Algebra & Number Theory

Volume 11
2017
No. 9

msp

Algebra & Number Theory

msp.org/ant

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

The subscription price for 2017 is US $325/year for the electronic version, and $520/year (+$55, if shipping outside the US) for print and electronic. Subscriptions, requests for back issues and changes of subscriber address should be sent to MSP.

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

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

PUBLISHED BY

mathematical sciences publishers

nonprofit scientific publishing

http://msp.org/

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.1967

A nonarchimedean Ax–Lindemann theorem

Antoine Chambert-Loir and François Loeser

À Daniel Bertrand, en témoignage d'amitié

Motivated by the André–Oort conjecture, Pila has proved an analogue of the Ax–Lindemann theorem for the uniformization of classical modular curves. In this paper, we establish a similar theorem in nonarchimedean geometry. Precisely, we give a geometric description of subvarieties of a product of hyperbolic Mumford curves such that the irreducible components of their inverse image by the Schottky uniformization are algebraic, in some sense. Our proof uses a p-adic analogue of the Pila–Wilkie theorem due to Cluckers, Comte and Loeser, and requires that the relevant Schottky groups have algebraic entries.

1. Introduction

1.1. The classical Lindemann–Weierstrass theorem states that if algebraic numbers $\alpha_1, \dots, \alpha_n$ are $\mathbf{Q}$-linearly independent, then their exponentials $\exp(\alpha_1), \dots, \exp(\alpha_n)$ are algebraically independent over $\mathbf{Q}$. More generally, if $\alpha_1, \dots, \alpha_n$ are any $\mathbf{Q}$-linearly independent complex numbers, no longer assumed to be algebraic, Schanuel's conjecture predicts that the field $\mathbf{Q}(\alpha_1, \dots, \alpha_n, \exp(\alpha_1), \dots, \exp(\alpha_n))$ has transcendence degree at least n over $\mathbf{Q}$. Ax [1971] established power series and differential field versions of Schanuel's conjecture. In particular, the part of Ax's results corresponding to the Lindemann–Weierstrass theorem can be recast into geometrical terms as follows:

Theorem 1.2 (exponential Ax–Lindemann). *Let* $\exp : \mathbf{C}^n \to (\mathbf{C}^\times)^n$ *be the morphism* $(z_1, \dots, z_n) \mapsto (\exp(z_1), \dots, \exp(z_n))$. *Let V be an irreducible algebraic subvariety of* $(\mathbf{C}^\times)^n$ *and let W be an irreducible component of a maximal algebraic subvariety of* $\exp^{-1}(V)$. *Then W is geodesic, that is, W is defined by a finite family of equations of the form* $\sum_{i=1}^n a_i z_i = b$ *with* $a_1, \dots, a_n \in \mathbf{Q}$ *and* $b \in \mathbf{C}$.

In a breakthrough paper, Pila [2011] succeeded in providing an unconditional proof of the André–Oort conjecture for products of modular curves. One of his

MSC2010: primary 11G18; secondary 03C98, 11D88, 11J91, 14G22, 14G35.
Keywords: Schottky group, Ax–Lindemann theorem, Pila–Wilkie theorem, nonarchimedean analytic geometry.

main ingredients was to prove a hyperbolic version of the above Ax–Lindemann theorem, which we now state in a simplified version.

Let $\boldsymbol{h}$ denote the complex upper half-plane and $j : \boldsymbol{h} \to \mathbf{C}$ the elliptic modular function. By an algebraic subvariety of $\boldsymbol{h}^n$, we mean the trace in $\boldsymbol{h}^n$ of an algebraic subvariety of $\mathbf{C}^n$. An algebraic subvariety of $\boldsymbol{h}^n$ is said to be geodesic if it can be defined by equations of the form $z_i = c_i$ and $z_k = g_{k\ell} z_\ell$, with $c_i \in \mathbf{C}$ and $g_{k\ell} \in \mathrm{GL}(2, \mathbf{Q})^+$.

Theorem 1.3 (hyperbolic Ax–Lindemann). *Let $j : \boldsymbol{h}^n \to \mathbf{C}^n$ be the morphism $(z_1, \dots, z_n) \mapsto (j(z_1), \dots, j(z_n))$. Let V be an irreducible algebraic subvariety of $\mathbf{C}^n$ and let W be an irreducible component of a maximal algebraic subvariety of $j^{-1}(V)$. Then W is geodesic.*

Pila's method to prove this Ax–Lindemann theorem is quite different from the differential approach of Ax. It follows a strategy initiated by Pila and Zannier [2008] in their new proof of the Manin–Mumford conjecture for abelian varieties; that approach makes crucial use of the bound on the number of rational points of bounded height in the transcendental part of sets definable in an o-minimal structure obtained in [Pila and Wilkie 2006]. Recently, still using the Pila and Zannier strategy, Klingler, Ullmo and Yafaev [Klingler et al. 2016] have succeeded in proving a very general form of the hyperbolic Ax–Lindemann theorem valid for any arithmetic variety; see also [Ullmo and Yafaev 2014] for the compact case.

1.4. In the recent paper [Cluckers et al. 2015], Cluckers, Comte and Loeser established a nonarchimedean analogue of the Pila–Wilkie theorem of [Pila and Wilkie 2006] in its block version of [Pila 2009]. The purpose of this paper is to use this result to prove a version of Ax–Lindemann for products of algebraic curves admitting a nonarchimedean uniformization and whose corresponding Schottky group is “arithmetic” and has rank at least 2 (Theorem 2.7). In particular, this theorem applies for products of Shimura curves admitting a p-adic uniformization à la Čerednik–Drinfel'd (see Section 3).

The basic strategy we use is strongly inspired by that of [Pila 2011] (see also [Pila 2015]), though some new ideas are required in order to adapt it to the nonarchimedean setting. Similarly as in Pila's approach one starts by working on some neighborhood of the boundary of our space (which, instead of a product of Poincaré upper half-planes, is a product of open subsets of the Berkovich projective line). Analytic continuation and monodromy arguments are replaced by more algebraic ones and explicit matrix computations by group theory considerations. We also take advantage of the fact that Schottky groups are free and of the geometric description of their fundamental domains. Compared with Pila's proof, where parabolic elements are used in a crucial way, one main difficulty of the nonarchimedean situation lies in the fact that all nontrivial elements of a Schottky groups are hyperbolic.

To conclude, let us note that there are cases where p-adic analogues of theorems in transcendental number theory seem to require other methods than those used to prove their complex counterparts. For instance, it is still an open problem to prove a p-adic analogue, for values of the p-adic exponential function, of the classical Lindemann–Weierstrass theorem.

Since his first works (see, for example, [Bertrand 1976]), Daniel Bertrand has shown deep insight into p-adic transcendental number theory, and disseminated his vision within the mathematical community. We are pleased to dedicate this paper to him.

2. Statement of the theorem

2.1. ***Nonarchimedean analytic spaces.*** Given a complete nonarchimedean valued field F, we consider in this paper F-analytic spaces in the sense of Berkovich [1990; 1993]. However, the statements, and essentially the proofs, can be carried on mutatis mutandis in the rigid analytic setting. In this context, there is a notion of irreducible component; see [Ducros 2009], or [Conrad 1999] for the rigid analytic version.

If V is an algebraic variety over F, we denote by V^{an} the corresponding F-analytic space. There is a canonical topological embedding of $V(F)$ in V^{an}, and its image is closed if F is locally compact.

If F' is a complete nonarchimedean extension of F, we denote by $X_{F'}$ the F'-analytic space deduced from an F-analytic space X by base change to F'.

2.2. ***Schottky groups.*** Let p be a prime number; we denote by $\mathbf{C}_p$ the completion of an algebraic closure of $\mathbf{Q}_p$ and let F be a finite extension of $\mathbf{Q}_p$ contained in $\mathbf{C}_p$. The group $\mathrm{PGL}(2, F)$ acts by homographies on the F-analytic projective line $\mathbf{P}_1^{\mathrm{an}}$. In the next paragraphs, we recall from [Gerritzen and van der Put 1980] a few definitions concerning Schottky groups in $\mathrm{PGL}(2, F)$, their limit sets and the associated uniformizations of algebraic curves.

One says that a discrete subgroup Γ of $\mathrm{PGL}(2, F)$ is a *Schottky group* if it is finitely generated, and if no element ($\neq$ id) has finite order [Gerritzen and van der Put 1980, I, (1.6)]. If Γ is a Schottky group, then Γ is free; moreover, any discrete finitely generated subgroup of $\mathrm{PGL}(2, F)$ possesses a normal subgroup of finite index which is a Schottky group [Gerritzen and van der Put 1980, I, (3.1)].

