pairwise distinct irreducible hypersurfaces, it has codimension at least 2 in X.

Now let $Z^{(1)} := \bigcup_{(n_1,\ldots,n_N)} Z_{(n_1,\ldots,n_N)}$. This is a countable union of irreducible subsets of X of codimension at least 2. We may choose one point lying on each component of $Z^{(1)} \smallsetminus \{z_m\}_{m\in\mathbb{N}}$. Now, for each n, choose a hypersurface $D_{n,N+1}$ such that $z_n \in D_{n,N+1}$, the local ideal of $D_{n,N+1}$ at z_n is contained in $\mathcal{J}_n$, and $D_{n,N+1}$ avoids all the (countably many) chosen points of components of $Z^{(1)}$ and all z_m, $m \ne n$. Then for each n, $Z^{(1)} \cap D_{n,N+1}$ is a countable union of irreducible algebraic subsets of codimension at least 3 (it is a union of proper intersections of $D_{n,N+1}$ with irreducible subsets of codimension at least 2). Let $Z^{(2)} := \bigcup_n (Z^{(1)} \cap D_{n,N+1})$.

Repeating the previous construction with $Z^{(2)}$, we get hypersurfaces $D_{n,N+2}$ such that each $Z^{(2)} \cap D_{n,N+2}$ is a countable union of irreducible subsets of codimension at least 4. Iterating, we eventually define hypersurfaces $D_{n,N+i}$, $i = 1, \ldots, d$, with the following properties:

(A′) For all $m \in \mathbb{N}$, a scheme-theoretic equality $\mathrm{Spec}(C/J_m) = D_{m,1} \cap \cdots \cap D_{m,N+d}$ exists.

(B′) For every sequence $(n_1, \ldots, n_{N+d})$ of positive integers, we have a set-theoretic equality

$$D_{n_1,1} \cap \cdots \cap D_{n_{N+d},N+d} = \begin{cases} z_m & \text{if } (n_1, \ldots, n_{N+d}) = (m, \ldots, m), \\ \varnothing & \text{otherwise.} \end{cases}$$

(C) For all n and each $m \ne n$, $z_m \notin D_{n,i}$ for any i.

For $0 \le n \le m - 1$, we abusively let $\tilde{D}_{n,i} \subset X_m^o$ denote the proper transform of $D_{n,i}$. By construction, $\tilde{D}_{n,1} \cap \cdots \cap \tilde{D}_{n,N+d} = \varnothing$.

For each $m \in \mathbb{N}$, the map $C^{N+d} \to J_m$, $e_i \mapsto d_{m,i}$ induces a closed immersion $Y_m \to \mathbb{P}_X^{N+d-1}$. Let $V_{m,i} \subseteq Y_m$ be the open affine given by $e_i \neq 0$. Note that $V_{m,i} \cap X_m^o = X_m^o \smallsetminus \tilde{D}_{m,i}$ (recall X_m^o here denotes the closure in X_m of the preimage in X_m of the generic point of X). Let

$$U_{n,i} := X_n \cap (V_{0,i} \times_X V_{1,i} \times_X \cdots \times_X V_{n-1,i}.)$$

The $U_{n,i}$ are open and affine. Since $D_{m,i} \not\ni z_n$ for $m \neq n$, the set $\bigcup_i U_{n,i}$ includes all irreducible components of X_n except possibly for X_n^o. But

$$X_n^o \smallsetminus \bigcup_{i=1}^{N+d} U_{n,i} = \bigcap_i \bigcup_{m=0}^{n-1} \tilde{D}_{m,i} = \bigcup_m \bigcap_i \tilde{D}_{m,i} = \varnothing.$$

Thus, the $U_{n,i}$ are an open affine cover of X_n.

Since $\pi_n|_{U_{n,i}}$ is obtained by base extension from the affine morphism $V_{n-1,i} \to X$, it is affine, and $\pi_n(U_{n,i}) \subseteq U_{n-1,i}$. Writing $C_i := \varinjlim C_{m,i}$ and $U_i := \operatorname{Spec} C_i$, we get $U_i = \varprojlim U_{m,i}$; all the U_i are affine schemes. By construction, we obtain induced maps $U_i \to X_\infty$. Let $U := \bigsqcup_i U_i$.

Claim 4.4. *The induced morphism $U \to X_\infty$ is representable and formally étale.*

Proof. Since each map $U_i \to X_\infty$ is a limit of formally étale morphisms, each is itself formally étale. We must show that if T is a scheme equipped with a morphism $T \to X_\infty$, then $T \times_{X_\infty} U \to T$ is a scheme over T. Each morphism $T \times_{X_m} U_{m,i} \to T$ is an affine open immersion since the morphisms $U_{m,i} \to X_m$ are affine open immersions. Hence, the morphism $\varprojlim(T \times_{X_m} U_{m,i}) \to T$ is an inverse limit of schemes affine over T and thus is itself a scheme affine over T [EGA IV.3 1966, Proposition 8.2.3]. The claim now follows from this result:

Lemma 4.5. *For any scheme T equipped with a morphism $T \to X_\infty$, we have $T \times_{X_\infty} U_i \cong \varprojlim(T \times_{X_m} U_{m,i})$.* $\qquad\square$

Claim 4.6. *The map $U \to X_\infty$ is surjective.*

Proof. Surjectivity for representable morphisms can be checked locally on the target by [Laumon and Moret-Bailly 2000, 3.10]. Thus, we may change base along a map $T \to X_\infty$ from a scheme T, and taking a point that is the image of a map $\operatorname{Spec}(K) \to T$ where K is a field containing $\Bbbk$, it suffices to find $\operatorname{Spec}(K) \to U$ making the diagram

$$\begin{array}{ccc} \operatorname{Spec}(K) & \longrightarrow & U \\ \| & & \downarrow \\ \operatorname{Spec}(K) & \longrightarrow & X_\infty \end{array} \qquad (4.7)$$

commute.

Thus, suppose we are given a map $\mathrm{Spec}(K) \to X_\infty$; let y_n denote its image (i.e., the image of the unique point of $\mathrm{Spec}(K)$) in X_n. Let I_n denote the (finite) set of those i so that $y_n \in U_{n,i}$. Since the $U_{n,i}$ cover X_n, each I_n is nonempty; further, $I_n \subseteq I_{n-1}$. The intersection $\bigcap_n I_n$ is thus nonempty and contains some i_0. The maps $\mathrm{Spec}(K) \to U_{m,i_0}$ for $m \gg 0$ define a map $f : \mathrm{Spec}(K) \to U_{i_0} = \varprojlim U_{m,i_0} \subset U$, and thus, defining the map in the top row of (4.7) to be f gives the desired commutative diagram. This proves the claim. $\qquad\square$

Returning to the proof of Proposition 4.2(1a), let $R := U \times_{X_\infty} U$. If we define $R_{ij} := U_i \times_{X_\infty} U_j$, then we have $R = \bigsqcup_{i,j} R_{ij}$. Note that R_{ij} is a scheme affine over U_i by the previous paragraph. Since affine schemes are quasicompact, this proves that the morphism $U \to X_\infty$ is quasicompact. Furthermore, $\mathbb{O}(R_{ij})$ is a localization of C_i (obtained by inverting the images of the elements $d_{n,j}$), so $R_{ij} \to U_i$ is flat. We have already proved that $U \to X_\infty$ is surjective, so we conclude that $U \to X_\infty$ is faithfully flat. It follows that $U \to X_\infty$ is fpqc using [Vistoli 2005, Proposition 2.33(iii)]. This completes the proof of (1a).

(1b) The diagonal $\Delta : X_\infty \to X_\infty \times_{\Bbbk} X_\infty$ is the inverse limit of the diagonals $\Delta_n : X_n \to X_n \times_{\Bbbk} X_n$. Similarly to Lemma 4.5, if $V \to X_\infty \times_{\Bbbk} X_\infty$ is any morphism from a scheme V, we get $X_\infty \times_{X_\infty \times_{\Bbbk} X_\infty} V \cong \varprojlim X_n \times_{X_n \times_{\Bbbk} X_n} V$. Since each X_n is separated over $\Bbbk$, each morphism $X_n \times_{X_n \times_{\Bbbk} X_n} V \to V$ is a closed immersion; hence, $X_\infty \times_{X_\infty \times_{\Bbbk} X_\infty} V \to V$ is a closed immersion. This proves (1b).

(2) Again, we may assume that X is affine. Then, as above, we have an fpqc cover $U \xrightarrow{\not\!\,{}^{B}} X_\infty$ by an affine scheme U. Since an affine scheme is quasicompact and a continuous image of a quasicompact space is quasicompact, X_∞ is quasicompact as desired.

(3) By our definition, it suffices to prove that U is noetherian or, equivalently, that each C_i is a noetherian ring. This follows as in [Artin et al. 1999, Theorem 1.5]. We will need the following:

Lemma 4.8. *Let A be a commutative noetherian ring, and (with M an $n \times m$ matrix acting by left multiplication) let J be an ideal of A with a resolution*

$$A^m \xrightarrow{M} A^n \to J \to 0.$$

Let $A' := A[t_1, \ldots, t_{n-1}]/(t_1, \ldots, t_{n-1}, 1)M$. Let P' be a prime of A', and let $P := P' \cap A$. If P and J are comaximal, then $PA' = P'$.

Note that A' is the coordinate ring of a chart of $\mathrm{Proj}\,\mathrm{Sym}_A J$.

Proof. We may localize at P, so without loss of generality, $J = A$. Then $A' \cong A[g^{-1}]$ for some $g \in A$. The result follows. $\qquad\square$

We return to the proof of (3). It suffices to show, by [Eisenbud 1995, Exercise

2.22], that each prime of C_i is finitely generated. Let $\tilde{P} \neq 0$ be a prime of C_i. Let $P := \tilde{P} \cap C$, and let $P_n := \tilde{P} \cap C_{n,i}$. By critical density, there is some $n \in \mathbb{N}$ so that if $m \geq n$, then J_m and P are comaximal. It follows from Lemma 4.8 that $P_m = P_n C_{m,i}$ for $m \geq n$. So $\tilde{P} = \bigcup P_m = P_n C_i$. This is finitely generated because $C_{n,i}$ is noetherian, so P_n is finitely generated.

Proposition 4.2 is now proved. $\qquad\square$

Corollary 4.9. *Let X be a projective variety, let $\sigma \in \mathrm{Aut}_{\mathbb{k}}(X)$, and let $\mathcal{L}$ be a σ-ample invertible sheaf on X. Let P be a zero-dimensional subscheme of X, all of whose points have critically dense σ-orbits. Let $\mathcal{I}_n := \mathcal{I}_P \mathcal{I}_P^\sigma \cdots \mathcal{I}_P^{\sigma^{n-1}}$. Let $a_n : X_n \to X$ be the blowup of X at $\mathcal{I}_n$ as in Corollary 3.3. Let $\alpha_n : X_n \to X_{n-1}$ be given by Corollary 3.3. Then the limit*

$$X_\infty := \varprojlim X_n$$

is a noetherian fpqc-algebraic stack.

Proof. This follows immediately from Proposition 4.2. $\qquad\square$

5. Moduli schemes for truncated point modules

Let X be a projective variety, let $\sigma \in \mathrm{Aut}_{\mathbb{k}}(X)$, and let $\mathcal{L}$ be a σ-ample invertible sheaf on X. Let P be a zero-dimensional subscheme of X, all of whose points have critically dense σ-orbits. We define $\mathcal{S} := \mathcal{S}(X, \mathcal{L}, \sigma, P)$ and $S := S(X, \mathcal{L}, \sigma, P)$ as in Section 2. As usual, we assume that S is generated in degree 1.

In this section, we construct moduli schemes of truncated point modules over $\mathcal{S}$ and S. In the next section, we compare them. We begin by constructing moduli schemes for shifted point modules for an arbitrary connected graded noetherian algebra generated in degree 1, generalizing slightly results of [Artin et al. 1990; Rogalski and Stafford 2009].

Let C be any commutative $\mathbb{k}$-algebra. Recall that we use subscript notation to denote changing base. Thus, if R is a $\mathbb{k}$-algebra, we write $R_C := R \otimes_{\mathbb{k}} C$. We write $X_C := X \times_{\mathbb{k}} \mathrm{Spec}\, C$. Recall that a *C-point module* (*over R*) is a graded factor M of R_C so that M_i is rank-1 projective for $i \geq 0$. An *l-shifted C-point module* (*over R*) is a factor of $(R_C)_{\geq l}$ that is rank-1 projective in degree $\geq l$. A *truncated l-shifted C-point module of length m* is a factor module of $(R_C)_{\geq l}$ so that M_i is rank-1 projective over C for $l \leq i \leq l+m-1$ and $M_i = 0$ for $i \geq l+m$. Since these modules depend on a finite number of parameters, they are clearly parametrized up to isomorphism by a projective scheme. For fixed $l \leq n$, we let ${}_l Y_n$ denote the l-shifted length-$(n-l+1)$ point scheme of R. A point in ${}_l Y_n$ gives a surjection $R_{\geq l} \to M$ (up to isomorphism) or equivalently a submodule of $R_{\geq l}$ with appropriate Hilbert series. Thus, we say that ${}_l Y_n$ parametrizes *embedded* (shifted truncated) point modules. The map $M \mapsto M/M_n$ induces a morphism $\chi_n : {}_l Y_n \to {}_l Y_{n-1}$.

For later use, we explicitly construct a projective embedding of $_lY_n$.

Proposition 5.1 [Artin et al. 1990, Section 3]. *Let R be a connected graded $\Bbbk$-algebra generated in degree 1.*

(1) *For all $l \leq n \in \mathbb{N}$, there is a closed immersion*

$$_l\Pi_n : {_l}Y_n \to \mathbb{P}((R_1^{\otimes l})^\vee) \times \mathbb{P}(R_1^\vee)^{\times(n-l)}.$$

(2) *Fix $l \leq n$, and let*

$$\pi : \mathbb{P}((R_1^{\otimes l})^\vee) \times \mathbb{P}(R_1^\vee)^{\times(n-l)} \to \mathbb{P}((R_1^{\otimes l})^\vee) \times \mathbb{P}(R_1^\vee)^{\times(n-l-1)}$$

be projection onto the first $n-l$ factors. Then the following diagram commutes:

$$
\begin{array}{ccc}
{_l}Y_n & \xrightarrow{\;{_l}\Pi_n\;} & \mathbb{P}((R_1^{\otimes l})^\vee) \times \mathbb{P}(R_1^\vee)^{\times(n-l)} \\
\chi_n \downarrow & & \downarrow \pi \\
{_l}Y_{n-1} & \xrightarrow{\;{_l}\Pi_{n-1}\;} & \mathbb{P}((R_1^{\otimes l})^\vee) \times \mathbb{P}(R_1^\vee)^{\times(n-l-1)}
\end{array}
\tag{5.2}
$$

Proof. (1) Let $T = T^\bullet(R_1)$ denote the tensor algebra on the finite-dimensional $\Bbbk$-vector space R_1. We identify T_1 canonically with R_1 and T_l with $R_1^{\otimes l}$.

Given an element $f \in T_n$, we get an $(l, 1, \ldots, 1)$-form

$$\tilde{f} : T_l^\vee \times (T_1^\vee)^{\times(n-l)} \to \Bbbk$$

by pairing with f. The map is $\Bbbk$-multilinear; hence, $\tilde{f}$ defines a hypersurface $Y(\tilde{f})$ in $\mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}$. More generally, given a collection $\{f_i\}$ of elements of T_n, we get a closed subscheme

$$Y(\{\tilde{f}_i\}) \subseteq \mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}.$$

Let I be the kernel of the natural surjection $T \twoheadrightarrow R$. Then the above construction gives a closed subscheme $Y(\tilde{I}_n) \subseteq \mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}$. We claim that $Y(\tilde{I}_n)$ is naturally isomorphic to $_lY_n$.

Let C be a commutative $\Bbbk$-algebra, and let $R_C = R \otimes_\Bbbk C$ and $T_C = T \otimes_\Bbbk C$ with the gradings induced from R and T, respectively. Suppose that $\bar{\alpha} : (R_C)_{\geq l} \to M$ is an embedded l-shifted truncated C-point module of length $n - l + 1$. We write $\alpha : (T_C)_{\geq l} \to M$ for the composite of the two surjections. Assume that $M = \bigoplus_{i=l}^n m_i \cdot C$ is a free graded C-module on generators m_i. Then α determines C-linear maps $a_j : T_1 \otimes_\Bbbk C \to C$ for $1 \leq j \leq n - l$ by $m_{l+j-1}x = m_{l+j}a_j(x)$ for $x \in T_1 \otimes_\Bbbk C$ and a C-linear map $b : T_l \otimes_\Bbbk C \to C$ by $\alpha(y) = m_l b(y)$ for $y \in T_l \otimes_\Bbbk C$. Since M is a (shifted truncated) point module, hence generated in degree l, these maps are

surjective. Hence, they determine a morphism

$$\Pi(\alpha) = (b, a_1, \ldots, a_{n-l}) : \operatorname{Spec}(C) \to \mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}.$$

We see immediately from the construction that if $f \in I_n \otimes_\Bbbk C$, then $\tilde{f}(\Pi(\alpha)) = 0$. In particular, $\Pi(\alpha)$ factors through $Y(\tilde{I}_n)$.

It follows immediately that the above construction defines a morphism Π from the moduli functor of shifted truncated point modules with *free* (as C-modules) graded components to $Y(\tilde{I}_n)$. Since the latter is a scheme, hence a sheaf in the fpqc topology, Π induces a morphism, which we denote $_l\Pi_n$, from the moduli functor $_lY_n$ for all shifted truncated C-point modules over R to $Y(\tilde{I}_n)$.

Claim 5.3. *The morphism $_l\Pi_n$ is an isomorphism; that is, $Y(\tilde{I}_n) \cong {}_lY_n$ represents the moduli functor of embedded truncated l-shifted C-point modules over R of length $n - l + 1$.*

Proof. A morphism $\operatorname{Spec}(C) \to Y(\tilde{I}_n) \subseteq \mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}$ gives a tuple $(b, a_1, \ldots, a_{n-l})$ where each a_j is a surjective C-linear map $a_j : T_1 \otimes_\Bbbk C \to N_j$ and each N_j is a finitely generated projective C-module of rank 1; and $b : T_l \otimes_\Bbbk C \to M_l$ is a surjective C-linear map onto a finitely generated projective C-module M_l of rank 1.

Assume first that M_l and each N_j is a free C-module, and choose basis elements. Define a T_C-module $M = \bigoplus_{j=l}^n m_j \cdot C$ by $m_{j-1}x = m_j a_{j-l}(x)$ for $x \in T_1 \otimes_\Bbbk C$. Moreover, define a map $\alpha : (T_C)_{\geq l} \to M$ by $\alpha(y) = m_l b(y)$ for $y \in T_l$ and extending linearly. It is a consequence of the construction of $Y(\tilde{I}_n)$ that the map α factors through $(R_C)_{\geq l}$ and makes M an l-shifted truncated C-point module over R.

Next, we observe that the functor Π (on shifted truncated point modules with free C-module components) and the above construction (on maps $\operatorname{Spec}(C) \to Y(\tilde{I}_n)$ for which the modules N_j and M_l are free C-modules) give mutual inverses. This follows from the argument of [Artin et al. 1990, 3.9], which uses only the freeness condition. In particular, the functor Π is injective.

To prove that the sheafification $_l\Pi_n$ is an isomorphism, then it suffices to show that Π is locally surjective; that is, for every morphism $\operatorname{Spec}(C) \to Y(\tilde{I}_n)$, there is a faithfully flat morphism $\operatorname{Spec}(C') \to \operatorname{Spec}(C)$ and a shifted truncated C'-point module with free C'-module components, whose image under Π is the composite map $\operatorname{Spec}(C') \to Y(\tilde{I}_n)$. But it is standard that such a homomorphism $C \to C'$ can be found that makes each N_j and M_l trivial, and now the construction of the previous paragraph proves the existence of the desired shifted truncated point module. This completes the proof of the claim. $\square$

Now part (1) of the proposition follows from the claim, and (2) follows by construction. $\square$

Proposition 5.4 [Artin et al. 1990, Proposition 3.6]. *Let R be a connected graded*
$\Bbbk$*-algebra generated in degree* 1. *Let $n > l$, and consider the truncation morphism*

$$\chi_n : {}_lY_n \to {}_lY_{n-1}.$$

Let $y \in {}_lY_{n-1}$, and suppose that $\dim \chi_n^{-1}(y) = 0$. *Then χ_n^{-1} is defined and is a local
isomorphism locally in a neighborhood of y.*

Proof. We consider the commutative diagram (5.2) of Proposition 5.1(2). By
Proposition 5.1(1), the horizontal maps are closed immersions. Since the defining
equations of $Y(\tilde{I}_n) \subseteq \mathbb{P}(T_l^\vee) \times (\mathbb{P}(T_1^\vee))^{\times(n-l)}$ are $(l, 1, \ldots, 1)$-forms and in partic-
ular are linear in the last coordinate, the fibers of χ_n are linear subspaces of $\mathbb{P}(T_1^\vee)$.
The result follows as in the proof of [Artin et al. 1990, Proposition 3.6(ii)]. $\qquad\square$

Proposition 5.5 [Rogalski and Stafford 2009, Proposition 2.5]. *Let R be a noe-
therian connected graded $\Bbbk$-algebra generated in degree* 1. *For $n > l \geq 0$, de-
fine $\chi_n : {}_lY_n \to {}_lY_{n-1}$ as in the beginning of the section. Let $n_0 \geq 0$, and let
$\{y_n \in {}_lY_n \mid n \geq n_0\}$ be a sequence of (not necessarily closed) points so that
$\chi_n(y_n) = y_{n-1}$ for all $n > n_0$. Then for all $n \gg n_0$, the fiber $\chi_n^{-1}(y_{n-1})$ is a
singleton and χ_n^{-1} is defined and is a local isomorphism at y_{n-1}.*

Proof. This follows as in the proof of [Rogalski and Stafford 2009, Proposition 2.5],
using Proposition 5.4 instead of [Artin et al. 1990, Proposition 3.6(ii)]. $\qquad\square$

We are interested in studying the limit ${}_lY_\infty := \varprojlim {}_lY_n$; however, we first study
the point schemes of $\mathcal{S}$. That is, for $n \geq l \geq 0$, it is clear that we may also define a
scheme that parametrizes factor modules $\mathcal{M}$ of $\mathcal{S}_{\geq l}$ so that, as graded $\mathcal{O}_X$-modules,
$\mathcal{M} \cong \Bbbk_x t^l \oplus \cdots \oplus \Bbbk_x t^n$ for some $x \in X$. We say that x is the *support* of $\mathcal{M}$. We let ${}_lZ_n$
denote this l-shifted length-$(n-l+1)$ truncated point scheme of $\mathcal{S}$. More formally,
a Spec(C)-point of ${}_lZ_n$ will be a factor module $\mathcal{M}$ of $\mathcal{S}_{\geq l} \otimes_\Bbbk C$ that is isomorphic
as a graded $\mathcal{O}_{X_C}$-module to a direct sum $P_l \oplus \cdots \oplus P_n$, where each P_i is a coherent
$\mathcal{O}_{X_C}$-module that is finite over C (in the sense that its support in X_C is finite over
Spec(C)) and is a rank-1 projective C-module (which is well defined because of
the finite support condition). We let $Z_n := {}_0Z_n$ be the unshifted length-$(n+1)$ point
scheme of $\mathcal{S}$.

For all $n > l \geq 0$, there are maps

$$\phi_n : {}_lZ_n \to {}_lZ_{n-1} \quad \text{defined by} \quad \mathcal{M} \mapsto \mathcal{M}/\mathcal{M}_n.$$

If $l = 0$ and $\mathcal{M}$ is a truncated point module of length n over $\mathcal{S}$, then $\mathcal{M}[1]_{\geq 0}$ is
also cyclic (since $\mathcal{S}$ is generated in degree 1) and so is a factor of $\mathcal{S}$ in a unique
way up to a scalar. This induces a map

$$\psi_n : Z_n \to Z_{n-1} \quad \text{defined by} \quad \mathcal{M} \mapsto \mathcal{M}[1]_{\geq 0}.$$

It is clear that ψ_n and ϕ_n map relevant components to relevant components.

Lemma 5.6. *Let* $f_n : {}_lZ_n \to X$ *be the map that sends a module $\mathcal{M}$ to its support. The following diagrams commute*:

$$
\begin{array}{ccc}
 & & {}_lZ_n \\
 & {}^f \nearrow & \downarrow \phi \\
X & & \\
 & {}_f \nwarrow & \downarrow \\
 & & {}_lZ_{n-1}
\end{array}
\tag{5.7}
$$

$$
\begin{array}{ccc}
{}_0Z_n & \xrightarrow{f=\phi^n} & X \\
\psi \downarrow & & \downarrow \sigma \\
{}_0Z_{n-1} & \xrightarrow{f=\phi^{n-1}} & X
\end{array}
\tag{5.8}
$$

Proof. It is clear by construction that if $\mathcal{M}$ is a shifted truncated point module and $\mathcal{M}'$ is a further factor of $\mathcal{M}$, then $\mathcal{M}$ and $\mathcal{M}'$ have the same support. Thus, (5.7) commutes.

Let $\mathcal{M}$ be a truncated point module corresponding to a point $z \in {}_0Z_n$. Let $x := f(z)$. By [Keeler et al. 2005, Lemma 5.5], we have $\mathcal{M}[1]_n \cong (\mathcal{M}_{n+1})^{\sigma^{-1}} \cong (\Bbbk_x)^{\sigma^{-1}} \cong \Bbbk_{\sigma(x)}$. Thus, $f\psi(z) = \sigma f(z)$ as claimed, and (5.8) commutes. $\qquad\square$

Recall that $\mathcal{S}_n = \mathcal{I}_n \mathcal{L}_n$.

Proposition 5.9. *For $n \geq 0$, let X_n be the blowup of X at $\mathcal{I}_n$. Let $\alpha_n, \beta_n : X_n \to X_{n-1}$ be as in Corollary 3.3.*

Then for all $n > l \geq 0$, there are isomorphisms $j_n : X_n \to {}_lZ_n^o \subseteq {}_lZ_n$ so that the following diagrams commute:

$$
\begin{array}{ccc}
X_n & \xrightarrow{j_n} & {}_lZ_n \\
\alpha_n \downarrow & & \downarrow \phi_n \\
X_{n-1} & \xrightarrow{j_{n-1}} & {}_lZ_{n-1}
\end{array}
\quad and \quad
\begin{array}{ccc}
X_n & \xrightarrow{j_n} & {}_0Z_n \\
\beta_n \downarrow & & \downarrow \psi_n \\
X_{n-1} & \xrightarrow{j_{n-1}} & {}_0Z_{n-1}
\end{array}
$$

Proof. We will do the case that $l = 0$; the general case is similar. Let $Z_n := {}_0Z_n$. For $0 \leq i \leq n$, let $W_i = \operatorname{Proj} \mathcal{S}ym_X \mathcal{I}_i$ be the scheme parametrizing colength-1 ideals inside $\mathcal{I}_i$, and let $c_i : X_i \to W_i$ be the map from Lemma 3.1. Let

$$
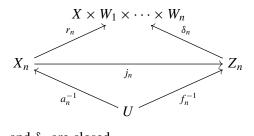r_n : X_n \to X \times W_1 \times \cdots \times W_n
$$

be the composition

$$
X_n \xrightarrow{\alpha^n \times \alpha^{n-1} \times \cdots \times 1} X \times X_1 \times \cdots \times X_n \xrightarrow{c_0 \times \cdots \times c_n} X \times W_1 \times \cdots \times W_n.
$$

Since this is the composition of the graph of a morphism with a closed immersion, it is also a closed immersion and is an isomorphism onto its image.

Now, a point in Z_n corresponds to an ideal $\mathcal{J} \subset \mathcal{S}$ so that the factor is a truncated point module of length $n+1$, and there is thus a canonical closed immersion $\delta_n : Z_n \to X \times W_1 \times \cdots \times W_n$. The map δ_n sends a graded right ideal $\mathcal{J}$ of $\mathcal{S}$ to the tuple $(\mathcal{J}_0, \mathcal{J}_1, \ldots, \mathcal{J}_n)$.

Conversely, a point $(\mathcal{J}_0, \ldots, \mathcal{J}_n) \in X \times W_1 \times \cdots \times W_n$ is in $\mathrm{Im}(\delta)$ if and only if we have $\mathcal{J}_i \mathcal{S}_j^{\sigma^i} \subseteq \mathcal{J}_{i+j}$ for all $i + j \leq n$. It follows from Lemma 3.2(c) that $\mathrm{Im}(r_n) \subseteq \mathrm{Im}(\delta_n)$. Since r_n and δ_n are closed immersions and X_n is reduced, we may define $j_n = \delta_n^{-1} r_n : X_n \to Z_n$.

Let $U := X \smallsetminus \mathrm{Cosupp}\, \mathcal{I}_n$. Then f_n^{-1} and a_n^{-1} are defined on U, and the diagram

commutes. Since r_n and δ_n are closed,

$$r_n(X_n) = r_n(\overline{a_n^{-1}(U)}) = \overline{r_n a_n^{-1}(U)} = \overline{\delta_n f_n^{-1}(U)} = \delta_n(\overline{f_n^{-1}(U)}) = \delta_n(Z_n^o).$$

Therefore, j_n is an isomorphism to Z_n^o.

Let $q : X \times W_1 \times \cdots \times W_n \to X \times W_1 \times \cdots \times W_{n-1}$ be projection onto the first n factors. Consider the diagram

$$
\begin{array}{ccccc}
X_n & \xrightarrow{\ r_n\ } & X \times W_1 \times \cdots \times W_n & \xleftarrow{\ \delta_n\ } & Z_n \\
{\scriptstyle \alpha_n}\big\downarrow & & {\scriptstyle q}\big\downarrow & & \big\downarrow{\scriptstyle \phi_n} \\
X_{n-1} & \xrightarrow[\ r_{n-1}\ ]{} & X \times W_1 \times \cdots \times W_{n-1} & \xleftarrow[\ \delta_{n-1}\ ]{} & Z_{n-1}
\end{array}
$$

From the definitions of r_n and δ_n, we see that this diagram commutes; since $j_n = \delta_n^{-1} r_n$, the following diagram commutes:

$$
\begin{array}{ccc}
X_n & \xrightarrow{\ j_n\ } & Z_n \\
{\scriptstyle \alpha_n}\big\downarrow & & \big\downarrow{\scriptstyle \phi_n} \\
X_{n-1} & \xrightarrow[\ j_{n-1}\ ]{} & Z_{n-1}
\end{array}
$$

Let $X_n' := \mathrm{Im}(r_n) \subset X \times W_1 \times \cdots \times W_n$, and let $\beta_n' : X_n' \to X_{n-1}'$ be the map induced from β_n. The proof of Corollary 3.3 shows that if $(\mathcal{J}_0, \mathcal{J}_1, \ldots, \mathcal{J}_n) \in X_n'$,

then its image under β'_n is

$$\beta'_n((\mathcal{J}_0, \mathcal{J}_1, \ldots, \mathcal{J}_n)) = (\mathcal{F}_0, \ldots, \mathcal{F}_{n-1}) \in X'_{n-1}, \qquad (5.10)$$

where $\mathcal{F}_i := (\mathcal{J}_{i+1} : \mathcal{J}_1)^{\sigma^{-1}} \cap \mathcal{J}_i$.

Now let $\mathcal{J}$ be the ideal defining a truncated point module of length $n+1$. By abuse of notation, we think of $\mathcal{J}$ as a point in Z_n. Let $\mathcal{M} := \mathcal{S}/\mathcal{J}$. Then we have $\psi(\mathcal{J})_i = (\mathrm{Ann}_{\mathcal{S}}(\mathcal{M}_1))_i$ for $0 \leq i \leq n-1$. Thus, $\mathcal{J}_1(\psi(\mathcal{J})_i)^{\sigma} \subseteq \mathcal{J}_{i+1}$ or $\psi(\mathcal{J})_i \subseteq (\mathcal{J}_{i+1} : \mathcal{J}_1)^{\sigma^{-1}} \cap \mathcal{J}_i$. If $\mathcal{J} \in \mathrm{Im}(j_n)$, then we have equality by the computation in (5.10). Thus, the diagram

commutes.   □

To end this section, we construct stacks $_l Z_\infty$ and $_l Y_\infty$ that are fine moduli spaces for (shifted) embedded point modules and give some of their properties. A version of the following result was known long ago to M. Artin; however, it does not seem to have appeared in the literature:

Theorem 5.11. *Fix* $l \in \mathbb{N}$. *For* $n \geq l$, *let* X_n *be the blowup of* X *at* $\mathcal{I}_n$. *Let* $_l Y_n$ *be the moduli space of l-shifted length-$(n-l+1)$ point modules over* S. *Let* $_l Z_n$ *be the moduli space of l-shifted length-$(n-l+1)$ point modules over* $\mathcal{S}$. *Define the morphisms* $\chi_n : {}_l Y_n \to {}_l Y_{n-1}$, $\phi_n : {}_l Z_n \to {}_l Z_{n-1}$, *and* $\alpha_n : X_n \to X_{n-1}$ *as above. Let*

$$_l Z_\infty := \varprojlim_{\phi_n} {}_l Z_n, \quad {}_l Y_\infty := \varprojlim_{\chi_n} {}_l Y_n, \quad \text{and} \quad X_\infty := \varprojlim_{\alpha_n} X_n.$$

Then the stack $_l Y_\infty$ *is a sheaf in the fpqc topology and is a fine moduli space for l-shifted embedded point modules over* S. *The stack* $_l Z_\infty$ *is noetherian fpqc-algebraic and is a fine moduli space for l-shifted embedded point modules over* $\mathcal{S}$. *The relevant component of* $_l Z_\infty$ *is isomorphic to* X_∞.

Proof. We suppress the subscript l in the proof.

For $n \geq l$, let F_n be the moduli functor for truncated l-shifted point modules over S, so $Y_n \cong F_n$. Define a contravariant functor

$$F : \text{Affine schemes} \to \text{Sets},$$

$$\mathrm{Spec}\, C \mapsto \{\text{Embedded } l\text{-shifted } C\text{-point modules over } S\}.$$

By descent theory, F is a sheaf in the fpqc topology. More precisely, recall that quasicoherent sheaves form a stack in the fpqc topology [Vistoli 2005, Section 4.2.2];

consequently, the (graded) quotients of $S_{\geq l}$ form a sheaf of sets in the fpqc topology. Moreover, as in the first paragraph of [Vistoli 2005, Section 4.2.3], those quotients of $S_{\geq l}$ that are S-module quotients form a subsheaf in the fpqc topology; this subsheaf is F. It is formal that F is isomorphic to the functor h_{Y_∞}.

Likewise, Z_∞ parametrizes l-shifted point modules over $\mathcal{S}$. We show that (i)–(v) of Proposition 4.2 apply to Z_∞. We have $\mathcal{S}_n = \mathcal{I}_n \mathcal{L}_n$; let $P_n \subset X$ be the subscheme defined by $\mathcal{I}_n$. Consider the maps

$$Z_n \xrightarrow{\phi_n} Z_{n-1} \underset{f_n}{\overset{f_{n-1}}{\rightrightarrows}} X$$

from Lemma 5.6. Now, f_n^{-1} is defined away from P_n, and $\bigcup_n P_n$ is a countable critically dense set. Thus, (i) and (ii) hold.

Let $x \in \bigcup P_n$. As the points in P have infinite orbits, there is some $m \in \mathbb{N}$ so that $x \notin \sigma^{-n}(P)$ for all $n \geq m$. Let $z_m \in Z_m$ with $f_m(z_m) = x$, corresponding to a right ideal $\mathcal{J} \subseteq \mathcal{S}_{\geq l}$ with $S_{\geq l}/\mathcal{J} \cong \bigoplus_{j=l}^m \mathbb{O}_x$. Let $\mathcal{J}' := \mathcal{J}_{\leq m} \cdot \mathcal{S}$. For any $j \geq 0$, we have $(\mathcal{S}_j^{\sigma^m})_x = (\mathcal{L}_j^{\sigma^m})_x$, and so $\mathcal{J}'$ gives the unique preimage of z_m in Z_∞. A similar uniqueness holds upon base extension, so the scheme-theoretic preimage of z_m in Z_∞ is a $\mathbb{k}$-point, and Φ_m^{-1} is defined and is a local isomorphism at z_m.

For $j \geq l$, let $W_j := \operatorname{Proj} \mathcal{S}ym_X \mathcal{I}_n$. As in the proof of Proposition 5.9, we may regard Z_n as a closed subscheme of $W_l \times \cdots \times W_n$, and (iv) and (v) follow from this and the fact that the orbits of points in P are infinite. By Proposition 4.2, then Z_∞ is a noetherian fpqc-algebraic stack.

Consider the morphisms $j_n : X_n \overset{\cong}{\rightrightarrows} Z_n^o \subseteq Z_n$ from Proposition 5.9. Commutativity of the first diagram there gives an induced isomorphism $j : X_\infty \to Z_\infty^o \subseteq Z_\infty$. $\square$

6. Comparing moduli of points

In this section, we prove that $_l Y_\infty$ is also noetherian fpqc-algebraic and that, at least for sufficiently large l, the stacks $_l Z_\infty$ and $_l Y_\infty$ are isomorphic.

In the following pages, we will always use the following notation. We write a commutative $\mathbb{k}$-algebra C as $p : \mathbb{k} \to C$ to indicate the structure map explicitly. We write $X_C := X \otimes_{\mathbb{k}} \operatorname{Spec} C$. We abuse notation and let p also denote the projection map $1 \otimes p : X_C \to X$. We let $q : X_C \to \operatorname{Spec} C$ be projection on the second factor.

Suppose that $p : \mathbb{k} \to C$ is a commutative $\mathbb{k}$-algebra and $y : \operatorname{Spec} C \to X$ is a C-point of X. Then y determines a section of q, which we also call y. This is a morphism $y : \operatorname{Spec} C \to X_C$. We define $\mathcal{I}_y \subseteq \mathbb{O}_{X_C}$ to be the ideal sheaf of the corresponding closed subscheme of X_C. We define $\mathbb{O}_y := \mathbb{O}_{X_C}/\mathcal{I}_y$.

We use the relative regularity results from Section 2 to study the pullbacks of the sheaves $\mathcal{S}_n$ to X_C. Fix a very ample invertible sheaf $\mathbb{O}_X(1)$ on X, which we will use to measure regularity.

Lemma 6.1. *Suppose $p : \Bbbk \to C$ is a commutative noetherian $\Bbbk$-algebra. Then $\{p^*\mathcal{S}_n\}_{n \geq 0}$ is a right ample sequence on X_C.*

Proof. Let $\mathcal{F}$ be a coherent sheaf on X_C. By [Rogalski and Stafford 2007, Corollary 3.14], $\lim_{n \to \infty} \operatorname{reg}(\mathcal{S}_n) = -\infty$. Thus, $\lim_{n \to \infty} \operatorname{reg}(p^*\mathcal{S}_n) = -\infty$ by Lemma 2.9. Since each $\mathcal{S}_n$ is invertible away from a dimension-0 set, $p^*\mathcal{S}_n$ is invertible away from a locus of relative dimension 0. By Proposition 2.13, $\mathcal{F} \otimes_{X_C} p^*\mathcal{S}_n$ is 0-regular for $n \gg 0$. Theorem 2.10 shows that (1) and (2) of Definition 2.7 apply.   □

We now prove a uniform regularity result for certain subsheaves of a pullback of some $\mathcal{S}_n$.

Lemma 6.2. *There exists $m \geq 0$ so that the following holds for any $n \geq m$: for any commutative noetherian $\Bbbk$-algebra $p : \Bbbk \to C$, for any C-point y of X, and for any coherent sheaf $\mathcal{K}$ on X_C so that $\mathcal{I}_y p^*\mathcal{S}_n \subseteq \mathcal{K} \subseteq p^*\mathcal{S}_n$, $\mathcal{K}$ is 0-regular. In particular, $\mathcal{K}$ is globally generated and $R^1 q_* \mathcal{K} = 0$.*

Proof. Let D be the constant from Proposition 2.13, and let $r := \operatorname{reg}(\mathcal{O}_X)$. By [Rogalski and Stafford 2007, Corollary 3.14], we have $\lim_{n \to \infty} \operatorname{reg}(\mathcal{S}_n) = -\infty$. Let m be such that for all $n \geq m$, $\operatorname{reg}(\mathcal{S}_n) \leq -r - D - 1$. We claim this m satisfies the conclusions of the lemma.

Fix a commutative noetherian $\Bbbk$-algebra $p : \Bbbk \to C$ and a C-point y of X. We first claim that $\operatorname{reg}(\mathcal{I}_y) \leq r + 1$. To see this, let $i \geq 1$ and consider the exact sequence

$$R^{i-1} q_* \mathcal{O}_{X_C}(r + 1 - i) \xrightarrow{\alpha} R^{i-1} q_* \mathcal{O}_y(r + 1 - i)$$
$$\to R^i q_* \mathcal{I}_y(r + 1 - i) \to R^i q_* \mathcal{O}_{X_C}(r + 1 - i).$$

The last term vanishes as $\mathcal{O}_{X_C}$ is $(r+1)$-regular by Lemma 2.9 and Theorem 2.10(1). If $i \geq 2$, then $R^{i-1} q_* \mathcal{O}_y(r+1-i) = 0$ for dimension reasons, so $R^i q_* \mathcal{I}_y(r+1-i) = 0$. On the other hand, if $i = 1$, then because $\mathcal{O}_{X_C}(r)$ is 0-regular, by Lemma 2.11, α is surjective. Again, $R^i q_* \mathcal{I}_y(r + 1 - i) = 0$. Thus, $\mathcal{I}_y$ is $(r + 1)$-regular as claimed.

Let $n \geq m$. By Lemma 2.9, $\operatorname{reg}(p^*\mathcal{S}_n) = \operatorname{reg}(\mathcal{S}_n)$. Note that $\mathcal{I}_y$ and $p^*\mathcal{S}_n$ are both locally free away from a set of relative dimension 0. Thus, the hypotheses of Proposition 2.13 apply, and by that result, we have

$$\operatorname{reg}(\mathcal{I}_y \otimes_{X_C} p^*\mathcal{S}_n) \leq \operatorname{reg}(\mathcal{I}_y) + \operatorname{reg}(p^*\mathcal{S}_n) + D$$
$$\leq r + 1 + D + \operatorname{reg}(p^*\mathcal{S}_n) = r + 1 + D + \operatorname{reg}(\mathcal{S}_n).$$

Our choice of n ensures this is nonpositive. In particular, $\mathcal{I}_y \otimes_{X_C} p^*\mathcal{S}_n$ is 0-regular.

Let $\mathcal{I}_y p^*\mathcal{S}_n \subseteq \mathcal{K} \subseteq p^*\mathcal{S}_n$. There is a natural map $f : \mathcal{I}_y \otimes_{X_C} p^*\mathcal{S}_n \to \mathcal{K}$ given by the composition $\mathcal{I}_y \otimes_{X_C} p^*\mathcal{S}_n \to \mathcal{I}_y \cdot p^*\mathcal{S}_n \subseteq \mathcal{K}$. The kernel and cokernel of f are supported on a set of relative dimension 0, and it is an easy exercise to show that $\mathcal{K}$ is therefore also 0-regular. By Theorem 2.10, $\mathcal{K}$ is globally generated and $R^1 q_* \mathcal{K} = 0$ as claimed.   □

Definition 6.3. We call a positive integer m satisfying the conclusion of Lemma 6.2 a *positivity parameter*.

The proof of Lemma 6.2 shows that if we are willing to replace $\mathcal{L}$ by a sufficiently ample invertible sheaf, we may in fact assume that $m = 1$ is a positivity parameter. (By [Keeler 2000, Theorem 1.2], the existence of a σ-ample sheaf means that any ample invertible sheaf is σ-ample.)

Corollary 6.4. *Let $p : \Bbbk \to C$ be a noetherian commutative $\Bbbk$-algebra. Let m be a positivity parameter (Definition 6.3), and let $n \geq m$.*

(1) *If $\mathcal{J} \subset p^*\mathcal{S}_n$ is a sheaf on X_C so that $p^*\mathcal{S}_n/\mathcal{J}$ has support on X_C that is finite over $\mathrm{Spec}(C)$ and is a rank-1 projective C-module, then $q_*\mathcal{J}$ is a C-submodule of $q_*p^*\mathcal{S}_n = S_n \otimes C$ such that the cokernel is rank-1 projective.*

(2) *If $\mathcal{K} \subsetneqq \mathcal{J} \subset p^*\mathcal{S}_n$ are sheaves on X_C so that $p^*\mathcal{S}_n/\mathcal{J}$ is a rank-1 projective C-module, then $q_*\mathcal{K} \subsetneqq q_*\mathcal{J}$.*

(3) *If $\mathcal{J}, \mathcal{J}' \subseteq p^*\mathcal{S}_n$ are sheaves on X_C so that $p^*\mathcal{S}_n/\mathcal{J}$ and $p^*\mathcal{S}_n/\mathcal{J}'$ are rank-1 projective C-modules, then $q_*\mathcal{J} = q_*\mathcal{J}'$ if and only if $\mathcal{J} = \mathcal{J}'$.*

Proof. (1) Let $x \in \mathrm{Spec}\, C$ be a closed point. Consider the fiber square

$$
\begin{array}{ccc}
X_x & \longrightarrow & X_C \\
\downarrow & \square & \downarrow q \\
\{x\} & \longrightarrow & \mathrm{Spec}\, C
\end{array}
$$

Let $\mathcal{J}_x := \mathcal{J}|_{X_x}$. Since $p^*\mathcal{S}_n/\mathcal{J}$ is flat over $\mathrm{Spec}\, C$, $\mathcal{J}_x \subseteq p^*\mathcal{S}_n|_{X_x} \cong \mathcal{S}_n \otimes_{\Bbbk} \Bbbk(x)$. Further, $(\mathcal{S}_n \otimes_{\Bbbk} \Bbbk(x))/\mathcal{J}_x \cong \mathcal{O}_x$. By our choice of n, therefore, $H^1(X_x, \mathcal{J}_x) = 0$.

Now $\mathcal{J}$ is the kernel of a surjective morphism of flat sheaves and so is flat over $\mathrm{Spec}\, C$. Since $H^1(X_x, \mathcal{J}_x) = 0$, by the theorem on cohomology and base change [Hartshorne 1977, Theorem III.12.11(a)], we have $R^1 q_*\mathcal{J} \otimes_C \Bbbk(x) = 0$. The C-module $R^1 q_*\mathcal{J}$ thus vanishes at every closed point and is therefore 0.

The complex

$$0 \to q_*\mathcal{J} \to q_*p^*\mathcal{S}_n \to q_*(p^*\mathcal{S}_n/\mathcal{J}) \to 0$$

is thus exact. By assumption, $q_*(p^*\mathcal{S}_n/\mathcal{J})$ is a rank-1 projective C-module. Since cohomology commutes with flat base change [Hartshorne 1977, Proposition III.9.3], we have $q_*p^*\mathcal{S}_n \cong H^0(X, \mathcal{S}_n) \otimes_{\Bbbk} C = S_n \otimes_{\Bbbk} C$.

(2) Since m is a positivity parameter, $\mathcal{J}$ is globally generated, and it follows immediately that $q_*\mathcal{K} \neq q_*\mathcal{J}$.

(3) From (2), we have

$$q_*(\mathcal{J} \cap \mathcal{J}') = q_*\mathcal{J} \iff \mathcal{J} \cap \mathcal{J}' = \mathcal{J} \iff \mathcal{J} \subseteq \mathcal{J}'.$$

It follows from our assumptions that this occurs if and only if $\mathscr{I} = \mathscr{I}'$. Further,

$$q_* \mathscr{I} = q_*(\mathscr{I} \cap \mathscr{I}') = q_*(\mathscr{I}) \cap q_* \mathscr{I}' \iff q_* \mathscr{I} \subseteq q_* \mathscr{I}'.$$

From (1), we obtain that this is equivalent to $q_* \mathscr{I} = q_* \mathscr{I}'$. □

We now apply these regularity results to show that $_l Y_\infty \cong {}_l Z_\infty$ for $l \gg 0$. We will need the following easy lemma:

Lemma 6.5. *Let A and B be commutative noetherian local $\Bbbk$-algebras with residue field $\Bbbk$ and $s : A \to B$ a local homomorphism. If $s^* : \mathrm{Hom}^{\mathrm{alg}}(B, C) \to \mathrm{Hom}^{\mathrm{alg}}(A, C)$ is surjective for all finite-dimensional commutative local $\Bbbk$-algebras C, then s is injective.*

Proof. Let $\mathfrak{m}$ be the maximal ideal of A, and let $\mathfrak{n}$ be the maximal ideal of B. Let $f \in \ker s$. Suppose first that there is some k so that $f \in \mathfrak{m}^{k-1} \setminus \mathfrak{m}^k$. Let $C := A/\mathfrak{m}^k$, and let $\pi : A \to C$ be the natural map. Then C is a finite-dimensional artinian local $\Bbbk$-algebra. Now, $\pi(f) \neq 0$ but $s(f) = 0$. Thus, $\pi \notin \mathrm{Im}(s^*)$, a contradiction.

We thus have $\ker s \subseteq \bigcap_k \mathfrak{m}^k$. By the Artin–Rees lemma, $\ker s = 0$. □

Proposition 6.6. *Let m be a positivity parameter, and let $n \geq l \geq m$. Let $_l Z_n$ be the l-shifted length-$(n - l + 1)$ point scheme of $\mathscr{S}$ with truncation morphism $\phi_n : Z_n \to Z_{n-1}$ as in Proposition 5.9. Let $_l Y_n$ be the l-shifted length-$(n - l + 1)$ point scheme of S with truncation morphism $\chi_n : {}_l Y_n \to {}_l Y_{n-1}$ as in Theorem 5.11. Let $_l Z_\infty := \varprojlim {}_l Z_n$, and let $_l Y_\infty := \varprojlim Y_n$.*

Then the global section functor induces a closed immersion $s_n : {}_l Z_n \to {}_l Y_n$ so that the following diagram commutes:

$$
\begin{array}{ccc}
{}_l Z_n & \xrightarrow{\;s_n\;} & {}_l Y_n \\
{\scriptstyle \phi_n} \downarrow & & \downarrow {\scriptstyle \chi_n} \\
{}_l Z_{n-1} & \xrightarrow{\;s_{n-1}\;} & {}_l Y_{n-1}
\end{array}
\tag{6.7}
$$

Proof. Let $p : \Bbbk \to C$ be a commutative noetherian $\Bbbk$-algebra. Let $p : X_C \to X$ and $q : X_C \to \mathrm{Spec}\, C$ be the two projection maps as usual. Note that if $\mathscr{I} \subset p^* \mathscr{S}_{\geq l}$ is the defining ideal of an l-shifted length-$(n - l + 1)$ truncated C-point module over $\mathscr{S}$, then by Corollary 6.4(1), $q_* \mathscr{I}$ is the defining ideal of an l-shifted length-$(n - l + 1)$ truncated C-point module over S. Thus, s_n is well defined. By Corollary 6.4(3), s_n is injective on $\Bbbk$-points and on $\Bbbk[\epsilon]$ points. It is standard [Harris 1992, proof of Theorem 14.9] that, because s_n is projective, s_n is a closed immersion. That (6.7) commutes is immediate. □

We will see that $_l Y_\infty$ and $_l Z_\infty$ are isomorphic. It does not seem to be generally true that $_l Y_n$ and $_l Z_n$ are isomorphic, but we will see that there is an induced isomorphism of certain naturally defined closed subschemes.

Let $l \in \mathbb{N}$. For any n, let $\Phi_n : {}_l Z_\infty \to {}_l Z_n$ and $\Upsilon_n : {}_l Y_\infty \to {}_l Y_n$ be the induced maps. For any $n \geq l$, define ${}_l Z'_n \subseteq {}_l Z_n$ to be the image of $\Phi_n : {}_l Z_\infty \to {}_l Z_n$. That is,

$$ {}_l Z'_n = \bigcap_{i \geq 0} \mathrm{Im}(\phi^i : {}_l Z_{n+i} \to {}_l Z_n). $$

Since ${}_l Z_n$ is noetherian and the ϕ_n are closed, this is a closed subscheme of ${}_l Z_n$ equal to $\mathrm{Im}(\phi^k : {}_l Z_{n+k} \to {}_l Z_n)$ for some k. Similarly, let ${}_l Y'_n = \mathrm{Im}(\Upsilon_n : {}_l Y_\infty \to {}_l Y_n)$. Clearly ${}_l Z_\infty = \varprojlim {}_l Z'_n$ and ${}_l Y_\infty = \varprojlim {}_l Y'_n$. We refer to ${}_l Z'_n$ and ${}_l Y'_n$ as *essential* point schemes as modules in ${}_l Z'_n$ and ${}_l Y'_n$ are truncations of honest (shifted) point modules.

Theorem 6.8. *Let m be a positivity parameter, and let $n \geq l \geq m$.*

(1) *The morphism $s_n : {}_l Z_n \to {}_l Y_n$ defined in [Proposition 6.6](#) induces an isomorphism of essential point schemes*

$$ s'_n : {}_l Z'_n \xrightarrow{\;\cong\;} {}_l Y'_n. $$

(2) *The limit $s : {}_l Z_\infty \to {}_l Y_\infty$ is an isomorphism of stacks.*

Proof. Since the subscript l will remain fixed, we suppress it in the notation. Let $s'_n := s_n|_{Z'_n}$. It follows from commutativity of (6.7) that $s'_n(Z'_n) \subseteq Y'_n$.

We next prove (2). The limit $s = \varprojlim s_n$ is clearly a morphism of stacks, that is, a natural transformation of functors. Let C be a commutative finite-dimensional local $\Bbbk$-algebra. We will show that s is bijective on C-points.

Let $y \in Y_\infty(C)$ be a C-point of Y_∞, which by [Theorem 5.11](#) corresponds to an exact sequence

$$ 0 \to J \to (S_C)_{\geq l} \to M \to 0, $$

where M is an l-shifted S_C-point module.

For $i \geq l$, let $\mathcal{J}_i \subseteq p^* \mathcal{S}_i$ be the subsheaf generated by $J_i \subseteq q_* p^* \mathcal{S}_i$. Let $\mathcal{J} := \bigoplus_{i \geq l} \mathcal{J}_i$. We will show that $\mathcal{M} := p^* \mathcal{S}_{\geq l} / \mathcal{J}$ is an l-shifted $p^* \mathcal{S}$-point module, that is, that there is a C-point y of X so that $\mathcal{M}_n \cong \mathbb{O}_y \otimes p^* \mathcal{L}_n$ for all $n \geq l$.

Since $p^* \mathcal{S}_j$ is globally generated for $j \geq m$, we have $\mathcal{J}_i p^* \mathcal{S}_j^{\sigma^i} \subseteq \mathcal{J}_{i+j}$ for $i \geq l$ and $j \geq m$. Therefore, $\mathcal{M}$ is a coherent right module over the bimodule algebra $\mathcal{S}' := \mathbb{O}_{X_C} \oplus \bigoplus_{j \geq m} p^* \mathcal{S}_j$. Further, each $\mathcal{M}_j$ is clearly torsion over X as $\mathrm{Spec}\, C$ is zero-dimensional. As in [Rogalski and Stafford 2007, Lemma 4.1(1)], it follows from critical density of the orbits of the points in P that there are a coherent X-torsion sheaf $\mathcal{F}$ on X_C and $n_0 \geq l$ so that $\mathcal{M}_j \cong \mathcal{F} \otimes_{X_C} p^* \mathcal{L}_j$ for $j \geq n_0$. By critical density again, there is $n_1 \geq n_0$ so that

$$ \mathrm{Supp}(\mathcal{F}) \cap p^{-1}(\sigma^{-j}(P)) = \varnothing $$

for $j \geq n_1$. This implies that for $j \geq n_1$ and $k \geq 1$, we have

$$ \mathcal{M}_j \cdot p^* \mathcal{S}_k^{\sigma^j} = \mathcal{M}_j \otimes_{X_C} p^* \mathcal{S}_k^{\sigma^j} = \mathcal{M}_j \otimes_{X_C} p^* \mathcal{L}_k^{\sigma^j} \cong \mathcal{M}_{j+k} $$

and $\mathcal{J}_j(p^*\mathcal{S}_k^{\sigma^j}) = \mathcal{J}_{j+k}$. (In particular, $\mathcal{J}_{\geq n_1}$ and $\mathcal{M}_{\geq n_1}$ are right $p^*\mathcal{S}$-modules.)

By Lemma 6.1, $\{p^*\mathcal{S}_k^{\sigma^j}\}_{k \geq 0}$ is right ample on X_C. As in [Rogalski and Stafford 2009, Lemma 9.3], for any $i \geq l$ and $k \gg 0$, we have $q_*(\mathcal{J}_i p^*\mathcal{S}_k^{\sigma^i}) = J_i(S_C)_k \subseteq J_{i+k}$. For $i \geq n_1$ and for $k \geq 1$, we have $q_*(\mathcal{J}_i p^*\mathcal{S}_k^{\sigma^i}) = q_*\mathcal{J}_{i+k} \supseteq J_{i+k}$. There is thus $n_2 \geq n_1$ so that $J_j = q_*\mathcal{J}_j$ for $j \geq n_2$.

It follows from Lemma 6.1 that there is $n_3 \geq n_2$ so that the top row of

$$0 \longrightarrow q_*\mathcal{J}_j \longrightarrow (S_C)_j \longrightarrow q_*\mathcal{M}_j \longrightarrow 0$$
$$\| \qquad\qquad\qquad\qquad\qquad \|$$
$$J_j \qquad\qquad\qquad q_*(\mathcal{F} \otimes_{X_C} p^*\mathcal{L}_j)$$

is exact for $j \geq n_3$. Thus, $q_*(\mathcal{F} \otimes_{X_C} p^*\mathcal{L}_j) \cong (S_C)_j/J_j \cong C$ for $j \geq n_3$. Since $\{p^*\mathcal{L}_j\}$ is right ample, this implies that $\mathcal{F} \cong \mathbb{O}_y$ for some C-point y of X.

For $l \leq i \leq n_3$, let $\mathcal{J}_i' := \mathcal{J}_i + \mathcal{I}_y p^*\mathcal{S}_i$. By choice of l, $\mathcal{J}_i'$ is 0-regular. Thus, the rows of the commutative diagram

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & J_i & \longrightarrow & (S_C)_i & \longrightarrow & C & \longrightarrow & 0 \\
& & \subseteq \downarrow & & \| & & \downarrow \alpha & & \\
0 & \longrightarrow & q_*\mathcal{J}_i' & \longrightarrow & q_*p^*\mathcal{S}_i & \longrightarrow & q_*(p^*\mathcal{S}_i/\mathcal{J}_i') & \longrightarrow & 0
\end{array}
\tag{6.9}
$$

are exact, and α is therefore surjective. This shows that, as a $(C \cong \mathbb{O}_y)$-module, $p^*\mathcal{S}_i/\mathcal{J}_i'$ is cyclic.

Let $N := \mathrm{Ann}_C(p^*\mathcal{S}_i/\mathcal{J}_i')$. For $i \gg 0$, $p^*\mathcal{S}_i/\mathcal{J}_i \cong \mathbb{O}_y \otimes_{X_C} p^*\mathcal{L}_i$ is killed by $\mathcal{I}_y$. Thus, for $i \gg 0$, we have $(\mathcal{J}_i')_y = (\mathcal{J}_i)_y + (\mathcal{I}_y p^*\mathcal{S}_i)_y = (\mathcal{J}_i)_y$, and

$$(\mathcal{J}_{i+j})_y \supseteq (\mathcal{J}_i)_y(p^*\mathcal{S}_j^{\sigma^i})_y = (\mathcal{J}_i')_y(p^*\mathcal{S}_j^{\sigma^i})_y$$
$$\supseteq N(p^*\mathcal{S}_i)_y(p^*\mathcal{S}_j^{\sigma^i})_y = N(p^*\mathcal{S}_{i+j})_y.$$

As $\mathrm{Ann}_C(p^*\mathcal{S}_{i+j}/\mathcal{J}_{i+j}) = 0$ for $j \gg 0$, we must have $N = 0$.

Thus, in fact $p^*\mathcal{S}_i/\mathcal{J}_i' \cong \mathbb{O}_y \cong C$. Looking again at (6.9), we see that α is an isomorphism, and so $J_i = q_*\mathcal{J}_i'$. As $\mathcal{J}_i'$ is 0-regular, it is globally generated: in other words, $\mathcal{J}_i = \mathcal{J}_i'$.

We still need to show $\mathcal{J}$ is a $p^*\mathcal{S}$-module. Let $i \geq l$, and suppose $\mathcal{J}_i p^*\mathcal{S}_1^{\sigma^i} \not\subseteq \mathcal{J}_{i+1}$, so $(\mathcal{J}_{i+1} + \mathcal{J}_i p^*\mathcal{S}_1^{\sigma^i})/\mathcal{J}_{i+1}$ is a nonzero submodule of $p^*\mathcal{S}_{i+1}/\mathcal{J}_{i+1} \cong \mathbb{O}_y$. Since $(\mathcal{J}_{i+1} + \mathcal{J}_i p^*\mathcal{S}_1^{\sigma^i})\mathcal{S}_j^{\sigma^{i+1}} \subseteq \mathcal{J}_{i+j+1}$ for all $j \gg 0$, a similar argument to the last paragraph but one produces a contradiction.

Thus, $p^*\mathcal{S}_{\geq l}/\mathcal{J}$ is an l-shifted C-point module for $\mathcal{S}$, and $q_*\mathcal{J} = J$. This shows that $s : Z_\infty \to Y_\infty$ induces a surjection on C-points. It follows from Corollary 6.4(3) that s is injective on C-points.

Consider the commutative diagram

$$
\begin{array}{ccc}
Z_\infty & \xrightarrow{\ s\ } & Y_\infty \\
\Phi_n \downarrow & & \downarrow \Upsilon_n \\
Z'_n & \xrightarrow{\ s'_n\ } & Y'_n.
\end{array}
\tag{6.10}
$$

Let $x \in X$ be a $\Bbbk$-point. There is some N so that for $n \geq N$, Φ_n is a local isomorphism at all points in the preimage of x. We claim that for $n \geq N$, in fact all maps in (6.10) are local isomorphisms at all points in the preimage of x. In particular, s is an isomorphism in the preimage of x; since x was arbitrary, s is therefore an isomorphism of stacks.

So it suffices to prove the claim. We may work locally. If $z \in Z_\infty$ is a point lying over x, let z_n, y, and y_n be the images of z in Z'_n, Y_∞, and Y'_n, respectively. Let $B := \mathcal{O}_{Z_\infty, z} \cong \mathcal{O}_{Z'_n, z_n}$. Let $A := \mathcal{O}_{Y_\infty, y}$, and let $A' := \mathcal{O}_{Y'_n, y_n}$. We have a commutative diagram of local homomorphisms

$$
\begin{array}{ccc}
B & \xleftarrow{\ s^{\#}\ } & A \\
\| & & \uparrow \Upsilon_n^{\#} \\
B & \xleftarrow{\ (s'_n)^{\#}\ } & A'
\end{array}
$$

As Υ_n is scheme-theoretically surjective, $\Upsilon_n^{\#}$ is injective. It follows from Proposition 5.5 that A is isomorphic to a local ring of some Y'_m and in particular is a noetherian $\Bbbk$-algebra. By Lemma 6.5, $s^{\#}$ is injective. Thus, $(s'_n)^{\#}$ is injective. As s_n is a closed immersion, $(s'_n)^{\#}$ is also surjective and thus an isomorphism; thus, all maps in (6.10) are local isomorphisms above x. This proves the claim as required.

We now prove (1). Consider the diagram (6.10). By Proposition 6.6, $s_n : Z_n \to Y_n$ is a closed immersion. Thus, the restriction $s'_n : Z'_n \to Y'_n$ is also a closed immersion.

On the other hand, Y'_n is the scheme-theoretic image of Υ_n and s is an isomorphism. Thus, the composition $\Upsilon_n s = s'_n \Phi_n$ is scheme-theoretically surjective, so s'_n is scheme-theoretically surjective. But a scheme-theoretically surjective closed immersion is an isomorphism. □

Of course, the defining ideal of a 1-shifted point module also defines a 0-shifted point module, so if the positivity parameter $m = 1$, the conclusions of Theorem 6.8 in fact hold for $m = 0$. In this situation, we will refer to $m = 0$ as a positivity parameter, by slight abuse of notation, since we need only the isomorphism $_l Y_\infty \cong {_l Z_\infty}$ for $l \geq m$ in the sequel.

Corollary 6.11. *Let $l \in \mathbb{N}$, and let $_l Y_\infty$ be the moduli stack of embedded l-shifted S-point modules as above. Then $_l Y_\infty$ is a noetherian fpqc-algebraic stack.*

Proof. We know that $_lY_\infty$ is a sheaf in the fpqc topology by Theorem 5.11. Let m be a positivity parameter. If $l \geq m$, then $_lY_\infty \cong {_lZ_\infty}$ is noetherian fpqc-algebraic by Theorem 5.11.

Suppose then that $l < m$. Let $T : {_lY_\infty} \to {_mY_\infty}$ be the morphism defined by $T(M) := M_{\geq m}$. It is straightforward that T is a morphism of functors and the product

$$\Phi_m \times T : {_lY_\infty} \to {_lY_m} \times_X ({_mY_\infty})$$

is a closed immersion. Since $_mY_\infty$ is noetherian by the first paragraph, it has an fpqc cover $U \to {_mY_\infty}$ by a noetherian affine scheme U. This cover can clearly be lifted and refined to induce an fpqc cover $V \to {_lY_m} \times_X ({_mY_\infty})$ where V is a noetherian affine scheme. But then the Cartesian product

$$
\begin{array}{ccc}
V' & \longrightarrow & {_lY_\infty} \\
\downarrow & \square & \downarrow {\scriptstyle \Phi_m \times T} \\
V & \longrightarrow & {_lY_m} \times_X ({_mZ_\infty})
\end{array}
$$

gives an fpqc cover $V' \to Y_\infty$. Since $\Phi_m \times T$ is a closed immersion, so is $V' \to V$. Thus, V' is isomorphic to a closed subscheme of V and is noetherian and affine. $\square$

Let m be a positivity parameter, and let $l \geq m$. We note that the relevant component of $_lY_\infty \cong {_lZ_\infty}$ is the component containing the $\Bbbk(X)$-point corresponding to the *generic point module* $\Bbbk(X)z^l \oplus \Bbbk(X)z^{l+1} \oplus \cdots$, which is isomorphic to $(Q_{\mathrm{gr}}(S))_{\geq l}$.

7. A coarse moduli space for point modules

In this section, we consider point modules up to module isomorphism in qgr-S and show that the scheme X is a coarse moduli scheme for this functor.

We define the following maps. For any l, let $\Phi : {_lZ_\infty} \to X$ be the map induced from the f_n. Let $\Psi : {_0Z_\infty} \to {_0Z_\infty}$ be the map induced from $\psi_n : {_0Z_n} \to {_0Z_{n-1}}$. Taking the limit of (5.8), we obtain that

$$\Phi\Psi = \sigma\Phi : {_0Z_\infty} \to X.$$

For any noetherian commutative $\Bbbk$-algebra C, there is a graded $(\mathcal{O}_{X_C}, \sigma \times 1)$-bimodule algebra $\mathcal{S}_C$ given by pulling back $\mathcal{S}$ along the projection map $X_C \to X$. Taking global sections gives a functor Gr-$\mathcal{S}_C \to$ Gr-S_C. If $C = \Bbbk$, this induces an equivalence qgr-$\mathcal{S}_C \to$ qgr-S_C by Theorem 2.8. In order to avoid the issues involved with extending this result to bimodule algebras over arbitrary base schemes, we work instead with point modules in Qgr-$\mathcal{S}_C$ and Qgr-S_C.

Let $l \geq m$, where m is a positivity parameter (Definition 6.3), and let F be the moduli functor of (embedded) l-shifted point modules over S as in the previous section. Define an equivalence relation $\sim$ on $F(C)$ by saying that $M \sim N$ if their images are

isomorphic in Qgr-S_C. Define a contravariant functor $G :$ Affine schemes $\to$ Sets by sheafifying, in the fpqc topology, the presheaf G^{pre} of sets Spec $C \mapsto F(C)/\sim$. Let $\mu : F \to G$ be the natural map, $\mathscr{F} \cong {}_l Z_\infty$ the moduli functor of l-shifted point modules over $\mathscr{S}$, and $\mathscr{G} := \mathscr{F}/\sim$ as above. Let $\mu : \mathscr{F} \to \mathscr{G}$ be the natural map.

We recall that $a_n : X_n \to X$ is the blowup of X at $\mathscr{I}_n$ and that by [Corollary 3.3](#) there are morphisms $\alpha : X_{n+1} \to X_n$ that intertwine with the maps a_n.

We briefly discuss point modules over local rings. We note that if C is a local ring of a point of $Z_\infty^o \cong X_\infty$ with maximal ideal $\mathfrak{m}$, then the map $h_{Z_\infty}(C) \to \mathscr{F}(C)$ has a particularly simple form. Let $\zeta : \mathrm{Spec}\, C \to X_\infty$ be the induced morphism. By critical density, there is some $n \geq m$ so that $\zeta_n(\mathfrak{m})$ is not a fundamental point of any of the maps $\alpha^i : X_{i+n} \to X_n$ for any $i > 0$. Let $x_n := \zeta_n(\mathfrak{m})$. Define

$$a_n^{-1}(\mathscr{S}_j) := a_n^{-1}(\mathscr{I}_j) \otimes_{X_n} a_n^* \mathscr{L}_j.$$

Then $a_n^{-1}(\mathscr{S}_j)$ is flat at x_n for all j. Let

$$\mathscr{M} := \bigoplus_{j \geq 0} a_n^{-1}(\mathscr{S}_j)_{x_n}.$$

Then $\mathscr{M}$ is flat over C and is the C-point module corresponding to ζ. The C-action on $\mathscr{M}$ is obvious; to define the $\mathscr{S}$-action on $\mathscr{M}$, let $x := a_n(x_n)$. Then there are maps

$$\mathscr{M}_j \otimes_\Bbbk \mathscr{S}_i(X) \to \mathscr{M}_j \otimes_\Bbbk \mathscr{S}_i^{\sigma^j}(X) \to \mathscr{M}_j \otimes_{\mathbb{O}_{X,x}} (\mathscr{S}_i^{\sigma^j})_x \to a_n^{-1}(\mathscr{S}_{j+i})_{x_n}.$$

This gives a right $\mathscr{S}$-action on $\mathscr{M}$; by letting C act naturally on the left and identifying C with C^{op}, we obtain an action of $\mathscr{S}_C$ on M.

If C is a local ring, we do not know if S_C is necessarily noetherian. However, C-point modules in Qgr-S_C are well behaved as follows.

Lemma 7.1. *Let C be a commutative noetherian local $\Bbbk$-algebra. Let N and M be l-shifted C-point modules with $M \cong N$ in Qgr-S_C. Then for some k, we have $M_{\geq k} \cong N_{\geq k}$.*

Proof. The torsion submodules of M and N are trivial. Thus,

$$\mathrm{Hom}_{\mathrm{Qgr}\text{-}S_C}(M, N) = \varinjlim \mathrm{Hom}_{\mathrm{Gr}\text{-}S_C}(M', N),$$

where the limit is taken over all submodules $M' \subseteq M$ with M/M' torsion. If $M \cong N$ in Qgr-S_C, then there is some submodule $M' \subseteq M$ so that M/M' is torsion and so that there is a homomorphism $f : M' \to N$ so that $\ker f$ and $N/f(M')$ are torsion. Since M and N are torsion-free, we must have $\ker f = 0$.

Thus, it suffices to show that if N is an l-shifted C-point module and $M \subseteq N$ is a graded submodule with $T := N/M$ torsion, then $T_n = 0$ for $n \gg 0$.

Let L be the residue field of C. By assumption, N is C-flat. Thus, there is an exact sequence

$$0 \to \mathrm{Tor}_1^C(T, L) \to M_L \to N_L \to T_L \to 0.$$

By [Artin et al. 1999, Theorem 5.1], the algebra R_L is noetherian. Thus, N_L and T_L are also noetherian. Since T_L is torsion, it is finite-dimensional. Thus, for $n \gg 0$, we have $(T_L)_n = (T_n) \otimes_C L = 0$. By Nakayama's lemma, $T_n = 0$.   □

Lemma 7.2. *Let C be a noetherian local ring, and let $\mathcal{M}$ and $\mathcal{N}$ be C-point modules over $\mathcal{G}$, corresponding to morphisms $f_{\mathcal{M}}, f_{\mathcal{N}} : \mathrm{Spec}\, C \to {}_0 Z_\infty$. If $\mathcal{N} \sim \mathcal{M}$, then there is some k so that $\Psi^k f_{\mathcal{M}} = \Psi^k f_{\mathcal{N}}$.*

Proof. Let m be a positivity parameter, and let $M := s(\mathcal{M}_{\geq m})$ and $N := s(\mathcal{N}_{\geq m})$. Then $M \sim N$; by Lemma 7.1, there is some k, which we may take to be at least m, so that $M_{\geq k} \cong N_{\geq k}$. Since ${}_k Z_\infty \cong {}_k Y_\infty$, we have $\mathcal{M}_{\geq k} \cong \mathcal{N}_{\geq k}$. Thus, the modules $\mathcal{M}[k]_{\geq 0} \cong \mathcal{N}[k]_{\geq 0}$ correspond to the same point of ${}_0 Z_\infty$; that is, $\Psi^k f_{\mathcal{M}} = \Psi^k f_{\mathcal{N}}$. □

We now show that X is a coarse moduli space for ${}_0 Z_\infty / \sim$. In fact, we prove this result in greater generality to be able to analyze the spaces ${}_l Y_\infty$.

Proposition 7.3. *Let $Z_\infty := {}_0 Z_\infty$. Let V_∞ be a closed algebraic substack of Z_∞ so that $X_\infty \subseteq V_\infty \subseteq Z_\infty$, and assume that $V_\infty = \varprojlim V_n$, where $V_n \subset Z_n$ is a closed subscheme that maps into V_{n-1} under $Z_n \to Z_{n-1}$ for all n. Then X is a coarse moduli space for $V_\infty / \sim$. More precisely, let $\mathcal{H}$ be the image of V_∞ under $\mu : Z_\infty \to \mathcal{G}$. Then*

(1) *the morphism $\Phi : V_\infty \to X$ factors via $V_\infty \xrightarrow{\mu} \mathcal{H} \xrightarrow{t} X$ and*

(2) *every morphism $\mathcal{H} \to A$ where A is a scheme (of finite type) factors uniquely through $\mathcal{H} \xrightarrow{t} X$.*

Proof. (1) It suffices to prove that if C is a commutative noetherian ring and $\mathcal{M} \sim \mathcal{N}$ are C-point modules over $\mathcal{G}$, corresponding to maps $f_{\mathcal{M}}, f_{\mathcal{N}} : \mathrm{Spec}\, C \to Z_\infty$, then we have $\Phi f_{\mathcal{M}} = \Phi f_{\mathcal{N}} : \mathrm{Spec}\, C \to X$. To show this, it suffices to consider the case that C is local. By Lemma 7.2, $\Psi^k f_{\mathcal{M}} = \Psi^k f_{\mathcal{N}}$ for some k. Thus, as required,

$$\Phi f_{\mathcal{M}} = \sigma^{-k} \Phi \Psi^k f_{\mathcal{M}} = \sigma^{-k} \Phi \Psi^k f_{\mathcal{N}} = \Phi f_{\mathcal{N}}.$$

(2) Let $\nu : \mathcal{H} \to h_A$ be a natural transformation for some scheme A. For all $n \in \mathbb{N}$, let P_n be the subscheme of X defined by $\mathcal{I}_n$. Fix any closed point $x \in X \setminus \bigcup P_n$; some such x exists since $\Bbbk$ is uncountable. Let $C := \mathcal{O}_{X,x}$. The induced map $\mathrm{Spec}\, C \to X_\infty \to V_\infty$ gives a C-point module $\mathcal{M}_x$ as described above; its $\sim$-equivalence class is an element of $\mathcal{H}(C)$. Applying ν, we therefore have a morphism $\mathrm{Spec}\, C \to A$. This extends to a morphism $g_x : U_x \to A$ for some open subset U_x of X. It follows from critical density of the orbits of points in P that $X \setminus \bigcup P_n$

is quasicompact. Thus, we may take finitely many U_x, say $U_1, \ldots, U_k$, that cover $X \smallsetminus \bigcup P_n$ with maps $g_i : U_i \to A$. These maps all agree on the generic point of X and so agree on overlaps $U_i \cap U_j$.

Let $U := \bigcup_{i=1}^{k} U_i$, and let $g : U \to A$ be the induced map. Then $X \smallsetminus U \subseteq \bigcup P_n$ is a closed subset of X, and so $X \smallsetminus U = \{z_1, \ldots, z_r\}$ for some $z_1, \ldots, z_r \in \bigcup P_n$. Let n be such that for any $i > 0$, the map $\phi^i : Z_{n+i} \to Z_n$ is a local isomorphism at all points in the preimage of $\{z_1, \ldots, z_r\}$.

Let $\Phi_n : Z_\infty \to Z_n$ be the map induced from the ϕ_m. There is an induced map $f_n^{-1}(U) \cap V_n \to U \to A$. Further, for every $y \in V_\infty \smallsetminus \Phi^{-1}(U)$, the map $\operatorname{Spec} \mathcal{O}_{V_\infty, y} \to V_\infty \to \mathcal{H} \to A$ factors through $V_\infty \to V_n$ as Φ_n is a local isomorphism at y. We thus obtain morphisms $\operatorname{Spec} \mathcal{O}_{V_n, y} \to A$ for all $y \in V_n$. Using these, we may extend g to give a morphism $\theta : V_n \to A$ so that the following diagram commutes:

$$
\begin{array}{ccc}
V_\infty & \xrightarrow{\ \mu\ } & \mathcal{H} \\
{\scriptstyle \Phi_n} \downarrow & & \downarrow {\scriptstyle \nu} \\
V_n & \xrightarrow[\ \theta\ ]{} & A \\
{\scriptstyle f_n} \big\downarrow & {\scriptstyle g} \nearrow & \\
U & &
\end{array}
$$

We claim that θ contracts each of the loci $f_n^{-1}(z_j) \cap V_n$ to a point. To see this, let $x, y \in \Phi^{-1}(z_j) \cap V_\infty$, corresponding to maps $f_x, f_y : \operatorname{Spec} \Bbbk \to V_\infty$. We must show that $\theta \Phi_n f_x = \theta \Phi_n f_y$.

Since for $k \gg 0$, $\sigma^k(z_j)$ is not in $\bigcup P_n$, Φ is a local isomorphism at $\Psi^k(\Phi^{-1}(z_j))$. We have

$$\Phi \Psi^k f_x = \sigma^k \Phi f_x = \sigma^k \Phi f_y = \Phi \Psi^k f_y$$

and so $\Psi^k f_x = \Psi^k f_y$. Therefore, $\mu f_x = \mu f_y$, and so

$$\theta \Phi_n f_x = \nu \mu f_x = \nu \mu f_y = \theta \Phi_n f_y$$

as we wanted.

The morphism $\theta : V_n \to A$ thus factors set-theoretically to give a map from X to A. Since X is smooth at all z_i by critical density of the orbits of the z_i, it is well known that θ also factors scheme-theoretically.

Consequently, we have the morphism $X \to A$ that we sought. This proves Proposition 7.3. $\qquad\square$

Theorem 7.4. *Fix a positivity parameter m (Definition 6.3), and let $l \geq m$. Then X is a coarse moduli space for $G = F/\sim$.*

Proof. As above, we let $\mathcal{G}$ denote the functor of l-shifted point modules over $\mathcal{S}$ modulo $\sim$. By Theorem 6.8(2), it is enough to show that X is a coarse moduli

space for $\mathcal{G}$. Let $L : {}_lZ_\infty \to {}_0Z_\infty$ be the map that sends $\mathcal{M} \mapsto \mathcal{M}[l]$. Notice that if $\mathcal{M}$ and $\mathcal{N}$ are l-shifted point modules, then $\mathcal{M} \sim \mathcal{N}$ if and only if $\mathcal{M}[l] \sim \mathcal{N}[l]$. That is, if we let $\mathcal{G}'$ be the functor of (unshifted) point modules over $\mathcal{S}$ modulo $\sim$, then L induces an inclusion $\mathcal{G} \to \mathcal{G}'$ so that the diagram

$$
\begin{array}{ccc}
{}_lZ_\infty & \xrightarrow{\ L\ } & {}_0Z_\infty \\
\mu \downarrow & & \downarrow \mu \\
\mathcal{G} & \xrightarrow[\ L\ ]{} & \mathcal{G}'
\end{array}
$$

commutes. Let $V_n := \mathrm{Im}({}_lZ_{l+n} \to {}_0Z_n)$ and $V_\infty := \varprojlim V_n$, so $V_\infty = L({}_lZ_\infty)$.

Note that L is injective on $X_\infty \subseteq {}_lY_\infty$. Thus, V_∞ satisfies the hypotheses of Proposition 7.3, so X is a coarse moduli scheme for $\mathcal{H} = \mu(V_\infty) \cong \mu({}_lZ_\infty) \cong \mathcal{G}$. $\qquad\square$

Acknowledgements

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 91k:14003 Zbl 0717.14001

[Artin and Zhang 2001] M. Artin and J. J. Zhang, "Abstract Hilbert schemes", *Algebr. Represent. Theory* **4**:4 (2001), 305–394. MR 2002h:16046 Zbl 1030.14003

[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. 1, edited by P. Cartier et al., Progr. Math. **86**, Birkhäuser, Boston, MA, 1990. MR 92e:14002 Zbl 0744.14024

[Artin et al. 1999] M. Artin, L. W. Small, and J. J. Zhang, "Generic flatness for strongly Noetherian algebras", *J. Algebra* **221**:2 (1999), 579–610. MR 2001a:16006 Zbl 0958.16024

[Van den Bergh 1996] M. Van den Bergh, "A translation principle for the four-dimensional Sklyanin algebras", *J. Algebra* **184**:2 (1996), 435–490. MR 98a:16047 Zbl 0876.17011

[EGA IV.2 1965] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, II", *Inst. Hautes Études Sci. Publ. Math.* **24** (1965), 5–231. MR 33 #7330 Zbl 0135.39701

[EGA IV.3 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 36 #178 Zbl 0144.19904

[EGA IV.4 1967] A. Grothendieck, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* **32** (1967), 5–361. MR 39 #220 Zbl 0153.22301

[Eisenbud 1995] D. Eisenbud, *Commutative algebra: With a view toward algebraic geometry*, Graduate Texts in Mathematics **150**, Springer, New York, 1995. MR 97a:13001 Zbl 0819.13001

[Gabriel 1962] P. Gabriel, "Des catégories abéliennes", *Bull. Soc. Math. France* **90** (1962), 323–448. MR 38 #1144 Zbl 0201.35602

[Harris 1992] J. Harris, *Algebraic geometry: A first course*, Graduate Texts in Mathematics **133**, Springer, New York, 1992. MR 93j:14001 Zbl 0779.14001

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Huybrechts and Lehn 1997] D. Huybrechts and M. Lehn, *The geometry of moduli spaces of sheaves*, Aspects of Mathematics **E31**, Friedr. Vieweg & Sohn, Braunschweig, 1997. MR 98g:14012 Zbl 0872.14002

[Keeler 2000] D. S. Keeler, "Criteria for σ-ampleness", *J. Amer. Math. Soc.* **13**:3 (2000), 517–532. MR 2001d:14003 Zbl 0952.14002

[Keeler 2010] D. S. Keeler, "Ample filters and Frobenius amplitude", *J. Algebra* **323**:10 (2010), 3039–3053. MR 2011f:14029 Zbl 1231.13005

[Keeler et al. 2005] D. S. Keeler, D. Rogalski, and J. T. Stafford, "Naïve noncommutative blowing up", *Duke Math. J.* **126**:3 (2005), 491–546. MR 2006g:14005

[Kleiman 1990] S. L. Kleiman, "Multiple-point formulas, II: The Hilbert scheme", pp. 101–138 in *Enumerative geometry* (Sitges, 1987), edited by S. Xambó-Descamps, Lecture Notes in Mathematics **1436**, Springer, Berlin, 1990. MR 92a:14062

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergeb. Math. Grenzgeb. (3) **39**, Springer, Berlin, 2000. MR 2001f:14006 Zbl 0945.14005

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting, line bundles and linear series*, Ergeb. Math. Grenzgeb. (3) **48**, Springer, Berlin, 2004. MR 2005k:14001a

[Nevins and Sierra 2012] T. A. Nevins and S. J. Sierra, "Canonical birationally commutative factors of noetherian graded algebras", in preparation, 2012.

[Rogalski 2004] D. Rogalski, "Generic noncommutative surfaces", *Adv. Math.* **184**:2 (2004), 289–341. MR 2005e:16047 Zbl 1068.16038

[Rogalski and Stafford 2007] D. Rogalski and J. T. Stafford, "Naïve noncommutative blowups at zero-dimensional schemes", *J. Algebra* **318**:2 (2007), 794–833. MR 2009a:16055 Zbl 1141.14001

[Rogalski and Stafford 2009] D. Rogalski and J. T. Stafford, "A class of noncommutative projective surfaces", *Proc. Lond. Math. Soc.* (3) **99**:1 (2009), 100–144. MR 2010j:14007 Zbl 1173.14005

[Rogalski and Zhang 2008] D. Rogalski and J. J. Zhang, "Canonical maps to twisted rings", *Math. Z.* **259**:2 (2008), 433–455. MR 2010b:16056 Zbl 1170.16021

[Sierra 2011] S. J. Sierra, "Geometric idealizer rings", *Trans. Amer. Math. Soc.* **363**:1 (2011), 457–500. MR 2012c:16085 Zbl 1227.16021

[Stafford and Van den Bergh 2001] J. T. Stafford and M. Van den Bergh, "Noncommutative curves and noncommutative surfaces", *Bull. Amer. Math. Soc.* (*N.S.*) **38**:2 (2001), 171–216. MR 2002d:16036 Zbl 1042.16016

[Vistoli 2005] A. Vistoli, "Grothendieck topologies, fibered categories and descent theory", pp. 1–104 in *Fundamental algebraic geometry: Grothendieck's FGA explained*, Math. Surveys Monogr. **123**, American Mathematical Society, Providence, RI, 2005. MR 2223406

nevins@illinois.edu   *Department of Mathematics, University of Illinois at*
                             Urbana–Champaign, 1409 West Green Street, MC-382,
                             Urbana, IL, 61801, United States

s.sierra@ed.ac.uk   *School of Mathematics, The University of Edinburgh,*
                             James Clerk Maxwell Building, The King's Buildings,
                             Mayfield Road, Edinburgh, EH9 3JZ, United Kingdom
                             http://www.maths.ed.ac.uk/~ssierra/

Density of rational points on certain surfaces

Sir Peter Swinnerton-Dyer

Let V be a nonsingular projective surface defined over $\mathbb{Q}$ and having at least two elliptic fibrations defined over $\mathbb{Q}$; the most interesting case, though not the only one, is when V is a K3 surface with these properties. We also assume that $V(\mathbb{Q})$ is not empty. The object of this paper is to prove, under a weak hypothesis, the Zariski density of $V(\mathbb{Q})$ and to study the closure of $V(\mathbb{Q})$ under the real and the p-adic topologies. The first object is achieved by the following theorem:

Let V be a nonsingular surface defined over $\mathbb{Q}$ and having at least two distinct elliptic fibrations. There is an explicitly computable Zariski closed proper subset X of V defined over $\mathbb{Q}$ such that if there is a point P_0 of $V(\mathbb{Q})$ not in X then $V(\mathbb{Q})$ is Zariski dense in V.

The methods employed to study the closure of $V(\mathbb{Q})$ in the real or p-adic topology demand an almost complete knowledge of V; a typical example of what they can achieve is as follows. Let V_c be

$$V_c : X_0^4 + cX_1^4 = X_2^4 + cX_3^4 \quad \text{for } c = 2, 4 \text{ or } 8;$$

then $V_c(\mathbb{Q})$ is dense in $V_c(\mathbb{Q}_2)$ for $c = 2, 4, 8$.

1. *Introduction.* Let V be a nonsingular projective surface defined over $\mathbb{Q}$ and having at least two elliptic fibrations defined over $\mathbb{Q}$; the most interesting case, though not the only one, is when V is a K3 surface with these properties. (By an elliptic fibration we mean a fibration by curves of genus 1; we do not assume that these curves have distinguished points, so they need not be elliptic curves in the number-theoretic sense.) We also assume that $V(\mathbb{Q})$ is not empty. The object of this paper is to prove, under a weak hypothesis, the Zariski density of $V(\mathbb{Q})$ and to study the closure of $V(\mathbb{Q})$ under the real and the p-adic topologies. These results can be found in Section 2; Sections 3 and 4 contain applications and examples. Note that the geometric terminology in this paper is that of Weil.