We say that Γ is *arithmetic* if its elements can be represented by matrices whose coefficients lie in a number field. In this case, it follows from the hypothesis that Γ is finitely generated that there exists a number field $K \subset F$ such that $\Gamma \subset \mathrm{PGL}(2, K)$.

2.3. ***Limit sets.*** Let Γ be a Schottky subgroup of $\mathrm{PGL}(2, F)$. Its *limit set* is the set $\mathscr{L}_\Gamma$ of all points in $\mathbf{P}_1(\mathbf{C}_p)$ of the form $\lim_n(\gamma_n \cdot x)$, where (γ_n) is a sequence of distinct elements of Γ and $x \in \mathbf{P}_1(\mathbf{C}_p)$ [Gerritzen and van der Put 1980, I, (1.3)].

By [Gerritzen and van der Put 1980, I, (1.6)], the limit set $\mathscr{L}_\Gamma$ is a compact subset of $\mathbf{P}_1(F)$. If the rank of Γ is at least 2, then $\mathscr{L}_\Gamma$ is a perfect (that is, closed and without isolated points) subset of $\mathbf{P}_1(F)$; see [Gerritzen and van der Put 1980, I, (1.6.3) and (1.7.2)].

Let $\Omega_\Gamma = (\mathbf{P}_1)^{\mathrm{an}} - \mathscr{L}_\Gamma$; it is a Γ-invariant open set of $\mathbf{P}_1^{\mathrm{an}}$. By Lemma 5.4 below, it is geometrically irreducible.

2.4. ***Quotients.*** Let us assume that Γ is a Schottky group and let g be its rank. From the explicit description of the action of the group Γ given by [Gerritzen and van der Put 1980, I.4] and recalled in Section 6.5 below (see also [Berkovich 1990, p. 86]), it follows that the group Γ acts freely on Ω_Γ, and the quotient space Ω_Γ/Γ admits a unique structure of an F-analytic space such that the projection $p_\Gamma : \Omega_\Gamma \to \Omega_\Gamma/\Gamma$ is both a topological covering and a local isomorphism. Moreover, Ω_Γ/Γ is the F-analytic space associated with a smooth, geometrically connected, projective F-curve X_Γ of genus g [Gerritzen and van der Put 1980, III, (2.2); Berkovich 1990, Theorem 4.4.1, p. 86], canonically determined by the GAGA theorem in this context, [Berkovich 1990, Theorem 3.4.12, p. 68].

2.5. Let us now consider a finite family $(\Gamma_i)_{1\le i\le n}$ of Schottky subgroups of $\mathrm{PGL}(2, F)$ of rank ≥ 2. Let us set $\Omega = \prod_{i=1}^n \Omega_{\Gamma_i}$ and $X = \prod_{i=1}^n X_{\Gamma_i}$, and let $p : \Omega \to X^{\mathrm{an}}$ be the morphism deduced from the morphisms $p_{\Gamma_i} : \Omega_{\Gamma_i} \to X_{\Gamma_i}^{\mathrm{an}}$.

2.6. ***Flat subvarieties.*** Let K be a complete extension of F and let W be a closed analytic subspace of Ω_K.

The following terminology is borrowed from the analogous notions in the differential geometry of hermitian symmetric domains.

We say that W is *irreducible algebraic* if there exists a K-algebraic subvariety Y of $(\mathbf{P}_1^n)_K$ such that W is an irreducible component of the analytic space $\Omega_K \cap Y^{\mathrm{an}}$. In this case, one can take for Y the Zariski closure of W in $(\mathbf{P}_1^n)_K$; it is irreducible and satisfies $\dim(Y) = \dim(W)$; see [Ducros 2009, Proposition 4.22].

We say that W is *flat* if it can be defined by equations of the following form:

(1) $z_i = c$ for some $i \in \{1, \dots, n\}$ and $c \in \Omega_{\Gamma_i}(K)$;

(2) $z_j = g \cdot z_i$ for some pair (i, j) of distinct elements of $\{1, \dots, n\}$ and some $g \in \mathrm{PGL}(2, F)$.

Assume that W is flat and let Y be the subvariety of $(\mathbf{P}_1^n)_K$ defined by equations of this form which define W on Ω_K. There exists a subset I of $\{1, \dots, n\}$ such that the projection $q_I : \mathbf{P}_1^n \to \mathbf{P}_1^I$ given by the coordinates in I induces an isomorphism of Y to $(\mathbf{P}_1^I)_K$. This implies that q_I induces an isomorphism from W to $\prod_{i\in I} \Omega_{i,K}$. In particular, W is irreducible, even geometrically irreducible, and hence is irreducible algebraic. Conversely, we observe that if W is geometrically irreducible and if there exists a complete extension L of K such that W_L is flat, then W is flat.

We say that W is *geodesic* if, moreover, the elements g in (2) can be taken such that $g\Gamma_i g^{-1}$ and Γ_j are commensurable (i.e., their intersection has finite index in both of them).

Here is the main result of this paper.

Theorem 2.7 (nonarchimedean Ax–Lindemann theorem). *Let F be a finite extension of $\mathbf{Q}_p$ and let $(\Gamma_i)_{1\leq i\leq n}$ be a finite family of* arithmetic *Schottky subgroups of* $\mathrm{PGL}(2, F)$ *of ranks ≥ 2. As above, let us set $\Omega=\prod_{i=1}^n \Omega_{\Gamma_i}$ and $X=\prod_{i=1}^n X_{\Gamma_i}$, and let $p:\Omega\to X^{\mathrm{an}}$ be the morphism deduced from the morphisms $p_{\Gamma_i}:\Omega_{\Gamma_i}\to X_{\Gamma_i}^{\mathrm{an}}$.*

Let V be an irreducible algebraic subvariety of X and let W be an irreducible algebraic subvariety of Ω, maximal among those contained in $p^{-1}(V^{\mathrm{an}})$. Then every irreducible component of $W_{\mathbf{C}_p}$ is flat.

The proof of this theorem is given in Section 8; it follows the strategy of Pila–Zannier. In the archimedean setting, this strategy relies crucially on a theorem of Pila–Wilkie about rational points on definable sets; we recall in Section 4 the nonarchimedean analogue of this theorem [Cluckers et al. 2015] which is used here. It is at this point that we need the assumption that the group Γ be arithmetic. This restriction is inherent to Pila–Zannier's strategy and we do not know whether it can be bypassed.

In Section 6, we recall a few more facts on p-adic Schottky groups and p-adic uniformization, essentially borrowed from [Gerritzen and van der Put 1980].

In a final section, we prove a characterization (Theorem 9.2) of geodesic subvarieties of Ω as the geometrically irreducible algebraic subvarieties whose projection to X is algebraic ("bialgebraic subvarieties"), in analogy with what happens in the context of Ax's theorem or of Shimura varieties.

3. The example of Shimura curves

We begin by recalling the definition of Shimura curves and their p-adic uniformization. The literature is unfortunately rather scattered; we refer to [Boutot and Carayol 1992] for more detail, as well as to [Clark 2003, Chapter 0].

3.1. ***Complex Shimura curves.*** Let B be a quaternion division algebra with center $\mathbf{Q}$; we assume that it is indefinite, namely $B\otimes_{\mathbf{Q}}\mathbf{R}\simeq \mathrm{M}_2(\mathbf{R})$. Let then $\mathscr{O}_B$ be a maximal order of B, that is a maximal subring of B which is isomorphic to $\mathbf{Z}^4$ as a $\mathbf{Z}$-module. Let H be the algebraic group of units of $\mathscr{O}_B$, modulo center, considered as a $\mathbf{Z}$-group scheme. For every field R containing $\mathbf{Q}$, one has $H(R)=(B\otimes_{\mathbf{Q}} R)^\times/Z((B\otimes_{\mathbf{Q}} R)^\times)$; in particular, the group $H(\mathbf{R})$ is isomorphic to $\mathrm{PGL}(2,\mathbf{R})$, and we fix such an isomorphism. Then the group $H(\mathbf{R})$ acts by homographies on the double Poincaré upper half-plane

$$\boldsymbol{h}^{\pm}=\mathbf{C}-\mathbf{R}.$$

Let also Δ be a congruence subgroup of $H(\mathbf{Z})$; recall that this means that there exists an integer $N \geq 1$ such that Δ contains the kernel of the canonical morphism $H(\mathbf{Z}) \to H(\mathbf{Z}/N\mathbf{Z})$. We assume that Δ has been chosen small enough so that the stabilizer of every point of $\boldsymbol{h}^{\pm}$ is trivial. The quotient $\boldsymbol{h}^{\pm}/\Delta$ has a natural structure of a compact Riemann surface and the projection $p : \boldsymbol{h}^{\pm} \to \boldsymbol{h}^{\pm}/\Delta$ is an étale covering.

This curve parameterizes triples (V, ι, ν), where V is a complex two-dimensional abelian variety, $\iota : \mathscr{O}_B \to \operatorname{End}(V)$ is a faithful action of $\mathscr{O}_B$ on V and ν is a level structure "of type Δ" on V. When Δ is the kernel of $H(\mathbf{Z})$ to $H(\mathbf{Z}/N\mathbf{Z})$, for some integer $N \geq 1$, such a level structure corresponds to an equivariant isomorphism of V_N, the subgroup of N-torsion of V, with $\mathscr{O}_B/N\mathscr{O}_B$.

By [Shimura 1961], it admits a canonical structure of an algebraic curve S which can be defined over a number field E in $\mathbf{C}$.

3.2. ***The p-adic uniformization of Shimura curves.*** Let p be a prime number at which B ramifies, which means that $B \otimes_{\mathbf{Q}} \mathbf{Q}_p$ is a division algebra. Let also F be the completion of the field E at a place dividing p; we denote by $\mathbf{C}_p$ the p-adic completion of an algebraic closure of F. We still denote by S the F-curve deduced from an E-model of the complex curve S.

Let $\Omega = (\mathbf{P}_1)_F^{\mathrm{an}} - \mathbf{P}_1(\mathbf{Q}_p)$ be the extension of scalars to F of Drinfel'd's upper half-plane. According to the theorem of Čerednik and Drinfel'd [Čerednik 1976; Drinfel'd 1976] (see also [Boutot and Carayol 1992] for a detailed exposition), and up to replacing F by a finite unramified extension, the F-analytic curve S^{an} admits a "p-adic uniformization" which takes the form of a surjective analytic morphism

$$j : \Omega \to S^{\mathrm{an}},$$

identifying S^{an} with the quotient of Ω by the action of a subgroup Γ of $\mathrm{PGL}(2, \mathbf{Q}_p)$. Up to replacing Δ by a smaller congruence subgroup, which replaces S by a finite (possibly ramified) covering, we may also assume that Γ is a p-adic Schottky subgroup acting freely on Ω, and that j is topologically étale. Then the morphism $j : \Omega \to S^{\mathrm{an}}$ is the universal cover of S^{an}.

Let us describe this subgroup. Let A be the quaternion division algebra over $\mathbf{Q}$ with the same invariants as B, except for those invariants at p and ∞ which are switched. In particular, $A \otimes_{\mathbf{Q}} \mathbf{R}$ is Hamilton's quaternion algebra, while $A \otimes_{\mathbf{Q}} \mathbf{Q}_p \simeq \mathrm{M}_2(\mathbf{Q}_p)$. Let G be the algebraic group of units of A, modulo center; in particular, $G(\mathbf{Q}_p) \simeq \mathrm{PGL}(2, \mathbf{Q}_p)$. As explained in [Boutot and Carayol 1992], the discrete subgroup Γ is the intersection of $G(\mathbf{Q})$ with a compact open subgroup of $G(\mathbf{A}_{\mathrm{f}})$, the adelic group associated with G where the place at ∞ is omitted.

Lemma 3.3. *The group Γ is conjugate to an arithmetic Schottky subgroup in* $\mathrm{PGL}(2, \mathbf{Q}_p)$, *its rank is at least* 2, *and its limit set is equal to* $\mathbf{P}_1(\mathbf{Q}_p)$.

Proof. The group Γ is a discrete subgroup of $\mathrm{PGL}(2, \mathbf{Q}_p)$, so its limit set $\mathscr{L}_\Gamma$ is a Γ-invariant subset of $\mathbf{P}_1(\mathbf{Q}_p)$. In other words, the Drinfeld upper half-plane $\Omega = \mathbf{P}_1^{\mathrm{an}} - \mathbf{P}_1(\mathbf{Q}_p)$ is an open subset of $\Omega_\Gamma = \mathbf{P}_1^{\mathrm{an}} - \mathscr{L}_\Gamma$. By the theory of Mumford curves and Schottky groups (see [Gerritzen and van der Put 1980]), the analytic curve $(\mathbf{P}_1^{\mathrm{an}} - \mathscr{L}_\Gamma)/\Gamma$ is algebraic, and admits the analytic curve $S^{\mathrm{an}} = \Omega/\Gamma$ as an open subset. According to the Čerednik–Drinfel'd theorem, the curve S^{an} is projective. This implies that $\Omega = \mathbf{P}_1^{\mathrm{an}} - \mathscr{L}_\Gamma$, and hence $\mathscr{L}_\Gamma = \mathbf{P}_1(\mathbf{Q}_p)$.

After base change to $\mathbf{Q}_p$, the algebraic $\mathbf{Q}$-group G becomes isomorphic to $\mathrm{PGL}(2)_{\mathbf{Q}_p}$. Consequently, there exists a finite algebraic extension K of $\mathbf{Q}$, contained in $\mathbf{Q}_p$, such that $G_K \simeq \mathrm{PGL}(2)_K$. By such an isomorphism, $G(\mathbf{Q})$ is mapped into $\mathrm{PGL}(2, K)$; this implies that the group Γ is conjugate to an arithmetic group.

Since Γ is a Schottky group, it is free. Since it is nonabelian, its rank is at least 2. □

By this lemma, the following result is a special case of our main theorem (Theorem 2.7).

Theorem 3.4. *Let F be a finite extension of $\mathbf{Q}_p$, let $\Omega = (\mathbf{P}_1)_F^{\mathrm{an}} - \mathbf{P}_1(\mathbf{Q}_p)$ and let $j : \Omega^n \to S^{\mathrm{an}}$ be the Čerednik–Drinfel'd uniformization of a product of Shimura curves. Let V be an irreducible algebraic subvariety of S and let $W \subset \Omega^n$ be a maximal irreducible algebraic subvariety of $j^{-1}(V^{\mathrm{an}})$. Then every irreducible component of $W_{\mathbf{C}_p}$ is flat.*

3.5. By the same arguments, one can show that Theorem 2.7 also applies to the uniformizations of Shimura curves associated with quaternion division algebras over totally real fields, as considered by Čerednik [1976] and Boutot and Zink [1995].

3.6. As suggested by J. Pila and explained to us by Y. André, Theorem 3.4 can also be deduced from its complex analogue, which is a particular case of [Ullmo and Yafaev 2014]. The crucial ingredient is a deep theorem of André [2003, III, 4.7.4] stating that the p-adic uniformization and the complex uniformization of Shimura curves satisfy the *same* nonlinear differential equation. His proof relies on a delicate description of the Gauss–Manin equation in terms of convergent crystals and on the tempered fundamental group introduced by him. From that point on, one can apply Seidenberg's embedding theorem [1958] in differential algebra to prove that both the complex and nonarchimedean Ax–Lindemann theorems are equivalent to a single statement in differential algebra, in the original spirit of [Ax 1971].

4. Definability — a p-adic Pila–Wilkie theorem

4.1. There are two distinct notions of p-adic analytic geometry: one is "naïve", and the other rigid analytic. (Regarding rigid analytic geometry, we work in the framework defined by Berkovich.) These two notions give rise to three classes of sets, and we use them all in this paper. Let F be a finite extension of $\mathbf{Q}_p$.

a) *Semialgebraic* and *subanalytic* subsets of $\mathbf{Q}_p^n$ are defined by Denef and van den Dries [1988]; see also [Cluckers et al. 2015, p. 26].

   Replacing $\mathbf{Q}_p$ by a finite extension F, this leads to an analogous notion of F-semialgebraic, or F-subanalytic, subset of F^n. Considering affine charts, one then defines F-semialgebraic or F-subanalytic subsets of $V(F)$, for every (quasiprojective, say) algebraic variety V defined over F.

   On the other hand, the Weil restriction functor assigns to V an algebraic variety W defined over $\mathbf{Q}_p$ together with a canonical identification $V(F) \to W(\mathbf{Q}_p)$; we say that a subset of $V(F)$ is $\mathbf{Q}_p$-semialgebraic or $\mathbf{Q}_p$-subanalytic if its image in $W(\mathbf{Q}_p)$ is $\mathbf{Q}_p$-semialgebraic or $\mathbf{Q}_p$-subanalytic, respectively. Observe that F-semialgebraic subsets of $V(F)$ are $\mathbf{Q}_p$-semialgebraic, and that F-subanalytic subsets of $V(F)$ are $\mathbf{Q}_p$-subanalytic.