For the real and Zariski topologies, such results have been proved for a particular family of such surfaces in [Logan et al. 2010], and for one such surface already in [Swinnerton-Dyer 1968]; see Section 3. I am indebted to the referee for drawing

my attention to [Bogomolov and Tschinkel 1998], where Proposition 2.13 proves that such V have the closely related property of potential density, and to René Pannekoek for pointing out a major error in an earlier draft.

2. *The main theorems.* Let $\mathscr{F}$ denote an elliptic fibration of V defined over $\mathbb{Q}$. Then $\mathscr{F}$ is given by a map $\theta : V \to \mathbb{P}^1$ defined over $\mathbb{Q}$ whose generic fibre $\theta^{-1}(t)$ is a curve of genus 1 defined over $\mathbb{Q}(t)$. The Jacobian of $\theta^{-1}(t)$ is also defined over $\mathbb{Q}(t)$. Let P be generic over $\mathbb{Q}(t)$ on this Jacobian and let W be the locus of P over $\mathbb{Q}$; then $P \mapsto t$ gives a natural map $\phi : W \to \mathbb{P}^1$ whose fibre over t is the Jacobian of $\theta^{-1}(t)$, and for any c such that $C = \theta^{-1}(c)$ is nonsingular the Jacobian of C is $J = \phi^{-1}(c)$. In other words, ϕ is the Jacobian elliptic fibration associated to θ.

In this situation C is an n-covering of J for some n, and by abuse of language we can claim that n only depends on the fibration and not on c. (The abuse of language here is that n is not unique.) However, this property by itself does not determine the covering map uniquely. We shall require that the covering map is the restriction to C of a rational map $\psi : V \to W$ defined over $\mathbb{Q}$. One way of ensuring this is as follows. Let P be generic on V, let C be the fibre through P and J the corresponding fibre of W, let $\mathfrak{a}$ be the section of C by a hyperplane Π defined over $\mathbb{Q}$, and write $n = \deg(\mathfrak{a})$. The divisor $nP - \mathfrak{a}$ on C has degree 0 and is defined over $\mathbb{Q}(P)$; so, it is represented by a point R on J, also defined over $\mathbb{Q}(P)$, and we can take ψ to be the map defined over $\mathbb{Q}$ that sends P to R. This ψ is defined except possibly at points of the singular fibres.

Theorem 1. *Let V be a nonsingular surface defined over $\mathbb{Q}$ and having at least two distinct elliptic fibrations. There is an explicitly computable Zariski closed proper subset X of V defined over $\mathbb{Q}$ such that if there is a point P_0 of $V(\mathbb{Q})$ not in X then $V(\mathbb{Q})$ is Zariski dense in V.*

Remarks. Here *explicitly computable* means that if we are given equations for V and the two fibrations θ_1 and θ_2 we can compute equations for X. If P_0 in $V(\mathbb{Q})$ is known, then in fact it is easy to test whether P_0 is in X without needing to calculate X. A similar theorem holds over any algebraic number field. For applications in this more general case, it will often be helpful to use the sufficient condition due to Lutz for a point on an elliptic curve not to be torsion; see [Silverman 1986, Corollary VIII.7.2].

Proof. Denote by $\mathscr{F}_\nu$ for $\nu = 1, 2$ the two given elliptic fibrations on V. We retain the notation introduced above, except that objects associated with $\mathscr{F}_\nu$ will have ν as a subscript. We take $\mu = 3 - \nu$, so that $\{\mu, \nu\} = \{1, 2\}$.

We start by defining a certain Zariski closed set X_ν on V. Let Y'_ν be the union of all the singular fibres of ϕ_ν. By a theorem of Mazur, if Γ is an elliptic curve defined over $\mathbb{Q}$ and R, also defined over $\mathbb{Q}$, is an m-torsion point of Γ, then $m \leq 10$

or $m = 12$. If t is generic on $\mathbb{P}^1$ with $C_\nu = \theta_\nu^{-1}(t)$ and $J_\nu = \phi_\nu^{-1}(t)$, let $\mathfrak{b}_\nu$ be the union of the m-division points on J_ν for these values of m. Thus $\mathfrak{b}_\nu$ is a positive divisor defined over $\mathbb{Q}(t)$, and there is a (reducible) curve Y_ν'' on W_ν defined over $\mathbb{Q}$ and such that the intersection of Y_ν'' with J_ν is $\mathfrak{b}_\nu$. Now let X_ν be the union of the support of $\psi_\nu^{-1}(Y_\nu' \cup Y_\nu'')$ and those fibres of $\mathscr{F}_\nu$ that lie entirely in $\psi_\mu^{-1}(Y_\mu'')$. Clearly X_ν is Zariski closed. Write $X = X_1 \cap X_2$. The effect of this construction is that if P_0 is a point of $V(\mathbb{Q})$ not in X_ν then the fibre $C_{0\nu}$ through P_0 is nonsingular and not in $\psi_\mu^{-1}(Y_\mu'')$, and $\psi_\nu(P_0)$ is not a torsion point of the corresponding $J_{0\nu}$. In applications we can replace X by the Zariski closure of $X(\mathbb{Q})$, which will usually be much smaller; in practice one would normally choose P_0 in $V(\mathbb{Q})$ and then verify that it has the properties stated in the previous sentence for at least one value of ν.

Now let P_0 be a point of V not in X; without loss of generality we can assume that P_0 is not in X_1. For any integer r, denote by P_r the point obtained from P_0 by translation along C_{01} by $[r]\psi_1(P_0)$, where $[r]$ denotes multiplication by r on J_{01}. By construction, the points P_r are distinct and defined over $\mathbb{Q}$, so they are Zariski dense in C_{01}. Moreover C_{01} is not contained in $\psi_2^{-1}(Y_2'')$ by definition, and it is not in $\psi_2^{-1}(Y_2')$ because a nonsingular fibre of $\mathscr{F}_1$ cannot lie in a fibre of $\mathscr{F}_2$, so only finitely many of the P_r lie in X_2. Let C_{r2} be the fibre of $\mathscr{F}_2$ through P_r. The number of P_r that lie on a given fibre of $\mathscr{F}_2$ is bounded, so there are infinitely many C_{r2}. By an argument like that above for C_{01}, if P_r is not in X_2 then $C_{r2}(\mathbb{Q})$ is Zariski dense in C_{r2}. Hence $V(\mathbb{Q})$, which contains all the $C_{r2}(\mathbb{Q})$, is Zariski dense in V. $\qquad\square$

For use in the proof of Theorem 2 below, we note that $J_{01}(\mathbb{Q})$ is not merely Zariski dense in J_{01} but dense in the real topology in that connected component of $J_{01}(\mathbb{R})$ which contains the identity element. Hence $C_{01}(\mathbb{Q})$ is not merely Zariski dense in C_{01} but dense in the real topology in that connected component of $C_{01}(\mathbb{R})$ which contains P_0. A similar remark applies to each C_{r2} for which P_r is not in X_2. Now let Z be the union of the singular points of the singular fibres of $\mathscr{F}_1$ and $\mathscr{F}_2$ and the other points at which the fibres of $\mathscr{F}_1$ and $\mathscr{F}_2$ have intersection multiplicity greater than 1; clearly Z is Zariski closed. Let X be as in Theorem 1 and write $U = V \setminus X$.

Theorem 2. *Let $\mathscr{R}$ denote the closure of $U(\mathbb{Q})$ in $V(\mathbb{R})$ under the real topology. Then the boundary of $\mathscr{R}$ is contained in $Z \cup X$.*

Proof. Let A be a point of $\mathscr{R}$ that is not in $Z \cup X$. The fibres of $\mathscr{F}_1$ and $\mathscr{F}_2$ through A are nonsingular at A and transversal there; by continuity there is a neighbourhood of A in $V(\mathbb{R})$ at every point of which the same properties hold. After contracting this neighbourhood if necessary, we can suppose that the map χ defined by $P \mapsto (\theta_1(P), \theta_2(P))$ is a homeomorphism of this neighbourhood to an

open subset of $\mathbb{P}^1 \times \mathbb{P}^1$ and that the neighbourhood does not meet X; without loss of generality we can suppose that it does not meet X_1. Choose open intervals I_1, I_2 of $\mathbb{P}^1$ such that this open subset contains $I_1 \times I_2$ and $I_1 \times I_2$ contains $\chi(A)$. Since A is in $\mathcal{R}$, there is a point P_0 of $U(\mathbb{Q})$ in $\mathcal{N} = \chi^{-1}(I_1 \times I_2)$, and P_0 is not in X_1. But if C_{01} is the fibre of $\mathcal{F}_1$ through P_0 then $C_{01} \cap \mathcal{N}$ is homeomorphic to I_2 and hence is connected. By the remark that follows the proof of Theorem 1, $C_{01}(\mathbb{Q})$ is dense in $C_{01} \cap \mathcal{N}$. If P' is any point of $C_{01}(\mathbb{Q}) \cap \mathcal{N}$ other than the finitely many which lie in X_2, and if C'_2 is the fibre of $\mathcal{F}_2$ that passes through P', then by the same remark $C'_2(\mathbb{Q}) \cap \mathcal{N}$ is dense in $C'_2 \cap \mathcal{N}$. Hence $V(\mathbb{Q}) \cap \mathcal{N}$ is dense in $\mathcal{N}$, and so is $U(\mathbb{Q}) \cap \mathcal{N}$ since X is nowhere dense in $\mathcal{N}$. In other words, $\mathcal{R} \supset \mathcal{N}$, and therefore A is not in the boundary of $\mathcal{R}$. $\qquad\square$

The same ideas can be applied to the p-adic topologies, though the results that they yield are weaker and the calculations needed to make use of them are more complicated. Let $\mathcal{S}$ be a finite nonempty set of finite primes and let $\mathcal{S}^+$ be the union of $\mathcal{S}$ and the infinite place of $\mathbb{Q}$. For the rest of this section we work in the topology associated with $\mathcal{S}$ or $\mathcal{S}^+$; unfortunately I have not been able to prove any density theorems for the adelic topology.

Lemma 1. *Let*

$$E : \eta^2 + a_1\xi\eta + a_3\eta = \xi^3 + a_2\xi^2 + a_4\xi + a_6$$

be an elliptic curve defined over $\mathbb{Q}_p$, with all the a_i in $p\mathbb{Z}_p$, and suppose further that $4 \mid a_1$ and $4 \mid a_3$ if $p = 2$, and that $9 \mid a_2$ if $p = 3$. Let $\mathbb{Y}_p$ be $\mathbb{Z}_p$ if $p = 2$ or 3 and $p\mathbb{Z}_p$ if $p > 3$, and write $\zeta = \xi/\eta$. Let E' be the union of the point at infinity and the subset of $E(\mathbb{Q}_p)$ for which ξ^{-1} is in $\mathbb{Y}_p$. Then E' is a subgroup of $E(\mathbb{Q}_p)$. Moreover the map

$$E' \to \{\zeta \text{ in } \mathbb{Y}_p\} \tag{1}$$

is one-one and bicontinuous. If $A_0 = (\xi_0, \eta_0)$ is a point of $E'(\mathbb{Q})$ other than the point at infinity, then $E'(\mathbb{Q})$ is dense in that part of $E'(\mathbb{Q}_p)$ in which $v_p(\xi) \le v_p(\xi(A_0))$.

Remark. A less ugly proof can be given by using the theory of formal groups; compare [Silverman 1986, Chapter IV]. The lemma is a weaker and easier version of a result due to René Pannekoek [2012]. I am indebted to him for communicating his result to me.

Proof. We can reduce to the case $a_1 = a_2 = a_3 = 0$ by an obvious linear transformation, which does not change the description of E'. If (ξ, η) is in E' then

$$3v(\xi) = 2v(\eta) \le 0 \text{ and therefore } v(\zeta) = v(\xi/\eta) = -\tfrac{1}{2}v(\xi) \ge 0.$$

The map (1) is defined and continuous. Conversely, if we write $u = \xi^{-1}$ then

$$u = \zeta^2(1 + a_4u^2 + a_6u^3). \tag{2}$$

For ζ in $\mathbb{Y}_p$ this has just one solution u in $\mathbb{Y}_p$, which is a continuous function of ζ with $v(u) = 2v(\zeta)$. Hence, the map (1) is a homeomorphism. If $\tilde{E}$ is the reduction of E mod p, then the image of E' is $\tilde{E}(\mathbb{F}_p) \setminus (0,0)$ if $p = 2$ or 3 and the point at infinity otherwise. This is a group, so the sum or difference of two points of E' is again a point of E'. Hence E' is a subgroup of E.

$\mathbb{Z}$ acts on E' by $P \mapsto [m]P$. We now show that this map is continuous, so that $\mathbb{Z}_p$ acts continuously on E'. Let $P_i = (\xi_i, \eta_i)$ for $i = 1, 2$ be points of E' and write $\zeta_i = \xi_i/\eta_i$ and $n_i = v(\zeta_i)$. The addition formula on E gives

$$\xi(P_1 + P_2) = \alpha^2 - \xi_1 - \xi_2, \quad \eta(P_1 + P_2) = -\alpha(\xi(P_1 + P_2) - \xi_1) - \eta_1, \quad (3)$$

where in general $\alpha = (\eta_1 - \eta_2)/(\xi_1 - \xi_2)$; if $P_1 = P_2$, then $\alpha = (3\xi_1^2 + a_4)/2\eta_1$.

If $p = 2$ and $P_1 = P_2$ then $v(\alpha) = -n_1 - 1$ so that $\zeta([2]P_1) \equiv 2\zeta_1 \bmod 2^{n_1+2}$. If $p = 3$ and $P_1 = P_2$ then $v(\alpha) \geq 1 - n_1$, so that $\xi([2]P_1) \equiv -2\xi_1 \bmod 3^{2-2n_1}$ and $\eta([2]P_1) \equiv -\eta_1 \bmod 3^{2-3n_1}$; thus $\zeta([2]P_1) \equiv 2\zeta_1 \bmod 3^{2+n_1}$. If instead $P_2 = [2]P_1$ then $\alpha \equiv 2\eta_1/3\xi_1 \bmod 3^{-n_1}$, so that $\xi([3]P_1) \equiv \alpha^2 \bmod 3^{-2n_1}$ and $\zeta([3]P_1) \equiv 3\zeta_1 \bmod 3^{2+n_1}$. If $p > 3$ then $\xi_1 \equiv \zeta_1^{-2} \bmod p^{2-2n_1}$ and $\eta_1 \equiv \zeta_1^{-3} \bmod p^{2-3n_1}$. If $P_1 = P_2$ then $\alpha \equiv \frac{3}{2}\zeta_1^{-1} \bmod p^{2-n_1}$ so that $\xi([2]P_1) \equiv \frac{1}{4}\zeta_1^{-2} \bmod p^{2-2n_1}$ and $\eta([2]P_1) \equiv \frac{1}{8}\zeta_1^{-3} \bmod p^{2-3n_1}$; thus $\zeta([2]P_1) \equiv 2\zeta_1 \bmod p^{2+n_1}$. For $3 \leq m \leq p$ we now take $P_2 = [m-1]P_1$; it follows by induction on m that (as p-adic numbers)

$$\alpha(m^2 - m) \equiv (m^2 - m + 1)\zeta_1^{-1} \bmod mp^{2-n_1},$$

$$m^2\xi([m]P_1) \equiv \zeta_1^{-2} \bmod m^2 p^{2-2n_1},$$

$$m^3\eta([m]P_1) \equiv \zeta_1^{-3} \bmod m^3 p^{2-3n_1},$$

and therefore $\zeta([m]P_1) \equiv m\zeta_1 \bmod mp^{2+n_1}$.

It follows by induction on r that $\zeta([p^r]P_1) \equiv p^r\zeta_1 \bmod p^{r+n_1+1}$ for any p and any $r > 0$. Now suppose that $n_2 > n_1$. By (2),

$$\xi_2 \equiv \zeta_2^{-2} \bmod p^{2n_2+2}, \quad \eta_2 \equiv \zeta_2^{-3} \bmod p^{3n_2+2}$$

and therefore $\alpha \equiv \zeta_2^{-1}(1 - \eta_1/\eta_2 + \xi_1/\xi_2) \bmod p^{3n_2-4n_1}$. Hence (as p-adic numbers)

$$\xi(P_1 + P_2) \equiv \xi_1 - 2\eta_1\zeta_2 \bmod 2p^{2n_2-4n_1} \quad (4)$$

by the first equation of (3). Now $\eta(P_1 + P_2) \equiv \eta_1 \bmod p^{-n_2-2n_1}$ by the second equation of (3). This tells us which square root to take in the equation for E, and combining this fact with (4) gives

$$\eta(P_1 + P_2) \equiv \eta_1 \bmod 2p^{n_2-4n_1}. \quad (5)$$

Now (4) and (5), taken with the first sentence of this paragraph, prove that $\mathbb{Z}$ acts continuously on E'. Hence $\mathbb{Z}_p$ acts on E' too, and this action is continuous. If A_0 is as in the lemma and A satisfies $v(\xi) \leq v(\xi(A_0))$, a successive approximation

argument shows that there is an m in $\mathbb{Z}_p$ such that $A = [m]A_0$. If n in $\mathbb{Z}$ is close to m then $[n]A_0$ will be close to A. □

Theorem 3. *Let X_1, X_2 be as in the proof of Theorem 1, and let P_0 be a point of $V(\mathbb{Q})$ not in $X_1 \cup X_2$. Then there is a constructable open neighbourhood $\mathscr{D}$ of P_0 in $\prod V(\mathbb{Q}_p)$, where the product is taken over all p in $\mathscr{S}$, such that $V(\mathbb{Q})$ is dense in $\mathscr{D}$. A similar result holds for $\mathscr{S}^+$.*

Proof. In the notation of Lemma 1, we can rescale the equations of W so that J_{01} satisfies the conditions imposed on E for each p in $\mathscr{S}$, where J_{01} is again the Jacobian of the fibre of $\mathscr{F}_1$ through P_0. For each such p, let r_p be the least positive integer r such that $[r]\psi_1(P_0)$ is in the set J'_{01} derived from J_{01} in the way that E' was derived from E and let r^* be a convenient common multiple of the r_p for p in $\mathscr{S}$. Let $(\xi^*, \eta^*) = [r^*]\psi_1(P_0)$ and let m_p be the exact power of p that divides ξ^*/η^*. Let $\mathscr{M}_{p1}(P_0) \subset J_{01}(\mathbb{Q}_p)$ be the subset of J'_{01} which corresponds to the ζ in $m_p\mathbb{Z}_p$ under the isomorphism corresponding to (1), and let $\mathscr{N}_{p1}(P_0) \subset C_{01}(\mathbb{Q}_p)$ consist of the points obtained from P_0 by translation by an element of $\mathscr{M}_{p1}(P_0)$. If we denote by P_r the translation of P_0 by $[r]\psi_1(P_0)$ then by the Chinese remainder theorem the P_r for which r is a multiple of r^* are distinct and dense in $\prod \mathscr{N}_{p1}(P_0)$ where the product is taken over all p in $\mathscr{S}$.

We can apply a similar construction with $\nu = 2$ to obtain a neighbourhood $\mathscr{A}$ of P_0 in $\prod C_{02}(\mathbb{Q}_p)$ in which points of $V(\mathbb{Q})$ are dense. If in the construction of the previous paragraph we replace P_0 by any point P' in $\mathscr{A}$, everything involved in the construction varies continuously with P'; so after reducing $\mathscr{A}$ if necessary, we can assume that the r_p and the m_p are independent of $\mathscr{A}$. One consequence of this is that if $P' = \prod P'_p$ then $\psi_1(P'_p)$ cannot be a torsion point; and it is now easy to see that we can choose $\mathscr{D}$ to be the union of the $\prod \mathscr{N}_{p1}(P')$ where the union is taken over all points P' of $\mathscr{A}$.

The corresponding result for $\mathscr{S}^+$ now follows easily. For let $\mathscr{D}_0$, containing P_0, be a small open subset of the $\mathscr{D}$ associated with $\mathscr{S}$; by diminishing $\mathscr{D}_0$ we can further assume that it is disjoint from $X_1 \cup X_2$. Let $\mathscr{R}(\mathscr{D}_0)$ be the closure of the projection of $\phi V(\mathbb{Q}) \cap (\mathscr{D}_0 \times V(\mathbb{R}))$ onto $V(\mathbb{R})$, where ϕ is the obvious diagonal map. Then it is enough to show that there is a nonempty open set contained in $\mathscr{R}(\mathscr{D}_0)$ and independent of $\mathscr{D}_0$. Suppose that r^* is even and divisible by a high enough power of every p in $\mathscr{S}$ and by all the r_p defined in the second paragraph of this proof. Then r^* depends only on $\mathscr{D}_0$, and the translation of P_0 by $[r^*]\psi_1(P_0)$ lies in $\mathscr{D}_0$ and is dense in the connected component containing P_0 of the fibre of $\mathscr{F}_1$ through P_0. There is a corresponding result for $\mathscr{F}_2$. So the argument used to prove Theorem 2 still applies in this situation, and the boundary of $\mathscr{R}(\mathscr{D}_0)$ lies in $Z \cup X$, which is independent of $\mathscr{D}_0$. Thus $\mathscr{R}(\mathscr{D}_0)$ contains that connected component of $V(\mathbb{R}) \setminus (Z \cup X)$ in which P_0 lies. □

In what circumstances will the machinery of Theorem 3 enable us to prove that $\overline{V(\mathbb{Q})}$, in the topology generated by $\mathscr{S}$, is at least as large as it actually is? (A similar question for $\mathscr{S}^+$ can be addressed in much the same way; the details are left to the reader.) For this purpose we make the temporary assumption that $\overline{V(\mathbb{Q})}$ is open as well as closed — which would certainly hold if for example the entire obstruction to weak approximation consists of finitely many Brauer–Manin conditions.

Suppose that P_0 is in $V(\mathbb{Q}_p)$ and not on a singular fibre of $\mathscr{F}_1$. Let r_p be the least positive r such that $[r]\psi_1(P_0)$ is in J'_{01} in the notation of the last proof; then $\psi_1(P_0)$ can only be a torsion point of J_{01} if $[r_p]\psi_1(P_0)$ is the identity element of J_{01}. Thus the points P_0 in $V(\mathbb{Q}_p)$ not on a singular fibre of $\mathscr{F}_1$ and for which $\psi_1(P_0)$ is torsion lie on a (reducible) curve Z_{1p}. Let $\mathscr{X}$ be a compact open subset of $\prod V(\mathbb{Q}_p)$ such that if $\prod A_p$ is in $\mathscr{X}$ then no A_p lies in Z_{1p} or on a singular fibre of $\mathscr{F}_1$. By compactness, the values of m_p and r_p associated with A_p are bounded as $\prod A_p$ varies in $\mathscr{X}$; thus if P_0 is a point of $\mathscr{X} \cap V(\mathbb{Q})$ the size of the domain $\mathscr{D}$ associated with P_0 in the sense of Theorem 3 is bounded below. This implies that provided $V(\mathbb{Q})$ is actually dense in $\mathscr{X}$ we can prove this by exhibiting finitely many points of $V(\mathbb{Q})$. What we want to prove for any particular V is of course stronger than this. But these considerations do suggest that it will be easiest to prove that $\overline{V(\mathbb{Q})}$ as a subset of $\prod V(\mathbb{Q}_p)$ is as large as it actually is when there are so many elliptic fibrations $\mathscr{F}_\nu$ that $\bigcap X_\nu(\mathbb{Q})$ is empty. Notice that for specific V we can usually modify the argument so as to make $\mathscr{D}$ substantially larger than the $\mathscr{D}$ constructed above. We illustrate this in Section 4.

3. *An example of Theorems 1 and 2.* The best known examples of K3 surfaces with at least two elliptic pencils are given by equations of the form

$$a_0 X_0^4 + a_1 X_1^4 + a_2 X_2^4 + a_3 X_3^4 = 0 \quad \text{where} \quad a_0 a_1 a_2 a_3 = b^2. \tag{6}$$

Theorems on the density of rational points in the real and the Zariski topologies on such a surface, assuming that it contains at least one rational point lying outside a certain subvariety, already appear in [Logan et al. 2010]. The earliest example of such a result, for the real topology and one special V, appears to be [Swinnerton-Dyer 1968]. The arguments in [Logan et al. 2010] do not appear to extend to the most general V considered in this paper, but the methods of this paper do yield, in the case of (6), a simpler proof of their result, which is stated here as Theorem 4. For this, and for later applications, we need to calculate the exceptional sets X and Z for (6), and for this it is convenient to repeat some formulae from [Swinnerton-Dyer 2000], though the reader is warned that there are some errors in that paper.

Denote by V the K3 surface given by (6). The union of the singular fibres of V is the union of the 48 lines on V. There is an obvious map from V to

$$T: a_0 Y_0^2 + a_1 Y_1^2 + a_2 Y_2^2 + a_3 Y_3^2 = 0. \tag{7}$$

Since $a_0a_1a_2a_3$ is a square, each of the two families of lines on T is defined over $\mathbb{Q}$, and the lifts of these families to V give two elliptic fibrations on V. The fibre through $P_0 = (x_0, x_1, x_2, x_3)$ of one of the fibrations is given by four equations of the form

$$d_{il}X_i^2 + d_{jl}X_j^2 + d_{kl}X_k^2 = 0 \tag{8}$$

where i, j, k, l are 0, 1, 2, 3 in some order. Any three of these equations are linearly dependent. A possible set of values of the coefficients is

$$\left.\begin{aligned}
d_{01} &= b(a_2x_2^4 + a_3x_3^4), & d_{23} &= (b/a_0a_1)d_{01}, \\
d_{02} &= -a_2(a_0a_3x_0^2x_3^2 + bx_1^2x_2^2), & d_{31} &= (b/a_0a_2)d_{02}, \\
d_{03} &= a_3(a_0a_2x_0^2x_2^2 - bx_1^2x_3^2), & d_{12} &= (b/a_0a_3)d_{03},
\end{aligned}\right\} \tag{9}$$

together with $d_{ji} = -d_{ij}$ for all i, j. The d_{ij} are nonzero since the fibre is absolutely irreducible. In these formulae we can permute the subscripts, provided that for an odd permutation we also change the sign of b. The Jacobian of this fibre can be written in the form

$$\eta^2 = (\xi - c_1)(\xi - c_2)(\xi - c_3) \tag{10}$$

where

$$c_1 - c_2 = d_{03}d_{21}, \quad c_2 - c_3 = d_{01}d_{32}, \quad c_3 - c_1 = d_{02}d_{13}. \tag{11}$$

If temporarily we write $P = (\xi, \eta)$, then

$$\xi([2]P) - c_2 = \frac{((\xi - c_2)^2 - d_{01}d_{03}d_{21}d_{23})^2}{4\eta^2}. \tag{12}$$

The map $\psi : V \to W$, which is effectively the canonical map from each fibre to its Jacobian, can be taken to be

$$\eta = d_{12}d_{23}d_{31}X_1X_2X_3/X_0^3, \quad \xi - c_i = d_{ij}d_{ki}X_i^2/X_0^2 \tag{13}$$

where i, j, k is any permutation of 1, 2, 3. This map has degree 4. There is also an identification of the fibre with its Jacobian which maps P_0 to the identity. This satisfies

$$\xi - c_i = d_{0j}d_{0k}\frac{d_{ji}x_jX_j + d_{ki}x_kX_k - d_{0i}x_0X_0}{d_{i0}x_iX_i + d_{j0}x_jX_j + d_{k0}x_kX_k} \tag{14}$$

where again i, j, k is any permutation of 1, 2, 3. The formula for η is too complicated to be useful. But an immediate deduction from the formula for ξ is that the translation along the fibre of either system by the 2-division point $(c_i, 0)$ is obtained by changing the signs of x_0 and x_i. We shall use this in the proof of Theorem 4. To obtain formulae for the other fibration we need only replace b by $-b$ in all the formulae of this paragraph.

Lemma 2. *Suppose that the point A_0 given by* (13) *with x_i for X_i is a rational torsion point, that $x_0 x_1 x_2 x_3 \neq 0$ and that P_0 does not lie on any of the lines of V. Then*

$$4a_j^2 d_{kl}^2 x_j^4 x_k^2 x_l^2 + d_{jk} d_{jl} (a_i x_i^4 + a_j x_j^4)^2 = 0 \tag{15}$$

for some permutation i, j, k, l of $0, 1, 2, 3$. This condition depends only on the value of i.

Proof. The point A_i that is the translation of A_0 by the 2-division point $(c_i, 0)$ on the Jacobian satisfies

$$\xi - c_i = \frac{d_{0j} d_{k0} x_0^2}{x_i^2}, \quad \xi - c_j = \frac{d_{0k} d_{kj} x_k^2}{x_i^2}, \quad \xi - c_k = \frac{d_{0j} d_{jk} x_j^2}{x_i^2} \tag{16}$$

where again i, j, k is a permutation of $1, 2, 3$. We can restore the symmetry by writing

$$c_i = c_{i0} = c_{0i} = c_{jk} = c_{kj};$$

thus we have $c_{ik} - c_{jk} = d_{ij} d_{kl}$ for any even permutation i, j, k, l of $0, 1, 2, 3$. The point A_i is now given by

$$\xi - c_{ij} = d_{jk} d_{lj} x_j^2 / x_i^2 \tag{17}$$

where i, j, k, l is any permutation of $0, 1, 2, 3$; in particular A_i is not a 2-division point. Such a value of ξ corresponds to two distinct rational values of η, so we appear to have constructed twelve torsion points on (10). If these are not all distinct then at least two of the four values of ξ given by (17) must be equal. Suppose for example that the values of ξ given by $i = 0$ and $i = 1$ are equal; then by comparing the two formulae for $\xi - c_1$ we obtain $a_0 x_0^4 + a_1 x_1^4 = 0$, which implies that P_0 lies on one of the lines of V.

   Suppose instead that the twelve torsion points exhibited above are all distinct, so that by Mazur's theorem they form a group of order 12. Without loss of generality we can assume that one of the two 3-division points is A_0. The condition that (ξ, η) is a 3-division point of (10) can be written in the form

$$4\eta^2 (\xi - c_1) = \{(\xi - c_1)^2 - (c_1 - c_2)(c_1 - c_3)\}^2,$$

which reduces by (10) and (11) to

$$4a_1^2 d_{23}^2 x_1^4 x_2^2 x_3^2 + d_{12} d_{13} (a_0 x_0^4 + a_1 x_1^4)^2 = 0.$$

Symmetry now completes the proof of the lemma. $\qquad\square$

Theorem 4. *Suppose that V is given by* (6). *If there is a rational point $P_0 = (x_0, x_1, x_2, x_3)$ on V with $x_0 x_1 x_2 x_3 \neq 0$ and not lying on one of the lines of V, then $V(\mathbb{Q})$ is Zariski dense on V and dense in $V(\mathbb{R})$.*

Proof. We retain the notation in the proof of Theorem 1. To deduce Zariski density from Theorem 1 we need to show that P_0 does not lie in X. The first step is to show that the fibre of $\mathcal{F}_\nu$ is not contained in $\psi_\mu^{-1}(Y_\mu'')$ — in other words, that if P' is generic on the fibre of $\mathcal{F}_\nu$ through P_0 then $\psi_\mu(P')$ is not a torsion point. Suppose otherwise; then the exact order of this torsion point cannot depend on the choice of P', provided that P' does not lie on a singular fibre of $\mathcal{F}_\mu$. By the proof of Lemma 2 this exact order must be 3 or 6. But by changing the signs of a suitable pair of x_i we can alter this order from 3 to 6 or from 6 to 3, and this is a contradiction.

It is now enough to show that P_0 does not lie in $(\psi_1^{-1}Y_1'')\cap(\psi_2^{-1}Y_2'')$; by Lemma 2 this is equivalent to showing that it does not satisfy (15) for both b and $-b$. Suppose otherwise; without loss of generality we can assume that for $b > 0$ it satisfies (15) with $i = 0$ and that a_0, a_2 are positive and a_1, a_3 negative. In (15) we take $j = 1, k = 2, l = 3$; by (9) we obtain

$$a_0 a_2 a_3 x_0^4 x_2^2 x_3^2 + a_2 b x_0^2 x_1^2 x_2^4 - a_3 b x_0^2 x_1^2 x_3^4 + 3a_1 a_2 a_3 x_1^4 x_2^2 x_3^2 = 0. \qquad (18)$$

On the left the first term is negative and the other three are all positive; so $a_0 a_2 x_0^2 x_2^2 > b x_1^2 x_3^2$ and $a_0 x_0^4 > -3a_1 x_1^4$. The first of these, taken with (6), implies that $a_0 x_0^4$ and $a_2 x_2^4$ lie between $-a_1 x_1^4$ and $-a_3 x_3^4$; so

$$-a_3 x_3^4 > (a_0 x_0^4 \text{ and } a_2 x_2^4) > -a_1 x_1^4. \qquad (19)$$

We can argue similarly for any one of the four equations (15) with $-b$ for b; the only one which yields a condition compatible with (19) is the one with $i = 2$. Remembering that we can cyclically permute the subscripts $1, 2, 3$ in (18), we therefore have

$$a_0 a_3 a_1 x_0^4 x_3^2 x_1^2 + a_3 b x_0^2 x_2^2 x_3^4 - a_1 b x_0^2 x_2^2 x_1^4 + 3a_2 a_3 a_1 x_2^4 x_3^2 x_1^2 = 0,$$
$$3a_0 a_3 a_1 x_0^4 x_3^2 x_1^2 + a_3 b x_0^2 x_2^2 x_3^4 - a_1 b x_0^2 x_2^2 x_1^4 + a_2 a_3 a_1 x_2^4 x_3^2 x_1^2 = 0.$$

The difference of these is $2a_1 a_3 x_1^2 x_3^2 (a_0 x_0^4 - a_2 x_2^4) = 0$. But now either of the last two displayed equations implies that $a_3 x_3^4 / a_1 x_1^4 = (2 + \sqrt{5})^2$, so that P_0 would not be rational. This completes the proof of Zariski density.

To deduce density in $V(\mathbb{R})$ from Theorem 2 we still need to find the points at which nonsingular fibres of $\mathcal{F}_1$ and $\mathcal{F}_2$ touch. These fibres are the pull-backs of the lines on (7), on which two lines of opposite systems are always transversal; so the points at which fibres of $\mathcal{F}_1$ and $\mathcal{F}_2$ touch are just those which have multiplicity greater than 1 in the pull-back of a point of (7). These are just those points of V at which at least one of the x_i vanishes. Thus a maximum connected component in $V(\mathbb{R})$ of the complement of Z consists of all the points at which the signs of the X_i/X_j take assigned values; and since we can vary the signs of the x_i, if one of

these connected components meets $V(\mathbb{Q})$ then each of them does. Thus the closure of $V(\mathbb{Q})$ contains the whole of $V(\mathbb{R}) \setminus Z$, and therefore the whole of $V(\mathbb{R})$. □

4. *An example of Theorem 3.* In this section we treat two particular examples of Theorem 3, for each of which V has the form

$$a_0 X_0^4 + a_1 X_1^4 = a_0 X_2^4 + a_1 X_3^4 \quad \text{with } a_0, a_1 \text{ positive.} \tag{20}$$

In constructing examples to illustrate the machinery of Theorem 3, we have to choose a suitable $\mathscr{S}$. It is not unreasonable to hope that the only places involved in describing the obstruction to weak approximation on a given V are those at which V has bad reduction and possibly the infinite place. (The corresponding result for nonsingular cubic surfaces is known; see [Swinnerton-Dyer 2001]. An analogous result for the Brauer–Manin obstruction on arbitrary diagonal quartic surfaces defined over $\mathbb{Q}$ is due to Colliot-Thélène and Skorobogatov [2013].) If we want V to have the form (6) and to have bad reduction only at 2, then there are just seven such V which are everywhere locally soluble, and there are just two of them for which the arithmetic part of the Brauer–Manin obstruction is trivial. (See Bright's table [2002, Appendix A].) These can conveniently be written as

$$V_c : \quad X_0^4 + c X_1^4 = X_2^4 + c X_3^4 \quad \text{for } c = a_1/a_0 = 2, 4 \text{ or } 8.$$

Here V_2 and V_8 are the same, but by considering both we can confine ourselves to the study of points for which the X_i are integers with X_0, X_2 odd. In Bright's table V_4 is case A75, and Ieronymou, Skorobogatov and Zarhin [Ieronymou et al. 2011] have shown that it has no transcendental Brauer–Manin obstruction either. V_2 and V_8 are instances of Bright's case A104. We shall prove

Theorem 5. $V_c(\mathbb{Q})$ *is dense in* $V_c(\mathbb{Q}_2)$ *for* $c = 2, 4, 8$.

There is a substantial disparity between the numbers of essentially distinct nontrivial solutions of V_2 and of V_4: for example, V_4 has 599 such solutions of height less than 25000, the smallest having height 9, whereas V_2 has 43 such solutions of height less than 25000, the smallest having height 139. This is perhaps accounted for by the fact that the Nèron–Severi group of V_4 has rank 9 whereas that of V_2 has rank 6.

The general surface (20) is case A45 in Bright's table. This surface contains four rational lines. If it has other parametric solutions, which I doubt, even the simplest nonsingular one has degree at least 17. It has three kinds of elliptic pencils of low degree, where the degree means the degree of the general curve of the pencil. Apart from these, which are described below, if there are any other elliptic pencils whose general fibre is nonsingular, they have degree at least 19.

There are four elliptic pencils of degree 3, the curves of such a pencil being those which lie in a plane through one of the rational lines and are residual to that line.

A typical such pencil, which we shall call $\mathcal{F}_3$, consists of the curves given by

$$\left. \begin{array}{r} X_0 - X_2 = \lambda(X_1 - X_3), \\ \lambda a_0(X_0^3 + X_0^2 X_2 + X_0 X_2^2 + X_2^3) + a_1(X_1^3 + X_1^2 X_3 + X_1 X_3^2 + X_3^3) = 0, \end{array} \right\} \quad (21)$$

where λ is a parameter. This curve contains the rational point $(\lambda, 1, -\lambda, -1)$ and can therefore be taken to be its own Jacobian. In general it has no rational torsion points, but if $\lambda = -a_1/a_0\alpha^3$ then it contains the point $(\alpha, 1, \alpha, 1)$, which is a 2-division point. By means of the transformation

$$\left. \begin{array}{ll} X_0 = -a_1 X + \lambda Y, & X_1 = a_0\lambda(a_1^2 - a_0^2\lambda^8)Z + a_0\lambda^3 X + Y, \\ X_2 = -a_1 X - \lambda Y, & X_3 = a_0\lambda(a_1^2 - a_0^2\lambda^8)Z + a_0\lambda^3 X - Y, \end{array} \right\} \quad (22)$$

whose inverse is

$$X = -\frac{X_0 + X_2}{2a_1}, \quad Y = \frac{X_1 - X_3}{2}, \quad Z = \frac{X_1 + X_3 - 2a_0\lambda^3 X}{2a_0\lambda(a_1^2 - a_0^2\lambda^8)}, \quad (23)$$

we can take this curve into the canonical form

$$Y^2 Z = X^3 - 3a_0^2\lambda^6 X^2 Z - 3\lambda^4 a_0^2\theta X Z^2 - a_0^2\lambda^2\theta^2 Z^3 \quad (24)$$

where $\theta = a_1^2 - a_0^2\lambda^8$. If we write $X = (U + a_0^2\lambda^6)Z$ and $Y = WZ$, this becomes

$$W^2 = U^3 - 3a_0^2 a_1^2\lambda^4 U - a_0^2 a_1^2\lambda^2(a_1^2 + \lambda^8 a_0^2). \quad (25)$$

$\mathcal{F}_3$ is the only elliptic pencil on (20) that we shall use in the argument that follows. However, for possible applications elsewhere we record some useful information about the other pencils of low degree on (20).

There are two elliptic pencils of degree 4, which were described for the more general surface (6) in Section 3. There are four elliptic pencils of degree 6; a typical one, parametrized by α, consists of the curves which are the intersections of the cubic surfaces

$$a_0(X_0^3 + X_0^2 X_2 + X_0 X_2^2 + X_2^3) + \alpha(X_0 X_3 - X_1 X_2)(X_1 - X_3) = 0,$$
$$a_1(X_1^3 + X_1^2 X_3 + X_1 X_3^2 + X_3^3) - \alpha(X_0 X_3 - X_1 X_2)(X_0 - X_2) = 0,$$

residual to the three common lines $X_0 = \kappa X_2$, $X_1 = \kappa X_3$ where $\kappa = -1$ or $\kappa = \pm\sqrt{-1}$. This curve is a 2-covering of its Jacobian, and in canonical form it can be written

$$Y^2 = \alpha^3(a_0 + a_1 X^4) + a_0 a_1(3\alpha^2 - a_0 a_1)X^2$$

where

$$X = \frac{X_1 + X_3}{X_0 + X_2}, \quad Y = \frac{a_1(a_0 + \alpha X^2)(X_1 + X_3)}{X_0 - X_2}.$$

We can recover the original variables by writing

$$X_0 = Y + a_1 X (a_0 + \alpha X^2), \quad X_1 = XY - a_0(\alpha + a_1 X^2),$$
$$X_2 = Y - a_1 X (a_0 + \alpha X^2), \quad X_3 = XY + a_0(\alpha + a_1 X^2).$$

The proof of Theorem 5 consists of a number of steps, each of which is described in a separate lemma. Throughout, v will be the normalized 2-adic valuation on $\mathbb{Q}_2$ and $v(c) = \gamma$. In Lemmas 3 and 4, $P_0 = (x_0, x_1, x_2, x_3)$ will be a point of $V_c(\mathbb{Q}_2)$, with the x_i odd; it then follows from the second equation of (21) that

$$v(\lambda) + v(x_0 + x_2) = \gamma + v(x_1 + x_3). \tag{26}$$

The next two lemmas, taken together, provide for V_c a quantitative version of Theorem 3. Let $\mathscr{F}_3$ be the fibration of degree 3 described by (21).

Lemma 3. *Let P_0 be in $V_c(\mathbb{Q})$ with all x_i odd, and let F_3 be the fibre of $\mathscr{F}_3$ through P_0.*

(i) *If $2 \parallel (x_0 - x_2)$ and $2 \parallel (x_1 - x_3)$, then $F_3(\mathbb{Q})$ is dense in that part of $F_3(\mathbb{Q}_2)$ in which the X_i are odd with $2 \parallel (X_1 - X_3)$ and $v(X_0 + X_2) \geq v(x_0 + x_2)$.*

(ii) *If $2 \parallel (x_0 + x_2)$ and $2 \parallel (x_1 - x_3)$, then $F_3(\mathbb{Q})$ is dense in that part of $F_3(\mathbb{Q}_2)$ in which the X_i are odd.*

Proof. Let $\mathscr{D}$ be that part of $F_3(\mathbb{Q}_2)$ in which we have to prove that $F_3(\mathbb{Q})$ is dense. In the notation of (22) and (23), let x, y, z be the values of X, Y, Z at P_0 and let $\lambda = (x_0 - x_2)/(x_1 - x_3)$ be the value of the parameter in (21) which defines F_3. We can take $a_0 = 1$, $a_1 = c$.

In case (i), we have $v(\lambda) = 0$; it now follows that $v(X_1 + X_3) = v(2a_0\lambda^3 X)$ and $v(Z) > v(X)$. Thus $3v(X/Z) = v(Y/X) \leq v(y/x) = 3v(x/z) < 0$, where the equalities follow from (24). The conclusion now follows from Lemma 1.

In case (ii), we have $v(\lambda) = \gamma + \delta - 1$ by (26), where $\delta = v(x_1 + x_3) \geq 2$; so if we write $\xi = 2^{-2\gamma} X/Z$ and $\eta = 2^{-3\gamma} Y/Z$, the curve F_3 takes the form $\eta^2 = \xi^3 + \cdots$ with coefficients satisfying the conditions of Lemma 1. At P_0 we have $v(x) = -\gamma$, $v(y) = 0$, $v(z) = -3\gamma$, where the last result again follows from (26); so $v(\xi) = 0$. Hence, by Lemma 1, $F_3(\mathbb{Q})$ is dense in that part of $F_3(\mathbb{Q}_2)$ in which $v(\xi) \leq 0$. But in $\mathscr{D}$ we have

$$v(X_0 + X_2) = v(X_1 + X_3) + \gamma - v(\lambda) = v(X_1 + X_3) + 1 - \delta.$$

Hence $v(X_1 + X_3) < v(2\lambda^3 X) = v(X_0 + X_2) + 2\gamma + 3\delta - 3$, so that

$$v(Z) = v(X_1 + X_3) - v(\lambda) - 2\gamma - 1 = v(X_0 + X_2) - 3\gamma - 1$$

whence $v(\xi) = v(X_0 + X_2) - v(Z) - 3\gamma - 1 = 0$. $\qquad\square$

Lemma 4. *Suppose there is a point P_0 of $V_c(\mathbb{Q})$ with all x_i odd and $2^2 \parallel (x_1 + x_3)$. Let $\mathcal{D}_1$ be the part of $V_c(\mathbb{Q}_2)$ in which the X_i are odd. Then $V_c(\mathbb{Q})$ is dense in $\mathcal{D}_1$.*

Proof. Let $Q = (u_0, u_1, u_2, u_3)$ be a point of $\mathcal{D}_1$; after changing the signs of u_2 and/or u_3 if necessary, we can assume that $2 \parallel (u_0 - u_2)$ and $2 \parallel (u_1 - u_3)$, and we write $\lambda = (u_0 - u_2)/(u_1 - u_3)$. Similarly we can assume that $2 \parallel (x_0 - x_2)$. Since $v(\lambda) = 0$, the analogue for Q of (26) shows that $v(u_0 + u_2) \geq \gamma + 2$. Now let F_3' be the fibre of $\mathscr{F}_3$ through Q. Write $\mu = (x_0 + x_2)/(x_1 - x_3)$ and let F_3'' be the intersection of V_c with $X_0 + X_2 = \mu(X_1 - X_3)$ residual to the line $X_0 + X_2 = X_1 - X_3 = 0$; thus P_0 lies on F_3''.

Let $R = (y_0, y_1, y_2, y_3)$ be a point of $F_3' \cap F_3''$. If we can ensure that the y_i are odd elements of $\mathbb{Z}_2$ with $2 \parallel (y_0 - y_2)$, $2 \parallel (y_1 - y_3)$ and $v(u_0 + u_2) \geq v(y_0 + y_2)$, we can argue as follows. By Lemma 3(ii) with the sign of X_2 reversed, $F_3''(\mathbb{Q})$ is dense in that part of $F_3''(\mathbb{Q}_2)$ in which all the X_i are odd integers; hence there is a point R' of $F_3''(\mathbb{Q})$ arbitrarily close to R. Let F_3^* be the fibre of $\mathscr{F}_3$ through R', so that it is arbitrarily close to F_3'. By Lemma 3(i) $F_3^*(\mathbb{Q})$ is dense in that part of $F_3^*(\mathbb{Q}_2)$ in which the X_i are odd with $2 \parallel (X_1 - X_3)$ and $v(X_0 + X_2) \geq v(y_0 + y_2)$, and in particular in that part of $F_3^*(\mathbb{Q}_2)$ which is close to Q. Hence there is a point of $V_c(\mathbb{Q})$ close to Q.

It remains to prove the assertions about R. The points of $F_3' \cap F_3''$ are those points (y_0, y_1, y_2, y_3) of V_c at which

$$y_0 - y_2 = \lambda(y_1 - y_3), \quad y_0 + y_2 = \mu(y_1 - y_3), \quad y_1 - y_3 \neq 0,$$

where $v(\lambda) = 0$, $v(\mu) = \gamma + \delta - 1$. Write $y_1 = y_3 + 2\eta$; then $y_0 = (\lambda + \mu)\eta$ and $y_2 = (\mu - \lambda)\eta$, so that the equation for R becomes

$$\lambda\mu(\lambda^2 + \mu^2)\eta^3 + c(y_3 + \eta)(y_3^2 + 2y_3\eta + 2\eta^2) = 0.$$

By Hensel's lemma, this has a solution in $\mathbb{Z}_2$ with y_3, η odd; and for this solution $v(y_3 + \eta) = \delta - 1$ and thus $v(y_1 + y_3) = \delta \leq v(u_1 + u_3)$. □

Suitable points P_0 exist for each of the three values of c. We can for example take P_0 to be $(849, 653, -969, 167)$ when $c = 2$, $(189, 677, -557, -657)$ when $c = 4$, and $(1197, 177, 499, 707)$ when $c = 8$.

The next two lemmas again provide a quantitative version of Theorem 3. In them we denote by $\mathcal{D}_3$ the part of $V_c(\mathbb{Q}_2)$ in which X_0, X_2 are odd and X_1, X_3 even, with just one of X_1, X_3 divisible by 4; and $P_0 = (x_0, x_1, x_2, x_3)$ will be a point of $\mathcal{D}_3$ defined over $\mathbb{Q}$. As before let x, y, z be the values of X, Y, Z at P_0 and let $\lambda = (x_0 - x_2)/(x_1 - x_3)$ be the value of the parameter in (21) which defines F_3. We can take $a_0 = 1$, $a_1 = c$.

Lemma 5. *Let F_3 be the fibre of $\mathscr{F}_3$ through P_0, where $2 \parallel (x_0 - x_2)$. Then $F_3(\mathbb{Q})$ is dense in $F_3 \cap \mathcal{D}_3$.*

Proof. Since $v(\lambda) = 0$, we have $v(X_0 - X_2) = 1$ at each point of $F_3 \cap \mathcal{D}_3$. Thus at each such point $v(X_0 + X_2) = \gamma + 2$ and $v(X) = 1$, $v(Y) = 0$, $v(Z) = 0$. In the notation of (25) we have $v(U) = v(W) = 0$; and (25) satisfies the conditions of Lemma 1. □

Lemma 6. *Suppose that there is a point P_0 in $\mathcal{D}_3$ defined over $\mathbb{Q}$. Then $V_c(\mathbb{Q})$ is dense in $\mathcal{D}_3$.*

Proof. Let $Q = (u_0, u_1, u_2, u_3)$ be a point of $\mathcal{D}_3$; after changing the sign of x_2 and/or u_2 if necessary, we can assume that $2 \parallel (x_0 - x_2)$ and $2 \parallel (u_0 - u_2)$. Write $\lambda = (u_0 - u_2)/(u_1 - u_3)$ and $\mu = (x_0 - x_2)/(x_1 + x_3)$. Let F_3' be the fibre of $\mathcal{F}_3$ through Q and F_3'' the intersection of V_c with $X_0 - X_2 = \mu(X_1 + X_3)$ residual to the line $X_0 - X_2 = X_1 + X_3 = 0$; thus P_0 lies on F_3''.

Let $R = (y_0, y_1, y_2, y_3)$ be a point of $F_3' \cap F_3''$. If we write $y_0 + y_3 = \chi(y_0 - y_2)$ and use $y_0 - y_2 = \lambda(y_1 - y_3) = \mu(y_1 + y_3)$, we find that χ satisfies

$$\chi(\chi^2 + 1) = c(\lambda^2 + \mu^2)(\lambda\mu)^{-3}.$$

Since $v(\lambda) = v(\mu) = 0$, this equation is satisfied by one value of χ in $\mathbb{Q}_2$ with $v(\chi) = \gamma + 1$, and the corresponding R is in $\mathcal{D}_3$. The lemma now follows by the same argument as in the second paragraph of the proof of Lemma 4. □

Suitable points P_0 exist for each of the three values of c. We can for example take P_0 to be $(489, 684, -577, 662)$ for $c = 2$, $(61, 168, -237, 58)$ for $c = 4$, and $(257, 22, 223, 124)$ for $c = 8$.

Henceforth, for any integer $\beta > 0$ let $\mathcal{D}_4^\beta$ consist of those points of $V_c(\mathbb{Q}_2)$ with X_0, X_2 odd and $v(X_1) = v(X_3) = \beta$ and let $\mathcal{D}_5^\beta$ consist of those points of $V_c(\mathbb{Q}_2)$ with X_0, X_2 odd and $v(X_1) > v(X_3) = \beta$ or $v(X_3) > v(X_1) = \beta$. Thus $\mathcal{D}_5^1$ is what we have previously called $\mathcal{D}_3$. To complete the proof of Theorem 5 we need to prove that $V_c(\mathbb{Q})$ is dense in each $\mathcal{D}_4^\beta$ with $\beta > 0$ and each $\mathcal{D}_5^\beta$ with $\beta > 1$.

Lemma 7. *If $V_c(\mathbb{Q})$ is dense in $\mathcal{D}_1$ then it is dense in $\mathcal{D}_5^\beta$ for each $\beta > 1$.*

Proof. Let $Q = (u_0, u_1, u_2, u_3)$ be a point of $\mathcal{D}_5^\beta$ where $\beta > 1$; after changing the sign of u_2 if necessary, we can assume that $2 \parallel (u_0 - u_2)$ and therefore $v(u_0 + u_2) = 4\beta - 2$. Let $\lambda = (u_0 - u_2)/(u_1 - u_3)$, so that $v(\lambda) = 1 - \beta$. At Q we have $v(X) = 4\beta - 3$, $v(Y) = \beta - 1$ and therefore $v(Z) = 10(\beta - 1)$; hence also $v(U) = 6(1 - \beta)$, $v(W) = 9(1 - \beta)$. If we write $\xi = 2^{4(\beta-1)}U$, $\eta = 2^{6(\beta-1)}W$ in (25), then the equation between ξ and η satisfies the conditions of Lemma 1. In the notation of that lemma, let $R = (y_0, y_1, y_2, y_3)$ be the point of E' which satisfies $\zeta(R) = 2^{1-\beta}\zeta(Q)$; R exists because $v(\zeta(Q)) = \beta - 1$. Thus $v(\xi/\eta) = 0$ at R, and so $v(\xi) = v(\eta) = 0$ since R is in E'. Moreover if we choose the representation of R so that $v(y_0 - y_2) = 1$, we have $v(y_1 - y_3) = 1 - v(\lambda) = \beta$; so $v(Y) = \beta - 1$, $v(Z) = 7(\beta - 1)$, $v(X) = \beta - 1$,

$v(y_0 + y_2) = \beta + \gamma > 1$ at R. The results for $y_0 \pm y_2$ show that y_0 and y_2 are odd integers. If we assume that $v(y_1 + y_3) \le 0$, we find that

$$\beta + \gamma + 2 = v(y_0^4 - y_2^4) = \gamma + v(y_1^4 - y_3^4) \le \beta + \gamma - 1,$$

and if instead we assume that $v(y_1 + y_3) \ge 2$, we find that

$$\beta + \gamma + 2 = v(y_0^4 - y_2^4) = \gamma + v(y_1^4 - y_3^4) \ge \beta + \gamma + 4,$$

both of which are absurd. So we must have $v(y_1 + y_3) = 1$, whence R is in $\mathcal{D}_1$. Hence there is a point R' in $V_c(\mathbb{Q})$ arbitrarily close to R. Let $Q' = [2^{\beta-1}]R'$, where the operation is to be carried out on the fibre of $\mathcal{F}_3$ through R'; then Q' is in $V_c(\mathbb{Q})$ and arbitrarily close to Q.   □

Lemma 8. *If $V_c(\mathbb{Q})$ is dense in $\mathcal{D}_1$ then it is dense in $\mathcal{D}_4^\beta$ for each $\beta > 0$.*

Proof. Let $Q = (u_0, u_1, u_2, u_3)$ be a point of $\mathcal{D}_4^\beta$ where $\beta > 0$; after changing the signs of u_2 and/or u_3 if necessary, we can assume that $v(u_0 - u_2) = 1$ and $v(u_1 - u_3) = \beta + 1$. Thus $v(u_1 + u_3) = \beta + \delta$ for some $\delta > 1$, whence $v(u_0 + u_2) = 4\beta + \gamma + \delta$. Let $\lambda = (u_0 - u_2)/(u_1 - u_3)$, so that $v(\lambda) = -\beta$. At Q we have $v(X) = 4\beta + \delta - 1$, $v(Y) = \beta$ and therefore $v(Z) > 10\beta + \delta - 1$ and it now follows from (24) that $v(Z) = 10\beta + 3\delta - 3$. Hence $v(U) = -6\beta - 2\delta + 2$ and $v(W) = -9\beta - 3\delta + 3$. If we write $\xi = 2^{4\beta}U$, $\eta = 2^{6\beta}W$ in (25) then the equation between ξ and η satisfies the conditions of Lemma 1. In the notation of that lemma, let $R = (y_0, y_1, y_2, y_3)$ be the point of E' which satisfies $\zeta(R) = 2^{1-\beta-\delta}\zeta(Q)$; R exists because $v(\zeta(Q)) = \beta + \delta - 1$. Thus $v(\xi/\eta) = 0$ at R, and so $v(\xi) = v(\eta) = 0$ since R is in E'. Moreover if we choose the representation of R so that $v(y_0 - y_2) = 1$ we have $v(y_1 - y_3) = 1 - v(\lambda) = 1 + \beta$; so $v(Y) = \beta$, $v(Z) = 7\beta$, $v(X) = \beta$, $v(y_0 + y_2) = \beta + \gamma + 1 > 2$ at R. The results for $y_0 \pm y_2$ show that y_0 and y_2 are odd integers. If we assume that $v(y_1 + y_3) \le 0$, we find that

$$\beta + \gamma + 3 = v(y_0^4 - y_2^4) = \gamma + v(y_1^4 - y_3^4) \le \beta + \gamma,$$

and if instead we assume that $v(y_1 + y_3) \ge 2$, we find that

$$\beta + \gamma + 3 = v(y_0^4 - y_2^4) = \gamma + v(y_1^4 - y_3^4) \ge \beta + \gamma + 5,$$

both of which are absurd. So we must have $v(y_1 + y_3) = 1$, whence R is in $\mathcal{D}_1$. Hence there is a point R' in $V_c(\mathbb{Q})$ arbitrarily close to R. Let $Q' = [2^{\beta+\delta-1}]R'$, where the operation is to be carried out on the fibre of $\mathcal{F}_3$ through R'; then Q' is in $V_c(\mathbb{Q})$ and arbitrarily close to Q.   □

Proof of Theorem 5. $V_c(\mathbb{Q})$ is dense in $\mathcal{D}_1$ by Lemma 4, and dense in $\mathcal{D}_3$ by Lemma 6. It is dense in the rest of that part of $V_c(\mathbb{Q}_2)$ in which the X_i are integers with X_0, X_2 odd, by Lemmas 7 and 8. This is all we need.   □

Very little extra effort is needed to take account also of the infinite place.

Corollary. *The image of $V_c(\mathbb{Q})$ is dense in $V_c(\mathbb{Q}_2) \times V_c(\mathbb{R})$.*

The argument in the last paragraph of the proof of Theorem 3 still works, and it shows that the image of $V_c(\mathbb{Q})$ is dense in at least one of the $V_c(\mathbb{Q}_2) \times \mathscr{R}$, where $\mathscr{R}$ is one of the connected components of $V_c(\mathbb{R}) \setminus (Z \cup X)$. As in the proof of Theorem 4, each of these connected components is given by fixing the signs of the X_i / X_j, and because we can change the sign of any x_i the corollary follows.

References

[Bogomolov and Tschinkel 1998] F. A. Bogomolov and Y. Tschinkel, "Density of rational points on Enriques surfaces", *Math. Res. Lett.* **5**:5 (1998), 623–628. MR 99m:14040 Zbl 0957.14016

[Bright 2002] M. Bright, *Computations on diagonal quartic surfaces*, Ph.D. thesis, University of Cambridge, 2002, Available at http://www.boojum.org.uk/maths/quartic-surfaces/thesis.pdf.

[Colliot-Thélène and Skorobogatov 2013] J.-L. Colliot-Thélène and A. N. Skorobogatov, "Good reduction of the Brauer–Manin obstruction", *Trans. Amer. Math. Soc.* **365**:2 (2013), 579–590. MR 2995366

[Ieronymou et al. 2011] E. Ieronymou, A. N. Skorobogatov, and Y. G. Zarhin, "On the Brauer group of diagonal quartic surfaces", *J. Lond. Math. Soc.* (2) **83**:3 (2011), 659–672. MR 2012e:14046 Zbl 1239.14013

[Logan et al. 2010] A. Logan, D. McKinnon, and R. van Luijk, "Density of rational points on diagonal quartic surfaces", *Algebra Number Theory* **4**:1 (2010), 1–20. MR 2011a:11126 Zbl 1206.11082

[Pannekoek 2012] R. Pannekoek, "On p-torsion of p-adic elliptic curves with additive reduction", preprint, 2012. arXiv 1211.5833

[Silverman 1986] J. H. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics **106**, Springer, New York, 1986. MR 87g:11070 Zbl 0585.14026

[Swinnerton-Dyer 1968] H. P. F. Swinnerton-Dyer, "$A^4 + B^4 = C^4 + D^4$ revisited", *J. London Math. Soc.* **43** (1968), 149–151. MR 37 #2685 Zbl 0206.33703

[Swinnerton-Dyer 2000] P. Swinnerton-Dyer, "Arithmetic of diagonal quartic surfaces, II", *Proc. London Math. Soc.* (3) **80**:3 (2000), 513–544. MR 2001d:11069 Zbl 1066.11029

[Swinnerton-Dyer 2001] P. Swinnerton-Dyer, "Weak approximation and R-equivalence on cubic surfaces", pp. 357–404 in *Rational points on algebraic varieties*, edited by E. Peyre and Y. Tschinkel, Progr. Math. **199**, Birkhäuser, Basel, 2001. MR 2003c:11070 Zbl 1079.11035

H.P.F.Swinnerton-Dyer@dpmms.cam.ac.uk

Department of Pure Mathematics and Mathematical Statistics, University of Cambridge, Cambridge, CB3 0WB, United Kingdom

msp

Albanese varieties with
modulus over a perfect field

Henrik Russell

Let X be a smooth proper variety over a perfect field k of arbitrary characteristic. Let D be an effective divisor on X with multiplicity. We introduce an Albanese variety $\mathrm{Alb}(X, D)$ of X of modulus D as a higher-dimensional analogue of the generalized Jacobian of Rosenlicht and Serre with modulus for smooth proper curves. Basing on duality of 1-motives with unipotent part (which are introduced here), we obtain explicit and functorial descriptions of these generalized Albanese varieties and their dual functors.

We define a relative Chow group of zero cycles $\mathrm{CH}_0(X, D)$ of modulus D and show that $\mathrm{Alb}(X, D)$ can be viewed as a universal quotient of $\mathrm{CH}_0(X, D)^0$.

As an application we can rephrase Lang's class field theory of function fields of varieties over finite fields in explicit terms.

0. Introduction

The generalized Jacobian variety with modulus of a smooth proper curve X over a field is a well-established object in algebraic geometry and number theory and has shown to be of great benefit, for instance, for the theory of algebraic groups, ramification theory and class field theory. In this work we extend this notion from [Serre 1959, V] to the situation of a higher-dimensional smooth proper variety X over a perfect field k. The basic idea of this construction comes from [Russell 2008] and is accomplished in [Kato and Russell 2012], both only for the case that k is of characteristic 0. Positive characteristic however requires distinct methods and turns out to be the difficult part of the story.

To a rational map $\varphi : X \dashrightarrow P$ from X to a torsor P under a commutative algebraic group G we assign an effective divisor $\mathrm{mod}(\varphi)$, the *modulus of φ* (Definition 3.11). Our definition from [Kato and Russell 2010] coincides with the classical definition in the curve case as in [Serre 1959, III, Section 1]. For an effective divisor D on X the generalized Albanese variety $\mathrm{Alb}^{(1)}(X, D)$ of X of modulus D and the Albanese map $\mathrm{alb}_{X,D}^{(1)} : X \dashrightarrow \mathrm{Alb}^{(1)}(X, D)$ are defined by the following universal property:

For every torsor P under a commutative algebraic group G and every rational map φ from X to P of modulus $\leq D$, there exists a unique homomorphism of torsors $h : \mathrm{Alb}^{(1)}(X, D) \to P$ such that $\varphi = h \circ \mathrm{alb}^{(1)}_{X,D}$. Every rational map to a torsor for a commutative algebraic group admits a modulus, and the effective divisors on X form an inductive system. Then the projective limit $\varprojlim \mathrm{Alb}^{(1)}(X, D)$ over all effective divisors D on X yields a torsor for a proalgebraic group that satisfies the universal mapping property for all rational maps from X to torsors for commutative algebraic groups.

The Albanese variety with modulus (Theorem 0.2) arises as a special case of a broader notion of generalized Albanese varieties defined by a universal mapping property for categories of rational maps from X to torsors for commutative algebraic groups. As the construction of these universal objects is based on duality, a notion of duality for smooth connected commutative algebraic groups over a perfect field k of arbitrary characteristic is required. For this purpose we introduce so called *1-motives with unipotent part* (Definition 1.18), which generalize Deligne 1-motives [1971, Définition (10.1.2)] and Laumon 1-motives [1996, Définition (5.1.1)]. In this context, we obtain explicit and functorial descriptions of these generalized Albanese varieties and their dual functors (Theorem 0.1).

In a geometric way we define a relative Chow group of 0-cycles $\mathrm{CH}_0(X, D)$ with respect to the modulus D (Definition 3.27). Then we can realize $\mathrm{Alb}^{(1)}(X, D)$ as a universal quotient of $\mathrm{CH}_0(X, D)^0$, the subgroup of $\mathrm{CH}_0(X, D)$ of cycles of degree 0 (Theorem 0.3), in the case that the base field is algebraically closed. The relation of $\mathrm{CH}_0(X, D)$ to the K-theoretic idèle class groups from [Kato and Saito 1983] gives rise to some future study, but is beyond the scope of this paper. Using these idèle class groups, Önsiper [1989] proved the existence of generalized Albanese varieties for smooth proper surfaces in characteristic $p > 0$.

Lang's class field theory of function fields of varieties over finite fields [Serre 1959, V] is written in terms of so called *maximal maps*, which appeared as a purely theoretical notion, apart from their existence very little seemed to be known about which. The Albanese map with modulus allows us to replace these black boxes by concrete objects (Theorem 0.4).

We present the main results by giving a summary of each section.

0.1. *Leitfaden.* Section 1 is devoted to the following generalization of 1-motives: A *1-motive with unipotent part* (Definition 1.18) is roughly a homomorphism $[\mathscr{F} \to G]$ in the category of sheaves of abelian groups over a perfect field k from a dual-algebraic commutative formal group $\mathscr{F}$ to an extension G of an abelian variety A by a commutative affine algebraic group L. Here a commutative formal group $\mathscr{F}$ is called *dual-algebraic* if its Cartier-dual $\mathscr{F}^\vee = \underline{\mathrm{Hom}}(\mathscr{F}, \mathbb{G}_\mathrm{m})$ is algebraic. 1-motives with unipotent part admit duality (Definition 1.21). The dual of $[0 \to G]$

is given by $[L^\vee \to A^\vee]$, where $L^\vee = \underline{\mathrm{Hom}}(L, \mathbb{G}_\mathrm{m})$ is the Cartier-dual of L and $A^\vee = \mathrm{Pic}^0_A = \underline{\mathrm{Ext}}^1(A, \mathbb{G}_\mathrm{m})$ is the dual abelian variety of A, and the homomorphism between them is the connecting homomorphism associated to $0 \to L \to G \to A \to 0$. In particular, every smooth connected commutative algebraic group over k has a dual in this category. Moreover, these 1-motives may contain torsion.

Section 2: Let X be a smooth proper variety over a perfect field k. In the framework of *categories of rational maps from X to torsors for commutative algebraic groups* (Definition 2.8), we ask for the existence of universal objects (Definition 2.14) for such categories, that is, objects having the universal mapping property with respect to the category they belong to. Assume for the moment k is an algebraically closed field. Then a torsor can be identified with the algebraic group acting on it. A necessary and sufficient condition for the existence of such universal objects is given in Theorem 2.16, as well as their explicit construction, using duality of 1-motives with unipotent part. (This was done in [Russell 2008] for char$(k) = 0$.) We pass to general perfect base field in Theorem 0.1.

In particular we show the following: Let $\underline{\mathrm{Div}}_X$ be the sheaf of relative Cartier divisors, that is, the sheaf of abelian groups that assigns to any k-algebra R the group $\underline{\mathrm{Div}}_X(R)$ of all Cartier divisors on $X \otimes_k R$ generated locally on Spec R by effective divisors which are flat over R. Let $\underline{\mathrm{Pic}}_X$ be the Picard functor and $\mathrm{Pic}^{0,\mathrm{red}}_X$ the Picard variety of X. Then let $\underline{\mathrm{Div}}^{0,\mathrm{red}}_X$ be the inverse image of $\mathrm{Pic}^{0,\mathrm{red}}_X$ under the class map $\mathrm{cl} : \underline{\mathrm{Div}}_X \to \underline{\mathrm{Pic}}_X$. A rational map $\varphi : X \dashrightarrow G$, where G is a smooth connected commutative algebraic group with affine part L, induces a natural transformation $\tau_\varphi : L^\vee \to \underline{\mathrm{Div}}^{0,\mathrm{red}}_X$ (Section 2.2.1). If $\mathcal{F}$ is a formal subgroup of $\underline{\mathrm{Div}}^{0,\mathrm{red}}_X$, denote by $\mathrm{Mr}_{\mathcal{F}}(X)$ the category of rational maps for which the image of this induced transformation lies in $\mathcal{F}$. If k is an arbitrary perfect base field, we define $\mathrm{Mr}_{\mathcal{F}}(X)$ via base change to an algebraic closure $\bar{k}$ (Definition 2.13).

Theorem 0.1. *Let $\mathcal{F}$ be a dual-algebraic formal k-subgroup of $\underline{\mathrm{Div}}^{0,\mathrm{red}}_X$. The category $\mathrm{Mr}_{\mathcal{F}}(X)$ admits a universal object $\mathrm{alb}^{(1)}_{\mathcal{F}} : X \dashrightarrow \mathrm{Alb}^{(1)}_{\mathcal{F}}(X)$. Here $\mathrm{Alb}^{(1)}_{\mathcal{F}}(X)$ is a torsor for an algebraic group $\mathrm{Alb}^{(0)}_{\mathcal{F}}(X)$, which arises as an extension of the classical Albanese $\mathrm{Alb}(X)$ by the Cartier-dual of $\mathcal{F}$. The algebraic group $\mathrm{Alb}^{(0)}_{\mathcal{F}}(X)$ is dual to the 1-motive $\left[\mathcal{F} \to \mathrm{Pic}^{0,\mathrm{red}}_X \right]$, the homomorphism induced by the class map $\mathrm{cl} : \underline{\mathrm{Div}}_X \to \underline{\mathrm{Pic}}_X$.*

Theorem 0.1 results from (the stronger) Theorem 2.16, which says that a category of rational maps to algebraic groups (over an algebraically closed field) admits a universal object if and only if it is of the shape $\mathrm{Mr}_{\mathcal{F}}(X)$ for some dual-algebraic formal subgroup $\mathcal{F}$ of $\underline{\mathrm{Div}}^{0,\mathrm{red}}_X$, and Galois descent (Theorem 2.21). The generalized Albanese varieties $\mathrm{Alb}^{(i)}_{\mathcal{F}}(X)$ ($i = 1, 0$) satisfy an obvious functoriality property (Proposition 2.22).

Section 3 is the main part of this work, where we establish a higher-dimensional analogue to the generalized Jacobian with modulus of Rosenlicht and Serre. Let X be a smooth proper variety over a perfect field k. We use the notion of modulus from [Kato and Russell 2010], which associates to a rational map $\varphi : X \dashrightarrow P$ an effective divisor $\mathrm{mod}(\varphi)$ on X (Definition 3.11). If D is an effective divisor on X, we define a formal subgroup $\mathscr{F}_{X,D} = (\mathscr{F}_{X,D})_{\text{ét}} \times_k (\mathscr{F}_{X,D})_{\text{inf}}$ of $\underline{\mathrm{Div}}_X$ (Definition 3.14) by the conditions

$$(\mathscr{F}_{X,D})_{\text{ét}} = \big\{ B \in \underline{\mathrm{Div}}_X(k) \,\big|\, \mathrm{Supp}(B) \subset \mathrm{Supp}(D) \big\},$$

and for $\mathrm{char}(k) = 0$,

$$(\mathscr{F}_{X,D})_{\text{inf}} = \exp\big(\widehat{\mathbb{G}}_a \otimes_k \Gamma\big(X, \mathbb{O}_X(D - D_{\text{red}})/\mathbb{O}_X\big)\big),$$

and for $\mathrm{char}(k) = p > 0$,

$$(\mathscr{F}_{X,D})_{\text{inf}} = \mathrm{Exp}\Big(\sum_{r>0} {}_r\widehat{W} \otimes_{W(k)} \Gamma\big(X, \mathrm{fil}^{\text{F}}_{D - D_{\text{red}}} W_r(\mathcal{K}_X)/W_r(\mathbb{O}_X)\big), 1\Big),$$

where D_{red} is the underlying reduced divisor of D, Exp denotes the Artin–Hasse exponential, ${}_r\widehat{W}$ is the kernel of the r-th power of the Frobenius on the completion $\widehat{W}$ of the Witt group W at 0 and $\mathrm{fil}^{\text{F}}_D W_r(\mathcal{K}_X)$ is a filtration of the Witt group (Definition 3.2). Let

$$\mathscr{F}^{0,\text{red}}_{X,D} = \mathscr{F}_{X,D} \times_{\underline{\mathrm{Div}}_X} \underline{\mathrm{Div}}^{0,\text{red}}_X$$

be the intersection of $\mathscr{F}_{X,D}$ and $\underline{\mathrm{Div}}^{0,\text{red}}_X$. The formal groups $\mathscr{F}_{X,D}$ and $\mathscr{F}^{0,\text{red}}_{X,D}$ are dual-algebraic (Proposition 3.15).

Then $\mathrm{mod}(\varphi) \leq D$ if and only if $\mathrm{im}(\tau_\varphi) \subset \mathscr{F}^{0,\text{red}}_{X,D}$ (Lemma 3.16). This yields (see Theorem 3.18 and Theorem 3.19):

Theorem 0.2. *The category* $\mathrm{Mr}(X, D)$ *of those rational maps* $\varphi : X \dashrightarrow P$ *such that* $\mathrm{mod}(\varphi) \leq D$ *admits a universal object* $\mathrm{alb}^{(1)}_{X,D} : X \dashrightarrow \mathrm{Alb}^{(1)}(X, D)$, *called the* Albanese *of* X *of modulus* D. *The algebraic group* $\mathrm{Alb}^{(0)}(X, D)$ *acting on* $\mathrm{Alb}^{(1)}(X, D)$ *is dual to the 1-motive* $[\mathscr{F}^{0,\text{red}}_{X,D} \to \mathrm{Pic}^{0,\text{red}}_X]$.

The Albanese varieties with modulus $\mathrm{Alb}^{(i)}(X, D)$ for $i = 1, 0$ are functorial (Proposition 3.22). In the case that $X = C$ is a curve, our Albanese with modulus $\mathrm{Alb}^{(i)}(C, D)$ coincide with the generalized Jacobians with modulus $J^{(i)}(C, D)$ of Rosenlicht and Serre (Theorem 3.25 and Galois descent).

A relative Chow group $\mathrm{CH}_0(X, D)$ of modulus D is introduced in Definition 3.27. We say a rational map $\varphi : X \dashrightarrow P$ to a torsor P under a commutative algebraic group G *factors through* $\mathrm{CH}_0(X, D)^0$ if the associated map $Z_0(U)^0 \to G(k)$, $\sum l_i p_i \mapsto \sum l_i \varphi(p_i)$ on 0-cycles of degree 0 (where U is the open set on which φ is defined) factors through a homomorphism of abstract groups $\mathrm{CH}_0(X, D)^0 \to G(k)$. We show (see Theorem 3.29):

Theorem 0.3. *Assume k is algebraically closed. A rational map $\varphi : X \dashrightarrow P$ factors through $\mathrm{CH}_0(X, D)^0$ if and only if it factors through $\mathrm{Alb}^{(1)}(X, D)$. In other words, $\mathrm{Alb}^{(0)}(X, D)$ is a universal quotient of $\mathrm{CH}_0(X, D)^0$.*

The theory of Albanese varieties with modulus has an application to the class field theory of function fields of varieties over finite fields. Let X be a geometrically irreducible projective variety over a finite field $k = \mathbb{F}_q$. Let $\bar{k}$ be an algebraic closure of k. Let K_X denote the function field of X, and $\mathrm{K}_X^{\mathrm{ab}}$ be the maximal abelian extension of K_X. From Lang's class field theory one obtains:

Theorem 0.4. *The geometric Galois group $\mathrm{Gal}(\mathrm{K}_X^{\mathrm{ab}} / \mathrm{K}_X \bar{k})$ is isomorphic to the projective limit of the k-rational points of the Albanese varieties of X with modulus D*

$$\mathrm{Gal}(\mathrm{K}_X^{\mathrm{ab}} / \mathrm{K}_X \bar{k}) \cong \varprojlim_{D} \mathrm{Alb}^{(0)}(X, D)(k),$$

where D ranges over all effective divisors on X rational over k.

The proof of Theorem 0.4 is analogous to the proof of Lang's class field theory given in [Serre 1959, VI, §4, nos. 16–19], replacing *maximal maps* by the universal objects $\mathrm{alb}_{X,D}^{(1)} : X \dashrightarrow \mathrm{Alb}^{(1)}(X, D)$ for the category of rational maps to k-torsors of modulus $\leq D$ from Theorem 0.2.

1. 1-motives

The aim of this section is to construct a category of generalized 1-motives that contains all smooth connected commutative algebraic groups over a perfect field and provides a notion of duality for them.

1.1. *Algebraic groups and formal groups.* I will use the language of group functors, algebraic groups and formal groups. References for algebraic groups are [Demazure and Gabriel 1970; Waterhouse 1979], and for formal groups and Cartier duality are [SGA3 1970, VII$_B$; Demazure 1972, II; Fontaine 1977, I].

By *algebraic group* and *formal group* I will always mean a *commutative* (algebraic and formal, respectively) group.

Let k be a ring (that is, associative, commutative and with unit). Set denotes the category of sets, Ab the category of abelian groups. Alg$/k$ denotes the category of k-algebras, and Art$/k$ the category of finite k-algebras (that is, of finite length). A *k-functor* is by definition a covariant functor from Alg$/k$ to Set. A *formal k-functor* is by definition a covariant functor from Art$/k$ to Set. A (formal) k-functor with values in Ab is called a *(formal) k-group functor*.

A *k-group* (or *k-group scheme*) is by definition a k-group functor with values in Ab whose underlying set-valued k-functor is represented by a k-scheme. The category of k-groups is denoted by $\mathscr{G}/k$, and the category of affine k-groups by

$\mathscr{G}_a/k$. An *algebraic k-group* (or just *algebraic group*) is a k-group whose underlying scheme is separated and of finite type over k. The category of algebraic k-groups is denoted by $a\mathscr{G}/k$, and the category of affine algebraic k-groups by $a\mathscr{G}_a/k$.

Now let k be a field. A *formal k-scheme* is by definition a formal k-functor with values in Set that is the limit of a directed inductive system of finite k-schemes. Let $\mathscr{A}$ be a *profinite k-algebra*. The *formal spectrum of $\mathscr{A}$* is the formal k-functor that assigns to $R \in \mathrm{Art}/k$ the set of continuous homomorphisms of k-algebras from the topological ring $\mathscr{A}$ to the discrete ring R: $\mathrm{Spf}\,\mathscr{A}(R) = \mathrm{Hom}^{\mathrm{cont}}_{k\text{-alg}}(\mathscr{A}, R)$.

A *formal k-group* (or just *formal group*) is a formal k-group functor with values in Ab whose underlying set-valued formal k-functor is represented by a formal k-scheme, or equivalently is isomorphic to $\mathrm{Spf}\,\mathscr{A}$ for some profinite k-algebra $\mathscr{A}$. The category of formal k-groups is denoted by $\mathscr{G}f/k$.

Remark 1.1. A formal k-group $\mathscr{F} : \mathrm{Art}/k \to \mathrm{Ab}$ extends in a natural way to a k-group functor $\widetilde{\mathscr{F}} : \mathrm{Alg}/k \to \mathrm{Ab}$, by defining $\widetilde{\mathscr{F}}(R)$ for $R \in \mathrm{Alg}/k$ as the inductive limit of the $\mathscr{F}(S)$, where S ranges over the finite k-subalgebras of R. If $\mathscr{F} = \mathrm{Spf}\,\mathscr{A}$ for some profinite k-algebra $\mathscr{A}$, then $\widetilde{\mathscr{F}}(R) = \mathrm{Hom}^{\mathrm{cont}}_{k\text{-alg}}(\mathscr{A}, R)$ for every $R \in \mathrm{Alg}/k$.

Theorem 1.2. *A formal k-group $\mathscr{F}$ is canonically an extension of an étale formal k-group $\mathscr{F}_{\text{ét}}$ by an infinitesimal ($=$ connected) formal k-group (that is, the formal spectrum of a local ring) $\mathscr{F}_{\text{inf}}$. Here*

$$\mathscr{F}_{\text{ét}}(R) = \mathscr{F}(R_{\text{red}}) \quad and \quad \mathscr{F}_{\text{inf}}(R) = \ker(\mathscr{F}(R) \to \mathscr{F}(R_{\text{red}}))$$

for $R \in \mathrm{Art}/k$, where $R_{\text{red}} = R/\mathrm{Nil}(R)$. If the base field k is perfect, there is a unique isomorphism $\mathscr{F} \cong \mathscr{F}_{\text{inf}} \times_k \mathscr{F}_{\text{ét}}$.

Proof. See [Demazure 1972, I, No. 7, Proposition on p. 34] or [Fontaine 1977, I, 7.2, p. 46]. □

Let R be a ring. An *R-sheaf* is a sheaf (of sets) on Alg/R for the topology fppf. An R-sheaf with values in Ab is called an *R-group sheaf*. The category of R-group sheaves is denoted by $\mathscr{A}b/R$.

Let k be a field. The category of k-groups $\mathscr{G}/k$ and the category of formal k-groups $\mathscr{G}f/k$ are full subcategories of $\mathscr{A}b/k$. This can be seen as follows: A k-functor that is represented by a scheme is a sheaf; see [Demazure and Gabriel 1970, III, §1, 1.3]. This gives the sheaf property for $\mathscr{G}/k$ by definition. For $\mathscr{G}f/k$ we can reduce to this case by Remark 1.1 and the fact that a formal k-group is the direct limit of finite k-schemes.

1.1.1. *Linear group associated to a ring.* Let k be a field.

Definition 1.3. Let R be a k-algebra. The *linear group associated to R* is the Weil restriction $\mathbb{L}_R := \Pi_{R/k}\mathbb{G}_{\mathrm{m},R} := \mathbb{G}_{\mathrm{m}}(\,\cdot \otimes R)$ of $\mathbb{G}_{\mathrm{m},R}$ from R to k.

If S is a finite k-algebra, then $\mathbb{L}_S$ is an affine algebraic k-group, according to [Demazure and Gabriel 1970, I, §1, 6.6].

Lemma 1.4. *Let k be a perfect field. Every affine algebraic k-group L is isomorphic to a closed subgroup of $\mathbb{L}_S$ for some $S \in \mathsf{Art}/k$.*

Proof. By Galois descent we can reduce to the case that k is algebraically closed. Every affine algebraic k-group L is isomorphic to a closed subgroup of GL_r for some $r \in \mathbb{N}$; see [Waterhouse 1979, 3.4 Theorem, p. 25]. Let $\rho : L \to \mathrm{GL}_r$ be a faithful representation. Define S to be the group algebra of $\rho(L)$, that is, the k-subalgebra of the algebra of $(r \times r)$-matrices $\mathrm{Mat}_{r \times r}(k)$ generated by $\rho(L)(k)$. In particular, S is finite-dimensional. Here we may assume that L is reduced, hence determined by its k-valued points; otherwise embed the multiplicative part into $(\mathbb{G}_m)^t$ for some $t \in \mathbb{N}$ (see [Demazure and Gabriel 1970, IV, §1, 1.5]) and the unipotent part into $(W_r)^n$ for some $r, n \in \mathbb{N}$ (see [ibid., V, §1, 2.5]), and replace L by $(\mathbb{G}_m)^t \times_k (W_r)^n$. Then $\rho(L)(k)$ is contained in the unit group of S, and $\rho : L \to \mathbb{G}_m (\,\cdot \otimes S) = \mathbb{L}_S$ is a monomorphism from L to $\mathbb{L}_S$. $\square$

1.1.2. *Cartier duality.* Let k be a field. We will use the functorial description of Cartier-duality as in [Demazure 1972, II, No. 4]. We may consider formal groups as objects of $\mathscr{A}b/k$; see Remark 1.1. Let G be a k-group sheaf. Let $\underline{\mathrm{Hom}}_{\mathscr{A}b/k}(G, \mathbb{G}_m)$ be the k-group functor defined by $R \mapsto \mathrm{Hom}_{\mathscr{A}b/R}(G_R, \mathbb{G}_{m,R})$, which assigns to a k-algebra R the group of homomorphisms of R-group sheaves from G_R to $\mathbb{G}_{m,R}$.

Theorem 1.5. *If G is an affine group or formal group), the k-group functor $\underline{\mathrm{Hom}}_{\mathscr{A}b/k}(G, \mathbb{G}_m)$ is represented by a formal group or affine group, respectively, $G^\vee$, which is called the* Cartier dual *of G.*

Cartier duality is an antiequivalence between the category of affine groups $\mathscr{G}a/k$ and the category of formal groups $\mathscr{G}f/k$. The functors $L \mapsto L^\vee$ and $\mathscr{F} \mapsto \mathscr{F}^\vee$ are quasiinverse to each other.

Proof. See [SGA3 1970, VII$_\mathrm{B}$, 2.2.2] or [Fontaine 1977, I, 5.4, p. 37] for a description of Cartier duality via bialgebras. See [Demazure 1972, II, No. 4, Theorem, p. 27] for one direction of the functorial description of Cartier duality (it is only one direction since formal groups and affine groups are not considered as objects of the same category there). According to Section 1.1 and the properties of the group functor $\underline{\mathrm{Hom}}_{\mathscr{A}b/k}(G, \mathbb{G}_m)$ as described in [Demazure and Gabriel 1970, II, §1, 2.10], it is an easy exercise to invert the given direction $L \mapsto L^\vee$ of the functorial description (one has to replace affine groups by formal groups, $\otimes_k$ by $\widehat{\otimes}_k$, $\mathrm{Hom}^{\mathrm{cont}}_{k\text{-alg}}$ by $\mathrm{Hom}_{k\text{-alg}}$ and Spf by Spec). $\square$

Lemma 1.6. *Let L be an affine group and R a k-algebra. The R-valued points of the Cartier-dual of L are given by $L^\vee(R) = \mathrm{Hom}_{\mathscr{A}b/k}(L, \mathbb{L}_R)$.*

Proof. The statement is due to the fact that Weil restriction is right-adjoint to base extension

$$L^\vee(R) = \text{Hom}_{\mathscr{A}b/R}(L_R, \mathbb{G}_{m,R}) = \text{Hom}_{\mathscr{A}b/k}(L, \mathbb{G}_{m,R}(\,\cdot\,\otimes R)). \qquad \square$$

Cartier dual of a multiplicative group.

Proposition 1.7 [Demazure 1972, II, No. 8]. *Let L be an affine k-group. Then L is multiplicative if and only if the Cartier-dual $L^\vee$ is an étale formal k-group.*

Example 1.8. In particular, the Cartier-dual of a split torus $T \cong (\mathbb{G}_m)^t$ is a lattice of the same rank: $T^\vee \cong \mathbb{Z}^t$, that is, a torsion-free étale formal group.

Proposition 1.9 [Demazure and Gabriel 1970, IV, §1, 1.2]. *Let L be a multiplicative k-group. Then L is algebraic if and only if $L^\vee(\bar{k})$ is of finite type.*

Cartier dual of a unipotent group.

Proposition 1.10 [Demazure 1972, II, No. 9]. *Let L be an affine k-group. Then L is unipotent if and only if the Cartier-dual $L^\vee$ is an infinitesimal formal k-group.*

Example 1.11 (Cartier duality of Witt vectors). Suppose $\text{char}(k) = p > 0$. Let W denote the k-group of Witt-vectors, W_r the k-group of Witt-vectors of finite length r. Let $\widehat{W}$ be the completion of W at 0, that is, $\widehat{W}$ is the subfunctor of W that associates to $R \in \text{Alg}/k$ the set of $(w_0, w_1, \dots) \in W(R)$ such that $w_\nu \in \text{Nil}(R)$ for all $\nu \in \mathbb{N}$ and $w_\nu = 0$ for almost all $\nu \in \mathbb{N}$. Moreover let $_r\widehat{W} = \ker(\text{F}^r : \widehat{W} \to \widehat{W}^{(p^r)})$ be the kernel of the r-th power of the Frobenius F. Let Λ denote the affine k-group that associates with $R \in \text{Alg}/k$ the multiplicative group $1 + tR[\![t]\!]$ of formal power series in R. Let E be the series

$$\text{E}(t) = \exp\Big(-\sum_{r \geq 0} \frac{t^{p^r}}{p^r}\Big) = \prod_{\substack{r \geq 1 \\ (r,p)=1}} (1 - t^r)^{\mu(r)/r},$$

where μ denotes the Möbius function. The *Artin–Hasse exponential* is the homomorphism of k-groups $\text{Exp} : W \to \Lambda$ defined by

$$\text{Exp}(w, t) := \text{Exp}(w)(t) := \prod_{r \geq 0} \text{E}(w_r t^{p^r}).$$

For details see for instance [Demazure 1972, III, Nos. 1 and 2].