   Recall that an F-subanalytic subset S is said to be smooth of dimension d at a point x if it possesses a neighborhood U which is isomorphic to the unit ball of F^d; then S is smooth of dimension d at every point of U.

b) Lipshitz [1993] defined a notion of *rigid subanalytic subset* of $\mathbf{C}_p^n$. We use in this paper the variant [Lipshitz and Robinson 2000a, Definition 2.1.1] where the coefficients of all polynomials and power series involved belong to F; we call them *rigid F-subanalytic.* The notion extends to subsets of $V(\mathbf{C}_p)$, where V is an algebraic variety defined over F.

These classes of sets are stable under boolean operations and projections [Lipshitz and Robinson 2000b, Corollary 4.3], admit cell decompositions [Cluckers et al. 2006, Theorem 7.4], a natural notion of dimension (in fact, they are b-minimal in the sense of [Cluckers and Loeser 2007]), as well as a natural notion of smoothness.

Lemma 4.2. *Let F be a finite extension of $\mathbf{Q}_p$ contained in $\mathbf{C}_p$ and let V be an algebraic variety over F. Let Z be a rigid F-subanalytic subset of $V(\mathbf{C}_p)$. Then $Z(F) = Z \cap V(F)$ is an F-subanalytic subset of $V(F)$.*

Proof. We may assume that $V = \mathbf{A}^n$. Then Z can be defined by a quantifier-free formula of the above-mentioned variant of Lipshitz's analytic language, and our claim follows from the very definition of this language. $\square$

4.3. A *block* in $\mathbf{Q}_p^n$ is either empty, or a singleton, or a smooth subanalytic subset of pure dimension $d > 0$ which is contained in a smooth semialgebraic subset of dimension d.

A *family of blocks* in $\mathbf{Q}_p^n \times \mathbf{Q}_p^s$ is a subanalytic subset W such that there exists an integer $t \geq 0$ and a semialgebraic set $Z \subset \mathbf{Q}_p^n \times \mathbf{Q}_p^t$ such that for every $\sigma \in \mathbf{Q}_p^s$, there exists $\tau \in \mathbf{Q}_p^t$ such that the fibers W_σ and Z_τ are smooth of the same dimension, and $W_\sigma \subset Z_\tau$. (In particular, the sets W_σ, for $\sigma \in \mathbf{Q}_p^s$, are blocks in $\mathbf{Q}_p^n$.)

Let F be a finite extension of $\mathbf{Q}_p$. Considering Weil restriction, we deduce from these notions the definition of a block in F^n, or of a family of blocks in $F^n \times \mathbf{Q}_p^t$.

4.4. Let H be the standard height function on $\overline{\mathbf{Q}}$; for $x \in \mathbf{Q}$, written as a fraction a/b in lowest terms, one has $H(x) = \max(|a|, |b|)$. We also write H for the height function on $\overline{\mathbf{Q}}^n$ defined by $H(x_1, \dots, x_n) = \max_i(H(x_i))$. Viewing $\mathrm{GL}(d, \overline{\mathbf{Q}})$ as a subspace of $\overline{\mathbf{Q}}^{d^2}$, it defines a height function on $\mathrm{GL}(d, \overline{\mathbf{Q}})$. There exists a strictly positive real number c such that $H(gg') \le cH(g)H(g')$ for every $g, g' \in \mathrm{GL}(d, \overline{\mathbf{Q}})$, and $H(g^{-1}) \ll H(g)^c$ for every $g \in \mathrm{GL}(d, \overline{\mathbf{Q}})$. When $d = 2$ and $g \in \mathrm{SL}(2, \overline{\mathbf{Q}})$, one even has $H(g^{-1}) = H(g)$.

Consider $g \in \mathrm{GL}(d, \overline{\mathbf{Q}})$. If g is diagonal, then $H(g^n) = H(g)^n$ for every $n \in \mathbf{Z}$. More generally, if g is *semisimple*, then we have upper and lower bounds $H(g)^n \ll H(g^n) \ll H(g)^n$ for every $n \in \mathbf{Z}$.

By abuse of language, if G is a linear algebraic $\overline{\mathbf{Q}}$-group, we implicitly choose an embedding in some linear group, which furnishes a height function H on $G(\overline{\mathbf{Q}})$.

The actual choice of this height function depends on the chosen embedding, but any other height function H' is equivalent, in the sense that there is a strictly positive real number c such that $H(x)^{1/c} \ll H'(x) \ll H(x)^c$ for every $x \in G(\overline{\mathbf{Q}})$.

4.5. Let Z be a subset of F^n and let K be a finite extension of $\mathbf{Q}$ contained in F. We write $Z(K) = Z \cap K^n$ (K-rational points of Z). For every real number T, we define $Z(K; T) = \{x \in Z(K) : H(x) \le T\}$; for every integer D, we also define $Z(D; T)$ to be the set of points $x \in Z(F)$ such that $[\mathbf{Q}(x_i) : \mathbf{Q}] \le D$ for every $i \in \{1, \dots, n\}$ and $H(x) \le T$. These are finite sets.

We say that Z has *many K-rational points* if there exist strictly positive real numbers c, α such that

$$\mathrm{Card}(Z(K; T)) \ge cT^\alpha$$

for all T large enough. This notion only depends on the equivalence class of the height.

4.6. In [Cluckers et al. 2015], Cluckers, Comte and Loeser established a p-adic analogue of a theorem of Pila and Wilkie [2006] concerning the rational points of a definable set. We will use the following variant of [Cluckers et al. 2015, Theorem 4.2.3].

Theorem 4.7. *Let F be a finite extension of $\mathbf{Q}_p$ and let K be a finite extension of $\mathbf{Q}$ contained in F. Let $Z \subset F^n$ be a $\mathbf{Q}_p$-subanalytic subset. Let $\varepsilon > 0$. There exist $s \in \mathbf{N}$, $c \in \mathbf{R}$ and a family of blocks $W \subset Z \times \mathbf{Q}_p^s$ satisfying the following property: for every $T > 1$, there exists a subset $S_T \subset \mathbf{Q}_p^s$ of cardinality $< cT^\varepsilon$ such that $Z(K; T) \subset \bigcup_{\sigma \in S_T} W_\sigma$.*

Proof. Let $d = [F : \mathbf{Q}_p]$. By Krasner's lemma, there exists an algebraic number $e \in F$ of degree d such that $F = \mathbf{Q}_p(e)$. Then the basis $(1, e, \dots, e^{d-1})$ defines a $\mathbf{Q}_p$-linear bijection $\psi : \mathbf{Q}_p^d \xrightarrow{\sim} F$, $(x_1, \dots, x_d) \mapsto \sum x_i e^{i-1}$. Let $\varphi : F \simeq \mathbf{Q}_p^d$ be its inverse.

By construction, if K is a number field contained in F and $x \in K^d$, then $\psi(x) \in K(e)$; in particular, $[\mathbf{Q}(\psi(x)) : \mathbf{Q}] \leq d[\mathbf{Q}(x) : \mathbf{Q}]$. Conversely, if $x \in K$, then the coordinates of $\varphi(x)$ in $\mathbf{Q}_p^d$ belong to the Galois closure $K(e)'$ of the compositum $K \cdot \mathbf{Q}(e)$, hence are algebraic numbers of degrees $\leq D = [K(e)' : \mathbf{Q}]$. In other words, φ and ψ induce bijections at the level of algebraic points. Since these maps are linear, there exists a positive real number $a > 0$ such that $a^{-1}H(x) \leq H(\varphi(x)) \leq aH(x)$ for every $x \in K$.

We deduce from φ a $\mathbf{Q}_p$-linear isomorphism $\varphi : F^n \to \mathbf{Q}_p^{nd}$. In particular, $Z' = \varphi(Z)$ is a subanalytic subset of $\mathbf{Q}_p^{nd}$. The morphism φ maps algebraic points of given degree to algebraic points of uniformly bounded degree, and there exists a positive real number $a > 0$ such that $a^{-1}H(x) \leq H(\varphi(x)) \leq aH(x)$ for every $x \in Z(K)$.

The definition of a family of blocks that we have adopted here is slightly stronger than the one used in Theorem 4.2.3 of [Cluckers et al. 2015]. However, all proofs go over without any modification, so that there exists a family of blocks $W' \subset Z' \times \mathbf{Q}_p^s$ such that for any $T > 1$, there exists a subset $S_T \subset \mathbf{Q}_p^s$ of cardinality $< cT^\varepsilon$ such that $Z'(D; T) \subset \bigcup_{\sigma \in S_T} W'_\sigma$. Let $\psi : F^n \times \mathbf{Q}_p^s \to \mathbf{Q}_p^{nd} \times \mathbf{Q}_p^s$ be the map $(x, y) \mapsto (\varphi(x), y)$ and let $W = \psi^{-1}(W') \subset F^n \times \mathbf{Q}_p^s$. By definition, W is a family of blocks in Z. Moreover, for any $T > 1$, one has

$$Z(F; T) \subset \psi^{-1}(Z'(D; aT)) \subset \bigcup_{\sigma \in S_{aT}} \varphi^{-1}(W'_\sigma) = \bigcup_{\sigma \in S_{aT}} W_\sigma.$$

Since $\mathrm{Card}(S_{aT}) \leq ca^\varepsilon T^\varepsilon$, the family of blocks W satisfies the requirements of the theorem. □

5. Zariski closures and analytic functions

5.1. Let F be a complete nonarchimedean valued field. Let V be an F-scheme of finite type. One says that a subset K of V^{an} is *sparse* if there exist a set T and a subset Z of $V^{\mathrm{an}} \times T$ such that for every $t \in T$, $Z_t = \{x \in V^{\mathrm{an}} : (x, t) \in Z\}$ is a Zariski-closed subset of V^{an} with empty interior, and $K = \bigcup_{t \in T} Z_t$.

Lemma 5.2. *A sparse set has empty interior.*

Proof. Let us say that a point $x \in V^{\mathrm{an}}$ is maximally Abhyankar if the rational rank of the value group of $\mathscr{H}(x)$ is equal to $\dim_x(V^{\mathrm{an}})$. If V is irreducible, then maximally Abhyankar points are dense in V^{an}; moreover, each of them is Zariski dense. Let K be a sparse set in V^{an}; write $K = \bigcup_t Z_t$ as above. Let us argue by contradiction and let U be a nonempty subset of V^{an} contained in K. By what precedes, there exists a maximally Abhyankar point $x \in U$. Let $t \in T$ be such that $x \in Z_t$. Then Z_t contains the Zariski closure of x in V^{an}, so that Z_t contains an irreducible component of V^{an}, contradicting the definition of a sparse set. □

Lemma 5.3. *Let F' be an algebraically closed complete extension of F and $q : V^{\mathrm{an}}_{F'} \to V^{\mathrm{an}}$ the base change morphism. Let K be a closed sparse subset of V^{an} and let $K' = q^{-1}(K)$. Then K' is sparse.*

Proof. Indeed, if $K = \bigcup_{t \in T} Z_t^{\mathrm{an}}$ is a description of the sparse set K, then the equality $K' = \bigcup_{t \in T} (Z_t)^{\mathrm{an}}_{F'}$ shows that K' is sparse as well. □

Lemma 5.4. *Let us assume that K is sparse, and let $C \subset V$ be a geometrically irreducible curve such that $C^{\mathrm{an}} \not\subset K$. Then $C^{\mathrm{an}} - K$ is connected.*

Proof. Using Lemma 5.3, we reduce to the case where F is algebraically closed; moreover, we may assume that C is reduced. Let $K = \bigcup_{t \in T} Z_t^{\mathrm{an}}$ be a description of K as above. Up to adding the singular locus of C to K, we may assume that C is smooth. By assumption, for every $t \in T$, $C \not\subset Z_t^{\mathrm{an}}$; consequently, $Z_t^{\mathrm{an}} \cap C^{\mathrm{an}}$ consists of rigid points of C^{an}, and hence $K \cap C^{\mathrm{an}}$ consists of rigid points of C^{an}. In the topological description of smooth geometrically irreducible analytic curves as real graphs [Berkovich 1990, Chapter 4], their rigid points are endpoints, so $C^{\mathrm{an}} - (K \cap C^{\mathrm{an}})$ is connected as well. □

Proposition 5.5. *Let F be a complete nonarchimedean valued field. Let V be an F-scheme of finite type which is geometrically connected* (*resp. geometrically irreducible*) *and let K be a closed sparse subset of V^{an}. Then $V^{\mathrm{an}} - K$ is a geometrically connected* (*resp. geometrically irreducible*) *analytic space.*

The particular case $K = \varnothing$ implies the "GAGA"-type consequence that if V is geometrically connected (or geometrically irreducible), then so is V^{an}.

Proof. Using Lemma 5.3, we reduce to the case where F is algebraically closed. By assumption, V is connected. Let us prove that $V^{\mathrm{an}} - K$ is connected. Let $x, y \in V^{\mathrm{an}} - K$. Let F' an algebraically closed complete valued field containing both $\mathscr{H}(x)$ and $\mathscr{H}(y)$, and view x, y as elements of $V(F')$. Let $q : V^{\mathrm{an}}_{F'} \to V^{\mathrm{an}}$ be the base change morphism and let $K' = q^{-1}(K)$; by Lemma 5.3, this is a sparse subset of $V^{\mathrm{an}}_{F'}$. By [Mumford 1970, p. 56], there exists an irreducible curve $C \subset V_{F'}$ which passes through x and y. Then C^{an} is connected. One has $C \not\subset K'$, by definition of K'; it follows from Lemma 5.4 that $C^{\mathrm{an}} - (K' \cap C^{\mathrm{an}})$ is connected. Consequently, x and y belong to the same component of $V^{\mathrm{an}}_{F'} - K'$, and hence their images in $V^{\mathrm{an}} - K$ belong to the same connected component. This proves that $V^{\mathrm{an}} - K$ is connected.

Let us now assume that V is geometrically irreducible. The normalization morphism $p : W \to V$ is finite, and W is geometrically connected. Since $p^{-1}(K)$ is a sparse subset of W^{an}, it follows from the first part of the lemma that $W^{\mathrm{an}} - p^{-1}(K)$ is geometrically connected. Since W^{an} is the normalization of V^{an} [Ducros 2016, Lemma 2.7.15], then $W^{\mathrm{an}} - p^{-1}(K) = p^{-1}(V^{\mathrm{an}} - K)$ is the normalization of $V^{\mathrm{an}} - K$. By Theorem 5.17 of [Ducros 2009], this implies that $V^{\mathrm{an}} - K$ is geometrically irreducible. □

Corollary 5.6. *Let F be a complete valued field, let V be an F-scheme of finite type and let K be a closed sparse subset of V^{an}. The set of irreducible components of $V^{\mathrm{an}} - K$ is finite. If V is equidimensional, then each of them has dimension* $\dim(V)$.

Proof. We may assume that V is irreducible. Let $\Omega = V^{\mathrm{an}} - K$. Let E be the completion of an algebraic closure of F. By Proposition 5.5, $\Omega_E \cap Z^{\mathrm{an}}$ is irreducible for every irreducible component Z of V_E, and the family of these intersections is the family of irreducible components of Ω_E. The finiteness statement then follows from [Ducros 2009, Lemme 4.25], while the one about dimension follows from [Ducros 2009, Proposition 4.22]. □

Corollary 5.7. *Let F be a complete valued field, let V be an irreducible F-scheme of finite type and let K be a closed sparse subset of V^{an}. Let W be an irreducible component of $V^{\mathrm{an}} - K$. If W is geometrically irreducible, then V is geometrically irreducible as well, one has $W = V^{\mathrm{an}} - K$ and W is topologically dense in V^{an}.*

Proof. Let E be a complete algebraically closed extension of F, and let $V_1, \dots, V_n$ be the irreducible components of V_E. Let L be the preimage of K in V_E; it is a closed sparse subset of V_E^{an} (Lemma 5.3). Consequently, $L_j = V_j^{\mathrm{an}} \cap L$ is a closed sparse subset of V_j^{an}, for every j. By Proposition 5.5, $W_j = V_j^{\mathrm{an}} - L_j$ is geometrically irreducible. The automorphism group $\mathrm{Aut}(E/F)$ acts transitively on the set $\{V_1, \dots, V_n\}$ of irreducible components of V_E, hence on the set $\{W_1, \dots, W_n\}$ of irreducible components of $V_E^{\mathrm{an}} - L$. Since V_E is geometrically irreducible, there exists an index j such that $W_E = W_j$; then $\mathrm{Aut}(E/F)$ fixes W_j, so that $n = 1$ and $j = 1$. This proves that V is geometrically irreducible. By Proposition 5.5, one has $W = V^{\mathrm{an}} - K$. By Lemma 5.2, W is topologically dense in V^{an}. □

Proposition 5.8. *Let F be a finite extension of $\mathbf{Q}_p$. Let A be an affine scheme of finite type over F and let $\Omega \subset A^{\mathrm{an}}$ be the complement of a closed sparse subset. Let X be a closed analytic subspace of Ω. Let V be a $\mathbf{Q}_p$-semialgebraic subset of $A(F)$, contained in $X(F)$, and let W be its Zariski closure in A. Then $W^{\mathrm{an}} \cap \Omega \subset X$.*

Proof. This proof is inspired by that of [Pila and Tsimerman 2013, Lemma 4.1].