Then $\widehat{W}$ and $_r\widehat{W}$ are Cartier-dual to W and W_r, respectively, and the pairings $\langle\,\cdot\,,\,\cdot\,\rangle : \widehat{W} \times W \to \mathbb{G}_m$ and $\langle\,\cdot\,,\,\cdot\,\rangle : {_r\widehat{W}} \times W_r \to \mathbb{G}_m$ are given by

$$\langle v, w \rangle = \text{Exp}(v \cdot w, 1) = \prod_{\substack{r \geq 0 \\ s \geq 0}} \text{E}(v_r^{p^s} w_s^{p^r}).$$

See [Demazure and Gabriel 1970, V, §4, Proposition 4.5 and Corollary 4.6].

Proposition 1.12. *Suppose* $\mathrm{char}(k) = p > 0$. *Let L be an affine k-group. The following conditions are equivalent*:

(i) *L is unipotent algebraic.*

(ii) *There is a monomorphism $L \hookrightarrow (W_r)^n$ for some $r, n \in \mathbb{N}$.*

(iii) *There is an epimorphism $({}_r\widehat{W})^n \twoheadrightarrow L^\vee$ for some $r, n \in \mathbb{N}$.*

(*Here we use the notation from [Example 1.11](#).*)

Proof. (i) $\Longrightarrow$ (ii) [Demazure and Gabriel 1970, V, §1, 2.5].

(ii) $\Longrightarrow$ (i) The underlying k-scheme of W_r is the affine space $\mathbb{A}^r$; thus W_r is algebraic. $0 = W_0 \subset W_1 \subset W_2 \subset \cdots \subset W_r$ is a filtration of W_r with quotients $W_\nu / W_{\nu-1} \cong W_1 = \mathbb{G}_a$; hence W_r is unipotent, according to [ibid., IV, §2, 2.5]. Products of unipotent groups and closed subgroups of a unipotent group are unipotent by [ibid., IV, §2, 2.3]. Since L is isomorphic to a closed subgroup of $(W_r)^n$, it is unipotent and algebraic.

(ii) $\Longleftrightarrow$ (iii) This is due to Cartier duality of Witt vectors; see [Example 1.11](#). $\square$

1.1.3. *Dual abelian variety.* Let k be a field. Let A be an abelian variety over k. The dual of A is given by $A^\vee = \mathrm{Pic}^0 A$. According to the generalized Barsotti–Weil formula (see [Oort 1966, III.18]), the dual abelian variety $A^\vee$ represents the k-group sheaf $\underline{\mathrm{Ext}}^1_{\mathscr{A}b/k}(A, \mathbb{G}_\mathrm{m})$ associated to $R \mapsto \mathrm{Ext}^1_{\mathscr{A}b/R}(A_R, \mathbb{G}_{\mathrm{m},R})$.

Proposition 1.13. *Let A be an abelian k-variety and S a finite k-algebra. There is a canonical isomorphism*

$$\mathrm{Ext}^1_{\mathscr{A}b/k}(A, \mathbb{L}_S) \xrightarrow{\sim} \mathrm{Ext}^1_{\mathscr{A}b/S}(A_S, \mathbb{G}_{\mathrm{m},S}).$$

Thus the S-valued points of the dual abelian variety are given by

$$A^\vee(S) = \mathrm{Ext}^1_{\mathscr{A}b/k}(A, \mathbb{L}_S).$$

Proof. Consider the following composition of functors on $\mathscr{A}b/S$:

$$\mathrm{Hom}_{\mathscr{A}b/k}(A, \cdot) \circ \Pi_{S/k} : G \mapsto \mathrm{Hom}_{\mathscr{A}b/k}(A, G\,(\cdot \otimes S)) = \mathrm{Hom}_{\mathscr{A}b/S}(A_S, G).$$

Since $\mathrm{Ext}^1_{\mathscr{A}b/k}(A, \mathbb{L}_S)$ and $\mathrm{Ext}^1_{\mathscr{A}b/S}(A_S, \mathbb{G}_{\mathrm{m},S})$ are identified with the sets of primitive elements in $\mathrm{H}^1(A, \mathbb{L}_S(\mathcal{O}_A))$ and $\mathrm{H}^1(A_S, \mathbb{G}_\mathrm{m}(\mathcal{O}_{A_S}))$, (see [Serre 1959, VII, no. 15, théorème 5] and [Oort 1966, III.17.6], respectively), we may compute these Ext-groups using the étale site instead of the flat site, according to [Milne 1980, III, Theorem 3.9]. As S is a finite k-algebra, the Weil restriction $\Pi_{S/k} : G \mapsto G\,(\cdot \otimes S)$ is exact for the étale topology (see [Milne 1980, II, Corollary 3.6]). Then the exact sequence of low degree terms of the Grothendieck spectral sequence yields a canonical isomorphism

$$(\mathrm{R}^1\,\mathrm{Hom}_{\mathscr{A}b/k}(A, \cdot))(\Pi_{S/k}(\mathbb{G}_\mathrm{m})) \xrightarrow{\sim} \mathrm{R}^1(\mathrm{Hom}_{\mathscr{A}b/k}(A, \cdot) \circ \Pi_{S/k})(\mathbb{G}_\mathrm{m})$$

(see [Milne 1980, Appendix B, Corollary 2]), showing the statement.   □

1.1.4. *Extensions of formal groups.* Let k be a field.

Lemma 1.14. $\underline{\mathrm{Ext}}^1_{\mathscr{A}b/k}(\mathscr{F}, \mathbb{G}_{\mathrm{m}}) = 0$ *for any dual-algebraic formal k-group $\mathscr{F}$.*

Proof. Let L be the affine algebraic group dual to $\mathscr{F}$. Let R be a k-algebra. We show that $\mathrm{Ext}^1_{\mathscr{A}b/R}(L^\vee_R, \mathbb{G}_{\mathrm{m},R}) = 0$ flat locally over Spec R. As L has a filtration $0 = L_0 \subset L_1 \subset \cdots \subset L_r = L$ with quotients equal to $\mathbb{G}_{\mathrm{m}}$, $\mathbb{G}_{\mathrm{a}}$ or a finite k-group, it suffices to show the statement for $L = \mathbb{G}_{\mathrm{m}}$, $\mathbb{G}_{\mathrm{a}}$ or a finite group.

If L is a finite k-group, the Cartier dual $L^\vee$ is again a finite k-group F, and $\mathrm{Ext}^1_{\mathscr{A}b/R}(F_R, \mathbb{G}_{\mathrm{m},R}) = 0$ flat locally according to [Milne 1980, III, §4, Lemma 4.17].

If $L = \mathbb{G}_{\mathrm{m}}$, then $L^\vee = \mathbb{Z}$, and $\mathrm{Ext}^1_{\mathscr{A}b/R}(\mathbb{Z}, \mathbb{G}_{\mathrm{m}}) = 0$ is clear.

If $L = \mathbb{G}_{\mathrm{a}}$ and char$(k) = 0$, then $L^\vee = \widehat{\mathbb{G}}_{\mathrm{a}}$. We have

$$\mathrm{Ext}^1_{\mathscr{A}b/R}(\widehat{\mathbb{G}}_{\mathrm{a}}, \mathbb{G}_{\mathrm{m}}) = \mathrm{Ext}^1_{\mathscr{A}b/R}(\widehat{\mathbb{G}}_{\mathrm{a}}, \widehat{\mathbb{G}}_{\mathrm{a}}) = 0;$$

see [Barbieri-Viale and Bertapelle 2009, Lemma A.4.6].

If $L = \mathbb{G}_{\mathrm{a}} = W_1$ and char$(k) = p > 0$, then $L^\vee = {}_1\widehat{W}$ (with notation from Example 1.11). Since ${}_1\widehat{W} = \ker(\mathrm{F} \colon \widehat{W} \to \widehat{W})$ is annihilated by the Frobenius F and since $\ker(\mathrm{F} \colon \mathbb{G}_{\mathrm{m}} \to \mathbb{G}_{\mathrm{m}}) = \mu_p$ is the group of p-th roots of unity over Spec R (this group is finite, and hence both an algebraic group and a formal group), any extension $E \in \mathrm{Ext}^1_{\mathscr{A}b/R}({}_1\widehat{W}, \mathbb{G}_{\mathrm{m}})$ is the push-out of an extension $\mathscr{F} \in \mathrm{Ext}^1_{\mathscr{A}b/R}({}_1\widehat{W}, \mu_p)$. As μ_p and ${}_1\widehat{W}$ are base changes of formal k-groups, the affine algebra $\mathbb{O}(\mu_p)$ of μ_p is a free R-module of finite rank and the affine algebra $\mathbb{O}({}_1\widehat{W})$ of ${}_1\widehat{W}$ is the projective limit of free R-modules of finite rank; I will refer to those algebras as *free pro-finite-rank R-algebras*. The underlying μ_p-bundle of $\mathscr{F}$ is flat locally trivial and hence flat locally the affine algebra of $\mathscr{F}$ is $\mathbb{O}(\mathscr{F}) = \mathbb{O}(\mu_p) \otimes_R \mathbb{O}({}_1\widehat{W})$, and this is a free pro-finite-rank R-algebra as well. In this case Cartier duality works in the same way as for formal k-groups,[1] so the exact sequence

$$0 \to \mu_p \to \mathscr{F} \to {}_1\widehat{W} \to 0$$

is turned into the exact sequence

$$0 \to \mathbb{G}_{\mathrm{a}} \to \mathscr{F}^\vee \to \mathbb{Z}/p\mathbb{Z} \to 0$$

of R-groups, where $\mathscr{F}^\vee = \underline{\mathrm{Hom}}_{\mathscr{A}b/R}(\mathscr{F}, \mathbb{G}_{\mathrm{m}})$. Applying $\underline{\mathrm{Hom}}_{\mathscr{A}b/R}(\,\cdot\,, \mathbb{G}_{\mathrm{m}})$ to the push-out diagram $\mathbb{G}_{\mathrm{m}} \leftarrow \mu_p \to \mathscr{F}$ of E shows that $E^\vee := \underline{\mathrm{Hom}}_{\mathscr{A}b/R}(E, \mathbb{G}_{\mathrm{m}})$ is the pull-back of the diagram $\mathbb{Z} \to \mathbb{Z}/p\mathbb{Z} \leftarrow \mathscr{F}^\vee$. In particular, since $\mathscr{F}^\vee \to \mathbb{Z}/p\mathbb{Z}$ is surjective, $E^\vee \to \mathbb{Z}$ is surjective as well. Thus we obtain an exact sequence

$$0 \to \mathbb{G}_{\mathrm{a}} \to E^\vee \to \mathbb{Z} \to 0,$$

---

[1]The category of flat locally free pro-finite-rank R-algebras is not abelian. The references for Cartier duality listed in the proof of Theorem 1.5 make additional assumptions on the base ring R in order to achieve that the category of R-formal groups is abelian.

which is obviously split. Dualizing again gives the split exact sequence

$$0 \to \mathbb{G}_m \to E^{\vee\vee} \to {}_1\widehat{W} \to 0,$$

where $E^{\vee\vee} = \underline{\mathrm{Hom}}_{\mathscr{A}b/R}(E^\vee, \mathbb{G}_m)$. The canonical map of any abelian sheaf $\mathscr{A}$ to its double dual $\mathscr{A}^{\vee\vee}$ yields the following commutative diagram with exact rows:

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & \mathbb{G}_m & \longrightarrow & E & \longrightarrow & {}_1\widehat{W} & \longrightarrow & 0 \\
& & \| & & \downarrow & & \| & & \\
0 & \longrightarrow & \mathbb{G}_m & \longrightarrow & E^{\vee\vee} & \longrightarrow & {}_1\widehat{W} & \longrightarrow & 0,
\end{array}
$$

where the vertical arrow in the middle is an isomorphism by the Five Lemma. Thus $E \cong E^{\vee\vee}$ is split. $\qquad\square$

1.2. *1-motives with unipotent part.* Let k be a field.

1.2.1. *Definition of a 1-motive with unipotent part.*

Definition 1.15. A formal k-group $\mathscr{F}$ is called *dual-algebraic* if its Cartier-dual $\mathscr{F}^\vee$ is algebraic. The category of dual-algebraic formal k-groups is denoted by $d\mathscr{G}f/k$.

Proposition 1.16. *A formal k-group $\mathscr{F}$ is dual-algebraic if and only if the following conditions are satisfied*:

(1) $\mathscr{F}(\bar{k})$ *is of finite type*,

(2) *for* $\mathrm{char}(k) = 0$, $\mathrm{Lie}(\mathscr{F})$ *is finite-dimensional, and*
   for $\mathrm{char}(k) > 0$, $\mathscr{F}_{\mathrm{inf}}$ *is a quotient of* $({}_r\widehat{W})^n$ *for some* $r, n \in \mathbb{N}$
   (see Example 1.11 for the definition of ${}_r\widehat{W}$).

Proof. The decomposition of $\mathscr{F}$ into étale part and infinitesimal part gives the decomposition of the affine group $\mathscr{F}^\vee$ into multiplicative part and unipotent part, according to Propositions 1.7 and 1.10. Then that statement follows directly from Propositions 1.9 and 1.12 for $\mathrm{char}(k) > 0$. For $\mathrm{char}(k) = 0$, the claim in (2) follows since the Lie functor yields an equivalence between the category of commutative infinitesimal formal k-groups and the category of k-vector spaces; see [SGA3 1970, VII$_B$, 3.2.2]. $\qquad\square$

Lemma 1.17. *Let $\mathscr{F}$ be a dual-algebraic formal group. Then any formal group $\mathscr{G}$ that is a subgroup or a quotient of $\mathscr{F}$ is also dual-algebraic.*

Proof. By Cartier-duality, this is equivalent to the dual statement about affine algebraic groups; see [Demazure 1972, II, No. 6, Corollary 4 of Theorem 2, p. 32]. $\qquad\square$

Definition 1.18. A *1-motive with unipotent part* is a tuple $M = (\mathscr{F}, L, A, G, \mu)$, where

(a) $\mathcal{F}$ is a dual-algebraic formal group (Definition 1.15),

(b) L is an affine algebraic group,

(c) A is an abelian variety,

(d) G is an extension of A by L,

(e) $\mu : \mathcal{F} \to G$ is a homomorphism in $\mathscr{A}b/k$.

A *homomorphism between 1-motives with unipotent part* $M = (\mathcal{F}, L, A, G, \mu)$ and $N = (\mathcal{E}, \Lambda, B, H, \nu)$ is a tuple $h = (\varphi, \lambda, \alpha, \gamma)$ of homomorphisms $\varphi : \mathcal{E} \to \mathcal{F}$, $\lambda : L \to \Lambda$, $\alpha : A \to B$, $\gamma : G \to H$, compatible with the structures of M and N as 1-motives with unipotent part, that is, giving an obvious commutative diagram.

For convenience, we will refer to a *1-motive with unipotent part* only as a *1-motive*.

If G is a smooth connected algebraic group, it admits a canonical decomposition $0 \to L \to G \to A \to 0$ as an extension of an abelian variety A by a connected affine algebraic group L, according to the theorem of Chevalley. Thus a homomorphism $\mu : \mathcal{F} \to G$ in $\mathscr{A}b/k$ gives rise to a 1-motive $M = (\mathcal{F}, L, A, G, \mu)$ that we will denote just by $M = [\mathcal{F} \xrightarrow{\mu} G]$.

1.2.2. *Duality of 1-motives.*

Theorem 1.19. *Let L be an affine algebraic group and A an abelian variety. There is a canonical isomorphism of abelian groups*

$$\Phi : \operatorname{Ext}^1_{\mathscr{A}b/k}(A, L) \xrightarrow{\sim} \operatorname{Hom}_{\mathscr{A}b/k}(L^{\vee}, A^{\vee}).$$

Proof. Consider the following left exact functor on $\mathscr{A}b/k$:

$$F : G \mapsto \operatorname{Bilin}_{\mathscr{A}b/k}(A, L^{\vee}; G) = \operatorname{Hom}_{\mathscr{A}b/k}(A, \underline{\operatorname{Hom}}_{\mathscr{A}b/k}(L^{\vee}, G))$$
$$= \operatorname{Hom}_{\mathscr{A}b/k}(L^{\vee}, \underline{\operatorname{Hom}}_{\mathscr{A}b/k}(A, G)),$$

where $\operatorname{Bilin}_{\mathscr{A}b/k}(A, L^{\vee}; G)$ is the group of $\mathbb{Z}$-bilinear maps $A \times L^{\vee} \to G$ of sheaves of abelian groups. The two ways of writing F as a composite yield the following two spectral sequences:

$$\operatorname{Ext}^p_{\mathscr{A}b/k}(A, \underline{\operatorname{Ext}}^q_{\mathscr{A}b/k}(L^{\vee}, G)) \Rightarrow \operatorname{R}^{p+q} F(G),$$
$$\operatorname{Ext}^p_{\mathscr{A}b/k}(L^{\vee}, \underline{\operatorname{Ext}}^q_{\mathscr{A}b/k}(A, G)) \Rightarrow \operatorname{R}^{p+q} F(G).$$

For $G = \mathbb{G}_m$, the associated exact sequences of low degree terms are

$$0 \to \operatorname{Ext}^1(A, \underline{\operatorname{Hom}}(L^{\vee}, \mathbb{G}_m)) \to \operatorname{R}^1 F(\mathbb{G}_m) \to \operatorname{Hom}(A, \underline{\operatorname{Ext}}^1(L^{\vee}, \mathbb{G}_m)) = 0,$$

where the last term vanishes due to Lemma 1.14, and

$$0 = \text{Ext}^1(L^\vee, \underline{\text{Hom}}(A, \mathbb{G}_\text{m})) \to \text{R}^1 F(\mathbb{G}_\text{m}) \to \text{Hom}(L^\vee, A^\vee)$$
$$\to \text{Ext}^2(L^\vee, \underline{\text{Hom}}(A, \mathbb{G}_\text{m})) = 0.$$

Putting these together we obtain isomorphisms

$$\text{Ext}^1_{\mathscr{A}b/k}(A, L) \xrightarrow{\sim} \text{R}^1 F(\mathbb{G}_\text{m}) \xrightarrow{\sim} \text{Hom}_{\mathscr{A}b/k}(L^\vee, A^\vee) \qquad \square$$

Remark 1.20 (explicit description of $\text{Ext}^1(A, L) \xrightarrow{\sim} \text{Hom}(L^\vee, A^\vee)$). The isomorphism Φ in Theorem 1.19 sends $G \in \text{Ext}^1(A, L)$ to the connecting homomorphism $\underline{\text{Hom}}_{\mathscr{A}b/k}(L, \mathbb{G}_\text{m}) \to \underline{\text{Ext}}^1_{\mathscr{A}b/k}(A, \mathbb{G}_\text{m})$ in the long exact cohomology sequence obtained from applying $\underline{\text{Hom}}_{\mathscr{A}b/k}(\cdot, \mathbb{G}_\text{m})$ to the short exact sequence $0 \to L \to G \to A \to 0$. Explicitly, this is the map $\Phi(G) : \lambda \mapsto \lambda_* G$, which sends $\lambda \in L^\vee(R) = \text{Hom}_{\mathscr{A}b/k}(L, \mathbb{L}_R)$ to the push-out $\lambda_* G \in \text{Ext}^1_{\mathscr{A}b/k}(A, \mathbb{L}_R) = A^\vee(R)$ of the diagram $\mathbb{L}_R \xleftarrow{\lambda} L \to G$.

If $L = \mathbb{L}_S$ for some $S \in \text{Art}/k$, the inverse map of Φ is given by the map $\Phi^{-1} : \vartheta \mapsto \vartheta(\text{id}_L)$, which sends a homomorphism $\vartheta : L^\vee \to A^\vee$ to the image $\vartheta(\text{id}_L) \in \text{Ext}^1_{\mathscr{A}b/k}(A, L) = A^\vee(S)$ of the identity $\text{id}_L \in \text{Hom}_{\mathscr{A}b/k}(L, L) = L^\vee(S)$. In general, a given homomorphism $\vartheta \in \text{Hom}_{\mathscr{A}b/k}(L^\vee, A^\vee)$ can be written as an element $[L^\vee \to A^\vee] \in \mathscr{C}^{[-1,0]}(\mathscr{A}b/k)$ of the category of two-term complexes in $\mathscr{A}b/k$, with $L^\vee$ placed in degree -1 and $A^\vee$ in degree 0. Then

$$\Phi^{-1}(\vartheta) = \underline{\text{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([L^\vee \to A^\vee], \mathbb{G}_\text{m}),$$

and this k-group sheaf is represented by an algebraic group: The short exact sequence $0 \to A^\vee \to [L^\vee \to A^\vee] \to L^\vee[1] \to 0$ gives rise to the exact sequence $0 \to \underline{\text{Hom}}(L^\vee, \mathbb{G}_\text{m}) \to \underline{\text{Ext}}^1([L^\vee \to A^\vee], \mathbb{G}_\text{m}) \to \underline{\text{Ext}}^1(A^\vee, \mathbb{G}_\text{m}) \to 0$ since $\underline{\text{Ext}}^1_{\mathscr{A}b/k}(L^\vee, \mathbb{G}_\text{m}) = 0$ by Lemma 1.14. Thus $\underline{\text{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([L^\vee \to A^\vee], \mathbb{G}_\text{m})$ is an extension of A by L.

Definition 1.21. The dual of a 1-motive $M = (\mathscr{F}, L, A, G, \mu)$ is the 1-motive $M^\vee = (L^\vee, \mathscr{F}^\vee, A^\vee, H, \eta)$, where

$$H = \underline{\text{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([\mathscr{F} \to A], \mathbb{G}_\text{m}) = \Phi^{-1}(\bar{\mu})$$

for $\bar{\mu} : \mathscr{F} \xrightarrow{\mu} G \to A$ the composite, and $\eta : L^\vee \to H$ the connecting homomorphism $\underline{\text{Hom}}_{\mathscr{A}b/k}(L, \mathbb{G}_\text{m}) \to \underline{\text{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([\mathscr{F} \to A], \mathbb{G}_\text{m})$ in the long exact cohomology sequence associated with $0 \to [0 \to L] \to [\mathscr{F} \to G] \to [\mathscr{F} \to A] \to 0$.

Remark 1.22. The double dual $M^{\vee\vee}$ of a 1-motive M is canonically isomorphic to M. If M is of the form $[0 \to G] := (0, L, A, G, 0)$, then the dual is

$$[L^\vee \xrightarrow{\Phi(G)} A^\vee] := (L^\vee, 0, A^\vee, A^\vee, \Phi(G)).$$

If M is of the form $[\mathscr{F} \xrightarrow{\mu} A] := (\mathscr{F}, 0, A, A, \mu)$, the dual is

$$[0 \to \Phi^{-1}(\mu)] := (0, \mathscr{F}^\vee, A^\vee, \Phi^{-1}(\mu), 0).$$

This is clear by Theorem 1.19, and these "pure 1-motives" are the only ones that we are concerned with in this note. For the general case, the proof carries over literally from [Laumon 1996, (5.2.4)].

Proposition 1.23. *Duality of* 1-*motives is functorial*, *that is*, *duality assigns to a homomorphism of* 1-*motives* $h : M \to N$ *a dual homomorphism* $h^\vee : N^\vee \to M^\vee$.

Proof. Let $M = (\mathscr{F}, L, A, G, \mu)$ and $M' = (\mathscr{F}', L', A', G', \mu')$ be 1-motives and $h : M \to M'$ a homomorphism of 1-motives. Applying $\underline{\mathrm{Hom}}_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}(\,\cdot\,, \mathbb{G}_{\mathrm{m}})$ to the commutative diagram with exact rows

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & [0 \to L] & \longrightarrow & [\mathscr{F} \to G] & \longrightarrow & [\mathscr{F} \to A] & \longrightarrow & 0 \\
& & \downarrow & & \downarrow h & & \downarrow & & \\
0 & \longrightarrow & [0 \to L'] & \longrightarrow & [\mathscr{F}' \to G'] & \longrightarrow & [\mathscr{F}' \to A'] & \longrightarrow & 0
\end{array}
$$

yields the homomorphism $h^\vee : [(L')^\vee \to H'] \to [L^\vee \to H]$, where

$$H = \underline{\mathrm{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([\mathscr{F} \to A], \mathbb{G}_{\mathrm{m}}) \quad \text{and} \quad H' = \underline{\mathrm{Ext}}^1_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}([\mathscr{F}' \to A'], \mathbb{G}_{\mathrm{m}}).$$

Applying $\underline{\mathrm{Hom}}_{\mathscr{C}^{[-1,0]}(\mathscr{A}b/k)}(\,\cdot\,, \mathbb{G}_{\mathrm{m}})$ to the commutative diagram with exact rows

$$
\begin{array}{ccccccccc}
0 & \longrightarrow & [0 \to A] & \longrightarrow & [\mathscr{F} \to A] & \longrightarrow & [\mathscr{F} \to 0] & \longrightarrow & 0 \\
& & \downarrow & & \downarrow & & \downarrow & & \\
0 & \longrightarrow & [0 \to A'] & \longrightarrow & [\mathscr{F}' \to A'] & \longrightarrow & [\mathscr{F}' \to 0] & \longrightarrow & 0
\end{array}
$$

shows that the image of $(\mathscr{F}')^\vee$ under $h^\vee$ is contained in $\mathscr{F}^\vee$, which implies that $h^\vee : (M')^\vee \to M^\vee$ is a homomorphism of 1-motives. $\qquad\square$

2. Universal rational maps

The classical Albanese variety $\mathrm{Alb}(X)$ of a variety X over a field k (as in [Lang 1959, II, §3]) is an abelian variety, defined together with the Albanese map $\mathrm{alb} : X \dashrightarrow \mathrm{Alb}(X)$ by the following universal mapping property: For every rational map $\varphi : X \dashrightarrow A$ to an abelian variety A there is a unique homomorphism $h : \mathrm{Alb}(X) \to A$ such that $\varphi = h \circ \mathrm{alb}$ up to translation by a constant $a \in A(k)$. Now we replace in this definition the category of abelian varieties by a subcategory C of the category of commutative algebraic groups. A result of Serre [1958–1959, No. 6, Théorème 8, p. 10–14] says that if the category C contains the additive group $\mathbb{G}_{\mathrm{a}}$ and X is a variety of dimension > 0, there does not exist an Albanese variety in C that is universal for all rational maps from X to algebraic groups in C. One is therefore led to restrict

the class of considered rational maps. This motivates the concept of *categories of rational maps from X to commutative algebraic groups*, or more generally, *categories of rational maps from X to torsors under commutative algebraic groups* (Definition 2.8), and to ask for the existence of universal objects for such categories.

For k an algebraically closed field with $\mathrm{char}(k) = 0$, in [Russell 2008, Section 2] I gave a criterion for which categories Mr of rational maps from a smooth proper variety X over k to algebraic groups there exists a universal object $\mathrm{Alb}_{\mathrm{Mr}}(X)$, as well as an explicit construction of these universal objects via duality of 1-motives. In this section we prove similar results for categories of rational maps, defined over a perfect field, from a smooth proper variety X to torsors for commutative algebraic groups.

2.1. *Relative Cartier divisors.*

The construction of such universal objects as above involves the functor $\underline{\mathrm{Div}}_X : \mathrm{Alg}/k \to \mathrm{Ab}$ of families of Cartier divisors, given by

$$\underline{\mathrm{Div}}_X(R) = \left\{ \begin{array}{l} \text{Cartier divisors } \mathscr{D} \text{ on } X \times_k \mathrm{Spec}\, R \\ \text{whose fibers } \mathscr{D}_p \text{ define Cartier divisors on } X \times_k \{p\} \\ \text{for all } p \in \mathrm{Spec}\, R \end{array} \right\}$$

for each k-algebra R, and for a homomorphism $h : R \to S$ in Alg/k the induced homomorphism $\underline{\mathrm{Div}}_X(h) : \underline{\mathrm{Div}}_X(R) \to \underline{\mathrm{Div}}_X(S)$ in Ab is the pull-back of Cartier divisors on $X \times_k \mathrm{Spec}\, R$ to those on $X \times_k \mathrm{Spec}\, S$. The elements of $\underline{\mathrm{Div}}_X(R)$ are called *relative Cartier divisors*. See [Russell 2008, No. 2.1] for more details on $\underline{\mathrm{Div}}_X$.

We will be mainly concerned with the completion $\widehat{\underline{\mathrm{Div}}_X} : \mathrm{Art}/k \to \mathrm{Ab}$ of $\underline{\mathrm{Div}}_X$, which is given for every finite k-algebra R by

$$\widehat{\underline{\mathrm{Div}}_X}(R) = \Gamma\big(X \otimes R, (\mathscr{K}_X \otimes_k R)^* / (\mathcal{O}_X \otimes_k R)^*\big).$$

We will regard $\widehat{\underline{\mathrm{Div}}_X}$ as a subsheaf of $\underline{\mathrm{Div}}_X$; see Remark 1.1.

Proposition 2.1. $\widehat{\underline{\mathrm{Div}}_X}$ *is a formal k-group.*

Proof. According to [Demazure 1972, I, No. 6; Fontaine 1977, I, §4] it suffices to show that $\widehat{\underline{\mathrm{Div}}_X}$ is left-exact (that is, commutes with finite projective limits). We are going to show that $\widehat{\underline{\mathrm{Div}}_X}$ is the composition of left-exact functors.

Let R be a finite k-algebra. $\widehat{\underline{\mathrm{Div}}_X}(R) = \Gamma\big(X, \mathscr{Q}(R)\big)$ is the abelian group of global sections of the sheaf $\mathscr{Q}(R) := (\mathrm{pr}_X)_*((\mathscr{K}_X \otimes R)^* / (\mathcal{O}_X \otimes R)^*)$, where $\mathrm{pr}_X : X \otimes R \to X$ is the projection. The global section functor $\Gamma(X, \cdot)$ is known to be left-exact. We show that the formal k-group functor $\mathscr{Q} : \mathrm{Art}/k \to \mathscr{Ab}/X$ (with values in the category of sheaves of abelian groups over X) commutes with finite projective limits (hence is left-exact):

Let (R_i) be a projective system of local finite k-algebras, with homomorphisms $h_{ij} : R_j \to R_i$ for $i < j$. We have projections $\mathrm{pr}_j : \varprojlim R_i \to R_j$ for each j, which commute with the h_{ij}. Functoriality of $\mathfrak{Q}$ in $R \in \mathrm{Art}/k$ induces homomorphisms $\mathfrak{Q}(h_{ij}) : \mathfrak{Q}(R_j) \to \mathfrak{Q}(R_i)$ and $\mathfrak{Q}(\mathrm{pr}_j) : \mathfrak{Q}(\varprojlim R_i) \to \mathfrak{Q}(R_j)$, which commute. The universal property of $\varprojlim \mathfrak{Q}(R_i)$ yields a unique homomorphism of sheaves $\mathfrak{Q}(\varprojlim R_i) \to \varprojlim \mathfrak{Q}(R_i)$. A homomorphism of sheaves is an isomorphism if and only if it is an isomorphism on stalks. Therefore it remains to show that the stalks $\mathfrak{Q}_q : \mathrm{Art}/k \to \mathrm{Ab}$ for $q \in X$ are left-exact in $R \in \mathrm{Art}/k$. We have

$$\mathfrak{Q}_q = \mathbb{G}_\mathrm{m}(\mathcal{K}_{X,q} \otimes_k \cdot)/\mathbb{G}_\mathrm{m}(\mathcal{O}_{X,q} \otimes_k \cdot).$$

The tensor product over a field $\mathcal{A} \otimes_k \cdot : \mathrm{Art}/k \to \mathrm{Alg}/k$ is exact for any k-algebra $\mathcal{A}$. Also the sheaf $\mathbb{G}_\mathrm{m} : \mathrm{Alg}/k \to \mathrm{Ab}$ is left-exact. Therefore the formal k-group functors $\mathbb{G}_\mathrm{m}(\mathcal{K}_{X,q} \otimes_k \cdot)$ and $\mathbb{G}_\mathrm{m}(\mathcal{O}_{X,q} \otimes_k \cdot)$ are formal k-groups. Since the category $\mathcal{G}f/k$ of formal k-groups is abelian (see [SGA3 1970, $\mathrm{VII_B}$, 2.4.2]), the quotient $\mathfrak{Q}_q$ of these two formal k-groups is again a formal k-group. $\qquad\square$

Definition 2.2. Let R be a finite k-algebra. If $D \in (\widehat{\underline{\mathrm{Div}}_X})_\mathrm{\acute{e}t}(R)$, then $\mathrm{Supp}(D)$ denotes the locus of zeroes and poles of local sections $(f_\alpha)_\alpha$ of $\mathbb{G}_\mathrm{m}(\mathcal{K}_X \otimes R_\mathrm{red})$ representing

$$D \in \Gamma(\mathbb{G}_\mathrm{m}(\mathcal{K}_X \otimes R_\mathrm{red})/\mathbb{G}_\mathrm{m}(\mathcal{O}_X \otimes R_\mathrm{red})).$$

If $\delta \in (\widehat{\underline{\mathrm{Div}}_X})_\mathrm{inf}(R)$, then $\mathrm{Supp}(\delta)$ denotes the locus of poles of local sections $(g_\alpha)_\alpha$ of $\mathbb{U}_R(\mathcal{K}_X)$ representing $\delta \in \Gamma(\mathbb{U}_R(\mathcal{K}_X)/\mathbb{U}_R(\mathcal{O}_X)) = (\widehat{\underline{\mathrm{Div}}_X})_\mathrm{inf}(R)$, where $\mathbb{U}_R = \ker(\mathbb{G}_\mathrm{m}(\cdot \otimes R) \to \mathbb{G}_\mathrm{m}(\cdot \otimes R_\mathrm{red}))$ is the unipotent part of $\mathbb{L}_R$ from Section 1.1.1

Definition 2.3. Let $\mathscr{F}$ be a formal subgroup of $\underline{\mathrm{Div}}_X$. The support of $\mathscr{F}$ is defined to be

$$\mathrm{Supp}(\mathscr{F}) = \bigcup_{\substack{R \in \mathrm{Art}/k \\ \mathscr{D} \in \mathscr{F}(R)}} \mathrm{Supp}(\mathscr{D})$$

where we use the decomposition $\mathscr{F} = \mathscr{F}_\mathrm{\acute{e}t} \times \mathscr{F}_\mathrm{inf}$ and Definition 2.2.

Suppose now that X is a geometrically irreducible smooth proper variety over a perfect field k. Then the Picard functor $\underline{\mathrm{Pic}}_X$ is represented by a separated algebraic space Pic_X, whose identity component Pic_X^0 is a proper scheme over k (see [Bosch et al. 1990, No. 8.4, Theorem 3]). The underlying reduced scheme $\mathrm{Pic}_X^{0,\mathrm{red}}$ of Pic_X^0 is an abelian variety, called the *Picard variety* of X. The subfunctor of $\underline{\mathrm{Pic}}_X$ that is represented by $\mathrm{Pic}_X^{0,\mathrm{red}}$ will be denoted by $\underline{\mathrm{Pic}}_X^{0,\mathrm{red}}$.

There is a natural transformation

$$\mathrm{cl} : \underline{\mathrm{Div}}_X \to \underline{\mathrm{Pic}}_X.$$

We define $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ to be the subfunctor of $\underline{\mathrm{Div}}_X$ given by

$$\underline{\mathrm{Div}}_X^{0,\mathrm{red}} = \underline{\mathrm{Div}}_X \times_{\underline{\mathrm{Pic}}_X} \underline{\mathrm{Pic}}_X^{0,\mathrm{red}}.$$

2.2. Categories of rational maps to torsors. Let X be a smooth proper variety over a perfect field k. Let $\bar{k}$ be an algebraic closure of k. In this note, algebraic groups and formal groups are commutative by definition (Section 1.1), and torsors are always torsors for commutative algebraic groups.

2.2.1. *Induced transformation.* Let G be a smooth connected algebraic group, and let $0 \to L \to G \to A \to 0$ be the canonical decomposition of G, where A is an abelian variety and L an affine smooth connected algebraic group (theorem of Chevalley). Write $L = T \times_k U$ where T is a torus and U is unipotent; see [SGA3 1970, XVII, 7.2.1]. If k is algebraically closed, $T \cong (\mathbb{G}_m)^t$ for some $t \in \mathbb{N}$. If k is of characteristic 0, one has $U \cong (\mathbb{G}_a)^s$ for some $s \in \mathbb{N}$ [Demazure and Gabriel 1970, IV, §2, 4.2]. If k is of characteristic $p > 0$, the unipotent group U is embedded into a finite direct sum $(W_r)^s$ of Witt vector groups for some $r, s \in \mathbb{N}$ [ibid., V, §1, 2.5].

Since $\mathrm{H}^1_{\mathrm{fppf}}(\mathrm{Spec}(\mathcal{O}_{X,q}), \mathbb{G}_m) = 0$ and $\mathrm{H}^1_{\mathrm{fppf}}(\mathrm{Spec}(\mathcal{O}_{X,q}), U) = 0$ for any point q of X, we have exact sequences

$$0 \to L(\mathcal{K}_{X,q}) \to G(\mathcal{K}_{X,q}) \to A(\mathcal{K}_{X,q}) \to 0,$$
$$0 \to L(\mathcal{O}_{X,q}) \to G(\mathcal{O}_{X,q}) \to A(\mathcal{O}_{X,q}) \to 0.$$

Since a rational map to an abelian variety is defined at every smooth point (see [Lang 1959, II, §1, Theorem 2]), we have $A(\mathcal{K}_{X,q}) = A(\mathcal{O}_{X,q})$ for every point q of X. Hence the canonical map

$$L(\mathcal{K}_{X,q})/L(\mathcal{O}_{X,q}) \to G(\mathcal{K}_{X,q})/G(\mathcal{O}_{X,q})$$

is bijective. By Cartier-duality, we have a pairing

$$\langle \cdot, \cdot \rangle : L^\vee \times \Gamma(L(\underline{\mathcal{K}}_X)/L(\underline{\mathcal{O}}_X)) \to \Gamma(\mathbb{G}_m(\underline{\mathcal{K}}_X)/\mathbb{G}_m(\underline{\mathcal{O}}_X)),$$

where $\underline{\mathcal{K}}_X := \mathcal{K}_X \otimes \cdot$ and $\underline{\mathcal{O}}_X := \mathcal{O}_X \otimes \cdot$.

Definition 2.4. Let $\varphi : X \dashrightarrow G$ be a rational map to a smooth connected algebraic group G, let L be the affine part of G. Then $\tau_\varphi : L^\vee \to \widehat{\underline{\mathrm{Div}}_X}$ denotes the induced transformation given by $\langle \cdot, \ell_\varphi \rangle$, where ℓ_φ is the image of $\varphi \in G(\mathcal{K}_X)$ in $\Gamma(G(\mathcal{K}_X)/G(\mathcal{O}_X)) \xrightarrow{\sim} \Gamma(L(\mathcal{K}_X)/L(\mathcal{O}_X))$. By construction, τ_φ is a homomorphism of formal k-group functors.

Lemma 2.5. *Let G be a smooth connected algebraic group, let L be the affine part of G. Let $\varphi : X \dashrightarrow G$ be a rational map. Let $\tau_\varphi : L^\vee \to \widehat{\underline{\mathrm{Div}}_X}$ be the induced transformation. Then $\mathrm{im}(\tau_\varphi)$ is a dual-algebraic formal group.*

Proof. $\widehat{\underline{\mathrm{Div}}_X}$ is a formal group by [Proposition 2.1](), and $\mathscr{G}f/k$ is a full subcategory of $\mathsf{Fctr}(\mathsf{Art}/k, \mathsf{Ab})$. Therefore $\tau_\varphi : L^\vee \to \widehat{\underline{\mathrm{Div}}_X}$ is a homomorphism of formal groups. Since $\mathscr{G}f/k$ is an abelian category, kernel and image of the homomorphism τ_φ are formal groups. Since L is algebraic, $L^\vee$ is dual-algebraic and hence $\mathrm{im}(\tau_\varphi)$, as a quotient of $L^\vee$, is dual-algebraic ([Lemma 1.17]()).   □

Lemma 2.6. *Let $G \in \mathrm{Ext}^1_{\mathscr{A}b/k}(A, L)$ be a smooth connected algebraic group. Let $\varphi : X \dashrightarrow G$ be a rational map. Let $\tau_\varphi : L^\vee \to \widehat{\underline{\mathrm{Div}}_X}$ be the induced transformation. Then $\mathrm{im}(\tau_\varphi)$ is contained in the completion of $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$.*

Proof. As A is an abelian variety, the composition $X \overset{\varphi}{\dashrightarrow} G \overset{\rho}{\to} A$ extends to a morphism $\overline{\varphi} : X \to A$. The description of the induced transformation τ_φ in terms of local sections into principal fiber bundles as given in [[Russell 2008](), No. 2.2] shows that the composition

$$L^\vee \overset{\tau_\varphi}{\to} \underline{\mathrm{Div}}_X \overset{\mathrm{cl}}{\to} \underline{\mathrm{Pic}}_X$$

is given by $\lambda \mapsto \lambda_* G_X$, where $\lambda_* G$ is the push-out of $G \in \mathrm{Ext}^1_{\mathscr{A}b/k}(A, L)$ via $\lambda \in L^\vee(R) = \mathrm{Hom}_{\mathscr{A}b/k}(L, \mathbb{L}_R)$, and $G_X = G \times_A X$ is the fiber-product of G and X over A. Hence it comes down to showing that for each $R \in \mathsf{Art}/k$, each $\lambda \in L^\vee(R)$ the $\mathbb{L}_R$-bundle $\lambda_* G_X$ yields an element of $\mathrm{Pic}_X^{0,\mathrm{red}}(R)$.

The universal mapping property of the classical Albanese $\mathrm{Alb}(X)$ yields that $\overline{\varphi}$ factors through $\mathrm{Alb}(X)$. Hence the pull-back $G_X = G \times_A X$ over X is a pull-back of $G_{\mathrm{Alb}} = G \times_A \mathrm{Alb}(X)$ over $\mathrm{Alb}(X)$. Then for each $\lambda \in L^\vee(R)$ the $\mathbb{L}_R$-bundle $\lambda_* G_{\mathrm{Alb}}$ over $\mathrm{Alb}(X)$ is an element of $\mathrm{Ext}^1_{\mathscr{A}b/k}(\mathrm{Alb}(X), \mathbb{L}_R)$, and hence gives an element of $\mathrm{Pic}^0_{\mathrm{Alb}(X)}(R)$. Since $\mathrm{Alb}(X) = (\mathrm{Pic}_X^{0,\mathrm{red}})^\vee$ is the dual abelian variety of $\mathrm{Pic}_X^{0,\mathrm{red}}$, we have an isomorphism

$$\mathrm{Pic}^0_{\mathrm{Alb}(X)} \overset{\sim}{\to} \mathrm{Pic}_X^{0,\mathrm{red}}, \quad P \mapsto P_X = P \times_{\mathrm{Alb}(X)} X.$$

As $\lambda_* G_X = \lambda_* G_{\mathrm{Alb}} \times_{\mathrm{Alb}(X)} X$, we have $\lambda_* G_X \in \mathrm{Pic}_X^{0,\mathrm{red}}(R)$.   □

Lemma 2.7. *Let L be an affine algebraic group and $\tau : L^\vee \to \widehat{\underline{\mathrm{Div}}_X^{0,\mathrm{red}}}$ a homomorphism of formal groups. Let $G \in \mathrm{Ext}^1_{\mathscr{A}b/k}(\mathrm{Alb}(X), L)$ be the extension corresponding to*

$$\mathrm{cl} \circ \tau : L^\vee \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}} \to \underline{\mathrm{Pic}}_X^{0,\mathrm{red}}$$

under the bijection Φ from [Theorem 1.19](). There is a rational map $\varphi : X \dashrightarrow G$ whose induced transformation is τ, and φ is determined uniquely up to translation by a constant $g \in G(k)$.

Proof. By [Lemma 1.4]() we may choose an embedding $\lambda : L \hookrightarrow \mathbb{L}_S$ for some finite ring $S \in \mathsf{Art}/k$. Let $\mathscr{D} \in \underline{\mathrm{Div}}_X^{0,\mathrm{red}}(S)$ be the image of $\mathrm{id}_{\mathbb{L}_S} \in \mathrm{Hom}_{\mathscr{A}b/k}(\mathbb{L}_S, \mathbb{L}_S) = \mathbb{L}_S^\vee(S)$ under the composition $\tau \circ \lambda^\vee : \mathbb{L}_S^\vee \to L^\vee \to \underline{\mathrm{Div}}_X$. [Remark 1.20]() shows that

$\mathbb{O}_{X \otimes S}(\mathcal{D}) \in \mathrm{Pic}_X^{0,\mathrm{red}}(S)$ is the line bundle corresponding to $G \in \mathrm{Ext}_{\mathcal{A}b/k}^1(\mathrm{Alb}(X), L)$ under the map

$$\mathrm{Ext}_{\mathcal{A}b/k}^1(\mathrm{Alb}(X), L) \to \mathrm{Ext}_{\mathcal{A}b/k}^1(\mathrm{Alb}(X), \mathbb{L}_S) = \mathrm{Pic}_{\mathrm{Alb}(X)}^0(S) \xrightarrow{\sim} \mathrm{Pic}_X^{0,\mathrm{red}}(S).$$

Let $\mathrm{G}_{\mathbb{L}_S}(\mathcal{D})$ be the image of G in $\mathrm{Ext}_{\mathcal{A}b/k}^1(\mathrm{Alb}(X), \mathbb{L}_S)$. Then the fiber product $\mathrm{P}_{\mathbb{L}_S}(\mathcal{D}) := \mathrm{G}_{\mathbb{L}_S}(\mathcal{D}) \times_{\mathrm{Alb}(X)} X$ is the $\mathbb{L}_S$-bundle on X associated to $\mathbb{O}_{X \otimes S}(\mathcal{D})$, and G and $P := G \times_{\mathrm{Alb}(X)} X$ are reductions of the $\mathbb{L}_S$-bundles $\mathrm{G}_{\mathbb{L}_S}(\mathcal{D})$ and $\mathrm{P}_{\mathbb{L}_S}(\mathcal{D})$ to the fiber L. The canonical 1-section of $\mathbb{O}_{X \otimes S}(\mathcal{D})$ yields a section $X \dashrightarrow P$, and composition with $P \to G$ yields the desired rational map $\varphi : X \dashrightarrow G$, which by construction satisfies $\tau_\varphi = \tau$. Then

$$\ell_\varphi \in \Gamma(G(\mathcal{H}_X)/G(\mathbb{O}_X)) \cong \Gamma(L(\mathcal{H}_X)/L(\mathbb{O}_X))$$
$$\subset \Gamma(\mathbb{L}_S(\mathcal{H}_X)/\mathbb{L}_S(\mathbb{O}_X)) = \Gamma(\mathbb{G}_m(\mathcal{H}_X \otimes S)/\mathbb{G}_m(\mathbb{O}_X \otimes S))$$

corresponding to $\mathcal{D}$ is uniquely determined by τ. The rational map $\varphi \in G(\mathcal{H}_X)$, as a lift of ℓ_φ, is determined up to a constant $g \in G(k) = \Gamma(G(\mathbb{O}_X))$, according to the exact sequence $0 \to \Gamma(G(\mathbb{O}_X)) \to \Gamma(G(\mathcal{H}_X)) \to \Gamma(G(\mathcal{H}_X)/G(\mathbb{O}_X))$. $\qquad\square$

2.2.2. *Definition of a category of rational maps.*

Definition 2.8. A *category* Mr *of rational maps from X to torsors* is a category satisfying the following conditions: The objects of Mr are rational maps $\varphi : X \dashrightarrow P$, where P is a torsor for a smooth connected algebraic group. The morphisms of Mr between two objects $\varphi : X \dashrightarrow P$ and $\psi : X \dashrightarrow Q$ are given by the set of those homomorphisms of torsors $h : P \to Q$ such that $h \circ \varphi = \psi$.

Remark 2.9. Let $\varphi : X \dashrightarrow P$ and $\psi : X \dashrightarrow Q$ be two rational maps from X to torsors. Then Definition 2.8 implies that for any category Mr of rational maps from X to torsors containing φ and ψ as objects the set of morphisms $\mathrm{Hom}_{\mathrm{Mr}}(\varphi, \psi)$ is the same. Therefore two categories Mr and Mr' of rational maps from X to torsors are equivalent if every object of Mr is isomorphic to an object of Mr'.

Remark 2.10. If a k-torsor P for an algebraic k-group G admits a k-rational point, then P may be identified with G. Thus for a rational map $\varphi : X \dashrightarrow P$ it makes sense to consider the base changed map $\varphi \otimes_k \bar{k} : X \otimes_k \bar{k} \dashrightarrow P \otimes_k \bar{k}$ as a rational map from $X \otimes_k \bar{k} = \overline{X}$ to an algebraic $\bar{k}$-group $P \otimes_k \bar{k} \cong G \otimes_k \bar{k}$.

Definition 2.11. The category of rational maps from X to abelian varieties is denoted by Mav.

Remark 2.12. The objects of Mav are in fact *morphisms* from X to abelian varieties, since a rational map from a smooth variety X to an abelian variety A extends to a morphism from X to A; see [Lang 1959, II, §1, Theorem 2].

Definition 2.13. Let $\mathscr{F}$ be a dual-algebraic formal k-subgroup of $\underline{\mathrm{Div}}_X$. If k is algebraically closed, then $\mathrm{Mr}_{\mathscr{F}}(X)$ denotes the category of all those rational maps $\varphi : X \dashrightarrow G$ from X to algebraic k-groups for which the image of the induced transformation $\tau_{\varphi} : L^{\vee} \to \underline{\mathrm{Div}}_X$ (Definition 2.4) lies in $\mathscr{F}$, that is, which induce a homomorphism of formal groups $L^{\vee} \to \mathscr{F}$, where L is the affine part of G. For general k, $\mathrm{Mr}_{\mathscr{F}}(X)$ denotes the category of all those rational maps $\varphi : X \dashrightarrow P$ from X to k-torsors for which the base changed map $\varphi \otimes \bar{k}$ is an object of $\mathrm{Mr}_{\mathscr{F} \otimes \bar{k}}$. (Here we use Remark 2.10.)

2.3. _Universal objects._ Let X be a smooth proper variety over a perfect field k. Algebraic groups are always assumed to be smooth and connected, and torsors are those for smooth connected algebraic groups, unless stated otherwise.

2.3.1. _Existence and construction._

Definition 2.14. Let Mr be a category of rational maps from X to torsors. Then $(u : X \dashrightarrow \mathcal{U}) \in \mathrm{Mr}$ is called a _universal object for_ Mr if it admits the universal mapping property in Mr: For all $(\varphi : X \dashrightarrow P) \in \mathrm{Mr}$ there is a unique homomorphism of torsors $h : \mathcal{U} \to P$ such that $\varphi = h \circ u$.

Remark 2.15. Universal objects are uniquely determined up to (unique) isomorphism.

Now assume that the base field k is algebraically closed. (Arbitrary perfect base field is considered from Section 2.3.2 on.) In this case we may identify a torsor with the algebraic group acting on it (Remark 2.10), and a homomorphism of torsors becomes a homomorphism of algebraic groups composed with a translation (which is an isomorphism of torsors).

For the category Mav of morphisms from X to abelian varieties (Definition 2.11) there exists a universal object, the _Albanese mapping_ to the _Albanese variety,_ denoted by $\mathrm{alb} : X \to \mathrm{Alb}(X)$. This is a classical result; see [Lang 1959; Matsusaka 1952; Serre 1958–1959]. The Albanese variety $\mathrm{Alb}(X)$ is an abelian variety, dual to the Picard variety $\mathrm{Pic}_X^{0,\mathrm{red}}$.

In the following we consider categories Mr of rational maps from X to algebraic groups satisfying the following conditions:

$(1^{\diamond})$ Mr contains the category Mav.

$(2^{\diamond})$ If $(\varphi : X \dashrightarrow G) \in \mathrm{Mr}$ and $h : G \to H$ is a homomorphism of torsors for smooth connected algebraic groups, then $h \circ \varphi \in \mathrm{Mr}$.

Theorem 2.16. _Let_ Mr _be a category of rational maps from X to algebraic groups that satisfies_ $(1^{\diamond})$ _and_ $(2^{\diamond})$ _Then the following conditions are equivalent_:

(i) _For_ Mr _there exists a universal object_ $(u : X \dashrightarrow \mathcal{U}) \in \mathrm{Mr}$.

(ii) *There is a dual-algebraic formal subgroup $\mathscr{F}$ of $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ such that* Mr *is equivalent to* $\mathrm{Mr}_{\mathscr{F}}(X)$.

(iii) *The formal group $\mathscr{F}_{\mathrm{Mr}} \subset \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ generated by* $\bigcup_{\varphi \in \mathrm{Mr}} \mathrm{im}(\tau_\varphi)$ *is dual-algebraic and* $\mathrm{Mr} = \mathrm{Mr}_{\mathscr{F}_{\mathrm{Mr}}}(X)$.

Here $\mathrm{Mr}_{\mathscr{F}}(X)$ is the category of rational maps that induce a homomorphism of formal groups to $\mathscr{F}$ (Definition 2.13).

Proof. (ii) $\implies$ (i) Assume that Mr is equivalent to $\mathrm{Mr}_{\mathscr{F}}(X)$, where $\mathscr{F}$ is a dual-algebraic formal group in $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$. The first step is the construction of an algebraic group $\mathscr{U}$ and a rational map $u : X \dashrightarrow \mathscr{U}$. In a second step the universality of $u : X \dashrightarrow \mathscr{U}$ for $\mathrm{Mr}_{\mathscr{F}}(X)$ will be shown.

Step 1: Construction of $u : X \dashrightarrow \mathscr{U}$. X is a smooth proper variety over k; thus the functor $\underline{\mathrm{Pic}}_X^0$ is represented by an algebraic group Pic_X^0 whose underlying reduced scheme $\mathrm{Pic}_X^{0,\mathrm{red}}$, the Picard variety of X, is an abelian variety. The class map $\underline{\mathrm{Div}}_X \to \underline{\mathrm{Pic}}_X$ induces a homomorphism $\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}$.

We obtain a 1-motive $M = [\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}]$. Since $\mathrm{Pic}_X^{0,\mathrm{red}}$ is an abelian variety, the dual 1-motive of M is of the form $M^\vee = [0 \to G]$, where G is a smooth connected algebraic group. Then define $\mathscr{U}$ to be this algebraic group. The canonical decomposition $0 \to \mathscr{L} \to \mathscr{U} \to \mathscr{A} \to 0$ is the extension of $\mathscr{A} = \mathrm{Alb}(X) = (\mathrm{Pic}_X^{0,\mathrm{red}})^\vee$ by $\mathscr{L} = \mathscr{F}^\vee$ induced by the homomorphism $\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}$ (Theorem 1.19).

Define the rational map $u : X \dashrightarrow \mathscr{U}$ by the condition that the induced transformation $\tau_u : \mathscr{F} \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ from Definition 2.4 is the inclusion. According to Lemma 2.7, $u : X \dashrightarrow \mathscr{U}$ is determined up to a constant $c \in \mathscr{U}(k)$.

Note that $u : X \dashrightarrow \mathscr{U}$ generates $\mathscr{U}$: Let H be the subgroup generated by the image of u, and let $\Lambda \subset \mathscr{L}$ be the affine part of H. Since $u : X \to \mathscr{U}$ factors through H, the induced transformation $\tau_u : \mathscr{L}^\vee \to \widehat{\mathrm{Div}_X}$ factors through the quotient $\mathscr{L}^\vee \twoheadrightarrow \Lambda^\vee$. Since τ_u is injective, this yields $\Lambda^\vee \cong \mathscr{L}^\vee$; hence $\Lambda \cong \mathscr{L}$. Since the composition $X \overset{u}{\dashrightarrow} \mathscr{U} \to \mathscr{A}$ generates $\mathscr{A}$, the abelian quotient of H is $\mathscr{A}$. These two conditions imply that $H \cong \mathscr{U}$ by the five lemma.

Step 2: Universality of $u : X \dashrightarrow \mathscr{U}$. Let $\varphi : X \dashrightarrow G$ be a rational map to a smooth connected algebraic group G with canonical decomposition
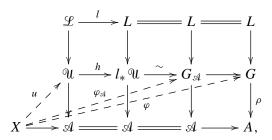
$$0 \to L \to G \overset{\rho}{\to} A \to 0,$$

inducing a homomorphism of formal groups $\tau_\varphi : L^\vee \to \mathscr{F} \subset \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$, $\lambda \mapsto \langle \lambda, \ell_\varphi \rangle$ (Definition 2.4). Let $l := (\tau_\varphi)^\vee : \mathscr{L} \to L$ be the dual homomorphism of affine groups. The composition

$$X \overset{\varphi}{\dashrightarrow} G \overset{\rho}{\to} A$$

extends to a morphism from X to an abelian variety. Translating φ by a constant $g \in G(k)$, if necessary, we may assume that $\rho \circ \varphi$ factors through $\mathscr{A} = \mathrm{Alb}(X)$. We

are going to show that we have a commutative diagram

$$
\begin{array}{ccccccc}
\mathscr{L} & \xrightarrow{\ l\ } & L & =\!=\!= & L & =\!=\!= & L \\
\downarrow & & \downarrow & & \downarrow & & \downarrow \\
\mathscr{U} & \xrightarrow{\ h\ } & l_*\mathscr{U} & \xrightarrow{\ \sim\ } & G_{\mathscr{A}} & \longrightarrow & G \\
& {}_{\varphi_{\mathscr{A}}} & & {}_{\varphi} & & & \downarrow \rho \\
X & =\!=\!= & \mathscr{A} & =\!=\!= & \mathscr{A} & =\!=\!= & \mathscr{A} & \longrightarrow & A ,
\end{array}
$$

where $G_{\mathscr{A}} = G \times_A \mathscr{A}$ is the fiber product, $l_*\mathscr{U} = \mathscr{U} \amalg_{\mathscr{L}} L$ is the amalgamated sum and $h : \mathscr{U} \to l_*\mathscr{U}$ is the map obtained from the amalgamated sum.

Denoting by $\bar{l} : \Gamma(\mathscr{L}(\mathscr{K}_X)/\mathscr{L}(\mathbb{O}_X)) \to \Gamma(L(\mathscr{K}_X)/L(\mathbb{O}_X))$ the map induced by $l : \mathscr{L} \to L$, we have $\ell_{h \circ u} = \bar{l}(\ell_u)$. This yields

$$
\tau_{h \circ u} = \langle\, \cdot\,, \ell_{h \circ u} \rangle = \langle\, \cdot\,, \bar{l}(\ell_u) \rangle = \langle\, \cdot \circ l, \ell_u \rangle = \tau_u \circ l^{\vee} = \tau_{\varphi}
$$

since $\tau_u : \mathscr{F} \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ is the inclusion by construction of u.

This implies that $l_*\mathscr{U}_X$ and G_X are isomorphic L-bundles over X. Then $l_*\mathscr{U}$ and $G_{\mathscr{A}}$ are isomorphic as extensions of $\mathscr{A}$ by L, using the isomorphism $\mathrm{Pic}_X^0 \xrightarrow{\sim} \mathrm{Pic}_{\mathscr{A}}^0$. Thus $\tau_{h \circ u} = \tau_{\varphi}$ shows that $h \circ u$ and $\varphi_{\mathscr{A}}$ coincide up to translation. Since $u : X \to \mathscr{U}$ generates $\mathscr{U}$, each $h' : \mathscr{U} \to G_{\mathscr{A}}$ fulfilling $h' \circ u = \varphi_{\mathscr{A}}$ coincides with h. Hence h is unique.

(i) $\Longrightarrow$ (iii) Assume $u : X \dashrightarrow \mathscr{U}$ is universal for Mr. Let $0 \to \mathscr{L} \to \mathscr{U} \to \mathscr{A} \to 0$ be the canonical decomposition of $\mathscr{U}$, and let $\mathscr{F}$ be the image of the induced transformation $\tau_u : \mathscr{L}^{\vee} \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$. For $\lambda \in \mathscr{L}^{\vee}(R)$ the uniqueness of the homomorphism $h_\lambda : \mathscr{U} \to \lambda_*\mathscr{U}$ fulfilling $u_\lambda = h_\lambda \circ u$ implies that the rational maps $u_\lambda : X \dashrightarrow \lambda_*\mathscr{U}$ are nonisomorphic to each other for distinct $\lambda \in \mathscr{L}^{\vee}(R)$. Hence $\mathrm{div}_R(u_{X,\nu}) \neq \mathrm{div}_R(u_{X,\lambda})$ for $\nu \neq \lambda \in \mathscr{L}^{\vee}(R)$. Therefore $\mathscr{L}^{\vee} \to \mathscr{F}$ is injective and hence an isomorphism.

Let $\varphi : X \dashrightarrow G$ be an object of Mr and $0 \to L \to G \to A \to 0$ be the canonical decomposition of G. Translating φ by a constant $g \in G(k)$, if necessary, we may assume that $\varphi : X \dashrightarrow G$ factors through a unique homomorphism $h : \mathscr{U} \to G$. The restriction of h to $\mathscr{L}$ gives a homomorphism of affine groups $l : \mathscr{L} \to L$. Then the dual homomorphism $l^{\vee} : L^{\vee} \to \mathscr{F}$ yields a factorization of $L^{\vee} \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ through $\mathscr{F}$. The properties $(1^{\diamond})$, $(2^{\diamond})$ and the existence of a universal object guarantee that Mr contains all rational maps that induce a transformation to $\mathscr{F}$; hence Mr is equal to $\mathrm{Mr}_{\mathscr{F}}(X)$.

(iii) $\Longrightarrow$ (ii) is evident. $\qquad\square$

Notation 2.17. If $\mathscr{F}$ is a dual-algebraic formal subgroup of $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$, then the universal object for $\mathrm{Mr}_{\mathscr{F}}(X)$ is denoted by $\mathrm{alb}_{\mathscr{F}} : X \dashrightarrow \mathrm{Alb}_{\mathscr{F}}(X)$.

Remark 2.18. By construction, $\mathrm{Alb}_{\mathscr{F}}(X)$ is generated by X. Since X is reduced, $\mathrm{Alb}_{\mathscr{F}}(X)$ is reduced as well and thus smooth. In the proof of Theorem 2.16 we have seen that $\mathrm{Alb}_{\mathscr{F}}(X)$ is an extension of the abelian variety $\mathrm{Alb}(X)$ by the affine group $\mathscr{F}^{\vee}$. More precisely, $[0 \to \mathrm{Alb}_{\mathscr{F}}(X)]$ is the dual 1-motive of $[\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}]$. The rational map $(\mathrm{alb}_{\mathscr{F}} : X \dashrightarrow \mathrm{Alb}_{\mathscr{F}}(X)) \in \mathrm{Mr}_{\mathscr{F}}(X)$ is characterized by the fact that the transformation $\tau_{\mathrm{alb}_{\mathscr{F}}} : L^{\vee} \to \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ is the identity $\mathscr{F} \xrightarrow{\mathrm{id}} \mathscr{F}$.

2.3.2. *Descent of the base field.* Let k be a perfect field. Let $\bar{k}$ be an algebraic closure of k. Let X be a smooth proper variety defined over k, and write $\overline{X} = X \otimes_k \bar{k}$. Let $\mathscr{F}$ be a formal k-subgroup of $\underline{\mathrm{Div}}_X^0$, and write $\overline{\mathscr{F}} = \mathscr{F} \widehat{\otimes} \bar{k}$.

The wish is to show that the universal object $\mathrm{alb}_{\overline{\mathscr{F}}} : \overline{X} \dashrightarrow \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$ for the category $\mathrm{Mr}_{\overline{\mathscr{F}}}$ can be defined over k. This will be accomplished by a Galois descent, as described in [Serre 1959, V, §4]. Due to Cartier duality between formal groups and affine groups (Theorem 1.5), Galois descent applies to formal groups as well.

When one does not assume that X is endowed with a k-rational point, one is led to two different descents of $\mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$:

First: The universal mapping property of $\mathrm{alb}_{\overline{\mathscr{F}}} : \overline{X} \dashrightarrow \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$ gives for every $\sigma \in \mathrm{Gal}(\bar{k}/k)$ transformations $h_{\sigma}^{(1)} : \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X}) \to \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})^{\sigma}$ between $\mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$ and its conjugate $\mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})^{\sigma}$, which are homomorphisms of torsors. Therefore the descent of $\mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$ by means of the $h_{\sigma}^{(1)}$ yields a k-torsor $\mathrm{Alb}_{\mathscr{F}}^{(1)}(X)$.

Second: To avoid translations or the reference to base points, one may reformulate the universal mapping property, replacing rational maps $\varphi : \overline{X} \dashrightarrow G$ from $\overline{X}$ to algebraic groups by its associated "difference maps" $\varphi^{(-)} : \overline{X} \times \overline{X} \dashrightarrow G$, $(p, q) \mapsto \varphi(q) - \varphi(p)$. In this way translations are eliminated and one obtains transformations $h_{\sigma}^{(0)} : \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X}) \to \mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})^{\sigma}$ that are homomorphisms of algebraic groups. Then the descent of $\mathrm{Alb}_{\overline{\mathscr{F}}}(\overline{X})$ by means of the $h_{\sigma}^{(0)}$ yields an algebraic k-group $\mathrm{Alb}_{\mathscr{F}}^{(0)}(X)$. This is the k-group acting on the k-torsor $\mathrm{Alb}_{\mathscr{F}}^{(1)}(X)$.

Notation 2.19. If $\varphi : X \dashrightarrow P$ is a rational map to a torsor P for an algebraic group G, then $\varphi^{(-)} : X \times X \dashrightarrow G$ denotes the rational map to the algebraic group G that assigns for $S \in \mathrm{Alg}/k$ to $(p, q) \in X(S) \times X(S)$ the unique $g \in G(S)$ such that $g \cdot \varphi(p) = \varphi(q)$.

Notation 2.20. If $\varphi : X \dashrightarrow P$ is a rational map to a torsor, then set $\varphi^{(1)} := \varphi$, $\varphi^{(0)} := \varphi^{(-)}$.

Theorem 2.21. *There exists a k-torsor $\mathrm{Alb}_{\mathscr{F}}^{(1)}(X)$ for an algebraic k-group $\mathrm{Alb}_{\mathscr{F}}^{(0)}(X)$ and rational maps defined over k*

$$\mathrm{alb}_{\mathscr{F}}^{(i)} : X^{2-i} \dashrightarrow \mathrm{Alb}_{\mathscr{F}}^{(i)}(X) \quad \text{for } i = 1, 0,$$

satisfying the following universal property:

If $\varphi : X \dashrightarrow G^{(1)}$ is a rational map defined over k to a k-torsor $G^{(1)}$ for an algebraic k-group $G^{(0)}$ that is an object of $\mathrm{Mr}_{\mathscr{F}}(X)$, then there exist unique homomorphisms of k-torsors and algebraic k-groups

$$h^{(1)} : \mathrm{Alb}_{\mathscr{F}}^{(1)}(X) \to G^{(1)} \quad and \quad h^{(0)} : \mathrm{Alb}_{\mathscr{F}}^{(0)}(X) \to G^{(0)},$$

respectively, defined over k, such that $\varphi^{(i)} = h^{(i)} \circ \mathrm{alb}_{\mathscr{F}}^{(i)}$ for $i = 1, 0$.

The algebraic group $\mathrm{Alb}_{\mathscr{F}}^{(0)}(X)$ is dual to the 1-motive $[\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}]$.

Proof. Galois descent. The same arguments as given in [Serre 1959, V, no. 22] work in our situation. □

2.3.3. *Functoriality.* Let $\mathscr{F} \subset \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ be a dual-algebraic formal k-group. Let $\psi : Y \to X$ be a k-morphism of smooth proper k-varieties, such that no irreducible component of $\psi(Y)$ is contained in $\mathrm{Supp}(\mathscr{F})$. For each dual-algebraic formal k-group $\mathscr{G} \subset \underline{\mathrm{Div}}_Y^{0,\mathrm{red}}$ containing $\psi^* \mathscr{F}$ the pull-back of relative Cartier divisors and of line bundles induces a homomorphism $[\mathscr{G} \to \mathrm{Pic}_Y^{0,\mathrm{red}}] \leftarrow [\mathscr{F} \to \mathrm{Pic}_X^{0,\mathrm{red}}]$ of 1-motives.

According to the construction of universal objects over $\bar{k}$ (Remark 2.18), we obtain via dualization of 1-motives and by Galois descent:

Proposition 2.22. *Using the assumptions above, ψ induces for every formal group $\mathscr{G} \subset \underline{\mathrm{Div}}_Y^{0,\mathrm{red}}$ containing $\psi^* \mathscr{F}$ a homomorphism of k-torsors $\mathrm{Alb}^{(1)}{}_{\mathscr{G}}^{\mathscr{F}}(\psi)$ and a homomorphism of algebraic k-groups $\mathrm{Alb}^{(0)}{}_{\mathscr{G}}^{\mathscr{F}}(\psi)$,*

$$\mathrm{Alb}^{(i)}{}_{\mathscr{G}}^{\mathscr{F}}(\psi) : \mathrm{Alb}_{\mathscr{G}}^{(i)}(Y) \to \mathrm{Alb}_{\mathscr{F}}^{(i)}(X) \quad for \ i = 1, 0.$$

3. Albanese with modulus

Let X be a smooth proper variety over a perfect field k. Let D be an effective divisor on X (with multiplicity). The Albanese $\mathrm{Alb}^{(1)}(X, D)$ of X of modulus D is a higher-dimensional analogue of the generalized Jacobian with modulus of Rosenlicht and Serre. $\mathrm{Alb}^{(1)}(X, D)$ is defined by the universal mapping property for morphisms from $X \setminus D$ to torsors of modulus $\leq D$ (Definition 3.11). Our definition of the modulus of rational maps coincides with the classical definition from [Serre 1959, III, §1] in the curve case. Therefore the Albanese with modulus agrees with the Jacobian with modulus of Rosenlicht and Serre for curves, which we review in Section 3.3.

In Section 3.4 we consider a Chow group $\mathrm{CH}_0(X, D)^0$ of 0-cycles relative to the modulus D (Definition 3.27). We give an alternative characterization of $\mathrm{Alb}(X, D)$ as a universal quotient of $\mathrm{CH}_0(X, D)^0$, when the base field is algebraically closed (Theorem 3.29).

3.1. *Filtrations of the Witt group.* Here we present a global version of some basic notions from [Kato and Russell 2010] that are needed for the construction of the Albanese with modulus.

Let (K, v) be a discrete valuation field of characteristic $p > 0$ with residue field k. The group of Witt vectors of length r is denoted by W_r (Example 1.11).

Definition 3.1. Let $\mathrm{fil}_n W_r(K)$ be the subgroup of $W_r(K)$ from [Brylinski 1983, Section 1, Proposition 1]:

$$\mathrm{fil}_n W_r(K) = \{(f_{r-1}, \ldots, f_0) \mid f_i \in K, \mathrm{v}(f_i) \geq -n/p^i \text{ for all } 0 \leq i \leq r - 1\}.$$

Let $\mathrm{fil}_n^{\mathrm{F}} W_r(K)$ be the subgroup of $W_r(K)$ generated by $\mathrm{fil}_n W_r(K)$ by means of the Frobenius F (see [Kato and Russell 2010, 2.2]):

$$\mathrm{fil}_n^{\mathrm{F}} W_r(K) = \sum_{\nu \geq 0} \mathrm{F}^\nu \mathrm{fil}_n W_r(K).$$

Let X be a variety over k, regular in codimension 1. Let $D = \sum_{q \in S} n_q D_q$ be an effective divisor on X, where S is a finite set of points of codimension 1 in X, where D_q are the prime divisors associated to $q \in S$ and n_q are positive integers for $q \in S$.

Definition 3.2. Let $\mathrm{fil}_D W_r(\mathcal{K}_X)$ and $\mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X)$ be the sheaves of subgroups of $W_r(\mathcal{K}_X)$ formed by the groups

$$(\mathrm{fil}_D W_r(\mathcal{K}_X))(U) = \left\{ w \in W_r(\mathcal{K}_X) \mid \begin{array}{ll} w \in \mathrm{fil}_{n_q} W_r(\mathcal{K}_{X,q}) & \text{for all } q \in S \cap U, \\ w \in W_r(\mathcal{O}_{X,p}) & \text{for all } p \in U \setminus S \end{array} \right\},$$

$$(\mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X))(U) = \left\{ w \in W_r(\mathcal{K}_X) \mid \begin{array}{ll} w \in \mathrm{fil}_{n_q}^{\mathrm{F}} W_r(\mathcal{K}_{X,q}) & \text{for all } q \in S \cap U, \\ w \in W_r(\mathcal{O}_{X,p}) & \text{for all } p \in U \setminus S \end{array} \right\},$$

respectively, for open $U \in X$, where $\mathrm{fil}_n W_r(\mathcal{K}_{X,q})$ and $\mathrm{fil}_n^{\mathrm{F}} W_r(\mathcal{K}_{X,q})$ denote the filtrations associated to the valuation v_q attached to the point $q \in S$.

Proposition 3.3. *Suppose X is a projective variety over k and D an effective divisor on X. Then $\Gamma(X, \mathrm{fil}_D W_r(\mathcal{K}_X))$ is a finite $W_r(k)$-module.*

Proof. The Verschiebung $\mathrm{V} : W_{r-1} \to W_r$ yields an exact sequence

$$0 \to \mathrm{fil}_D W_{r-1}(\mathcal{K}_X) \to \mathrm{fil}_D W_r(\mathcal{K}_X) \to \mathrm{fil}_{\lfloor D/p^{r-1} \rfloor} W_1(\mathcal{K}_X) \to 0,$$

where $\lfloor D/p^{r-1} \rfloor = \sum_{q \in S} \lfloor n_q/p^{r-1} \rfloor D_q$. This induces the exact sequence

$$0 \to \Gamma(\mathrm{fil}_D W_{r-1}(\mathcal{K}_X)) \to \Gamma(\mathrm{fil}_D W_r(\mathcal{K}_X)) \to \Gamma(\mathrm{fil}_{\lfloor D/p^{r-1} \rfloor} W_1(\mathcal{K}_X)).$$

By induction over $r \geq 1$ and since $W_1(k) = k$ is noetherian, it is sufficient to show the statement for $r = 1$. Now $\mathrm{fil}_D W_1(\mathcal{K}_X) = \mathcal{O}_X(D)$ is a coherent sheaf, and hence $\Gamma(X, \mathrm{fil}_D W_1(\mathcal{K}_X))$ is a finite module over $W_1(k) = k$. $\square$

Definition 3.4. Let R be a commutative ring over $\mathbb{F}_p$. We let $R[\mathrm{F}]$ be the noncommutative polynomial ring defined by

$$R[\mathrm{F}] = \left\{ \sum_{i=1}^{n} \mathrm{F}^i \, r_i \;\middle|\; r_i \in R, n \in \mathbb{N} \right\}, \quad \text{where } \mathrm{F}\,r = r^p\,\mathrm{F} \quad \text{for all } r \in R.$$

Definition 3.5. If $\Omega_{\mathcal{K}_X} = \Omega_{\mathcal{K}_X/k}$ is the module of differentials of $\mathcal{K}_X$ over k, we let δ be the homomorphism

$$\delta : W_r(\mathcal{K}_X) \to \Omega_{\mathcal{K}_X}, \quad (f_{r-1}, \ldots, f_0) \mapsto \sum_{i=0}^{r-1} f_i^{p^i-1} \, \mathrm{d} f_i.$$

Definition 3.6. If E is a reduced effective divisor on X with normal crossings, we let $\Omega_X(\log E)$ be the sheaf of differentials on X with log-poles along E, that is, the $\mathcal{O}_X$-module generated locally by $\mathrm{d} f$ for $f \in \mathcal{O}_X$ and $\mathrm{d}\log t = t^{-1}\mathrm{d}t$, where t is a local equation for E.

Proposition 3.7. *Suppose D_{red} is a normal crossing divisor. The homomorphism δ from Definition 3.5 induces injective homomorphisms*

$$\eth_D : \mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X)/\mathrm{fil}_{\lfloor D/p \rfloor}^{\mathrm{F}} W_r(\mathcal{K}_X) \to \mathfrak{D}_D,$$
$$\overline{\eth}_D : \mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X)/\mathrm{fil}_{D-D_{\mathrm{red}}}^{\mathrm{F}} W_r(\mathcal{K}_X) \to \overline{\mathfrak{D}}_D,$$

where $\mathfrak{D}_D$ and $\overline{\mathfrak{D}}_D$ are the $\mathcal{O}_X$-modules

$$\mathfrak{D}_D = k[\mathrm{F}] \otimes_k (\Omega_X(\log D_{\mathrm{red}}) \otimes_{\mathcal{O}_X} \mathcal{O}_X(D)/\mathcal{O}_X(\lfloor D/p \rfloor)),$$
$$\overline{\mathfrak{D}}_D = k[\mathrm{F}] \otimes_k (\Omega_X(\log D_{\mathrm{red}}) \otimes_{\mathcal{O}_X} \mathcal{O}_X(D)/\mathcal{O}_X(D - D_{\mathrm{red}})),$$

and $\lfloor D/p \rfloor$ means the largest divisor E such that $pE \leq D$.

Proof. This is the global formulation of [Kato and Russell 2010, 4.6]. $\qquad\square$

Definition 3.8. Let ${}^{\flat}\overline{\mathfrak{D}}_D$ be the image in $\overline{\mathfrak{D}}_D$ of the $\mathcal{O}_X$-module

$$k[\mathrm{F}] \otimes_k (\Omega_X \otimes_{\mathcal{O}_X} \mathcal{O}_X(D)/\mathcal{O}_X(D - D_{\mathrm{red}}))$$

(without log-poles). Then

$${}^{\flat}\overline{\mathfrak{D}}_D \cong k[\mathrm{F}] \otimes_k (\Omega_{D_{\mathrm{red}}} \otimes_{\mathcal{O}_{D_{\mathrm{red}}}} \mathcal{O}_X(D)/\mathcal{O}_X(D - D_{\mathrm{red}}))$$

since $t^{-n_q}\mathrm{d}t = t^{1-n_q}\mathrm{d}\log t$ vanishes in $\overline{\mathfrak{D}}_D$ for any local equation t of D_{red}. Then we let ${}^{\flat}\mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X) \subset \mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X)$ be the inverse image of ${}^{\flat}\overline{\mathfrak{D}}_D$ under the map $\overline{\eth}_D$ from Proposition 3.7. According to [Kato and Russell 2010, 4.7], this is a global version of the following alternative definition:

Definition 3.9. Let ${}^{\flat}\mathrm{fil}_n W_r(K)$ be the subgroup of $\mathrm{fil}_n W_r(K)$ consisting of all elements $(f_{r-1}, \ldots, f_0)$ satisfying the following condition: If the p-adic order ν of n is less than r, then $p^{\nu} \mathrm{v}(f_{\nu}) > -n$. Then ${}^{\flat}\mathrm{fil}_n^{\mathrm{F}} W_r(K)$ is the subgroup of $W_r(K)$ generated by ${}^{\flat}\mathrm{fil}_n W_r(K)$ by means of the Frobenius F,

$$
{}^{\flat}\mathrm{fil}_n^{\mathrm{F}} W_r(K) = \sum_{\nu \geq 0} \mathrm{F}^{\nu} \, {}^{\flat}\mathrm{fil}_n W_r(K).
$$

Lemma 3.10. *Let $\psi : Y \to X$ be a morphism of varieties over k, such that $\psi(Y)$ intersects $\mathrm{Supp}(D)$ properly. Let $D \cdot Y$ denote the pull-back of D to Y. Suppose that D_{red} and $(D \cdot Y)_{\mathrm{red}}$ are normal crossing divisors. There is a commutative diagram of homomorphisms with injective rows*

$$
\begin{array}{ccc}
\mathrm{fil}_D^{\mathrm{F}} W_r(\mathcal{K}_X)/\mathrm{fil}_{\lfloor D/p \rfloor}^{\mathrm{F}} W_r(\mathcal{K}_X) & \xrightarrow{\partial_{X,D}} & \mathfrak{D}_{X,D} \\
\downarrow & & \downarrow \\
\mathrm{fil}_{D \cdot Y}^{\mathrm{F}} W_r(\mathcal{K}_Y)/\mathrm{fil}_{\lfloor D \cdot Y/p \rfloor}^{\mathrm{F}} W_r(\mathcal{K}_Y) & \xrightarrow{\partial_{Y,D \cdot Y}} & \mathfrak{D}_{Y,D \cdot Y},
\end{array}
$$

where the vertical arrows are the obvious pull-back maps from X to Y.

Proof. Straightforward. $\qquad\square$

3.2. *Albanese with modulus.*

3.2.1. *Existence and construction.* Let X be a smooth proper variety over a perfect field k.

Definition 3.11. First assume k is algebraically closed. Let $\varphi : X \dashrightarrow G$ be a rational map from X to a smooth connected algebraic group G. Let L be the affine part of G and U the unipotent part of L. The *modulus of φ* from [Kato and Russell 2010, §3] is the following effective divisor

$$
\mathrm{mod}(\varphi) = \sum_{\mathrm{ht}(q)=1} \mathrm{mod}_q(\varphi) D_q
$$

where q ranges over all points of codimension 1 in X, and D_q is the prime divisor associated to q. For each $q \in X$ of codimension 1, the canonical map $L(\mathcal{K}_{X,q})/L(\mathcal{O}_{X,q}) \to G(\mathcal{K}_{X,q})/G(\mathcal{O}_{X,q})$ is bijective; see Section 2.2.1. Take an element $l_q \in L(\mathcal{K}_{X,q})$ whose image in $G(\mathcal{K}_{X,q})/G(\mathcal{O}_{X,q})$ coincides with the class of $\varphi \in G(\mathcal{K}_{X,q})$. If $\mathrm{char}(k) = 0$, let $(u_{q,i})_{1 \leq i \leq s}$ be the image of l_q in $\mathbb{G}_a(\mathcal{K}_{X,q})^s$ under $L \to U \cong (\mathbb{G}_a)^s$. If $\mathrm{char}(k) = p > 0$, let $(u_{q,i})_{1 \leq i \leq s}$ be the image of l_q in $W_r(\mathcal{K}_{X,q})^s$ under $L \to U \subset (W_r)^s$.

$$
\mathrm{mod}_q(\varphi) = \begin{cases} 0 & \text{if } \varphi \in G(\mathcal{O}_{X,q}), \\ 1 + \max\{\mathrm{n}_q(u_{q,i}) \mid 1 \leq i \leq s\} & \text{if } \varphi \notin G(\mathcal{O}_{X,q}), \end{cases}
$$

where for $u \in \mathbb{G}_a(\mathcal{H}_{X,q})$ if $\mathrm{char}(k) = 0$, or $W_r(\mathcal{H}_{X,q})$ if $\mathrm{char}(k) = p > 0$, we denote

$$
\mathrm{n}_q(u) = \begin{cases} -\mathrm{v}_q(u) & \text{if } \mathrm{char}(k) = 0, \\ \min\{n \in \mathbb{N} \mid u \in \mathrm{fil}_n^{\mathrm{F}} W_r(\mathcal{H}_{X,q})\} & \text{if } \mathrm{char}(k) = p > 0. \end{cases}
$$

Note that $\mathrm{mod}_v(\varphi)$ is independent of the choice of the isomorphism $U \cong (\mathbb{G}_a)^s$ or, respectively, of the embedding $U \subset (W_r)^s$; see [Kato and Russell 2010, Theorem 3.3].

For arbitrary perfect base field k and G a torsor for a smooth connected algebraic group we obtain $\mathrm{mod}(\varphi)$ by means of a Galois descent from $\mathrm{mod}(\varphi \otimes_k \bar{k})$, where $\bar{k}$ is an algebraic closure of k; see [Kato and Russell 2010, No. 3.4] and Remark 2.10.

Definition 3.12. Let D be an effective divisor on X. Then $\mathrm{Mr}(X, D)$ denotes the category of those rational maps φ from X to torsors such that $\mathrm{mod}(\varphi) \le D$. The universal object of $\mathrm{Mr}(X, D)$ (if it exists) is called the *Albanese of X of modulus D* and denoted by $\mathrm{Alb}^{(1)}(X, D)$, or just $\mathrm{Alb}(X, D)$, if it admits a k-rational point (cf. Remark 2.10).

Remark 3.13. In the definition of $\mathrm{mod}(\varphi)$ (Definition 3.11) we used, instead of the original filtration $\mathrm{fil}_\bullet W$ of Brylinski, the saturation $\mathrm{fil}_\bullet^{\mathrm{F}} W$ of $\mathrm{fil}_\bullet W$ with respect to the Frobenius. This is motivated as follows: Let $\mathrm{mod}^\dashv(\varphi)$ be the modulus of a rational map φ using the filtration $\mathrm{fil}_\bullet W$ instead of $\mathrm{fil}_\bullet^{\mathrm{F}} W$. If $\varphi : X \dashrightarrow \mathbb{G}_a$ is a nonconstant rational map, that is, the multiplicity of $\mathrm{mod}(\varphi) =: D$ is greater than 1, then $\mathrm{mod}^\dashv(\mathrm{F}^\nu \circ \varphi) = p^\nu(D - D_{\mathrm{red}}) + D_{\mathrm{red}}$, where D_{red} is the reduced part of D. On the other hand, if $u : X \dashrightarrow \mathcal{U}$ is a universal object for a certain category of rational maps Mr, then clearly u satisfies the universal mapping property as well for all maps of the form $\mathrm{F}^\nu \circ \varphi$, $\varphi \in \mathrm{Mr}$ (cf. condition $(2^\diamond)$ before Theorem 2.16). This shows that $\mathrm{mod}^\dashv(\varphi)$ is not compatible with the notion of universal objects.

Definition 3.14. Let D be an effective divisor on X, and let D_{red} be the reduced part of D. Then $\mathscr{F}_{X,D}$ denotes the formal subgroup of $\underline{\mathrm{Div}}_X$ characterized by

$$
(\mathscr{F}_{X,D})_{\mathrm{\acute{e}t}} = \big\{ B \in \underline{\mathrm{Div}}_X(k) \mid \mathrm{Supp}(B) \subset \mathrm{Supp}(D) \big\},
$$

and for $\mathrm{char}(k) = 0$,

$$
(\mathscr{F}_{X,D})_{\mathrm{inf}} = \exp\big(\widehat{\mathbb{G}}_a \otimes_k \Gamma(\mathbb{O}_X(D - D_{\mathrm{red}})/\mathbb{O}_X) \big),
$$

for $\mathrm{char}(k) = p > 0$,

$$
(\mathscr{F}_{X,D})_{\mathrm{inf}} = \mathrm{Exp}\Big(\sum_{r>0} {}_r\widehat{W} \otimes_{W_r(k)} \Gamma\big(\mathrm{fil}_{D-D_{\mathrm{red}}}^{\mathrm{F}} W_r(\mathcal{H}_X)/W_r(\mathbb{O}_X) \big), 1 \Big).
$$

Let $\mathscr{F}_{X,D}^{0,\mathrm{red}} = \mathscr{F}_{X,D} \times_{\underline{\mathrm{Div}}_X} \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$ be the intersection of $\mathscr{F}_{X,D}$ and $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$.

Proposition 3.15. *The formal groups $\mathscr{F}_{X,D}$ and $\mathscr{F}_{X,D}^{0,\mathrm{red}}$ are dual-algebraic.*

Proof. The statement is obvious for $\mathrm{char}(k) = 0$, so we suppose $\mathrm{char}(k) = p > 0$. The proof is done in two steps. Let $\mathscr{F}^{\dashv}_{X,D}$ be the formal subgroup of $\underline{\mathrm{Div}}_X$ defined in the same way as $\mathscr{F}_{X,D}$, but using the filtration $\mathrm{fil}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)$ instead of $\mathrm{fil}^{\mathrm{F}}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)$. In the first step, we show that for any effective divisor D the formal group $\mathscr{F}^{\dashv}_{X,D}$ is dual-algebraic. In the second step, we show that for any D there exists $D' \geq D$ such that $\mathscr{F}_{X,D}$ is contained in the image of $\mathscr{F}^{\dashv}_{X,D'}$ in $\mathscr{F}_{X,D'}$. Thus $\mathscr{F}_{X,D}$ is a formal subgroup of a quotient of a dual-algebraic formal group, and hence dual-algebraic by Lemma 1.17. Then also the formal subgroup $\mathscr{F}^{0,\mathrm{red}}_{X,D}$ of $\mathscr{F}_{X,D}$ is dual-algebraic.

Step 1. Let D be an effective divisor on X. Write $D = \sum_{\mathrm{ht}(q)=1} n_q D_q$, where q ranges over all points of codimension 1 in X, and D_q is the prime divisor associated to q. Let S be the finite set of those q with $n_q > 0$. Let

$$m = \min\{r \mid p^r > n_q - 1 \text{ for all } q \in S\}.$$

Hence for $r \geq m$, if $(f_{r-1}, \ldots, f_0) \in \mathrm{fil}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)$, then $f_i \in \mathcal{O}_X$ for $r > i \geq m$, according to Definition 3.2. Then the Verschiebung $\mathrm{V}^{r-m} : W_m(\mathcal{K}_X) \to W_r(\mathcal{K}_X)$ yields a surjective homomorphism

$$\mathrm{fil}_{D-D_{\mathrm{red}}} W_m(\mathcal{K}_X)/W_m(\mathcal{O}_X) \twoheadrightarrow \mathrm{fil}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X).$$

Thus $(\mathscr{F}^{\dashv}_{X,D})_{\mathrm{inf}}$ is already generated by a finite sum via Exp:

$$(\mathscr{F}^{\dashv}_{X,D})_{\mathrm{inf}} = \mathrm{Exp}\Big(\sum_{1 \leq r \leq m} {}_r\widehat{W} \otimes_{W_r(k)} \Gamma\big(\mathrm{fil}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X)\big), 1\Big).$$

Each $\Gamma(X, \mathrm{fil}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X))$ is a finitely generated $W_r(k)$-module, by the same proof as for Proposition 3.3. Hence $(\mathscr{F}^{\dashv}_{X,D})_{\mathrm{inf}}$ is a quotient of the direct sum of finitely many ${}_r\widehat{W}$.

Moreover, $(\mathscr{F}^{\dashv}_{X,D})_{\mathrm{ét}} = (\mathscr{F}_{X,D})_{\mathrm{ét}}$ is an abelian group of finite type, since D has only finitely many components. Thus $\mathscr{F}^{\dashv}_{X,D}$ is dual-algebraic, according to Proposition 1.16.

Step 2. We show that for any effective divisor D there exists an effective divisor $D' \geq D$ such that $\mathscr{F}_{X,D}$ is generated by $\mathscr{F}^{\dashv}_{X,D'}$. We will find an effective divisor $D' \geq D$ such that $\Gamma(\mathrm{fil}^{\mathrm{F}}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X))$ is generated by

$$\sum_{\nu \geq 0} \mathrm{F}^\nu \, \Gamma(\mathrm{fil}_{D'-D'_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X)).$$

This is sufficient because

$$\mathrm{Exp}\big(v \otimes \textstyle\sum_i \mathrm{F}^{\nu_i} \omega_i, 1\big) = \mathrm{Exp}\big(\textstyle\sum_i \mathrm{V}^{\nu_i} v \otimes \omega_i, 1\big).$$

Since the homomorphism

$$\mathrm{V}^{r-m} : \mathrm{fil}^{\mathrm{F}}_{D-D_{\mathrm{red}}} W_m(\mathcal{K}_X)/W_m(\mathcal{O}_X) \to \mathrm{fil}^{\mathrm{F}}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X)$$

is surjective for $r \geq m$, we only need to consider $r = m$.

The exact sequence

$$0 \to W_r(\mathcal{O}_X) \to W_r(\mathcal{K}_X) \to W_r(\mathcal{K}_X)\big/W_r(\mathcal{O}_X) \to 0$$

yields the exact sequence

$$\Gamma(W_r(\mathcal{K}_X)) \to \Gamma(W_r(\mathcal{K}_X)\big/W_r(\mathcal{O}_X)) \to \mathrm{H}^1(W_r(\mathcal{O}_X)) \to 0.$$

Here $\mathrm{H}^1(W_r(\mathcal{K}_X)) = 0$ since $W_r(\mathcal{K}_X)$ is a flasque sheaf. Since $\mathrm{H}^1(W_r(\mathcal{O}_X))$ is a finite $W_r(k)$-module, there is an effective divisor E such that the map

$$\Gamma(\mathrm{fil}_E W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X)) \to \mathrm{H}^1(W_r(\mathcal{O}_X))$$

is surjective. Hence for any $\sigma \in \Gamma(\mathrm{fil}^{\mathrm{F}}_{D-D_{\mathrm{red}}} W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X))$ there is

$$\rho \in \Gamma(\mathrm{fil}_E W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X))$$

such that $\sigma - \rho$ lies in the image of $\Gamma(W_r(\mathcal{K}_X))$, and hence in the image of $\Gamma(\mathrm{fil}^{\mathrm{F}}_{E'} W_r(\mathcal{K}_X))$, where $E' = \max\{E, D - D_{\mathrm{red}}\}$. Therefore we are reduced to showing that for any D there exists $D' \geq D$ such that $\Gamma(\mathrm{fil}^{\mathrm{F}}_D W_r(\mathcal{K}_X))$ is generated by

$$\sum_{\nu \geq 0} \mathrm{F}^\nu \, \Gamma(\mathrm{fil}_{D'} W_r(\mathcal{K}_X)).$$

Consider the exact sequence

$$0 \to \bigoplus_{\nu \geq 0} \mathrm{fil}_{\lfloor D/p \rfloor} W_r(\mathcal{K}_X) \to \bigoplus_{\nu \geq 0} \mathrm{fil}_D W_r(\mathcal{K}_X) \to \mathrm{fil}^{\mathrm{F}}_D W_r(\mathcal{K}_X) \to 0$$

where the third arrow is $(w_\nu)_\nu \mapsto \sum_\nu \mathrm{F}^\nu w_\nu$, and the second arrow is $(w_\nu)_\nu \mapsto (\mathrm{F}\, w_\nu - w_{\nu-1})_\nu$, where we set $w_{-1} = 0$. Here $\lfloor D/p \rfloor$ means the largest divisor E such that $pE \leq D$. This yields an exact sequence

$$\bigoplus_{\nu \geq 0} \Gamma(\mathrm{fil}_D W_r(\mathcal{K}_X)) \to \Gamma(\mathrm{fil}^{\mathrm{F}}_D W_r(\mathcal{K}_X)) \to \bigoplus_{\nu \geq 0} \mathrm{H}^1(\mathrm{fil}_{\lfloor D/p \rfloor} W_r(\mathcal{K}_X)).$$

$W_r(\mathcal{K}_X)$ is the inductive limit of $\mathrm{fil}_E W_r(\mathcal{K}_X)$, where E ranges over all effective divisors on X, and hence

$$0 = \mathrm{H}^1(W_r(\mathcal{K}_X)) = \mathrm{H}^1\left(\varinjlim_E \mathrm{fil}_E W_r(\mathcal{K}_X)\right) = \varinjlim_E \mathrm{H}^1(\mathrm{fil}_E W_r(\mathcal{K}_X)).$$

As $H^1(\mathrm{fil}_{\lfloor D/p \rfloor} W_r(\mathcal{H}_X))$ is a finite $W_r(k)$-module, there is an effective divisor $D' \geq D$ such that the image of $H^1(\mathrm{fil}_{\lfloor D/p \rfloor} W_r(\mathcal{H}_X))$ in $H^1(\mathrm{fil}_{\lfloor D'/p \rfloor} W_r(\mathcal{H}_X))$ is 0. Thus the image of $\Gamma(\mathrm{fil}^F_D W_r(\mathcal{H}_X))$ in $\Gamma(\mathrm{fil}^F_{D'} W_r(\mathcal{H}_X))$ is contained in

$$\sum_{\nu \geq 0} F^\nu \, \Gamma(\mathrm{fil}_{D'} W_r(\mathcal{H}_X)). \qquad \square$$

Lemma 3.16. *Let $\varphi : X \dashrightarrow G$ be a rational map from X to a smooth connected algebraic group G. Then the following conditions are equivalent*:

(i) $\mathrm{mod}(\varphi) \leq D$.

(ii) $\mathrm{im}(\tau_\varphi) \subset \mathcal{F}_{X,D}$.

Proof. Write $D = \sum_{\mathrm{ht}(q)=1} n_q D_q$, where q ranges over all points in X of codimension 1, and D_q is the prime divisor associated to q. Condition (i) is thus expressed by the condition that for all $q \in X$ of codimension 1 we have

(i)$_q$ $\mathrm{mod}_q(\varphi) \leq n_q$.

Using the canonical splitting of a formal group into an étale and an infinitesimal part, condition (ii) is equivalent to the condition that the following (ii)$_{\text{ét}}$ and (ii)$_{\text{inf}}$ are satisfied:

(ii)$_{\text{ét}}$ $\mathrm{im}(\tau_{\varphi,\text{ét}}) \subset (\mathcal{F}_{X,D})_{\text{ét}}$.

(ii)$_{\text{inf}}$ $\mathrm{im}(\tau_{\varphi,\text{inf}}) \subset (\mathcal{F}_{X,D})_{\text{inf}}$.

Let L be the affine part of G. Remember from Section 2.2.1 that the transformation $\tau_\varphi : L^\vee \to \widehat{\mathrm{Div}}_X$ is given by $\langle \cdot, \ell_\varphi \rangle$, where ℓ_φ is the image of $\varphi \in G(\mathcal{H}_X)$ in $\Gamma(G(\mathcal{H}_X)/G(\mathcal{O}_X)) \xrightarrow{\sim} \Gamma(L(\mathcal{H}_X)/L(\mathcal{O}_X))$, and the pairing

$$\langle \cdot, \cdot \rangle : L^\vee \times \Gamma(L(\mathcal{H}_X)/L(\mathcal{O}_X)) \to \Gamma\big(\mathbb{G}_m(\mathcal{H}_X)/\mathbb{G}_m(\mathcal{O}_X)\big).$$

is obtained from Cartier duality. Write $L = T \times_k U$ as a product of a torus T and a unipotent group U. Fix an isomorphism $T \cong (\mathbb{G}_m)^m$ and an isomorphism $U \cong (\mathbb{G}_a)^a$ if $\mathrm{char}(k) = 0$, or an embedding $U \subset (W_r)^a$ if $\mathrm{char}(k) = p > 0$.

Let $(t_j)_{1 \leq j \leq m}$ be the image of ℓ_φ under

$$\Gamma(L(\mathcal{H}_X)/L(\mathcal{O}_X)) \to \Gamma(T(\mathcal{H}_X)/T(\mathcal{O}_X)) \to \Gamma(\mathbb{G}_m(\mathcal{H}_X)/\mathbb{G}_m(\mathcal{O}_X))^m$$

and $(u_i)_{1 \leq i \leq a}$ be the image of ℓ_φ under

$$\Gamma(L(\mathcal{H}_X)/L(\mathcal{O}_X)) \to \Gamma(U(\mathcal{H}_X)/U(\mathcal{O}_X)) \to \begin{cases} \Gamma(\mathbb{G}_a(\mathcal{H}_X)/\mathbb{G}_a(\mathcal{O}_X))^a, \\ \Gamma(W_r(\mathcal{H}_X)/W_r(\mathcal{O}_X))^a. \end{cases}$$

The étale part of τ_φ is

$$\tau_{\varphi,\text{ét}} : \mathbb{Z}^m \to \Gamma(\mathbb{G}_m(\mathcal{K}_X)/\mathbb{G}_m(\mathcal{O}_X)),$$

$$(e_j)_{1 \le j \le m} \mapsto \prod_{j=1}^{m} t_j^{e_j}.$$

The image of the infinitesimal part of τ_φ is given by the image of

$$\widetilde{\tau}_{\varphi,\text{inf}} : \left. \begin{matrix} (\widehat{\mathbb{G}}_a)^a \\ (_r\widehat{W})^a \end{matrix} \right\} \to \Gamma(\mathbb{G}_m(\mathcal{K}_X \otimes \cdot)/\mathbb{G}_m(\mathcal{O}_X \otimes \cdot)),$$

$$(v_i)_{1 \le i \le a} \mapsto \begin{cases} \prod_{i=1}^{a} \exp(v_i\, u_i), \\ \prod_{i=1}^{a} \text{Exp}(v_i \cdot u_i, 1); \end{cases}$$

see Example 1.11 for the pairing $_r\widehat{W} \times W_r \to \mathbb{G}_m$. For each $q \in X$ of codimension 1 let $(t_{q,i})_{1 \le j \le m}$ be a representative in $\mathbb{G}_m(\mathcal{K}_{X,q})^m$ of the image of $(t_j)_{1 \le j \le m}$ under

$$\Gamma(\mathbb{G}_m(\mathcal{K}_X)/\mathbb{G}_m(\mathcal{O}_X))^m \to \mathbb{G}_m(\mathcal{K}_{X,q})^m/\mathbb{G}_m(\mathcal{O}_{X,q})^m,$$

and let $(u_{q,i})_{1 \le i \le a}$ be a representative in $\mathbb{G}_a(\mathcal{K}_{X,q})^a$ or $W_r(\mathcal{K}_{X,q})^a$ of the image of $(u_i)_{1 \le i \le a}$ under

$$\Gamma(\mathbb{G}_a(\mathcal{K}_X)/\mathbb{G}_a(\mathcal{O}_X))^a \to \mathbb{G}_a(\mathcal{K}_{X,q})^a/\mathbb{G}_a(\mathcal{O}_{X,q})^a$$

or

$$\Gamma(W_r(\mathcal{K}_X)/W_r(\mathcal{O}_X))^a \to W_r(\mathcal{K}_{X,q})^a/W_r(\mathcal{O}_{X,q})^a,$$

respectively. Then (ii)$_\text{ét}$ is equivalent to the following condition being satisfied for every point $q \in X$ of codimension 1:

(ii)$_{\text{ét},q}$ If $n_q = 0$, then $t_{q,j} \in \mathbb{G}_m(\mathcal{O}_{X,q})$ for $1 \le j \le m$.

On the other hand, condition (ii)$_\text{inf}$ is equivalent to the following condition being satisfied for every point $q \in X$ of codimension 1, according to Definition 3.14 of $\mathcal{F}_{X,D}$:

(ii)$_{\text{inf},q}$ If $n_q = 0$, then $u_{q,i} \in \mathbb{G}_a(\mathcal{O}_{X,q})$ or $W_r(\mathcal{O}_{X,q})$ for $1 \le i \le a$.
      If $n_q > 0$, then $\text{n}_q(u_{q,i}) \le n_q - 1$.

Note that $\varphi \in G(\mathcal{O}_{X,q})$ if and only if $t_{q,j} \in \mathbb{G}_m(\mathcal{O}_{X,q})$ for $1 \le j \le m$ and $u_{q,i} \in \mathbb{G}_a(\mathcal{O}_{X,q})$ or $W_r(\mathcal{O}_{X,q})$ for $1 \le i \le a$. By Definition 3.11, for each $q \in X$ of codimension 1

(i)$_q$ $\text{mod}_q(\varphi) \le n_q$

if and only if (ii)$_{\text{ét},q}$ and (ii)$_{\text{inf},q}$ are satisfied.   □

Now assume k is algebraically closed. Arbitrary perfect base field is considered in Sections 3.2.2 and 3.2.3.

Theorem 3.17. *The category* $\mathrm{Mr}(X, D)$ *of rational maps of modulus* $\leq D$ *is equivalent to the category* $\mathrm{Mr}_{\mathscr{F}_{X,D}}(X)$ *of rational maps that induce a transformation to* $\mathscr{F}_{X,D}$.

Proof. According to the definitions of $\mathrm{Mr}(X, D)$ and $\mathrm{Mr}_{\mathscr{F}_{X,D}}(X)$, the statement is due to Lemma 3.16. $\qquad\square$

Theorem 3.18. *The Albanese* $\mathrm{Alb}(X, D)$ *of* X *of modulus* D *exists and is dual (in the sense of* 1*-motives) to the* 1*-motive* $\left[\mathscr{F}_{X,D}^{0,\mathrm{red}} \to \mathrm{Pic}_X^{0,\mathrm{red}}\right]$.

Proof. By Theorem 3.17, $\mathrm{Alb}(X, D)$ is the universal object of $\mathrm{Mr}_{\mathscr{F}_{X,D}}(X)$ (if it exists). A rational map from X to an algebraic group induces a transformation to $\mathscr{F}_{X,D}$ if and only if it induces a transformation to $\mathscr{F}_{X,D}^{0,\mathrm{red}}$, by Lemma 2.6. Since $\mathscr{F}_{X,D}^{0,\mathrm{red}}$ is dual-algebraic (Proposition 3.15), the category $\mathrm{Mr}_{\mathscr{F}_{X,D}}(X)$ admits a universal object (Theorem 2.16), and this universal object is dual to $\left[\mathscr{F}_{X,D}^{0,\mathrm{red}} \to \mathrm{Pic}_X^{0,\mathrm{red}}\right]$ (Remark 2.18). $\qquad\square$

3.2.2. *Descent of the base field.* Let k be a perfect field. Let $\bar{k}$ be an algebraic closure of k. Let X be a smooth proper variety defined over k, and let D be an effective divisor on X rational over k.

Theorem 3.19. *There exists a k-torsor* $\mathrm{Alb}^{(1)}(X, D)$ *for an algebraic k-group* $\mathrm{Alb}^{(0)}(X, D)$ *and rational maps defined over k*

$$\mathrm{alb}_{X,D}^{(i)} : X^{2-i} \dashrightarrow \mathrm{Alb}^{(i)}(X, D)$$

for $i = 1, 0$, satisfying the following universal property:

If $\varphi : X \dashrightarrow G^{(1)}$ is a rational map defined over k to a k-torsor $G^{(1)}$ for an algebraic k-group $G^{(0)}$, such that $\mathrm{mod}(\varphi) \leq D$, then there exist a unique homomorphism of k-torsors $h^{(1)} : \mathrm{Alb}^{(1)}(X, D) \to G^{(1)}$ and a unique homomorphism of algebraic k-groups $h^{(0)} : \mathrm{Alb}^{(0)}(X, D) \to G^{(0)}$, defined over k, such that $\varphi^{(i)} = h^{(i)} \circ \mathrm{alb}_{X,D}^{(i)}$ for $i = 1, 0$.

Here $\mathrm{Alb}^{(0)}(X, D)$ is dual to the 1*-motive* $\left[\mathscr{F}_{X,D}^{0,\mathrm{red}} \to \mathrm{Pic}_X^{0,\mathrm{red}}\right]$.

Proof. It follows directly from Theorem 2.21 and the definition of the modulus via Galois descent (Definition 3.11). $\qquad\square$

Corollary 3.20. *For every rational map $\varphi : X \dashrightarrow P$ from X to a torsor P there exists an effective divisor D, namely $D = \mathrm{mod}(\varphi)$, such that φ factors through* $\mathrm{Alb}^{(1)}(X, D)$.

Proposition 3.21. *Let $\mathscr{F}$ be a formal k-subgroup of* $\underline{\mathrm{Div}}_X^{0,\mathrm{red}}$. *Then $\mathscr{F}$ is dual-algebraic if and only if there exists an effective divisor D, rational over k, such that* $\mathscr{F} \subset \mathscr{F}_{X,D}$.

Proof. ($\Longrightarrow$) A formal subgroup of a dual-algebraic group is also dual-algebraic, according to Lemma 1.17.

($\Longleftarrow$) By Galois descent (possible for formal groups due to Cartier duality) we may assume that k is algebraically closed. Let $D = \mathrm{mod}(\mathrm{alb}_{\mathscr{F}})$ be the modulus of the universal rational map $\mathrm{alb}_{\mathscr{F}} : X \to \mathrm{Alb}_{\mathscr{F}}(X)$ associated to $\mathscr{F} \subset \underline{\mathrm{Div}}_X^{0,\mathrm{red}}$. Then by Lemma 3.16, we have $\mathscr{F} = \mathrm{im}(\tau_{\mathrm{alb}_{\mathscr{F}}}) \subset \mathscr{F}_{X,D}$. $\qquad\square$

3.2.3. *Functoriality.* We specialize the results from Section 2.3.3 to the case of Albanese varieties with modulus.

Proposition 3.22. *Let $\psi : Y \to X$ be a morphism of smooth proper varieties. Let D be an effective divisor on X intersecting $\psi(Y)$ properly. Then ψ induces a homomorphism of torsors $\mathrm{Alb}^{(1)}{}_{Y,E}^{X,D}(\psi)$ and a homomorphism of algebraic groups $\mathrm{Alb}^{(0)}{}_{Y,E}^{X,D}(\psi)$,*

$$\mathrm{Alb}^{(i)}{}_{Y,E}^{X,D}(\psi) : \mathrm{Alb}^{(i)}(Y, E) \to \mathrm{Alb}^{(i)}(X, D),$$

for each effective divisor E on Y satisfying $E \geq (D - D_{\mathrm{red}}) \cdot Y + (D \cdot Y)_{\mathrm{red}}$, where $B \cdot Y$ denotes the pull-back of a Cartier divisor B on X to Y.

Proof. According to Proposition 2.22, for the existence of $\mathrm{Alb}_{Y,E}^{X,D}(\psi)$ it is sufficient to show $\mathscr{F}_{Y,E} \supset \mathscr{F}_{X,D} \cdot Y$. Definition 3.14 of $\mathscr{F}_{X,D}$ implies that this is the case if and only if $\mathrm{Supp}(E) \supset \mathrm{Supp}(D \cdot Y)$ and $E - E_{\mathrm{red}} \geq (D - D_{\mathrm{red}}) \cdot Y$. But this is equivalent to $E \geq (D - D_{\mathrm{red}}) \cdot Y + (D \cdot Y)_{\mathrm{red}}$. $\qquad\square$

Corollary 3.23. *If D and E are effective divisors on X with $E \geq D$, then there are canonical surjective homomorphisms $\mathrm{Alb}^{(i)}(X, E) \twoheadrightarrow \mathrm{Alb}^{(i)}(X, D)$ for $i = 1, 0$, given by $\mathrm{Alb}^{(i)}{}_{X,E}^{X,D}(\mathrm{id}_X)$.*

Proof. If $E \geq D$, it is evident that $\mathrm{Alb}^{(i)}(X, E)$ generates $\mathrm{Alb}^{(i)}(X, D)$; thus $\mathrm{Alb}^{(i)}{}_{X,E}^{X,D}(\mathrm{id}_X)$ is surjective. $\qquad\square$

3.3. *Jacobian with modulus.* Let C be a smooth proper curve over a perfect field k, which we assume to be algebraically closed for convenience. Let $D = \sum_{q \in S} n_q q$ be an effective divisor on C, where S is a finite set of closed points on C and n_q are positive integers for $q \in S$. The Jacobian $J(C, D)$ of C of modulus D is by definition the universal object for the category of those morphisms φ from $C \setminus S$ to algebraic groups such that $\varphi(\mathrm{div}(f)) = 0$ for all $f \in \mathscr{K}_C$ with $f \equiv 1 \bmod D$. Here we used the definition $\varphi(\sum l_j c_j) = \sum l_j \varphi(c_j)$ for a divisor $\sum l_j c_j$ on C with $c_j \in C \setminus S$, and "$f \equiv 1 \bmod D$" means $\mathrm{v}_q(1 - f) \geq n_q$ for all $q \in S$, where v_q is the valuation attached to the point $q \in C$.

Theorem 3.24. *The generalized Jacobian $J(C, D)$ of C of modulus D is an extension*

$$0 \to L(C, D) \to J(C, D) \to J(C) \to 0$$

of the classical Jacobian $J(C) \cong \mathrm{Pic}_C^0$ of C, which is an abelian variety, by the affine algebraic group $L(C, D)$, which is characterized by

$$L(C, D)(k) = \frac{\prod_{q \in S} k(q)^*}{k^*} \times \prod_{q \in S} \frac{1 + \mathfrak{m}_q}{1 + \mathfrak{m}_q^{n_q}}$$

where $k(q)$ denotes the residue field and $\mathfrak{m}_q$ the maximal ideal at $q \in C$.

Proof. [Serre 1959, V, §3]; see also the summary [ibid., I, no. 1]. $\qquad\square$

Theorem 3.25. *The Jacobian with modulus $J(C, D)$ is dual (in the sense of 1-motives) to the 1-motive $\left[\mathscr{F}_{C,D}^0 \to \mathrm{Pic}_C^0 \right]$, where $\mathscr{F}_{C,D}^0 = \mathscr{F}_{C,D}^{0,\mathrm{red}}$ is the formal subgroup of $\underline{\mathrm{Div}}_C^0$ from Definition 3.14, and $\mathscr{F}_{C,D}^0 \to \mathrm{Pic}_C^0$ is the homomorphism induced by the class map $\underline{\mathrm{Div}}_C^0 \to \underline{\mathrm{Pic}}_C^0$.*

Proof. We have to ensure that the category for which $J(C, D)$ is universal is characterized by the formal group $\mathscr{F}_{C,D}$. The Jacobian $J(C, D)$ of modulus D is by definition the universal object for morphisms φ from $C \setminus S$ to algebraic groups satisfying

(i) $\varphi(\mathrm{div}(f)) = 0$ for all $f \in \mathscr{K}_C$ with $f \equiv 1 \bmod D$.

Condition (i) is equivalent to

(ii) $(\varphi, f)_q = 0$ for all $q \in S$, for all $f \in \mathscr{K}_C$ with $f \equiv 1 \bmod D$ at q,

where $(\varphi, \cdot) . : \mathscr{K}_C^* \times C \to G(k)$ is the local symbol associated to the morphism $\varphi : C \setminus S \to G$, according to [Serre 1959, I, no. 1, thèoréme 1 and III, §1]. It is shown in [Kato and Russell 2010, Sections 6.1–6.3] that condition (ii) is equivalent to

(iii) $\mathrm{mod}(\varphi) \leq D$.

Then the assertion is due to Theorems 3.17 and 3.18. $\qquad\square$

3.4. *Relative Chow group with modulus.* Let X be a smooth proper variety over an algebraically closed field k, and let D be an effective divisor on X and D_{red} the reduced part of D.

Notation 3.26. If C is a curve in X, then $\nu : \widetilde{C} \to C$ denotes the normalization. For $f \in \mathscr{K}_C$, we write $\tilde{f} := \nu^* f$ for the image of f in $\mathscr{K}_{\widetilde{C}}$. If $\varphi : X \dashrightarrow G$ is a rational map, we write $\varphi|_{\widetilde{C}} := \varphi|_C \circ \nu$ for the composition of φ and ν. If B is a Cartier divisor on X intersecting C properly, then $B \cdot \widetilde{C}$ denotes the pull-back of B to $\widetilde{C}$.

Definition 3.27. Let $Z_0(X \setminus D)$ be the group of 0-cycles on $X \setminus D$, set

$$\mathfrak{R}_0(X, D) = \left\{ (C, f) \,\middle|\, \begin{array}{l} C \text{ a curve in } X \text{ intersecting } \mathrm{Supp}(D) \text{ properly, } f \in \mathscr{K}_C^* \\ \text{such that } \tilde{f} \equiv 1 \bmod (D - D_{\mathrm{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\mathrm{red}} \end{array} \right\}$$

and let $R_0(X, D)$ be the subgroup of $Z_0(X \setminus D)$ generated by the elements $\text{div}(f)_C$ with $(C, f) \in \mathfrak{R}_0(X, D)$. Then define

$$\text{CH}_0(X, D) = Z_0(X \setminus D) \big/ R_0(X, D).$$

Let $\text{CH}_0(X, D)^0$ be the subgroup of $\text{CH}_0(X, D)$ of cycles ζ with $\deg \zeta|_W = 0$ for all irreducible components W of $X \setminus D$.

Definition 3.28. Let $\text{Mr}^{\text{CH}}(X, D)$ be the category of rational maps from X to algebraic groups defined as follows: The objects of $\text{Mr}^{\text{CH}}(X, D)$ are morphisms $\varphi : X \setminus D \to G$ whose associated map on 0-cycles of degree zero,

$$Z_0(X \setminus D)^0 \to G(k), \quad \sum l_i\, p_i \mapsto \sum l_i \varphi(p_i), \quad \text{where } l_i \in \mathbb{Z},$$

factors through a homomorphism of groups $\text{CH}_0(X, D)^0 \to G(k)$. The morphisms are the ones as in Definition 2.8. We refer to the objects of $\text{Mr}^{\text{CH}}(X, D)$ as rational maps from X to algebraic groups *factoring through* $\text{CH}_0(X, D)^0$.

Theorem 3.29. *The category* $\text{Mr}(X, D)$ *of rational maps of modulus* $\leq D$ *is equivalent to the category* $\text{Mr}^{\text{CH}}(X, D)$ *of rational maps factoring through* $\text{CH}_0(X, D)^0$. *In particular, the Albanese* $\text{Alb}(X, D)$ *of* X *of modulus* D *is the universal quotient of* $\text{CH}_0(X, D)^0$.

Proof. According to the definitions of $\text{Mr}(X, D)$ and $\text{Mr}^{\text{CH}}(X, D)$ the task is to show that for a morphism $\varphi : X \setminus D \to G$ from $X \setminus D$ to a smooth connected algebraic group G the following conditions are equivalent:

(i) $\text{mod}(\varphi) \leq D$,

(ii) $\varphi(\text{div}(f)_C) = 0$ for all $(C, f) \in \mathfrak{R}_0(X, D)$.

Since $\varphi(\text{div}(f)_C) = \varphi|_{\widetilde{C}}(\text{div}(\widetilde{f})_{\widetilde{C}})$ (see [Russell 2008, Lemma 3.32]), condition (ii) is equivalent to the condition

(iii) $\text{mod}(\varphi|_{\widetilde{C}}) \leq (D - D_{\text{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\text{red}}$ for all curves C in X intersecting $\text{Supp}(D)$ properly,

as was seen in the proof of Theorem 3.25, substituting D by $(D - D_{\text{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\text{red}}$. The equivalence of (i) and (iii) is the content of Lemma 3.30. $\square$

Lemma 3.30. *Let* $\varphi : X \dashrightarrow G$ *be a rational map from* X *to a smooth connected algebraic group* G. *Then the following conditions are equivalent*:

(i) $\text{mod}(\varphi) \leq D$,

(ii) $\text{mod}(\varphi|_{\widetilde{C}}) \leq (D - D_{\text{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\text{red}}$ *for all curves* C *in* X *intersecting* $\text{Supp}(D)$ *properly.*

Proof. (i) $\Longrightarrow$ (ii) Let C be a curve in X intersecting D properly. As φ is regular away from D, the restriction $\varphi|_{\widetilde{C}}$ of φ to $\widetilde{C}$ is regular away from $D \cdot \widetilde{C}$. Hence $\mathrm{Supp}(\mathrm{mod}(\varphi|_{\widetilde{C}})) \subset \mathrm{Supp}(D \cdot \widetilde{C}) = \mathrm{Supp}((D - D_{\mathrm{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\mathrm{red}})$. According to Definition 3.11 of the modulus, it is easy to see that $\mathrm{mod}(\varphi) \le D = (D - D_{\mathrm{red}}) + D_{\mathrm{red}}$ implies $\mathrm{mod}(\varphi|_{\widetilde{C}}) \le (D - D_{\mathrm{red}}) \cdot \widetilde{C} + (D \cdot \widetilde{C})_{\mathrm{red}}$.

(ii) $\Longrightarrow$ (i) Let $E := \mathrm{mod}(\varphi)$ and $q \in \mathrm{Supp}(E)$ be a point of codimension 1 in X. We are going to construct a family of smooth curves $\{C_e\}_e$ intersecting E at a fixed point $x \in E_q = \overline{\{q\}}$ such that

$$\lim_{e \to \infty} \frac{\mathrm{mod}_x(\varphi|_{C_e})}{\mu_x((E - E_{\mathrm{red}}) \cdot C_e) + 1} = 1,$$

where $\mu_x(E \cdot C)$ denotes the intersection multiplicity of E and C at x.

After the construction we will show that the existence of such a family of curves for each $q \in \mathrm{Supp}(E)$ of codimension 1 in X yields the implication (ii) $\Longrightarrow$ (i).

If $\mathrm{char}(k) = 0$, it is easy to see that a general curve C in X intersecting E_q at a point x satisfies $\mathrm{mod}_x(\varphi|_C) = \mu_x((E - E_{\mathrm{red}}) \cdot C) + 1$. Therefore we suppose that $\mathrm{char}(k) = p > 0$. Using the notation of Definition 3.11, let $(u_{q,i})_{1 \le i \le a} \in W_r(\mathcal{K}_{X,q})^a$ be a representative of the unipotent part of the class of $\varphi \in G(\mathcal{K}_{X,q})$ in

$$G(\mathcal{K}_{X,q})/G(\mathcal{O}_{X,q}) = L(\mathcal{K}_{X,q})/L(\mathcal{O}_{X,q}).$$

Then $\mathrm{mod}_q(\varphi) = 1 + \mathrm{n}_q(u_{q,i})$ for some $1 \le i \le a$. Set $n := \mathrm{n}_q(u_{q,i})$. Let $t \in \mathfrak{m}_{X,q}$ be a uniformizer at q. Let

$$\sum_{\nu} \mathrm{F}^{\nu} \otimes \omega_{\nu} \otimes t^{-n} \in k[\mathrm{F}] \otimes_k \Omega_{X,q}(\log q) \otimes_{\mathcal{O}_{X,q}} \mathfrak{m}_{X,q}^{-n}$$

be a representative of $\overline{\partial}_{nq}(u_{q,i}) \in \overline{\mathfrak{D}}_{nq}$ (Proposition 3.7). Choose a regular closed point $x \in E_q$ such that t is a local equation for E_q at x and ω_{ν} is regular and nonzero at x for some ν. We may assume that $\dim X = 2$ via cutting down by hyperplanes through x transversal to E_q. Let $s \in \mathfrak{m}_{X,x}$ be a local parameter at x that gives a uniformizer of $\mathcal{O}_{E_q,x}$. Define a curve C_e locally around x by the equation $t = s^e$ for $e \ge 1$. Note that $E - E_{\mathrm{red}}$ is locally defined by the equation $t^n = 0$. Then

$$\mu_x((E - E_{\mathrm{red}}) \cdot C_e) = \dim_k \frac{\mathcal{O}_{X,x}}{(t^n, t - s^e)} = ne.$$

We can write $\omega_{\nu} = g \, \mathrm{d}s + h \, \mathrm{d}\log t$ with $g, h \in \mathcal{O}_{X,q}$ and the values at x are $g(x) \ne 0$ if $\overline{\partial}_{nq}(u_{q,i}) \in {}^{\flat}\overline{\mathfrak{D}}_{nq}$, and $h(x) \ne 0$ if $\overline{\partial}_{nq}(u_{q,i}) \in \overline{\mathfrak{D}}_{nq} \setminus {}^{\flat}\overline{\mathfrak{D}}_{nq}$ and x in general position (what we assume), for some ν. The restriction of $t^{-n}\omega_{\nu}$ to C_e is

$$t^{-n}\omega_{\nu}|_{C_e} = s^{-ne} g \, \mathrm{d}s + s^{-ne} h \, \mathrm{d}\log s^e = s^{1-ne} g \, \mathrm{d}\log s + e \, s^{-ne} h \, \mathrm{d}\log s,$$

and the class of $t^{-n}\omega_\nu|_{C_e}$ is nonzero in

$$\begin{cases} \Omega_{C_e,x}(\log x) \otimes_{\mathbb{O}_{C_e,x}} \mathfrak{m}_{C_e,x}^{-ne}/\mathfrak{m}_{C_e,x}^{1-ne} & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in \overline{\mathfrak{D}}_{nq} \setminus {}^\flat\overline{\mathfrak{D}}_{nq} \text{ and } p \nmid e, \\ \Omega_{C_e,x}(\log x) \otimes_{\mathbb{O}_{C_e,x}} \mathfrak{m}_{C_e,x}^{1-ne}/\mathfrak{m}_{C_e,x}^{2-ne} & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in {}^\flat\overline{\mathfrak{D}}_{nq}. \end{cases}$$

[Lemma 3.10](#) assures that the modulus of $\varphi|_{C_e}$ is computed from the restriction (of a representative) of $\overline{\mathfrak{d}}_{nq}(u_{q,i})$ to C_e, for e large enough such that $ne - 1 > \lfloor ne/p \rfloor$ (this is satisfied for $e > 2$). Thus we have

$$\mathrm{n}_x(u_{q,i}|_{C_e}) = \begin{cases} ne & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in \overline{\mathfrak{D}}_{nq} \setminus {}^\flat\overline{\mathfrak{D}}_{nq} \text{ and } p \nmid e, \\ ne - 1 & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in {}^\flat\overline{\mathfrak{D}}_{nq}, \end{cases}$$

$$\mathrm{mod}_x(\varphi|_{C_e}) = \begin{cases} ne + 1 & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in \overline{\mathfrak{D}}_{nq} \setminus {}^\flat\overline{\mathfrak{D}}_{nq} \text{ and } p \nmid e, \\ ne & \text{if } \overline{\mathfrak{d}}_{nq}(u_{q,i}) \in {}^\flat\overline{\mathfrak{D}}_{nq}. \end{cases}$$

Then

$$\lim_{\substack{e \to \infty \\ p \nmid e}} \frac{\mathrm{mod}_x(\varphi|_{C_e})}{\mu_x((E - E_{\mathrm{red}}) \cdot C_e) + 1} = 1.$$

Now we show that "not (i) implies not (ii)". Suppose $E := \mathrm{mod}(\varphi) \not\leq D$. Then there is a point $q \in \mathrm{Supp}(E)$ of codimension 1 in X such that $\mu_q(E) > \mu_q(D)$, where μ_q is the multiplicity at q. By the construction above there is a sequence of curves $\{C_e\}_e$ in X intersecting E at a fixed point $x \in E_q$ such that

$$\lim_{\substack{e \to \infty \\ p \nmid e}} \frac{\mathrm{mod}_x(\varphi|_{C_e})}{\mu_x((E - E_{\mathrm{red}}) \cdot C_e) + 1} = 1.$$

If $\mu_q(D) \neq 0$, then since

$$\sup_{e \geq 0} \frac{\mu_x((D - D_{\mathrm{red}}) \cdot C_e) + 1}{\mu_x((E - E_{\mathrm{red}}) \cdot C_e) + 1} < 1,$$

there is e such that $\mathrm{mod}_x(\varphi|_{C_e}) > \mu_x((D - D_{\mathrm{red}}) \cdot C_e) + 1$. If $\mu_q(D) = 0$, then

$$0 \neq \mathrm{mod}(\varphi|_{C_e})_x > \mu_x((D - D_{\mathrm{red}}) \cdot C_e + (D \cdot C_e)_{\mathrm{red}}) = 0.$$

Thus $\mathrm{mod}(\varphi|_{C_e}) \not\leq (D - D_{\mathrm{red}}) \cdot C_e + (D \cdot C_e)_{\mathrm{red}}$. $\square$

Acknowledgement

References

[Barbieri-Viale and Bertapelle 2009] L. Barbieri-Viale and A. Bertapelle, "Sharp de Rham realization", *Adv. Math.* **222**:4 (2009), 1308–1338. MR 2010i:14002 Zbl 1216.14006

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **21**, Springer, Berlin, 1990. MR 91i:14034 Zbl 0705.14001

[Brylinski 1983] J.-L. Brylinski, "Théorie du corps de classes de Kato et revêtements abéliens de surfaces", *Ann. Inst. Fourier* (*Grenoble*) **33**:3 (1983), 23–38. MR 85f:11088 Zbl 0524.12008

[Deligne 1971] P. Deligne, "Théorie de Hodge, II", *Inst. Hautes Études Sci. Publ. Math.* 40 (1971), 5–57. MR 58 #16653a Zbl 0219.14007

[Demazure 1972] M. Demazure, *Lectures on p-divisible groups*, Lecture Notes in Mathematics **302**, Springer, Berlin, 1972. MR 88a:14049 Zbl 0247.14010

[Demazure and Gabriel 1970] M. Demazure and P. Gabriel, *Groupes algébriques, I: Géométrie algébrique, généralités, groupes commutatifs*, Masson & Cie, Paris, 1970. MR 46 #1800

[Fontaine 1977] J.-M. Fontaine, *Groupes p-divisibles sur les corps locaux* (Orsay, 1992), Astérisque **47-48**, Société Mathématique de France, Paris, 1977. MR 58 #16699 Zbl 0377.14009

[Kato and Russell 2010] K. Kato and H. Russell, "Modulus of a rational map into a commutative algebraic group", *Kyoto J. Math.* **50**:3 (2010), 607–622. MR 2011j:14098 Zbl 1206.14069

[Kato and Russell 2012] K. Kato and H. Russell, "Albanese varieties with modulus and Hodge theory", *Ann. Inst. Fourier* **62**:2 (2012), 783–806.

[Kato and Saito 1983] K. Kato and S. Saito, "Two-dimensional class field theory", pp. 103–152 in *Galois groups and their representations* (Nagoya, 1981), edited by Y. Ihara, Adv. Stud. Pure Math. **2**, North-Holland, Amsterdam, 1983. MR 87a:11060 Zbl 0544.12011

[Lang 1959] S. Lang, *Abelian varieties*, Interscience Tracts in Pure and Applied Mathematics **7**, Interscience Publishers, New York, 1959. MR 21 #4959 Zbl 0099.16103

[Laumon 1996] G. Laumon, "Transformation de Fourier généralisée", preprint, 1996. arXiv alg-geom/9603004v1

[Matsusaka 1952] T. Matsusaka, "On the algebraic construction of the Picard variety, II", *Jap. J. Math.* **22** (1952), 51–62. MR 15,983b Zbl 0049.22801

[Milne 1980] J. S. Milne, *Étale cohomology*, Princeton Mathematical Series **33**, Princeton University Press, 1980. MR 81j:14002 Zbl 0433.14012

[Önsiper 1989] H. Önsiper, "Generalized Albanese varieties for surfaces in characteristic $p > 0$", *Duke Math. J.* **59**:2 (1989), 359–364. MR 90h:14025 Zbl 0753.14039

[Oort 1966] F. Oort, *Commutative group schemes*, Lecture Notes in Mathematics **15**, Springer, Berlin, 1966. MR 35 #4229 Zbl 0216.05603

[Russell 2008] H. Russell, "Generalized Albanese and its dual", *J. Math. Kyoto Univ.* **48**:4 (2008), 907–949. MR 2010b:14008 Zbl 1170.14005

[Serre 1958–1959] J.-P. Serre, "Morphismes universels et variétés d'Albanese", pp. 1–22, Exp. No. 10 in *Variétés de Picard*, Séminaire C. Chevalley **4**, Secrétariat mathématique, Paris, 1958–1959. Zbl 0123.13903

[Serre 1959] J.-P. Serre, *Groupes algébriques et corps de classes*, Publications de l'institut de mathématique de l'université **7**, Hermann, Paris, 1959. MR 21 #1973 Zbl 0097.35604

[SGA3 1970] M. Demazure and A. Grothendieck (editors), *Schémas en groupes, I: Propriétés générales des schémas en groupes* (SGA3), Lecture Notes in Mathematics **151**, Springer, Berlin, 1970. MR 43 #223a Zbl 0207.51401

[Waterhouse 1979] W. C. Waterhouse, *Introduction to affine group schemes*, Graduate Texts in
  Mathematics **66**, Springer, New York, 1979. MR 82e:14003 Zbl 0442.14017

henrik.russell@math.fu-berlin.de

                         *Freie Universität Berlin, Mathematik und Informatik,
                         Arnimallee 3, 14195 Berlin, Germany*

msp

Chai's conjecture and Fubini properties of dimensional motivic integration

Raf Cluckers, François Loeser and Johannes Nicaise

We prove that a conjecture of Chai on the additivity of the base change conductor for semiabelian varieties over a discretely valued field is equivalent to a Fubini property for the dimensions of certain motivic integrals. We prove this Fubini property when the valued field has characteristic zero.

1. Introduction

Let R be a henselian discrete valuation ring with quotient field K and perfect residue field k. Let G be a semiabelian variety over K, i.e., an extension of an abelian K-variety by a K-torus. Then G can be canonically extended to a smooth separated commutative group scheme $\mathcal{G}$ over R, the so-called Néron lft-model of G [Bosch et al. 1990, 10.1.1]. We say that G has semiabelian reduction if the identity component of the special fiber of $\mathcal{G}$ is a semiabelian k-variety.

Chai [2000] introduced the *base change conductor* $c(G)$ of G, a positive rational number that measures the defect of semiabelian reduction of G. Its precise definition is recalled in Definition 2.3.1. The base change conductor vanishes if and only if G has semiabelian reduction. For algebraic tori, this invariant had previously been defined and studied by Chai and Yu [2001]. They proved the deep result that the base change conductor of a K-torus T is invariant under isogeny. Applying an argument from [Gross and Gan 1999], they deduced that $c(T)$ equals one half of the Artin conductor of the cocharacter module of T. For semiabelian varieties, however, no similar cohomological interpretation is known to hold in general; in fact, the base change conductor is not even invariant under isogeny [Chai 2000, §6.10], and many of its properties remain mysterious. One of the main open questions is the following conjecture:

Conjecture 1.1 [Chai 2000, §8.1]. *Let G be a semiabelian K-variety that fits into an exact sequence of algebraic K-groups*

$$0 \to T \to G \to A \to 0$$

with T a K-torus and A an abelian K-variety. Then we have

$$c(G) = c(A) + c(T).$$

The fundamental difficulty underlying this conjecture is that an exact sequence of semiabelian varieties does not give rise to an exact sequence of Néron lft-models in general. Chai proved the conjecture if k is finite, using Fubini's theorem for integrals with respect to the Haar measure on the completion of K. He also proved the conjecture when K has mixed characteristic, using a different method and applying the property that $c(T)$ only depends on the isogeny class of T. If k has characteristic zero (more generally, if G obtains semiabelian reduction after a *tame* finite extension of K), Chai's conjecture can be proven in an elementary way; see [Halle and Nicaise 2011, 4.23].

In the first part of the present paper, we show that, in arbitrary characteristic, Chai's conjecture is equivalent to a Fubini property for the dimensions of certain *motivic* integrals (Equation (4.2-1) in Theorem 4.2.1). We then prove in the second part of the paper that this Fubini property holds when K has characteristic zero (Theorem 4.2.3). This yields a new proof of the conjecture in that case, which is close in spirit to Chai's proof of the finite residue field case.

The strength of our approach lies in the fact that we combine insights of two theories of motivic integration, namely, the geometric theory of motivic integration on rigid varieties of Loeser and Sebag [2003] and the model-theoretic approach of Cluckers and Loeser [2008; 2012]. Let us emphasize that the Fubini property in (4.2-1) is not an immediate corollary of the Fubini results from [Cluckers and Loeser 2012]; see Remark 4.2.5. We need to combine the theory in [Cluckers and Loeser 2012] with a new result (Theorem 5.2.1 and its corollary), which roughly states that the virtual dimension of a motivic integral over a fixed space only depends on the dimensions of the values of the integrand. This theorem may be of independent interest.

We hope that our reformulation of Chai's conjecture in terms of motivic integrals will also shed new light on the open case of the conjecture, when k is infinite, K has positive characteristic and G is wildly ramified.

2. Preliminaries

2.1. *Notation.* Throughout this article, R denotes a henselian discrete valuation ring with quotient field K and perfect residue field k. We denote by $\mathfrak{m}$ the maximal

ideal of R, R^{sh} a strict henselization of R and K^{sh} its field of fractions. The residue field k^s of R^{sh} is an algebraic closure of k. We denote by $\widehat{R}$ the $\mathfrak{m}$-adic completion of R and $\widehat{K}$ its field of fractions.

For every ring A, we denote by (Sch/A) the category of A-schemes. We consider the *special fiber functor*

$$(\cdot)_k : (\mathrm{Sch}/R) \to (\mathrm{Sch}/k) : \mathscr{X} \mapsto \mathscr{X}_k = X \times_R k$$

and the *generic fiber functor*

$$(\cdot)_K : (\mathrm{Sch}/R) \to (\mathrm{Sch}/K) : \mathscr{X} \mapsto \mathscr{X}_K = \mathscr{X} \times_R K.$$

A variety over a ring A is a reduced separated A-scheme of finite type.

2.2. *Néron models and semiabelian reduction.* A semiabelian variety over a field F is an extension of an abelian F-variety by an algebraic F-torus. Let G be a semiabelian variety over K. It follows from [Bosch et al. 1990, 10.2.2] that G admits a Néron lft-model $\mathscr{G}$ in the sense of [Bosch et al. 1990, 10.1.1]. It is the minimal extension of G to a smooth separated group scheme over R. We say that G has semiabelian reduction if the identity component $\mathscr{G}_k^o$ of the special fiber of $\mathscr{G}$ is a semiabelian k-variety. There always exists a finite separable extension L of K such that $G \times_K L$ has semiabelian reduction. If G is an abelian variety, then this is Grothendieck's semistable reduction theorem [Grothendieck et al. 1972, IX.3.6]. If G is a torus, then one can take for L the splitting field of G. The general case is easily deduced from these special cases; see [Halle and Nicaise 2010, 3.11].

Let K' be a finite separable extension of K, and denote by R' the integral closure of R in K'. We set $G' = G \times_K K'$, and we denote by $\mathscr{G}'$ the Néron lft-model of G'. By the universal property of the Néron lft-model, there exists a unique morphism of R'-schemes

$$h : \mathscr{G} \times_R R' \to \mathscr{G}' \tag{2.2-1}$$

that extends the natural isomorphism between the generic fibers. If G has semiabelian reduction, then h is an open immersion [Grothendieck et al. 1972, 3.1(e)], which induces an isomorphism

$$(\mathscr{G} \times_R R')^o \to (\mathscr{G}')^o$$

between the identity components of $\mathscr{G} \times_R R'$ and $\mathscr{G}'$ [Demazure and Grothendieck 1970a, VI$_B$.3.11].

2.3. *The base change conductor.* Let G be a semiabelian variety over K. Let K' be a finite separable extension of K such that $G' = G \times_K K'$ has semiabelian reduction, and denote by $e(K'/K)$ the ramification index of K' over K. The

morphism (2.2-1) induces an injective morphism

$$\mathrm{Lie}(h) : \mathrm{Lie}(\mathcal{G}) \otimes_R R' \to \mathrm{Lie}(\mathcal{G}') \qquad (2.3\text{-}1)$$

of free R'-modules of rank $\dim(G)$.

Definition 2.3.1 [Chai 2000, §1]. The base change conductor of G is defined by

$$c(G) = \frac{1}{e(K'/K)} \cdot \mathrm{length}_{R'}(\mathrm{coker}(\mathrm{Lie}(h))).$$

This definition does not depend on the choice of K'. The base change conductor is a positive rational number that vanishes if and only if G has semiabelian reduction [Halle and Nicaise 2011, 4.16]. One can view $c(G)$ as a measure for the defect of semiabelian reduction of G.

2.4. *A generalization of Chai's conjecture.* Chai [2000, 8.1] asks whether one can generalize Conjecture 1.1 as follows.

Question 2.4.1. Do we have $c(G_2) = c(G_1) + c(G_3)$ for every exact sequence of semiabelian K-varieties

$$0 \to G_1 \to G_2 \to G_3 \to 0?$$

If G_1, G_2 and G_3 are tori, this can be easily deduced from the deep fact that the base change conductor of a torus is one half of the Artin conductor of the cocharacter module [Chai and Yu 2001] in the following way:

Proposition 2.4.2. *Let*

$$0 \to G_1 \to G_2 \to G_3 \to 0$$

be an exact sequence of K-tori. Then $c(G_2) = c(G_1) + c(G_3)$.

Proof. The sequence of cocharacter modules

$$0 \to X_\bullet(G_1) \to X_\bullet(G_2) \to X_\bullet(G_3) \to 0$$

is exact. Tensoring with $\mathbb{Q}$, we get a split exact sequence of $\mathbb{Q}[\mathrm{Gal}(L/K)]$-modules

$$0 \to X_\bullet(G_1) \otimes_\mathbb{Z} \mathbb{Q} \to X_\bullet(G_2) \otimes_\mathbb{Z} \mathbb{Q} \to X_\bullet(G_3) \otimes_\mathbb{Z} \mathbb{Q} \to 0$$

where L is the splitting field of G_2. Thus, the Artin conductor of $X_\bullet(G_2) \otimes_\mathbb{Z} \mathbb{Q}$ is the sum of the Artin conductors of $X_\bullet(G_1) \otimes_\mathbb{Z} \mathbb{Q}$ and $X_\bullet(G_3) \otimes_\mathbb{Z} \mathbb{Q}$. Since the base change conductor of a torus is one half of the Artin conductor of the cocharacter module [Chai and Yu 2001], we find that $c(G_2) = c(G_1) + c(G_3)$. $\qquad \square$

Corollary 2.4.3. *If Conjecture 1.1 holds, then Question 2.4.1 has a positive answer when G_1 is a torus.*

Proof. Assume that G_1 is a torus. For every semiabelian K-variety G, we denote by G_{tor} its maximal subtorus and $G_{\text{ab}} = G/G_{\text{tor}}$ its abelian part. We consider the closed subgroup $\widetilde{G}_2 = (G_3)_{\text{tor}} \times_{G_3} G_2$ of G_2. We have a short exact sequence of K-groups

$$0 \to G_1 \to \widetilde{G}_2 \to (G_3)_{\text{tor}} \to 0 \tag{2.4-1}$$

so that $\widetilde{G}_2$ is an extension of K-tori and thus a torus. Moreover, the morphism

$$G_2/\widetilde{G}_2 \to G_3/(G_3)_{\text{tor}} = (G_3)_{\text{ab}}$$

is an isomorphism so that $\widetilde{G}_2 = (G_2)_{\text{tor}}$ and $(G_2)_{\text{ab}} \cong (G_3)_{\text{ab}}$. By Conjecture 1.1, we have $c(G_i) = c((G_i)_{\text{tor}}) + c((G_i)_{\text{ab}})$ for $i = 2, 3$. Applying Proposition 2.4.2 to the sequence (2.4-1), we find that $c((G_2)_{\text{tor}}) = c(G_1) + c((G_3)_{\text{tor}})$. It follows that $c(G_2) = c(G_1) + c(G_3)$. $\qquad\square$

Below, we will follow a different approach. We will use the invariance of the base change conductor of a torus under isogeny to reduce Question 2.4.1 to the case where the maximal split subtorus $(G_3)_{\text{sp}}$ of G_3 is trivial (of course, this is always the case if G_3 is an abelian variety as in Conjecture 1.1). Then we prove that, if G_1 is a torus and $(G_3)_{\text{sp}}$ is trivial, the additivity property of the base change conductor in Question 2.4.1 is equivalent to a certain Fubini property for motivic integrals. We prove this Fubini property when K has characteristic zero. These arguments do not use the invariance of the base change conductor of a torus under isogeny.

3. Motivic Haar measures on semiabelian varieties

3.1. *The Grothendieck ring of varieties.* Let F be a field. We denote by $K_0(\text{Var}_F)$ the *Grothendieck ring of varieties over F*. As an abelian group, $K_0(\text{Var}_F)$ is defined by the following presentation:

generators: isomorphism classes $[X]$ of separated F-schemes of finite type X,

relations: if X is a separated F-scheme of finite type and Y is a closed subscheme of X, then

$$[X] = [Y] + [X \setminus Y].$$

These relations are called *scissor relations*.

By the scissor relations, one has $[X] = [X_{\text{red}}]$ for every separated F-scheme of finite type X, where X_{red} denotes the maximal reduced closed subscheme of X. We endow the group $K_0(\text{Var}_F)$ with the unique ring structure such that

$$[X] \cdot [X'] = [X \times_F X']$$

for all separated F-schemes of finite type X and X'. The identity element for the multiplication is the class $[\text{Spec } F]$ of the point. To any constructible subset C of a

separated F-scheme of finite type X, one can associate an element $[C]$ in $K_0(\mathrm{Var}_F)$ by choosing a finite partition of C into subvarieties $C_1, \ldots, C_r$ of X and setting $[C] = [C_1] + \cdots + [C_r]$. The scissor relations imply that this definition does not depend on the choice of the partition. For a detailed survey on the Grothendieck ring of varieties, we refer to [Nicaise and Sebag 2011a].

We denote by $K_0^{\mathrm{mod}}(\mathrm{Var}_F)$ the *modified Grothendieck ring of varieties over F* [Nicaise and Sebag 2011a, §3.8]. This is the quotient of $K_0(\mathrm{Var}_F)$ by the ideal $\mathcal{I}_F$ generated by elements of the form $[X] - [Y]$ where X and Y are separated F-schemes of finite type such that there exists a finite, surjective, purely inseparable F-morphism $Y \to X$. If F has characteristic zero, then it is easily seen that $\mathcal{I}_F$ is the zero-ideal [Nicaise and Sebag 2011a, 3.11] so that $K_0(\mathrm{Var}_F) = K_0^{\mathrm{mod}}(\mathrm{Var}_F)$. It is not known if $\mathcal{I}_F$ is nonzero if F has positive characteristic. In particular, if F' is a nontrivial finite purely inseparable extension of F, it is not known whether $[\mathrm{Spec}\, F'] \neq 1$ in $K_0(\mathrm{Var}_F)$.

There exists a canonical isomorphism from $K_0^{\mathrm{mod}}(\mathrm{Var}_F)$ to the Grothendieck ring $K_0(\mathrm{ACF}_F)$ of the theory ACF_F of algebraically closed fields over F [Nicaise and Sebag 2011a, 3.13]. One may also consider the semiring variant $K_0^+(\mathrm{ACF}_F)$ of the ring $K_0(\mathrm{ACF}_F)$, defined as follows. Let $\mathcal{L}_{\mathrm{ring}}(F)$ be the ring language with coefficients from F. As a semigroup, $K_0^+(\mathrm{ACF}_F)$ is the quotient of the free commutative semigroup generated by a symbol $[X]$ for each $\mathcal{L}_{\mathrm{ring}}(F)$-definable set, with zero-element $[\varnothing]$, and divided out by the following relations:

- if X and Y are $\mathcal{L}_{\mathrm{ring}}(F)$-definable subsets of a common $\mathcal{L}_{\mathrm{ring}}(F)$-definable set, then
$$[X \cup Y] + [X \cap Y] = [X] + [Y],$$

- if there exists an $\mathcal{L}_{\mathrm{ring}}(F)$-definable bijection $X \to Y$ for the theory ACF_F, then $[X] = [Y]$.

The semigroup $K_0^+(\mathrm{ACF}_F)$ carries a structure of semiring, induced by taking Cartesian products, $[X][Y] = [X \times Y]$.

If R has equal characteristic, then we put
$$K_0^R(\mathrm{Var}_k) = K_0(\mathrm{Var}_k).$$

If R has mixed characteristic, then we put
$$K_0^R(\mathrm{Var}_k) = K_0^{\mathrm{mod}}(\mathrm{Var}_k).$$

We denote by $\mathbb{L}$ the class of the affine line $\mathbb{A}_k^1$ in $K_0^R(\mathrm{Var}_k)$ and also in $K_0^+(\mathrm{ACF}_k)$. We will write $\mathcal{M}_k^R$ for the localization of $K_0^R(\mathrm{Var}_k)$ with respect to $\mathbb{L}$ and $\mathcal{M}_k^+$ for the localization of $K_0^+(\mathrm{ACF}_k)$ with respect to $\mathbb{L}$ and the elements $\mathbb{L}^i - 1$ for all $i > 0$.

For every element α of $K_0^R(\mathrm{Var}_k)$, we denote by $P(\alpha)$ its *Poincaré polynomial* [Nicaise and Sebag 2011a, 4.13]. This is an element of $\mathbb{Z}[T]$, and the map

$$P : K_0^R(\mathrm{Var}_k) \to \mathbb{Z}[T] : \alpha \mapsto P(\alpha)$$

is a ring morphism. Hence, the map

$$P^+ : K_0^+(\mathrm{ACF}_k) \to \mathbb{Z}[T],$$

obtained by composing P with the canonical morphism

$$K_0^+(\mathrm{ACF}_F) \to K_0(\mathrm{ACF}_F) \cong K_0^R(\mathrm{Var}_k),$$

is a morphism of semirings. When α is the class of a separated k-scheme of finite type X, then for every $i \in \mathbb{N}$, the coefficient of T^i in $P(\alpha)$ is $(-1)^i$ times the i-th *virtual Betti number* of X. The degree of $P(\alpha)$ is twice the dimension of X [Nicaise 2011, 8.7].

We have $P(\mathbb{L}) = T^2$ so that P localizes through a ring morphism

$$P : \mathcal{M}_k^R \to \mathbb{Z}[T, T^{-1}]$$

and P^+ localizes through a semiring morphism

$$P^+ : \mathcal{M}_k^+ \to \mathbb{Z}[T, T^{-1}, (T^{2i} - 1)^{-1}]_{i > 0}.$$

Definition 3.1.1. Let α be an element of $\mathcal{M}_k^R$ or $\mathcal{M}_k^+$. We define the *virtual dimension* of α as $1/2$ times the degree of the Poincaré polynomial $P(\alpha)$ or $P^+(\alpha)$, respectively, where the degree of the zero-polynomial is $-\infty$ and $(1/2) \cdot (-\infty) = -\infty$ by convention. We denote the virtual dimension of α by $\dim(\alpha)$.

By definition, the virtual dimension is an element of $(1/2) \cdot \mathbb{Z} \cup \{-\infty\}$. For every separated k-scheme of finite type X and every integer i, we have

$$\dim([X]\mathbb{L}^i) = \dim(X) + i.$$

3.2. *Motivic integration on K-varieties.* Let X be a K-variety. We say that X is *bounded* if $X(K^{sh})$ is bounded in X in the sense of [Bosch et al. 1990, 1.1.2]. If X is a smooth K-variety, then by [Bosch et al. 1990, 3.4.2 and 3.5.7], X is bounded if and only if X admits a *weak Néron model* $\mathscr{X}$. This means that $\mathscr{X}$ is a smooth R-variety endowed with an isomorphism $\mathscr{X}_K \to X$ such that the natural map

$$\mathscr{X}(R^{sh}) \to X(K^{sh})$$

is a bijection.

The theory of motivic integration on rigid varieties was developed in [Loeser and Sebag 2003] and further extended in [Nicaise and Sebag 2008; Nicaise 2009]. We refer to [Nicaise and Sebag 2011b] for a survey; see in particular [Nicaise and Sebag

2011b, §2.4] for an erratum to the previous papers. One of the main results can be reformulated for algebraic varieties as follows. Let X be bounded smooth K-variety of pure dimension, and let ω be a gauge form on X, i.e., a nowhere-vanishing differential form of degree $\dim(X)$. Let $\mathscr{X}$ be a weak Néron model for X. For every connected component C of $\mathscr{X}_k = \mathscr{X} \times_R k$, we denote by $\operatorname{ord}_C \omega$ the order of ω along C. If ϖ is a uniformizer in R, then $\operatorname{ord}_C \omega$ is the unique integer n such that $\varpi^{-n}\omega$ extends to a generator of $\Omega_{\mathscr{X}/R}^{\dim(X)}$ at the generic point of C.

Theorem-Definition 3.2.1. *The object*

$$\int_X |\omega| = \mathbb{L}^{-\dim(X)} \sum_{C \in \varpi_0(\mathscr{X}_k)} [C]\mathbb{L}^{-\operatorname{ord}_C \omega} \in \mathcal{M}_k^R \qquad (3.2\text{-}1)$$

only depends on X and ω and not on the choice of a weak Néron model $\mathscr{X}$. We call it the motivic integral of ω on X.

Proof. By [Nicaise 2011, 4.9], the formal $\mathfrak{m}$-adic completion of $\mathscr{X}$ is a formal weak Néron model of the rigid analytification X^{rig} of $X \times_K \widehat{K}$ so that the result follows from [Halle and Nicaise 2011, 2.3]. □

It is clear from the definition that the motivic integral of ω on X remains invariant if we multiply ω with a unit in R.

Remark 3.2.2. In the literature, the factor $\mathbb{L}^{-\dim(X)}$ in the right-hand side of (3.2-1) is sometimes omitted (for instance in [Nicaise and Sebag 2011b]); this depends on the choice of a normalization for the motivic measure.

3.3. *Motivic Haar measures.* Consider a semiabelian K-variety G of dimension g. We denote by $\mathscr{G}$ the Néron lft-model of G and Ω_G the free rank-one R-module of translation-invariant differential forms in $\Omega_{\mathscr{G}/R}^g(\mathscr{G})$. Note that $\Omega_G \otimes_R K$ is canonically isomorphic to the K-vector space of translation-invariant differential forms of maximal degree on G so that we can view Ω_G as an R-lattice in this vector space. We denote by ω_G a generator of Ω_G. It is unique up to multiplication with a unit in R.

Let K' be a finite separable extension of K such that $G' = G \times_K K'$ has semi-abelian reduction, and let d be the ramification index of K' over K. Denote by R' the normalization of R in K'. Dualizing the morphism (2.3-1) and taking determinants, we find a morphism of free rank-one R'-modules

$$\det(\operatorname{Lie}(h))^\vee : \Omega_{G'} \to \Omega_G \otimes_R R'$$

that induces an isomorphism

$$\Omega_{G'} \otimes_{R'} K' \cong \Omega_G \otimes_R K'$$

by tensoring with K'. Thus, we can view Ω_G as a sub-R-module of $\Omega_{G'} \otimes_{R'} K'$. This yields the following alternative description of the base change conductor:

Proposition 3.3.1. *Let ϖ' be a uniformizer in R'. The base change conductor $c(G)$ of G is the unique element r of $\mathbb{Z}[1/d]$ such that*

$$(\varpi')^{rd} \omega_G$$

generates the R'-module $\Omega_{G'}$.

Proof. Denote by $\mathcal{G}'$ the Néron lft-model of G'. By definition, the length of the cokernel of the natural morphism

$$\mathrm{Lie}(h) : \mathrm{Lie}(\mathcal{G} \times_R R') \to \mathrm{Lie}(\mathcal{G}')$$

from (2.3-1) is equal to $c(G)d$. Writing $\mathrm{Lie}(h)$ in Smith normal form, it is easily seen that the cokernel of

$$\det(\mathrm{Lie}(h))^\vee : \Omega_{G'} \to \Omega_G \otimes_R R'$$

is isomorphic to $R'/(\varpi')^{c(G)d}$. $\square$

Proposition 3.3.2. *Let $R \to S$ be a flat local homomorphism of discrete valuation rings of ramification index one (in the sense of [Bosch et al. 1990, 3.6.1]), and denote by L the quotient field of S. We denote by $\mathcal{G}^L$ the Néron lft-model of $G \times_K L$.*

(1) *The natural morphism*

$$\mathcal{G} \times_R S \to \mathcal{G}^L$$

is an isomorphism. In particular, it induces an isomorphism of S-modules

$$\Omega_G \otimes_R S \cong \Omega_{G \times_K L}.$$

(2) *We have $c(G \times_K L) = c(G)$.*

This applies in particular to the case $S = \widehat{R^{sh}}$.

Proof. (1) The formation of Néron lft-models commutes with the base change $R \to S$ by [Bosch et al. 1990, 3.6.1].

(2) This follows easily from (1). $\square$

The semiabelian K-variety G is bounded if and only if its Néron lft-model $\mathcal{G}$ is of finite type over R [Bosch et al. 1990, 10.2.1]. In that case, $\mathcal{G}$ is called the Néron model of G. If G is bounded, then for every gauge form ω on G, we can consider the motivic integral

$$\int_G |\omega| \in \mathcal{M}_k^R.$$

In particular, we can consider the motivic integral of the "motivic Haar measure" $|\omega_G|$ associated to G. It does not depend on the choice of ω_G since ω_G is unique up to multiplication with a unit in R.

Proposition 3.3.3. *Let G be a bounded semiabelian K-variety of dimension g with Néron model $\mathscr{G}$. Let ϖ be a uniformizer in R. Then for every integer γ, we have*

$$\int_G |\varpi^\gamma \omega_G| = \mathbb{L}^{-\gamma-g}[\mathscr{G}_k] \tag{3.3-1}$$

in $\mathcal{M}_k^R$. In particular, the virtual dimension of this motivic integral is equal to $-\gamma$.

Proof. Since ω_G generates Ω_G, we have

$$\mathrm{ord}_C(\varpi^\gamma \omega_G) = \gamma$$

for every connected component C of $\mathscr{G}_k$. Thus, formula (3.2-1) becomes

$$\int_G |\varpi^\gamma \omega_G| = \mathbb{L}^{-\gamma-g} \sum_{C \in \varpi_0(\mathscr{G}_k)} [C] = \mathbb{L}^{-\gamma-g}[\mathscr{G}_k]$$

in $\mathcal{M}_k^R$, where the last equality follows from the scissor relations in the Grothendieck ring. $\qquad\square$

3.4. *Split subtori and bounded varieties.* We mentioned in Section 3.3 that a semi-abelian K-variety G is bounded if and only if the Néron lft-model $\mathscr{G}$ of G is quasicompact. If R is excellent (e.g., complete) and k algebraically closed, then this is also equivalent to the property that G does not contain a split torus [Bosch et al. 1990, 10.2.1]. Since the boundedness condition plays an important role in the definition of the motivic integral, we'll now take a closer look at split subtori of semi-abelian varieties. The results in this section will allow us to establish an equivalence between Question 2.4.1 and a Fubini property of motivic integrals (Theorem 4.2.1).

Let F be any field. We denote by (SpT/F) the category of split F-tori and (SAb/F) the category of semiabelian F-varieties (the morphisms in these categories are morphisms of algebraic F-groups).

For every semiabelian F-variety G, we denote by G_{sp} the maximal split subtorus of G [Halle and Nicaise 2010, 3.6]. If T is a split F-torus, then every morphism of F-groups $T \to G$ factors through G_{sp} by [Halle and Nicaise 2010, 3.5]. Thus, we can define a functor

$$(\,\cdot\,)_{\mathrm{sp}} : (\mathrm{SAb}/F) \to (\mathrm{SpT}/F) : G \mapsto G_{\mathrm{sp}}.$$

For every semiabelian F-variety G, we put $G^b = G/G_{\mathrm{sp}}$. Then $(G^b)_{\mathrm{sp}}$ is trivial by the remark after [Halle and Nicaise 2010, 3.6]. It follows that every morphism of semiabelian F-varieties $f : G \to H$ induces a morphism of semiabelian F-varieties

$$f^b : G^b \to H^b$$

so that we obtain a functor

$$(\,\cdot\,)^b : (\mathrm{SAb}/F) \to (\mathrm{SAb}/F) : G \mapsto G^b.$$

Lemma 3.4.1. *Let F be a field, and let $f : G \to H$ be a smooth morphism of semiabelian K-varieties. Then the morphism $f_{\mathrm{sp}} : G_{\mathrm{sp}} \to H_{\mathrm{sp}}$ is surjective.*

Proof. The identity component of $G \times_H H_{\mathrm{sp}}$ is a smooth and connected closed subgroup of G and thus a semiabelian F-variety [Halle and Nicaise 2011, 5.2]. The morphism

$$(G \times_H H_{\mathrm{sp}})^o \to H_{\mathrm{sp}}$$

is still smooth. Therefore, we may assume that H is a split torus. It follows from [Demazure and Grothendieck 1970a, VI_B.1.2] that the image of f is closed in H, and it is also open by flatness of f. Thus, f is surjective.

We denote by I the schematic image of the morphism $g : G_{\mathrm{sp}} \to H$. This is a closed subgroup of the split torus H. The quotient H/I is again a split F-torus (it is geometrically connected because G_{sp} is geometrically connected, and it is smooth [Demazure and Grothendieck 1970a, VI_B.9.2(xii)] and diagonalizable [Demazure and Grothendieck 1970b, IX.8.1] so that it is a split torus). The quotient $Q = G/G_{\mathrm{sp}}$ is an extension of an abelian F-variety and an anisotropic F-torus so that the morphism of F-groups $Q \to H/I$ induced by f is trivial. But this morphism is surjective by surjectivity of f so that H/I must be trivial and $H = I$. Since the image of $G_{\mathrm{sp}} \to H$ is closed [Demazure and Grothendieck 1970a, VI_B.1.2], it follows that $G_{\mathrm{sp}} \to H$ is surjective. □

Proposition 3.4.2. *Let F be a field, and let*

$$0 \to G_1 \to G_2 \to G_3 \to 0$$

be an exact sequence of semiabelian F-varieties.

(1) *The schematic image of $(G_1)^\mathrm{b} \to (G_2)^\mathrm{b}$ is a semiabelian subvariety H of G_2, and the morphism $(G_1)^\mathrm{b} \to H$ is an isogeny. Moreover, the sequence*

$$0 \to H \to (G_2)^\mathrm{b} \to (G_3)^\mathrm{b} \to 0 \tag{3.4-1}$$

   is exact.

(2) *If $(G_3)_{\mathrm{sp}}$ is trivial, then*

$$0 \to (G_1)^\mathrm{b} \to (G_2)^\mathrm{b} \to (G_3)^\mathrm{b} \to 0 \tag{3.4-2}$$

   is exact.

Proof. Dividing G_1 and G_2 by $(G_1)_{\mathrm{sp}}$, we may assume that $(G_1)_{\mathrm{sp}}$ is trivial (here we use that $G^\mathrm{b} = (G/T)^\mathrm{b}$ for every semiabelian F-variety G and every split subtorus T of G). Then $(G_1)^\mathrm{b} = G_1$.

First, we prove (1). The kernel of the morphism $G_1 \to (G_2)^\mathrm{b}$ is the closed subgroup $\widetilde{G}_1 = G_1 \times_{G_2} (G_2)_{\mathrm{sp}}$ of G_1. It is also a closed subgroup of $(G_2)_{\mathrm{sp}}$. By [Demazure and Grothendieck 1970a, VI_B.9.2(xii)], the quotient $H = G_1/\widetilde{G}_1$ is

smooth over F. Since $(G_2)_{sp}$ is a split F-torus, we know that $\widetilde{G}_1$ is a diagonalizable F-group [Demazure and Grothendieck 1970b, IX.8.1]. Since $(G_1)_{sp}$ is trivial, the F-group $\widetilde{G}_1$ must be finite so that the projection $G_1 \to H$ is an isogeny. The morphism $G_1 \to G_2$ induces a morphism of F-groups $H \to (G_2)^b$ that is a closed immersion [Demazure and Grothendieck 1970a, VI$_B$.1.4.2]. It identifies H with the schematic image of $G_1 \to (G_2)^b$. It follows from [Halle and Nicaise 2011, 5.2] that H is a semiabelian F-variety because it is a connected smooth closed subgroup of the semiabelian F-variety $(G_2)^b$.

It is clear that (3.4-1) is exact at the right so that it remains to prove that this sequence is also exact in the middle. By the natural isomorphism

$$(G_2/G_1)/((G_2)_{sp}/\widetilde{G}_1) \cong (G_2/(G_2)_{sp})/(G_1/\widetilde{G}_1),$$

it is enough to show $\widetilde{G}_2 = (G_2)_{sp}/\widetilde{G}_1$ is the maximal split subtorus of $G_3 = G_2/G_1$. But $(G_2)_{sp} \to (G_3)_{sp}$ is surjective by Lemma 3.4.1, and its kernel is precisely $\widetilde{G}_1$, so we see that $\widetilde{G}_2 = (G_3)_{sp}$.

Now we prove (2). Assume that $(G_3)_{sp}$ is trivial. Then the closed immersion $(G_2)_{sp} \to G_2$ factors through G_1, and since $(G_1)_{sp}$ is trivial, we find that $(G_2)_{sp}$ must be trivial. Thus, $G_i = (G_i)^b$ for $i = 1, 2, 3$, and the result is obvious. $\qquad\square$

If $(G_3)_{sp}$ is not trivial, it can happen that the sequence

$$0 \to (G_1)^b \to (G_2)^b \to (G_3)^b \to 0$$

in Proposition 3.4.2 is not left exact as is shown by the following:

Example 3.4.3. Let K be the field $\mathbb{C}((t))$ of complex Laurent series, and put $K' = K((\sqrt{t}))$. The Galois group $\Gamma = \mathrm{Gal}(K'/K)$ is isomorphic to $\mathbb{Z}/2\mathbb{Z}$, and it is generated by the automorphism σ that maps $\sqrt{t}$ to $-\sqrt{t}$. Let G_2 be the K-torus with splitting field K' and character module

$$X(G_2) = \mathbb{Z} \cdot e_1 \oplus \mathbb{Z} \cdot e_2$$

where σ permutes e_1 and e_2.

Let G_1 be the maximal anisotropic subtorus of G_2. Its character module is $X(G_1) = X(G_2)/X(G_2)^\Gamma$. We put $G_3 = G_2/G_1$. This is a split K-torus with character module $X(G_3) = X(G_2)^\Gamma = \mathbb{Z} \cdot (e_1 + e_2)$.

For every K-torus T that splits over K', we can consider the trace map

$$\mathrm{tr}_T : X(T) \to X(T)^\Gamma : x \mapsto x + \sigma \cdot x.$$

It follows from the duality between tori and their character modules that the maximal split subtorus of T has character module $X(T)/\ker(\mathrm{tr}_T)$ and that T^b is the K-torus with character module $\ker(\mathrm{tr}_T)$. In this way, we see that $(G_1)_{sp}$ is trivial and that

$(G_2)^b$ is the K-torus with character module

$$\ker(\mathrm{tr}_{G_2}) = \mathbb{Z} \cdot (e_1 - e_2).$$

Thus, applying the functor $(\cdot)^b$ to the exact sequence of K-tori

$$0 \to G_1 \to G_2 \to G_3 \to 0,$$

we obtain the sequence

$$0 \to G_1 \to (G_2)^b \to 0 \to 0 \tag{3.4-3}$$

and the morphism of K-tori $G_1 \to (G_2)^b$ corresponds to the morphism of character modules

$$\alpha : \mathbb{Z} \cdot (e_1 - e_2) \to X(G_2)/X(G_2)^\Gamma.$$

The morphism α is injective but not surjective; its cokernel is

$$X(G_2)/(\mathbb{Z} \cdot (e_1 - e_2) + \mathbb{Z} \cdot (e_1 + e_2)) \cong \mathbb{Z}/2\mathbb{Z}$$

with trivial Γ-action. Therefore, (3.4-3) is not exact. More precisely, the morphism

$$G_1 \to (G_2)^b$$

is an isogeny with kernel $\mu_{2,K}$.

Proposition 3.4.4. *Assume that R is excellent and that k is algebraically closed. For every semiabelian K-variety G, the quotient G^b is a bounded semiabelian K-variety.*

Proof. Since $(G^b)_{\mathrm{sp}}$ is trivial, this follows immediately from [Bosch et al. 1990, 10.2.1]. □

4. Chai's conjecture and Fubini properties of motivic integrals

4.1. *Chai's conjecture and Haar measures.* Let

$$0 \to T \to G \to A \to 0$$

be a short exact sequence of semiabelian K-varieties (as the notation suggests, the main example we have in mind is the Chevalley decomposition of a semiabelian K-variety G as in Conjecture 1.1, but we will work in greater generality). The sequence of K-vector spaces

$$0 \to \mathrm{Lie}(T) \to \mathrm{Lie}(G) \to \mathrm{Lie}(A) \to 0$$

is exact, and by dualizing and taking determinants, we find a canonical isomorphism of K-vector spaces

$$\Omega_G \otimes_R K \cong (\Omega_T \otimes_R \Omega_A) \otimes_R K.$$

In this way, we can view $\Omega_T \otimes_R \Omega_A$ as an R-lattice in $\Omega_G \otimes_R K$.

The following proposition is implicit in the proof on pages 724–725 of [Chai 2000] (proof of Proposition 4.1 in [loc. cit.] when the residue field is finite):

Proposition 4.1.1. *Assume that T is a torus. Let ω_T and ω_A be generators of Ω_T and Ω_A, respectively. Let ϖ be a uniformizer of R, and denote by γ the unique integer such that $\varpi^{-\gamma}(\omega_T \otimes \omega_A)$ generates the R-module Ω_G. Then*

$$c(G) = c(T) + c(A) + \gamma.$$

In particular, $c(G) - c(T) - c(A)$ belongs to $\mathbb{Z}$.

Proof. By Proposition 3.3.2, we may assume that R is complete and that k is algebraically closed. Suppose that

$$\omega_G := \varpi^{-\gamma}(\omega_T \otimes \omega_A)$$

generates Ω_G. Let K' be a finite separable extension of K such that $G' = G \times_K K'$ has semiabelian reduction, and denote by R' the normalization of R in K'. Then $A' = A \times_K K'$ has semiabelian reduction and $T' = T \times_K K'$ is split [Halle and Nicaise 2010, 4.1].

We denote by ϖ' a uniformizer of R' and d the ramification degree of the extension K'/K. By Proposition 3.3.1, the R'-module $\Omega_{G'}$ is generated by

$$(\varpi')^{c(G)d}\omega_G,$$

and the analogous property holds for A and T. We denote by $\mathcal{G}'$, $\mathcal{T}'$ and $\mathcal{A}'$ the Néron lft-models of G', T' and A', respectively. By the universal property of the Néron lft-model, the exact sequence

$$0 \to T' \to G' \to A' \to 0$$

extends uniquely to a sequence of R'-group schemes

$$0 \to \mathcal{T}' \to \mathcal{G}' \to \mathcal{A}' \to 0$$

and this sequence is exact by [Chai 2000, 4.8(a)]. It follows that

$$\Omega_{G'} = \Omega_{T'} \otimes_{R'} \Omega_{A'} \subset \Omega_{G'} \otimes_{R'} K'$$

so that both $(\varpi')^{c(G)d}\omega_G$ and

$$(\varpi')^{(c(T)+c(A))d}(\omega_T \otimes \omega_A) = (\varpi')^{(c(T)+c(A))d}\varpi^\gamma \omega_G$$

are generators of the free R'-module $\Omega_{G'}$. Thus, we find that

$$(\varpi')^{(c(G)-c(T)-c(A))d}\varpi^{-\gamma}$$

is a unit in R'. This means that its ϖ'-adic valuation is zero so that

$$c(G) = c(T) + c(A) + \gamma. \qquad \square$$

Remark 4.1.2. Let

$$0 \to T \to G \to A \to 0$$

be an exact sequence of semiabelian K-varieties, and let ω_T and ω_A be generators of Ω_T and Ω_A, respectively. Chai [2000, §8.1] considers the following statement:

> *One has $c(G) = c(T) + c(A)$ if and only if $\omega_T \otimes \omega_A$ generates Ω_G.* (∗)

If T is a torus, then this is a corollary of Proposition 4.1.1. However, if T is not a torus, it is not clear to us how statement (∗) can be proven although Chai hints that it may be implicit in the proof on pages 724–725 of [Chai 2000]. If G, T and A have semiabelian reduction, then $c(G) = c(T) = c(A) = 0$ so that statement (∗) contains the following special case:

> *If G, T and A have semiabelian reduction, then $\omega_T \otimes \omega_A$ generates Ω_G.* (∗∗)

Even this property does not seem obvious because the sequence of identity components of Néron lft-models

$$0 \to \mathcal{T}^o \to \mathcal{G}^o \to \mathcal{A}^o \to 0$$

might not be exact; see [Bosch et al. 1990, 7.5.8] for an example where T, G and A are abelian varieties with good reduction and $\mathcal{T}^o \to \mathcal{G}^o$ is not a monomorphism. If statement (∗∗) is true, then the proof of Proposition 4.1.1 shows that Proposition 4.1.1, and thus statement (∗), are valid without the assumption that T is a torus.

Lemma 4.1.3. *Let G be a semiabelian K-variety, and let T be a split subtorus of G. Then $c(G) = c(G/T)$. In particular, $c(G) = c(G^{\flat})$.*

Proof. We set $H = G/T$. By [Chai 2000, 4.8(a)], the canonical sequence of group schemes

$$0 \to \mathcal{T} \to \mathcal{G} \to \mathcal{H} \to 0$$

is exact so that $\Omega_T \otimes \Omega_H = \Omega_G$. Now the result follows from Proposition 4.1.1 and the fact that $c(T) = 0$ because T has semiabelian reduction. $\square$

4.2. *Main results.* We can now state our main results.

Theorem 4.2.1. *Let*

$$0 \to T \to G \to A \to 0$$

be an exact sequence of semiabelian K-varieties with T a torus. We put

$$\widetilde{G} = (G \times_K \widehat{K^{sh}})^{\mathrm{b}} \quad and \quad \widetilde{A} = (A \times_K \widehat{K^{sh}})^{\mathrm{b}},$$

and we denote by $\widetilde{T}$ the schematic image of the morphism

$$(T \times_K \widehat{K^{sh}})^{\mathrm{b}} \to \widetilde{G}.$$

Then $\widetilde{T}$ is a $\widehat{K^{sh}}$-subtorus of $\widetilde{G}$,

$$0 \to \widetilde{T} \to \widetilde{G} \to \widetilde{A} \to 0$$

is an exact sequence of bounded semiabelian $\widehat{K^{sh}}$-varieties and

$$c(G) = c(T) + c(A)$$

if and only if

$$\dim \int_{\widetilde{G}} |\omega_{\widetilde{T}} \otimes \omega_{\widetilde{A}}| = \dim \int_{\widetilde{T}} |\omega_{\widetilde{T}}| + \dim \int_{\widetilde{A}} |\omega_{\widetilde{A}}| = 0. \tag{4.2-1}$$

Proof. By Proposition 3.3.2, we may assume that K is complete and k algebraically closed so that $\widehat{K^{sh}} = K$. By Proposition 3.4.2, we know that T^{b} and $\widetilde{T}$ are isogenous K-tori so that $c(T^{\mathrm{b}}) = c(\widetilde{T})$ by [Chai and Yu 2001, 11.3 and 12.1]. Thus, by Proposition 3.4.2 and Lemma 4.1.3, we may assume that $\widetilde{T} = T$, $\widetilde{G} = G$ and $\widetilde{A} = A$. Then T, G and A are bounded, by Proposition 3.4.4, so that we can take motivic integrals of gauge forms on T, G and A.

Let ϖ be a uniformizer in R. It follows from Proposition 3.3.3 that

$$\dim \int_T |\omega_T| + \dim \int_A |\omega_A| = 0$$

and that

$$\dim \int_G |\omega_T \otimes \omega_A|$$

is equal to the unique integer γ such that

$$\varpi^\gamma (\omega_T \otimes \omega_A)$$

generates the R-module Ω_G. By Proposition 4.1.1, we know that $\gamma = 0$ if and only if $c(G) = c(T) + c(A)$. □

Remark 4.2.2. In Conjecture 1.1, T is a torus and A is an abelian variety. This implies that A_{sp} is trivial so that $A = A^{\mathrm{b}}$ and

$$0 \to T^{\mathrm{b}} \to G^{\mathrm{b}} \to A \to 0$$

is exact by Proposition 3.4.2. In this case, in the proof of Theorem 4.2.1, we do not need the fact that the base change conductor of a torus is invariant under isogeny.

Theorem 4.2.3. *Assume that K is complete and of characteristic zero and that k is algebraically closed. Let*

$$0 \to T \to G \to A \to 0$$

be an exact sequence of bounded semiabelian K-varieties with T a torus. Then

$$\dim \int_G |\omega_T \otimes \omega_A| = \dim \int_T |\omega_T| + \dim \int_A |\omega_A| = 0.$$

We will prove Theorem 4.2.3 in Section 5.4, using the model-theoretic approach to motivic integration in [Cluckers and Loeser 2012] and a new result on dimensions of motivic integrals (Theorem 5.2.1). As a corollary, we obtain a new proof of the following theorem:

Theorem 4.2.4 (Chai). *Let*

$$0 \to T \to G \to A \to 0$$

be an exact sequence of semiabelian K-varieties with T a torus. Assume that K is of characteristic zero. Then

$$c(G) = c(T) + c(A).$$

Proof. This follows at once from Theorem 4.2.1 and Theorem 4.2.3.   □

Remark 4.2.5. Theorem 4.2.3 is not a direct corollary of the Fubini theorem in [Cluckers and Loeser 2012] and the above results. We need to combine the Fubini result [Cluckers and Loeser 2012, 12.5] with the new result in Theorem 5.2.1 and its corollary below, which compares dimensions of motivic parameter integrals under rather general conditions. By the lack of a definable section for the morphism $G \to A$ as in Theorem 4.2.3, the motivic integral of $|\omega_G|$ over G may not be equal to the product of the integrals of $|\omega_T|$ over T and of $|\omega_A|$ over A. By the corollary to Theorem 5.2.1 and by the change of variables in [Cluckers and Loeser 2012, 12.4], this product survives at the rough level of virtual dimensions, which is sufficient to prove Theorem 4.2.3.