We argue by noetherian induction on W, assuming that if W' is the Zariski closure of a $\mathbf{Q}_p$-semialgebraic subset V' of $A(F)$ contained in $X(F)$, and if $W' \subsetneq W$, then $(W')^{\mathrm{an}} \cap \Omega \subset X$.

First assume that W is not irreducible. Then any irreducible component W' of W is the Zariski closure in A of $V \cap W'(F)$, a $\mathbf{Q}_p$-semialgebraic subset of $A(F)$; by induction, $(W')^{\mathrm{an}} \cap \Omega \subset X$, so that $W^{\mathrm{an}} \cap \Omega \subset X$.

We may thus assume that W is irreducible; since its subset $W(F)$ of F-rational points contains V, it is Zariski-dense in W, so that W is geometrically irreducible.

Let $K = A^{\mathrm{an}} - \Omega$. By assumption, K is closed and sparse. Let $K = \bigcup S_t^{\mathrm{an}}$ be a presentation of K, where for every t, S_t is a Zariski-closed subset with empty interior of A. Since W is irreducible and not contained in S_t, $W \cap S_t$ is a strict Zariski-closed subset of W. Consequently, $W^{\mathrm{an}} \cap K$ is a sparse subset of W^{an}. By Proposition 5.5, $W^{\mathrm{an}} \cap \Omega$ is thus a geometrically irreducible analytic space.

Let R be the Weil restriction functor from F to $\mathbf{Q}_p$. By definition, $A(F)$ is identified with $\mathrm{R}(A)(\mathbf{Q}_p)$ and we write $\mathrm{R}(V)$ for the image of V inside $\mathrm{R}(A)(\mathbf{Q}_p)$. Let then Z be the Zariski closure of $\mathrm{R}(V)$ inside $\mathrm{R}(A)$.

Let Z' be an irreducible component of Z. Then $Z' \cap \mathrm{R}(V)$ is a semialgebraic subset of $\mathrm{R}(A)$, of the form $\mathrm{R}(V')$, for a unique $\mathbf{Q}_p$-semialgebraic subset V' of V. When Z' varies, the corresponding subsets V' cover V; we may thus choose Z' such that V' is Zariski dense in W. Replacing V by V', we may assume that Z is irreducible; then it is geometrically irreducible, because its set of $\mathbf{Q}_p$-points is Zariski dense.

Since V is $\mathbf{Q}_p$-semialgebraic, the subset $\mathrm{R}(V)$ of $\mathrm{R}(A)(\mathbf{Q}_p)$ is semialgebraic; hence, the dimension of Z coincides with the dimension of V as a $\mathbf{Q}_p$-semialgebraic subset of $A(F)$. Consequently, $\dim_{\mathrm{Zar}}(Z) = \dim(Z(\mathbf{Q}_p)) = \dim(\mathrm{R}(V))$.

Since W is a Zariski closed subset of A containing V, the subscheme $\mathrm{R}(W)$ is Zariski closed in $\mathrm{R}(A)$ and contains $\mathrm{R}(V)$, so that $Z \subset \mathrm{R}(W)$. By Weil restriction, the inclusion $Z \to \mathrm{R}(W)$ corresponds to a morphism $g : Z_F \to W$. Let $x \in A(F)$ and let $\tilde{x} \in \mathrm{R}(A)(\mathbf{Q}_p)$ be the corresponding point; if $x \in V$, then $\tilde{x} \in \mathrm{R}(V) \subset Z(\mathbf{Q}_p)$, and hence $\tilde{x} \in Z_F(F)$. By the definition of the Weil restriction functor, one has $g(\tilde{x}) = x$. In particular, the image of $Z_F(F)$ under g contains V. Hence, g is dominant, by definition of W.

The morphism g induces an analytic morphism $g^{\mathrm{an}} : Z_F^{\mathrm{an}} \to W^{\mathrm{an}} \subset A^{\mathrm{an}}$. The inverse image of $W^{\mathrm{an}} \cap \Omega$ is the complement of a closed sparse subset of Z_F^{an}; since Z_F^{an} is geometrically irreducible, Corollary 5.6 implies that $(g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap \Omega)$ is geometrically irreducible, of dimension $\dim(Z_F^{\mathrm{an}})$. Let $Y = (g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap X)$; it is a Zariski closed analytic subset of $(g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap \Omega)$.

Let us admit for a moment that $\dim(Y) = \dim(Z_F)$ and let us conclude that $W^{\mathrm{an}} \cap \Omega \subset X$. Since $\dim(Z_F^{\mathrm{an}}) = \dim(Z_F) = \dim((g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap \Omega))$, we see that

$$Y = (g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap X) = (g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap \Omega).$$

The morphism $g : Z_F \to W$ being dominant, its image contains a nonempty open subset W' of W. Since W is geometrically irreducible, $(W')^{\mathrm{an}}$ is dense in W^{an};

in particular, the image of g^{an} meets any nonempty open subset of W^{an}. Since $(g^{\mathrm{an}})^{-1}(W^{\mathrm{an}} \cap (\Omega - X))$ is empty, by the preceding equality, this implies that $W^{\mathrm{an}} \cap (\Omega - X)$ is empty; hence, $W^{\mathrm{an}} \cap \Omega = W^{\mathrm{an}} \cap X$.

It remains to prove the equality $\dim(Y) = \dim(Z_F)$.

Let us consider a semialgebraic cell decomposition of $\mathrm{R}(A)(\mathbf{Q}_p)$ which is adapted to $\mathrm{R}(V)$, $Z(\mathbf{Q}_p)$, $Z_{\mathrm{sing}}(\mathbf{Q}_p)$, and to their singular loci: a finite partition of $\mathrm{R}(A)(\mathbf{Q}_p)$ into "open cells" such that these $\mathbf{Q}_p$-semialgebraic subsets are unions of cells; see [Denef 1986] and also [Cluckers and Loeser 2007].

Let $\widetilde{C}$ be a cell of dimension $\dim(\mathrm{R}(V))$ which is contained in $\mathrm{R}(V)$. Since

$$\dim(Z_{\mathrm{sing}}(\mathbf{Q}_p)) \leq \dim(Z_{\mathrm{sing}}) < \dim(Z) = \dim(\mathrm{R}(V)),$$

the cell $\widetilde{C}$ is disjoint from $Z_{\mathrm{sing}}(\mathbf{Q}_p)$. By definition of a cellular decomposition, $\widetilde{C}$ is open in $\mathrm{R}(V)$ and in $(Z - Z_{\mathrm{sing}})(\mathbf{Q}_p)$.

Let C be the subset of V corresponding to $\widetilde{C}$. Since the identification of C with $\widetilde{C}$ provided by the Weil restriction functor is a homeomorphism which respects the singular loci, C is an open subset of V.

Let x be a point of C and let $\tilde{x}$ be the corresponding point of $\widetilde{C}$. By what precedes, $\mathrm{R}(V)$, $Z(\mathbf{Q}_p)$ and Z are smooth at $\tilde{x}$, so that $\mathrm{T}_{\tilde{x}}(\mathrm{R}(V)) = \mathrm{T}_{\tilde{x}}(Z(\mathbf{Q}_p)) = \mathrm{T}_{\tilde{x}}(Z)$. In particular, these three $\mathbf{Q}_p$-vector spaces have the same dimension, equal to $\dim(T_x(V)) = \dim(V)$.

Since $g(\tilde{x}) = x \in X$, one has $\tilde{x} \in Y$; more generally, $\widetilde{C} \subset Y$. The tangent space $\mathrm{T}_{\tilde{x}}(Y)$ of Y at $\tilde{x}$ is an F-vector subspace of $\mathrm{T}_{\tilde{x}}(Z_F) = (\mathrm{T}_{\tilde{x}}(Z))_F$ which contains $\mathrm{T}_{\tilde{x}}(\widetilde{C}) = \mathrm{T}_{\tilde{x}}(Z)$. Consequently, $\mathrm{T}_{\tilde{x}}(Y) = \mathrm{T}_{\tilde{x}}(Z_F)$. This implies that the analytic space Y has dimension $\dim(Z_F)$, and concludes the proof. □

6. Complements on *p*-adic Schottky groups and uniformization

Let F be a finite extension of $\mathbf{Q}_p$. Unless specified otherwise, analytic spaces are F-analytic spaces.

6.1. Let $a \in F$ and $r \in \mathbf{R}_{>0}$; as usual, we let $B(a,r)$ and $E(a,r)$ be the subsets of $(\mathbf{A}^1)^{\mathrm{an}}$ of points x such that $|T(x) - a| < r$ and $|T(x) - a| \leq r$, respectively. The subspace $B(a,r)$ is called a *bounded open disk*; we say that $E(a,r)$ is the corresponding *bounded closed disk*. If B is a bounded open disk, we write B^+ for the corresponding bounded closed disk. We say that such a disk is strict if its *radius* r belongs to $|F^\times|^{\mathbf{Q}}$.