5. A comparison result for the dimensions of motivic integrals

In this section, we will work with a specific context falling under [Cluckers and Loeser 2012] and define the dimension of motivic constructible functions at each point. These functions play an important role in motivic integration and in general Fubini results of [Cluckers and Loeser 2012; 2008]. In order to control dimensions as desired for Equation (4.2-1) in Theorem 4.2.1, we will compare the dimensions of the integrals of possibly different functions F and G when we are given that F and G have the same dimension in every point. Theorem 5.2.1 provides such

a comparison result with parameters, and its corollary gives a similar comparison result for integrals on an algebraic variety with a volume form.

5.1. *Dimensions of motivic constructible functions.* In this section, we suppose R is a complete discrete valuation ring of characteristic zero with quotient field K and algebraically closed residue field k of characteristic $p \geq 0$. We fix a uniformizer ϖ in K. We will use some terminology and results from [Cluckers and Loeser 2012] with precise references. Let $\mathscr{L}_{\text{high}}(K)$ be the Denef–Pas language $\mathscr{L}_{\text{high}}$ as in [Cluckers and Loeser 2012, §2.3] enriched with coefficients from K and where the angular component maps $\overline{\text{ac}}_n$ for $n > 0$ are given by $\overline{\text{ac}}_n(x) = x\varpi^{-\operatorname{ord} x} \bmod (\varpi)^n$ for nonzero $x \in K$. Let $\mathcal{T}$ be the $\mathscr{L}_{\text{high}}(K)$-theory of K.

Since $\mathcal{T}$ falls under the combined Examples 1 and 4 of [Cluckers and Loeser 2012, §3.1], we can use the theory of motivic integration of [Cluckers and Loeser 2012]. Also, since $\mathcal{T}$ is a complete theory, any definable subassignment X is uniquely determined by the definable set $X(K)$ with notation from [Cluckers and Loeser 2012, §4.1]. We will sometimes say "definable set" instead of "definable subassignment".

We first define how to take (virtual) dimensions of several objects appearing in [Cluckers and Loeser 2012]. Write R_n for the ring $R/(\varpi^n)$ and ϖ_n for the image of ϖ in R_n. Let $\mathscr{L}_r$ be the multisorted language with sorts R_n for integers $n > 0$, on each R_n the ring language with coefficients from R_n, and with the natural projection maps $p_{n,m} : R_n \to R_m$ for $n \geq m$. It follows from the quantifier elimination results of [Pas 1989; 1990] that any $\mathscr{L}_{\text{high}}(K)$-definable set $X \subset \prod_{i=1}^{s} R_{n_i}$ is already $\mathscr{L}_r$-definable with parameters; see also [Cluckers and Loeser 2012, Theorem 3.10]. To each $\mathscr{L}_r$-definable set $X \subset \prod_{i=1}^{s} R_{n_i}$, we associate an $\mathscr{L}_{\text{ring}}(k)$-definable set $\delta(X)$ as follows. If R has mixed characteristic, the projection $p_{n,1}$ induces a bijection from the set of p^n-th powers in R_n to k by Hensel's lemma, Newton's binomial theorem and the hypotheses on K. Let us write P_{p^n} for the set of p^n-th powers in R_n. Then any x in R_n can be written uniquely as

$$\sum_{i=0}^{n-1} x_i \varpi_n^i,$$

with $x_i \in P_{p^n}$, yielding a bijection $R_n \to k^n : x \mapsto (p_{n,1}(x_i))_i$ that is, in fact, $\mathscr{L}_r$-definable. If R has equal characteristic zero, we choose a retraction $k \to R$ of the ring morphism $R \to k$. This choice determines an isomorphism $R \to k[[\varpi]]$, and we identify $R_n \cong k[\varpi]/(\varpi^n)$ with the k-vector space k^n by means of the basis $1, \varpi, \ldots, \varpi^{n-1}$ of R_n. In both cases, we obtain a bijection $\prod_{i=1}^{s} R_{n_i} \to k^N$ with $N = \sum_{i=1}^{s} n_i$. This identification maps the $\mathscr{L}_r$-definable subset X of $\prod_{i=1}^{s} R_{n_i}$ onto an $\mathscr{L}_{\text{ring}}(k)$-definable subset of k^N that we denote by $\delta(X)$.

Recall from [Cluckers and Loeser 2012, §7.1] that $\mathscr{C}_+(\text{Point})$ is the Grothendieck semiring of $\mathscr{L}_{\text{high}}(K)$-definable subsets of Cartesian products of the form $\prod_{i=1}^s R_{n_i}$ up to definable isomorphisms with scissor relations, with zero-element $[\varnothing]$ and localized with respect to $\mathbb{L}$ and the elements $\mathbb{L}^i - 1$ for all $i > 0$, where $\mathbb{L}$ stands for the class of the affine line over k. Clearly, δ induces a semiring morphism

$$\mathscr{C}_+(\text{Point}) \to \mathscr{M}_k^+,$$

which we also denote by δ. Recall that objects in $\mathscr{M}_k^+$ have a dimension by Definition 3.1.1.

For an $\mathscr{L}_{\text{high}}(K)$-definable set Z, $\mathscr{C}_+(Z)$ is a relative variant of $\mathscr{C}_+(\text{Point})$ over Z; see [Cluckers and Loeser 2012, §7.1]. An object $\varphi \in \mathscr{C}_+(Z)$ is called a motivic constructible function on Z. Moreover, for every $z \in Z(K)$, there is the evaluation map $i_z^* : \mathscr{C}_+(Z) \to \mathscr{C}_+(\text{Point})$ at z, and $i_z^*(\varphi)$ is called the evaluation of φ at z. For an $\mathscr{L}_{\text{high}}(K)$-definable set Z, a point $z \in Z(K)$ and a function φ in $\mathscr{C}_+(Z)$, the dimension of φ at z is defined as $\dim(\delta(i_z^*(\varphi)))$ and is denoted by $\dim_z(\varphi)$. If Z is the point and $\varphi \in \mathscr{C}_+(\text{Point})$, we write $\dim(\varphi)$ instead of $\dim_{\text{Point}}(\varphi)$.

5.2. *A comparison result.* In this section, definable will mean for the language $\mathscr{L}_{\text{high}}(K)$. Recall that, for definable sets X, Y and $Z \subset X \times Y$, under integrability conditions in the fibers of the projection $Z \to X$, called X-integrability, one can integrate $\varphi \in \mathscr{C}_+(Z)$ in the fibers of the projection $Z \to X$ to obtain a function $\mu_{/X}(\varphi)$ in $\mathscr{C}_+(X)$; see [Cluckers and Loeser 2012, 9.1]. The method of [Cluckers and Loeser 2008; 2012] for calculating integrals goes back to ideas by Denef [1984] in the p-adic case and to Pas [1989; 1990] in a premotivic setting.

Now we can state and prove our comparison result, stating that the dimension of a motivic integral only depends on the dimensions of the values of the integrand at each point.

Theorem 5.2.1. *Let F and G be in $\mathscr{C}_+(Z)$, and suppose that $Z \subset X \times Y$ for some definable sets X, Y and Z. Suppose that F and G are X-integrable and that*

$$\dim_z(F) = \dim_z(G) \text{ for each point } z \text{ on } Z(K).$$

Then one has

$$\dim_x(\mu_{/X}(F)) = \dim_x(\mu_{/X}(G)) \text{ for each point } x \text{ on } X(K).$$

Proof. For some integers $n, r, s \geq 0$ and for some tuple $m = (m_1, \ldots, m_s)$ of nonnegative integers, Y is contained in $K^n \times \prod_{1 \leq i \leq s} R_{m_i} \times \mathbb{Z}^r$. By projecting one variable at a time and by iterating the one variable result, it suffices to consider the case where two of the three values n, m and r are zero and either $n = 1, r = 1$ or $s = 1$ and $m_1 = 1$. By the cell decomposition theorem of [Pas 1989; 1990], we may suppose that $n = 0$. Indeed, via cell decomposition, each integral over a valued

field variable is precisely calculated as a sum over $\mathbb{Z}$-variables and a subsequent integral over residue ring variables; see [Cluckers and Loeser 2012, §8].

Recall that F is a finite sum of terms of the form $a_i \otimes b_i$, with $a_i \in \mathscr{P}_+(Z)$ and $b_i \in \mathscr{Q}_+(Z)$, and similarly for G with notation from [Cluckers and Loeser 2012, §7.1]. (The semiring $\mathscr{P}_+(Z)$ is related to the value group and $\mathscr{Q}_+(Z)$ to the residue field, and $\mathscr{C}_+(Z)$ is a tensor product of both.)

If $n = r = 0$, then we may suppose that the a_i lie in $\mathscr{P}_+(X)$, and similarly for G, by Proposition 7.5 of [Cluckers and Loeser 2012]. The result of the theorem now follows from the definition in [Cluckers and Loeser 2012, §6] of $\mu_{/X}$ in this case and the following simple comparison property for dimensions of constructible sets $A_i \subset \mathbb{A}_k^{n_i}$. If, for certain morphisms $f_i : \mathbb{A}_k^{n_i} \to \mathbb{A}_k^{n_3}$ for $i = 1, 2$, one has that $f_1(A_1) = f_2(A_2)$ and, for each $x \in \mathbb{A}_k^{n_3}(k)$, the dimension of $f_1^{-1}(x) \cap A_1$ equals the dimension of $f_2^{-1}(x) \cap A_2$, then one has $\dim(A_1) = \dim(A_2)$.

Let us finally consider the case that $n = m = 0$ and $r = 1$. In this case, we may suppose that the b_i lie in $\mathscr{Q}_+(X)$, and similarly for G, again by Proposition 7.5 of [Cluckers and Loeser 2012]. In the considered case, the theorem follows from the definition of $\mu_{/X}$ of [Cluckers and Loeser 2012] and the following two observations. For any $a \in \mathbb{A}$, where $a = a(\mathbb{L})$ is thus a rational function in $\mathbb{L}$ of a specific kind, one has that $\dim(a)$ equals the degree of the rational function $a(\mathbb{L})$, where the degree of a rational function is the degree of its numerator minus the degree of its denominator and where the degree of 0 is defined as $-\infty$. Secondly, there is the following elementary comparison property for the degrees of rational functions. Consider, for each $i \in \mathbb{Z}$, an integer $n_i \geq 0$ and a polynomial $a_i(x)$ over $\mathbb{Z}$ in one variable x such that $a_i(q) \geq 0$ for each real $q > 1$. If there is a rational function $r(x)$ such that $\sum_{i \in \mathbb{Z}} a_i(q)/q^{n_i}$ converges and equals $r(q)$ for each real $q > 1$, then

$$\deg(r(x)) = \max_i \deg\left(\frac{a_i(x)}{x^{n_i}}\right),$$

where deg stands for the degree. Since summation of nonnegative functions over $\mathbb{Z}$ in [Cluckers and Loeser 2012, §5] is calculated and defined by considering specific sums of rational functions in $\mathbb{L}$ and by evaluating in real numbers $q > 1$, the result follows. □

By working with affine charts over K, one may consider a variety V over K as a definable subassignment, and one defines $\mathscr{C}_+(V)$ correspondingly; see [Cluckers and Loeser 2012, §12.3]. Let us write $\int^{\mathrm{CL}}$ for the integral as defined there, to distinguish them from the integrals from Section 3.2 of this paper. The definition of these integrals is based on finite affine covers of V over K, on finite additivity for motivic integrals and on the change of variables formula.

Corollary 5.2.2. *Let V be an algebraic variety over K with a volume form ω_V. Let F and G be integrable functions in $\mathscr{C}_+(V)$ such that, for each $x \in V(K)$, one has*

$\dim_x(F) = \dim_x(G)$. *Then their integrals have the same dimension*:

$$\dim \int_V^{CL} F|\omega_V| = \dim \int_V^{CL} G|\omega_V|.$$

Proof. This follows immediately from Theorem 5.2.1 and the definition of the integrals $\int^{CL}$ in [Cluckers and Loeser 2012, §12.3]. □

5.3. *Gel′fand–Leray residues.* Let $f : X \to Y$ be a smooth morphism between smooth equidimensional varieties over K. Let m be the dimension of Y, and let $m + n$ be the dimension of X. Let ω_X and ω_Y be differential forms of maximal degree on X and Y, respectively. Assume that ω_Y is a gauge form, that is, a generator of the line bundle $\Omega_{Y/K}^m$ at each point of Y.

Since f is smooth, the fundamental sequence of locally free coherent $\mathcal{O}_X$-modules

$$0 \to f^*\Omega_{Y/K}^1 \to \Omega_{X/K}^1 \to \Omega_{X/Y}^1 \to 0$$

is exact [Grothendieck and Dieudonné 1967, 17.2.3]. Taking maximal exterior powers, we obtain an isomorphism

$$f^*\Omega_{Y/K}^m \otimes \Omega_{X/Y}^n \to \Omega_{X/K}^{m+n}.$$

Locally, this isomorphism is defined by

$$\varphi \otimes \eta \mapsto \varphi \wedge \widetilde{\eta}$$

where $\widetilde{\eta}$ is any lift of η to $\Omega_{X/K}^n$. Since $f^*\omega_Y$ generates the line bundle $f^*\Omega_{Y/K}^m$, we obtain an isomorphism $\Omega_{X/Y}^n \to \Omega_{X/K}^{m+n}$ that is locally defined by

$$\eta \mapsto f^*\omega_Y \wedge \widetilde{\eta}.$$

The inverse image of ω_X under this isomorphism is called the Gel′fand–Leray form associated to ω_X and ω_Y and denoted by ω_X/ω_Y. It induces a differential form of maximal degree on each of the fibers of f.

5.4. *Proof of Theorem 4.2.3.* It follows from Proposition 3.3.3 that

$$\dim \int_T |\omega_T| = \dim \int_A |\omega_A| = 0.$$

It is proven in Proposition 12.6 of [Cluckers and Loeser 2012] that the theory of motivic integration developed there can be used to compute the motivic integrals defined by the formula (3.2-1). In particular, the dimensions of the respective motivic integrals are the same so that we can use the corollary of Theorem 5.2.1 and the results in [Cluckers and Loeser 2012] to prove Theorem 4.2.3.

We denote the projection morphism $G \to A$ by f. Since $H^1(K, T) = 0$ by [Chai 2000, 4.3], the map $f(K) : G(K) \to A(K)$ is surjective. For every $a \in A(K)$,

we set $G_a = f^{-1}(a)$. If we choose a point x in $G_a(K)$, then the multiplication by x defines an isomorphism $\tau_x : T \to G_a$. Since the relative differential form $(\omega_T \otimes \omega_A)/\omega_A$ on G is invariant under translation, the pullback through τ_x of its restriction to G_a equals ω_T. Thus, by the change of variables formula in [Cluckers and Loeser 2012, 12.4], we have

$$\int_{G_a}^{\mathrm{CL}} |(\omega_T \otimes \omega_A)/\omega_A| = \int_T^{\mathrm{CL}} |\omega_T|$$

for each a in $A(K)$. Hence, by the Fubini property in [Cluckers and Loeser 2012, 12.5], we find that

$$\int_G^{\mathrm{CL}} |\omega_T \otimes \omega_A| = \int_A^{\mathrm{CL}} \psi |\omega_A|$$

where ψ is a motivic constructible function on A such that $\dim_a(\psi) = 0$ for each $a \in A(K)$. Now Corollary 5.2.2 with $V = A$ implies that

$$\dim \int_A^{\mathrm{CL}} \psi |\omega_A| = \dim \int_A^{\mathrm{CL}} |\omega_A| = 0.$$

Combining the above equations, we find

$$\dim \int_G |\omega_T \otimes \omega_A| = \dim \int_G^{\mathrm{CL}} |\omega_T \otimes \omega_A| = 0,$$

which concludes the proof.

References

[Bosch et al. 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik und ihrer Grenzgebiete (3) **21**, Springer, Berlin, 1990. MR 91i:14034 Zbl 0705.14001

[Chai 2000] C.-L. Chai, "Néron models for semiabelian varieties: Congruence and change of base field", *Asian J. Math.* **4**:4 (2000), 715–736. MR 2002i:14025 Zbl 1100.14511

[Chai and Yu 2001] C.-L. Chai and J.-K. Yu, "Congruences of Néron models for tori and the Artin conductor", *Ann. of Math.* (2) **154**:2 (2001), 347–382. MR 2003e:11126 Zbl 1098.14014

[Cluckers and Loeser 2008] R. Cluckers and F. Loeser, "Constructible motivic functions and motivic integration", *Invent. Math.* **173**:1 (2008), 23–121. MR 2009g:14018 Zbl 1179.14011

[Cluckers and Loeser 2012] R. Cluckers and F. Loeser, "Motivic integration in all residue field characteristics for Henselian discretely valued fields of characteristic zero", preprint, 2012. arXiv 1102.3832

[Demazure and Grothendieck 1970a] M. Demazure and A. Grothendieck (editors), *Schémas en groupes, I: Propriétés générales des schémas en groupes*, Lecture Notes in Mathematics **151**, Springer, Berlin, 1970. MR 43 #223a Zbl 0207.51401

[Demazure and Grothendieck 1970b] M. Demazure and A. Grothendieck (editors), *Schémas en groupes, II: Groupes de type multiplicatif, et structure des schémas en groupes généraux*, Lecture Notes in Mathematics **152**, Springer, Berlin, 1970. MR 43 #223b Zbl 0209.24201

[Denef 1984] J. Denef, "The rationality of the Poincaré series associated to the p–adic points on a variety", *Invent. Math.* **77**:1 (1984), 1–23. MR 86c:11043 Zbl 0537.12011

[Gross and Gan 1999] B. H. Gross and W. T. Gan, "Haar measure and the Artin conductor", *Trans. Amer. Math. Soc.* **351**:4 (1999), 1691–1704. MR 99f:20078 Zbl 0991.20033

[Grothendieck and Dieudonné 1967] A. Grothendieck and J. Dieudonné, "Éléments de géométrie algébrique, IV: Étude locale des schémas et des morphismes de schémas, IV", *Inst. Hautes Études Sci. Publ. Math.* 32 (1967), 361. MR 39 #220 Zbl 0153.22301

[Grothendieck et al. 1972] A. Grothendieck, M. Raynaud, and D. S. Rim (editors), *Groupes de monodromie en géométrie algébrique, I*, Lecture Notes in Mathematics **288**, Springer, Berlin, 1972. MR 50 #7134 Zbl 0237.00013

[Halle and Nicaise 2010] L. H. Halle and J. Nicaise, "The Néron component series of an abelian variety", *Math. Ann.* **348**:3 (2010), 749–778. MR 2011j:11105 Zbl 1245.11072

[Halle and Nicaise 2011] L. H. Halle and J. Nicaise, "Motivic zeta functions of abelian varieties, and the monodromy conjecture", *Adv. Math.* **227**:1 (2011), 610–653. MR 2012c:14050 Zbl 1230.11076

[Loeser and Sebag 2003] F. Loeser and J. Sebag, "Motivic integration on smooth rigid varieties and invariants of degenerations", *Duke Math. J.* **119**:2 (2003), 315–344. MR 2004g:14026 Zbl 1078.14029

[Nicaise 2009] J. Nicaise, "A trace formula for rigid varieties, and motivic Weil generating series for formal schemes", *Math. Ann.* **343**:2 (2009), 285–349. MR 2010b:14043 Zbl 1177.14050

[Nicaise 2011] J. Nicaise, "A trace formula for varieties over a discretely valued field", *J. Reine Angew. Math.* **650** (2011), 193–238. MR 2012d:14039 Zbl 1244.14017

[Nicaise and Sebag 2008] J. Nicaise and J. Sebag, "Motivic Serre invariants and Weil restriction", *J. Algebra* **319**:4 (2008), 1585–1610. MR 2009e:14041 Zbl 1211.14015

[Nicaise and Sebag 2011a] J. Nicaise and J. Sebag, "The Grothendieck ring of varieties", pp. 145–188 in *Motivic integration and its interactions with model theory and non-Archimedean geometry, I*, edited by R. Cluckers et al., London Math. Soc. Lecture Note Ser. **383**, Cambridge Univ. Press, 2011. MR 2885336 Zbl 06059495

[Nicaise and Sebag 2011b] J. Nicaise and J. Sebag, "Motivic invariants of rigid varieties, and applications to complex singularities", pp. 244–304 in *Motivic integration and its interactions with model theory and non-Archimedean geometry, I*, edited by R. Cluckers et al., London Math. Soc. Lecture Note Ser. **383**, Cambridge Univ. Press, 2011. MR 2885338 Zbl 06059497

[Pas 1989] J. Pas, "Uniform *p*–adic cell decomposition and local zeta functions", *J. Reine Angew. Math.* **399** (1989), 137–172. MR 91g:11142 Zbl 0666.12014

[Pas 1990] J. Pas, "Cell decomposition and local zeta functions in a tower of unramified extensions of a *p*–adic field", *Proc. London Math. Soc.* (3) **60**:1 (1990), 37–67. MR 91g:11143 Zbl 0659.12017

raf.cluckers@math.univ-lille1.fr   *Laboratoire Painlevé, Université Lille 1, UMR CNRS 8524, Cité Scientifique, 59655 Villeneuve d'Ascq, France* http://math.univ-lille1.fr/~cluckers

loeser@math.jussieu.fr   *Institut de Mathématiques de Jussieu, Université Pierre et Marie Curie, UMR CNRS 7586, 4 place Jussieu, 75252 Paris, France* http://www.math.jussieu.fr/~loeser

johannes.nicaise@wis.kuleuven.be   *Department of Mathematics, Katholieke Universiteit Leuven, Celestijnenlaan 200B, 3001 Heverlee, Belgium* https://perswww.kuleuven.be/~u0025871/

Adjoint ideals and a correspondence between log canonicity and *F*-purity

Shunsuke Takagi

Dedicated to Professor Shihoko Ishii on the occasion of her sixtieth birthday

This paper presents three results on F-singularities. First, we give a new proof of Eisenstein's restriction theorem for adjoint ideal sheaves using the theory of F-singularities. Second, we show that a conjecture of Mustață and Srinivas implies a conjectural correspondence of F-purity and log canonicity. Finally, we prove this correspondence when the defining equations of the variety are very general.

Introduction

This paper deals with the theory of F-singularities, which are singularities defined using the Frobenius morphism in positive characteristic. We present three main results. First, we give a new proof of a restriction theorem for adjoint ideal sheaves. Second, we show that a certain arithmetic conjecture implies a conjectural correspondence of F-purity and log canonicity. Finally, we prove this correspondence when the defining equations of the variety are very general.

The notion of the adjoint ideal sheaf along a normal $\mathbb{Q}$-Gorenstein closed subvariety X of a smooth complex variety A with codimension c was introduced in [Takagi 2010] (see Definition 1.8 for its definition). It is a modification of the multiplier ideal sheaf associated to the pair (A, cX) and encodes much information on the singularities of X. Eisenstein [2010] recently proved a restriction theorem for these adjoint ideal sheaves. In this paper, we give a new proof of his result using the theory of F-singularities.

Building on earlier results [Hara and Yoshida 2003; Takagi 2004b; 2008], we introduced in [Takagi 2010] a positive characteristic analogue of the adjoint ideal sheaf called the test ideal sheaf (see Proposition-Definition 1.1). We conjectured that the adjoint ideal sheaf coincides after reduction to characteristic $p \gg 0$ with the test ideal sheaf and some partial results were obtained in [loc. cit.]. Making use of these results, we reduce the problem to an ideal theoretic problem on a normal

$\mathbb{Q}$-Gorenstein ring essentially of finite type over a perfect field of characteristic $p > 0$. The desired restriction formula is then obtained by adapting the argument of [Schwede 2009] (which can be traced back to [Fedder 1983]) to our setting (see Theorem 3.2). As a corollary, we show the correspondence between adjoint ideal sheaves and test ideal sheaves in a full generality (Corollary 3.4 on page 934).

The other ingredients of this paper are on a correspondence between F-pure singularities and log canonical singularities. F-pure singularities are defined via splitting of Frobenius morphisms (see Definition 1.3). Log canonical singularities form a class of singularities associated to the minimal model program (see Definition 1.7). It is known that the pair $(X; tZ)$ is log canonical if its modulo p reduction $(X_p; tZ_p)$ is F-pure for infinitely many primes p, and the converse is conjectural (see Conjecture 2.4 for the precise statement). This conjecture is widely open, and only a few special cases are known. On the other hand, Mustață and Srinivas [2011, Conjecture 1.1] proposed the following more arithmetic conjecture to study a behavior of test ideal sheaves: if V is a d-dimensional smooth projective variety over an algebraically closed field of characteristic zero, the action induced by the Frobenius morphism on the cohomology group $H^d(V_p, \mathcal{O}_{V_p})$ of its modulo p reduction V_p is bijective for infinitely many primes p. In this paper, we show that their conjecture implies the correspondence of F-purity and log canonicity (see Theorem 2.11). Our result can be viewed as strong evidence in favor of this conjectural correspondence although the conjecture of Mustață–Srinivas is also largely open.

As additional evidence of this correspondence, we consider the case when the defining equations of X are very general. Shibuta and Takagi [2009] proved the correspondence if $X = \mathbb{C}^n$ and Z is a complete intersection binomial subscheme or a space monomial curve. Using a similar idea, Hernández [2011] recently proved the case when $X = \mathbb{C}^n$ and Z is a hypersurface of X such that the coefficients of terms of its defining equation are algebraically independent over $\mathbb{Q}$. Using the techniques we have developed for Theorem 3.2, we generalize his result in Theorem 4.1 (page 935).

1. Preliminaries

Test ideals and F-singularities of pairs. In this subsection, we briefly review the definitions of test ideal sheaves and F-singularities of pairs. The reader is referred to [Schwede 2008; 2009; 2011] and [Takagi 2004a; 2010] for the details.

Throughout this paper, all schemes are Noetherian, excellent and separated, and all sheaves are coherent. Let A be an integral scheme of prime characteristic p. For each integer $e \geq 1$, we denote by $F^e : A \to A$ or $F^e : \mathcal{O}_A \to F^e_* \mathcal{O}_A$ the e-th iteration of the absolute Frobenius morphism on A. We say that A is *F-finite* if $F : A \to A$ is a

finite morphism. For example, every scheme essentially of finite type over a perfect field is *F*-finite. Given an ideal sheaf $I \subseteq \mathbb{O}_A$, for each $q = p^e$, we let $I^{[q]} \subseteq \mathbb{O}_A$ denote the ideal sheaf identified with $I \cdot F_*^e \mathbb{O}_A$ via the identification $F_*^e \mathbb{O}_A \cong \mathbb{O}_A$. For a closed subscheme Y of A, we let $\mathcal{I}_Y$ denote the defining ideal sheaf of Y in X.

The notion of test ideal sheaves along arbitrary subvarieties was introduced in [Takagi 2010]. Below we give an alternate description of these sheaves based on the ideas of [Schwede 2009]. Let A be a normal $\mathbb{Q}$-Gorenstein variety over an *F*-finite field of characteristic $p > 0$ and $X \subseteq A$ be a reduced equidimensional closed subscheme of codimension c. Suppose that the Gorenstein index of A is not divisible by p. There then exists infinitely many e such that $(p^e - 1)K_A$ is Cartier, and we fix such an integer $e_0 \geq 1$. Grothendieck duality yields an isomorphism of $F_*^{e_0} \mathbb{O}_A$-modules

$$F_*^{e_0} \mathbb{O}_A \cong \mathcal{H}om_{\mathbb{O}_A}(F_*^{e_0}((1 - p^{e_0})K_A), \mathbb{O}_A),$$

and we let

$$\varphi_{A,e_0} : F_*^{e_0} \mathbb{O}_A((1 - p^{e_0})K_A) \to \mathbb{O}_A$$

denote the map corresponding to the global section 1 of $\mathbb{O}_A$ via this isomorphism. When A is Gorenstein, we can describe φ_{A,e_0} more explicitly: it is obtained by tensoring the canonical dual $(F^{e_0})^\vee : F_*^{e_0} \omega_A \to \omega_A$ of the e_0-times iterated Frobenius morphism $F^{e_0} : \mathbb{O}_A \to F_*^{e_0} \mathbb{O}_A$ with $\mathbb{O}_A(-K_A)$. Also, the composite map

$$\varphi_{A,e_0} \circ F_*^{e_0} \varphi_{A,e_0} \circ \cdots \circ F_*^{(n-1)e_0} \varphi_{A,e_0} : F_*^{ne_0} \mathbb{O}_A((1 - p^{ne_0})K_A) \to \mathbb{O}_A$$

is denoted by φ_{A,ne_0} for all integers $n \geq 1$. Just for convenience, $\varphi_{A,0}$ is defined to be the identity map $\mathbb{O}_A \to \mathbb{O}_A$.

Proposition-Definition 1.1 (cf. [Takagi 2010, Definition 2.2]). *Let the notation be as above, and let $Z := \sum_{i=1}^m t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of A that do not contain any component of X in their support.*

(1) *There exists a unique smallest ideal sheaf $J \subseteq \mathbb{O}_A$ whose support does not contain any component of X and that satisfies*

$$\varphi_{A,ne_0}\left(F_*^{ne_0}\left(J \mathcal{I}_X^{c(p^{ne_0}-1)} \mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil} \cdots \mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil} \mathbb{O}_A((1-p^{ne_0})K_A)\right)\right) \subseteq J$$

for all integers $n \geq 1$. This ideal sheaf is denoted by $\tilde{\tau}_X(A, Z)$. When $X = \varnothing$ (resp. $Z = \varnothing$), we write simply $\tilde{\tau}(A; Z)$ (resp. $\tilde{\tau}_X(A)$).

(2) *(A, Z) is said to be* purely *F*-regular *along X if $\tilde{\tau}_X(A, Z) = \mathbb{O}_A$.*

Proof. We will prove that $\tilde{\tau}_X(A, Z)$ always exists. First, we suppose that A is affine, $\mathbb{O}_A((1 - p^{ne_0})K_A) \cong \mathbb{O}_A$ and $\mathrm{Hom}_{\mathbb{O}_A}(F_*^{e_0}\mathbb{O}_A, \mathbb{O}_A)$ is generated by φ_{A,e_0} as an $F_*^{e_0}\mathbb{O}_A$-module. Then $\mathrm{Hom}_{\mathbb{O}_A}(F_*^{ne_0}\mathbb{O}_A, \mathbb{O}_A)$ is generated by φ_{A,ne_0} as an $F_*^{ne_0}\mathbb{O}_A$-module for all $n \geq 1$. Here we use the following fact:

Claim. *There exists an element* $\gamma \in \mathbb{O}_A$ *not contained in any minimal prime ideal of* $\mathcal{I}_X$ *and satisfying the following property*: *for every* $\delta \in \mathbb{O}_A$ *not contained in any minimal prime of* $\mathcal{I}_X$, *there exists an integer* $n \geq 1$ *such that*

$$\gamma \in \varphi_{A,ne_0}\big(F_*^{ne_0}(\delta\mathcal{I}_X^{c(p^{ne_0}-1)}\mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil}\cdots\mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil})\big).$$

Proof. Suppose that $g \in \bigcap_i \mathcal{I}_{Z_i}$ is an element not contained in any minimal prime of $\mathcal{I}_X$ such that $D(g)|_X \subseteq X$ is regular. By [Takagi 2010, Example 2.6], $D(g)$ is purely F-regular along $D(g)|_X$. It then follows from an argument similar to [Schwede 2011, Proposition 3.21] that some power of g satisfies the condition of the claim. $\square$

Let $\gamma \in \mathbb{O}_A$ be an element satisfying the conditions of the above claim. Then we will show that

$$\tilde{\tau}_X(A, Z) = \sum_{n \geq 0} \varphi_{A,ne_0}\big(F_*^{ne_0}(\gamma\mathcal{I}_X^{c(p^{ne_0}-1)}\mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil}\cdots\mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil})\big).$$

It is easy to check that $\sum_{n \geq 0} \varphi_{A,ne_0}\big(F_*^{ne_0}(\gamma\mathcal{I}_X^{c(p^{ne_0}-1)}\mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil}\cdots\mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil})\big)$ is the smallest ideal $J \subseteq \mathbb{O}_A$ containing γ and satisfying

$$\varphi_{A,ne_0}\big(F_*^{ne_0}(J\mathcal{I}_X^{c(p^{ne_0}-1)}\mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil}\cdots\mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil})\big) \subseteq J$$

for all $n \geq 1$. On the other hand, if an ideal $I \subseteq \mathbb{O}_A$ is not contained in any minimal prime of $\mathcal{I}_X$ and satisfying

$$\varphi_{A,ne_0}\big(F_*^{ne_0}(I\mathcal{I}_X^{c(p^{ne_0}-1)}\mathcal{I}_{Z_1}^{\lceil t_1(p^{ne_0}-1)\rceil}\cdots\mathcal{I}_{Z_m}^{\lceil t_m(p^{ne_0}-1)\rceil})\big) \subseteq I$$

for all $n \geq 1$, then γ is forced to be in I by definition. This completes the proof when A is affine and $\mathrm{Hom}_{\mathbb{O}_A}(F_*^{e_0}\mathbb{O}_A, \mathbb{O}_A)$ is generated by φ_{A,e_0} as an $F_*^{e_0}\mathbb{O}_A$-module.

In the general case, $\tilde{\tau}_X(A, Z)$ is obtained by gluing the constructions on affine charts. $\square$

Remark 1.2. The definition of $\tilde{\tau}_X(A, Z)$ is independent of the choice of e_0.

Next, we will give a definition of F-singularities of pairs and F-pure thresholds.

Definition 1.3 ([Takagi 2006, Definition 3.1; Schwede 2008, Proposition 3.3], cf. [Schwede 2008, Proposition 5.3]). Let X be an F-finite integral normal scheme of characteristic $p > 0$ and D be an effective $\mathbb{Q}$-divisor on X. Let $Z = \sum_{i=1}^m t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of X. Fix an arbitrary point $x \in X$.

(i) $((X, D); Z)$ is said to be *strongly F-regular* at x if, for every nonzero $\gamma \in \mathbb{O}_{X,x}$, there exist an integer $e \geq 1$ and $\delta \in \mathscr{I}_{Z_1,x}^{\lceil t_1(p^e-1)\rceil} \cdots \mathscr{I}_{Z_m,x}^{\lceil t_m(p^e-1)\rceil}$ such that

$$\gamma \delta F^e : \mathbb{O}_{X,x} \to F_*^e \mathbb{O}_X(\lceil (p^e-1)D\rceil)_x, \quad a \mapsto \gamma \delta a^{p^e},$$

splits as an $\mathbb{O}_{X,x}$-module homomorphism.

(ii) $((X, D); Z)$ is said to be *sharply F-pure* at x if there exist an integer $e \geq 1$ and $\delta \in \mathscr{I}_{Z_1,x}^{\lceil t_1(p^e-1)\rceil} \cdots \mathscr{I}_{Z_m,x}^{\lceil t_m(p^e-1)\rceil}$ such that

$$\delta F^e : \mathbb{O}_{X,x} \to F_*^e \mathbb{O}_X(\lceil (p^e-1)D\rceil)_x, \quad a \mapsto \delta a^{p^e},$$

splits as an $\mathbb{O}_{X,x}$-module homomorphism.

We simply say that $(X; Z)$ is strongly *F*-regular (resp. sharply *F*-pure) at x if $((X, 0); Z)$ is. We say that (X, D) is strongly *F*-regular (resp. sharply *F*-pure) if $((X, D); \varnothing)$ is. Also, we say that $((X, D); Z)$ is strongly *F*-regular (resp. sharply *F*-pure) if it is for all $x \in X$.

(iii) Suppose that (X, D) is sharply *F*-pure at x. Then the *F-pure threshold* $\mathrm{fpt}_x((X, D); Z)$ of Z at x is defined to be

$$\mathrm{fpt}_x((X, D); Z) := \sup\{\, t \in \mathbb{R}_{\geq 0} \mid ((X, D); tZ) \text{ is sharply } F\text{-pure at } x \,\}.$$

We write simply $\mathrm{fpt}_x(X; Z)$ when $D = 0$.

Remark 1.4. Let A and Z be as in Proposition-Definition 1.1. Then $(A; Z)$ is strongly *F*-regular at a point $x \in A$ if and only if $\tilde{\tau}(A, Z)_x = \mathbb{O}_{A,x}$.

There exists a criterion for sharp *F*-purity called the Fedder type criterion, which we will use later.

Lemma 1.5 [Fedder 1983, Lemma 1.6; Schwede 2008, Theorem 4.1]. *Let A be an F-finite regular integral affine scheme of characteristic $p > 0$ and $X \subseteq A$ be a reduced equidimensional closed subscheme.*

(1) *For each nonnegative integer e, the natural morphism*

$$F_*^e(\mathscr{I}_X^{[p^e]} : \mathscr{I}_X) \cdot \mathscr{H}om_{\mathbb{O}_A}(F_*^e \mathbb{O}_A, \mathbb{O}_A) \to \mathscr{H}om_{\mathbb{O}_X}(F_*^e \mathbb{O}_X, \mathbb{O}_X)$$

sending $s \cdot \varphi_A$ to $\overline{\varphi_A \circ F_*^e(\times s)}$ induces the isomorphism

$$\frac{F_*^e(\mathscr{I}_X^{[p^e]} : \mathscr{I}_X) \cdot \mathscr{H}om_{\mathbb{O}_A}(F_*^e \mathbb{O}_A, \mathbb{O}_A)}{F_*^e \mathscr{I}_X^{[p^e]} \cdot \mathscr{H}om_{\mathbb{O}_A}(F_*^e \mathbb{O}_A, \mathbb{O}_A)} \cong \mathscr{H}om_{\mathbb{O}_X}(F_*^e \mathbb{O}_X, \mathbb{O}_X).$$

(2) *Let $Z = \sum_{i=1}^m t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of A that do not contain any component of X in their support. Let $x \in X$ be an arbitrary point. Then the following conditions are equivalent*:

(a) $(X; Z|_X)$ *is sharply F-pure at x.*

(b) *There exists an integer $e_0 \geq 1$ such that*

$$(\mathscr{I}_{X,x}^{[p^{e_0}]} : \mathscr{I}_{X,x}) \mathscr{I}_{Z_1,x}^{\lceil t_1(p^{e_0}-1) \rceil} \cdots \mathscr{I}_{Z_m,x}^{\lceil t_m(p^{e_0}-1) \rceil} \not\subseteq \mathfrak{m}_{A,x}^{[p^{e_0}]},$$

which is equivalent to saying that

$$(\mathscr{I}_{X,x}^{[p^{ne_0}]} : \mathscr{I}_{X,x}) \mathscr{I}_{Z_1,x}^{\lceil t_1(p^{ne_0}-1) \rceil} \cdots \mathscr{I}_{Z_m,x}^{\lceil t_m(p^{ne_0}-1) \rceil} \not\subseteq \mathfrak{m}_{A,x}^{[p^{ne_0}]}$$

for all integers $n \geq 1$. Here, $\mathfrak{m}_{A,x} \subseteq \mathbb{O}_{A,x}$ denotes the maximal ideal of x.

We remark that (2) is an easy consequence of (1) in Lemma 1.5.

Singularities of the minimal model program. In this subsection, we recall the definitions of adjoint ideal sheaves, multiplier ideal sheaves and singularities of pairs. The reader is referred to [Lazarsfeld 2004] for basic theory of multiplier ideal sheaves and to [Eisenstein 2010; Takagi 2010] for that of adjoint ideal sheaves.

Let X be a normal variety over an algebraically closed field K of characteristic 0, and let $Z = \sum_i t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of X. A *log resolution* of the pair (X, Z) is a proper birational morphism $\pi : \widetilde{X} \to X$ with X a smooth variety such that all scheme theoretic inverse images $\pi^{-1}(Z_i)$ are divisors and in addition $\bigcup_i \operatorname{Supp} \pi^{-1}(Z_i) \cup \operatorname{Exc}(\pi)$ is a simple normal crossing divisor. The existence of log resolutions is guaranteed by the desingularization theorem of Hironaka [1964].

Definition 1.6. Let X and Z be as above, and let $D := \sum_k d_k D_k$ be a boundary divisor on X, that is, D is a $\mathbb{Q}$-divisor on X with $0 \leq d_k \leq 1$ for all k. In addition, we assume that $K_X + D$ is $\mathbb{Q}$-Cartier and no component of $\lfloor D \rfloor$ is contained in the support of the Z_i. Fix a log resolution $\pi : \widetilde{X} \to X$ of $(X, D+Z)$ such that $\pi_*^{-1}\lfloor D \rfloor$ is smooth. Then the *adjoint ideal sheaf* $\operatorname{adj}_D(X, Z)$ of (X, Z) along D is defined to be

$$\operatorname{adj}_D(X, Z) := \pi_* \mathbb{O}_{\widetilde{X}}\left(\left\lceil K_{\widetilde{X}} - \pi^*(K_X + D) - \sum_i t_i \pi^{-1}(Z_i)\right\rceil + \pi_*^{-1}\lfloor D \rfloor\right) \subseteq \mathbb{O}_X.$$

When $D = 0$, we denote this ideal sheaf by $\mathscr{I}(X, Z)$ and call it the *multiplier ideal sheaf* associated to (X, Z).

Definition 1.7. Let X and Z be as above, and let D be a $\mathbb{Q}$-divisor on X such that $K_X + D$ is $\mathbb{Q}$-Cartier. Fix a log resolution $\pi : \widetilde{X} \to X$ of $(X, D + Z)$, and then we can write

$$K_{\widetilde{X}} = \pi^*(K_X + D) + \sum_i t_i \pi^{-1}(Z_i) + \sum_j a_j E_j,$$

where the a_j are real numbers and the E_j are prime divisors on $\widetilde{X}$. Fix an arbitrary point $x \in X$.

(i) $((X, D); Z)$ is said to be *klt* at x if $a_j > -1$ for all j such that $x \in \pi(E_j)$.

(ii) $((X, D); Z)$ is said to be *log canonical* at x if $a_j \geq -1$ for all j such that $x \in \pi(E_j)$.

When X is $\mathbb{Q}$-Gorenstein and $D = 0$, we simply say that $(X; Z)$ is klt (resp. log canonical) at x instead of saying that $((X, 0); Z)$ is klt (resp. log canonical) at x. When $Z = \varnothing$, we simply say that (X, D) is klt (resp. log canonical) at x instead of saying that $((X, D); \varnothing)$ is klt (resp. log canonical) at x. Also, we say that $((X, D); Z)$ is klt (resp. log canonical) if it is so for all $x \in X$.

(iii) Suppose that (X, D) is log canonical at x. Then the *log canonical threshold* $\text{lct}_x((X, D); Z)$ of Z at x is defined to be

$$\text{lct}_x((X, D); Z) := \sup\{\, t \in \mathbb{R}_{\geq 0} \mid ((X, D); tZ) \text{ is log canonical at } x \,\}.$$

We simply denote this threshold $\text{lct}_x(X; Z)$ if X is $\mathbb{Q}$-Gorenstein and $D = 0$.

When the ambient variety is smooth, we can generalize the notion of adjoint ideal sheaves to the higher codimension case. Let A be a smooth variety over an algebraically closed field of characteristic 0 and $X \subseteq A$ be a reduced equidimensional closed subscheme of codimension c.

Definition 1.8 ([Takagi 2010, Definition 1.6]; cf. [Eisenstein 2010, Definition 3.4]). Let the notation be as above. Let $Z = \sum_i t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of A that do not contain any component of X in their support.

(i) Let $f : A' \to A$ be the blow-up of A along X and E be the reduced exceptional divisor of f that dominates X. Let $g : \widetilde{A} \to A'$ be a log resolution of $(A', f^{-1}(X) + \sum_i f^{-1}(Z_i))$ so that the strict transform $g_*^{-1}E$ is smooth, and set $\pi = f \circ g$. Then the *adjoint ideal sheaf* $\text{adj}_X(A, Z)$ of the pair (A, Z) along X is defined to be

$$\text{adj}_X(A, Z) := \pi_* \mathcal{O}_{\widetilde{A}}\left(K_{\widetilde{A}/A} - c\pi^{-1}(X) - \left\lfloor \sum_i t_i \pi^{-1}(Z_i) \right\rfloor + g_*^{-1}E \right).$$

(ii) $(A; Z)$ is said to be *plt* along X if $\text{adj}_X(A, Z) = \mathcal{O}_A$.

Remark 1.9.   (1) Definitions 1.6, 1.7 and 1.8 are independent of the choice of a log resolution used to define them.

(2) Let X and Z be as in Definition 1.7, and assume that X is $\mathbb{Q}$-Gorenstein. Then $(X; Z)$ is klt at a point $x \in X$ if and only if $\mathcal{J}(X, Z)_x = \mathcal{O}_{X,x}$.

2. Reduction from characteristic 0 to characteristic *p*

In this section, we briefly review how to reduce things from characteristic 0 to characteristic $p > 0$. Our main references are [Hochster and Huneke 1999, Chapter 2; Mustaţă and Srinivas 2011, Section 3.2].

Let X be a scheme of finite type over a field K of characteristic 0 and $Z = \sum_i t_i Z_i$ be a formal combination, where the t_i are real numbers and the Z_i are proper closed subschemes of X. Choosing a suitable finitely generated $\mathbb{Z}$-subalgebra B of K, we can construct a scheme X_B of finite type over B and closed subschemes $Z_{i,B} \subsetneq X_B$ such that there exist isomorphisms

$$
\begin{array}{ccc}
X & \overset{\cong}{\longrightarrow} & X_B \times_{\operatorname{Spec} B} \operatorname{Spec} K \\
\uparrow & & \uparrow \\
Z_i & \overset{\cong}{\longrightarrow} & Z_{i,B} \times_{\operatorname{Spec} B} \operatorname{Spec} K
\end{array}
$$

Note that we can enlarge B by localizing at a single nonzero element and replacing X_B and $Z_{i,B}$ with the corresponding open subschemes. Thus, applying the generic freeness [Hochster and Huneke 1999, (2.1.4)], we may assume that X_B and the $Z_{i,B}$ are flat over $\operatorname{Spec} B$. Letting $Z_B := \sum_i t_i Z_{i,B}$, we refer to (X_B, Z_B) as a *model* of (X, Z) over B. Given a closed point $\mu \in \operatorname{Spec} B$, we let X_μ (resp. $Z_{i,\mu}$) denote the fiber of X_B (resp. $Z_{i,B}$) over μ and define $Z_\mu := \sum_i t_i Z_{i,\mu}$. Then X_μ is a scheme of finite type over the residue field $\kappa(\mu)$ of μ, which is a finite field of characteristic $p(\mu)$. If X is regular, then after possibly enlarging B, we may assume that X_B is regular. In particular, there exists a dense open subset $W \subseteq \operatorname{Spec} B$ such that X_μ is regular for all closed points $\mu \in W$. Similarly, if X is normal (resp. reduced, irreducible, locally a complete intersection, Gorenstein, $\mathbb{Q}$-Gorenstein of index r, Cohen–Macaulay), then so is X_μ for general closed points $\mu \in \operatorname{Spec} B$. Also, $\dim X = \dim X_\mu$ and $\operatorname{codim}(Z_i, X) = \operatorname{codim}(Z_{i,\mu}, X_\mu)$ for general closed points $\mu \in \operatorname{Spec} B$. In particular, if X is normal and Z is a $\mathbb{Q}$-Weil (resp. $\mathbb{Q}$-Cartier) divisor on X, then Z_μ is a $\mathbb{Q}$-Weil (resp. $\mathbb{Q}$-Cartier) divisor on X_μ for general closed points $\mu \in \operatorname{Spec} B$. If K_X is a canonical divisor on X, then $K_{X,\mu}$ gives a canonical divisor K_{X_μ} on X_μ for general closed points $\mu \in \operatorname{Spec} B$.

Given a morphism $f : X \to Y$ of schemes of finite type over K and a model (X_B, Y_B) of (X, Y) over B, after possibly enlarging B, we may assume that f is induced by a morphism $f_B : X_B \to Y_B$ of schemes of finite type over B. Given a closed point $\mu \in \operatorname{Spec} B$, we obtain a corresponding morphism $f_\mu : X_\mu \to Y_\mu$ of schemes of finite type over $\kappa(\mu)$. If f is projective (resp. finite), then so is f_μ for general closed points $\mu \in \operatorname{Spec} B$.

Definition 2.1. Let P be a property defined for a triple (X, D, Z), where X is a scheme of finite type over a finite field, D is an effective $\mathbb{Q}$-divisor on X and Z is an $\mathbb{R}_{\geq 0}$-linear combination of closed subschemes of X.

(i) $((X, D); Z)$ is said to be of P *type* if, for a model of (X, D, Z) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense open subset $W \subseteq \operatorname{Spec} B$ such that $((X_\mu, D_\mu); Z_\mu)$ satisfies P for all closed points $\mu \in W$.

(ii) $((X, D); Z)$ is said to be of *dense **P** type* if, for a model of (X, D, Z) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that $((X_\mu, D_\mu); Z_\mu)$ satisfies **P** for all $\mu \in W$.

Remark 2.2. (1) By enlarging B, $((X, D); Z)$ is of **P** type if and only if for some model over B, **P** holds for all closed points $\mu \in \operatorname{Spec} B$.

(2) When **P** is strong *F*-regularity, pure *F*-regularity or sharp *F*-purity, the above definition is independent of the choice of a model.

There exists a correspondence between adjoint ideal sheaves and test ideal sheaves.

Theorem 2.3 ([Takagi 2008, Theorem 5.3]; cf. [Hara and Yoshida 2003; Takagi 2004b]). *Let X be a normal variety over a field K of characteristic 0, and let $Z = \sum_i t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of X. Let $D = \sum_j d_j D_j$ be a boundary divisor on X such that $K_X + D$ is $\mathbb{Q}$-Cartier and no component of $\lfloor D \rfloor$ is contained in the support of the Z_i. Given any model of (X, Z, D) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense open subset $W \subseteq \operatorname{Spec} B$ such that*

$$\operatorname{adj}_D(X, Z)_\mu = \tilde{\tau}_{D_\mu}(X_\mu, Z_\mu)$$

for every closed point $\mu \in W$. In particular, $((X, D); Z)$ is klt at x if and only if it is of strongly F-regular type at x.

An analogous correspondence between log canonicity and *F*-purity, that is, the equivalence of log canonical pairs and pairs of dense sharply *F*-pure type, is largely conjectural.

Conjecture 2.4. *Let X be a normal variety over an algebraically closed field K of characteristic 0 and D be an effective $\mathbb{Q}$-divisor on X such that $K_X + D$ is $\mathbb{Q}$-Cartier. Let $Z = \sum_i t_i Z_i$ be a formal combination, where the t_i are nonnegative rational numbers and the Z_i are proper closed subschemes of X. Fix an arbitrary point $x \in X$.*

(i) *$((X, D); Z)$ is log canonical at x if and only if it is of dense sharply F-pure type at x.*

(ii) *Suppose that (X, D) is log canonical at x. Given any model of (X, D, Z, x) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that*

$$\operatorname{lct}_x((X, D); Z) = \operatorname{fpt}_{x_\mu}((X_\mu, D_\mu); Z_\mu)$$

for all $\mu \in W$.

Remark 2.5. (1) It is easy to see that (i) implies (ii) in Conjecture 2.4.

(2) If $((X, D); Z)$ is of dense sharply F-pure type at x, then by [Hara and Watanabe 2002, Theorem 3.3; Takagi 2004a, Proposition 3.8], it is log canonical at x.

Remark 2.6. Conjecture 2.4 is known to hold in the following cases (see also Theorem 4.1):

(i) X is a $\mathbb{Q}$-Gorenstein toric variety, $D = 0$ and the Z_i are monomial subschemes.

(ii) X is the affine space $\mathbb{A}_K^n$, $D = 0$ and $Z = t_1 Z_1$, where Z_1 is a binomial complete intersection subscheme or a space monomial curve (in the latter case, $n = 3$).

(iii) X is a normal surface, D is an integral effective divisor on X and $Z = \varnothing$.

(iv) X is the affine space $\mathbb{A}_K^n$, $D = 0$ and Z is a hypersurface of X such that the coefficients of terms of its defining equation are algebraically independent over $\mathbb{Q}$.

Case (i) follows from [Blickle 2004, Theorem 3], (ii) from [Shibuta and Takagi 2009, Theorem 0.1] and (iv) from [Hernández 2011, Theorem 5.16]. We explain here how to check the case (iii). If $D \neq 0$, then it follows from comparing [Hara and Watanabe 2002, Theorem 4.5] with [Kawamata 1988, Theorem 9.6]. So we consider the case when $D = 0$. By Remark 2.5, it suffices to show that a two-dimensional log canonical singularity (X, x) is of dense F-pure type. Passing to an index-1 cover, we may assume that (X, x) is Gorenstein. If it is log terminal, then by [Hara 1998, Theorem 5.2] (see also Theorem 2.3), it is of F-regular type and, in particular, of dense F-pure type. Hence, we can assume that (X, x) is not log terminal, that is, (X, x) is a cusp singularity or a simple elliptic singularity. By [Mehta and Srinivas 1991, Theorem 1.2; Watanabe 1988, Theorem 1.7], cusp singularities are of dense F-pure type. Also, by [Mehta and Srinivas 1991], a simple elliptic singularity with exceptional elliptic curve E is of dense F-pure type if and only if for a model E_B of E over a finitely generated $\mathbb{Z}$-subalgebra $B \subseteq K$, there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that E_μ is ordinary for all $\mu \in W$. Applying the same argument as the proof of [Mustață and Srinivas 2011, Proposition 5.3], we may assume that E is defined over $\overline{\mathbb{Q}}$. It then follows from the ordinary reduction theorem of Serre [1966] that such W always exists. Thus, simple elliptic singularities are of dense F-pure type.

Lemma 2.7. *In order to prove Conjecture 2.4, it is enough to consider the case when $Z = \varnothing$.*

Proof. Since the question is local, we work in a sufficiently small neighborhood of x. By Remark 2.5, it suffices to show that if $((X, D); Z)$ is log canonical, then it is of dense sharply F-pure type.

Suppose that $((X, D); Z)$ is log canonical. Let $h_{i,1}, \ldots, h_{i,m_i}$ be a system of generators for $\mathscr{I}_{Z_i}$ for each i. Let $g_{i,1}, \ldots, g_{i,m_i}$ be general linear combinations of $h_{i,1}, \ldots, h_{i,m_i}$ with coefficients in K, and set $g_i := \prod_{j=1}^{m_i} g_{i,j}$ so that

$$\left(X, D + \sum_i \frac{t_i}{m_i} \operatorname{div}_X(g_i) \right) \tag{$\dagger$}$$

is log canonical. On the other hand, since $g_i \in \mathscr{I}_{Z_i}^{m_i}$, if ($\dagger$) is of dense sharply *F*-pure type, then so is $((X, D); Z)$. Therefore, it is enough to show that the log canonical pair ($\dagger$) is of dense sharply *F*-pure type. $\qquad\square$

Mustață and Srinivas [2011] recently proposed the following more arithmetic conjecture and related it to another conjecture on a comparison between multiplier ideal sheaves and test ideal sheaves:

Conjecture 2.8 [Mustață and Srinivas 2011, Conjecture 1.1]. *Let X be an n-dimensional smooth projective variety over $\overline{\mathbb{Q}}$. Given a model of X over a finitely generated $\mathbb{Z}$-subalgebra B of $\overline{\mathbb{Q}}$, there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that the action induced by Frobenius on $H^n(X_\mu, \mathcal{O}_{X_\mu})$ is bijective for all $\mu \in W$.*

Remark 2.9. Conjecture 2.8 is known to be true when X is a smooth projective curve of genus less than or equal to 2 (see [Mustață and Srinivas 2011, Example 5.5], which can be traced back to [Ogus 1982; Serre 1966]) or a smooth projective surface of Kodaira dimension 0; see [Jang 2011, Proposition 2.3].

Example 2.10. We check that Conjecture 2.8 holds for the Fermat hypersurface X of degree d in $\mathbb{P}_K^n$ over a field K of characteristic 0. Given a prime number p, set $S_p := \mathbb{F}_p[x_0, \ldots, x_n]$, $\mathfrak{m}_p := (x_0, \ldots, x_n) \subseteq S_p$, $f_p := x_0^d + \cdots + x_n^d \in S_p$ and $X_p := \operatorname{Proj} S_p/f_p$. Since $H^{n-1}(X_p, \mathcal{O}_{X_p}) = 0$ for almost all p when $d \leq n$, we consider the case when $d \geq n + 1$. Note that

$$H^{n-1}(X_p, \mathcal{O}_{X_p}) \cong \left[H_{\mathfrak{m}_p}^n(S_p/f_p) \right]_0 \cong \left[(0 : f_p)_{H_{\mathfrak{m}_p}^{n+1}(S_p)} \right]_{-d}.$$

Via this isomorphism, the action induced by Frobenius on $H^{n-1}(X_p, \mathcal{O}_{X_p})$ is identified with

$$f_p^{p-1} F : \left[(0 : f_p)_{H_{\mathfrak{m}_p}^{n+1}(S_p)} \right]_{-d} \to \left[(0 : f_p)_{H_{\mathfrak{m}_p}^{n+1}(S_p)} \right]_{-d},$$

where $F : H_{\mathfrak{m}_p}^{n+1}(S_p) \to H_{\mathfrak{m}_p}^{n+1}(S_p)$ is the map induced by Frobenius on $H_{\mathfrak{m}_p}^{n+1}(S_p)$. Let

$$\xi := [z/(x_0 \cdots x_n)^m] \in H_{\mathfrak{m}_p}^{n+1}(S_p)$$

be a homogeneous element such that $f_p^{p-1} F(\xi)$ vanishes; that is, such that $f_p^{p-1} z^p$ lies in $(x_0^{mp}, \ldots, x_n^{mp})$. Set $W := \{ p \in \operatorname{Spec} \mathbb{Z} \mid p \equiv 1 \bmod d \}$, which is a dense

subset of Spec $\mathbb{Z}$, and suppose that $p \in W$. Let $a_0, \ldots, a_n$ be nonnegative integers such that $\sum_{i=0}^{n} a_n = d - n - 1$. Then the term

$$\left(x_0^{d(a_0+1)} \cdots x_n^{d(a_n+1)}\right)^{(p-1)/d} = (x_0^{a_0} \cdots x_n^{a_n})^p x_0^{p-a_0-1} \cdots x_n^{p-a_n-1}$$

appears in the expansion of f_p^{p-1}. Since $\{x_1^{i_1} \cdots x_n^{i_n}\}_{0 \leq i_1, \ldots, i_n \leq p-1}$ is a free basis of S_p as an S_p^p-module, $f^{p-1} z^p$ can be written as

$$f^{p-1} z^p = u(x_0^{a_0} \cdots x_n^{a_n} z)^p x_0^{p-a_0-1} \cdots x_n^{p-a_n-1} + \sum_{i_j \neq p-a_j-1} g_{i_0, \ldots, i_n}^p x_0^{i_0} \cdots x_n^{i_n},$$

where $u \in \mathbb{F}_p$ is a nonzero element and $g_{i_0, \ldots, i_n} \in S_p$ for each $0 \leq i_0, \ldots, i_n \leq p-1$. Let $\varphi : F_* S_p \to S_p$ be the S-linear map sending $x_0^{p-a_0-1} \cdots x_n^{p-a_n-1}$ to 1 and the other part of the basis to 0. Then

$$u x_0^{a_0} \cdots x_n^{a_n} z = \varphi(f^{p-1} z^p) \in \varphi((x_0^{mp}, \ldots, x_n^{mp})) \subseteq (x_0^m, \ldots, x_n^m).$$

By the definition of the a_i, one has $\mathfrak{m}_p^{d-n-1} z \subseteq (x_0^m, \ldots, x_n^m)$, that is, $\mathfrak{m}_p^{d-n-1} \xi = 0$ in $H_{\mathfrak{m}_p}^{n+1}(S_p)$. This means that $\deg \xi \geq -d + 1$, and we conclude that, for all $p \in W$, $f_p^{p-1} F : [(0 : f_p)_{H_{\mathfrak{m}_p}^{n+1}(S_p)}]_{-d} \to [(0 : f_p)_{H_{\mathfrak{m}_p}^{n+1}(S_p)}]_{-d}$ is injective.

The following result comes from a discussion with Karl Schwede, whom the author thanks:

Theorem 2.11. *If Conjecture 2.8 holds, then Conjecture 2.4 holds as well.*

To prove it, we use a notion of sharp F-purity for noneffective integral divisors.

Definition 2.12. Let X be an F-finite normal integral scheme of characteristic $p > 0$ and D be a (not necessarily effective) integral divisor on X. We assume that $K_X + D$ is $\mathbb{Q}$-Cartier with index not divisible by p. Let $x \in X$ be an arbitrary point. We decompose D as $D = D_+ - D_-$, where D_+ and D_- are effective integral divisors on X that have no common irreducible components. We then say that the pair (X, D) is *sharply F-pure* at x if there exists an integer $e_0 > 0$ such that $(p^{e_0} - 1)(K_X + D)$ is Cartier and that for all positive multiples $e = n e_0$ of e_0, one has an $\mathcal{O}_{X,x}$-linear map $\varphi : F_*^e \mathcal{O}_X((p^e - 1)D_+ + D_-)_x \to \mathcal{O}_X(D_-)_x$ whose image of $F_*^e \mathcal{O}_X(D_-)_x$ contains 1. We say that (X, D) is sharply F-pure if it is sharply F-pure at every closed point of X.

If D is an effective integral divisor, this definition coincides with Definition 1.3(ii). We need a variant of [Schwede and Tucker 2012, Theorem 6.28] involving sharp F-purity in the sense of Definition 2.12.

Lemma 2.13 (cf. [Schwede and Tucker 2012, Theorem 6.28]). *Let $\pi : Y \to X$ be a finite separable morphism of F-finite normal integral schemes of characteristic $p > 0$. Let Δ_X be an effective $\mathbb{Q}$-divisor on X such that $K_X + \Delta_X$ is $\mathbb{Q}$-Cartier with index not divisible by p. Suppose that Δ_Y is an integral divisor on Y such that*

$K_Y + \Delta_Y = \pi^*(K_X + \Delta_X)$. *Also, we assume that the trace map* $\mathrm{Tr}_{Y/X} : \pi_* \mathcal{O}_Y \to \mathcal{O}_X$ *is surjective. Then* (X, Δ_X) *is sharply F-pure if and only if* (Y, Δ_Y) *is sharply F-pure in the sense of* Definition 2.12.

Proof. The statement is local on X, so we assume that $X = \mathrm{Spec}\, A$ and $Y = \mathrm{Spec}\, B$, where A is a local ring and B is a semilocal ring. There exists $e_0 \in \mathbb{N}$ such that $(p^{e_0} - 1)(K_X + \Delta)$ is Cartier. Then $\mathrm{Hom}_A(F_*^e A((p^e - 1)\Delta_X), A)$ is a free $F_*^e A$-module of rank 1 for all positive multiples $e = ne_0$ of e_0. Let $\varphi_X : F_*^e A \to A$ be its generator. We decompose Δ_Y as $\Delta_{Y,+} - \Delta_{Y,-}$, where $\Delta_{Y,+}$ and $\Delta_{Y,-}$ are effective integral divisors on Y that have no common components. Then the $F_*^e B$-module

$$\mathrm{Hom}_B\big(F_*^e B((p^e - 1)\Delta_{Y,+} + \Delta_{Y,-}), B(\Delta_{Y,-})\big) \cong F_*^e B((1 - p^e)(K_Y + \Delta_Y))$$
$$= F_*^e \pi^* A((1 - p^e)(K_X + \Delta_X))$$
$$\cong F_*^e \pi^* A = F_*^e B,$$

and we pick its generator $\varphi_Y : F_*^e B(\Delta_{Y,-}) \to B(\Delta_{Y,-})$ extending $\varphi_X : F_*^e A \to A$.

Suppose that (X, Δ_X) is sharply F-pure. By the definition of sharp F-purity, after possibly enlarging e, we have that $1 \in \mathrm{Im}\, \varphi_X \subseteq \mathrm{Im}\, \varphi_Y$, and hence, (Y, Δ_Y) is sharply F-pure.

Conversely, suppose that (Y, Δ_Y) is sharply F-pure. Making e larger if necessary, we may assume that $1 \in \mathrm{Im}\, \varphi_Y$. Note that $\Delta_Y = \pi^* \Delta_X - R$ and $R \geq \Delta_{Y,-}$, where R denotes the ramification divisor of π. Then the $F_*^e B$-module

$$\mathrm{Hom}_B\big(F_*^e B((p^e - 1)\pi^* \Delta_X + R), B(R)\big)$$
$$\cong F_*^e \pi^* A((1 - p^e)(K_X + \Delta_X))$$
$$\cong \mathrm{Hom}_B(F_*^e B((p^e - 1)\Delta_{Y,+} + \Delta_{Y,-}), B(\Delta_{Y,-})).$$

We pick its generator $\widetilde{\varphi}_Y : F_*^e B(R) \to B(R)$ extending $\varphi_Y : F_*^e B(\Delta_{Y,-}) \to B(\Delta_{Y,-})$. Since the trace map $\mathrm{Tr}_{Y/X}$ corresponds to the ramification divisor R, we have the following commutative diagram:

$$
\begin{array}{ccc}
F_*^e A & \xrightarrow{\ \varphi_X\ } & A \\
{\scriptstyle F_*^e \mathrm{Tr}_{Y/X}}\big\uparrow & & \big\uparrow{\scriptstyle \mathrm{Tr}_{Y/X}} \\
F_*^e B(R) & \xrightarrow{\ \widetilde{\varphi}_Y\ } & B(R)
\end{array}
$$

The surjectivity of the trace map $\mathrm{Tr}_{Y/X} : B \to A$ implies that

$$1 \in \mathrm{Tr}_{Y/X}(\mathrm{Im}\, \varphi_Y) \subseteq \mathrm{Tr}_{Y/X}(\mathrm{Im}\, \widetilde{\varphi}_Y) = \varphi_X(\mathrm{Im}\, F_*^e \mathrm{Tr}_{Y/X}) = \mathrm{Im}\, \varphi_X$$

because $B \subseteq \mathrm{Im}\, \varphi_Y$. Thus, (X, Δ_X) is sharply F-pure.   □

Proof of Theorem 2.11. Let the notation be as in Conjecture 2.4. By Lemma 2.7, we may assume that $K_X + D$ is Cartier and $Z = \varnothing$. Since the question is local, we

work in a sufficiently small neighborhood of x. By Remark 2.5, it suffices to show that if (X, D) is log canonical, then it is of dense sharply F-pure type.

Suppose that (X, D) is log canonical. By [Kollár and Mori 1998, Section 2.4], there exists a finite morphism $f : X' \to X$ from a normal variety X' over K such that $f^*(K_X + D)$ is Cartier. Let D' be a (not necessarily effective) integral divisor on X' such that $K_{X'} + D' = f^*(K_X + D)$. It then follows from [Kollár and Mori 1998, Proposition 5.20] that (X', D') is log canonical. We decompose D' as $D' = D'_+ - D'_-$, where D'_+ and D'_- are effective integral divisors on X that have no common components. Take a log resolution $\pi : \widetilde{X} \to X'$ of (X', D'), and let E denote the reduced divisor supported on the π-exceptional locus $\mathrm{Exc}(\pi)$. Let $(X_B, D_B, X'_B, D'_B = D'_{+,B} - D'_{-,B}, \pi_B, E_B)$ be a model of $(X, D, X', D' = D'_+ - D'_-, \pi, E)$ over a finitely generated $\mathbb{Z}$-subalgebra B of K. After possibly enlarging B, we may assume that $K_{X_\mu} + D_\mu$ is $\mathbb{Q}$-Cartier with index not divisible by the characteristic $p(\mu)$ and that the trace map $\mathrm{Tr}_{X'_\mu/X_\mu} : f_{\mu*}\mathcal{O}_{X'_\mu} \to \mathcal{O}_{X_\mu}$ is surjective for all closed points $\mu \in \mathrm{Spec}\, B$.

By virtue of [Mustaţă and Srinivas 2011, Theorem 5.10], there exists a dense subset of closed points $W \subseteq \mathrm{Spec}\, B$ such that for every integer $e \geq 1$ and every $\mu \in W$, the map

$$\pi_{\mu*}F^e_*(\mathcal{O}_{\widetilde{X}_\mu}(K_{\widetilde{X}_\mu} + \pi_{\mu*}^{-1}D'_{+,\mu} + E_\mu)) \to \pi_{\mu*}\mathcal{O}_{\widetilde{X}_\mu}(K_{\widetilde{X}_\mu} + \pi_{\mu*}^{-1}D'_{+,\mu} + E_\mu), \quad (\diamond)$$

induced by the canonical dual of the e-times iterated Frobenius map $\mathcal{O}_{\widetilde{X}_\mu} \to F^e_*\mathcal{O}_{\widetilde{X}_\mu}$, is surjective. Tensoring $(\diamond)$ with $\mathcal{O}_{X'_\mu}(-K_{X'_\mu} - D'_\mu)$, one can see that the map

$$\rho : \pi_{\mu*}F^e_*\big(\mathcal{O}_{\widetilde{X}_\mu}\big(M + (1 - p(\mu)^e)\pi_\mu^*(K_{X'_\mu} + D'_\mu)\big)\big) \to \pi_{\mu*}\mathcal{O}_{\widetilde{X}_\mu}(M)$$

is surjective, where $M = K_{\widetilde{X}_\mu} + \pi_{\mu*}^{-1}D'_{+,\mu} - \pi_\mu^*(K_{X'_\mu} + D'_\mu) + E_\mu$. Since (X', D') is log canonical, $1 \in \pi_{\mu*}\mathcal{O}_{\widetilde{X}_\mu}(M) \subseteq \mathcal{O}_{X'_\mu}(D'_{-,\mu})$. By Grothendieck duality, ρ is identified with the evaluation map

$$F^e_*\mathcal{O}_{X'_\mu}(D'_{-,\mu}) \otimes \mathcal{H}om_{\mathcal{O}_{X'_\mu}}\big(F^e_*\mathcal{O}_{X'_\mu}\big((p(\mu)^e - 1)D'_{+,\mu} + D'_{-,\mu}\big), \mathcal{O}_{X'_\mu}(D'_{-,\mu})\big)$$
$$\to \mathcal{O}_{X'_\mu}(D'_{-,\mu}).$$

The subjectivity of ρ then implies that there exists an $\mathcal{O}_{X'}$-linear map

$$\varphi_{X'} : F^e_*\mathcal{O}_{X'_\mu}((p(\mu)^e - 1)D'_{+,\mu} + D'_{-,\mu}) \to \mathcal{O}_{X'_\mu}(D'_{-,\mu})$$

such that $1 \in \varphi_{X'}(F^e_*\mathcal{O}_{X'_\mu}(D'_{-,\mu}))$. That is, (X'_μ, D'_μ) is sharply F-pure in the sense of Definition 2.12. Applying Lemma 2.13, we conclude that (X_μ, D_μ) is sharply F-pure for all $\mu \in W$. $\square$

Remark 2.14. Let Y be an $S2$, $G1$ and seminormal variety over an algebraically closed field K of characteristic 0 and Γ be an effective $\mathbb{Q}$-Weil divisorial sheaf on Y such that $K_Y + \Gamma$ is $\mathbb{Q}$-Cartier. Combining Theorem 2.11 with [Miller and

Schwede 2012, Corollary 4.4], we can conclude that if Conjecture 2.8 holds, then the pair (Y, Γ) is semilog canonical if and only if it is of dense sharply *F*-pure type.

3. Restriction theorem for adjoint ideal sheaves

In this section, building on an earlier work [Takagi 2010], we give a new proof of Eisenstein's restriction theorem for adjoint ideal sheaves using test ideal sheaves.

Definition 3.1. Let A be a smooth variety over an algebraically closed field K of characteristic 0 and $X \subseteq A$ be a normal $\mathbb{Q}$-Gorenstein closed subvariety of codimension c. Let r denote the Gorenstein index of X, that is, the smallest positive integer m such that mK_X is Cartier. Then the *l.c.i. defect ideal sheaf*[1] $J_X \subseteq \mathbb{O}_X$ is defined as follows. Since the construction is local, we may consider the germ at a closed point $x \in X \subseteq A$. We take generally a closed subscheme Y of A that contains X and is locally a complete intersection (l.c.i. for short) of codimension c. By Bertini's theorem, Y is the scheme-theoretic union of X and another variety C^Y of codimension c. Then the closed subscheme $D^Y := C^Y|_X$ of X is a Weil divisor such that rD^Y is Cartier and $\mathbb{O}_X(rK_X) = \mathbb{O}_X(-rD^Y)\omega_Y^{\otimes r}$. The l.c.i. defect ideal sheaf J_X is defined by

$$J_X := \sum_Y \mathbb{O}_X(-rD^Y),$$

where Y runs through all the general l.c.i. closed subschemes of codimension c containing X. Note that the support of J_X exactly coincides with the non-l.c.i. locus of X. In particular, $J_X = \mathbb{O}_X$ if and only if X is l.c.i. The reader is referred to [Kawakita 2008, Section 2; Ein and Mustaţă 2009, Section 9.2] for further properties of l.c.i. defect ideal sheaves.

Now we give a new proof of the theorem of Eisenstein [2010, Corollary 5.2].

Theorem 3.2. *Let A be a smooth variety over an algebraically closed field K of characteristic 0 and $Z = \sum_{i=1}^m t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the Z_i are proper closed subschemes of A. If X is a normal $\mathbb{Q}$-Gorenstein closed subvariety of A that is not contained in the support of any Z_i, then*

$$\mathcal{J}\left(X, Z|_X + \tfrac{1}{r}V(J_X)\right) = \mathrm{adj}_X(A, Z)|_X,$$

where r is the Gorenstein index of X and J_X is the l.c.i. defect ideal sheaf of X.

Proof. We refine the proof of [Takagi 2010, Theorem 3.1]. The inclusion

---

[1]We follow a construction due to Kawakita [2008], but our terminology is slightly different from his. We warn the reader that the ideal sheaf called the l.c.i. defect ideal in [Kawakita 2008] is different from our J_X. Also, Ein and Mustaţă [2009] introduced a very similar ideal, which coincides with our J_X up to integral closure.

$$\mathscr{J}(X, Z|_X + \frac{1}{r}V(J_X)) \supseteq \mathrm{adj}_X(A, Z)|_X$$

follows from a combination of [Ein and Mustaţă 2009, Remark 8.5] and [Takagi 2010, Lemma 1.7]. Hence, we will prove the converse inclusion.

We consider the germ at a closed point $x \in X \cap \bigcap_{i=1}^m Z_i \subset A$ since the question is local. Let c denote the codimension of X in A. Take generally a subscheme Y of A that contains X and is l.c.i. of codimension c, so Y is the scheme-theoretic union of X and a variety C^Y. Then $D^Y := C^Y|_X$ is a Weil divisor on X such that rD^Y is Cartier. By a general choice of Y, one has

$$\mathscr{J}\big(X, Z|_X + \tfrac{1}{r}V(J_X)\big) = \mathrm{adj}_{D^Y}(X, Z|_X) \qquad\qquad (\star)$$

(which follows from an argument similar to the claim in the proof of [Takagi 2010, Theorem 3.1]). Therefore, it is enough to show that $\mathrm{adj}_{D^Y}(X, Z|_X) \subseteq \mathrm{adj}_X(A, Z)|_X$.

By Theorem 2.3 and [Takagi 2010, Theorem 2.7], in order to prove this inclusion, it suffices to show that given any model of (A, X, Y, Z, C^Y, D^Y) over a finitely generated $\mathbb{Z}$-subalgebra B of K, one has

$$\tilde{\tau}_{D_\mu^Y}(X_\mu, Z_\mu|_{X_\mu}) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu} \qquad\qquad (\star\star)$$

for general closed points $\mu \in \mathrm{Spec}\, B$. Since μ is a general point of $\mathrm{Spec}\, B$ and the formation of test ideal sheaves commutes with localization, we may assume that $\mathbb{O}_{A_\mu}$ is an F-finite regular local ring of characteristic $p = p(\mu) > r$, $X_\mu = V(I)$ is a normal $\mathbb{Q}$-Gorenstein closed subscheme of A_μ with Gorenstein index r and $Y_\mu = V((f_1, \ldots, f_c))$ is a complete intersection closed subscheme of codimension c containing X_μ. We may assume in addition that D_μ^Y is a Weil divisor on X_μ such that rD_μ^Y is Cartier and $\mathbb{O}_{X_\mu}(rK_{X_\mu}) = \mathbb{O}_{X_\mu}(-rD_\mu^Y)\omega_{Y_\mu}^{\otimes r}$. We take a germ $g \in \mathbb{O}_{A_\mu}$ whose image $\bar{g}$ is the local equation of rD_μ^Y on $\mathbb{O}_{X_\mu}$. Let $\mathfrak{a}_i \subseteq \mathbb{O}_{A_\mu}$ be the defining ideal of $Z_{i,\mu}$ for each $i = 1, \ldots, m$. Fix an integer $e_0 \geq 1$ such that $p^{e_0} - 1$ is divisible by r, and set $q_0 := p^{e_0}$.

Claim. *For all powers $q = q_0^n$ of q_0, one has*

$$g^{(q-1)/r}(I^{[q]} : I) = (f_1 \cdots f_c)^{q-1} \text{ in } \mathbb{O}_{A_\mu}/I^{[q]}.$$

Proof. Since $q - 1$ is divisible by r,

$$\mathbb{O}_{X_\mu}((1-q)(K_{X_\mu} + D_\mu^Y)) = \mathbb{O}_{Y_\mu}((1-q)K_{Y_\mu})|_{X_\mu}$$
$$= \mathbb{O}_{A_\mu}\big((1-q)\big(K_{A_\mu} + \sum_{i=1}^c \mathrm{div}_{A_\mu}(f_i)\big)\big)\big|_{X_\mu}.$$

Set $e := ne_0$. By making use of Grothendieck duality, this implies that the natural map of $F_*^e\mathbb{O}_{A_\mu}$-modules

$$\mathrm{Hom}_{\mathbb{O}_{A_\mu}}\big(F_*^e\mathbb{O}_{A_\mu}\big((q-1)\sum_{i=1}^c \mathrm{div}_{A_\mu}(f_i)\big), \mathbb{O}_{A_\mu}\big) \to \mathrm{Hom}_{\mathbb{O}_{X_\mu}}\big(F_*^e\mathbb{O}_{X_\mu}((q-1)D_\mu^Y), \mathbb{O}_{X_\mu}\big)$$

induced by restriction is surjective. It then follows from Lemma 1.5(1) that the $\mathbb{O}_{A_\mu}$-linear map

$$(f_1 \cdots f_c)^{q-1}\mathbb{O}_{A_\mu} \to \frac{g^{(q-1)/r}(I^{[q]} : I)}{I^{[q]}}$$

induced by the natural quotient map $\mathbb{O}_{A_\mu} \to \mathbb{O}_{A_\mu}/I^{[q]}$ is surjective. Thus, we obtain the assertion. □

Let $\varphi_{X_\mu, e_0} : F_*^{ne_0}\mathbb{O}_{X_\mu} \to \mathbb{O}_{X_\mu}$ be a generator for the rank-1 free $F_*^{ne_0}\mathbb{O}_{X_\mu}$-module $\mathrm{Hom}_{\mathbb{O}_{X_\mu}}(F_*^{ne_0}\mathbb{O}_{X_\mu}, \mathbb{O}_{X_\mu})$. Then $\tilde{\tau}_{D_\mu^Y}(X_\mu, Z_\mu|_{X_\mu})$ is the unique smallest ideal J whose support does not contain any component of D_μ^Y and that satisfies

$$\varphi_{X_\mu, ne_0}\big(F_*^{ne_0}(Jg^{(q_0^n-1)/r}\mathfrak{a}_1^{\lceil t_1(q_0^n-1)\rceil} \cdots \mathfrak{a}_m^{\lceil t_m(q_0^n-1)\rceil})\big) \subseteq J$$

for all integers $n \geq 1$. By Lemma 1.5(1), there exist an $\mathbb{O}_{A_\mu}$-linear map

$$\varphi_{A_\mu, ne_0} : F_*^{ne_0}\mathbb{O}_{A_\mu} \to \mathbb{O}_{A_\mu}$$

and a germ $h_n \in \mathbb{O}_{A_\mu}$ whose image is a generator for the cyclic $\mathbb{O}_{X_\mu}$-module $(I^{[q_0^n]} : I)/I^{[q_0]}$ such that we have the commutative diagram

$$
\begin{array}{ccc}
F_*^{ne_0}\mathbb{O}_{A_\mu} & \xrightarrow{\varphi_{A_\mu, ne_0} \circ F_*^{ne_0}h_n} & \mathbb{O}_{A_\mu} \\
\downarrow & & \downarrow \\
F_*^{ne_0}\mathbb{O}_{X_\mu} & \xrightarrow{\varphi_{X_\mu, ne_0}} & \mathbb{O}_{X_\mu}
\end{array}
$$

where the vertical maps are natural quotient maps. By the definition of $\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)$, one has

$$\varphi_{A_\mu, ne_0}\big(F_*^{ne_0}(\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)I^{c(q_0^n-1)}\mathfrak{a}_1^{\lceil t_1(q_0^n-1)\rceil} \cdots \mathfrak{a}_m^{\lceil t_m(q_0^n-1)\rceil})\big) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu).$$

Since $g^{(q_0^n-1)/r}h_n \in I^{c(q_0^n-1)} + I^{[q_0^n]}$ by the claim,

$$\varphi_{A_\mu, ne_0}\big(F_*^{ne_0}(\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)g^{(q_0^n-1)/r}h_n\mathfrak{a}_1^{\lceil t_1(q_0^n-1)\rceil} \cdots \mathfrak{a}_m^{\lceil t_m(q_0^n-1)\rceil})\big) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)+I.$$

It then follows from the commutativity of the above diagram that

$$\varphi_{X_\mu, ne_0}\big(F_*^{ne_0}(\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu}\bar{g}^{(q_0^n-1)/r}\overline{\mathfrak{a}_1}^{\lceil t_1(q_0^n-1)\rceil} \cdots \overline{\mathfrak{a}_m}^{\lceil t_m(q_0^n-1)\rceil})\big)$$
$$\subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu},$$

where, for each $i = 1, \ldots, m$, $\overline{\mathfrak{a}_i}$ is the image of $\mathfrak{a}_i$ in $\mathbb{O}_{X_\mu}$.

On the other hand, note that $\mathfrak{a}_1^{\lceil t_1\rceil} \cdots \mathfrak{a}_m^{\lceil t_m\rceil}\tilde{\tau}_{X_\mu}(A_\mu) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)$. By [Takagi 2010, Example 2.6], the support of $\tilde{\tau}_{X_\mu}(A_\mu)$ is contained in the singular locus of X_μ, which does not contain any component of D_μ^Y because X_μ is normal. Also, by a general choice of Y, we may assume that no component of D_μ^Y is contained in the support of $Z_{i,\mu}$ for all $i = 1, \ldots, m$. Thus, the support of $\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu}$

does not contain any component of D_μ^Y. By the minimality of $\tilde{\tau}_{D_\mu^Y}(X_\mu, Z_\mu|_{X_\mu})$, we conclude that $\tilde{\tau}_{D_\mu^Y}(X_\mu, Z_\mu|_{X_\mu}) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu}$. □

Remark 3.3. Let the notation be as in Theorem 3.2, and fix an arbitrary point $x \in X$. Employing the same strategy as the proof of [Kawakita 2007, Theorem], we can use Theorem 3.2 to prove that the pair $(X; Z|_X + \frac{1}{r}V(J_X))$ is log canonical at x if and only if $(A; cX + Z)$ is. This result is a special case of [Kawakita 2008, Theorem 1.1; Ein and Mustață 2009, Theorem 1.1], but our proof does not depend on the theory of jet schemes.

As a corollary, we prove the conjecture proposed in [Takagi 2010, Conjecture 2.8] when X is normal and $\mathbb{Q}$-Gorenstein.

Corollary 3.4. *Let A be a smooth variety over an algebraically closed field K of characteristic 0 and $X \subseteq A$ be a normal $\mathbb{Q}$-Gorenstein closed subvariety of A. Let $Z = \sum_{i=1}^m t_i Z_i$ be a formal combination, where the t_i are nonnegative real numbers and the $Z_i \subseteq A$ are proper closed subschemes that do not contain X in their support. Given any model of (A, X, Z) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense open subset $W \subseteq \mathrm{Spec}\, B$ such that*

$$\mathrm{adj}_X(A, Z)_\mu = \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)$$

for every closed point $\mu \in W$. In particular, the pair $(A; Z)$ is plt along X if and only if it is of purely F-regular type along X.

Proof. Let r be the Gorenstein index of X and $J_X \subseteq \mathbb{O}_X$ be the l.c.i. defect ideal sheaf of X. Let $(A_B, X_B, Z_B, J_{X,B})$ be any model of (A, X, Z, J_X) over a finitely generated $\mathbb{Z}$-subalgebra B of K. By [Takagi 2010, Theorem 2.7], there exists a dense open subset $W \subseteq \mathrm{Spec}\, B$ such that

$$\tilde{\tau}_{X_\mu}(A_\mu, Z_\mu) \subseteq \mathrm{adj}_X(A, Z)_\mu$$

for all closed points $\mu \in W$. Therefore, we will prove the reverse inclusion.

As an application of Theorem 2.3 to (⋆) and (⋆⋆) in the proof of Theorem 3.2, after replacing W by a smaller dense open subset if necessary, we may assume that

$$\mathrm{adj}_X(A, Z)_\mu|_{X_\mu} = \mathcal{J}\big(X, Z|_X + \tfrac{1}{r}V(J_X)\big)_\mu \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)|_{X_\mu},$$

that is,

$$\mathrm{adj}_X(A, Z)_\mu \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu) + \mathcal{I}_{X_\mu}$$

for all closed points $\mu \in W$. It, however, follows from Theorem 2.3 and [Eisenstein 2010, Theorem 5.1] that we may assume that, for all closed points $\mu \in W$,

$$\mathrm{adj}_X(A, Z)_\mu \cap \mathcal{I}_{X_\mu} = \mathcal{J}(A, cX + Z)_\mu = \tilde{\tau}(A_\mu, cX_\mu + Z_\mu) \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu).$$

Thus, $\mathrm{adj}_X(A, Z)_\mu \subseteq \tilde{\tau}_{X_\mu}(A_\mu, Z_\mu)$ for all closed points $\mu \in W$. □

4. The correspondence of log canonicity and F-purity when the defining equations are very general

Using the argument developed in the previous section and involving the l.c.i. defect ideal sheaf, we will show that Conjecture 2.4 holds true if the defining equations of the variety are very general. The following result is a generalization of a result of Hernández [2011] to the singular case:

Theorem 4.1. *Let* $\mathbb{A}_K^n := \operatorname{Spec} K[x_1, \ldots, x_n]$ *be the affine n-space over an algebraically closed field K of characteristic 0 and* $X \subseteq \mathbb{A}_K^n$ *be a normal* $\mathbb{Q}$*-Gorenstein closed subvariety of codimension c passing through the origin 0. Let r denote the Gorenstein index of X and* J_X *denote the l.c.i. defect ideal of X. Let* $\mathfrak{a} \subseteq \mathcal{O}_X$ *be a nonzero ideal and* $t > 0$ *be a real number. Suppose that there exist a system of generators* $h_1, \ldots, h_l$ *for the defining ideal* $\mathcal{I}_X$ *of X and a system of generators* $h_{l+1}, \ldots, h_v$ *for* $\mathfrak{a}$ *with the following property: for each* $i = 1, \ldots, v$, *we can write*

$$h_i = \sum_{j=1}^{\rho_i} \gamma_{ij} x_1^{\alpha_{ij}^{(1)}} \cdots x_n^{\alpha_{ij}^{(n)}} \in K[x_1, \ldots, x_n] \quad \left((\alpha_{ij}^{(1)}, \ldots, \alpha_{ij}^{(n)}) \in \mathbb{Z}_{\geq 0}^n \setminus \{\mathbf{0}\}, \gamma_{ij} \in K^* \right)$$

with $\gamma_{i1}, \ldots, \gamma_{i\rho_i}$ *algebraically independent over* $\mathbb{Q}$. *Then* $(X; tV(\mathfrak{a}) + \frac{1}{r}V(J_X))$ *is log canonical at 0 if and only if it is of dense sharply F-pure type at 0.*

Remark 4.2. By the definition of J_X, X is l.c.i. if and only if $J_X = \mathcal{O}_X$. Thus, if $X = \operatorname{Spec} K[x_1, \ldots, x_n]/(h_1, \ldots, h_l)$ is a normal complete intersection variety, $\mathfrak{a} \subseteq \mathcal{O}_X$ is the image of the ideal generated by $h_{l+1}, \ldots, h_v$ and the $h_i \in K[x_1, \ldots, x_n]$ satisfy the same property as that in Theorem 4.1, then Theorem 4.1 says that $(X, tV(\mathfrak{a}))$ is log canonical at 0 if and only if it is of dense sharply *F*-pure type.

Proof. By Remark 2.5, it suffices to show that if $\left(X; tV(\mathfrak{a}) + \frac{1}{r}V(J_X)\right)$ is log canonical at 0, then it is of dense sharply F-pure type.

Suppose that $\left(X; tV(\mathfrak{a}) + \frac{1}{r}V(J_X)\right)$ is log canonical at 0. Since the log canonical threshold $\operatorname{lct}_0\left((X; \frac{1}{r}V(J_X)); V(\mathfrak{a})\right)$ is a rational number, we may assume that t is a rational number. Take a sufficiently general complete intersection closed subscheme $Y := V((f_1, \ldots, f_c))$ of codimension c containing X, and let $s := c - l + v$ and $f_{c+j} := h_{l+j}$ for every $j = 1, \ldots, s - c$. For each $i = 1, \ldots, s$, we can write

$$f_i = \sum_{j=1}^{m_i} u_{ij} x_1^{a_{ij}^{(1)}} \cdots x_n^{a_{ij}^{(n)}} \in K[x_1, \ldots, x_n] \quad \left((a_{ij}^{(1)}, \ldots, a_{ij}^{(n)}) \in \mathbb{Z}_{\geq 0}^n \setminus \{\mathbf{0}\}, u_{ij} \in K^* \right),$$

where $u_{11}, \ldots, u_{1m_1}, \ldots, u_{s1}, \ldots, u_{sm_s}$ are algebraically independent over $\mathbb{Q}$. We decompose Y into the scheme-theoretic union of X and a variety C^Y and let D^Y denote the Weil divisor on X obtained by restricting C^Y to X. Let $g \in K[x_1, \ldots, x_n]$ be a polynomial whose image is a local equation of the Cartier divisor rD^Y in a

neighborhood of 0. Using the standard decent theory of [Hochster and Huneke 1999, Chapter 2], we can choose a model

$$\left(\mathbb{A}^n_B = \operatorname{Spec} B[x_1, \ldots, x_n], \ X_B, \ Y_B = V((f_{1,B}, \ldots, f_{c,B})), \ D^Y_B, \ \mathfrak{a}_B, \ J_{X,B}, \ g_B\right)$$

of $(\mathbb{A}^n_K, X, Y, D^Y, \mathfrak{a}, J_X, g)$ over a finitely generated $\mathbb{Z}$-subalgebra B of K satisfying these conditions, for all closed points $\mu \in \operatorname{Spec} B$:

(i) $\mathbb{Z}\left[u_{11}, \ldots, u_{1m_1}, \ldots, u_{s1}, \ldots, u_{sm_s}, 1/\left(\prod_{i,j} u_{ij}\right)\right] \subseteq B$.

(ii) The image of g_B lies in J_B.

(iii) X_μ is a normal $\mathbb{Q}$-Gorenstein closed subvariety of codimension c passing the origin 0 with Gorenstein index r.

(iv) Y_μ is a complete intersection closed subscheme of codimension c containing X_μ.

(v) $r D^Y_\mu$ is a Cartier divisor on X_μ and $\mathcal{O}_{X_\mu}(r K_{X_\mu}) = \mathcal{O}_{X_\mu}(-r D^Y_\mu) \omega^{\otimes r}_{Y_\mu}$.

(vi) The image of g_μ is a local equation of $r D^Y_\mu$ at 0.

It is then enough to show that there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that $(X_\mu; t V(\mathfrak{a}_\mu) + \frac{1}{r} V(J_{X,\mu}))$ is sharply F-pure at 0 for all $\mu \in W$.

Since $(X; t V(\mathfrak{a}) + \frac{1}{r} V(J_X))$ is log canonical at 0, it follows from [Kawakita 2008, Theorem 1.1; Ein and Mustață 2009, Theorem 1.1] (see also Remark 3.3) that $(\mathbb{A}^n_K; t V(\mathfrak{a}) + c X)$ is log canonical at 0. By a general choice of $f_1, \ldots, f_c$, it is equivalent to saying that

$$\left(\mathbb{A}^n_K; \sum_{i=1}^c \operatorname{div}(f_i) + t V(f_{c+1}, \ldots, f_s)\right)$$

is log canonical at 0. By making use of the summation formula for multiplier ideals [Takagi 2006, Theorem 3.2], for any $\epsilon > 0$, there exist nonnegative rational numbers $\lambda_{c+1}(\epsilon), \ldots, \lambda_s(\epsilon)$ with $\lambda_{c+1}(\epsilon) + \cdots + \lambda_s(\epsilon) = t(1 - \epsilon)$ such that

$$\left(\mathbb{A}^n_K; \sum_{i=1}^c (1 - \epsilon) \operatorname{div}(f_i) + \sum_{j=c+1}^s \lambda_j(\epsilon) \operatorname{div}(f_j)\right)$$

is klt at 0. Let $\mathfrak{a}_{f_i}$ be the term ideal of f_i (that is, the monomial ideal generated by the terms of f_i) for each $i = 1, \ldots, s$. Since $\mathfrak{a}_{f_i}$ contains f_i, the monomial ideal $\mathcal{J}(\mathbb{A}^n_K, \sum_{i=1}^c (1 - \epsilon) V(\mathfrak{a}_{f_i}) + \sum_{j=c+1}^s \lambda_j(\epsilon) V(\mathfrak{a}_{f_j}))$ is trivial. Then by the Main Theorem of [Howald 2001], the vector **1** lies in the interior of

$$\sum_{i=1}^c (1 - \epsilon) P(\mathfrak{a}_{f_i}) + \sum_{j=c+1}^s \lambda_j(\epsilon) P(\mathfrak{a}_{f_j}),$$

where $P(\mathfrak{a}_{f_i})$ is the Newton polyhedron of $\mathfrak{a}_{f_i}$ for each $i = 1, \ldots, s$. This is equivalent to saying that there exists

$$\sigma(\epsilon) = (\sigma_{11}(\epsilon), \ldots, \sigma_{1m_1}(\epsilon), \ldots, \sigma_{s1}(\epsilon), \ldots, \sigma_{sm_s}(\epsilon)) \in \mathbb{R}_{\geq 0}^{\sum_{i=1}^{s} m_i}$$

such that

(1) $A\sigma(\epsilon)^{\mathrm{T}} < \mathbf{1}$,

(2) $\sum_{j=1}^{m_i} \sigma_{ij}(\epsilon) = 1 - \epsilon$ for every $i = 1, \ldots, c$, and

(3) $\sum_{j=1}^{m_i} \sigma_{ij}(\epsilon) = \lambda_i(\epsilon)$ for every $i = c+1, \ldots, s$,

where A is the $n \times \sum_{i=1}^{s} m_i$ matrix

$$\begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1m_1}^{(1)} & a_{21}^{(1)} & \cdots & a_{s1}^{(1)} & \cdots & a_{sm_s}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{11}^{(n)} & \cdots & a_{1m_1}^{(n)} & a_{21}^{(n)} & \cdots & a_{s1}^{(n)} & \cdots & a_{sm_s}^{(n)} \end{pmatrix}.$$

Since such a $\sigma(\epsilon)$ exists for every $\epsilon > 0$, by the continuity of real numbers and the convexity of the solution space

$$\{\tau \in \mathbb{R}_{\geq 0}^{\sum_{i=1}^{s} m_i} \mid \tilde{A}\tau^{\mathrm{T}} \leq \mathbf{1}\},$$

there exists $\sigma = (\sigma_{11}, \ldots, \sigma_{1m_1}, \ldots, \sigma_{s1}, \ldots, \sigma_{sm_s}) \in \mathbb{Q}_{\geq 0}^{\sum_{i=1}^{s} m_i}$ such that

(1̃) $\tilde{A}\sigma^{\mathrm{T}} \leq \mathbf{1}$,

(2̃) $\sum_{j=1}^{m_i} \sigma_{ij} = 1$ for every $i = 1, \ldots, c$, and

(3̃) $\sum_{i=c+1}^{s} \sum_{j=1}^{m_i} \sigma_{ij} = t$,

where $\tilde{A}$ is the $(n+s) \times \sum_{i=1}^{s} m_i$ matrix

$$\begin{pmatrix} a_{11}^{(1)} & \cdots & a_{1m_1}^{(1)} & a_{21}^{(1)} & \cdots & a_{2m_2}^{(1)} & a_{31}^{(1)} & \cdots & a_{s1}^{(1)} & \cdots & a_{sm_s}^{(1)} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots & \ddots & \vdots \\ a_{11}^{(n)} & \cdots & a_{1m_1}^{(n)} & a_{21}^{(n)} & \cdots & a_{2m_2}^{(n)} & a_{31}^{(n)} & \cdots & a_{s1}^{(n)} & \cdots & a_{sm_s}^{(n)} \\ 1 & \cdots & 1 & 0 & \cdots & 0 & 0 & \cdots & 0 & \cdots & 0 \\ & & & 1 & \cdots & 1 & 0 & \cdots & 0 & \cdots & 0 \\ & & & & & & 1 & \cdots & 0 & \cdots & 0 \\ & & & & & & & & \vdots & & \vdots \\ & & \mathbf{0} & & & & & & 0 & \cdots & 0 \\ & & & & & & & & 1 & \cdots & 1 \end{pmatrix}. \quad (\Delta)$$

We take the least common multiple N of the denominators of the σ_{ij} so that $\sigma_{ij}(p-1)$ is an integer for all $i = 1, \ldots, s$ and all $j = 1, \ldots, m_i$ whenever $p \equiv 1 \mod N$.

Let p be a prime such that $p \equiv 1 \mod Nr$, and let $e_1, \ldots, e_n$ be nonnegative

integers such that

$$(p-1)\widetilde{A}\sigma^{\mathrm{T}} = \left(e_1, \ \cdots, \ e_n, \ \sum_{j=1}^{m_1} \sigma_{1j}(p-1), \ \cdots, \ \sum_{j=1}^{m_s} \sigma_{sj}(p-1)\right)^{\mathrm{T}}.$$

Then $e_k \leq p-1$ for all $k = 1, \ldots, n$. The coefficient of the monomial $x_1^{e_1} \cdots x_n^{e_n}$ in the expansion of

$$f_1^{\sum_{j=1}^{m_1} \sigma_{1j}(p-1)} \cdots f_s^{\sum_{j=1}^{m_s} \sigma_{sj}(p-1)}$$

is

$$\theta_{\sigma,p}(\boldsymbol{u}) := \sum_{\tau_{ij}} \prod_{i=1}^{s} \binom{\sum_{j=1}^{m_i} \sigma_{ij}(p-1)}{\tau_{i1}, \ldots, \tau_{im_i}} u_{i1}^{\tau_{i1}} \cdots u_{im_i}^{\tau_{im_i}} \in \mathbb{Z}[u_{ij}]_{\substack{i=1,\ldots,s \\ j=1,\ldots,m_i}} \subseteq B,$$

where the summation runs over all $\tau = (\tau_{11}, \ldots, \tau_{1m_1}, \ldots, \tau_{s1}, \ldots, \tau_{sm_s}) \in \mathbb{Z}_{\geq 0}^{\sum_{i=1}^{s} m_i}$ such that

$$\widetilde{A}\tau^{\mathrm{T}} = \left(e_1, \ \cdots, \ e_n, \ \sum_{j=1}^{m_1} \sigma_{1j}(p-1), \ \cdots, \ \sum_{j=1}^{m_s} \sigma_{sj}(p-1)\right)^{\mathrm{T}}.$$

Since $\widetilde{A}\sigma^{\mathrm{T}} \leq \mathbf{1}$, one has $\sum_{j=1}^{m_i} \sigma_{ij}(p-1) \leq p-1$ for all $i = 1, \ldots, s$, so the coefficient

$$\prod_{i=1}^{s} \binom{\sum_{j=1}^{m_i} \sigma_{ij}(p-1)}{\sigma_{i1}(p-1), \ldots, \sigma_{im_i}(p-1)}$$

of the monomial $\prod_{i=1}^{s} u_{i1}^{\sigma_{i1}(p-1)} \cdots u_{im_i}^{\sigma_{im_i}(p-1)}$ in $\theta_{\sigma,p}(\boldsymbol{u})$ is nonzero in $\mathbb{F}_p$. This means that $\theta_{\sigma,p}(\boldsymbol{u})$ is nonzero in $\mathbb{F}_p[u_{ij}]_{i=1,\ldots,s,\, j=1,\ldots,m_i} \subseteq B/pB$ because, by assumption, the u_{ij} are algebraically independent over $\mathbb{F}_p$, so $D(\theta_{\sigma,p}(\boldsymbol{u})) \cap \operatorname{Spec} B/pB$ is a dense open subset of $\operatorname{Spec} B/pB$.

We now set

$$W := \bigcup_{p \equiv 1 \bmod Nr} D(\theta_{\sigma,p}(\boldsymbol{u})) \cap \operatorname{Spec} B/pB \subseteq \operatorname{Spec} B.$$

Then W is a dense subset of $\operatorname{Spec} B$. Fix any closed point $\mu \in W$, and let p denote the characteristic of the residue field $\kappa(\mu) = B/\mu$ from now on. Since the image of $\theta_{\sigma,p}(\boldsymbol{u})$ is nonzero in B/μ, the monomial $x_1^{e_1} \cdots x_n^{e_n}$ appears in the expansion of

$$f_{1,\mu}^{(\sum_{j=1}^{m_1} \sigma_{1j})(p-1)} \cdots f_{s,\mu}^{(\sum_{j=1}^{m_s} \sigma_{sj})(p-1)}$$

in $(B/\mu)[x_1, \ldots, x_n]$. Since $\sum_{j=1}^{m_i} \sigma_{ij}(p-1) = p-1$ for all $i = 1, \ldots, c$ and $e_k \leq p-1$ for all $k = 1, \ldots, n$, one has

$$f_{1,\mu}^{p-1} \cdots f_{c,\mu}^{p-1} f_{c+1,\mu}^{(\sum_{j=1}^{m_{c+1}} \sigma_{c+1j})(p-1)} \cdots f_{s,\mu}^{(\sum_{j=1}^{m_s} \sigma_{sj})(p-1)} \notin (x_1^p, \ldots, x_n^p)$$

in $(B/\mu)[x_1, \ldots, x_n]_{(x_1,\ldots,x_n)}$. By Lemma 1.5(2), this is equivalent to saying that for all powers $q = p^e$ of p,

$$f_{1,\mu}^{q-1} \cdots f_{c,\mu}^{q-1} f_{c+1,\mu}^{(\sum_{j=1}^{m_{c+1}} \sigma_{c+1j})(q-1)} \cdots f_{s,\mu}^{(\sum_{j=1}^{m_s} \sigma_{sj})(q-1)} \notin (x_1^q, \ldots, x_n^q)$$

in $(B/\mu)[x_1, \ldots, x_n]_{(x_1,\ldots,x_n)}$. Applying the claim in the proof of Theorem 3.2, one has

$$(\mathscr{I}_{X,\mu}^{[q]} : \mathscr{I}_{X,\mu}) g_\mu^{(q-1)/r} f_{c+1,\mu}^{(\sum_{j=1}^{m_{c+1}} \sigma_{c+1j})(q-1)} \cdots f_{s,\mu}^{(\sum_{j=1}^{m_s} \sigma_{sj})(q-1)} \notin (x_1^q, \ldots, x_n^q)$$

in $(B/\mu)[x_1, \ldots, x_n]_{(x_1,\ldots,x_n)}$. Since $\sum_{i=c+1}^{s} \sum_{j=1}^{m_i} \sigma_{ij} = t$ and the image of g_μ lies in $J_{X,\mu}$, it follows from Lemma 1.5(2) again that the pair $(X_\mu; \frac{1}{r}V(J_{X,\mu}) + tV(\mathfrak{a}_\mu))$ is sharply *F*-pure at 0. $\qquad \square$

Remark 4.3. Using the same arguments as the proof of Theorem 4.1, we can prove the following. Let $X = \operatorname{Spec} K[x_1, \ldots, x_n]/(f_1, \ldots, f_c)$ be a normal complete intersection over a field K of characteristic 0 passing through the origin 0. Let $Z \subset X$ be a proper closed subscheme passing through 0 and $f_{c+1}, \ldots, f_s$ be a system of polynomials whose image generates the defining ideal $\mathscr{I}_Z \subseteq \mathbb{O}_X$ of Z. We write

$$f_i = \sum_{j=1}^{m_i} u_{ij} x_1^{a_{ij}^{(1)}} \cdots x_n^{a_{ij}^{(n)}} \in K[x_1, \ldots, x_n] \quad ((a_{ij}^{(1)}, \ldots, a_{ij}^{(n)}) \in \mathbb{Z}_{\geq 0}^n \setminus \{\mathbf{0}\}, u_{ij} \in K^*)$$

for each $i = 1, \ldots, s$ and set A to be the $(n + s) \times \sum_{i=1}^{s} m_i$ matrix from ($\triangle$). Then we consider the following linear programming problem:

$$\text{Maximize} \quad \sum_{i=c+1}^{s} \sum_{j=1}^{m_i} \sigma_{ij}$$

$$\text{subject to} \quad A(\sigma_{11}, \ldots, \sigma_{1m_1}, \ldots, \sigma_{s1}, \ldots, \sigma_{sm_s})^{\mathrm{T}} \leq \mathbf{1},$$

$$\sum_{i=1}^{c} \sum_{j=1}^{m_i} \sigma_{ij} = c,$$

$$\sigma_{ij} \in \mathbb{Q}_{\geq 0} \text{ for all } i = 1, \ldots, s \text{ and all } j = 1, \ldots, m_i.$$

Assume that there exists an optimal solution $\sigma = (\sigma_{11}, \ldots, \sigma_{1m_1}, \ldots, \sigma_{s1}, \ldots, \sigma_{sm_s})$ such that $A\sigma^{\mathrm{T}} \neq A\sigma'^{\mathrm{T}}$ for all other optimal solutions $\sigma' \neq \sigma$. In addition, we assume that X is log canonical at 0. Then:

(1) $\operatorname{lct}_0(X, Z)$ is equal to the optimal value $\sum_{i=c+1}^{s} \sum_{j=1}^{m_i} \sigma_{ij}$.

(2) Given any model of (X, Z) over a finitely generated $\mathbb{Z}$-subalgebra B of K, there exists a dense subset of closed points $W \subseteq \operatorname{Spec} B$ such that

$$\operatorname{lct}_0(X; Z) = \operatorname{fpt}_0(X_\mu; Z_\mu) \quad \text{for all } \mu \in W.$$

Shibuta and Takagi [2009] showed that the assumption of Remark 4.3 is satisfied if $X = \mathbb{A}_K^n$ and Z is a complete intersection binomial subscheme or a space monomial curve (in the latter case, $n = 3$). However, in general, there exists a binomial subscheme that does not satisfy the assumption.

Example 4.4. Let $X := \mathbb{A}_K^6 = \operatorname{Spec} K[x_1, x_2, x_3, y_1, y_2, y_3]$ be the affine 6-space over a field K of characteristic 0 and $Z \subseteq X$ be the closed subscheme defined by the binomials $x_1 y_2 - x_2 y_1$, $x_2 y_3 - x_3 y_2$ and $x_1 y_3 - x_3 y_1$. Then Z does not satisfy the assumption of Remark 4.3. Indeed, $\operatorname{lct}_0(X, Z) = 2$, but the optimal value of the linear programming problem in Remark 4.3 is equal to 3. Given a prime number p, let $X_p := \mathbb{A}_{\mathbb{F}_p}^6 = \operatorname{Spec} \mathbb{F}_p[x_1, x_2, x_3, y_1, y_2, y_3]$ and $Z_p \subseteq X_p$ be the reduction modulo p of Z. Since $\operatorname{fpt}_0(X_p, Z_p) = 2$ for all primes p, Conjecture 2.4 holds for this example.

Acknowledgments

References

[Blickle 2004] M. Blickle, "Multiplier ideals and modules on toric varieties", *Math. Z.* **248**:1 (2004), 113–121. MR 2006a:14082 Zbl 1061.14055

[Ein and Mustaţă 2009] L. Ein and M. Mustaţă, "Jet schemes and singularities", pp. 505–546 in *Algebraic geometry. Part 2* (Seattle, 2005), edited by D. Abramovich et al., Proc. Sympos. Pure Math. **80**, Amer. Math. Soc., Providence, RI, 2009. MR 2010h:14004 Zbl 1181.14019

[Eisenstein 2010] E. Eisenstein, "Generalization of the restriction theorem for multiplier ideals", preprint, 2010. arXiv 1001.2841

[Fedder 1983] R. Fedder, "*F*-purity and rational singularity", *Trans. Amer. Math. Soc.* **278**:2 (1983), 461–480. MR 84h:13031 Zbl 0519.13017

[Hara 1998] N. Hara, "A characterization of rational singularities in terms of injectivity of Frobenius maps", *Amer. J. Math.* **120**:5 (1998), 981–996. MR 99h:13005 Zbl 0942.13006

[Hara and Watanabe 2002] N. Hara and K.-I. Watanabe, "F-regular and F-pure rings vs. log terminal and log canonical singularities", *J. Algebraic Geom.* **11**:2 (2002), 363–392. MR 2002k:13009 Zbl 1013.13004

[Hara and Yoshida 2003] N. Hara and K.-I. Yoshida, "A generalization of tight closure and multiplier ideals", *Trans. Amer. Math. Soc.* **355**:8 (2003), 3143–3174. MR 2004i:13003 Zbl 1028.13003

[Hernández 2011] D. Hernández, "*F*-purity versus log canonicity for polynomials", preprint, 2011. arXiv 1112.2423

[Hironaka 1964] H. Hironaka, "Resolution of singularities of an algebraic variety over a field of characteristic zero. I, II", *Ann. of Math.* **79** (1964), 205–326. MR 33 #7333 Zbl 0122.38603

[Hochster and Huneke 1999] M. Hochster and C. Huneke, "Tight closure in equal characteristic zero", preprint, 1999, http://www.math.lsa.umich.edu/~hochster/tcz.ps.

[Howald 2001] J. A. Howald, "Multiplier ideals of monomial ideals", *Trans. Amer. Math. Soc.* **353**:7 (2001), 2665–2671. MR 2002b:14061 Zbl 0979.13026

[Jang 2011] J. Jang, "The ordinarity of an isotrivial elliptic fibration", *Manuscripta Math.* **134**:3-4 (2011), 343–358. MR 2012f:14014 Zbl 1225.14027

[Kawakita 2007] M. Kawakita, "Inversion of adjunction on log canonicity", *Invent. Math.* **167**:1 (2007), 129–133. MR 2008a:14025 Zbl 1114.14009

[Kawakita 2008] M. Kawakita, "On a comparison of minimal log discrepancies in terms of motivic integration", *J. Reine Angew. Math.* **620** (2008), 55–65. MR 2010i:14021 Zbl 1151.14014

[Kawamata 1988] Y. Kawamata, "Crepant blowing-up of 3-dimensional canonical singularities and its application to degenerations of surfaces", *Ann. of Math.* (2) **127**:1 (1988), 93–163. MR 89d:14023 Zbl 0651.14005

[Kollár and Mori 1998] J. Kollár and S. Mori, *Birational geometry of algebraic varieties*, Cambridge Tracts in Mathematics **134**, Cambridge University Press, 1998. MR 2000b:14018 Zbl 0926.14003

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry. II*, Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics **49**, Springer, Berlin, 2004. MR 2005k:14001b Zbl 1093.14500

[Mehta and Srinivas 1991] V. B. Mehta and V. Srinivas, "Normal *F*-pure surface singularities", *J. Algebra* **143**:1 (1991), 130–143. MR 92j:14044 Zbl 0760.14012

[Miller and Schwede 2012] L. E. Miller and K. Schwede, "Semi-log canonical vs *F*-pure singularities", *J. Algebra* **349** (2012), 150–164. MR 2853631 Zbl 1239.13010

[Mustață and Srinivas 2011] M. Mustață and V. Srinivas, "Ordinary varieties and the comparison between multiplier ideals and test ideals", *Nagoya Math. J.* **204** (2011), 125–157. MR 2863367 Zbl 1239.14011

[Ogus 1982] A. Ogus, "Hodge cycles and crystalline cohomology", pp. 357–414 in *Hodge cycles, motives, and Shimura varieties*, edited by P. Deligne et al., Lecture Notes in Mathematics **900**, Springer, Berlin, 1982. MR 84m:14046 Zbl 0538.14010

[Schwede 2008] K. Schwede, "Generalized test ideals, sharp *F*-purity, and sharp test elements", *Math. Res. Lett.* **15**:6 (2008), 1251–1261. MR 2010e:13004 Zbl 1185.13010

[Schwede 2009] K. Schwede, "*F*-adjunction", *Algebra Number Theory* **3**:8 (2009), 907–950. MR 2011b:14006 Zbl 1209.13013

[Schwede 2011] K. Schwede, "Test ideals in non-ℚ-Gorenstein rings", *Trans. Amer. Math. Soc.* **363**:11 (2011), 5925–5941. MR 2012c:13011 Zbl 05986706

[Schwede and Tucker 2012] K. Schwede and K. Tucker, "On the behavior of test ideals under finite morphisms", preprint, 2012. To appear in *J. Algebraic Geom.* arXiv 1003.4333

[Serre 1966] J.-P. Serre, "Groupes de Lie *l*-adiques attachés aux courbes elliptiques", pp. 239–256 in *Les tendances géométriques en algèbre et théorie des nombres*, Centre National de la Recherche Scientifique, Paris, 1966. MR 36 #1453 Zbl 0148.41502

[Shibuta and Takagi 2009] T. Shibuta and S. Takagi, "Log canonical thresholds of binomial ideals", *Manuscripta Math.* **130**:1 (2009), 45–61. MR 2010j:14031 Zbl 1183.13007

[Takagi 2004a] S. Takagi, "F-singularities of pairs and inversion of adjunction of arbitrary codimension", *Invent. Math.* **157**:1 (2004), 123–146. MR 2006g:14028 Zbl 1121.13008

[Takagi 2004b] S. Takagi, "An interpretation of multiplier ideals via tight closure", *J. Algebraic Geom.* **13**:2 (2004), 393–415. MR 2005c:13002 Zbl 1080.14004

[Takagi 2006] S. Takagi, "Formulas for multiplier ideals on singular varieties", *Amer. J. Math.* **128**:6 (2006), 1345–1362. MR 2007i:14006 Zbl 1109.14005

[Takagi 2008] S. Takagi, "A characteristic p analogue of plt singularities and adjoint ideals", *Math. Z.* **259**:2 (2008), 321–341. MR 2009b:13004 Zbl 1143.13007

[Takagi 2010] S. Takagi, "Adjoint ideals along closed subvarieties of higher codimension", *J. Reine Angew. Math.* **641** (2010), 145–162. MR 2011f:14032 Zbl 1193.14024

[Watanabe 1988] K. Watanabe, "Study of F-purity in dimension two", pp. 791–800 in *Algebraic geometry and commutative algebra*, vol. II, edited by H. Hijikata et al., Kinokuniya, Tokyo, 1988. MR 90b:14005 Zbl 0780.13004

stakagi@ms.u-tokyo.ac.jp   *Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan*
http://www.ms.u-tokyo.ac.jp/~stakagi/

Finitely presented exponential fields

Jonathan Kirby

We develop the algebra of exponential fields and their extensions. The focus is on ELA-fields, which are algebraically closed with a surjective exponential map. In this context, we define finitely presented extensions, show that finitely generated strong extensions are finitely presented, and classify these extensions. We give an algebraic construction of Zilber's pseudoexponential fields. As applications of the general results and methods of the paper, we show that Zilber's fields are not model-complete, answering a question of Macintyre, and we give a precise statement explaining how Schanuel's conjecture answers all transcendence questions about exponentials and logarithms. We discuss connections with the Kontsevich–Zagier, Grothendieck, and André transcendence conjectures on periods, and suggest open problems.

1. Introduction

An *exponential field* (or *E-field*) is a field F of characteristic zero equipped with a homomorphism $\exp_F$ (also written $\exp$, or $x \mapsto e^x$) from the additive group $\mathbb{G}_a(F) = \langle F; + \rangle$ to the multiplicative group $\mathbb{G}_m(F) = \langle F^\times; \cdot \rangle$. The main examples are the real and complex exponential fields, $\mathbb{R}_{\exp}$ and $\mathbb{C}_{\exp}$, where the exponential map is given by the familiar power series.

Zilber [2005] gave axioms describing particular exponential fields, which he called "pseudoexponential fields". His construction is model-theoretic and focuses mainly on the uncountable setting. In this paper we develop the algebra of exponential fields leading, amongst other things, to an algebraic construction of the pseudoexponential fields that gives some more information about them.

Some of the concepts in this paper appear also in Zilber's, but here we present them in a wider and more natural context. In particular, we do not assume that our exponential fields satisfy the Schanuel property, so much of what we do applies unconditionally to the complex setting. The main method of the paper is the use of a predimension function δ. These functions were introduced by Hrushovski [1993] for various model-theoretic constructions, but it appears they could have a significant use in transcendence theory as well.

We consider mainly those exponential fields that are algebraically closed and have a surjective exponential map, in this paper called ELA-fields. In Section 2 we show that an exponential field F, or even a field with a partially defined exponential map, can be extended in a free way to an ELA-field and, under some extra assumptions such as F being finitely generated, this free extension is unique up to isomorphism.

In Section 3 we give a definition of an extension of ELA-fields being *finitely presented*. The finite presentations take the form of algebraic varieties that are the locus of a suitable generating set. Compare the situation of a finitely generated extension of pure fields, where the extension is determined up to isomorphism by the ideal of polynomials satisfied by the generators, or equivalently by their algebraic locus over the base field. The most important extensions of exponential fields are the so-called *strong extensions*. In [Kirby 2010a] it was shown that these are the extensions that preserve the notion of exponential algebraicity. We prove that a finitely generated, kernel-preserving, strong extension of ELA-fields is a finitely presented extension. This theorem could be viewed as the analogue for exponential fields of the Hilbert basis theorem, which implies that any finitely generated extension of fields is finitely presented.

The brief Section 4 explains the convention for defining finitely presented ELA-fields (as opposed to finitely presented extensions). With this convention, it follows at once that, if Schanuel's conjecture is true, every finitely generated ELA-subfield of $\mathbb{C}_{\exp}$ is finitely presented. We can give a similar, unconditional result. By [Kirby 2010a, Theorem 1.2], we know that $\mathbb{C}_{\exp}$ is a strong extension of its countable subfield $\mathbb{C}_0$ of exponentially algebraic numbers. It is therefore an immediate consequence of Theorem 3.11 that every finitely generated ELA-extension of $\mathbb{C}_0$ within $\mathbb{C}_{\exp}$ is a finitely presented extension.

In Section 5 we show that whether or not a finitely presented extension is strong can be detected from the algebraic variety that gives the presentation, and a classification is given of all finitely generated strong ELA-extensions.

The analogue of the algebraic closure of a field is the strong exponential-algebraic closure $F^{\sim}$ of an exponential field F. Zilber's pseudoexponential fields are the simplest examples of this construction. The main claim of [Zilber 2005] was that the uncountable pseudoexponential fields are determined up to isomorphism by their cardinality. Unfortunately there is a mistake in the proof there. In the proof of [ibid., Proposition 5.15], there is no reason why $A'B'$ should not lie in C, and then V' would not contain V_0. Indeed, that proposition as stated is false, because the definition of *finitary* used there does not give sufficiently strong hypotheses. The stronger hypotheses of [ibid., Lemma 5.14] would be enough to prove the main result, but no correct proof is known to me, even with these hypotheses.[1]

---

[1] Added in proof: A proof now appears in [Bays and Kirby 2013].

In Section 6 we construct $F^{\sim}$ and, under some basic assumptions including countability, show that it is unique. In particular, we prove that the countable pseudoexponential fields are determined up to isomorphism by their exponential transcendence degree. In fact the uniqueness of the pseudoexponential field $B_{\aleph_1}$ of cardinality $\aleph_1$ then follows by Zilber's methods, as explained for example in [Kirby 2010b, Theorem 2.1], but the higher cardinalities are still problematic.

In Section 7 we answer a question of Macintyre by showing that Zilber's pseudo-exponential fields are not model-complete, and in Section 8 we show that ELA-fields satisfying the Schanuel Nullstellensatz are not necessarily strongly exponentially-algebraically closed, in contrast to the situation for pure fields where the Hilbert Nullstellensatz characterises algebraically closed fields.

In Section 9 we reflect on what the ideas of this paper show for transcendence problems, and try to give a formal statement expressing the generally accepted principle that Schanuel's conjecture answers all transcendence problems about exponentials and logarithms. We write $\mathbb{B}$ to mean a pseudoexponential field of cardinality $2^{\aleph_0}$. Zilber's conjecture is that $\mathbb{C}_{\exp} \cong \mathbb{B}$, which on the face of it makes sense only if $\mathbb{B}$ is well-defined, currently proved only under the continuum hypothesis, but the substance of the conjecture is the two assertions that Schanuel's conjecture is true and that $\mathbb{C}_{\exp}$ is strongly exponentially-algebraically closed, both of which make sense independently of the uniqueness of $\mathbb{B}$. We explore connections between Schanuel's conjecture and conjectures on periods.

Finally, in Section 10 we suggest some open problems.

2. Free extensions

As an intermediate stage in constructing exponential fields we need the notion of a *partial E-field*.

Definition 2.1. A *partial E-field* F consists of a field $\langle F; +, \cdot \rangle$ of characteristic zero, a $\mathbb{Q}$-linear subspace $D(F)$ of the additive group of the field, and a homomorphism

$$\langle D(F); + \rangle \xrightarrow{\exp_F} \langle F; \cdot \rangle.$$

$D(F)$ is the *domain* of the exponential map of F, and $I(F) = \exp_F(D(F))$ is the *image* of the exponential map.

A homomorphism of partial E-fields is a field embedding $F \xrightarrow{\theta} F_1$ such that $\theta(D(F)) \subseteq D(F_1)$ and $\exp_{F_1}(\theta(x)) = \theta(\exp_F(x))$ for every $x \in D(F)$.

If X is a subset of a partial E-field F, we define the partial E-subfield of F generated by X, written $\langle X \rangle_F$, to have $D(\langle X \rangle_F)$ equal to the $\mathbb{Q}$-span of $D(F) \cap X$, and the underlying field of $\langle X \rangle_F$ to be the subfield of F generated by $D(\langle X \rangle_F) \cup I(\langle X \rangle_F) \cup X$. Thus $\langle X \rangle_F$ contains all the exponentials in F of elements of X, but does not contain

iterated exponentials. A different but equivalent definition of partial E-fields is given in [Kirby 2010a], where $D(F)$ is given as a separate sort.

In this paper we consider only those partial E-fields F that are algebraic over $D(F) \cup I(F)$.

Now let F be a partial E-field, $\bar{x}$ a finite tuple from $D(F)$, and B a subset of $D(F)$. We define the *relative predimension function* to be

$$\delta(\bar{x}/B) = \text{td}(\bar{x}, \exp(\bar{x})/B, \exp(B)) - \text{ldim}_{\mathbb{Q}}(\bar{x}/B)$$

where $\text{td}(X/Y)$ means the transcendence degree of the field extension $\mathbb{Q}(XY)/\mathbb{Q}(Y)$ and by $\text{ldim}_{\mathbb{Q}}(X/Y)$ we mean the dimension of the $\mathbb{Q}$-vector space spanned by $X \cup Y$, quotiented by the subspace spanned by Y.

Definition 2.2. An extension $F \subseteq F_1$ of partial E-fields is *strong*, written $F \lhd F_1$, if and only if $\delta(\bar{x}/D(F)) \geqslant 0$ for every tuple $\bar{x}$ from $D(F_1)$.

If B is a subset of $D(F)$, we define $B \lhd F$ if and only if $\langle B \rangle_F \lhd F$.

As explained in [Kirby 2010a], strong extensions are essentially those for which the notion of exponential algebraicity is preserved, and are thus the most useful extensions to consider. In this paper we see they are intimately connected with free or finitely presented extensions.

The following basic properties are easy to verify.

Lemma 2.3 (basic properties of δ and strong extensions).

(1) (*Addition property.*) If $\bar{x}, \bar{y} \in D(F)$ are finite tuples and $B \subseteq D(F)$, then

$$\delta(\bar{x} \cup \bar{y}/B) = \delta(\bar{y}/B) + \delta(\bar{x}/\bar{y} \cup B).$$

(2) Given a finite tuple $\bar{x}$ from $D(F)$ and $B \subseteq D(F)$, there is a finite tuple $\bar{b}$ from B such that $\delta(\bar{x}/B) = \delta(\bar{x}/\bar{b})$.

(3) The identity $F \subseteq F$ is strong.

(4) If $F_1 \lhd F_2$ and $F_2 \lhd F_3$ then $F_1 \lhd F_3$. (That is, the composite of strong extensions is strong.)

(5) An extension $F \subseteq F_1$ is strong if and only if, the subextension $F \subseteq \langle F, \bar{x} \rangle_{F_1}$ is strong for every tuple $\bar{x}$ from F_1.

(6) If $F_1 \lhd F_2 \lhd \cdots \lhd F_n \lhd \cdots$ is an ω-chain of strong extensions then $F_1 \lhd \bigcup_{n < \omega} F_n$.

(7) If in addition each $F_n \lhd M$, then $\bigcup_{n < \omega} F_n \lhd M$. □

We now explain how exponential maps can be constructed abstractly. Let F be a field of characteristic zero, and $D(F)$ a $\mathbb{Q}$-subspace. We will construct an exponential map defined on $D(F)$.

Construction 2.4. Choose a $\mathbb{Q}$-basis $\{b_i \mid i \in I\}$ of $D(F)$. For each $i \in I$ we will choose $c_{i,1} \in F$, and we will define $\exp(b_i) = c_{i,1}$. The value of $\exp(b_i/m)$ must be an m-th root of $c_{i,1}$, so we have to specify which. Furthermore, as m varies, we must choose these roots coherently. So in fact for each $i \in I$ and $m \in \mathbb{N}$ we must choose $c_{i,m} \in F$ such that $c_{i,rm}^r = c_{i,m}$ for any $r, m \in \mathbb{N}$. Every element of $D(F)$ can be written as a finite sum $\sum r_i b_i / m$ for some $m \in \mathbb{N}$ and $r_i \in \mathbb{Z}$, and we define $\exp(\frac{1}{m} \sum r_i b_i) = \prod c_{i,m}^{r_i}$. The coherence condition shows that $\exp$ is well-defined.

This coherence property for the roots is important enough that we introduce some terminology for it.

Definition 2.5. Given c_1, a *coherent system of roots of c_1* is a sequence $(c_m)_{m \in \mathbb{N}}$ such that for every $r, m \in \mathbb{N}$ we have $c_{rm}^r = c_m$.

Of course, for the exponential map to be nontrivial we need to have some elements other than 1 (and 0) that have n-th roots for all n. In this case F will have to be infinite-dimensional as a $\mathbb{Q}$-vector space, so there will be a vast number (indeed $2^{|F|}$) of different total exponential maps that can be defined on F. Thus, for example there is no hope of classifying or understanding even all the exponential maps on $\mathbb{Q}^{\mathrm{alg}}$.

We will now explain how to construct exponential fields in as *free* a way as possible.

Construction 2.6. Let F be any partial E-field. We construct an extension F^e of F such that $D(F^e) = F$. First, embed F in a large algebraically closed field, $\mathscr{C}$. Let $\{b_i \mid i \in I\}$ be a $\mathbb{Q}$-linear basis for $F/D(F)$. Choose $\{c_{i,n} \mid i \in I, n \in \mathbb{N}\} \subseteq \mathscr{C}$ such that the $c_{i,1}$ are algebraically independent over F, and $(c_{i,n})_{n \in \mathbb{N}}$ is a coherent system of roots of $c_{i,1}$ for each i. Each $r \in F$ is a finite sum of the form $r_0 + \sum m_i b_i / n$ for some $r_0 \in D(F), n \in \mathbb{N}$, and some $m_i \in \mathbb{Z}$; we define $\exp(r_0 + \frac{1}{n} \sum m_i b_i) = \exp_F(r_0) \prod c_{i,n}^{m_i}$. Then let F^e be the subfield of $\mathscr{C}$ generated by $F \cup \{c_{i,n} \mid i \in I, n \in \mathbb{N}\}$.

A straightforward calculation shows that the isomorphism type of the extension F^e of F does not depend on the choice of $\mathscr{C}$, the choice of the b_i, or the choice of the $c_{i,n}$.

The exponential map on F^e will be a total map only when F is already a total E-field (and so $F^e = F$). However, we can iterate the construction to get a total E-field.

Construction 2.7. We write F^E for the union of the chain

$$F \hookrightarrow F^e \hookrightarrow F^{ee} \hookrightarrow F^{eee} \hookrightarrow \cdots$$

and call it the *free (total) E-field extension of F*.

We can also produce E-rings, and algebraically closed E-fields by slight variations on this method. It is convenient (albeit rather ugly) to introduce some terminology for the latter.

Definition 2.8. An *EA-field* is an E-field whose underlying field is algebraically closed.

Construction 2.9. For any partial E-field F, let F^a be the algebraic closure of F, with $D(F^a) = D(F)$.

We write F^{EA} for the union of the chain

$$F \hookrightarrow F^e \hookrightarrow F^{ea} \hookrightarrow F^{eae} \hookrightarrow F^{eaea} \hookrightarrow \cdots$$

and call it the *free EA-field extension* of F.

These constructions can intuitively be seen to be free in that at each stage there are no unnecessary algebraic or exponentially algebraic relations introduced. In the case of exponential rings (rather than fields), the analogous construction of the free E-ring extension can be seen to have the right category-theoretic universal property of a free object. In [Macintyre 1991], a universal property of the free E-field is given in terms of E-ring specialisations. The extension F^{EA} has nontrivial automorphisms over F, so cannot have a category-theoretic universal property, but later we prove uniqueness statements about these extensions making the intuitive notion of freeness precise.

Logarithms. A logarithm of an element b of an exponential field F is just some a such that $\exp(a) = b$. Of course such a logarithm will only exist if b is in the image of the exponential map, and will be defined only up to a coset of the kernel. In this algebraic setting there is no topology to make sense of a branch of the logarithm function, as in the complex case. We want to consider exponential fields, like $\mathbb{C}_{\exp}$, in which every nonzero element has a logarithm, so we extend our terminology conventions.

Definition 2.10. An *L-field* is a partial exponential field in which every nonzero element has a logarithm. An *EL-field* is a (total) exponential field in which every nonzero element has a logarithm. It is an *LA-field* or *ELA-field*, respectively, if, in addition, it is algebraically closed.

The additive group of a field of characteristic zero is just a $\mathbb{Q}$-vector space, whereas the multiplicative group has torsion, the roots of unity, so an L-field must have nontrivial kernel. The most important case is when the kernel is an infinite cyclic group.

Construction 2.11. Let $\mathbb{Q}_0$ be the partial E-field with underlying field $\mathbb{Q}$, and $D(\mathbb{Q}_0) = \{0\}$. Write $\mathbb{Q}^{\mathrm{ab}}$ for the maximal abelian extension of $\mathbb{Q}$, given by adjoining

all roots of unity. Let $\mathbb{Q}^{ab}(\tau)$ be a field extension with τ a single element, possibly in $\mathbb{Q}^{ab}$ but nonzero. Let CK_τ be the partial E-field with underlying field $\mathbb{Q}^{ab}(\tau)$, with $D(CK_\tau)$ the $\mathbb{Q}$-vector space spanned by τ and the $\exp(\tau/m)$ forming a coherent system of primitive m-th roots of unity. Then CK_τ is defined uniquely up to isomorphism by the minimal polynomial of τ over $\mathbb{Q}$. The letters "CK" stand for "cyclic kernel". In the special case where τ is transcendental, we write SK for CK_τ, meaning "standard kernel".

More generally, following Zilber we say that a partial exponential field F has *full kernel* if the image of the exponential map contains the subgroup μ of all roots of unity (so, in particular, F extends $\mathbb{Q}^{ab}$). The next proposition is implicit in [Zilber 2005] and shows that the terminology is justified because the property of F having full kernel depends only on the isomorphism type of the kernel of the exponential map as an abelian group.

Proposition 2.12. *Let F be a partial E-field extending $\mathbb{Q}^{ab}$, and let K be the kernel of its exponential map. Then the following are equivalent.*

(1) *F has full kernel.*

(2) *$\mathbb{Q}K/K \cong \mu$.*

(3) *For each $n \in \mathbb{N}^+$, K/nK is a cyclic group of order n.*

(4) *For each $n \in \mathbb{N}^+$, $|K/nK| = n$.*

(5) *$\langle K; + \rangle$ is elementarily equivalent to $\langle \mathbb{Z}; + \rangle$.*

Furthermore, if F is a field extending $\mathbb{Q}^{ab}$, and K is a subgroup of its additive group that satisfies the equivalent properties (2)–(5), then there is a partial exponential map on F with kernel K.

We give the proof for the sake of completeness.

Proof. Note that for $x \in D(F)$, we have $\exp_F(x) \in \mu$ if and only if x lies in the $\mathbb{Q}$-linear span of the kernel. Thus (1) $\Longrightarrow$ (2). But also μ has no proper self-embeddings, so (2) $\Longrightarrow$ (1).

Consider the "multiply by n map" $n : \mathbb{Q}K \to \mathbb{Q}K$. For any $x \in \mathbb{Q}K$, $\exp(x)$ lies in the n-torsion of $\mathbb{Q}K/K$ if and only if $nx \in K$, so the n-torsion group of $\mathbb{Q}K/K$ is isomorphic to $n^{-1}K/K$. Since $\mathbb{Q}K$ is divisible and torsion-free, this is isomorphic under the multiply by n map to K/nK. But the n-torsion of μ is the cyclic group of order n, so we have (2) $\Longrightarrow$ (3). In fact, μ is defined up to isomorphism by being a torsion abelian group with this n-torsion for each n, so (3) $\Longrightarrow$ (2). Clearly (3) $\Longrightarrow$ (4). For the converse, it suffices to prove it where $n = p^r$, a prime power. But then we have p^r elements of $K/p^r K$ of order dividing p^r and only p^{r-1} have order dividing p^{r-1}, and hence there is an element of order p^r, so $K/p^r K$ is cyclic of order p^r.

Property (4), together with being a torsion-free abelian group, gives a complete axiomatisation of the elementary theory of $\langle \mathbb{Z}; + \rangle$ by Szmielew's theorem [Hodges 1993, Theorem A.2.7], so (4) $\Longleftrightarrow$ (5).

For the "furthermore" statement, by property (2) there is a homomorphism from $\mathbb{Q}K$ onto μ with kernel K that makes F into a partial E-field with full kernel. $\square$

In this paper we are mainly interested in exponential fields with a surjective exponential map, so most partial E-fields we consider will have full kernel. We also assume that extensions of partial E-fields are kernel-preserving (that is, do not add new kernel elements) unless otherwise stated.

Any partial E-field F with full kernel can be extended to an ELA-field without adding new kernel elements. Indeed, we can produce free L-field, LA-field, EL-field, and ELA-field extensions of F, written F^L, F^{LA}, F^{EL}, and F^{ELA} in analogy to before.

Construction 2.13. Let F be a partial E-field with full kernel. We start by constructing a partial E-field extension F^l of F in which every element of F has a logarithm, and there are no new kernel elements. Embed F in a large algebraically closed field, $\mathcal{C}$. Inside $\mathcal{C}$ we have F^{rad}, the field extension of F obtained by adjoining all roots of all elements of F and iterating this process. The multiplicative group $(F^{\mathrm{rad}})^{\times}$ is divisible, and the image $\exp_F(D(F))$ contains the torsion and is divisible, so the quotient $(F^{\mathrm{rad}})^{\times} / \exp_F(D(F))$ is a $\mathbb{Q}$-vector space.

Choose $(b_i)_{i \in I}$ from F such that the cosets $b_i \cdot \exp_F(D(F))$ form a $\mathbb{Q}$-linear basis of $(F^{\mathrm{rad}})^{\times} / \exp_F(D(F))$. In other words, the b_i form a multiplicative basis of $(F^{\mathrm{rad}})^{\times}$ over $\exp_F(D(F))$. Now choose $(a_i)_{i \in I}$ from $\mathcal{C}$, algebraically independent over F, and for each $i \in I$, choose a coherent system of roots $(b_{i,m})_{m \in \mathbb{N}}$ of b_i.

Let $D(F^l)$ be the $\mathbb{Q}$-subspace of $\mathcal{C}$ spanned by $D(F)$ and the a_i. Define $\exp(a_i / m) = b_{i,m}$ and extend the exponential map appropriately. Let F^l be the subfield of $\mathcal{C}$ generated by $D(F^l)$ and $\exp(D(F^l))$. Then every element of F has a logarithm in F^l. The isomorphism type of F^l may depend on the choices made, but we write F^l for any resulting partial E-field.

Now we define F^{ELA} to be the union of any chain

$$F \hookrightarrow F^e \hookrightarrow F^{el} \hookrightarrow F^{ela} \hookrightarrow F^{elae} \hookrightarrow F^{elael} \hookrightarrow \cdots$$

iterating the three operations. The chain and its union are not necessarily uniquely defined because the operation $F \mapsto F^l$ is not necessarily uniquely defined. Where the union is uniquely defined we call it the *free ELA-field extension* of F. The extensions F^L, F^{LA}, and F^{EL} of F are defined in the obvious way.

Lemma 2.14. *For any partial E-field, F, the extensions $F \hookrightarrow F^e$, $F \hookrightarrow F^a$, $F \hookrightarrow F^E$ and $F \hookrightarrow F^{EA}$ are strong. If F^l, F^{ELA} are any results of Construction 2.13 then the extensions $F \hookrightarrow F^l$ and $F \hookrightarrow F^{ELA}$ are strong.*

Proof. By construction, $\delta(\bar{y}/D(F)) = 0$ for any $\bar{y}$ from $D(F^e)$. Hence $F \lhd F^e$. $F \lhd F^l$ by the same argument. It is immediate that $F \lhd F^a$ because the domain of the exponential map does not extend. The rest follows from Lemma 2.3. $\qquad\square$

In Construction 2.6 of F^e from F we made choices, but in fact the isomorphism type of F^e as an extension of F did not depend on those choices. In Construction 2.13 of F^l and F^{ELA} we again made choices, but in this case the isomorphism types of the extensions do in general depend on those choices. Before giving conditions where the extensions do not depend on the choices, so are well-defined, we illustrate the problem. Let $F = CK_\tau^a$, so $D(F)$ is spanned by τ. We want to define an extension F_1 of F in which 2 has a logarithm. So let $F_1 = F(a)$ as a field, with a transcendental over F. We define $\exp(a/m)$ to be an m-th root of 2. There is no problem in doing this, but all of these roots lie in F because it is algebraically closed, so if we make one choice of roots and produce F_1, and then make a different choice of roots and produce F_2, then F_1 and F_2 will not be isomorphic as partial E-field extensions of F. In fact these different choices will all be isomorphic as partial exponential fields and even as extensions of CK_τ. The problem is just that we had fixed all the roots of 2 in F before we defined the logarithms of 2. The way to solve the problem is to put in the logarithms earlier in the construction. In fact it is often possible to do this because of an important fact about pure fields known as the thumbtack lemma. (An explanation of the name can be found in [Baldwin 2009, p. 19].)

The thumbtack lemma was proved by Zilber [2006, Theorem 2] (with a correction to the statement and proof by Bays and Zilber [2011, Theorem 3]). We will give three versions of it in this paper as we need them. All are special cases of Zilber's theorem (cases that are not affected by the correction in the later paper), but we prefer to state exactly the form we need each time. Given an element b of a field, we write $\sqrt{b}$ for the set of all the m-th roots of b for all $m \in \mathbb{N}$.

Fact 2.15 (thumbtack lemma, version 1). *Let $F = \mathbb{Q}^{ab}(a_1, \ldots, a_r, \sqrt{b_1}, \ldots, \sqrt{b_r})$, an extension of $\mathbb{Q}^{ab}$ by finitely many generators together with all the roots of some of those generators. Now suppose that c lies in some field extension of F and is multiplicatively independent from $b_1, \ldots, b_r$. Then there is $m \in \mathbb{N}$ and an m-th root c_m of c such that there is exactly one isomorphism type of a coherent system of roots of c_m over F. That is, if F_1 and F_2 are both obtained from F by adjoining c_m and any coherent system of roots of c_m, then there is an isomorphism from F_1 to F_2 over F that sends the chosen system of roots in F_1 to the chosen system in F_2.*

Proof. This is the special case of [Zilber 2006, Theorem 2] with $n = 0$ and $l = 1$. $\square$

Note that if c is transcendental over F then the result is trivial. However, when c is algebraic over F then there is something to prove, and the condition that c is multiplicatively independent of the b_i is essential. Note also that we cannot

necessarily take $m = 1$. For example, if $F = \mathbb{Q}^{ab}$ and $c = 9$ then F certainly knows the difference between ± 3, so we must take $m \geqslant 2$. Another version of the thumbtack lemma applies to extensions of an algebraically closed field.

Fact 2.16 (thumbtack lemma, version 2). *Let $F = K(a_1, \ldots, a_r, \sqrt{b_1}, \ldots, \sqrt{b_r})$, where K is an algebraically closed field of characteristic zero. Suppose that c lies in some field extension of F and is multiplicatively independent from $K^{\times} \cdot \langle b_1, \ldots, b_r \rangle$. Then there is $m \in \mathbb{N}$ and an m-th root c_m of c such that there is exactly one isomorphism type of a coherent system of roots of c_m over F.*

Proof. This is the case $n = 1$ of Fact 3.7. See the proof there. $\qquad \square$

Definition 2.17. Let $F \subseteq F_1$ be an extension of partial E-fields. Then F_1 is *finitely generated* as an extension of F if and only if there is a finite subset $X \subseteq F_1$ such that $F_1 = \langle F \cup X \rangle_{F_1}$.

Now let F be an ELA-field, and $X \subseteq F$ a subset. We define $\langle X \rangle_F^{ELA}$ to be the smallest ELA-subfield of F that contains X. Note that it always exists, as the intersection of ELA-subfields of F is again an ELA-subfield of F.

Note also that $\langle X \rangle_F^{ELA}$ and $(\langle X \rangle_F)^{ELA}$ have different meanings. The first is the smallest ELA-subfield of F that contains X, and the second is a free ELA-field extension of the smallest partial E-subfield of F containing X, which may not be uniquely defined. In favourable circumstances (as below) the latter is well-defined and then the two ELA-fields will sometimes be isomorphic, but neither is generally true.

We now give sufficient conditions on F for F^{ELA} to be well-defined. For example, from the first case we deduce that CK_{τ}^{ELA} is well-defined. We only consider the case where F is countable here. The general case seems to be more difficult.

Theorem 2.18. *If F is a partial E-field with full kernel that is either finitely generated or a finitely generated extension of a countable LA-field, F_0, and $F \lhd K$, $F \lhd M$ are two strong extensions of F to ELA-fields that do not extend the kernel, then $\langle F \rangle_K^{ELA} \cong \langle F \rangle_M^{ELA}$ as extensions of F. In particular*:

(1) *The free ELA-closure F^{ELA} of F is well-defined.*

(2) *The extension $F \lhd K$ factors as $F \lhd F^{ELA} \lhd K$.*

Proof. Statements (1) and (2) are immediate from the main part of the theorem. For the main part, enumerate $\langle F \rangle_K^{ELA}$ as $s_1, s_2, s_3, \ldots$, such that for each $n \in \mathbb{N}$, either

(i) s_{n+1} is algebraic over $F \cup \{s_1, \ldots, s_n\}$, or

(ii) $s_{n+1} = \exp_K(a)$ for some $a \in F \cup \{s_1, \ldots, s_n\}$, or

(iii) $\exp_K(s_{n+1}) = b$ for some $b \in F \cup \{s_1, \ldots, s_n\}$.

This is possible by the definition of $\langle F \rangle_K^{ELA}$. We will inductively construct chains of partial E-subfields

$$F = K_0 \subseteq K_1 \subseteq K_2 \subseteq \cdots \text{ of } K \quad \text{and} \quad F = M_0 \subseteq M_1 \subseteq M_2 \subseteq \cdots \text{ of } M,$$

and nested isomorphisms $\theta_n : K_n \to M_n$ such that for each $n \in \mathbb{N}^+$ we have $s_n \in K_n$, $K_n \lhd K$ and $M_n \lhd M$. We also ensure that, as a pure field, each K_n has the form $F_0(\bar{\alpha}, \sqrt{\bar{\beta}})$, where F_0 is either $\mathbb{Q}^{ab}$ or a countable algebraically closed field and $\bar{\alpha}$ and $\bar{\beta}$ are finite tuples.

We start by taking θ_0 to be the identity map on F. Now assume we have K_n, M_n, and θ_n.

Case 1. s_{n+1} is algebraic over K_n (including the case where $s_{n+1} \in K_n$). Let $p(X)$ be the minimal polynomial of s_{n+1} over K_n. The image p^θ of p is an irreducible polynomial over M_n, so let t be any root of p^θ in M. Let $K_{n+1} = K_n(s_{n+1})$, $M_{n+1} = M_n(t)$, and let θ_{n+1} be the unique field isomorphism extending θ_n and sending s_{n+1} to t. We make K_{n+1} and M_{n+1} into partial exponential fields by taking the graph of exponentiation to be the graph of $\exp_K$ or $\exp_M$ intersected with K_{n+1}^2 or M_{n+1}^2, respectively. Suppose that $(a, \exp_K(a)) \in K_{n+1}^2$. Since $K_n \lhd K$, we have $\mathrm{td}(a, \exp_K(a)/K_n) - \mathrm{ldim}_\mathbb{Q}(a/D(K_n)) \geqslant 0$. But K_{n+1} is an algebraic extension of K_n, so it follows that $\mathrm{ldim}_\mathbb{Q}(a/D(K_n)) = 0$, that is, that $a \in D(K_n)$. Hence $D(K_{n+1}) = D(K_n)$. The same argument shows that $D(M_{n+1}) = D(M_n)$. Now if $\bar{x}$ is any tuple from K, we have $\delta(\bar{x}/D(K_{n+1})) = \delta(\bar{x}/D(K_n)) \geqslant 0$, and hence $K_{n+1} \lhd K$, and similarly $M_{n+1} \lhd M$. It is immediate that the pure field K_{n+1} is of the form $F_0(\bar{\alpha}, \sqrt{\bar{\beta}})$ because K_n is of that form.

Case 2. s_{n+1} is transcendental over K_n and $s_{n+1} = \exp_K(a)$ for some $a \in K_n$. Let $K_{n+1} = K_n(\sqrt{s_{n+1}})$ and $M_{n+1} = M_n(\sqrt{\exp_M(\theta_n(a))})$. Extend θ_n by defining $\theta_{n+1}(\exp_K(a/m)) = \exp_M(\theta_n(a)/m)$, and extending to a field isomorphism. This is possible because s_{n+1} is transcendental over K_n and $\exp_M(\theta_n(a))$ is transcendental over M_n (the latter because $M_n \lhd M$), and so there is a unique isomorphism type of a coherent system of roots of s_{n+1} over K_n, and of $\exp_M(\theta_n(a))$ over M_n. Then $\mathrm{td}(K_{n+1}/K_n) = 1$, $a \in D(K_{n+1}) \smallsetminus D(K_n)$, and $K_n \lhd K$, so $D(K_{n+1})$ is spanned by $D(K_n)$ and a. Similarly, $D(M_{n+1})$ is spanned by $\theta_n(a)$ over $D(M_n)$, so θ_{n+1} is an isomorphism of partial E-fields.

Now if $\bar{x}$ is any tuple from K, we have

$$\delta(\bar{x}/D(K_{n+1})) = \delta(\bar{x}, a/D(K_n)) - \delta(a/D(K_n)) = \delta(\bar{x}, a/D(K_n)) - 0 \geqslant 0$$

as $K_n \lhd K$, so $K_{n+1} \lhd K$. The same argument shows that $M_{n+1} \lhd M$. Again, it is immediate that the pure field K_{n+1} is of the required form.

Case 3. s_{n+1} is transcendental over K_n, not of the form $\exp_K(a)$ for any $a \in K_n$, but $\exp_K(s_{n+1}) = b$ for some $b \in K_n$. By hypothesis, K_n has the form $F_0(\bar{\alpha}, \sqrt{\bar{\beta}})$

for some finite tuples $\bar{\alpha}$, $\bar{\beta}$, and F_0 either $\mathbb{Q}^{ab}$ or a countable algebraically closed field. Hence, by either version 1 or version 2 of the thumbtack lemma, there is $N \in \mathbb{N}^+$ and c such that $c^N = b$ and there is a unique isomorphism type of a coherent sequence of roots of c over K_n. Let $t \in M$ be such that $\exp_M(t) = \theta_n(c)$. Let $K_{n+1} = K_n(s_{n+1}, \sqrt{c})$ and $M_{n+1} = M_n(t, \sqrt{\theta_n(c)})$. Extend θ_n by defining $\theta_{n+1}(s_{n+1}) = Nt$, and $\theta_{n+1}(\exp_K(s_{n+1}/Nm)) = \exp_M(t/m)$, and extending to a field isomorphism. This is possible by the choice of N, the fact that s_{n+1} is transcendental over K_n and (since $M_n \lhd M$) the fact that t is transcendental over M_n. As in Case 2 above, we have $K_{n+1} \lhd K$, $M_{n+1} \lhd M$, and the pure field K_{n+1} of the required form.

Conclusion. That completes the induction. Let $K_\omega = \bigcup_{n \in \mathbb{N}} K_n$. Then $K_\omega = \langle F \rangle_K^{ELA}$ because K_ω is an ELA-subfield of K containing F and is the smallest such because at each stage we add only elements of K that must lie in every ELA-subfield of K containing F. The union of the maps θ_n gives an embedding of K_ω into M, and, for the same reason, the image must be $\langle F \rangle_M^{ELA}$. Hence $\langle F \rangle_K^{ELA} \cong \langle F \rangle_M^{ELA}$ as required. $\square$

3. Finitely presented extensions

We say that a partial E-field F is *finitely generated* if there is a finite subset X of F such that $F = \langle X \rangle_F$. We restrict now to those partial E-fields F that are generated as fields by $D(F) \cup I(F)$ (call them *exponential-graph-generated*). Similarly, an ELA-field F is finitely generated as an ELA-field if $F = \langle X \rangle_F^{ELA}$ for some finite subset X of F. An extension $F \subseteq F_1$ of ELA-fields is finitely generated if and only if there is a finite subset X of F_1 such that $F_1 = \langle F \cup X \rangle_{F_1}^{ELA}$, and similarly for partial E-fields.

Let $F \subseteq F_1$ be a finitely generated extension of exponential-graph-generated partial E-fields, say generated by $a_1, \ldots, a_n \in D(F_1)$. Then the isomorphism type of the extension is given by the algebraic type of the infinite tuple $(\bar{a}, \exp(\bar{a}/m))_{m \in \mathbb{Z}}$ over F. Let

$$I(\bar{a}) = \left\{ f \in F[\bar{X}, (Y_{m,i})_{m \in \mathbb{Z}, i=1,\ldots,n}] \mid f(\bar{a}, (e^{a_i/m})) = 0 \right\}$$

and for $m \in \mathbb{N}^+$, let

$$I_m(\bar{a}) = \left\{ f \in F[\bar{X}, \bar{Y}_m, \bar{Y}_m^{-1}] \mid f(\bar{a}/m, e^{\bar{a}/m}, e^{-\bar{a}/m}) = 0 \right\}.$$

The ideal $I(\bar{a})$ contains all the *coherence polynomials* of the form $Y_{mr,i}^r - Y_{m,i}$ for each $m, r \in \mathbb{Z}$, and each $i = 1, \ldots, n$, which force each sequence $e^{a_i}, e^{a_i/2}, e^{a_i/3}, \ldots$ to be a coherent system of roots. Including the negative powers ensures that they are nonzero. The ideal $I(\bar{a})$ determines all the ideals $I_m(\bar{a})$ and if m_1 divides m_2 then $I_{m_1}(\bar{a})$ is determined by $I_{m_2}(\bar{a})$.

Definition 3.1. An ideal I of the polynomial ring $F[\bar{X}, (Y_{m,i})_{m\in\mathbb{Z}, i=1,\ldots,n}]$ is *additively free* if and only if it does not contain any polynomial of the form $\sum_{i=1}^{n} r_i X_i - c$ with the $r_i \in \mathbb{Z}$, not all zero, and $c \in D(F)$. It is *multiplicatively free* if and only if it does not contain any polynomial of the form $\prod_{i=1}^{n} Y_{1,i}^{r_i} - d$ with the $r_i \in \mathbb{Z}$, not all zero, and $d \in \exp(D(F))$. Similarly we say that I_m is additively free or multiplicatively free if it does not contain any polynomials of these forms.

If $F \subseteq F_1$ is a finitely generated extension of exponential-graph-generated partial E-fields, we may choose the generators $a_1, \ldots, a_n$ to be $\mathbb{Q}$-linearly independent over $D(F)$, and this corresponds to the ideal $I(\bar{a})$ being additively free. Conversely, if I is any prime ideal of the polynomial ring $F[\bar{X}, \bar{Y}_1, \bar{Y}_2, \bar{Y}_3, \ldots]$ that contains the coherence polynomials and is additively free, then it defines an extension F_I of F, the field of fractions of the ring $F[\bar{X}, \bar{Y}_1, \bar{Y}_2, \bar{Y}_3, \ldots]/I$, with exponentiation defined in the obvious way. All we have really done is translated Construction 2.4 into the language of ideals.

Lemma 3.2. *If I is a prime ideal containing the coherence polynomials and is additively free, then the extension F_I it defines has the same kernel as F if and only if I is multiplicatively free.*

Proof. Write a_i for the image of X_i in F_I. If I is not multiplicatively free then for some $r_i \in \mathbb{Z}$, not all zero, and some $c \in D(F)$, we have $\prod_{i=1}^{n} e^{r_i a_i} = e^c$, so $c - \sum_{i=1}^{n} r_i a_i$ lies in the kernel of $\exp_{F_I}$. Since I is additively free, this element does not lie in $D(F)$, in particular it does not lie in the kernel of $\exp_F$. Conversely, if I is multiplicatively free and $\exp\left(c + 1/m \sum_{i=1}^{n} r_i a_i\right) = 1$ with $c \in D(F)$ and $m, r_i \in \mathbb{Z}$, then $\prod_{i=1}^{n} \exp(a_i)^{r_i} = \exp(c)^m$, so $r_i = 0$ for each i, and c lies in the kernel of $\exp_F$. $\square$

Definition 3.3. We say that an extension $F \subseteq F_1$ of partial E-fields is *finitely presented* if and only if it has a finite generating set $a_1, \ldots, a_n$, which is $\mathbb{Q}$-linearly independent from $D(F)$, such that $I(\bar{a})$ is generated as an ideal by the coherence polynomials together with a finite set of other polynomials.

Definition 3.4. An additively free prime ideal J of $F(\bar{X}, \bar{Y}_1, \bar{Y}_1^{-1})$ is said to be *Kummer-generic* if and only if there is only one additively free prime ideal I of

$$F[\bar{X}, (Y_{m,i})_{m\in\mathbb{Z}, i=1,\ldots,n}],$$

containing the coherence polynomials, such that $J = I_1$, as defined above.

The term *Kummer-generic* is due to Martin Hils [2012, p. 10]. The usage here is not exactly the same as in that paper, because there they consider only adding new points to the multiplicative group, whereas here we are adding $\bar{a}$ to the additive group as well as $e^{\bar{a}}$ to the multiplicative group. The connection with Kummer theory can be seen from [Bays and Zilber 2011, Lemma 5.1].

Lemma 3.5. *If $F \subseteq F_1$ is a finitely presented extension of partial E-fields, then it has a generating set $a'_1, \ldots, a'_n$ such that the ideal $I_1(\bar{a}')$ is Kummer-generic.*

Proof. Let $g_1, \ldots, g_r \in I(\bar{a})$, together with the coherence polynomials, be a generating set for $I(\bar{a})$. Let $N \in \mathbb{N}$ be the least common multiple of the m such that some variable $Y_{m,i}$ occurs in some g_j. Then $I(\bar{a})$ is determined by $I_N(\bar{a})$. Take $\bar{a}' = \bar{a}/N$, so $I_1(\bar{a}') = I_N(\bar{a})$. Then $I_1(\bar{a}')$ is Kummer-generic, as required. $\square$

Example 3.6. Take an extension of an EA-field F generated by a_1, a_2, such that $e^{a_1/2} = a_2$, $e^{a_2} = a_1 + 1$. Then

$$I_1 = \langle Y_{1,1} = X_2^2, \, Y_{1,2} = X_1 + 1 \rangle \quad \text{and} \quad I_2 = \langle Y_{2,1} = X_2, \, Y_{2,2}^2 = X_1 + 1 \rangle.$$

In this case, I_1 is not Kummer-generic because it does not resolve whether $e^{a_1/2} = \pm a_2$.

There are finitely generated kernel-preserving extensions of some partial E-fields that are not finitely presented. However, another version of the thumbtack lemma gives conditions when this pathology does not occur.

Fact 3.7 (thumbtack lemma, version 3). *Let $F = K(a_1, \ldots, a_r, \sqrt{b_1}, \ldots, \sqrt{b_r})$, where K is an algebraically closed field of characteristic zero. Suppose that $c_1, \ldots, c_n$ lie in some field extension of F and are multiplicatively independent from $K^\times \cup \{b_1, \ldots, b_r\}$. Then there is $N \in \mathbb{N}$ and N-th roots c'_i of c_i such that there is exactly one isomorphism type over F of an n-tuple of coherent systems of roots of the (c'_i).*

Proof. It is enough to prove it in the case where K has finite transcendence degree, since if two tuples of coherent systems of roots are not isomorphic over F then that will be witnessed over a finite transcendence degree subfield. We show how it follows from [Zilber 2006, Theorem 2] in this case. Let P be the field $\mathbb{Q}(\bar{a})$ with a transcendence base of K adjoined. We take our $\bar{b}$ as the $\bar{a}$ from Zilber's theorem and our $\bar{c}$ as Zilber's $\bar{b}$. Then we take $L_1 = K$, and apply Zilber's theorem with $n = 1$. $\square$

As an immediate corollary, we have:

Corollary 3.8. *If F is an LA-field, F_1 is a finitely generated partial E-field extension of F, and F_2 is a finitely generated partial E-field extension of F_1, which does not extend the kernel, then F_2 is a finitely presented extension of F_1. In particular, every finitely generated kernel-preserving partial E-field extension of an ELA-field is finitely presented.*

Our main interest is not with partial E-fields, but with ELA-fields.

Definition 3.9. A finitely generated extension $F \subseteq F_1$ of countable ELA-fields is said to be *finitely presented* if and only if there is a finite generating set $\bar{a}$ such that,

taking $K = \langle F, \bar{a} \rangle_{F_1}$, the partial E-field extension of F generated by $\bar{a}$, we have $F_1 \cong K^{ELA}$.

Note that K^{ELA} is well-defined by Theorem 2.18. From Construction 2.4 it is clear that most finitely generated extensions of ELA-fields are not finitely presented. Indeed there are only countably many finitely presented extensions of a given countable ELA-field, but $2^{\aleph_0}$ finitely generated extensions.

We introduce a notation for finitely presented extensions. Since these are given by Kummer-generic ideals I_1, which are ideals in a polynomial ring with finitely many indeterminates, we can consider instead their associated varieties as subvarieties of $(\mathbb{G}_a \times \mathbb{G}_m)^n$.

Definition 3.10. Let F be an ELA-field. An irreducible subvariety V of $(\mathbb{G}_a \times \mathbb{G}_m)^n$ defined over F is said to be *additively free, multiplicatively free, and Kummer-generic* if and only if the corresponding ideal $I(V)$ is.

Suppose that V satisfies all three conditions. Then there is a uniquely determined partial E-field extension K of F that is generated by a tuple $(\bar{a}, e^{\bar{a}})$ that is generic in V over F. We write $F \mid V$, read "F extended by V", for the ELA-extension K^{ELA} of F.

Theorem 3.11. *Let $F \lhd F_1$ be a finitely generated kernel-preserving strong extension of ELA-fields. Then F_1 is a finitely presented extension of F.*

Proof. Let $\bar{a}$ be a finite set of generators of F_1 over F. By extending $\bar{a}$ if necessary, we may assume that $\langle F, \bar{a} \rangle_{F_1} \lhd F_1$. By Corollary 3.8, the extension $F \subseteq \langle F, \bar{a} \rangle_{F_1}$ is a finitely presented extension of partial E-fields. By Theorem 2.18, $F_1 \cong (\langle F, \bar{a} \rangle_{F_1})^{ELA}$, so is a finitely presented ELA-extension of F. $\qquad\square$

Note that there are finitely generated strong extensions of partial E-fields, of E-fields, of EA-fields, and of EL-fields that are not finitely presented, due to the issue of uniqueness of coherent systems of roots. This is the main reason why we work with ELA-fields. It is also important that the kernel does not extend, since if a is a new kernel element then the values of $\exp(a/m)$ for $m \in \mathbb{N}^+$ cannot all be specified by a finite list of equations.

4. Finitely presented ELA-fields

So far we have defined finitely presented *extensions* of ELA-fields, but it is natural also to ask about finitely presented ELA-fields. The useful convention is as follows.

Definition 4.1. An ELA-field F is said to be *finitely presented* if and only if there is a finitely generated partial E-field F_0 (with full kernel) such that $F = F_0^{ELA}$.

Note that a finitely presented ELA-extension of a finitely presented ELA-field is still finitely presented, since if $F = F_0^{ELA}$, $V \subseteq (\mathbb{G}_a \times \mathbb{G}_m)^n$ is defined over F,

additively free, multiplicatively free, and Kummer-generic, and $(\bar{a}, e^{\bar{a}}) \in V$ generates the extension $F \subseteq F \mid V$ and $F_1 = \langle F_0 \cup \bar{a} \rangle_{F \mid V}$, then $F \mid V \cong F_1^{ELA}$.

The definition is just a convention since there is no way to specify any partial E-field with full kernel within the category of partial E-fields, just by finitely many equations between a given set of generators. Within the subcategory of partial E-fields with full kernel, one might view the CK_τ as finitely presented, with explicit finite presentations

$$\exp(\tau/2) = -1, \qquad f(\tau) = 0,$$

where f is the minimal polynomial of the cyclic generator τ. However it does not follow that $\exp(\tau/m)$ is a primitive m-th root of 1 for each m and this cannot be specified by finitely many polynomial equations, for example τ/p could be the cyclic generator for any odd prime p. On the other hand, within the category of partial E-fields with cyclic kernel and named generator τ, the minimal polynomial of τ does indeed determine CK_τ precisely. So the matter of what constitutes a finite presentation is somewhat dependent on the axioms specifying the category, and the convention in Definition 4.1 is the useful one for the purposes of this paper.

5. Classification of strong extensions

It follows from Theorem 3.11 that finitely generated kernel-preserving strong extensions of ELA-fields are all of the form $F \lhd F \mid V$, where V is additively and multiplicatively free, and Kummer-generic. We next discuss the properties of the varieties V that occur in this way.

Let $G = \mathbb{G}_a \times \mathbb{G}_m$. Each matrix $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$ defines a homomorphism $M : G^n \to G^n$ by acting as a linear map on $\mathbb{G}_a^n$ and as a multiplicative map on $\mathbb{G}_m^n$. If $V \subseteq G^n$, we write $M \cdot V$ for its image. Note that if V is a subvariety of G^n, then so is $M \cdot V$.

Definition 5.1. An irreducible subvariety V of G^n is *rotund* if and only if for every matrix $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$,

$$\dim M \cdot V \geqslant \mathrm{rk}\, M.$$

A reducible subvariety is rotund if and only if at least one of its irreducible components is rotund.

A subvariety V of G^n is *perfectly rotund* if and only if it is irreducible, $\dim V = n$, for every $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$ with $0 < \mathrm{rk}\, M < n$,

$$\dim M \cdot V \geqslant \mathrm{rk}\, M + 1,$$

and also V is additively free, multiplicatively free, and Kummer-generic.

Note that a reducible subvariety may satisfy the dimension property of rotundity without being rotund. For example, take $n = 2$, V_1 given by $x_1 = y_1 = 1$, V_2 given by $x_2 = y_2 = 1$, and $V = V_1 \cup V_2$.

Proposition 5.2. *Let $F \subseteq F|V$ be an extension of ELA-fields, with V additively and multiplicatively free, and Kummer-generic. Then the extension is strong if and only if V is rotund.*

In the proof, and subsequently, we will use the concept of the locus of a tuple. If $F \subseteq F_1$ is a field extension and $\bar{a} \in F_1^n$, then the locus of $\bar{a}$ over F, written $\mathrm{Loc}_F(\bar{a})$ or $\mathrm{Loc}(\bar{a}/F)$, is the smallest Zariski-closed subset of F_1^n containing $\bar{a}$ that is defined over F.

Proof of Proposition 5.2. Let $\bar{a}$ be the tuple generating $F|V$ over F such that $(\bar{a}, e^{\bar{a}}) \in V$. Suppose $F \lhd F|V$, let $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$, and let $\bar{b} = M\bar{a}$. Then $\mathrm{Loc}_F(\bar{b}, e^{\bar{b}}) = M \cdot V$ and $\mathrm{ldim}_{\mathbb{Q}}(\bar{b}) = \mathrm{rk}\, M$, so

$$\dim M \cdot V - \mathrm{rk}\, M = \delta(\bar{b}/F) \geqslant 0$$

and hence V is rotund.

Conversely, suppose that V is rotund, let $F_1 = \langle F, \bar{a} \rangle_{F|V}$, the partial E-field extension of F generated by $\bar{a}$, and let $\bar{b}$ be any tuple from $D(F_1)$. The tuple $\bar{a}$ spans $D(F_1)/F$, so there is $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$ such that there is an equality of $\mathbb{Q}$-vector spaces $\langle M\bar{a} \rangle / F = \langle \bar{b} \rangle / F$. Then

$$\delta(\bar{b}/F) = \delta(M\bar{a}/F) = \dim M \cdot V - \mathrm{rk}\, M \geqslant 0$$

so $F \lhd F_1$. But $F|V = F_1^{ELA}$, so $F \lhd F|V$ as required. $\qquad \square$

Definition 5.3. A strong extension $F \lhd F_1$ of ELA-fields is *simple* if and only if whenever $F \lhd F_2 \lhd F_1$ is an intermediate ELA-field then $F_2 = F$ or $F_2 = F_1$.

It is easy to see that every simple extension of ELA-fields is finitely generated. For, suppose $\bar{a}$ is a nonempty finite tuple from $F_1 \setminus F$. Then there is a finite tuple $\bar{a}'$, extending $\bar{a}$, such that $\langle F, \bar{a}' \rangle_{F_1} \lhd F_1$. Then $F \lhd F_2 := \langle F, \bar{a}' \rangle_{F_1}^{ELA}$ and $F_2 \lhd F_1$, so by simplicity $F_2 = F_1$ and the extension is finitely generated. However, simple extensions are not necessarily generated by a single element.

It is important to distinguish between exponentially algebraic and exponentially transcendental extensions. The full definition of exponential algebraicity is given in [Kirby 2010a], but all we will use is the following fact:

Fact 5.4 [Kirby 2010a, Theorem 1.3]. *Let F be an E-field and suppose $C \lhd F$ is some strong subset, and $\bar{a}$ is a finite tuple from F. Then the exponential transcendence degree of $\bar{a}$ over C in F satisfies*

$$\mathrm{etd}^F(\bar{a}/C) = \min\left\{\delta(\bar{a}, \bar{b}/C) \mid \bar{b} \text{ is a finite tuple from } F\right\}.$$

Exponential transcendence degree is the dimension notion of a pregeometry, analogous to transcendence degree in pure fields. An element a is exponentially algebraic over C if and only if $\mathrm{etd}^F(a/C) = 0$.

Lemma 5.5. *There is a unique simple exponentially transcendental extension of any ELA-field.*

Proof. Let $F \lhd F_1$ be simple, with $a \in F_1$, exponentially transcendental over F. Then $\mathrm{td}(a, e^a/F) = 2$, so the partial E-field extension $\langle F, a \rangle_{F_1}$ is determined uniquely up to isomorphism. But $\langle F, a \rangle_{F_1} \lhd F_1$ by the above characterisation of exponential transcendence degree so, by Theorem 2.18, $\langle F, a \rangle_{F_1}^{ELA} \cong (\langle F, a \rangle_{F_1})^{ELA}$. Then $\langle F, a \rangle_{F_1}^{ELA} \lhd F_1$, so $\langle F, a \rangle_{F_1}^{ELA} = F_1$ because the extension is simple. $\square$

Note that if a is exponentially transcendental over F then $\mathrm{Loc}(a, e^a/F) = G$ (recall that $G = \mathbb{G}_a \times \mathbb{G}_m$), so the simple exponentially transcendental extension of F can be written in our notation as $F | G$.

Proposition 5.6. *If V is perfectly rotund then the strong extension of ELA-fields $F \lhd F | V$ is simple and exponentially algebraic.*

Conversely, if $F \lhd F'$ is a simple, exponentially algebraic extension of ELA-fields then $F' \cong_F F | V$ for some perfectly rotund V.

Proof. Let $\bar{a}$ be the tuple generating $F | V$ over F such that $(\bar{a}, e^{\bar{a}}) \in V$, and let $F_1 = \langle F, \bar{a} \rangle_{F | V}$, the partial E-field extension of F generated by $\bar{a}$.

Since V is rotund and additively and multiplicatively free, $F \lhd F | V$ is exponentially algebraic if and only if $\dim V = n$. Now suppose $F \lhd F_2 \lhd F | V$, a proper intermediate ELA-field. Then $(F_2 \cap D(F_1))/F$ is a nontrivial proper subspace of $D(F_1)/F$, which must be the span of $M\bar{a}$ for some $M \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$, with $0 < \mathrm{rk}\, M < n$. Since V is rotund, $\dim M \cdot V \geqslant \mathrm{rk}\, M$. Extend $M\bar{a}$ to a spanning set $M\bar{a}, \bar{b}$ of $D(F_1)/F$. Then $\delta(\bar{b}/F, M\bar{a}) \geqslant 0$, because $F_2 \lhd F | V$. But

$$\delta(\bar{b}/F, M\bar{a}) = \mathrm{td}(\bar{b}, e^{\bar{b}}/F, M\bar{a}, e^{M\bar{a}}) - \mathrm{ldim}_{\mathbb{Q}}(\bar{b}/F, M\bar{a})$$
$$= [n - \dim M \cdot V] - [n - \mathrm{rk}\, M]$$

so $\dim M \cdot V \leqslant \mathrm{rk}\, M$. Thus $\dim M \cdot V = \mathrm{rk}\, M$, and V is not perfectly rotund.

For the converse, choose $\bar{a}$ a tuple of smallest length that generates F' over F and such that $F_1 := \langle F, \bar{a} \rangle_{F'} \lhd F'$, and let $V = \mathrm{Loc}(\bar{a}, e^{\bar{a}}/F) \subseteq G^n$. Then $F' \cong F_1^{ELA}$. Since n is minimal, V is additively and multiplicatively free. By replacing $\bar{a}$ by $\bar{a}/m$ for some $m \in \mathbb{N}$, we may assume V is Kummer-generic. Since the extension is strong and exponentially algebraic, V is rotund and $\dim V = n$. If V is not perfectly rotund then there is a matrix M with $0 < \mathrm{rk}\, M < n$ such that $\dim M \cdot V = \mathrm{rk}\, M$. Let $F_2 = \langle F, M\bar{a} \rangle_{F'}^{ELA}$. Then $F \lhd F_2 \lhd F'$, but $F_2 \neq F$ and $F \lhd F'$ is simple, so $F_2 = F'$. But F_2 is generated by $M\bar{a}$, which is $\mathbb{Q}$-linearly dependent over F, so by

a basis for it that is a tuple shorter than $\bar{a}$. This contradicts the choice of $\bar{a}$. So V is perfectly rotund. $\square$

We now consider the problem of when two extensions $F|V$ and $F|W$ are isomorphic. Suppose $\bar{a}$ is a generator of $F|V$, with $(\bar{a}, e^{\bar{a}}) \in V$. Then if $\bar{b}$ is a different choice of basis of the extension, so $F \cup \bar{b}$ has the same $\mathbb{Q}$-linear span as $F \cup \bar{a}$, and $W = \mathrm{Loc}(\bar{b}, e^{\bar{b}}/F)$, then clearly $F|W \cong F|V$, but there is no reason why W should be equal to V. Essentially this is the only way an isomorphism can happen.

Definition 5.7. Suppose $V \subseteq G^n$ and $W \subseteq G^m$ are two perfectly rotund varieties, defined over F. Write $V \sim_F W$ if and only if $n = m$, there are $M_1, M_2 \in \mathrm{Mat}_{n \times n}(\mathbb{Z})$ of rank n, and there is $\bar{c} \in F^n$, such that $M_1 \cdot V = M_2 \cdot W + (\bar{c}, e^{\bar{c}})$ (where $+$ means the group operation in G^n, so multiplication on the $\mathbb{G}_m$ coordinates), and furthermore $M_1 \cdot V$ is Kummer-generic.

Proposition 5.8. *If V and W are perfectly rotund and defined over F then*

$$F|V \cong_F F|W \quad \textit{if and only if} \quad V \sim_F W.$$

Proof. Firstly suppose that $V \sim_F W$, and let $V' = M_1 \cdot V$, where M_1 is as above. Let $K = \langle F, \bar{a} \rangle_{F|V}$, where $(\bar{a}, e^{\bar{a}}) \in V$ is the generating tuple. Let $\bar{b} = M_1 \bar{a}$. Then $\langle F, \bar{b} \rangle_{F|V} = K$, and $(\bar{b}, e^{\bar{b}})$ is generic in V'. Furthermore, since V' is Kummer-generic (by assumption), K is well defined by V'. Hence $F|V \cong F|V'$. Similarly, translating V' to $V' - (\bar{c}, e^{\bar{c}})$ for some $\bar{c} \in F^n$ does not change K. So $F|V \cong_F F|W$.

Conversely, suppose $F|V \cong F|W$. Let $(\bar{a}, e^{\bar{a}}) \in V$ be a generating tuple for $F|V$. Let $F_1 = \langle F, \bar{a} \rangle_{F|V}$, and write $F|V$ as the union of a chain of partial E-fields

$$F \lhd F_1 \lhd F_2 \lhd F_3 \lhd \cdots,$$

where for $n \in \mathbb{N}^+$ we have $\mathrm{ldim}_{\mathbb{Q}}(D(F_{n+1})/D(F_n)) = 1$, which is possible since $F|V = F_1^{ELA}$. There is $\bar{b} \in F|V$ such that $\mathrm{Loc}(\bar{b}, e^{\bar{b}}/F) = W$. Suppose that $\bar{b}$ is $\mathbb{Q}$-linearly independent over $D(F_1)$. Then $\mathrm{Loc}(\bar{b}, e^{\bar{b}}/F_1) = W$ because $F_1 \lhd F|V$. Now each $D(F_{n+1})$ is generated over $D(F_n)$ by a single element c_{n+1} such that either c_{n+1} or $e^{c_{n+1}}$ is algebraic over F_n. By perfect rotundity of W, no b in the $\mathbb{Q}$-linear span of $\bar{b}$ satisfies this, so inductively we see that $\bar{b}$ is linearly independent over $D(F_n)$ for all n, a contradiction. So $\bar{b}$ is not $\mathbb{Q}$-linearly independent over $D(F_1)$. Write B for the $\mathbb{Q}$-linear span of $F \cup \bar{b}$. Then $D(F_1) \lhd F|V$ and $B \lhd F|V$, so $D(F_1) \cap B \lhd F|V$, and hence, since V and W are perfectly rotund, we must have $B = D(F_1)$, and $V \sim_F W$ as required. $\square$

We can give a normal form for a finitely generated strong ELA-extension $F \lhd F'$. The key is that the order of making simple extensions can often be interchanged.

Lemma 5.9. *Let F be a countable ELA-field, and let $V \subseteq G^n$ and $W \subseteq G^r$ be two additively and multiplicatively free, irreducible, Kummer-generic subvarieties, defined over F. Then*

$$(F\,|\,V)\,|\,W \cong (F\,|\,W)\,|\,V \cong F\,|\,(V \times W)$$

as extensions of F.

Proof. First note that the extension $(F\,|\,V)\,|\,W$ makes sense, since in the base change from F to $F\,|\,V$, the variety W remains free, irreducible, and Kummer-generic, because both F and $F\,|\,V$ are algebraically closed. Similarly $(F\,|\,W)\,|\,V$ makes sense. Now let $\bar{a}, \bar{b}$ be the tuples in $F_1 = (F\,|\,V)\,|\,W$ such that $(\bar{a}, e^{\bar{a}}) \in V$ determines the first extension and $(\bar{b}, e^{\bar{b}}) \in W$ determines the second extension. Let $\bar{a}', \bar{b}'$ be the equivalent tuples in $F_2 = (F\,|\,W)\,|\,V$. Then the partial E-fields $K_1 = \langle F\bar{a}, \bar{b}\rangle_{F_1}$ and $K_2 = \langle F\bar{a}', \bar{b}'\rangle_{F_2}$ are isomorphic extensions of F, as both $(\bar{a}, e^{\bar{a}}, \bar{b}, e^{\bar{b}})$ and $(\bar{a}', e^{\bar{a}'}, \bar{b}', e^{\bar{b}'})$ are generic in $V \times W$ over F. Now $K_1 \lhd F_1$ and $K_2 \lhd F_2$; hence the result follows by Theorem 2.18. $\qquad\square$

Indeed the extensions $F\,|\,V$ and $F\,|\,W$ can be freely amalgamated over F, and the free amalgam is in fact given by $F\,|\,(V \times W)$.