To these disks, we also add the unbounded open disks $\mathbf{P}_1^{\mathrm{an}} - E(a,r)$ and the unbounded closed disks $\mathbf{P}_1^{\mathrm{an}} - B(a,r)$. An unbounded disk is said to be strict if its complementary disk is strict.

The image by an homography $\gamma \in \mathrm{PGL}(2, F)$ of an open (resp. closed, strict) disk is again an open (resp. closed, strict) disk.

6.2. We endow $\mathbf{P}_1(\mathbf{C}_p)$ with the distance given by

$$\delta(x, y) = \frac{|x-y|}{\max(1, |x|)\max(1, |y|)}$$

for $x, y \in \mathbf{C}_p$ — it is invariant under the action of $\mathrm{PGL}(2, \mathscr{O}_{\mathbf{C}_p})$. Moreover, an elementary calculation shows that every element $g \in \mathrm{PGL}(2, \mathbf{C}_p)$ is Lipschitz for this distance; see also Theorem 1.1.1 of [Rumely 1989].

6.3. Let Γ be a Schottky group in $\mathrm{PGL}(2, F)$, $\mathscr{L}_\Gamma \subset \mathbf{P}_1(F)$ its limit set and $\Omega_\Gamma = \mathbf{P}_1^{\mathrm{an}} - \mathscr{L}_\Gamma$. For any rigid point $x \in \Omega_\Gamma$, let $\delta_\Gamma(x)$ be the δ-distance of x to $\mathscr{L}_\Gamma$.

For every $\gamma \in \mathrm{PGL}(2, F)$, there exists a real number $c \geq 1$ such that $c^{-1}\delta_\Gamma(z) \leq \delta_\Gamma(\gamma \cdot z) \leq c\delta_\Gamma(z)$ for every rigid point $z \in \Omega_\Gamma$.

Lemma 6.4. *Let $\mathfrak{G}$ be a compact subset of Ω_Γ. There exists a strictly positive real number c such that $\delta_\Gamma(x) \geq c$ for every rigid point $x \in \mathfrak{G}$.*

Proof. Arguing by contradiction, we assume that there exists a sequence (x_n) of rigid points of $\mathfrak{G}$ such that $\delta_\Gamma(x_n) \to 0$. For every n, let $\xi_n \in \mathscr{L}_\Gamma$ such that $\delta_\Gamma(x_n) = \delta(x_n, \xi_n)$; it exists since $\mathscr{L}_\Gamma$ is compact. Extracting a subsequence if necessary, we assume that the sequence (ξ_n) converges to a point ξ of $\mathscr{L}_\Gamma$. Then $\delta(x_n, \xi) \to 0$. This implies that the sequence (x_n) converges to ξ in the Berkovich space $\mathbf{P}_1^{\mathrm{an}}$. Since $\mathfrak{G}$ is compact, one has $\xi \in \mathfrak{G}$, a contradiction. □

6.5. Let Γ be a Schottky subgroup of $\mathrm{PGL}(2, F)$. Let us assume that the point at infinity ∞ does not belong to its limit set $\mathscr{L}_\Gamma$. Then, by [Gerritzen and van der Put 1980, I, (4.3)], the group Γ admits a basis $(\gamma_1, \dots, \gamma_g)$ and a *good fundamental domain* $\mathfrak{F}_\Gamma$ with respect to this basis, in the following sense:

(1) There exists a finite family $(B_1, \dots, B_g, C_1, \dots, C_g)$ of strict bounded open disks in $\mathbf{P}_1^{\mathrm{an}}$ such that $\mathfrak{F}_\Gamma = \mathbf{P}_1^{\mathrm{an}} - \left(\bigcup B_i \cup \bigcup C_i\right)$.

(2) The corresponding bounded closed disks $B_1^+, \dots, B_g^+, C_1^+, \dots, C_g^+$ are pairwise disjoint.

Let then $\mathfrak{F}_\Gamma^\circ = \mathbf{P}_1^{\mathrm{an}} - \left(\bigcup B_i^+ \cup \bigcup C_i^+\right)$.

(3) The elements $\gamma_1, \dots, \gamma_g$ satisfy $\gamma_i(\mathbf{P}_1^{\mathrm{an}} - B_i) = C_i^+$ and $\gamma_i(\mathbf{P}_1^{\mathrm{an}} - B_i^+) = C_i$ for every $i \in \{1, \dots, g\}$.

With this notation, let $W = \mathbf{P}_1^{\mathrm{an}} - \bigcup B_i$; this is an affinoid domain of $\mathbf{P}_1^{\mathrm{an}}$ containing $\mathfrak{F}$, stable under each γ_i. Indeed, one has $W \subset \mathbf{P}_1^{\mathrm{an}} - B_i$. Hence, $\gamma_i W \subset \gamma_i(\mathbf{P}_1^{\mathrm{an}} - B_i) = C_i^+$, and hence the claim since C_j^+ is disjoint from each B_i.

Moreover, the following properties are satisfied:

(4) One has $\bigcup_{\gamma \in \Gamma} \gamma \cdot \mathfrak{F}_\Gamma = \mathbf{P}_1 - \mathscr{L}_\Gamma$.

(5) For $\gamma \in \Gamma$, one has $\mathfrak{F}_\Gamma \cap \gamma \cdot \mathfrak{F}_\Gamma \neq \varnothing$ if and only if $\gamma \in \{\mathrm{id}, \gamma_1^{\pm 1}, \dots, \gamma_g^{\pm 1}\}$.

(6) For every $\gamma \in \Gamma - \{\mathrm{id}\}$, one has $\mathfrak{F}_\Gamma^\circ \cap \gamma \cdot \mathfrak{F}_\Gamma = \varnothing$.

In this context, we identify an element γ of Γ with a reduced word in the letters $\{\gamma_1^\pm, \dots, \gamma_g^\pm\}$ and denote its length by $\ell_\Gamma(\gamma)$.

For every $\gamma \in \Gamma - \{\mathrm{id}\}$, [Gerritzen and van der Put 1980, I, §4, p. 29] define a bounded open disk $B(\gamma)$, equal either to $\gamma \cdot (\mathbf{P}_1^{\mathrm{an}} - B_i^+)$ or to $\gamma \cdot (\mathbf{P}_1^{\mathrm{an}} - C_i^+)$, according to whether the last letter of the reduced word representing γ is γ_i or γ_i^{-1}; in any case, one has $\gamma \cdot \infty \in B(\gamma)$. Moreover, they prove:

(7) $B(\gamma') \subset B(\gamma)$ if and only if γ is an initial subword of γ'.

(8) For every integer n, one has

$$\mathbf{P}_1^{\mathrm{an}} - \bigcup_{\ell_\Gamma(\gamma) < n} \gamma \cdot \mathfrak{F} = \bigcup_{\ell_\Gamma(\gamma) = n} B(\gamma).$$

(9) There exists a real number $c > 1$ such that for every γ, the radius of the disk $B(\gamma)$ is $\ll c^{-\ell_\Gamma(\gamma)}$.

(10) The intersection of every decreasing sequence of open disks $(B(\gamma_n))$, where $\ell_\Gamma(\gamma_n) = n$, is reduced to a limit point of Γ, and every limit point can be obtained in this way.

Proposition 6.6. *Let Γ be a Schottky group in* $\mathrm{PGL}(2, F)$ *and let $\mathfrak{G}$ be a compact analytic domain of Ω_Γ. There exist positive real numbers a, b such that for every $\gamma \in \Gamma$ and every rigid point $x \in \gamma \cdot \mathfrak{G}$, one has*

$$\ell_\Gamma(\gamma) \le a - b \log(\delta_\Gamma(x)).$$

Proof. To prove this proposition, we may extend the scalars to a finite extension of F and henceforth assume that the limit set $\mathscr{L}_\Gamma$ is not equal to $\mathbf{P}_1(F)$. Placing a point of $\mathbf{P}_1(F) - \mathscr{L}_\Gamma$ at infinity, Section 6.5 furnishes a basis $(\gamma_1, \dots, \gamma_g)$ and a good fundamental domain with respect to this basis of the form $\mathfrak{F} = \mathbf{P}_1^{\mathrm{an}} - \left(\bigcup_{i=1}^g B_i \cup \bigcup_{i=1}^g C_i\right)$. Let b and $c > 1$ be positive real numbers such that the diameter of $B(\gamma)$ is bounded by $bc^{-\ell_\Gamma(\gamma)}$, for every $\gamma \in \Gamma - \{\mathrm{id}\}$.