Now consider a finitely generated strong extension of countable ELA-fields $F \lhd F'$. Let $\bar{a}_1$ be some tuple from F', $\mathbb{Q}$-linearly independent over F, such that $V_1 := \operatorname{Loc}(\bar{a}_1, e^{\bar{a}_1}/F)$ is perfectly rotund. If it does not exists, then $F = F'$. So we have $F = F_0 \lhd F_1 = \langle F, \bar{a}_1\rangle^{ELA}_{F'} \lhd F'$. Now iteratively choose tuples $\bar{a}_i$, which are $\mathbb{Q}$-linearly independent over F_{i-1}, such that $V_i := \operatorname{Loc}(\bar{a}_i, e^{\bar{a}_i}/F)$ is perfectly rotund and defined over F_j, where j is as small as possible, and define $F_i = \langle F_{i-1}, \bar{a}_i\rangle^{ELA}_{F'}$. At some finite stage we will exhaust F'. The previous propositions show there is only a very limited scope for choosing the tuples $\bar{a}_i$. Thus we have a Jordan–Hölder-type theorem, showing how a finitely generated extension decomposes as a chain of simple extensions, and the extent to which the chain is unique.

Theorem 5.10. *If $F \lhd F'$ is a finitely generated strong extension of countable ELA-fields, then it can be decomposed as*

$$F = K_0 \lhd K_1 \lhd K_2 \lhd \cdots \lhd K_r = F'$$

such that $K_i = K_{i-1}\,|\,V_i$ with $V_i = V_{i,1} \times \cdots \times V_{i,m_i}$ with each $V_{i,j}$ perfectly rotund and defined over K_{i-1} but not defined over K_{i-2}. Furthermore, if there is another decomposition

$$F = K_0 \lhd K_1' \lhd K_2' \lhd \cdots \lhd K_s' = F'$$

such that $K_i' = K_{i-1}'\,|\,V_i'$ with $V_i' = V_{i,1}' \times \cdots \times V_{i,q_i}'$, then $s = r$ and, for each i, $q_i = m_i$ and there is a permutation σ of $\{1, \ldots, m_i\}$ such that $V_{i,j}' \sim_{F'} V_{i,\sigma(j)}$.

A finer analysis is possible, in which one takes into account for each $V_{i,j}$ precisely which of the $V_{s,t}$ for $s < i$ are involved in the field of definition of $V_{i,j}$, to produce a partial order on the simple extensions.

6. The strong exponential-algebraic closure

We now consider the analogue for exponential fields of the algebraic closure of a field.

Definition 6.1. An exponential field F is said to be *strongly exponentially-algebraically closed* if and only if it is an ELA-field and for every finitely generated partial E-subfield A of F, and every finitely generated kernel-preserving exponentially algebraic strong partial E-field extension $A \lhd B$, there is an embedding $B \hookrightarrow F$ fixing A.

The word *strongly* in this definition actually does not refer to the strong extensions, but rather signifies that the property is stronger than another property, called *exponential-algebraic closedness*, which was also considered by Zilber. Exponential-algebraic closedness is a model-theoretic approximation to strong exponential-algebraic closedness, which is first-order axiomatisable, but strong exponential-algebraic closedness is the sensible notion from the algebraic point of view taken in this paper.

We now show that every countable ELA-field has a well-defined strong exponential-algebraic closure.

Theorem 6.2. *Let F be a countable ELA-field. Then there is a strongly exponentially-algebraically closed $F^{\sim}$ with $F \lhd F^{\sim}$ such that if $F \lhd K$, K is strongly exponentially-algebraically closed, and $\ker(K) = \ker(F)$, then there is a strong embedding $F^{\sim} \lhd K$ such that*

$$F \overset{\hookrightarrow}{\dashrightarrow} F^{\sim}$$

$$K$$

commutes. Furthermore, $F^{\sim}$ is unique up to isomorphism as an extension of F. We call it the strong exponential-algebraic closure of F.

The key property we need is the amalgamation property, which follows from Lemma 5.9.

Proof of Theorem 6.2. Let F be a countable ELA-field. List the triples

$$(n_\alpha, V_\alpha, A_\alpha)_{\alpha < \omega}$$

such that $n_\alpha \in \mathbb{N}^+$, V_α is a perfectly rotund subvariety of $G^{n_\alpha}(F)$, A_α is a finitely generated subfield of F over which V_α is defined, and F does not contain $\bar{b}$ such

that $(\bar{b}, e^{\bar{b}})$ is generic in V_α over A_α. Note that if F is not strongly exponentially-algebraically closed then there will be $\aleph_0$ such triples.

Let F_1 be the ELA-extension of F obtained by making the simple ELA-extensions determined by each V_α in turn. By Lemma 5.9 (and a back and forth argument), F_1 does not depend on the choice of well-ordering. Now iterate the process to produce a chain

$$F \lhd F_1 \lhd F_2 \lhd F_3 \lhd \cdots$$

and let $F^\sim$ be the union of this chain.

By construction, $F^\sim$ is strongly exponentially-algebraically closed, and $F \lhd F^\sim$. Furthermore if F is strongly exponentially-algebraically closed then $F = F^\sim$. The primality property and the uniqueness of $F^\sim$ follow from a standard back-and-forth argument.   □

If F is a partial E-field such that F^{ELA} is well-defined, then we also write $F^\sim$ for $(F^{ELA})^\sim$.

Note that if $F^\sim \neq F$ then $F^\sim$ will not be *minimal* over F, that is, it will be isomorphic over F to a proper subfield of itself. This is because we adjoin $\aleph_0$ copies of the extension of F defined by V_1 in constructing F_1, and if we miss out cocountably many of those realisations, we get a proper ELA-subfield of F_1 that is isomorphic over F to F_1.

Definition 6.3. Zilber's *pseudoexponential fields* are precisely the exponential fields F satisfying the following properties:

(1) F is an ELA-field.

(2) F has standard kernel.

(3) The Schanuel property holds.

(4) F is strongly exponentially-algebraically closed.

(5) For any countable subset $A \subseteq F$, the exponential-algebraic closure $\mathrm{ecl}^F(A)$ is countable.

Of course these are genuine exponential fields in our algebraic sense. The prefix "pseudo" refers to Zilber's programme of *pseudoanalytic* or *pseudocomplex* structures.

Construction 6.4 (Zilber's pseudoexponential fields). Let $B_0 = SK^\sim$. For each ordinal α, define $B_{\alpha+1} = (B_\alpha | G)^\sim$ (where $G = \mathbb{G}_a \times \mathbb{G}_m$ as before). For limit α, take unions. It is easy to see that the exponential transcendence degree of B_α is $|\alpha|$, and that the isomorphism type of B_α depends only $|\alpha|$. For a cardinal κ we write B_κ for the model of exponential transcendence degree κ.

By construction, the B_κ satisfy Zilber's axioms, and hence are pseudoexponential fields. Although B_κ exists for all cardinals κ, we have only proved that F^{ELA} and hence $F^{\sim}$ are uniquely defined for countable F, and hence the arguments of this paper only show that B_κ is well-defined for countable κ.

We now proceed to examine strong exponential-algebraic closedness in more detail before proving that the B_κ for countable κ are the only countable pseudoexponential fields.

The property of strong exponential-algebraic closedness is most useful when there is a proper subset of F that is strongly embedded in F, and especially when a finite such subset exists.

Definition 6.5. An E-field F is said to have ASP, for *almost the Schanuel property*, if and only if there is a finite tuple $\bar{c}$ from F such that $\langle \bar{c} \rangle_F \lhd F$.

Lemma 6.6. *Any strong extension of a finitely presented ELA-field has ASP.*

Proof. If F is a strong extension of a finitely presented ELA-field, then it is a strong extension of a finitely generated partial E-field F_0, and we can take $\bar{c}$ to be a generating tuple for F_0. □

Example 6.7. Consider the exponential field $\mathbb{C}_{2^x} = \langle \mathbb{C}; +, \cdot, 2^x \rangle$. Then $\mathbb{C}_{2^x}$ does not satisfy the Schanuel property because $2^1 = 2$, but if Schanuel's conjecture is true then it does satisfy ASP.

ASP is a slight weakening of the Schanuel property that allows for some extra generality such as this example, but such that the theory works almost unchanged.

Lemma 6.8. *Suppose F is an ELA-field. Then the following are equivalent.*

(1) *F is strongly exponentially-algebraically closed.*

(2) *For each $n \in \mathbb{N}$, every perfectly rotund subvariety $V \subseteq G^n(F)$, and every finitely generated subfield A of F over which V is defined, there is $(\bar{b}, \exp(\bar{b}))$ in F, generic in V over A.*

(3) *For each $n \in \mathbb{N}$, every additively and multiplicatively free, rotund subvariety $V \subseteq G^n(F)$, and every finitely generated subfield A of F over which V is defined, there are infinitely many distinct $(\bar{b}, \exp(\bar{b}))$ in F, generic in V over A.*

Furthermore, if F satisfies ASP then the next three conditions are also equivalent to the first three.

(4) *For each $n \in \mathbb{N}$, every perfectly rotund subvariety $V \subseteq G^n(F)$, and every finitely generated ELA-subfield A of F over which V is defined, there is $(\bar{b}, \exp(\bar{b}))$ in F, generic in V over A.*

(5) *For each finitely generated ELA-subfield A of F, and each finitely generated exponentially algebraic strong ELA-extension $A \lhd B$, there is an embedding $B \hookrightarrow F$ fixing A.*

(6) *For each finitely generated strong ELA-subfield $A \lhd F$, and each simple exponentially algebraic strong ELA-extension $A \lhd B$, there is an embedding $B \hookrightarrow F$ (necessarily strong) fixing A.*

Proof. (1) $\Longleftrightarrow$ (2) by Proposition 5.6. (3) $\Longrightarrow$ (2) is trivial. To show (2) $\Longrightarrow$ (3), first note that every finitely generated strong extension is the union of a chain of simple strong extensions, so to find a point in an additively and multiplicatively free rotund subvariety it is enough to find points in perfectly rotund subvarieties. Now we show by induction on $r \in \mathbb{N}$ that there are at least r many such $\bar{b}$. The case $r = 1$ is (2). Now suppose we have $\bar{b}_1, \ldots, \bar{b}_r$. Then by (2) there is a $\bar{b}_{r+1}$ such that $(\bar{b}_{r+1}, \exp(\bar{b}_{r+1}))$ is generic in V over $A \cup \{\bar{b}_1, \exp(\bar{b}_1), \ldots, \bar{b}_r, \exp(\bar{b}_r)\}$ In particular, $\bar{b}_1, \ldots, \bar{b}_{r+1}$ are distinct.

It is immediate that (4) implies (2), that (4) implies (5), and that (5) implies (6). We now assume that there is a finite $\bar{c} \lhd F$. Assume (2), and let $A = \langle \bar{a}, \bar{c} \rangle_F^{ELA}$ be a finitely generated ELA-subfield of F. Since $\bar{c} \lhd F$, we may assume that $A \lhd F$ by extending the tuple $\bar{a}$ if necessary. By (2), there is $(\bar{b}, \exp(\bar{b}))$ in F, generic in V over $\bar{a}$. By Lemma 5.9, the ELA-subfield $\langle \bar{a}, \bar{b} \rangle_F^{ELA}$ of F is isomorphic to $A \mid V$, and $(\bar{b}, \exp(\bar{b}))$ is generic in V over A. Hence (4) holds.

Now assume (6), let V be perfectly rotund, and let A be a finitely generated ELA-subfield over which V is defined. Then there is a finitely generated ELA-subfield A' of F containing A and $\bar{c}$ such that $A' \lhd F$. By (6), there is a realisation of $A' \mid V$ in F over A', say generated by $(\bar{b}, \exp(\bar{b}))$, generic in V over A'. But then $(\bar{b}, \exp(\bar{b}))$ is generic in V over A as $A \subseteq A'$, hence (4) holds. $\qquad \square$

Proposition 6.9. *Let F_0 be a finitely generated partial E-field with full kernel (or a finitely presented ELA-field), and let $F_0 \lhd F$ be a countable, kernel-preserving, strongly exponentially-algebraically closed strong extension of F_0. Then F is determined up to isomorphism as an extension of F_0 by the exponential transcendence degree* $\mathrm{etd}(F/F_0)$.

Proof. Suppose F_0 is as above and let $\mathscr{C}(F_0)$ be the category of all countable strong kernel-preserving ELA-extensions of F_0, with strong embeddings fixing F_0 as the morphisms. Let $\mathscr{C}^{<\aleph_0}(F_0)$ be the full subcategory of finitely generated ELA-extensions of F_0. Then $\mathscr{C}(F_0)$ is an $\aleph_0$-amalgamation category, that is:

- Every arrow is a monomorphism.
- $\mathscr{C}_0(F)$ has unions of ω-chains (by Lemma 2.3).
- $\mathscr{C}^{<\aleph_0}(F_0)$ has only countably many objects up to isomorphism (Theorem 3.11).
- For each $A \in \mathscr{C}^{<\aleph_0}(F_0)$, there are only countably many extensions of A in $\mathscr{C}^{<\aleph_0}(F_0)$ up to isomorphism (also by Theorem 3.11).
- $\mathscr{C}^{<\aleph_0}(F_0)$ has the amalgamation property (by Lemma 5.9).

- $\mathscr{C}^{<\aleph_0}(F_0)$ has the joint embedding property (since F_0^{ELA} embeds in all of the strong ELA-extensions of F_0, by Theorem 2.18).

Thus by the Fraïssé amalgamation theorem, specifically the version in [Kirby 2009, Theorem 2.18], there is a unique $\mathscr{C}^{<\aleph_0}(F_0)$-saturated extension $F_0 \lhd F$ in $\mathscr{C}(F_0)$, that is, one such that for any finitely generated ELA-extension A of F_0 inside F, and any finitely generated strong ELA extension $A \lhd B$, there is a (necessarily strong) embedding of B into F over A. Using part (6) of Lemma 6.8, this property is the same as being strongly exponentially-algebraically closed together with $\text{etd}(F/F_0)$ being infinite. Thus the proposition is proved in the case where $\text{etd}(F/F_0) = \aleph_0$.

Now suppose $F_0 \lhd F$ is as in the proposition with $\text{etd}(F/F_0) = n \in \mathbb{N}$. Let $a_1, \ldots, a_n$ be an exponential transcendence base for F over F_0, and let $F_1 = \langle F_0, a_1, \ldots, a_n \rangle_F^{ELA}$. Then $F_1 \cong_{F_0} F_0|G^n$, and $\text{etd}(F/F_1) = 0$. So it is enough to consider the case $\text{etd}(F/F_0) = 0$. Let $\mathscr{C}_0(F_0)$ be the subcategory of $\mathscr{C}(F_0)$ consisting of the exponentially algebraic extensions. The same argument as above shows that $\mathscr{C}(F_0)$ is an $\aleph_0$-amalgamation category, and we deduce that up to isomorphism there is a unique countable, kernel-preserving, strongly exponentially-algebraically closed strong extension of F_0, which of course is $F_0^{\sim}$. $\square$

Corollary 6.10. *The countable pseudoexponential fields are precisely B_κ for κ a countable cardinal.*

Proof. The pseudoexponential fields are all kernel-preserving strongly exponentially-algebraically closed strong extensions of SK. $\square$

From the proof of Proposition 6.9 one can see that much of the machinery of $\aleph_0$-stable first-order theories can be brought to bear on the category $\mathscr{C}(F_0)$ for any finitely presented ELA-field F_0. Indeed, the strongly exponentially-algebraically closed kernel-preserving strong extensions of F_0 (at least those with the countable closure property) should be thought of as analogous to the algebraically closed pure field extensions of a finitely generated field. They are the "universal domains" that are saturated and $\aleph_0$-homogeneous for the category $\mathscr{C}(F_0)$. Of course this is not saturation nor $\aleph_0$-stability in the sense of first-order model theory, because we are only considering extensions that do not extend the kernel. Also the $\aleph_0$-stability is with respect to counting types over strong ELA-subfields of F, not over arbitrary subsets. Developing the homogeneity property further, we make some observations about extending automorphisms.

Proposition 6.11. *Suppose that F is a partial E-field with full kernel, which is finitely generated or a finitely generated extension of a countable LA-field, and that σ is an automorphism of F.*

(1) *σ extends uniquely to an automorphism of F^E.*

(2) σ *extends to automorphisms of* F^{EA}, F^{ELA}, *and to any countable strongly exponentially-algebraically closed kernel-preserving strong extension of* F, *including* $F^\sim$.

Proof. To extend an automorphism σ of F to an automorphism of F^e we must have $\sigma(c_{i,n}) = \exp(\sigma(b_i)/n)$, in the notation of Construction 2.6. This does define a partial automorphism since the c_i are algebraically independent over F, and it extends uniquely to an automorphism of F^e because F^e is generated over F by the $c_{i,n}$. Thus, by induction, σ extends uniquely to an automorphism of F^E.

We have an extension $\theta : F \to F^{ELA}$, where θ is the inclusion map, and a second extension $\theta \circ \sigma : F \to F^{ELA}$. The partial E-field F satisfies the hypothesis of Theorem 2.18, so by that theorem there is a map $\bar\sigma : F^{ELA} \to F^{ELA}$ that restricts to σ on F. The image of $\bar\sigma$ is an ELA-subfield of F^{ELA} containing F, so must be all of F^{ELA}. Hence $\bar\sigma$ is an automorphism of F^{ELA} extending σ. The restriction of $\bar\sigma$ to F^{EA} is an automorphism of F^{EA} extending σ. Similarly, we can use the $\mathscr{C}^{<\aleph_0}(F_0)$-saturation and $\mathscr{C}_0^{<\aleph_0}(F_0)$-saturation properties to extend $\bar\sigma$ from F^{ELA} to an automorphism of $F^\sim$ or of another countable strongly exponentially-algebraically closed kernel-preserving strong extension of F. $\qquad\square$

The partial E-field SK embeds in $\mathbb{C}_{\exp}$, so the restriction, σ_0, of complex conjugation is an automorphism of SK, and it is easy to see that it and the identity map are the only automorphisms of SK. By Proposition 6.11, σ_0 extends to automorphisms of B_κ for any countable κ, and in [Kirby et al. 2012], these extensions of σ_0 are used to identify the algebraic numbers that are pointwise definable in pseudoexponential fields. However, the extensions of σ_0 are far from being unique, so this argument does not give an analogue of complex conjugation on any B_κ.

7. Nonmodel completeness

In this section we show that the B_κ, and other strongly exponentially-algebraically closed E-fields, are not model complete. We use the submodularity property of δ, which is well known from Hrushovski's amalgamation constructions, and some simple consequences.

Lemma 7.1 (submodularity). *Let* F *be a partial E-field, and let* C, X, Y *be* $\mathbb{Q}$-*subspaces of* $D(F)$ *such that* $C \subseteq X \cap Y$ *and* $\mathrm{ldim}_{\mathbb{Q}}(X \cup Y/C)$ *is finite. Let* $\bar x, \bar y, \bar z$ *be finite tuples such that* $\bar x \cup C$ *spans* X, $\bar y \cup C$ *spans* Y, *and* $\bar z \cup C$ *spans* $X \cap Y$. *Then*

$$\delta(\bar x \cup \bar y/C) + \delta(\bar z/C) \leqslant \delta(\bar x/C) + \delta(\bar y/C). \tag{1}$$

More prosaically, the predimension function $\delta(\,\cdot\,/C)$ *is submodular on the lattice of* $\mathbb{Q}$-*linear subspaces of* $D(F)$. *We write*

$$\delta(XY/C) + \delta(X \cap Y/C) \leqslant \delta(X/C) + \delta(Y/C).$$

Proof. If δ is replaced by td then (1) holds, and if δ is replaced by ldim$_\mathbb{Q}$ then it holds with $\leqslant$ replaced by $=$. Subtracting the latter from the former gives (1). □

Lemma 7.2. *Suppose F is a partial E-field, and $C \lhd F$. For each finite tuple $\bar{a}$ from F, there is a smallest $\mathbb{Q}$-vector subspace $\lceil C, \bar{a} \rceil_F$ of F containing $\bar{a}$ and C, called the* hull *of $C \cup \bar{a}$, such that $\lceil C, \bar{a} \rceil_F \lhd F$. Furthermore, $\lceil C, \bar{a} \rceil_F$ is finite-dimensional as an extension of C.*

Proof. Since $C \lhd F$, there is a finite-dimensional extension $C \subseteq A \subseteq D(F)$ with $\bar{a} \in A$ such that $\delta(A/C)$ is minimal, say equal to d. If A_1 and A_2 are two such, then by submodularity we see that $\delta(A_1 \cap A_2/C) \leqslant d$, and hence we can take $\lceil C, \bar{a} \rceil_F$ to be the intersection of all such A. □

Remark 7.3. Often in amalgamation-with-predimension constructions, the analogue of what is here called the *hull* is called the *strong closure* or, when *self-sufficient* is used in place of strong, the *self-sufficient closure*. While the notion of a self-sufficient subset makes semantic sense (X is self-sufficient in F if no witnesses outside X are needed to realise its full type in F), the sense is lost when dealing with extensions rather than subsets because in "F is a self-sufficient extension of X", the "self" should semantically refer to X rather than F, in conflict with the syntactic construction of the phrase. Since the focus here is on extensions rather than subsets, we do not use the terminology of self-sufficiency. Similarly, the terminology "strong closure" conflicts with the notion here of strong exponential-algebraic closure. The simplest amalgamation-with-predimension construction is that of the universal acyclic graph, and there the concept corresponding to our hull is exactly the convex hull of a set in the sense of the graph, that is, the hull of X is the union of all paths between elements of X.

Proposition 7.4. *Suppose F is an E-field, $C \lhd F$, and $\bar{a}$ is a tuple from F. Suppose $K \subseteq F$ is an E-subfield of F, containing C, such that $\lceil C, \bar{a} \rceil_F \cap K$ is spanned by $C \cup \bar{a}$. Let $r = \delta(\bar{a}/C) - \delta(\lceil C, \bar{a} \rceil_F/C)$. Then $\mathrm{etd}^K(\bar{a}/C) = \mathrm{etd}^F(\bar{a}/C) + r$.*

Proof. By Fact 5.4 and the definition of the hull,

$$\mathrm{etd}^K(\bar{a}/C) = \min\left\{ \delta(\bar{a}, \bar{b}/C) \mid \bar{b} \subseteq K \right\} = \delta(\lceil C, \bar{a} \rceil_K/C),$$

and similarly,

$$\mathrm{etd}^F(\bar{a}/C) = \delta(\lceil C, \bar{a} \rceil_F/C).$$

So we must show that $\delta(\lceil C, \bar{a} \rceil_K/C) = \delta(\bar{a}/C)$, or equivalently that for any $\bar{b}$ from K, $\delta(\bar{a}, \bar{b}/C) \geqslant \delta(\bar{a}/C)$.

Let A be the $\mathbb{Q}$-span of $C \cup \bar{a}$, $H = \lceil \bar{a}, C \rceil_F$, and let $B \subseteq K$ be an extension of A, generated by some tuple $\bar{b}$. Then, by the assumption on K, $B \cap H = A$. By the submodularity of δ,

$$\delta(B/C) - \delta(A/C) \geqslant \delta(BH/C) - \delta(H/C),$$

but the right hand side is positive as $H = \lceil A \rceil_F$. Hence $\delta(B/C) \geqslant \delta(A/C)$ as required. □

Proposition 7.5. *Let F be a countable strongly exponentially-algebraically closed E-field satisfying ASP, and of exponential transcendence degree at least 1. Then there is $K \subseteq F$, a proper E-subfield such that $K \cong F$ but the inclusion $K \hookrightarrow F$ is not an elementary embedding. In particular, F is not model-complete.*

Proof. (My thanks to Alf Onshuus who noticed a mistake in an earlier version of this proof.) Using the ASP assumption, let $\bar{c}$ be a finite tuple such that $\bar{c} \lhd F$ and $\mathrm{etd}(F/\bar{c}) \geqslant 1$, and let $F_0 = \langle \bar{c} \rangle_F^{ELA}$. So F_0 is a finitely generated strong ELA-subfield of F.

The precise variety V we use is not so important so we take a simple example, the intersection of three generic hyperplanes in G^3. That is, let $\alpha_1, \ldots, \alpha_{18} \in F_0$ be algebraically independent and let V be the subvariety of G^3 given by

$$\alpha_1 X_1 + \alpha_2 X_2 + \alpha_3 X_3 + \alpha_4 Y_1 + \alpha_5 Y_2 + \alpha_6 Y_3 = 1,$$
$$\alpha_7 X_1 + \alpha_8 X_2 + \alpha_9 X_3 + \alpha_{10} Y_1 + \alpha_{11} Y_2 + \alpha_{12} Y_3 = 1,$$
$$\alpha_{13} X_1 + \alpha_{14} X_2 + \alpha_{15} X_3 + \alpha_{16} Y_1 + \alpha_{17} Y_2 + \alpha_{18} Y_3 = 1,$$

where X_1, X_2, X_3 are the coordinates in $\mathbb{G}_a$ and Y_1, Y_2, Y_3 are the coordinates in $\mathbb{G}_m$.

Claim. *V is perfectly rotund.*

Proof. Certainly V is irreducible and has dimension 3. The projections to $\mathbb{G}_a^3$ and to $\mathbb{G}_m^3$ are dominant, so V is additively and multiplicatively free. Similarly, for any $M \in \mathrm{Mat}_{3\times 3}(\mathbb{Z})$, if $\mathrm{rk}\, M = 2$ then $\dim M \cdot V = 3$ and if $\mathrm{rk}\, M = 1$ then $\dim M \cdot V = 2$. V must be Kummer-generic from its simple structure, but in any case we could replace it by the locus of $(X_1/m, X_2/m, X_3/m, \sqrt[m]{Y_1}, \sqrt[m]{Y_1}, \sqrt[m]{Y_1})$ for a suitably large integer m (where $(X_1, X_2, X_3, Y_1, Y_2, Y_3)$ is generic in V) without affecting the rest of the argument. □

Choose $(a_1, a_2, a_3) \in F^3$ such that $(a_1, a_2, a_3, e^{a_1}, e^{a_2}, e^{a_3})$ is generic in V over F_0. Since F is strongly exponentially-algebraically closed and has ASP, such a point exists by Lemma 6.8. Now let $t \in F$ be exponentially transcendental over F_0, let $F_1 = \langle F_0(t) \rangle_F^{ELA}$, and let $K_1 = \langle F_0(a_1) \rangle_F^{ELA}$.

Claim. *$a_2, a_3 \notin K_1$.*

Proof. The intuition here is that V already gives the maximum three constraints between a_1, a_2, and a_3. If a_2 or a_3 were to lie in K_1 that would be an extra constraint, or perhaps $r + 1$ extra constraints with r extra witnesses, which would contradict F_0 being strong in F.

Suppose for a contradiction that $a_2 \in K_1$. Then there is a shortest chain of subfields of K_1, given by

$$\mathrm{acl}^F(F_0(a_1, e_1^a)) = L_0 \subset L_1 \subset L_2 \subset \cdots \subset L_r,$$

such that $a_2 \in L_r$ and, for each $i \in \{1, \ldots, r\}$, there are $x_i, e^{x_i} \in L_i$ such that $L_i = \mathrm{acl}^F(L_{i-1}(x_i, e^{x_i}))$ and either $x_i \in L_{i-1}$ or $e^{x_i} \in L_{i-1}$.

For each $i \in \{1, \ldots, r\}$, we have $\mathrm{td}(L_i/L_{i-1}) = 1$. We can consider each L_i as a partial exponential field by taking the intersection of the graph of $\exp_F$ with L_i^2. Then $L_i \neq L_{i-1}$ for each i, so $x_i \in D(L_i) \smallsetminus D(L_{i-1})$, so in particular $a_1, x_1, \ldots, x_r$ are $\mathbb{Q}$-linearly independent over F_0, and $e^{a_1}, e^{x_1}, \ldots, e^{x_r}$ are multiplicatively independent over F_0.

Let $V' \subseteq G^2$ be the fibre of V given by fixing the coordinates $X_1 = a_1$ and $Y_1 = e^{a_1}$. So V' is the locus of $(a_2, a_3, e^{a_2}, e^{a_3})$ over L_0. Also $\dim V' = 1$, and V' projects dominantly to each coordinate, so $a_2, a_3, e^{a_2}, e^{a_3}$ are interalgebraic over L_0. In particular, they all lie in L_r, and their locus over L_{r-1} is V'. Since V is additively and multiplicatively free, so is V'. So a_2, a_3 are $\mathbb{Q}$-linearly independent over L_{r-1} and e^{a_2}, e^{a_3} are multiplicatively independent over L_{r-1}.

Thus if $x_r \in L_{r-1}$ then $a_1, a_2, a_3, x_1, \ldots, x_r$ are $\mathbb{Q}$-linearly independent over F_0. Otherwise, $e^{x_r} \in L_{r-1}$, and $e^{a_1}, e^{a_2}, e^{a_3}, e^{x_1}, \ldots, e^{x_r}$ are multiplicatively independent over F_0, but then again (since the kernel of the exponential map lies in F_0) $a_1, a_2, a_3, x_1, \ldots, x_r$ are $\mathbb{Q}$-linearly independent over F_0.

So we have

$$\mathrm{td}(a_1, a_2, a_3, x_1, \ldots, x_r, e^{a_1}, e^{a_2}, e^{a_3}, e^{x_1}, \ldots, e^{x_r}/F_0)$$
$$= \mathrm{td}(L_r/L_0) + \mathrm{td}(L_0/F_0) = r + 2,$$

and thus

$$\delta(a_1, a_2, a_3, x_1, \ldots, x_r/F_0) = r + 2 - (r + 3) = -1,$$

which contradicts $F_0 \lhd F$. Hence $a_2, a_3 \notin K_1$. $\square$

Indeed, the proof of the claim shows that a_2 and a_3 must be $\mathbb{Q}$-linearly independent over K_1 since their locus over K_1 is the same as over L_0. Now $\lceil F_0, a_1 \rceil^F$ is spanned by F_0, a_1, a_2, a_3, so by ,

$$\mathrm{etd}^{K_1}(a_1/F_0) = \mathrm{etd}^F(a_1/F_0) + \delta(a_1/F_0) - \delta(a_1, a_2, a_3/F_0) = 0 + 1 - 0 = 1.$$

Thus $\mathrm{etd}^{K_1}(a_1/F_0) = \mathrm{etd}^{F_1}(t/F_0) = 1$, so there is an isomorphism $\theta_1 : F_1 \to K_1$ taking t to a_1 and fixing F_0 pointwise. Now choose an ω-chain of ELA-subfields of F

$$F_1 \lhd F_2 \lhd F_3 \lhd \cdots \lhd F$$

such that F_{n+1} is a simple strong ELA-extension of F_n, for each n, and $\bigcup_{n\in\mathbb{N}} F_n = F$. Inductively we construct chains of ELA-subfields $(K_n)_{n\in\mathbb{N}}$ of F and isomorphisms $\theta_n : F_n \to K_n$, and we also prove that $\mathrm{etd}(F/F_n) + 1 = \mathrm{etd}(F/K_n)$. (If $\mathrm{etd}(F)$ is infinite this is trivially true since both sides will be equal to $\aleph_0$.) We already have K_1 and θ_1. Note that $\mathrm{etd}(F/F_1) + 1 = \mathrm{etd}(F/K_1)$ since $\{t\}$ is an exponential transcendence base for $\mathrm{ecl}^F(F_1)$ over $\mathrm{ecl}^F(K_1)$.

Suppose we have K_n and θ_n. If $F_n \lhd F_{n+1}$ is an exponentially transcendental simple extension, then choose any $b \in F$ that is exponentially transcendental over K_n, and take $K_{n+1} = \langle K_n(b) \rangle_F^{ELA}$. This b exists because $\mathrm{etd}(F/F_n) \leqslant \mathrm{etd}(F/K_n)$. Also $\mathrm{etd}(F/F_{n+1}) + 1 = \mathrm{etd}(F/F_n)$ and $\mathrm{etd}(F/K_{n+1}) + 1 = \mathrm{etd}(F/K_n)$, so

$$\mathrm{etd}(F/F_{n+1}) + 1 = \mathrm{etd}(F/K_{n+1}).$$

By Lemma 5.5, θ_n extends to an isomorphism $\theta_{n+1} : F_{n+1} \to K_{n+1}$. Now suppose that $F_n \lhd F_{n+1}$ is exponentially algebraic. Let V_{n+1} be the corresponding perfectly rotund subvariety, say given by some equations $f_i = 0$, with coefficients in F_n. Let W_{n+1} be the subvariety obtained from V_{n+1} by applying θ_n^{-1} to all the coefficients of the f_i. Then W_{n+1} is a perfectly rotund subvariety defined over K_n, and K_n is a finitely generated ELA-subfield of F, which satisfies ASP, so by Lemma 6.8 there is a realisation of the ELA-extension of K_n corresponding to W_{n+1} in F. Let K_{n+1} be such a realisation, and let θ_{n+1} be any isomorphism from F_{n+1} to K_{n+1} extending θ_n. Also $\mathrm{etd}(F/F_{n+1}) = \mathrm{etd}(F/F_n)$ and $\mathrm{etd}(F/K_{n+1}) = \mathrm{etd}(F/K_n)$.

Let $K = \bigcup_{n\in\mathbb{N}} K_n$ and $\theta = \bigcup_{n\in\mathbb{N}} \theta_n$. Then $\theta : F \to K$ is an isomorphism. But the inclusion $K \subseteq F$ is not an elementary embedding, because

$$F \models \exists x_2, x_3[(a_1, x_2, x_3, e^{a_1}, e^{x_2}, e^{x_3}) \in V]$$

but $K_1 \lhd K$, so $\mathrm{etd}^K(a_1) = 1$, and hence

$$K \models \neg\exists x_2, x_3[(a_1, x_2, x_3, e^{a_1}, e^{x_2}, e^{x_3}) \in V] \qquad \square$$

Theorem 7.6. *Zilber's pseudoexponential fields* (*of exponential transcendence degree at least* 1) *are not model-complete.*

Proof. Proposition 7.5 shows that B_κ is not model-complete when $1 \leqslant \kappa \leqslant \aleph_0$. By [Zilber 2005, Theorem 5.13], every pseudoexponential field of infinite exponential transcendence degree is $L_{\omega_1,\omega}$-equivalent to $B_{\aleph_0}$, so in particular elementarily equivalent, and hence also not model-complete. $\qquad \square$

Remark 7.7. We know that the complex exponential field $\mathbb{C}_{\exp}$ is not model complete [Marker 2006, Proposition 1.1], but that proof uses topological methods and the definability of $\mathbb{Z}$ and $\mathbb{Q}$. Here we use exclusively algebraic methods, and in fact we have shown more than nonmodel-completeness in the language of exponential fields. We have shown that even if one adds symbols for every definable subset of

the kernel, then the result is still not model-complete. I believe that it is not known whether $\mathbb{C}_{\exp}$ is model complete after adding symbols for every definable subset of the kernel.

8. The Schanuel Nullstellensatz

D'Aquino, Macintyre and Terzo [D'Aquino et al. 2010] and also Shkop [2012] have shown that every strongly exponentially-algebraically closed exponential field satisfies the *Schanuel Nullstellensatz*:

Definition 8.1. An ELA-field F is said to satisfy the Schanuel Nullstellensatz if and only if whenever $f \in F[X_1, \ldots, X_n]^E$ is an exponential polynomial over F, not equal to $\exp(g)$ for any exponential polynomial g, then there are $a_1, \ldots, a_n \in F$ such that $f(a_1, \ldots, a_n) = 0$.

This statement was conjectured by Schanuel to hold in $\mathbb{C}_{\exp}$, and Henson and Rubel [1984, Theorem 5.4] proved that it does indeed hold there.

To show that a pure field is algebraically closed it is enough to know that every nontrivial polynomial has a root. The Schanuel Nullstellensatz is an analogue of that statement, but it does not characterise strongly exponentially-algebraically closed exponential fields.

Theorem 8.2. *There are ELA-fields satisfying the Schanuel Nullstellensatz that are not strongly exponentially-algebraically closed.*

Proof. Suppose that F is an ELA-field and $F \hookrightarrow F'$ is a partial E-field extension generated by a solution $a_1, \ldots, a_n$ to an exponential polynomial f (allowing iterations of exponentiation), not of the form $\exp(g)$. Following Shkop, we find F' is also generated over F by a tuple $\bar{b}$ such that $V_f := \mathrm{Loc}(\bar{b}, e^{\bar{b}}/F) \subseteq G^m$ is rotund, additively and multiplicatively free, and of dimension $m + n - 1$. In particular, the extension is strong. The method is to add extra variables to remove the iterations of exponentiation, and then to remove variables to ensure freeness. It follows that some choice of $n - 1$ of the a_i are exponentially-algebraically independent over F, and the remaining one (say a_1) satisfies an exponential polynomial equation in one variable over $F \cup \{a_2, \ldots, a_n\}$. Thus if F has infinite exponential transcendence degree, then it satisfies the Schanuel Nullstellensatz if and only if it satisfies the same statement just for exponential polynomials in one variable.

Now if $F \lhd F'$ is an E-field extension given by adjoining a root a of an exponential polynomial f in one variable, then $F' = \langle F, a \rangle_{F'}^E$. That is, as an E-field extension it is generated by the single element a.

Define a perfectly rotund variety V to have *depth 1* if and only if in an extension $F \lhd F | V$ with generating tuple $\bar{a}$, there is a single element c such that $\bar{a}$ is contained

in $\langle F, c \rangle^E$. Equivalently, $F \lhd \langle F, c \rangle^E \lhd F \,|\, V$. Since $\langle F, c \rangle^E$ is an E-field but not an ELA-field, such an intermediate field is possible.

Let $\mathscr{C}_1^{<\aleph_0}$ be the smallest category of finitely generated strong ELA-extensions F of SK^{ELA} that is closed under simple extensions that are either exponentially transcendental or given by perfectly rotund varieties of depth 1. Let $\mathscr{C}_1$ be the closure of $\mathscr{C}_1^{<\aleph_0}$ under unions of ω-chains. Then, as in the proof of Proposition 6.9, $\mathscr{C}_1$ is an amalgamation category and hence has a unique Fraïssé limit, say $\mathscr{U}$. Then $\mathscr{U}$ satisfies the Schanuel Nullstellensatz. However, there are perfectly rotund varieties V that do not have depth 1 such as the intersection of generic hypersurfaces in G^3 used in the proof of Proposition 7.5. By Theorem 5.10, for such V there is no $(\bar{a}, e^{\bar{a}})$ in $\mathscr{U}$ that is generic in V over a field of definition of V, and hence $\mathscr{U}$ is not strongly exponentially-algebraically closed. □

9. Transcendence problems

Schanuel's conjecture has many consequences in transcendence theory. Ribenboim [2000, pp. 323–326] gives a few examples of easy consequences, one being that the numbers e, π, e^{π}, $\log \pi$, e^e, π^e, π^{π}, $\log 2$, 2^{π}, 2^e, 2^i, e^i, π^i, $\log 3$, $\log \log 2$, $(\log 2)^{\log 3}$, and $2^{\sqrt{2}}$ are all algebraically independent.

When Lang [1966, p. 31] first published Schanuel's conjecture, he wrote:

> From this statement, one would obtain most statements about algebraic independence of values of e^t and $\log t$ which one feels to be true.

We strengthen this empirical observation, and make it precise. To make a precise statement we need a definition.

Definition 9.1. A *particular transcendence problem* is a problem of the following form:

> Given complex numbers $a_1, \ldots, a_n$ in some way, what is the transcendence degree of the subfield $\mathbb{Q}(a_1, \ldots, a_n)$ of $\mathbb{C}$?

Since we are concerned with open problems, the way in which the complex numbers are given is important. For example, it may be that e^e is a rational number r. The problem of finding the transcendence degree of $\mathbb{Q}(r)$ given r as an explicit rational number is not the same as the problem of finding the transcendence degree of $\mathbb{Q}(\exp(\exp(1)))$. This might lead one to ask how one is allowed to specify a complex number, but we do not need to address this question in any generality. Note that the example above of Ribenboim is a particular transcendence problem where all the numbers are explicitly constructed from the rationals $\mathbb{Q}$ by the operations of exponentiation, taking logarithms, taking roots of polynomials, and field operations.

Let $\mathbb{C}_0 = \mathrm{ecl}^{\mathbb{C}_{\exp}}(\varnothing)$ be the field of exponentially algebraic complex numbers. By Fact 5.4, for any $\bar{a} = (a_1, \ldots, a_n) \in \mathbb{C}^n$ such that no $\mathbb{Q}$-linear combination

of them lies in $\mathbb{C}_0$, we have $\mathrm{td}(\bar{a}, e^{\bar{a}}/\mathbb{C}_0) \geqslant n + 1$. Thus Schanuel's conjecture (for $\mathbb{C}_{\exp}$) is equivalent to its restriction to $\mathbb{C}_0$. Recall that SK embeds in $\mathbb{C}_0$, and so Schanuel's conjecture for $\mathbb{C}_0$ is equivalent to the assertion that $SK \lhd \mathbb{C}_0$. By Proposition 6.9, if $SK \lhd \mathbb{C}_0$ then $B_0 \cong \mathbb{C}_0^{\sim}$ (recall $B_0 = SK^{\sim}$). Thus Schanuel's conjecture is equivalent to the assertion that $\mathbb{C}_0$ embeds in B_0. If in addition $\mathbb{C}_0$ were strongly exponentially-algebraically closed, that is, $\mathbb{C}_0 = \mathbb{C}_0^{\sim}$, then there would be an isomorphism $\mathbb{C}_0 \cong B_0$. Since the automorphism group of B_0 is very large, such an isomorphism would be very far from being unique.

Theorem 9.2. *Schanuel's conjecture decides all particular transcendence problems where the complex numbers $a_1, \ldots, a_n \in \mathbb{C}$ are given by an explicit construction from $\mathbb{Q}$ by the operations of exponentiation, taking logarithms, taking roots of polynomials, field operations, and taking implicit solutions of systems of exponential polynomial equations.*

Proof. The conditions on the a_i are equivalent to them all lying in $\mathbb{C}_0$, that is, being exponentially algebraic complex numbers. Assuming Schanuel's conjecture, $\mathbb{C}_0$ embeds in B_0. Any explicit description of the a_i defines a finitely generated partial E-subfield F of B_0, the smallest one containing all the coefficients of the exponential polynomial equations used in the given descriptions of the a_i. F is necessarily strong in B_0, since it contains witnesses of all of its elements being exponentially algebraic. When taking logarithms or, more generally, taking implicit solutions of systems of equations, there are countably many solutions in B_0, but the homogeneity of B_0 for strong partial E-subfields (which follows from the Fraïssé theorem used in the proof of Proposition 6.9) shows that these choices do not affect the isomorphism type of F. Thus Schanuel's conjecture determines the isomorphism type of F as a partial E-field, and hence it determines the transcendence degree of its subfield $\mathbb{Q}(a_1, \ldots, a_n)$. $\qquad\square$

Note that if we do not allow taking implicit solutions of systems of exponential polynomial equations then the construction stays inside the field SK^{ELA}, and the proof depends only on Section 2 of this paper. In particular this covers the field SK^{EL}, which, under Schanuel's conjecture, is the field of all of what Chow [1999] calls *EL-numbers*, that is, those complex numbers that have a closed-form representation using $0, +, \cdot, -, \div$, exp and the principal branch of the logarithm.

The construction will produce a generating set $\bar{b}$ for $D(F)$, and polynomial equations with rational coefficients determining the locus V of $(\bar{b}, \exp(\bar{b}))$. If we had an algorithm to determine the $\mathbb{Q}$-linear relations holding on $\bar{b}$ and the multiplicative relations holding on $\exp(\bar{b})$, that would give an algorithm for answering particular transcendence problems of this form (assuming Schanuel's conjecture of course).

There are other transcendence problems that are more general in nature, for example the four exponentials conjecture that states that if $x_1, x_2, y_1, y_2 \in \mathbb{C}$ and

$\mathrm{ldim}_{\mathbb{Q}}(x_1, x_2) = \mathrm{ldim}_{\mathbb{Q}}(y_1, y_2) = 2$, then

$$\mathrm{td}(e^{x_1 y_1}, e^{x_1 y_2}, e^{x_2 y_1}, e^{x_2 y_2}) \geqslant 1.$$

The four exponentials conjecture is not a particular transcendence problem as defined above, but nonetheless it can easily be seen to follow from Schanuel's conjecture. So the statement of Theorem 9.2 is not a complete answer to formalising Lang's observation. Nonetheless, the method of proof above does apply. The four exponentials conjecture can be viewed as the conjunction of a set of particular transcendence problems, namely every specific instance of the problem. More generally, suppose $\mathcal{P}$ is a transcendence problem, such as the four exponentials conjecture, which asserts that some transcendence degree is large given suitable conditions (about exponentials, logarithms, and algebraic equations). Then either (every instance of) $\mathcal{P}$ is true in B_0 so it follows from Schanuel's conjecture that it is true in $\mathbb{C}_0$, or $\mathcal{P}$ is false in B_0, in which case, since B_0 is constructed in as free a way as possible, $\mathcal{P}$ cannot be true in any exponential field F (unless it is true trivially because the hypotheses are not satisfied by any numbers in F).

Connection with conjectures on periods. The two main conjectures about $\mathbb{C}_{\exp}$ are:

(1) Schanuel's conjecture, or equivalently, $\mathbb{C}_{\exp}$ embeds in $\mathbb{B}$, or equivalently, $\mathbb{C}_0$ embeds in B_0.

(2) $\mathbb{C}_{\exp}$ is strongly exponentially-algebraically closed, or equivalently, $\mathbb{C}_{\exp} = \mathbb{C}_{\exp}{}^{\sim}$.

Together, they form Zilber's conjecture that $\mathbb{C}_{\exp} \cong \mathbb{B}$ (at least assuming the continuum hypothesis, as discussed in the introduction). As discussed above, Schanuel's conjecture is equivalent to its restriction to $\mathbb{C}_0$. In the light of Lemma 7.2 and Proposition 7.4, the restriction of the conjecture to $\mathbb{C}_0$ is equivalent to the assertion that if a is an exponentially algebraic complex number then there is a unique reason for that, meaning a unique smallest finite-dimensional $\mathbb{Q}$-vector subspace $\lceil a \rceil$ of $\mathbb{C}$ containing a such that $\delta(\lceil a \rceil) = 0$.

This formulation of Schanuel's conjecture makes a visible connection with the conjecture of Kontsevich and Zagier [2001, Section 1.2] on periods. They conjecture that if a complex number is a period then there is a unique reason for that, up to three rules for manipulating integrals: additivity, change of variables, and Stokes' formula. Kontsevich and Zagier [2001, Section 4.1] give an alternative formulation of their conjecture. There is a canonical surjective homomorphism from a formal object, the vector space of effective periods, to the space of complex periods. The periods conjecture is equivalent to this homomorphism being an isomorphism. In the exponential case, the existence of automorphisms of B_0 means there can be no

canonical isomorphism from the formal object B_0 to $\mathbb{C}_0$. Furthermore, since the objects in question are fields rather than vector spaces, there cannot be a noninjective map between them so if the conjecture is false then there is no map at all from B_0 to $\mathbb{C}_0$, although one could repair this by taking suitable subrings of B_0 instead. Finally, the open question of strong exponential-algebraic closedness of $\mathbb{C}_0$ means that any map should go from $\mathbb{C}_0$ to B_0 rather than the other way round, or that the subrings of B_0 chosen should be restricted in some way. The power of the predimension method, as used in this paper, is that such considerations are not necessary.

The Kontsevich–Zagier conjecture does not imply Schanuel's conjecture, because for example e is (conjecturally) not a period. Even the expanded conjecture on exponential periods [Kontsevich and Zagier 2001, Section 4.3] does not say much about Schanuel's conjecture, because (again conjecturally) e^e is not an exponential period. Furthermore Schanuel's conjecture does not just refer to $\mathbb{C}_0$ but to all of $\mathbb{C}$ whereas periods form a countable subset of $\mathbb{C}$. André [2009, Section 4.4] has observed that the Kontsevich–Zagier conjecture is equivalent to Grothendieck's conjecture on periods, and André [2009, Section 5.8.1] himself proposed a conjecture that encompasses both Grothendieck's periods conjecture and Schanuel's conjecture, and applies to all of $\mathbb{C}$.

10. Open problems

We end with some open problems. Schanuel's conjecture is known to be very difficult, and the conjecture that $\mathbb{C}_{\exp}$ is strongly exponentially-algebraically closed is also widely open (even assuming Schanuel's conjecture). We suggest some questions about complex exponentiation that may be easier.

(1) Define an ELA-field F to be *locally finitely presented* if and only if every finitely generated ELA-subfield of F is finitely presented. Is $\mathbb{C}_{\exp}$ locally finitely presented?

(2) Is there *any* finitely presented exponential subfield of $\mathbb{C}_{\exp}$?

(3) Is there an exponential subfield C of $\mathbb{C}$, and a finitely presented proper extension of C realised inside $\mathbb{C}_0$, the subfield of exponentially algebraic numbers in $\mathbb{C}$? Since $\mathbb{C}_0 \lhd \mathbb{C}_{\exp}$, the question is resolved outside $\mathbb{C}_0$.

(4) Let $V \subseteq G^n(\mathbb{C})$ be perfectly rotund. From the Schanuel Nullstellensatz for $\mathbb{C}_{\exp}$ we see that if $n = 1$ then there is $(a, e^a) \in V$ in $\mathbb{C}_{\exp}$. How about $n = 2$, or $n = 3$? Indeed, for which V can one show there are any solutions in $\mathbb{C}_{\exp}$?

(5) Is there any perfectly rotund V that is not of depth 1 with $(a, e^a) \in V$ in $\mathbb{C}_{\exp}$?

An apparently difficult problem is to construct an ordered analogue of pseudo-exponentiation that should be conjecturally elementarily equivalent to the real exponential field $\mathbb{R}_{\exp}$. Since the real exponential function is determined just by it

being a homomorphism that is order-preserving, continuous, and by the cut in the reals of e, one would have to assume Schanuel's conjecture for $\mathbb{R}_{\exp}$ to construct an Archimedean model. The following problem is of the same nature, but may perhaps be more straightforward.

(6) Can the automorphism σ_0 on SK be extended to an automorphism of order 2 on a subfield of $B_{\aleph_0}$ larger than SK^E, such as SK^{EA}, SK^{ELA}, B_0, or even $B_{\aleph_0}$ itself, in such a way that the exponential map is order-preserving on the fixed field (which will necessarily be real-closed, and hence ordered)?

Mantova [2011] has shown that σ_0 can be extended to an automorphism of order 2 on any B_κ for $\kappa \leqslant 2^{\aleph_0}$, but in his constructions the exponential map is not order-preserving on the fixed field.

Finally, the predimension method used in this paper is very powerful, and can be extended beyond the exponential setting, for example to the exponential maps of semiabelian varieties [Kirby 2009] and to sufficiently generic holomorphic functions known as Liouville functions [Zilber 2002; Koiran 2003; Wilkie 2005]. The periods conjecture of André encompasses the first of these settings and also the Grothendieck–Kontsevich–Zagier periods conjecture.

(7) Is there a way to formulate André's conjecture as the nonnegativity of some predimension function, satisfying the essential properties such as the addition formula and submodularity?

Acknowledgements

References

[André 2009] Y. André, "Galois theory, motives and transcendental numbers", pp. 165–177 in *Renormalization and Galois theories*, edited by A. Connes et al., IRMA Lect. Math. Theor. Phys. **15**, Eur. Math. Soc., Zürich, 2009. MR 2011b:11101 Zbl 1219.11109

[Baldwin 2009] J. T. Baldwin, *Categoricity*, University Lecture Series **50**, American Mathematical Society, Providence, RI, 2009. MR 2010m:03068 Zbl 1183.03002

[Bays and Kirby 2013] M. Bays and J. Kirby, "Excellence and uncountable categoricity of Zilber's exponential fields", preprint, 2013. arXiv 1305.0493

[Bays and Zilber 2011] M. Bays and B. Zilber, "Covers of multiplicative groups of algebraically closed fields of arbitrary characteristic", *Bull. Lond. Math. Soc.* **43**:4 (2011), 689–702. MR 2012e:12008 Zbl 1229.12006

[Chow 1999] T. Y. Chow, "What is a closed-form number?", *Amer. Math. Monthly* **106**:5 (1999), 440–448. MR 2000e:11156 Zbl 1003.11063

[D'Aquino et al. 2010] P. D'Aquino, A. Macintyre, and G. Terzo, "Schanuel Nullstellensatz for Zilber fields", *Fund. Math.* **207**:2 (2010), 123–143. MR 2011c:03086 Zbl 1207.03044

[Henson and Rubel 1984] C. W. Henson and L. A. Rubel, "Some applications of Nevanlinna theory to mathematical logic: identities of exponential functions", *Trans. Amer. Math. Soc.* **282**:1 (1984), 1–32. MR 85h:03015 Zbl 0533.03015

[Hils 2012] M. Hils, "Generic automorphisms and green fields", *J. Lond. Math. Soc.* (2) **85**:1 (2012), 223–244. MR 2876317 Zbl 1243.03051

[Hodges 1993] W. Hodges, *Model theory*, Encyclopedia of Mathematics and its Applications **42**, Cambridge University Press, 1993. MR 94e:03002 Zbl 0789.03031

[Hrushovski 1993] E. Hrushovski, "A new strongly minimal set", *Ann. Pure Appl. Logic* **62**:2 (1993), 147–166. MR 94d:03064 Zbl 0804.03020

[Kirby 2009] J. Kirby, "The theory of the exponential differential equations of semiabelian varieties", *Selecta Math.* (*N.S.*) **15**:3 (2009), 445–486. MR 2011c:12005 Zbl 05640212

[Kirby 2010a] J. Kirby, "Exponential algebraicity in exponential fields", *Bull. Lond. Math. Soc.* **42**:5 (2010), 879–890. MR 2011k:03070 Zbl 1203.03050

[Kirby 2010b] J. Kirby, "On quasiminimal excellent classes", *J. Symbolic Logic* **75**:2 (2010), 551–564. MR 2011g:03081 Zbl 1192.03006

[Kirby et al. 2012] J. Kirby, A. Macintyre, and A. Onshuus, "The algebraic numbers definable in various exponential fields", *J. Inst. Math. Jussieu* **11**:4 (2012), 825–834. MR 2979823 Zbl 06101599

[Koiran 2003] P. Koiran, "The theory of Liouville functions", *J. Symbolic Logic* **68**:2 (2003), 353–365. MR 2004f:03065 Zbl 1059.03023

[Kontsevich and Zagier 2001] M. Kontsevich and D. Zagier, "Periods", pp. 771–808 in *Mathematics unlimited—2001 and beyond*, edited by B. Engquist and W. Schmid, Springer, Berlin, 2001. MR 2002i:11002 Zbl 1039.11002

[Lang 1966] S. Lang, *Introduction to transcendental numbers*, Addison-Wesley, Reading, MA, 1966. MR 35 #5397 Zbl 0144.04101

[Macintyre 1991] A. Macintyre, "Schanuel's conjecture and free exponential rings", *Ann. Pure Appl. Logic* **51**:3 (1991), 241–246. MR 92d:11071 Zbl 0724.13008

[Mantova 2011] V. Mantova, "Involutions on Zilber fields", *Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei* (9) *Mat. Appl.* **22**:2 (2011), 237–244. MR 2012h:03110 Zbl 1233.03040

[Marker 2006] D. Marker, "A remark on Zilber's pseudoexponentiation", *J. Symbolic Logic* **71**:3 (2006), 791–798. MR 2007d:03061 Zbl 1112.03029

[Ribenboim 2000] P. Ribenboim, *My numbers, my friends: Popular lectures on number theory*, Springer, New York, 2000. MR 2002d:11001 Zbl 0947.11001

[Shkop 2012] A. C. Shkop, "Henson and Rubel's theorem for Zilber's pseudoexponentiation", *J. Symbolic Logic* **77**:2 (2012), 423–432. MR 2963014 Zbl 06047769

[Wilkie 2005] A. J. Wilkie, "Liouville functions", pp. 383–391 in *Logic Colloquium 2000*, edited by R. Cori et al., Lect. Notes Log. **19**, Assoc. Symbol. Logic, Urbana, IL, 2005. MR 2006f:03060 Zbl 1098.03044

[Zilber 2002] B. Zilber, "A theory of a generic function with derivations", pp. 85–99 in *Logic and algebra*, edited by Y. Zhang, Contemp. Math. **302**, Amer. Math. Soc., Providence, RI, 2002. MR 2003g:03059 Zbl 1013.03043

[Zilber 2005] B. Zilber, "Pseudo-exponentiation on algebraically closed fields of characteristic zero", *Ann. Pure Appl. Logic* **132**:1 (2005), 67–95. MR 2006a:03051 Zbl 1076.03024

[Zilber 2006] B. Zilber, "Covers of the multiplicative group of an algebraically closed field of characteristic zero", *J. London Math. Soc.* (2) **74**:1 (2006), 41–58. MR 2008e:12003 Zbl 1104.03030

jonathan.kirby@uea.ac.uk   *School of Mathematics, University of East Anglia, Norwich Research Park, Norwich, NR4 7TJ, United Kingdom*
http://www.uea.ac.uk/~ccf09tku/

On a problem of Arnold:
The average multiplicative order
of a given integer

Pär Kurlberg and Carl Pomerance

For coprime integers g and n, let $\ell_g(n)$ denote the multiplicative order of g modulo n. Motivated by a conjecture of Arnold, we study the average of $\ell_g(n)$ as $n \leq x$ ranges over integers coprime to g, and x tending to infinity. Assuming the generalized Riemann Hypothesis, we show that this average is essentially as large as the average of the Carmichael lambda function. We also determine the asymptotics of the average of $\ell_g(p)$ as $p \leq x$ ranges over primes.

1. Introduction

Given coprime integers g and n with $n > 0$ and $|g| > 1$, let $\ell_g(n)$ denote the multiplicative order of g modulo n, that is, the smallest integer $k \geq 1$ such that $g^k \equiv 1 \bmod n$. For $x \geq 1$ an integer, let

$$T_g(x) := \frac{1}{x} \sum_{\substack{n \leq x \\ (n,g)=1}} \ell_g(n),$$

essentially the average multiplicative order of g. Arnold [2005] conjectured that if $|g| > 1$, then

$$T_g(x) \sim c(g) \frac{x}{\log x},$$

as $x \to \infty$, for some constant $c(g) > 0$. However, Shparlinski [2007] showed that if the generalized Riemann Hypothesis[1] (GRH) is true, then

$$T_g(x) \gg \frac{x}{\log x} \exp\big(C(g)(\log \log \log x)^{3/2}\big),$$

[1]What is needed is that the Riemann Hypothesis holds for Dedekind zeta functions $\zeta_{K_n}(s)$ for all $n > 1$, where K_n is the Kummer extension $\mathbb{Q}(e^{2\pi i/n}, g^{1/n})$.

where $C(g) > 0$. He also suggested that it should be possible to obtain, again assuming GRH, a lower bound of the form

$$T_g(x) \geq \frac{x}{\log x} \exp\big((\log\log\log x)^{2+o(1)}\big) \quad \text{as } x \to \infty.$$

Let

$$B = e^{-\gamma} \prod_p \Big(1 - \frac{1}{(p-1)^2(p+1)}\Big) = 0.3453720641\ldots, \tag{1}$$

the product being over primes, and where γ is the Euler–Mascheroni constant. The principal aim of this paper is to prove the following result.

Theorem 1. *Assuming the GRH,*

$$T_g(x) = \frac{x}{\log x} \exp\Big(\frac{B \log\log x}{\log\log\log x}(1+o(1))\Big) \quad \textit{as } x \to \infty,$$

uniformly in g with $1 < |g| \leq \log x$. The upper bound implicit in this result holds unconditionally.

Let $\lambda(n)$ denote the exponent of the group $(\mathbb{Z}/n\mathbb{Z})^\times$, which is commonly known as Carmichael's function. We have $\ell_g(n) \leq \lambda(n)$ when $(g, n) = 1$, so we immediately obtain that

$$T_g(x) \leq \frac{1}{x} \sum_{n \leq x} \lambda(n),$$

and it is via this inequality that we are able to unconditionally establish the upper bound implicit in Theorem 1. Indeed, Erdős, Pomerance, and Schmutz [Erdős et al. 1991] determined the average order of $\lambda(n)$ showing that, as $x \to \infty$,

$$\frac{1}{x} \sum_{n \leq x} \lambda(n) = \frac{x}{\log x} \exp\Big(\frac{B \log\log x}{\log\log\log x}(1+o(1))\Big). \tag{2}$$

Theorem 1 thus shows under assumption of the GRH that the mean values of $\lambda(n)$ and $\ell_g(n)$ are of a similar order of magnitude. We know, on assuming the GRH, that $\lambda(n)/\ell_g(n)$ is very small for almost all n (for instance, see [Kurlberg 2003; Li and Pomerance 2003]; in the latter paper it was in fact shown that $\lambda(n)/\ell_g(n) \leq (\log n)^{o(\log\log\log n)}$ as $n \to \infty$ on a set of relative asymptotic density 1 among integers coprime to g), so perhaps Theorem 1 is not very surprising. *However*, in [Erdős et al. 1991] it was also shown that the normal order of $\lambda(n)$ is quite a bit smaller than the average order: There exists a subset S of the positive integers of asymptotic density 1 such that for $n \in S$ and $n \to \infty$,

$$\lambda(n) = \frac{n}{(\log n)^{\log\log\log n + A + (\log\log\log n)^{-1+o(1)}}},$$

where $A > 0$ is an explicit constant. Thus the main contribution to the average of $\lambda(n)$ comes from a *density-zero subset* of the integers, and to obtain our result on the average multiplicative order, we must show that $\ell_g(n)$ is large for many n for which $\lambda(n)$ is large.

If one averages over g as well, then a result like our Theorem 1 holds unconditionally. In particular, it follows from [Luca and Shparlinski 2003, Theorem 6] that

$$\frac{1}{x^2} \sum_{n \le x} \sum_{\substack{1 < g < n \\ (g,n)=1}} \ell_g(n) = \frac{x}{\log x} \exp\left(\frac{B \log \log x}{\log \log \log x}(1 + o(1)) \right) \quad \text{as } x \to \infty.$$

We also note that our methods give that Theorem 1 still holds for $g = a/b$ a rational number, with uniform error for $|a|, |b| \le \log x$, and n ranging over integers coprime to ab.

1.1. *Averaging over prime moduli.*

We shall always have the letters p, q denoting prime numbers. Given a rational number $g \ne 0, \pm 1$ and a prime p not dividing the numerator or denominator of g, let $\ell_g(p)$ denote the multiplicative order of g modulo p. For simplicity, when p does divide the numerator or denominator of g, we let $\ell_g(p) = 1$.

Further, given $k \in \mathbb{Z}^+$, let

$$D_g(k) := \left[\mathbb{Q}(g^{1/k}, e^{2\pi i / k}) : \mathbb{Q} \right]$$

denote the degree of the Kummer extension obtained by taking the splitting field of $X^k - g$. Let $\text{rad}(k)$ denote the largest squarefree divisor of k and let $\omega(k)$ be the number of primes dividing $\text{rad}(k)$.

Theorem 2. *Given $g \in \mathbb{Q}$, $g \ne 0, \pm 1$, define*

$$c_g := \sum_{k=1}^{\infty} \frac{\phi(k) \, \text{rad}(k)(-1)^{\omega(k)}}{k^2 D_g(k)}.$$

The series for c_g converges absolutely, and, assuming the GRH,

$$\frac{1}{\pi(x)} \sum_{p \le x} \ell_g(p) = \tfrac{1}{2} c_g \cdot x + O\left(\frac{x}{(\log x)^{2-4/\log \log \log x}} \right).$$

Furthermore, with $g = a/b$, where $a, b \in \mathbb{Z}$, the error estimate holds uniformly for $|a|, |b| \le x$.

At the heart of our claims of uniformity, both in Theorems 1 and 2, is our Theorem 6 in Section 2.

Though perhaps not obvious from the definition, $c_g > 0$ for all $g \neq 0, \pm 1$. In order to determine c_g, define

$$c := \prod_p \left(1 - \frac{p}{p^3 - 1}\right) = 0.5759599689\ldots,$$

the product being over primes; c_g turns out to be a positive *rational* multiple of c. To sum the series that defines c_g we will need some further notation. For p a prime and $\alpha \in \mathbb{Q}^*$, let $v_p(\alpha)$ be the exponential p-valuation at α, that is, it is the integer for which $p^{-v_p(\alpha)}\alpha$ is invertible modulo p. Write $g = \pm g_0^h$, where h is a positive integer and $g_0 > 0$ is not an exact power of a rational number, and write $g_0 = g_1 g_2^2$, where g_1 is a squarefree integer and g_2 is a rational. Let $e = v_2(h)$ and define $\Delta(g) = g_1$ if $g_1 \equiv 1 \bmod 4$, and $\Delta(g) = 4g_1$ if $g_1 \equiv 2$ or $3 \bmod 4$. For $g > 0$, define $n = \mathrm{lcm}(2^{e+1}, \Delta(g))$. For $g < 0$, define $n = 2g_1$ if $e = 0$ and $g_1 \equiv 3 \bmod 4$, or $e = 1$ and $g_1 \equiv 2 \bmod 4$; let $n = \mathrm{lcm}(2^{e+2}, \Delta(g))$ otherwise.

Consider the multiplicative function $f(k) = (-1)^{\omega(k)}\mathrm{rad}(k)(h, k)/k^3$. We note that for p prime and $j \geq 1$,

$$f(p^j) = -p^{1-3j+\min(j, v_p(h))}.$$

Given an integer $t \geq 1$, define $F(p, t)$ and $F(p)$ by

$$F(p, t) := \sum_{j=0}^{t-1} f(p^j) \quad \text{and} \quad F(p) := \sum_{j=0}^{\infty} f(p^j).$$

In particular, we note that if $p \nmid h$, then

$$F(p) = 1 - \sum_{j=1}^{\infty} p^{1-3j} = 1 - \frac{p}{p^3 - 1}. \tag{3}$$

Proposition 3. *With notation as above, if $g < 0$ and $e > 0$, we have*

$$c_g = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - \frac{p}{p^3-1}} \cdot \left(1 - \frac{F(2, e+1) - 1}{2F(2)} + \prod_{p \mid n}\left(1 - \frac{F(p, v_p(n))}{F(p)}\right)\right);$$

otherwise

$$c_g = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - \frac{p}{p^3-1}} \cdot \left(1 + \prod_{p \mid n}\left(1 - \frac{F(p, v_p(n))}{F(p)}\right)\right).$$

For example, if $g = 2$, then $h = 1$, $e = 0$, and $n = 8$. Thus

$$c_2 = c \cdot \left(1 + 1 - \frac{F(2, 3)}{F(2)}\right) = c \cdot \left(2 - \frac{1 - 2/(2^1)^3 - 2/(2^2)^3}{1 - 2/(8 - 1)}\right) = c \cdot \frac{159}{160}.$$

We remark that the universal constant

$$c = \prod_p \left(1 - \frac{p}{p^3 - 1}\right)$$

is already present in the work of Stephens on prime divisors of recurrence sequences. Motivated by a conjecture of Laxton, Stephens [1976] showed that on GRH, the limit

$$\lim_{x \to \infty} \frac{1}{\pi(x)} \sum_{p \le x} \frac{\ell_g(p)}{p - 1}$$

exists and equals c times a rational correction factor depending on g. In fact, from the result it is easy deduce our Theorem 2 with a somewhat better error term. However, Stephens only treats integral g that are not powers, the error term is not uniform in g, and, as noted by Moree and Stevenhagen [2000], the correction factors must be adjusted in certain cases.

Theorem 2 might also be compared with the work of Pappalardi [1995]. In fact, his method suggests an alternate route to our Theorem 2, and would allow the upper bound

$$\frac{1}{\pi(x)} \sum_{p \le x} \ell_g(p) \le \tfrac{1}{2}(c_g + o(1))x,$$

as $x \to \infty$ to be established unconditionally. The advantage of our method is that it avoids computing the density of those primes for which g has a given index.

Finally, Theorem 2 should also be contrasted with the unconditional result of Luca [2005] that

$$\frac{1}{\pi(x)} \sum_{p \le x} \frac{1}{(p-1)^2} \sum_{g=1}^{p-1} \ell_g(p) = c + O(1/(\log x)^\kappa)$$

for any fixed $\kappa > 0$. By partial summation one can then obtain

$$\frac{1}{\pi(x)} \sum_{p \le x} \frac{1}{p-1} \sum_{g=1}^{p-1} \ell_g(p) \sim \tfrac{1}{2}c \cdot x \quad \text{as } x \to \infty,$$

a result that is more comparable to Theorem 2.

2. Some preliminary results

For an integer $m \ge 2$, we let $P(m)$ denote the largest prime dividing m, and we let $P(1) = 1$.

Given a rational number $g \ne 0, \pm 1$, we recall the notation h, e, n described in Section 1.1, and for a positive integer k, we recall that $D_g(k)$ is the degree of the splitting field of $X^k - g$ over $\mathbb{Q}$. We record a result of Wagstaff on $D_g(k)$; see

Proposition 4.1 and the second paragraph in the proof of Theorem 2.2 in [Wagstaff 1982].

Proposition 4. *With notation as above,*

$$D_g(k) = \frac{\phi(k) \cdot k}{(k, h) \cdot \epsilon_g(k)}, \tag{4}$$

where ϕ is Euler's function and $\epsilon_g(k)$ is defined as follows: If $g > 0$, then

$$\epsilon_g(k) := \begin{cases} 2 & \text{if } n \mid k, \\ 1 & \text{if } n \nmid k. \end{cases}$$

If $g < 0$, then

$$\epsilon_g(k) := \begin{cases} 2 & \text{if } n \mid k, \\ \frac{1}{2} & \text{if } 2 \mid k \text{ and } 2^{e+1} \nmid k, \\ 1 & \text{otherwise}. \end{cases}$$

We also record a GRH-conditional version of the Chebotarev density theorem for Kummerian fields over $\mathbb{Q}$; see [Hooley 1967, Section 5; Lagarias and Odlyzko 1977, Theorem 1]. Let $i_g(p) = (p - 1)/\ell_g(p)$, the index of $\langle g \rangle$ in $(\mathbb{Z}/p\mathbb{Z})^*$ when $g \in (\mathbb{Z}/p\mathbb{Z})^*$.

Theorem 5. *Assume the GRH. Suppose $g = a/b \neq 0, \pm 1$, where a, b are integers of absolute value at most x. For each integer $k \leq x$, the number of primes $p \leq x$ for which $k \mid i_g(p)$ is*

$$\frac{1}{D_g(k)} \pi(x) + O(x^{1/2} \log x).$$

Note that $k \mid i_g(p)$ if and only if $x^k - g$ splits completely modulo p. Also note that the trivial bound x/k is majorized by the error term in Theorem 5 when $k \geq x^{1/2}/\log x$. In fact, the error term majorizes the main term for $k \geq x^{1/4}$.

We will need the following *uniform* version of [Kurlberg and Pomerance 2005, Theorem 23].

Theorem 6. *If the GRH is true, then for x, L with $1 \leq L \leq \log x$ and $g = a/b \neq 0, \pm 1$, where a, b are integers with $|a|, |b| \leq x$, we have*

$$\left| \left\{ p \leq x : \ell_g(p) \leq \frac{p - 1}{L} \right\} \right| \ll \frac{\pi(x)}{L} \cdot \frac{h\tau(h)}{\phi(h)} + \frac{x \log \log x}{\log^2 x}$$

uniformly, where $\tau(h)$ is the number of divisors of h.

Proof. Since the proof is rather similar to the proofs of the main theorem in [Hooley 1967], Theorem 2 in [Kurlberg 2003], and Theorem 23 in [Kurlberg and Pomerance 2005], we only give a brief outline. We see that $\ell_g(p) \leq (p - 1)/L$ implies that $i_g(p) \geq L$. Further, in the case that $p \mid ab$, where we are defining $\ell_g(p) = 1$ and hence $i_g(p) = p - 1$, the number of primes p is $O(\log x)$. So we assume that $p \nmid ab$.

First step: Consider primes $p \leq x$ such that $i_g(p) > x^{1/2} \log^2 x$. Such a prime p divides $a^k - b^k$ for some positive integer $k < x^{1/2}/\log^2 x$. Since $\omega(|a^k - b^k|) \ll k \log x$, it follows that the number of primes p in this case is

$$O((x^{1/2}/\log^2 x)^2 \log x) = O(x/\log^3 x).$$

Second step: Consider primes p such that $q \mid i_g(p)$ for some prime q in the interval $I := [x^{1/2}/\log^2 x, x^{1/2} \log^2 x]$. We may bound this by considering primes $p \leq x$ such that $p \equiv 1 \mod q$ for some prime $q \in I$. The Brun–Titchmarsh inequality then gives that the number of such primes p is at most a constant times

$$\sum_{q \in I} \frac{x}{\phi(q) \log(x/q)} \ll \frac{x}{\log x} \sum_{q \in I} \frac{1}{q} \ll \frac{x \log \log x}{\log^2 x}.$$

Third step: Now consider primes p such that $q \mid i_g(p)$ for some prime q in the interval $[L, x^{1/2}/\log^2 x)$. In this range we use Proposition 4 and Theorem 5 to get on the GRH that

$$|\{p \leq x : q \mid i_g(p)\}| \ll \frac{\pi(x)(q, h)}{q\phi(q)} + x^{1/2} \log x.$$

Summing over primes q, we find that the number of such p is bounded by a constant times

$$\sum_{q \in [L, x^{1/2}/\log^2 x)} \left(\frac{\pi(x)(q, h)}{q^2} + x^{1/2} \log x \right) \ll \frac{\pi(x)\omega(h)}{L} + \frac{x}{\log^2 x}.$$

Fourth step: For the remaining primes p, any prime divisor $q \mid i_g(p)$ is smaller than L. Hence $i_g(p)$ must be divisible by some integer d in the interval $[L, L^2]$. By Proposition 4 and Theorem 5, assuming the GRH, we have

$$\left|\{p \leq x : d \mid i_g(p)\}\right| \leq 2\frac{\pi(x)(d, h)}{d\phi(d)} + O(x^{1/2} \log x). \tag{5}$$

Hence the total number of such p is bounded by

$$\sum_{d \in [L, L^2]} \left(2\frac{\pi(x)(d, h)}{d\phi(d)} + O(x^{1/2} \log x) \right) \ll \frac{\pi(x)}{L} \frac{h\tau(h)}{\phi(h)},$$

where the last estimate follows from

$$\sum_{d \in [L, L^2]} \frac{(d, h)}{d\phi(d)} \leq \sum_{m \mid h} \sum_{\substack{d \in [L, L^2] \\ m \mid d}} \frac{m}{d\phi(d)} \leq \sum_{m \mid h} \sum_{k \geq L/m} \frac{1}{\phi(m)k\phi(k)}$$

$$\ll \sum_{m \mid h} \frac{m}{L\phi(m)} = \frac{h}{L\phi(h)} \sum_{m \mid h} \frac{m}{\phi(m)} \cdot \frac{\phi(h)}{h} \leq \frac{h\tau(h)}{L\phi(h)}. \tag{6}$$

Here we used the bound

$$\sum_{k \geq T} \frac{1}{k\phi(k)} \ll 1/T$$

for $T > 0$, which follows by an elementary argument from the bound

$$\sum_{k \geq T} \frac{1}{k^2} \ll 1/T$$

and the identity

$$k/\phi(k) = \sum_{j \mid k} \frac{\mu^2(j)}{\phi(j)}.$$

Indeed,

$$\sum_{k \geq T} \frac{1}{k\phi(k)} = \sum_{j} \frac{\mu^2(j)}{\phi(j)j^2} \sum_{l \geq T/j} \frac{1}{l^2} \ll \frac{1}{T} \sum_{j \leq T} \frac{1}{\phi(j)j} + \sum_{j > T} \frac{1}{j^2} \ll \frac{1}{T}. \qquad \square$$

Corollary 7. *Assume the GRH is true. Let $m \geq 2$ be an integer and $x \geq 10^7$ a real number. Let $y = \log\log x$ and assume that $m \leq \log y / \log\log y$. Let $g = a/b \neq 0, \pm 1$, where a, b are integers with $|a|, |b| \leq \exp((\log x)^{3/m})$, and let h be as above. Then uniformly,*

$$\sum_{\substack{p \leq x \\ P(i_g(p)) > m}} \frac{1}{p} \ll y\left(\frac{1}{m} + \sum_{\substack{q \mid h \\ q > m}} \frac{1}{q}\right).$$

Proof. This result is more a corollary of the proof of Theorem 6 than its statement. We consider intervals $I_j := (e^j, e^{j+1}]$ for $j \leq \log x$, with j a nonnegative integer. The sum of reciprocals of all primes $p \leq \exp((\log x)^{1/m})$ is $y/m + O(1)$, so this contribution to the sum is under control. We thus may restrict to the consideration of primes $p \in I_j$ for $j > (\log x)^{1/m}$. For such an integer j, let $t = e^{j+1}$. If $q \mid i_g(p)$ for some prime $q > t^{1/2} \log^2 t$, then $\ell_g(p) \leq t^{1/2}/\log^2 t$, and the number of such primes is

$$O\left(\sum_{k \leq t^{1/2}/\log^2 t} k \log|ab|\right) = O(t \log|ab|/\log^4 t),$$

so that the sum of their reciprocals is $O(\log|ab|/\log^4 t) = O((\log x)^{3/m}/j^4)$. Summing this for $j > (\log x)^{1/m}$, we get $O(1)$, which is acceptable.

For $J := (t^{1/2}/\log^2 t, t^{1/2} \log^2 t]$, with $t = e^{j+1}$, we have that the reciprocal sum of the primes $p \in I_j$ with some $q \in J$ dividing $i_g(p)$ (so that $q \mid p - 1$) is $O(\log\log t/\log^2 t) = O(\log j/j^2)$. Summing this for $j > (\log x)^{1/m}$ is $o(1)$ as $x \to \infty$ and is acceptable.

For $q \leq t^{1/2}/\log^2 t$ we need the GRH. As in the proof of Theorem 6, the number of primes $p \in I_j$ with $q \mid i_g(p)$ is bounded by a constant times

$$\frac{t}{\log t} \frac{(q, h)}{q^2} + t^{1/2} \log t.$$

Thus, the reciprocal sum of these primes p is

$$O\left(\frac{(q,h)}{q^2 \log t} + \frac{\log t}{t^{1/2}}\right) = O\left(\frac{(q,h)}{q^2 j} + \frac{j}{e^{j/2}}\right).$$

We sum this expression over primes q with $m < q \ll e^{j/2}/j^2$, getting

$$O\left(\frac{1}{jm \log m} + \frac{1}{j} \sum_{q \mid h, q > m} \frac{1}{q} + \frac{1}{j^2}\right).$$

Summing on $j \leq \log x$ completes the proof.   □

3. Proof of Theorem 1

Let x be large and let g be an integer with $1 < |g| \leq \log x$. Define

$$y = \log \log x, \quad m = \lfloor y / \log^3 y \rfloor, \quad D = m!, \quad S_k = \{p \leq x : (p-1, D) = 2k\}.$$

Then $S_1, S_2, \ldots, S_{D/2}$ are disjoint sets of primes whose union equals $\{2 < p \leq x\}$. Let

$$\tilde{S}_k = \left\{p \in S_k : p \nmid g, \; \frac{p-1}{2k} \;\Big|\; \ell_g(p)\right\} \tag{7}$$

be the subset of S_k, where $\ell_g(p)$ is "large". Note that if $k \leq \log y$, $p \in S_k \setminus \tilde{S}_k$, and $p \nmid g$, there is some prime $q > m$ with $q \mid (p-1)/\ell_g(p)$, so that $P(i_g(p)) > m$. Indeed, for x sufficiently large, we have $\log y \leq m/2$, and thus $k \leq \log y$ implies that each prime dividing D also divides $D/(2k)$, so that $(p-1, D) = 2k$ implies that the least prime factor of $(p-1)/(2k)$ exceeds m.

Thus, from Theorem 6,

$$|S_k \setminus \tilde{S}_k| \leq \left|\{p \leq x : \ell_g(p) < p/m\}\right| + \sum_{p \mid g} 1 \ll \frac{\pi(x)}{m} \cdot \frac{h\tau(h)}{\phi(h)}$$

uniformly for $k \leq \log y$. Using this it is easy to see that S_k and $\tilde{S}_k$ are of similar size when k is small. However, we shall essentially measure the "size" of S_k or $\tilde{S}_k$ by the sum of the reciprocals of its members and for this we will use Corollary 7. We define

$$E_k := \sum_{\substack{p \in S_k \\ 1 < p^\alpha \leq x}} \frac{1}{p^\alpha} \quad \text{and} \quad \tilde{E}_k := \sum_{\substack{p \in \tilde{S}_k \\ 1 < p^\alpha \leq x}} \frac{1}{p^\alpha}.$$

By Lemma 1 of [Erdős et al. 1991],

$$E_k = \frac{y}{\log y} \cdot P_k \cdot (1 + o(1)) \tag{8}$$

uniformly for $k \leq \log^2 y$, where

$$P_k = \frac{e^{-\gamma}}{k} \prod_{q>2}\left(1 - \frac{1}{(q-1)^2}\right) \prod_{q \mid k,\, q>2} \frac{q-1}{q-2}. \tag{9}$$

Note that, with B given by (1),

$$\sum_{k=1}^{\infty} \frac{P_k}{2k} = B. \tag{10}$$

The next lemma shows that not much is lost when restricting to primes $p \in \tilde{S}_k$.

Lemma 8. *For $k \leq \log y$, we uniformly have*

$$\tilde{E}_k = E_k \cdot \left(1 + O\left(\frac{\log^5 y}{y}\right)\right).$$

Proof. By (8) and (9), we have

$$E_k \gg \frac{y}{k \log y} \geq \frac{y}{\log^2 y}, \tag{11}$$

and it is thus sufficient to show that $\sum_{p \in S_k \setminus \tilde{S}_k} 1/p \ll \log^3 y$ since the contribution from prime powers p^α for $\alpha \geq 2$ is $O(1)$. As we have seen, if $k \leq \log y$ and $p \in S_k \setminus \tilde{S}_k$, then either $p \mid g$ or $P(i_g(p)) > m$. Hence, using Corollary 7 and noting that the hypothesis $|g| \leq \log x$ implies that $h \ll y$ and so h has at most one prime factor $q > m$, we have

$$\sum_{p \in E_k \setminus \tilde{E}_k} \frac{1}{p} \ll \frac{y}{m} = \frac{y}{\lfloor y/\log^3 y\rfloor} \ll \log^3 y. \qquad \square$$

Lemma 9. *We have*

$$\sum_{k \leq \log y} \frac{E_k}{2k} = \frac{By}{\log y}(1 + o(1)),$$

where B is given by (1).

Proof. This follows immediately from (8), (9), and (10). $\square$

Given a vector $j = (j_1, j_2, \ldots, j_{D/2})$ with each $j_i \in \mathbb{Z}_{\geq 0}$, let

$$\|j\| := j_1 + j_2 + \cdots + j_{D/2}.$$

Paralleling the notation $\Omega_i(x; j)$ from [Erdős et al. 1991], we let

- $\tilde{\Omega}_1(x; j)$ be the set of integers that can be formed by taking products of $v = \|j\|$ distinct primes $p_1, p_2, \ldots, p_v$ so that
    - for each i, $p_i < x^{1/y^3}$, and
    - the first j_1 primes are in $\tilde{S}_1$, the next j_2 are in $\tilde{S}_2$, etc.;

- $\tilde{\Omega}_2(x; \boldsymbol{j})$ be the set of integers $u = p_1 p_2 \cdots p_v \in \tilde{\Omega}_1(x; \boldsymbol{j})$ where (p_i-1, p_j-1) divides D for all $i \neq j$;

- $\tilde{\Omega}_3(x; \boldsymbol{j})$ be the set of integers of the form $n = up$, where $u \in \tilde{\Omega}_2(x; \boldsymbol{j})$ and p satisfies $(p - 1, D) = 2$, $\max(x/(2u), x^{1/y}) < p \leq x/u$ and $\ell_g(p) > p/y^2$;

- $\tilde{\Omega}_4(x; \boldsymbol{j})$ be the set of integers $n = (p_1 p_2 \cdots p_v)p$ in $\tilde{\Omega}_3(x; \boldsymbol{j})$ with the additional property that $(p - 1, p_i - 1) = 2$ for all i.

(In the third bullet, note that the max is not strictly necessary since when x is sufficiently large, $x/(2u) > x^{1/y}$.)

3.1. *Some lemmas.* We shall also need the following analogues of [Erdős et al. 1991, Lemmas 2–4]. Let

$$J := \{j : 0 \leq j_k \leq E_k/k \text{ for } k \leq \log y, \text{ and } j_k = 0 \text{ for } k > \log y\}.$$

Lemma 10. *If $j \in J$, $n \in \tilde{\Omega}_4(x; j)$, and $x \geq x_1$, then*

$$\ell_g(n) \geq c_1 \frac{x}{y^3} \prod_{k \leq \log y} (2k)^{-j_k},$$

where $x_1, c_1 > 0$ are absolute constants.

Proof. Suppose that $n = (p_1 p_2 \cdots p_v)p \in \tilde{\Omega}_4(x; j)$. Let $d_i = (p_i - 1, D)$, and let $u_i := (p_i - 1)/d_i$. By (7), u_i divides $\ell_g(p_i)$ for all i, and by the definition of $\tilde{\Omega}_3(x; j)$ we also have $\ell_g(p) > p/y^2$. Since $(p-1)/2$ is coprime to $(p_i - 1)/2$ for each i and each $(p_i - 1, p_j - 1) \mid D$ for $i \neq j$, we have $u_1, \ldots, u_v, p - 1$ pairwise coprime. But

$$\ell_g(n) = \text{lcm}[\ell_g(p_1), \ell_g(p_2), \ldots, \ell_g(p_v), \ell_g(p)],$$

so we find that, using the minimal order of Euler's function and $\ell_g(p) > p/y^2$,

$$\ell_g(n) \geq u_1 u_2 \cdots u_v \ell_g(p) \geq \frac{\phi(n)}{y^2 \cdot \prod_{i=1}^{v} d_i}$$

$$\gg \frac{n}{y^2 \cdot \log \log n \cdot \prod_{k=1}^{l} (2k)^{j_k}} \gg \frac{x}{y^3 \cdot \prod_{k=1}^{l} (2k)^{j_k}},$$

where we recall that $d_i = (p_i - 1, D) = 2k$ if $p_i \in \tilde{S}_k$, and that $n \in \tilde{\Omega}_4(x; j)$ implies that $n > x/2$. $\square$

Lemma 11. *If $j \in J$, $u \in \tilde{\Omega}_2(x; j)$, and $x \geq x_2$, then*

$$\left| \{p : up \in \tilde{\Omega}_4(x; j)\} \right| > c_2 x/(uy \log x),$$

where $x_2, c_2 > 0$ are absolute constants.

Proof. Note that $\|j\| \leq \sum_{k=1}^{l} E_k/k \ll y/\log y$ for $j \in J$, by (8) and (9). For such vectors j, Lemma 3 of [Erdős et al. 1991] implies that the number of primes p with $\max(x/2u, x^{1/y}) < p \leq x/u$, $(p-1, D) = 2$, and $(p-1, p_i - 1) = 2$ for all $p_i \mid u$ is $\gg x/(uy \log x)$. Thus it suffices to show that

$$\left|\{p \leq x/u : (p-1, D) = 2, \ \ell_g(p) \leq p/y^2\}\right| = o(x/(uy \log x)).$$

As we have seen, $\|j\| \ll y/\log y$ for $j \in J$, so that $u \in \tilde{\Omega}_2(x; j)$ has $u \leq x^{1/y^2}$ for all large x. Thus, Theorem 6 implies that

$$\sum_{\substack{p \leq x/u \\ \ell_g(p) \leq p/y^2}} 1 \ll \frac{\pi(x/u)}{y^2} \ll \frac{x}{uy^2 \log x} = o\left(\frac{x}{uy \log x}\right).$$

The result follows. $\qquad\square$

Lemma 12. *If $j \in J$, then for $x \geq x_3$,*

$$\sum_{u \in \tilde{\Omega}_2(x; j)} \frac{1}{u} > \exp\left(\frac{-c_3 y \log \log y}{\log^2 y}\right) \prod_{k \leq \log y} \frac{E_k^{j_k}}{j_k!},$$

where $x_3, c_3 > 0$ are absolute constants.

Proof. The sum in the lemma is equal to

$$\frac{1}{j_1! j_2! \cdots j_{\lfloor \log y \rfloor}!} \sum_{\langle p_1, p_2, \ldots, p_v \rangle} \frac{1}{p_1 p_2 \cdots p_v},$$

where the sum is over sequences of distinct primes for which the first j_1 are in $\tilde{S}_1$, the next j_2 are in $\tilde{S}_2$, and so on, and also each $(p_i - 1, p_j - 1) \mid D$ for $i \neq j$. Such a sum is estimated from below in Lemma 4 of [Erdős et al. 1991] but without the extra conditions that differentiate $\tilde{S}_k$ from S_k. The key prime reciprocal sum there is estimated on pages 381–383 to be

$$E_k\left(1 + O\left(\frac{\log \log y}{\log y}\right)\right).$$

In our case we have the extra conditions that $p \nmid g$ and $(p-1)/2k \mid \ell_g(p)$, which alters the sum by a factor of $1 + O(\log^5 y/y)$ by Lemma 8. But the factor $1 + O(\log^5 y/y)$ is negligible compared with the factor $1 + O(\log \log y/\log y)$, so we have exactly the same expression in our current case. $\qquad\square$

3.2. *Conclusion.* For brevity, let $l = \lfloor \log y \rfloor$. We clearly have

$$T_g(x) \geq \frac{1}{x} \sum_{j \in J} \sum_{n \in \tilde{\Omega}_4(x; j)} \ell_g(n).$$

By Lemma 10, we thus have

$$T_g(x) \gg \frac{1}{y^3} \sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \sum_{n \in \tilde{\Omega}_4(x;j)} 1.$$

Now,

$$\sum_{n \in \tilde{\Omega}_4(x;j)} 1 = \sum_{u \in \tilde{\Omega}_2(x;j)} \sum_{up \in \tilde{\Omega}_4(x;j)} 1,$$

and by Lemma 11, this is

$$\gg \sum_{u \in \tilde{\Omega}_2(x;j)} \frac{x}{uy \log x},$$

which in turn by Lemma 12 is

$$\gg \frac{x}{y \log x} \exp\left(\frac{-c_3 y \log \log y}{\log^2 y}\right) \prod_{k=1}^{l} \frac{E_k^{j_k}}{j_k!}.$$

Hence

$$T_g(x) \gg \frac{x}{y^4 \log x} \exp\left(\frac{-c_3 y \log \log y}{\log^2 y}\right) \sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!}.$$

Now,

$$\sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!} = \prod_{k=1}^{l} \left(\sum_{j_k=0}^{[E_k/k]} \frac{(E_k/2k)^{j_k}}{j_k!} \right).$$

Note that $\sum_{j=0}^{2w} w^j/j! > e^w/2$ for $w \geq 1$ and also that $E_k/2k \geq 1$ for x sufficiently large, as $E_k \gg y/(k \log y)$ by (11). Thus,

$$\sum_{j \in J} \prod_{k=1}^{l} (2k)^{-j_k} \frac{E_k^{j_k}}{j_k!} > 2^{-l} \exp\left(\sum_{k=1}^{l} \frac{E_k}{2k} \right).$$

Hence

$$T_g(x) \gg \frac{x}{y^4 \log x} \exp\left(\frac{-c_3 y \log \log y}{\log^2 y}\right) 2^{-l} \exp\left(\sum_{k=1}^{l} \frac{E_k}{2k} \right).$$

By Lemma 9 we thus have the lower bound in the theorem. The proof is concluded.

4. Averaging over prime moduli — the proofs

4.1. *Proof of Theorem 2.* Let $z = \log x$ and abbreviate $\ell_g(p)$ and $i_g(p)$ by $\ell(p)$ and $i(p)$, respectively. We have

$$\sum_{p \le x} \ell(p) = \sum_{\substack{p \le x \\ i(p) \le z}} \ell(p) + \sum_{\substack{p \le x \\ i(p) > z}} \ell(p) = A + E,$$

say. Writing $\ell(p) = (p-1)/i(p)$ and using the identity $1/i(p) = \sum_{uv|i(p)} \mu(v)/u$, we find that

$$A = \sum_{\substack{p \le x \\ i(p) \le z}} (p-1) \sum_{uv \,|\, i(p)} \frac{\mu(v)}{u}$$

$$= \sum_{p \le x} (p-1) \sum_{\substack{uv \,|\, i(p) \\ uv \le z}} \frac{\mu(v)}{u} - \sum_{\substack{p \le x \\ i(p) > z}} (p-1) \sum_{\substack{uv \,|\, i(p) \\ uv \le z}} \frac{\mu(v)}{u}$$

$$= A_1 - E_1,$$

say. The main term A_1 is

$$A_1 = \sum_{uv \le z} \frac{\mu(v)}{u} \sum_{\substack{p \le x \\ uv | i(p)}} (p-1).$$

By a simple partial summation using Theorem 5, the inner sum here is

$$\frac{\text{Li}(x^2)}{D_g(uv)} + O(x^{3/2} \log x)$$

assuming the GRH. Thus,

$$A_1 = \text{Li}(x^2) \left(\sum_{uv \le z} \frac{\mu(v)}{u D_g(uv)} \right) + O\left(x^{3/2} \log x \sum_{n \le z} \left| \sum_{uv=n} \frac{\mu(v)}{u} \right| \right).$$

The inner sum in the O-term is bounded by $\phi(n)/n$, so the O-term is $O(x^{3/2} \log^2 x)$. Recalling that $\text{rad}(n)$ denotes the largest squarefree divisor of n, we note that $\sum_{v|k} \mu(v)v = \prod_{p|k}(1-p) = (-1)^{\omega(k)} \phi(\text{rad}(k))$, and hence

$$\sum_{uv=k} \frac{\mu(v)}{u D_g(uv)} = \sum_{v|k} \frac{\mu(v)v}{D_g(k)k} = \frac{(-1)^{\omega(k)} \phi(\text{rad}(k))}{D_g(k)k}.$$

On noting that $\phi(\text{rad}(k)) = \phi(k)\text{rad}(k)/k$, we have

$$\sum_{u,v} \frac{\mu(v)}{u D_g(uv)} = \sum_{k \ge 1} \frac{(-1)^{\omega(k)} \text{rad}(k)\phi(k)}{D_g(k)k^2} = c_g.$$

Thus, with $\psi(h) := h\tau(h)/\phi(h)$,

$$\sum_{uv \leq z} \frac{\mu(v)}{uv D_g(uv)} = c_g - \sum_{k>z} \frac{(-1)^{\omega(k)} \mathrm{rad}(k)\phi(k)}{D_g(k)k^2} = c_g + O(\psi(h)/z),$$

by Proposition 4 and the same argument as in the fourth step of the proof of Theorem 6 (in particular, see (6)). It now follows that

$$A_1 = \mathrm{Li}(x^2)(c_g + O(\psi(h)/z)).$$

It remains to estimate the two error terms E, E_1. Using Theorem 6, we have

$$E \ll \frac{x}{z} \cdot \frac{x \log \log x}{\log^2 x} \psi(h) \ll \frac{x^2 \psi(h)}{\log^2 x}.$$

Toward estimating E_1, we note that

$$f_z(n) := \left| \sum_{\substack{uv \mid n \\ uv \leq z}} \frac{\mu(v)}{u} \right| \leq \sum_{\substack{d \mid n \\ d \leq z}} \left| \sum_{v \mid d} \frac{\mu(v)v}{d} \right| = \sum_{\substack{d \mid n \\ d \leq z}} \frac{\phi(\mathrm{rad}(d))}{d} \leq z.$$

Further, from the last sum we get

$$f_z(n) \leq \prod_{\substack{p^a \| n \\ p \leq z}} \left(1 + \frac{p-1}{p} + \cdots + \frac{p-1}{p^a} \right) < 2^{\omega(n_z)},$$

where n_z denotes the largest divisor of n composed of primes in $[1, z]$. We have

$$|E_1| \leq \sum_{\substack{p \leq x \\ i(p)>z}} (p-1) f_z(i(p)) \leq x \sum_{\substack{p \leq x \\ i(p)>z}} f_z(i(p)).$$

Let $w := 4 \log z / \log \log z$. We break the sum above into three possibly overlapping parts:

$$E_{1,1} := x \sum_{\substack{p \leq x \\ i(p)>z \\ \omega(i(p)_z) \leq w}} f_z(i(p)), \quad E_{1,2} := x \sum_{\substack{p \leq x \\ z < i(p) \leq x^{1/2} \log^2 x \\ \omega(i(p)_z)>w}} f_z(i(p)),$$

$$E_{1,3} := x \sum_{\substack{p \leq x \\ i(p)>x^{1/2} \log^2 x}} f_z(i(p)).$$

Using Theorem 6, we have

$$E_{1,1} \leq x 2^w \sum_{\substack{p \leq x \\ i(p)>z}} 1 \ll 2^w \psi(h) \frac{x^2 \log \log x}{\log^2 x}.$$

The estimate for $E_{1,3}$ is similarly brief, this time using the "first step" in the proof of [Theorem 6](). We have

$$E_{1,3} \leq xz \sum_{\substack{p \leq x \\ i(p) > x^{1/2} \log^2 x}} 1 \ll \frac{x^2}{\log^2 x}.$$

The estimate for $E_{1,2}$ takes a little work. By the Brun–Titchmarsh inequality,

$$E_{1,2} \leq xz \sum_{\substack{z < n \leq x^{1/2} \log^2 x \\ \omega(n_z) > w}} \pi(x; n, 1) \ll \frac{x^2 z}{\log x} \sum_{\substack{z < n \leq x^{1/2} \log^2 x \\ \omega(n_z) > w}} \frac{1}{\phi(n)}$$

$$\leq x^2 \sum_{\substack{P(m) \leq z \\ \omega(m) > w}} \frac{1}{\phi(m)} \sum_{n \leq x^{1/2} \log^2 x} \frac{1}{\phi(n)} \ll x^2 \log x \sum_{\substack{P(m) \leq z \\ \omega(m) > w}} \frac{1}{\phi(m)}.$$

This last sum is smaller than

$$\sum_{k > w} \frac{1}{k!} \left(\sum_{p \leq z} \left(\frac{1}{p-1} + \frac{1}{p(p-1)} + \cdots \right) \right)^k = \sum_{k > w} \frac{1}{k!} \left(\sum_{p \leq z} \frac{p}{(p-1)^2} \right)^k$$

$$= \sum_{k > w} \frac{1}{k!} (\log \log z + O(1))^k.$$

The terms in this series are decaying at least geometrically by a large factor, so by a weak form of Stirling's formula, we have

$$\sum_{\substack{P(m) \leq z \\ \omega(m) > w}} \frac{1}{m} \ll \exp\big(w \log \log \log z - w \log w + w + O(w/\log \log z) \big).$$

By our choice for w, this last expression is smaller than $\exp(-3 \log z) = (\log x)^{-3}$ for all large values of x. Hence, $E_{1,2} \ll x^2/\log^2 x$.

Noting that $\psi(h) \ll \tau(h) \log \log x$, we conclude that

$$\sum_{p \leq x} \ell(p) = A + E = A_1 + E + O\big(E_{1,1} + E_{1,2} + E_{1,3}\big)$$

$$= c_g \mathrm{Li}(x^2) + O\left(\frac{x^2}{\log^2 x} \big(\psi(h) + 2^w \psi(h) \log \log x + 1 + 1 \big) \right)$$

$$= c_g \mathrm{Li}(x^2) + O\left(2^w \tau(h) \cdot \frac{x^2 (\log \log x)^2}{\log^2 x} \right)$$

$$= \tfrac{1}{2} c_g x \pi(x) + O\left(\frac{x^2}{(\log x)^{2 - 4/\log \log \log x}} \right),$$

using that $\mathrm{Li}(x^2) = \frac{1}{2}x\pi(x) + O(x^2/\log^2 x)$, the definition of w, and $h \leq \log x$ together with Wigert's theorem for the maximal order of the divisor function $\tau(h)$. This completes the proof.

4.2. *Proof of Proposition 3.* We begin with the cases $g > 0$, or $g < 0$ and $e = 0$. Recalling that $D_g(k) = \phi(k)k/(\epsilon_g(k)(k,h))$, we find that

$$c_g = \sum_{k \geq 1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)\phi(k)}{D_g(k)k^2} = \sum_{k \geq 1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(k,h)\epsilon_g(k)}{k^3}. \quad (12)$$

Now, since $\epsilon_g(k)$ equals 1 if $n \nmid k$, and 2 otherwise, (12) equals

$$\sum_{k \geq 1} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(h,k)}{k^3} + \sum_{n \mid k} \frac{(-1)^{\omega(k)}\mathrm{rad}(k)(h,k)}{k^3} = \sum_{k \geq 1}(f(k) + f(kn)), \quad (13)$$

where the function $f(k) = (-1)^{\omega(k)}\mathrm{rad}(k)(h,k)/k^3$ is multiplicative.

If $p \nmid h$ and $j \geq 1$, we have $f(p^j) = -p/p^{3j}$. On the other hand, writing $h = \prod_{p \mid h} p^{e_{h,p}}$ we have $f(p^j) = -p^{1+\min(j,e_{h,p})}/p^{3j}$ for $p \mid h$ and $j \geq 1$. Since f is multiplicative,

$$\sum_{k \geq 1}(f(k) + f(kn)) = \sum_{k:\mathrm{rad}(k) \mid hn}(f(k) + f(kn)) \cdot \sum_{(k,hn)=1} f(k).$$

Now, for $p \nmid h$ and $j \geq 1$, we have $f(p^j) = -\mathrm{rad}(p^j)/p^{3j} = -p/p^{3j}$; hence

$$\sum_{j \geq 0} f(p^j) = 1 - \frac{p}{p^3(1 - 1/p^3)} = 1 - \frac{p}{p^3 - 1}$$

and thus

$$\sum_{(k,hn)=1} f(k) = \prod_{p \nmid hn} F(p) = \prod_{p \nmid hn}\left(1 - \frac{p}{p^3 - 1}\right) = \frac{c}{\prod_{p \mid hn}\left(1 - \frac{p}{p^3 - 1}\right)}.$$

Similarly, $\sum_{\mathrm{rad}(k) \mid hn} f(k) = \prod_{p \mid hn} F(p)$ and

$$\sum_{\mathrm{rad}(k) \mid hn} f(kn) = \prod_{p \mid hn}\left(\sum_{j \geq e_{n,p}} f(p^j)\right) = \prod_{p \mid hn}(F(p) - F(p, e_{n,p})).$$

Hence

$$\sum_{\mathrm{rad}(k) \mid hn} f(k) + \sum_{\mathrm{rad}(k) \mid hn} f(kn) = \prod_{p \mid hn} F(p) + \prod_{p \mid hn}(F(p) - F(p, e_{n,p}))$$

$$= \prod_{p \mid hn} F(p) \cdot \left(1 + \prod_{p \mid hn}\left(1 - \frac{F(p, e_{n,p})}{F(p)}\right)\right).$$

Thus

$$c_g = \frac{c}{\prod_{p \mid hn}(1 - \frac{p}{p^3-1})} \cdot \prod_{p \mid hn} F(p) \cdot \left(1 + \prod_{p \mid hn}\left(1 - \frac{F(p, e_{n,p})}{F(p)}\right)\right),$$

which, by (3), simplifies to

$$c_g = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - \frac{p}{p^3-1}} \cdot \left(1 + \prod_{p \mid hn}\left(1 - \frac{F(p, e_{n,p})}{F(p)}\right)\right).$$

The case $g < 0$ and $e > 0$ is similar: using the multiplicativity of f together with the definition of $\epsilon_g(k)$, we find that

$$
\begin{aligned}
c_g &= \sum_{k \geq 1}(f(k) + f(kn)) - \tfrac{1}{2}\sum_{j=1}^{e}\sum_{(k,2)=1} f(2^j k) \\
&= \prod_p F(p) + \prod_p (F(p) - F(p, e_{n,p})) - \tfrac{1}{2} \cdot (F(2, e+1) - 1) \cdot \prod_{p>2} F(p) \\
&= \prod_p F(p)\left(1 + \prod_{p \mid n}\left(1 - \frac{F(p, e_{n,p})}{F(p)}\right) - \frac{F(2, e+1) - 1}{2F(2)}\right).
\end{aligned}
$$

Again using the fact that

$$\prod_p F(p) = \prod_{p \nmid h}\left(1 - \frac{p}{p^3+1}\right)\prod_{p \mid h} F(p) = c \cdot \prod_{p \mid h} \frac{F(p)}{1 - p/(p^3+1)},$$

the proof is concluded.

Acknowledgments

References

[Arnold 2005] V. Arnold, "Number-theoretical turbulence in Fermat–Euler arithmetics and large Young diagrams geometry statistics", *J. Mathematical Fluid Mech.* **7**:suppl. 1 (2005), S4–S50. MR 2006g:11199 Zbl 1134.11344

[Erdős et al. 1991] P. Erdős, C. Pomerance, and E. Schmutz, "Carmichael's lambda function", *Acta Arith.* **58**:4 (1991), 363–385. MR 92g:11093 Zbl 0734.11047

[Hooley 1967] C. Hooley, "On Artin's conjecture", *J. Reine Angew. Math.* **225** (1967), 209–220. MR 34 #7445 Zbl 0221.10048

[Kurlberg 2003] P. Kurlberg, "On the order of unimodular matrices modulo integers", *Acta Arith.* **110**:2 (2003), 141–151. MR 2005a:11146 Zbl 1030.11048

[Kurlberg and Pomerance 2005] P. Kurlberg and C. Pomerance, "On the periods of the linear congruential and power generators", *Acta Arith.* **119**:2 (2005), 149–169. MR 2006k:11153 Zbl 1080.11059

[Lagarias and Odlyzko 1977] J. C. Lagarias and A. M. Odlyzko, "Effective versions of the Chebotarev density theorem", pp. 409–464 in *Algebraic number fields: L-functions and Galois properties* (Durham, 1975), edited by A. Fröhlich, Academic Press, London, 1977. MR 56 #5506 Zbl 0362.12011

[Li and Pomerance 2003] S. Li and C. Pomerance, "On generalizing Artin's conjecture on primitive roots to composite moduli", *J. Reine Angew. Math.* **556** (2003), 205–224. MR 2004c:11177 Zbl 1022.11049

[Luca 2005] F. Luca, "Some mean values related to average multiplicative orders of elements in finite fields", *Ramanujan J.* **9**:1-2 (2005), 33–44. MR 2006i:11111 Zbl 1155.11344

[Luca and Shparlinski 2003] F. Luca and I. E. Shparlinski, "Average multiplicative orders of elements modulo n", *Acta Arith.* **109**:4 (2003), 387–411. MR 2004i:11113 Zbl 1043.11067

[Moree and Stevenhagen 2000] P. Moree and P. Stevenhagen, "A two-variable Artin conjecture", *J. Number Theory* **85**:2 (2000), 291–304. MR 2001k:11188 Zbl 0966.11042

[Pappalardi 1995] F. Pappalardi, "On Hooley's theorem with weights", *Rend. Sem. Mat. Univ. Politec. Torino* **53**:4 (1995), 375–388. MR 98c:11102 Zbl 0883.11042

[Shparlinski 2007] I. E. Shparlinski, "On some dynamical systems in finite fields and residue rings", *Discrete Contin. Dyn. Syst.* **17**:4 (2007), 901–917. MR 2007j:11098 Zbl 1127.11052

[Stephens 1976] P. J. Stephens, "Prime divisors of second-order linear recurrences, I", *J. Number Theory* **8**:3 (1976), 313–332. MR 54 #5142 Zbl 0334.10018

[Wagstaff 1982] S. S. Wagstaff, Jr., "Pseudoprimes and a generalization of Artin's conjecture", *Acta Arith.* **41**:2 (1982), 141–150. MR 83m:10004 Zbl 0496.10001

kurlberg@math.kth.se   *Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden*
http://www.math.kth.se/~kurlberg/

carl.pomerance@dartmouth.edu *Mathematics Department, Kemeny Hall, Dartmouth College, Hanover NH 03755, United States*
www.math.dartmouth.edu/~carlp

An analogue of Sturm's theorem for Hilbert modular forms

Yuuki Takai

In this paper, we consider congruences of Hilbert modular forms. Sturm showed that mod ℓ elliptic modular forms of weight k and level $\Gamma_1(N)$ are determined by the first $(k/12)[\Gamma_1(1) : \Gamma_1(N)]$ mod ℓ Fourier coefficients. We prove an analogue of Sturm's result for Hilbert modular forms associated to totally real number fields. The proof uses the positivity of ample line bundles on toroidal compactifications of Hilbert modular varieties.

1. Introduction

In this paper, we consider congruences of Hilbert modular forms. Sturm [1987, Theorem 1] showed that mod ℓ modular forms of weight k and level $\Gamma_1(N)$ are determined by the first $(k/12)[\Gamma_1(1) : \Gamma_1(N)]$ mod ℓ Fourier coefficients. We prove an analogue of Sturm's result for Hilbert modular forms associated to totally real number fields not equal to $\mathbb{Q}$.

Doi and Ohta [1977, Lemma 2.1] showed a result similar to Sturm's theorem for elliptic cusp forms of weight 2 by a geometric method. Sturm improves the result for general weights and general levels by a technical method. For the case that the coefficient field is $\mathbb{C}$, the similar result was long known [Miyake 1989, Corollary 2.3.4]. Recently, Baba, Chakraborty and Petridis [Baba et al. 2002, Theorem 3] obtained its generalization for complex Hilbert modular forms by using the Rayleigh quotient for the Laplace operator. It seems difficult to apply their method to the mod ℓ case. For the mod ℓ case, Burgos Gil and Pacetti [Dieulefait et al. 2010, Appendix B] showed a generalization for Hilbert modular forms associated to $\mathbb{Q}(\sqrt{5})$ and level $\Gamma_0(6\sqrt{5})$ by a method similar to ours.

As mentioned above, the aim of the article is to prove an analogue of Sturm's theorem for Hilbert modular forms. In other words, we obtain an upper bound of

the order at zeros of Hilbert modular forms at the exceptional locus of resolution of cusp singularities.

To explain our main result, we prepare several notions (see Sections 2.1–2.2 for more precise definitions). Let $N \geq 3$ be an integer, F a totally real number field of finite degree $g \geq 2$, $\mathbb{O}_F$ the ring of integers of F, d_F the discriminant of F, and $\mathfrak{c}$ a nonzero integral ideal of F. Let $\widetilde{F}$ be the Galois closure of F and $\mathbb{O}_{\widetilde{F}}$ the ring of integers of $\widetilde{F}$. For a field K that is an $\mathbb{O}_{\widetilde{F}}[1/(Nd_F), \mu_N]$-algebra, $M_K = M_K(\mathfrak{c}, N)$ denotes the open connected Hilbert modular variety over K defined as a moduli of $\mathfrak{c}$-polarized Hilbert–Blumenthal abelian varieties with $\Gamma(N)$-structure. Then $\overline{M}_K = \overline{M}_{K,\Sigma}$ denotes the toroidal compactification associated to a collection Σ of cone decompositions, $\underline{\omega}^k = \otimes_i \underline{\omega}_i^{k_i}$ denotes the automorphic line bundle on $\overline{M}_K$ of weight $k = (k_1, \ldots, k_g) \in \mathbb{Z}_{\geq 0}^g$, and M^* denotes the minimal compactification of M. When $k = (1, 1, \ldots, 1)$, we write $\underline{\omega}$ instead of $\underline{\omega}^k$. Our main result is the following:

Theorem 1. *Let $N \geq 3$ be an integer, $k = (k_1, k_2, \ldots, k_g) \in \mathbb{Z}_{\geq 0}^g$, $\mathfrak{c}$ a nonzero integral ideal of F, K a field that is an $\mathbb{O}_{\widetilde{F}}[1/(Nd_F), \mu_N]$-algebra (a $\mathbb{Z}[1/(Nd_F), \mu_N]$-algebra if k is parallel), and S a nonempty finite set of irreducible components of codimension 1 in $\overline{M}_K \setminus M_K$. Let f be a $\mathfrak{c}$-polarized geometric Hilbert modular form over K of weight k and level $\Gamma(N)$, i.e., $f \in H^0(M_K(\mathfrak{c}, N), \underline{\omega}^k)$. Then*

$$f \not\equiv 0 \Rightarrow \min_{E \in S}\{\mathrm{ord}_E(f)\} < \kappa$$

for

$$\kappa = \kappa_S = \kappa_S(k, N) = C^{g-1} \sum_{i=1}^g \frac{k_i\{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_i) + (\mathfrak{I}^{(g-1)} \cdot \underline{\omega}_i)\}}{\left(\mathfrak{I}^{(g-1)} \cdot \sum_{E \in S} E\right)},$$

where C is a positive integer independent of k and N, $\mathfrak{I}$ is the inverse image of the ideal sheaf defining $M_K^* \setminus M_K$ by $\pi : \overline{M}_K \to M_K^*$, and the dot $(\cdot)$ denotes the intersection number (see Section 2.1).

The reason why we call Theorem 1 an "analogue" of Sturm's theorem is that the constant κ includes a strange constant C. When $g = 2$ and the canonical divisor $K_{\overline{M}}$ of $\overline{M}$ is nef (numerically effective), we may take an explicit constant as κ (Theorem 16). Moreover, when the field K is of positive characteristic and $\overline{M}$ is a minimal surface of general type, using Ekedahl's result [1988, Chapter III, Proposition 1.13], we may take a slightly better form of κ than Theorem 16 (Theorem 19). Applying Theorem 1 to classical Hilbert modular forms, we obtain the more useful results for the complex case (Corollary 9) and the mod ℓ case (Corollary 12). Corollary 9 gives another proof of [Baba et al. 2002, Theorem 3]. As another application of Theorem 1, we obtain a rough upper bound of the dimension of vector space of Hilbert modular forms (Corollary 15). Because the dimension for weight 1 is unknown, Corollary 15 is not trivial. We remark that

we can also show the results for the congruence subgroup Γ_1 by changing Γ in the proofs to Γ_1 and using theory of the arithmetic compactifications of Hilbert modular varieties for Γ_1 by Dimitrov [2004].

Theorem 1 is proved by a method extending Doi and Ohta's algebraic geometric one. To obtain an upper bound of the order of zeros of modular forms at cusps, Doi and Ohta used Riemann–Roch's theorem on modular curves over finite fields. However, to obtain the bound for Hilbert modular forms, we use the positivity of ample line bundles. The key point of the proof is the construction of a specific ample line bundle on the toroidal compactification. To do this, we use the semiampleness of the automorphic line bundle on the minimal compactification proved by Moret-Bailly [1985, Chapter V, Theorem 2.1]. Combining the inverse image of the ample line bundle onto a toroidal compactification with a certain relatively ample line bundle, a specific ample line bundle on the toroidal compactification is constructed.

This article is organized as follows. In Section 2.1, the notation of intersection numbers and the two facts of ampleness are explained. In Section 2.2, the definitions and the properties of integral models of Hilbert modular varieties, these arithmetic toroidal and minimal compactifications, and the geometric Hilbert modular forms are recalled. In Section 3.1, for Hilbert modular forms associated to totally real number fields that are not $\mathbb{Q}$, Theorem 1 is proved. As consequences of Theorem 1, Corollaries 9–15 are obtained. In Section 3.2, for the case that F is a real quadratic field, we obtain more explicit forms for κ in Theorems 16 and 19.

2. Preliminaries

Notation and conventions.

- Let F denote a totally real number field and $g = [F : \mathbb{Q}] < \infty$. Let $\mathbb{O}_F$, $\mathfrak{d} = \mathfrak{d}_F$, and d_F be the ring of integers of F, the different ideal of $F/\mathbb{Q}$, and the discriminant of $F/\mathbb{Q}$, respectively. Furthermore, $I = I_F$ denotes the set of the embeddings of F into $\mathbb{R}$, and $F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$.

- For a nonzero fractional ideal $\mathfrak{a}$ of F, $\mathfrak{a}^* = \mathfrak{a}^{-1}\mathfrak{d}^{-1}$, and $(\mathfrak{a})_+$ denotes the subset of $\mathfrak{a}$ consisting of the totally positive elements.

- Let $\mathbf{Sch}_{/R}$ denote the category of the schemes over a ring R, $R\text{-}\mathbf{Alg}$ the category of the R-algebras, and $\mathbf{Sets}$ the category of the sets.

- Let X be a normal variety. For a Cartier divisor D on X, $[D]$ denotes the Weil divisor associated to D. For a rational function f on X, $\operatorname{div}(f)$ denotes the divisor associated to f. A $\mathbb{Q}$-Cartier divisor D on X is a divisor such that mD is a Cartier divisor for a nonzero integer m.

- For two functions $f, g : \mathbb{R}^n \to \mathbb{R}_{\geq 0}$, $f \ll g$ denotes that there is a positive constant A such that $f(x) \leq Ag(x)$ for all $x \in \mathbb{R}^n$.

2.1. *Intersection numbers and ampleness.* In this section, we recall some facts in intersection theory [Fulton 1998, Chapter 2; Lazarsfeld 2004, Chapter 1].

Let d and n be positive integers such that $d < n$, K a field, and X a normal proper variety of dimension n over K. For Cartier divisors $D_1, D_2, \ldots, D_d$ on X and a d-dimensional irreducible closed subvariety V of X, the intersection number of $D_1, D_2, \ldots, D_d$ and V, denoted $(D_1 \cdots D_d \cdot V)$, is defined by several methods. But it is unique [Hartshorne 1977, Appendix A]. If $D = D_1 = \cdots = D_d$, then we write $(D^{(d)} \cdot V)$ instead of $(D_1 \cdots D_d \cdot V)$. For a Cartier divisor E, the intersection number $(D_1 \cdots D_{n-1} \cdot E)$ of $D_1, \ldots, D_{n-1}$ and E is defined by $(D_1 \cdots D_{n-1} \cdot [E])$ and linearity. If $D = E = D_1 = \cdots = D_{n-1}$, we write $(D^{(n)})$ instead of $(D^{(n-1)} \cdot D)$. For line bundles $\mathscr{L}_i \simeq \mathbb{O}_X(D_i)$ for $1 \leq i \leq n-1$ and $\mathscr{L}' \simeq \mathbb{O}_X(E)$, the intersection number $(\mathscr{L}_1 \cdots \mathscr{L}_{n-1} \cdot \mathscr{L}')$ is defined as $(D_1 \cdots D_{n-1} \cdot E)$. For $\mathbb{Q}$-Cartier divisors $D_1, \ldots, D_{n-1}$ and a $\mathbb{Q}$-Cartier divisor E, their intersection number $(D_1 \cdots D_{n-1} \cdot E)$ is defined by $(m_1 D_1 \cdots m_{n-1} D_{n-1} \cdot m_n E)/m_1 m_2 \cdots m_n$, where m_i are integers such that $m_i D_i$ and $m_n E$ are Cartier.

We recall a fact on the ampleness of line bundles:

Lemma 2 [Lazarsfeld 2004, Proposition 1.7.10]. *Let X and Y be proper varieties, $f : X \to Y$ a proper morphism, $\mathscr{L}$ an f-ample line bundle on X, and $\mathscr{M}$ an ample line bundle on Y. Then the line bundle $(f^*\mathscr{M})^{\otimes m} \otimes \mathscr{L}$ is ample on X for a sufficiently large positive integer m.*

We also recall that ampleness of line bundles is preserved by the pullback of finite morphisms.

Lemma 3 [Lazarsfeld 2004, Proposition 1.2.13, Corollary 1.2.28]. *Let X and Y be two projective varieties, and let $f : X \to Y$ be a finite morphism. If a line bundle $\mathscr{L}$ on Y is ample, then $f^*\mathscr{L}$ is ample. Moreover, when f is finite and surjective, a line bundle $\mathscr{L}$ on Y is ample if and only if $f^*\mathscr{L}$ is ample.*

2.2. *Hilbert modular varieties and geometric modular forms.* Let $\mathfrak{c} \subset F$ be a fixed nonzero integral ideal, N a positive integer, and R a $\mathbb{Z}[1/(Nd_F)]$-algebra. For an R-algebra B, let $\underline{A} = (A, \iota, \lambda, \phi_N)$ be a $\mathfrak{c}$-polarized Hilbert–Blumenthal abelian variety with $\Gamma(N)$-structure over B; i.e.,

(1) $\rho : A \to \operatorname{Spec}(B)$ is an abelian scheme,

(2) $\iota : \mathbb{O}_F \hookrightarrow \operatorname{End}_B(A)$ is an injective ring homomorphism taking 1 to the identity,

(3) λ is a $\mathfrak{c}$-polarization (see [Hida 2004, Section 4.1.1] for the definition),

(4) $\phi_N : A[N] \simeq (\mathbb{O}_F/N\mathbb{O}_F)^2$ is a $\Gamma(N)$-structure (i.e., an isomorphism as $\mathbb{O}_F$-modules), and

(5) $\rho_* \Omega^1_{A/\operatorname{Spec} B}$ is a locally free $\mathbb{O}_F \otimes_{\mathbb{Z}} B$-module of rank 1.

A suitable rule defines the isomorphism on such schemes [Hida 2004, Section 4.1.1].

We consider the contravariant functor $\mathscr{E} = \mathscr{E}(\mathfrak{c}, \Gamma(N)) : \mathbf{Sch}_{/R}^{\mathrm{op}} \to \mathbf{Sets}$ defined by

$$\mathscr{E}(\mathfrak{c}, \Gamma(N))(B) = \{(A, \iota, \lambda, \phi_N)_{/B}/\mathrm{isom.}\}.$$

The functor $\mathscr{E}(\mathfrak{c}, \Gamma(N))$ admits a coarse moduli scheme over R. When $N \geq 3$, $\mathscr{E}(\mathfrak{c}, \Gamma(N))$ is representable by a scheme over R. Its fine moduli scheme is denoted $M = M_R = M_N = M_{R,N} = M(\mathfrak{c}, \Gamma(N)) = M_R(\mathfrak{c}, \Gamma(N))$, and the universal object is denoted $A^U \to M_N$. Remark that M is naturally defined over $R[\mu_N] = R[x]/(\Phi_N(x))$ [Deligne and Rapoport 1973, Section 3.20], where Φ_N is the Nth cyclotomic polynomial. Furthermore, M is a smooth, geometrically irreducible scheme of finite type and relative dimension g over $R[\mu_N]$.

Next we define the cusps of M_N. Let $\mathfrak{a}$ and $\mathfrak{b}$ be nonzero fractional ideals of F such that $\mathfrak{c} = \mathfrak{a}\mathfrak{b}^{-1}$, and let

$$\phi_N : \mathfrak{b} \oplus \mathfrak{a}^* / N(\mathfrak{b} \oplus \mathfrak{a}^*) \simeq (\mathbb{O}_F / N\mathbb{O}_F)^2$$

be an isomorphism as $\mathbb{O}_F$-modules. We set

$$\Gamma(N; \mathfrak{a}, \mathfrak{b}) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL_2(F) \ \middle| \ \begin{matrix} a, d \in 1 + N\mathbb{O}_F, \\ b \in N(\mathfrak{a}\mathfrak{b})^*, \ c \in N(\mathfrak{a}\mathfrak{b}\mathfrak{d}_F) \end{matrix} \right\}.$$

We define the action of $\gamma \in \Gamma(N; \mathfrak{a}, \mathfrak{b})$ on $\mathfrak{b} \oplus \mathfrak{a}^*$ by

$$\gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \mathfrak{b} \oplus \mathfrak{a}^* \to \mathfrak{b} \oplus \mathfrak{a}^* : (\alpha, \beta) \mapsto (\alpha, \beta)\gamma^{-1} = (d\alpha - c\beta, -b\alpha + a\beta).$$

A cusp of $M(\mathfrak{c}, \Gamma(N))$ is defined to be $(\mathfrak{a}, \mathfrak{b}, \phi_N) \bmod \mathscr{B}_{\mathbb{Q}} \cap \Gamma(N; \mathfrak{a}, \mathfrak{b})$, where $\mathscr{B}_{\mathbb{Q}}$ is the standard Borel subgroup of $\mathrm{Res}_{\mathbb{Q}}^F GL_2$.

Let $\Sigma = \{\Sigma_s\}$ be a collection of $\Gamma(N)$-admissible polyhedral cone decompositions [Hida 2004, Section 4.1.4], and $\overline{M} = \overline{M}_R = \overline{M}_{R,N} = \overline{M}_{R,N,\Sigma}$ denotes the toroidal compactification of $M_R(\mathfrak{c}, \Gamma(N))$ associated to Σ. Then $\overline{M}$ is normal over $R[\mu_N]$. We can take Σ such that $\overline{M}$ is smooth or projective over $R[\mu_N]$ [Hida 2004, Section 4.1.4].

Let $\widetilde{F}$ be the Galois closure of F and $\mathbb{O}_{\widetilde{F}}$ the ring of integers. Then, for an $\mathbb{O}_{\widetilde{F}}[1/(Nd_F), \mu_N]$-algebra R, we have the isomorphism

$$\mathbb{O}_F \otimes_{\mathbb{Z}} R \simeq \bigoplus_{i=1}^g R : a \otimes b \mapsto (\sigma_1(a)b, \sigma_2(a)b, \ldots, \sigma_g(a)b).$$

For the structure morphism $\rho : A^U \to M_R(\mathfrak{c}, \Gamma(N))$, $\overline{\omega}$ denotes the locally free $\mathbb{O}_F \otimes_{\mathbb{Z}} \mathbb{O}_M$-module $\rho_* \Omega^1_{A/R}$ of rank 1. By the above isomorphism, $\overline{\omega}$ is decomposed $\overline{\omega} \simeq \bigoplus_i \underline{\omega}_i$, where $\underline{\omega}_i$ is the locally free $\mathbb{O}_M$-module of rank 1 corresponding to $\sigma_i \in I$. For $k = (k_1, k_2, \ldots, k_g) \in \mathbb{Z}^g$, we define the line bundle $\underline{\omega}_R^k = \underline{\omega}_{N,R}^k = \otimes_i \underline{\omega}_i^{\otimes k_i}$ called

by the automorphic line bundle of weight k. When the weight $k = (k_0, k_0, \ldots, k_0)$ is parallel, we can construct the automorphic line bundle $\underline{\omega}^k_{R'} = (\bigwedge^g \overline{\omega})^{\otimes k_0} = \det(\overline{\omega})^{\otimes k_0}$ over every $\mathbb{Z}[1/(Nd_F), \mu_N]$-algebra R'.

A $\mathfrak{c}$-polarized holomorphic *geometric Hilbert modular form* (abbreviated as GHMF) associated to F of weight $k = (k_1, k_2, \ldots, k_g)$ and level $\Gamma(N)$ defined over R is an element of $H^0(M_R, \underline{\omega}^k_R)$.

There is the semiabelian scheme $\mathcal{G}$ over $\overline{M}_R$ that is an extension of the universal abelian scheme $\underline{A}^U \to M_R$. Thus, $\underline{\omega}^k$ is extended on $\overline{M}$. When $[F : \mathbb{Q}] \geq 2$, by the Koecher principle [Chai 1990, Section 4.3], we have

$$H^0(M_R, \underline{\omega}^k_R) = H^0(\overline{M}_R, \underline{\omega}^k_R).$$

Let $\rho : \mathcal{G} \to \overline{M}_R$ be the structure morphism and $\underline{\omega} = \det(\rho_* \Omega^1_{\mathcal{G}/\overline{M}_R})$.

By Moret-Bailly [1985, Chapter V, Theorem 2.1], $\underline{\omega}$ is semiample; i.e., there is a positive integer n_0 such that $\underline{\omega}^{\otimes n_0}$ is generated by global sections. Thus, the canonical rational map

$$\phi_{\underline{\omega}^{\otimes n_0}} : \overline{M} \to \mathbb{P}^r_R, \quad P \mapsto (s_0(P) : s_1(P) : \cdots : s_r(P))$$

is a morphism, where $s_0, s_1, \ldots, s_r \in H^0(\overline{M}, \underline{\omega}^{\otimes n_0})$ are global sections generating $\underline{\omega}^{\otimes n_0}$. We set $\pi = \pi_N = \phi_{\underline{\omega}^{\otimes n_0}}$. The minimal compactification, denoted $M^* = M^*_R = M^*_{R,N} = M^*_R(\mathfrak{c}, \Gamma(N))$, of M is defined to be the image of π, and

$$M^* \simeq \mathrm{Proj}\Big(\bigoplus_{n \geq 0} H^0(\overline{M}, \underline{\omega}^{\otimes n_0 n})\Big).$$

By the semiampleness of $\underline{\omega}$, the graded ring $\bigoplus_{n \in \mathbb{Z}_{\geq 0}} H^0(\overline{M}, \underline{\omega}^{\otimes n_0 n})$ is finitely generated over R [Chai 1990, Section 4.4]; in particular, M^* is of finite type over R. And M is isomorphic to an open dense subscheme of M^*, also denoted M. When $g = [F : \mathbb{Q}] \geq 2$, by the Koecher principle M^* does not depend on the choice of cone decompositions. And M^* is also normal. The connected components $M^* \setminus M$ are in one-to-one correspondence with the cusps of M. The direct image $\pi_* \underline{\omega}$ is $\mathbb{Q}$-Cartier; i.e., there is a positive integer n_0 such that $\pi_* \underline{\omega}^{\otimes n_0}$ is a line bundle. Then $\pi_* \underline{\omega}^{\otimes n_0}$ is ample [Chai 1990, Section 4.3].

There is another useful definition equivalent to the above definition. We consider the covariant functor $\mathcal{P} = \mathcal{P}_R(\mathfrak{c}, \Gamma(N)) : R\text{-}\mathbf{Alg} \to \mathbf{Sets}$ with

$$\mathcal{P}(B) = \big\{ (\underline{A}, \omega)_{/B}/\text{isom.} \,\big|\, \underline{A} \in \mathcal{E}(B), \ \rho : A \to \mathrm{Spec}(B),$$
$$H^0(A, \rho_* \Omega^1_{A/B}) \simeq (\mathcal{O}_F \otimes B)\omega \big\}.$$

Then for every R-algebra B, $\mathcal{T}_{\mathbb{Z}}(B) = (\mathcal{O}_F \otimes_{\mathbb{Z}} B)^{\times}$ acts on $\mathcal{P}(B)$, and $\mathcal{P}/\mathcal{T}_{\mathbb{Z}} \simeq \mathcal{E}$. If $\mathcal{E}$ is representable, so is $\mathcal{P}$. And $\mathcal{M}$ denotes the moduli scheme over R representing $\mathcal{P}$. Then $\mathcal{M}$ is a $\mathcal{T}_{\mathbb{Z}}$-torsor over M. It is known that $H^0(M, \underline{\omega}^k) \subset H^0(\mathcal{M}, \mathcal{O}_{\mathcal{M}})$ [Dimitrov and Tilouine 2004, Remarque 4.5]. Thus, $f \in H^0(M, \underline{\omega}^k)$ is regarded as

a function such that a pair $(\underline{A}, \omega)_{/B}$, for every R-algebra B, associates an element $f((\underline{A}, \omega)_{/B}) \in B$ that satisfies the following:

(1) The value of f depends only on the isomorphism class of $(\underline{A}, \omega)$.

(2) Given a base change map $\rho : B \to B'$, f satisfies

$$\rho(f((\underline{A}, \omega)_{/B})) = f((\underline{A}, \omega)_{/B'}).$$

(3) For every $\alpha \in (\mathbb{O}_F \otimes_{\mathbb{Z}} B)^{\times} = \mathcal{T}_{\mathbb{Z}}(B)$,

$$f(\underline{A}, \alpha\omega) = \prod_{i=1}^{g} \sigma_i(\alpha)^{-k_i} f(\underline{A}, \omega).$$

Let $R = \mathbb{O}_{\widetilde{F}}[1/(Nd_F), \mu_N]$, and let B be an R-algebra and f a $\mathfrak{c}$-polarized GHMF of weight $k = (k_1, k_2, \ldots, k_g)$ and level $\Gamma(N)$ defined over B. Let $\mathfrak{a}$ and $\mathfrak{b} \subset F$ be nonzero fractional ideals such that $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$, $\phi_N : (\mathfrak{b} \oplus \mathfrak{a}^*)/N(\mathfrak{b} \oplus \mathfrak{a}^*) \simeq (\mathbb{O}_F/N\mathbb{O}_F)^2$ an isomorphism, s a cusp of $\Gamma(N)$ parametrized by $(\mathfrak{a}, \mathfrak{b}, \phi_N)$, and σ a cone in the cone decomposition Σ_s. We set

$$R_\sigma(N) = R[q^\xi]_{\xi \in N^{-1}\mathfrak{a}\mathfrak{b} \cap \sigma^\vee} \quad \text{and} \quad R_0(N) = R[q^\xi]_{\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}},$$

where $\sigma^\vee$ is the dual of σ. Let $S_\sigma(N) = \operatorname{Spec} R_\sigma(N)$ and $S_0(N) = \operatorname{Spec} R_0(N)$. Let $\hat{S}_\sigma(N)$ be the formal completion of $S_\sigma(N)$ along $S_\sigma^\infty(N) = S_\sigma(N) \setminus S_0(N)$. The formal scheme $\hat{S}_\sigma(N)$ is affine, and we define $\hat{R}_\sigma(N)$ as its coordinate ring. Then we can show that

$$R[\![q^\xi]\!]_{\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}} \hookrightarrow \hat{R}_\sigma(N).$$

The Tate object $Tate_{(\mathfrak{a}, \mathfrak{b})}(q)$ corresponding to the cusp s is a Hilbert–Blumenthal abelian variety defined over the quotient field $Q(\hat{R}_\sigma(N) \otimes_R B)$. The q-expansion of f at the cusp s is defined to be the value of f at the Tate object corresponding to s:

$$f(Tate_{(\mathfrak{a}, \mathfrak{b})}(q), \phi_N, \omega) = a_0 + \sum_{\xi \in N^{-1}\mathfrak{a}\mathfrak{b} \cap \sigma^\vee} a_\xi q^\xi \in Q(\hat{R}_\sigma(N) \otimes_R B),$$

where ω is the canonical nonvanishing differential form on $Tate_{(\mathfrak{a}, \mathfrak{b})}(q)$.

The following fact is known as the q-expansion principle:

Lemma 4. *Let B' be a subring of B such that B' is an R-algebra, and let f be as above. Then f is defined over B' if and only if $a_\xi \in B'$ for all $\xi \in (\mathfrak{a}\mathfrak{b})_+ \cup \{0\}$.*

Proof. See [Rapoport 1978, Theorem 6.7]. □

There is a correspondence between the classical Hilbert modular forms and the geometric Hilbert modular forms defined over $\mathbb{C}$. Let $\mathfrak{H} = \mathscr{H}^g$ be the g-tuple product of the complex upper half-plane. Then the holomorphic function $f : \mathfrak{H} \to \mathbb{C}$ is a

classical Hilbert modular form of weight $k = (k_1, \ldots, k_g)$ of level $\Gamma(N; \mathfrak{a}, \mathfrak{b})$ if f satisfies the following equation:

$$f(\gamma z) = \prod_{\sigma_i \in I} (\sigma_i(c) z_i + \sigma_i(d))^{k_i} f(z) \quad \text{for all } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma(N; \mathfrak{a}, \mathfrak{b}).$$

Here $M_k(\Gamma(N; \mathfrak{a}, \mathfrak{b}))$ denotes the complex vector space of the classical Hilbert modular forms of weight k and level $\Gamma(N; \mathfrak{a}, \mathfrak{b})$.

To show corollaries of our main result, we need the following lemma:

Lemma 5. *As $\mathbb{C}$-vector spaces, $M_k(\Gamma(N; \mathfrak{a}, \mathfrak{b}))$ and $H^0(\overline{M}_\mathbb{C}(\mathfrak{a}\mathfrak{b}^{-1}, \Gamma(N)), \underline{\omega}_\mathbb{C}^k)$ are canonically isomorphic.*

Proof. See [Rapoport 1978, Lemme 6.12]. □

For $f \in M_k(\Gamma(N; \mathfrak{a}, \mathfrak{b}))$, let $f' \in H^0(\overline{M}_\mathbb{C}, \underline{\omega}_\mathbb{C}^k)$ be the form corresponding to f. We remark that replacing q^ξ by $e^{2\pi i \, \mathrm{Tr}(\xi z)}$, the q-expansion of f' at a cusp s corresponds to the Fourier expansion of f at the cusp s.

Remark 6. Let K be a field that is a $\mathbb{Z}[1/(N d_F), \mu_N]$-algebra. For the open Hilbert modular variety $M_K \subset \overline{M}_K$, the Kodaira–Spencer isomorphism [Katz 1978, Section 1.0] gives $\overline{\omega} \otimes_{\mathcal{O}_F \otimes \mathcal{O}_M} \overline{\omega} \simeq \Omega^1_{M/K}$. Therefore, $\underline{\omega}_K^{\otimes 2} = \underline{\omega}_K \otimes_{\mathcal{O}_M} \underline{\omega}_K \simeq \det(\overline{\omega} \otimes_{\mathcal{O}_F \otimes \mathcal{O}_M} \overline{\omega}) \simeq \Omega^g_{M/K}$. Observing the behaviors of sections at cusps, we have

$$\underline{\omega}^{\otimes 2} \simeq \Omega^g_{\overline{M}/K}(\log D_\infty) \simeq \mathcal{O}_{\overline{M}}(K_{\overline{M}} + D_\infty),$$

where $K_{\overline{M}}$ is the canonical divisor of $\overline{M}$, and the Cartier divisor $D_\infty = \sum E$, where E runs over the irreducible components of codimension 1 in $\overline{M}_K \setminus M_K$.

3. Proofs of the main theorems

3.1. *Main result.* We assume $g = [F : \mathbb{Q}] \geq 2$ so that we use intersection theory. Let N be a positive integer such that $N \geq 3$, $k = (k_1, \ldots, k_g) \in \mathbb{Z}_{\geq 0}^g$, $\mathfrak{c}$ a nonzero integral ideal of F, and K a field that is a $\mathbb{Z}[1/(N d_F), \mu_N]$-algebra if k is parallel or an $\mathcal{O}_{\widetilde{F}}[1/(N d_F), \mu_N]$-algebra otherwise. Let $M_K = M_{K,N} = M_K(\mathfrak{c}, \Gamma(N))$ be the moduli scheme over K defined in Section 2.2. We choose a collection of projective $\Gamma(N)$-admissible polyhedral cone decompositions $\Sigma = \{\Sigma_s\}$. Let $\overline{M}_K = \overline{M}_{K,N} = \overline{M}_{K,N,\Sigma}$ be the toroidal compactification of M_K associated to Σ, $M_K^* = M_{K,N}^*$ the minimal compactification of M_K, and $\pi = \pi_N : \overline{M}_K \to M_K^*$ the canonical morphism defined in Section 2.2. We set $D_\infty = \sum E$, where E runs over the irreducible components of codimension 1 in $\overline{M}_K \setminus M_K$. Let S be a nonempty finite set of irreducible components of codimension 1 in $\overline{M}_K \setminus M_K$. For a closed point $P \in M_K^*$, $\mathscr{I}_P$ denotes the ideal sheaf on M_K^* defining P. And $\mathfrak{I} = \mathfrak{I}_N$ denotes the inverse image ideal sheaf $\pi_N^{-1}(\otimes_{P \in M_K^* \setminus M_K} \mathscr{I}_P) \cdot \mathcal{O}_{\overline{M}_K}$.

For convenience, we restate Theorem 1.

Theorem 1. *Retain the notation above and let f be a $\mathfrak{c}$-polarized geometric Hilbert modular form over K of weight k and level $\Gamma(N)$, i.e., $f \in H^0(\overline{M}_K(\Gamma(N), \mathfrak{c}), \underline{\omega}^k)$. Then, if $f \not\equiv 0$, we have $\min_{E \in S}\{\mathrm{ord}_E(f)\} < \kappa$ for*

$$\kappa = \kappa_S = \kappa_S(k, N) = C^{g-1} \sum_{i=1}^{g} \frac{k_i\{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_i) + (\mathfrak{J}^{(g-1)} \cdot \underline{\omega}_i)\}}{\left(\mathfrak{J}^{(g-1)} \cdot \sum_{E \in S} E\right)},$$

where C is a positive integer which is independent of k or N.

Remark 6. If f is a cusp form, κ can be smaller. More generally, for a nonnegative integer a, if f has zeros of order a at all irreducible components of codimension 1 in $\overline{M} \setminus M$, we may take κ as

$$\kappa = C^{g-1} \sum_{i=1}^{g} \frac{k_i\{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_i) + (\mathfrak{J}^{(g-1)} \cdot \underline{\omega}_i)\}}{\left(\mathfrak{J}^{(g-1)} \cdot \sum_{E \in S} E\right)} - a\left(\frac{(\mathfrak{J}^{(g-1)} \cdot D_\infty)}{\left(\mathfrak{J}^{(g-1)} \cdot \sum_{E \in S} E\right)} - 1\right).$$

Remark 7. When the weight is parallel, $k = (k_0, \ldots, k_0)$, we have

$$\sum_i k_0(\mathfrak{J}^{(g-1)} \cdot \underline{\omega}_i) = k_0(\mathfrak{J}^{(g-1)} \cdot \underline{\omega}) = 0$$

by the projection formula. This implies

$$\kappa = \frac{k_0 C^{g-1}(\underline{\omega}^{(g)})}{\left(\mathfrak{J}^{(g-1)} \cdot \sum_{E \in S} E\right)}.$$

By Hirzebruch's proportionality theorem [Mumford 1977, Theorem 3.2], we have

$$(\underline{\omega}^{(g)}) = A \, \mathrm{vol}(\overline{M})c_1(\Omega^1_{\mathbb{P}^g_\mathbb{C}})^g = A'(-1)^g \zeta_F(-1)[\Gamma(1; \mathfrak{a}, \mathfrak{b}) : \Gamma(N; \mathfrak{a}, \mathfrak{b})],$$

where A and A' are positive constants not depending on the level N. In this case, we obtain

$$\kappa = k_0 C^{g-1} A'(-1)^g \zeta_F(-1) \frac{[\Gamma(1; \mathfrak{a}, \mathfrak{b}) : \Gamma(N; \mathfrak{a}, \mathfrak{b})]}{\left(\mathfrak{J}^{(g-1)} \cdot \sum_{E \in S} E\right)}.$$

To prove Theorem 1, we introduce the following lemma:

Lemma 8. *For a sufficiently large integer C, $\underline{\omega}^{\otimes C} \otimes \mathfrak{J}$ is ample on $\overline{M}_K$.*

Proof. We know that $\pi_* \underline{\omega}$ is $\mathbb{Q}$-Cartier; i.e., $\pi_* \underline{\omega}^{\otimes k_0}$ is a line bundle on M_K^* for a sufficiently large k_0. Then this is ample [Chai 1990, Section 4.3]. Since $\overline{M}_K$ is the normalization of the blowing-up of M_K^* along the ideal sheaf $\otimes_{P \in M_K^* \setminus M_K} \mathfrak{I}_P$ [Ash et al. 1975, Chapter IV; Faltings and Chai 1990, Chapter V, Theorem 5.8], $\mathfrak{J}$ is π-ample.

Lemma 2 implies the line bundle $(\pi^* \pi_* \underline{\omega})^{\otimes k_0 n} \otimes \mathfrak{J}$ is ample for an integer $n \gg 0$. Because $\underline{\omega}$ is semiample, $\pi^* \pi_* \underline{\omega}^{\otimes k_0} \simeq \underline{\omega}^{\otimes k_0}$. Thus, $\underline{\omega}^{\otimes k_0 n} \otimes \mathfrak{J}$ is also ample for n. Replacing $k_0 n$ with C, we obtain the lemma. $\square$

Proof of Theorem 1. We assume that $f \not\equiv 0$ and set

$$\nu = \min_{E \in S}\{\mathrm{ord}_E(f)\}.$$

Then $f \in H^0(\overline{M}_K, \underline{\omega}^k(-\nu \sum_{E \in S} E))$. Thus,

$$\mathbb{O}_{\overline{M}_K}(\mathrm{div}(f)) \subset \underline{\omega}^k\left(-\nu \sum_{E \in S} E\right),$$

and they are effective. By positivity of ample line bundles,

$$\left((\underline{\omega}^{\otimes C} \otimes \mathfrak{I})^{(g-1)} \cdot \underline{\omega}^k\left(-\nu \sum_{E \in S} E\right)\right) > 0$$

for the integer C in Lemma 8. Thus, we have

$$\nu < C^{g-1} \sum_{i=1}^{g} \frac{k_i\{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_i) + (\mathfrak{I}^{(g-1)} \cdot \underline{\omega}_i)\}}{(\mathfrak{I}^{(g-1)} \cdot \sum_{E \in S} E)}.$$

The independence of C and k and N refers to Lemma 14. $\square$

Next we apply Theorem 1 for classical Hilbert modular forms. For this purpose, we start with a review of the relation between smooth cone decompositions and local structures of $\overline{M}$ at cusps.

We assume that the cone decompositions in Σ are smooth. Let s be a cusp of $M(\mathfrak{c}, \Gamma(N))$ parametrized by $(\mathfrak{a}, \mathfrak{b}, \phi_N)$, where $\mathfrak{a}$ and $\mathfrak{b}$ are two nonzero fractional ideals such that $\mathfrak{a}\mathfrak{b}^{-1} = \mathfrak{c}$ and $\phi_N : \mathfrak{b} \oplus \mathfrak{a}^* / N(\mathfrak{b} \oplus \mathfrak{a}^*) \simeq (\mathbb{O}_F/N\mathbb{O}_F)^2$ is an isomorphism (see Section 2.2). We can take the quotient

$$\Sigma_s/U_N = \bigcup_{i=1}^{g} \{\overline{\sigma_{i,j}}\}_{j=1}^{r_i},$$

where $U_N = \{\epsilon \in \mathbb{O}_F^\times \mid \epsilon \equiv 1 \mod N\}$, r_i is a positive integer, and $\overline{\sigma_{i,j}} \equiv \sigma_{i,j} \mod U_N$ with an i-dimensional cone $\sigma_{i,j} \in \Sigma_s$. Let $1 \leq d \leq g$ and $\sigma \in \Sigma_s$ be a smooth d-dimensional cone. Since σ is smooth, we can take part of a $\mathbb{Z}$-basis $\alpha_1, \ldots, \alpha_d$ of $N^{-1}\mathfrak{a}\mathfrak{b} \cap \sigma^\vee$ such that $N^{-1}\mathfrak{a}\mathfrak{b} \cap \sigma^\vee = \mathbb{Z}_{\geq 0}\alpha_1 + \cdots + \mathbb{Z}_{\geq 0}\alpha_d + \mathbb{Z}\beta_{d+1} + \cdots + \mathbb{Z}\beta_g$ with some part of a $\mathbb{Z}$-basis $\beta_{d+1}, \ldots, \beta_g$ [Hida 2004, Section 4.1.4]. We regard 0 as a cone in Σ_s. Then $0^\vee = F_\infty = F \otimes_{\mathbb{Q}} \mathbb{R}$. We remark that $0^\vee = \bigcup_{\sigma \in \Sigma_s} \sigma^\vee$; in particular, $0^\vee \supset \sigma^\vee$. Thus, we have $N^{-1}\mathfrak{a}\mathfrak{b} \cap 0^\vee = \mathbb{Z}\alpha_1 + \cdots + \mathbb{Z}\alpha_d + \mathbb{Z}\beta_{d+1} + \cdots + \mathbb{Z}\beta_g$ for the above $\mathbb{Z}$-basis $\alpha_1, \ldots, \beta_g$.

For the above cone $\sigma \in \Sigma_s$ and the ring $R = \mathbb{Z}[1/(Nd_F), \mu_N]$, the ring $\hat{R}_\sigma(N)$ defined in Section 2.2 is the completion of $R_\sigma(N)$ by the principal ideal $(q^{\alpha_1} \cdots q^{\alpha_d})$. We set

$$\hat{R}_\sigma^0(N) = \hat{R}_\sigma(N) \otimes_{R_\sigma(N)} R_0(N).$$

It is easy to show that, if $\tau \in \Sigma_s$ is a face of σ, $\text{Spec}(\hat{R}_\tau(N)) \subset \text{Spec}(\hat{R}_\sigma(N))$. The exceptional locus of $\pi : \overline{M} \to M^*$ is defined by gluing $\text{Spec}(\hat{R}_\sigma(N)) \setminus \text{Spec}(\hat{R}_\sigma^0(N))$ along Σ_s / U_N by the above rule.

For one element α_i of the basis, the divisor E_i associated to α_i is defined by the closed irreducible subvariety of codimension 1 in $\overline{M}$ that contains the affine subvariety defined by $\{q^{\alpha_i} = 0\}$ in the exceptional locus.

For a GHMF f over a field K that is an $\mathbb{O}_{\widetilde{F}}[1/(Nd_F), \mu_N]$-algebra and its q-expansion

$$f = \sum_{\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}} a_\xi q^\xi$$

at the cusp s, the order of f at $E_{i,K} = E_i \times \text{Spec}(K)$ is

$$\text{ord}_{E_{i,K}}(f) = \min\{ m_i \in \mathbb{Z} \mid a_{m_1\alpha_1 + \cdots + m_g\alpha_g} \neq 0 \text{ in } K \}.$$

Corollary 9. *Let $f \in M_k(\Gamma(N; \mathfrak{a}, \mathfrak{b}))$ and*

$$f(z) = \sum_{\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}} a_\xi e^{2\pi i \, \text{Tr}(\xi z)}$$

be the Fourier expansion of f at cusp s. We fix a g-dimensional cone $\sigma \in \Sigma_s$. Let $\{\alpha_1, \ldots, \alpha_g\}$ be a $\mathbb{Z}$-basis of $N^{-1}(\mathfrak{a}\mathfrak{b})$ corresponding to σ and E_i the divisor associated to α_i. For a fixed $i \in \{1, 2, \ldots, g\}$, we set

$$\kappa_i = C^{g-1} \sum_{j=1}^{g} \frac{k_j \{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_j) + (\mathfrak{J}^{(g-1)} \cdot \underline{\omega}_j)\}}{(\mathfrak{J}^{(g-1)} \cdot E_{i,\mathbb{C}})},$$

$$\mathcal{S}_i = \left\{ \xi \in m_1\alpha_1 + \cdots + m_g\alpha_g \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\} \mid 0 \leq m_i \leq \kappa_i \right\},$$

where C is the integer in [Theorem 1](). Then, if $a_\xi = 0$ for every $\xi \in \mathcal{S}_i$, we have $a_\xi = 0$ for every $\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}$.

Proof. By [Lemma 5](), f is regarded as a global section of $H^0(\overline{M}_{\mathbb{C}}, \underline{\omega}^k)$, and the q-expansion of f at cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ is

$$f(Tate_{(\mathfrak{a},\mathfrak{b})}, \phi_N, \omega) = \sum_{\xi \in N^{-1}(\mathfrak{a}\mathfrak{b})_+ \cup \{0\}} a_\xi q^\xi,$$

where ω is a nonvanishing differential form on $Tate_{(\mathfrak{a},\mathfrak{b})}$. Applying [Theorem 1]() for f as $K = \mathbb{C}$ and $S = \{E_{i,\mathbb{C}}\}$, the corollary is proved. □

Remark 10. We can easily show that the subset $\mathcal{S}_i$ is a finite set. Indeed, by [Proposition 13]() and an easy calculation, we have an upper bound

$$\#\mathcal{S}_i \ll \max\{k_i\}^g N^{3g^2}.$$

Remark 11. [Corollary 9]() gives a different proof of the result of [Baba et al. 2002].

Next we discuss the case of positive characteristic. The following corollary is an analogue of Sturm's theorem:

Corollary 12. *Let* $f \in M_k(\Gamma(N; \mathfrak{a}, \mathfrak{b}))$ *and*

$$f(z) = \sum_{\xi \in N^{-1}(\mathfrak{ab})_+ \cup \{0\}} a_\xi e^{2\pi i \, \mathrm{Tr}(\xi z)}$$

be the Fourier expansion of f *at the cusp* $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$. *We assume* $g = [F : \mathbb{Q}] \geq 2$ *and fix a* g-*dimensional cone* $\sigma \in \Sigma_s$. *Let* $\{\alpha_1, \ldots, \alpha_g\}$ *be a* $\mathbb{Z}$-*basis of* $N^{-1}(\mathfrak{ab})$ *corresponding to* σ *and* E_i *the divisor associated to* α_i. *Let* L *be a number field,* $\mathbb{O}_L$ *the ring of integers of* L, *and* λ *a prime ideal in* $\mathbb{O}_L$ *such that* $\lambda \nmid N d_F \mathbb{O}_L$. *Assume* $a_\xi \in \mathbb{O}_L$ *for every* ξ. *For a fixed* $i \in \{1, 2, \ldots, g\}$, *we set*

$$\kappa_i = C^{g-1} \sum_{j=1}^{g} \frac{k_j \{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_j) + (\mathfrak{J}^{(g-1)} \cdot \underline{\omega}_j)\}}{(\mathfrak{J}^{(g-1)} \cdot E_{i,\mathbb{F}})},$$

where C *is the integer in Theorem 1 and* $\mathbb{F}$ *is a finite extension of* $\mathbb{O}_L/\lambda$ *(defined in the proof), and set* $\mathcal{G}_i$ *to be same as in Corollary 9.*

Then, if $a_\xi \equiv 0 \bmod \lambda$ *for every* $\xi \in \mathcal{G}_i$, *we have* $a_\xi \equiv 0 \bmod \lambda$ *for every* ξ *in* $N^{-1}(\mathfrak{ab})_+ \cup \{0\}$.

Proof. Lemma 5 implies that f is regarded as an element of $H^0(\overline{M}_\mathbb{C}, \underline{\omega}^k)$. Let $\widetilde{L}$ be the composition field of L and $\widetilde{F}$. By the assumption, all the Fourier coefficients of f are in $\mathbb{O}_L$ and thus in $\mathbb{O}_{\widetilde{L}}[1/(N d_F), \mu_N]$. Let λ' be a maximal ideal of ring $\mathbb{O}_{\widetilde{L}}[1/(N d_F), \mu_N]$ such that $\lambda' \mid \lambda \mathbb{O}_{\widetilde{L}}[1/(N d_F), \mu_N]$. Remark that $\lambda' \cap \mathbb{O}_L = \lambda$. By Lemma 4 and the commutativity with base change maps of GHMF (see Section 2.2), we regard f as a GHMF defined over the field $\mathbb{O}_{\widetilde{L}}[1/(N d_F), \mu_N]/\lambda'$, and the q-expansion of f at cusp $s = (\mathfrak{a}, \mathfrak{b}, \phi_N)$ is

$$f(\mathrm{Tate}_{(\mathfrak{a}, \mathfrak{b})}, \phi_N, \omega_{\mathrm{can}}) = \sum_{\xi \in N^{-1}(\mathfrak{ab})_+ \cup \{0\}} (a_\xi \bmod \lambda') q^\xi.$$

Applying Theorem 1 for f as $K = \mathbb{F} = \mathbb{O}_{\widetilde{L}}[1/(N d_F), \mu_N]/\lambda'$ and $S = \{E_{i,\mathbb{F}}\}$, if $f \not\equiv 0 \bmod \lambda'$, an integer m_i such that $0 \leq m_i \leq \kappa_i$ and $a_{m_1 \alpha_1 + \cdots + m_g \alpha_g} \not\equiv 0 \bmod \lambda'$ exists. Thus, $a_{m_1 \alpha_1 + \cdots + m_g \alpha_g} \notin \lambda' \cap \mathbb{O}_K = \lambda$, and the contrapositive of the corollary is proved. $\square$

Next we examine the growth of $\kappa = \kappa(k, N)$ associated to weight $k = (k_1, \ldots, k_g)$ and level N.

Proposition 13. *We may take* κ *in Theorem 1 as*

$$\kappa(k, N) \ll \max_i \{k_i\} N^{3g}.$$

Let $\underline{\omega}_N$ be the automorphic line bundle on $\overline{M}_N$. To show Proposition 13, we introduce the following lemma:

Lemma 14. *Let ℓ be the characteristic of K, N and N' two positive integers such that $N \geq 3$ and $N \mid N'$ and $\ell \nmid N'$, and C a positive integer. Then $\underline{\omega}_N^{\otimes C} \otimes \mathfrak{I}_N$ is ample if and only if $\underline{\omega}_{N'}^{\otimes C} \otimes \mathfrak{I}_{N'}$ is ample.*

Proof. We can take a finite étale morphism $h : M_{N',K} \to M_{N,K}$ over $\mathbb{Z}[1/(N')]$ [Chai 1990, Section 2.2]. Let Σ be a collection of cone decompositions of level N and $\overline{M}_{N,K,\Sigma}$ the toroidal compactification associated to Σ. We define $\overline{M}_{N',K}$ to be the normalization of $\overline{M}_{N,K}$ in $M_{N',K}$, and $\tilde{h} : \overline{M}_{N',K} \to \overline{M}_{N,K}$ denotes the normalization morphism. Then the normalization $\overline{M}_{N',K}$ is the toroidal compactification of $M_{N',K}$ associated to Σ [Faltings and Chai 1990, Chapter IV, Theorem 6.7(1)]. Here Σ is regarded as a collection of cone decompositions of level N'. Remark that $\overline{M}_{N',K}$ may not be smooth. Then $\tilde{h}$ is an extension of h to toroidal compactifications [Faltings and Chai 1990, Chapter IV, Theorem 6.7(2)]. In particular, $\tilde{h}|_{M_{N',K}}$ is étale. We can show that $\tilde{h}^* \underline{\omega}_N^{\otimes 2} \simeq \underline{\omega}_{N'}^{\otimes 2}$. Indeed by Remark 6,

$$\underline{\omega}_N^{\otimes 2} \simeq \mathbb{O}_{\overline{M}_{N,K}}(K_{\overline{M}_{N,K}} + D_{\infty,N}),$$

where $D_{\infty,N} = \sum E$ such that E runs over the irreducible components of codimension 1 in $\overline{M}_{N,K} \setminus M_{N,K}$. Since $\tilde{h}|_{M_{N',K}}$ is étale, the ramification of $\tilde{h}$ occurs only at $\overline{M}_{N,K} \setminus M_{N,K}$. Thus, we can show that

$$\tilde{h}^* \underline{\omega}_N^{\otimes 2} \simeq \tilde{h}^* \mathbb{O}_{\overline{M}_{N,K}}(K_{\overline{M}_{N,K}} + D_{\infty,N}) \simeq \mathbb{O}_{\overline{M}_{N',K}}(K_{\overline{M}_{N',K}} + D_{\infty,N'}) \simeq \underline{\omega}_{N'}^{\otimes 2}.$$

It is known that $\tilde{h}^* \mathfrak{I}_N \simeq \mathfrak{I}_{N'}$ [Faltings and Chai 1990, Chapter V, proof of Corollary 5.14]. Assume that C is even and that $\underline{\omega}_N^{\otimes C} \otimes \mathfrak{I}_N$ is ample. Since $\tilde{h}$ is finite and surjective, $\underline{\omega}_N^{\otimes C} \otimes \mathfrak{I}_N$ is ample if and only if $\underline{\omega}_{N'}^{\otimes C} \otimes \mathfrak{I}_{N'}$ is ample by Lemma 3. $\square$

Proof of Proposition 13. Let N and N' be positive integers such that $N, N' \geq 3$ and $\mathrm{char}(K) \nmid NN'$ and C a positive integer such that $\underline{\omega}_N^{\otimes C} \otimes \mathfrak{I}_N$ is ample. According to Lemma 14, $\underline{\omega}_{NN'}^{\otimes C} \otimes \mathfrak{I}_{NN'}$ is ample. Thus, $\underline{\omega}_{N'}^{\otimes C} \otimes \mathfrak{I}_{N'}$ is also ample. This implies that we may take C independent of N. By Remark 7, we have

$$\begin{aligned}
\kappa(k, N) &= C^{g-1} \sum_{i=1}^{g} \frac{k_i \{(\underline{\omega}^{(g-1)} \cdot \underline{\omega}_i) + (\mathfrak{I}^{(g-1)} \cdot \underline{\omega}_i)\}}{(\mathfrak{I}_N^{(g-1)} \cdot \sum_{E \in S} E)} \\
&\leq C^{g-1} \max_i \{k_i\} (\underline{\omega}^{(g)}) \\
&= C^{(g-1)} A \max_i \{k_i\} \zeta_F(-1) [\Gamma(1; \mathfrak{a}, \mathfrak{b}) : \Gamma(N; \mathfrak{a}, \mathfrak{b})] \\
&\ll \max_i \{k_i\} N^{3g}. \qquad\qquad \square
\end{aligned}$$

As a consequence of Theorem 1, Proposition 13, and Remark 10, we obtain an upper bound for the dimension of the vector space of the Hilbert modular forms.

Corollary 15. *Let N be an integer such that $N \geq 3$, $k = (k_1, \ldots, k_g) \in \mathbb{Z}_{\geq 0}^g$ a weight vector, and K a field that is an $\mathbb{O}_{\tilde{F}}[1/(Nd_F), \mu_N]$-algebra (a $\mathbb{Z}[1/(Nd_F), \mu_N]$-algebra if k is parallel). Then*

$$\dim_K H^0(\overline{M}_K, \underline{\omega}^k) \ll (\max_i \{k_i\})^g N^{3g^2}.$$

3.2. *The case that F is a real quadratic field.* In this section, for some special situations, we investigate the strange constant C appearing in Theorem 1.

Assume that $g = 2$ and the canonical bundle $K_{\overline{M}_K}$ is nef. For the invertible sheaf $\mathfrak{I}_N$ in Theorem 1, we set

$$\mathbb{O}_{\overline{M}_K}\left(-\sum_i n_i E_i\right) = \mathfrak{I}_N,$$

where $\{E_i\}$ are the exceptional curves of $\pi : \overline{M}_K \to M_K^*$ and n_i are positive integers. We set $n_{\max} = \max_i\{n_i\}$. Then we may take $2n_{\max}$ as the constant C.

Theorem 16. *In the above setting, we may take $C = 2n_{\max}$. Thus, in this situation, we may take*

$$\kappa = 2n_{\max} \frac{\sum_{j=1}^2 k_j \{(\underline{\omega} \cdot \underline{\omega}_j) - (\sum_i n_i E_i \cdot \underline{\omega}_j)\}}{-\left(\sum_i n_i E_i \cdot \sum_{E \in S} E\right)}.$$

Proof. For our purpose, we need a line bundle $\mathcal{L}$ on $\overline{M}$ that is nef and $(\mathcal{L} \cdot E) > 0$ for all exceptional curves E. Thus, it is sufficient to prove that $(\omega^{\otimes 2n_{\max}} \otimes \mathfrak{I}_N \cdot \mathscr{C}) \geq 0$ for every irreducible curve $\mathscr{C}$ in $\overline{M}_K$. In particular, it is not necessary to prove that the line bundle is ample. First, assume that the curve $\mathscr{C}$ is exceptional. Then, by the π-ampleness of $\mathfrak{I}_N$, we have

$$(\omega^{\otimes 2n_{\max}} \otimes \mathfrak{I}_N \cdot \mathscr{C}) = (\pi^* \pi_* \omega^{\otimes 2n_{\max}} \cdot \mathscr{C}) + (\mathfrak{I}_N \cdot \mathscr{C})$$
$$= (\pi_* \omega^{\otimes 2n_{\max}} \cdot \pi_* \mathscr{C}) + (\mathfrak{I}_N \cdot \mathscr{C}) = (\mathfrak{I}_N \cdot \mathscr{C}) > 0.$$

Next assume $\mathscr{C}$ is an irreducible curve such that $\pi(\mathscr{C})$ is an irreducible curve. Then

$$(\omega^{\otimes 2n_{\max}} \otimes \mathfrak{I}_N \cdot \mathscr{C}) = \left(\mathbb{O}_{\overline{M}_K}\left(n_{\max} K_{\overline{M}_K} + n_{\max} D_\infty - \sum_i n_i E_i\right) \cdot \mathscr{C}\right)$$

$$= (n_{\max} K_{\overline{M}_K} \cdot \mathscr{C}) + \left(\sum_i (n_{\max} - n_i) E_i \cdot \mathscr{C}\right) \geq 0. \qquad \square$$

We give an example for Theorem 16 in the setting of [Dieulefait et al. 2010, Appendix B].

Example 17. Let $F = \mathbb{Q}(\sqrt{5})$, $N = 3$, and $k_1 = k_2 = 2k \in 2\mathbb{Z}_{>0}$. Then there are the exceptional curves E_i ($i = 1, 2, \ldots, 10$) and the curve F_1 defined in [van der

Geer 1988, page 88]. These intersection numbers are as follows:

$$(E_i \cdot E_j) = \begin{cases} 0 & \text{if } i \neq j, \\ -4 & \text{if } i = j, \end{cases}$$
$$(F_1 \cdot F_1) = -60,$$
$$(E_i \cdot F_1) = 12 \quad \text{for } i = 1, 2, \ldots, 10.$$

We set

$$D' = \frac{1}{5}\left(\sum_{i=1}^{10} E_i + 2F_1\right),$$

then D' is nef and $(D' \cdot D') = 8$, and it is known that D' equals the canonical divisor $K_{\overline{M}}$. In particular, $\overline{M}$ is a minimal surface of general type. Then we may take 1 as $n_{\max}$ because we may use $\mathbb{O}(-D_\infty)$ as the ideal sheaf $\mathfrak{I}_3$, where $D_\infty = \sum_{i=1}^{10} E_i$. Thus, by easy calculation, we have $\kappa = 12k$. This estimate is the same as [Dieulefait et al. 2010, Theorem B.3].

Next we consider a more particular case. Assume that the toroidal compactification $\overline{M}_{\mathbb{F}}$ is a minimal surface of general type (i.e., the canonical divisor $K_{\overline{M}_{\mathbb{F}}}$ of $\overline{M}_{\mathbb{F}}$ is nef and $(K_{\overline{M}_{\mathbb{F}}} \cdot K_{\overline{M}_{\mathbb{F}}}) > 0$) over a field $\mathbb{F}$ of positive characteristic ℓ. Under the above assumption, we have:

Lemma 18 [Ekedahl 1988, Chapter III, Proposition 1.13]. *Let $\mathbb{F}$ be a field of positive characteristic, X a minimal surface of general type over $\mathbb{F}$, and Z the fundamental cycle on X. Then $2K_X - Z$ is numerically positive.*

Here the fundamental cycle of X is the fundamental cycle associated to the canonical morphism $\phi = \phi_{mK_X} : X \to X_{\text{can}}$ with a sufficiently large integer m, where X_{can} is the canonical model of X. The Weil divisor Z is the fundamental cycle of ϕ if Z is the smallest element in the set

$$\left\{ D = \sum m_i E_i \;\Big|\; E_i \in \text{Ex}(\phi),\; m_i > 0,\; (D \cdot E) \leq 0 \text{ for all } E \in \text{Ex}(\phi) \right\},$$

where $\text{Ex}(\phi)$ is the set of the exceptional curves of ϕ (i.e., the irreducible curves contracted by ϕ) on X. For an irreducible curve $\mathscr{C} \subset X$, $\mathscr{C}$ is contracted by ϕ if and only if $(K_X \cdot \mathscr{C}) = 0$. If $\mathscr{C} \simeq \mathbb{P}^1$ and $\mathscr{C}$ is contracted by ϕ, by the adjunction formula

$$p_a(\mathscr{C}) = 1 + \tfrac{1}{2}((\mathscr{C} \cdot \mathscr{C}) + (K_X \cdot \mathscr{C})),$$

we have $(\mathscr{C} \cdot \mathscr{C}) = -2$. Thus, X_{can} is obtained by contracting all (-2)-curves on X.

Return to our case. Let π be the morphism $\overline{M}_{\mathbb{F}} \to M_{\mathbb{F}}^*$, $\phi = \phi_{mK_{\overline{M}_{\mathbb{F}}}} : \overline{M}_{\mathbb{F}} \to \overline{M}_{\mathbb{F},\text{can}}$ with a sufficiently large integer m, and $\text{Ex}(\pi)$ and $\text{Ex}(\phi)$ the sets of the exceptional curves of π and ϕ, respectively. Then we have

$$\text{Ex}(\phi) = \{ E \in \text{Ex}(\pi) \mid (E \cdot E) = -2 \}.$$

We can take $Z = \sum_{E \in \text{Ex}(\phi)} E$ as the fundamental cycle of ϕ. Remark 6 implies

$$\mathbb{O}_{\overline{M}_{\mathbb{F}}}(2K_{\overline{M}_{\mathbb{F}}} - Z) \simeq \underline{\omega}^{\otimes 4} \otimes \mathbb{O}_{\overline{M}_{\mathbb{F}}}(-2D_\infty - Z).$$

Thus, we obtain:

Theorem 19. *Let $N \geq 3$ be an integer, $k = (k_1, k_2) \in \mathbb{Z}^2_{\geq 0}$, ℓ an odd prime number such that $\ell \nmid N d_F$, $\mathfrak{c}$ a nonzero integral ideal of F, and $\mathbb{F}$ a field of characteristic ℓ. Assume that $K_{\overline{M}_{\mathbb{F}}}$ is nef and $(K_{\overline{M}_{\mathbb{F}}} \cdot K_{\overline{M}_{\mathbb{F}}}) > 0$. Let S be nonempty finite set of exceptional curves of $\pi : \overline{M}_{\mathbb{F}} \to M^*_{\mathbb{F}}$. Let f be a $\mathfrak{c}$-polarized geometric Hilbert modular form over $\mathbb{F}$ of weight (k_1, k_2) and level $\Gamma(N)$. Then, if $f \not\equiv 0$, we have $\min_{E \in S}\{\text{ord}_E(f)\} < \kappa$. Here*

$$\kappa = \sum_{i=1}^2 \frac{k_i \{4(\underline{\omega} \cdot \underline{\omega}_i) - ((2D_\infty + Z) \cdot \underline{\omega}_j)\}}{\left((-2D_\infty - Z) \cdot \sum_{E \in S} E\right)},$$

where D_∞ and Z are the above Weil divisors.

Remark 20. In the settings of Theorem 19, $2n_{max} = 6$ if $\text{Ex}(\phi)$ is nonempty. Thus, when the weight is parallel (k, k) and $\text{Ex}(\phi)$ is not empty, Theorem 16 implies

$$\kappa = \frac{6k(\underline{\omega} \cdot \underline{\omega})}{\left((-2D_\infty - Z) \cdot \sum_{E \in S} E\right)}.$$

But on the same assumptions, Theorem 19 implies

$$\kappa = \frac{4k(\underline{\omega} \cdot \underline{\omega})}{\left((-2D_\infty - Z) \cdot \sum_{E \in S} E\right)}.$$

Therefore, the estimate of κ in Theorem 19 is slightly better than in Theorem 16.

Remark 21. In the situation of Theorems 16 and 19, if f is a cusp form, we can take a smaller bound as Remark 6. For example, under the assumption of Theorem 16, for f having zeros of order a at all the irreducible components of codimension 1 in $\overline{M} \setminus M$, the bound can be taken as

$$\kappa - a \left(\frac{(\sum_i n_i E_i \cdot D_\infty)}{\left(\sum_i n_i E_i \cdot \sum_{E \in S} E\right)} - 1 \right).$$

Acknowledgements

References

[Ash et al. 1975] A. Ash, D. Mumford, M. Rapoport, and Y. Tai, *Smooth compactification of locally symmetric varieties*, Lie Groups: History, Frontiers and Applications **4**, Math. Sci. Press, Brookline, MA, 1975. MR 56 #15642 Zbl 0334.14007

[Baba et al. 2002] S. Baba, K. Chakraborty, and Y. N. Petridis, "On the number of Fourier coefficients that determine a Hilbert modular form", *Proc. Amer. Math. Soc.* **130**:9 (2002), 2497–2502. MR 2003c:11037 Zbl 1125.11323

[Chai 1990] C.-L. Chai, "Arithmetic minimal compactification of the Hilbert–Blumenthal moduli spaces", *Ann. of Math.* (2) **131**:3 (1990), 541–554. MR 91i:11063 Zbl 0754.14030

[Deligne and Rapoport 1973] P. Deligne and M. Rapoport, "Les schémas de modules de courbes elliptiques", pp. 143–316 in *Modular functions of one variable* (Antwerp, 1972), vol. 2, edited by P. Deligne and W. Kuyk, Lecture Notes in Mathematics **349**, Springer, Berlin, 1973. MR 49 #2762 Zbl 0281.14010

[Dieulefait et al. 2010] L. Dieulefait, A. Pacetti, and M. Schuett, "Modularity of the Consani–Scholten quintic", preprint, 2010. arXiv 1005.4523

[Dimitrov 2004] M. Dimitrov, "Compactifications arithmétiques des variétés de Hilbert et formes modulaires de Hilbert pour $\Gamma_1(\mathfrak{c}, \mathfrak{n})$", pp. 527–554 in *Geometric aspects of Dwork theory*, edited by A. Adolphson et al., de Gruyter, Berlin, 2004. MR 2006e:11063 Zbl 1076.14029

[Dimitrov and Tilouine 2004] M. Dimitrov and J. Tilouine, "Variétés et formes modulaires de Hilbert arithmétiques pour $\Gamma_1(\mathfrak{c}, \mathfrak{n})$", pp. 555–614 in *Geometric aspects of Dwork theory*, edited by A. Adolphson et al., de Gruyter, Berlin, 2004. MR 2006e:11064 Zbl 1127.11036

[Doi and Ohta 1977] K. Doi and M. Ohta, "On some congruences between cusp forms on $\Gamma_0(N)$", pp. 91–105 in *Modular functions of one variable* (Bonn, 1976), vol. 5, Lecture Notes in Mathematics **601**, Springer, Berlin, 1977. MR 57 #235 Zbl 0361.10023

[Ekedahl 1988] T. Ekedahl, "Canonical models of surfaces of general type in positive characteristic", *Inst. Hautes Études Sci. Publ. Math.* **67** (1988), 97–144. MR 89k:14069 Zbl 0674.14028

[Faltings and Chai 1990] G. Faltings and C.-L. Chai, *Degeneration of abelian varieties*, Ergeb. Math. Grenzgeb. (3) **22**, Springer, Berlin, 1990. MR 92d:14036 Zbl 0744.14031

[Fulton 1998] W. Fulton, *Intersection theory*, 2nd ed., Ergeb. Math. Grenzgeb. (3) **2**, Springer, Berlin, 1998. MR 99d:14003 Zbl 0885.14002

[van der Geer 1988] G. van der Geer, *Hilbert modular surfaces*, Ergeb. Math. Grenzgeb. (3) **16**, Springer, Berlin, 1988. MR 89c:11073 Zbl 0634.14022

[Hartshorne 1977] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics **52**, Springer, New York, 1977. MR 57 #3116 Zbl 0367.14001

[Hida 2004] H. Hida, *p-adic automorphic forms on Shimura varieties*, Springer, New York, 2004. MR 2005e:11054 Zbl 1055.11032

[Katz 1978] N. M. Katz, "*p*-adic *L*-functions for CM fields", *Invent. Math.* **49**:3 (1978), 199–297. MR 80h:10039 Zbl 0417.12003

[Lazarsfeld 2004] R. Lazarsfeld, *Positivity in algebraic geometry, I: Classical setting, line bundles and linear series*, Ergeb. Math. Grenzgeb. (3) **48**, Springer, Berlin, 2004. MR 2005k:14001a Zbl 1093.14501

[Miyake 1989] T. Miyake, *Modular forms*, Springer, Berlin, 1989. MR 90m:11062 Zbl 0701.11014

[Moret-Bailly 1985] L. Moret-Bailly, *Pinceaux de variétés abéliennes*, Astérisque **129**, Société Mathématique de France, Paris, 1985. MR 87j:14069 Zbl 0595.14032

[Mumford 1977] D. Mumford, "Hirzebruch's proportionality theorem in the noncompact case", *Invent. Math.* **42** (1977), 239–272. MR 81a:32026 Zbl 0365.14012

[Rapoport 1978] M. Rapoport, "Compactifications de l'espace de modules de Hilbert–Blumenthal", *Compositio Math.* **36**:3 (1978), 255–335. MR 80j:14009 Zbl 0386.14006

[Sturm 1987] J. Sturm, "On the congruence of modular forms", pp. 275–280 in *Number theory* (New York, 1984–1985), edited by D. V. Chudnovsky et al., Lecture Notes in Mathematics **1240**, Springer, Berlin, 1987. MR 88h:11031 Zbl 0615.10035

[Takai 2010] Y. Takai, *An analogy of Sturm's theorem for real quadratic fields*, Ph.D. thesis, Nagoya University, 2010.

takai@ms.u-tokyo.ac.jp   *Graduate School of Mathematical Sciences, University of Tokyo, 3-8-1 Komaba, Meguro, Tokyo, 153-8914, Japan*

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.

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 7   No. 4   2013