Let $x \in \Omega_\Gamma$ and let $\gamma \in \Gamma$ be such that $x \in \gamma \cdot \mathfrak{F}$. Let $\xi \in \mathscr{L}_\Gamma(x)$ be such that $\delta_\Gamma(x) = \delta(x, \xi)$. As the disk $B(\gamma)$ contains both x and ξ, one has $\delta_\Gamma(x) \le bc^{-\ell_\Gamma(\gamma)}$, that is,

$$\ell_\Gamma(\gamma) \le \frac{1}{\log(c)}(-\log(\delta_\Gamma(x)) + \log(b)),$$

since $\log(c) > 0$. This proves the proposition in the particular case where $\mathfrak{G} = \mathfrak{F}$.

Let us now prove the general case. Let a be a real number such that $\delta_\gamma(x) \ge a > 0$ for every rigid point of $\mathfrak{G}$ (Lemma 6.4). The preceding inequality shows that there exists a finite subset S of Γ such that $\mathfrak{G}$ meets $\gamma \cdot \mathfrak{F}$ if and only if $\gamma \in S$. It then follows from property (8) that $\mathfrak{G}$ is contained in the finite union $\bigcup_{s \in S} s \cdot \mathfrak{F}$. To conclude the proof, we observe that if $x \in \gamma \cdot \mathfrak{G}$, then there exists $s \in S$ such that

$x \in \gamma s \cdot \mathfrak{F}$. The proposition then follows from the particular case already treated and from the inequality $\ell_\Gamma(\gamma) \leq \ell_\Gamma(\gamma s) + \ell_\Gamma(s)$. □

Corollary 6.7. *Let $\mathfrak{G}$ and $\mathfrak{G}'$ be compact analytic domains of Ω_Γ. The set of $\gamma \in \Gamma$ such that $\gamma \cdot \mathfrak{G} \cap \mathfrak{G}' \neq \varnothing$ is finite.*

Proof. Let S be this set. For $\gamma \in S$, the intersection $\gamma \cdot \mathfrak{G} \cap \mathfrak{G}'$ is a nonempty affinoid domain of $\mathbf{P}_1^{\mathrm{an}}$; hence, it contains a rigid point x_γ. With a and b as in the statement of Proposition 6.6, one has $\ell_\Gamma(\gamma) \leq a - b \log(\delta_\Gamma(x_\gamma))$. Since $x_\gamma \in \mathfrak{G}'$, $\delta_\Gamma(x_\gamma)$ is bounded from below by Lemma 6.4. This shows that $\ell_\Gamma(\gamma)$ is bounded above when γ runs over S. □

Proposition 6.8. *Let Γ be a Schottky group in* $\mathrm{PGL}(2, F)$ *and let g be its rank. Let $\xi \in \mathscr{L}_\Gamma$ and let U be an open neighborhood of ξ in $\mathbf{P}_1^{\mathrm{an}}$.*

There exist an open neighborhood U' of ξ, contained in U, a basis $\gamma_1, \dots, \gamma_g$ of Γ, an affinoid domain $\mathfrak{F} \subset \Omega_\Gamma$ such that the following properties hold:

(1) *One has $\mathfrak{F} \subset U'$.*

(2) *For every i, one has $\gamma_i(U') \subset U'$.*

(3) *One has $\bigcup_{\gamma \in \Gamma} \gamma \mathfrak{F} = \Omega_\Gamma$.*

Such an affinoid domain will be called a *fundamental set.*

Proof. We first treat the case where $\mathscr{L}_\Gamma \neq \mathbf{P}_1(F)$. Placing a point of $\mathbf{P}_1(F) - \mathscr{L}_\Gamma$ at infinity, Section 6.5 furnishes a basis $(\gamma_1, \dots, \gamma_g)$ and a good fundamental domain $\mathfrak{F}$ with respect to this basis of the form $\mathfrak{F} = \mathbf{P}_1^{\mathrm{an}} - \left(\bigcup_{i=1}^g B_i \cup \bigcup_{i=1}^g C_i\right)$.

By (10), for every integer $n \geq 1$, there is an element $\gamma \in \Gamma$ of length n such that $\xi \in B(\gamma)$; if n is large enough, one has $B(\gamma)^+ \subset U$, because the diameter of $B(\gamma)^+$ tends to 0 when $n = \ell_\Gamma(\gamma)$ tends to ∞. Since $\gamma \cdot \mathfrak{F} \subset B(\gamma)^+$, this implies that $\gamma \cdot \mathfrak{F} \subset U$.

Up to changing the basis $(\gamma_1, \dots, \gamma_g)$ into $(\gamma_1^{-1}, \dots, \gamma_g^{-1})$, and exchanging B_i and C_i for every i, we may assume that the last letter of γ is γ_s, for some $s \in \{1, \dots, g\}$. Set $W = \mathbf{P}_1^{\mathrm{an}} - \bigcup_{i=1}^g B_i$; recall that W is an affinoid domain of $\mathbf{P}_1^{\mathrm{an}}$ containing $\mathfrak{F}$ and stable under $\gamma_1, \dots, \gamma_g$. By definition, one has

$$B(\gamma)^+ = \gamma \cdot (\mathbf{P}_1^{\mathrm{an}} - B_s) \supset \gamma \cdot W,$$

since $W \subset \mathbf{P}_1^{\mathrm{an}} - B_s$.

Let us now set $\mathfrak{F}' = \gamma \cdot \mathfrak{F}$, $W' = \gamma \cdot W$ and $\gamma_i' = \gamma \gamma_i \gamma^{-1}$ for $i \in \{1, \dots, g\}$. By construction, $\mathfrak{F}'$ and W' are affinoid domains of $\mathbf{P}_1^{\mathrm{an}}$ such that $\mathfrak{F}' \subset W' \subset B(\gamma)^+ \subset U$, the translates of $\mathfrak{F}'$ under Γ cover Ω_Γ, and W' is stable under the basis $(\gamma_1', \dots, \gamma_g')$ of Γ.

This almost proves (1–3), except that W' is affinoid and not open. To conclude the construction, one sets U' to be the interior of W' and redoes the construction starting from U' instead of U. The second paragraph of the proof shows that there exists $\gamma' \in \Gamma$ such that $\gamma' \cdot \mathfrak{F}'$ is contained in U'. The affinoid $\gamma' \cdot \mathfrak{F}'$, the open subset U' and the basis $(\gamma_1', \dots, \gamma_g')$ satisfy the requirements of the proposition.

Let us now treat the case where $\mathscr{L}_\Gamma = \mathbf{P}_1(F)$. Let F' be a finite extension of F of degree > 1. The preceding construction can be applied starting with a point of $\mathbf{P}_1(F') \setminus \mathscr{L}_\Gamma$ and furnishes an open neighborhood V' of ξ in $(\mathbf{P}_1^{\mathrm{an}})_{F'}$, contained in $U_{F'}$, a basis $(\gamma_1, \dots, \gamma_g)$ of Γ and an affinoid domain $\mathfrak{F}'$ of $\Omega_{\Gamma, F'}$ satisfying properties (1–3). The images U' of V' and $\mathfrak{F}$ of $\mathfrak{F}'$ by the projection $(\mathbf{P}_1^{\mathrm{an}})_{F'} \to \mathbf{P}_1^{\mathrm{an}}$ satisfy the required properties. □

Lemma 6.9. *Let Γ be an arithmetic Schottky group in* $\mathrm{PGL}(2, F)$ *and let H be a height function on* $\mathrm{PGL}(2, \overline{\mathbf{Q}})$. *There exists a positive real number c such that $H(\gamma) \le c^{\ell_\Gamma(\gamma)+1}$ for every $\gamma \in \Gamma$.*

Proof. Let $(\gamma_1, \dots, \gamma_g)$ be a basis of Γ as above. Let c_1 be a positive real number such that $H(hh') \le c_1 H(h) H(h')$ for every $h, h' \in \mathrm{PGL}(2, \overline{\mathbf{Q}})$. Let $c = c_1 \sup(H(\mathrm{id}), H(\gamma_1), \dots, H(\gamma_g))$. One proves by induction on $\ell_\Gamma(\gamma)$ that

$$c_1 H(\gamma) \le \sup(c_1 H(\gamma_1^{\pm}), \dots, c_1 H(\gamma_g^{\pm}))^{\ell_\Gamma(\gamma)} c_1 H(\mathrm{id}) \le c_1 c^{\ell_\Gamma(\gamma)+1}$$

for every $\gamma \in \Gamma$, as was to be shown. □

Lemma 6.10. *Let Γ be a Schottky subgroup of* $\mathrm{PGL}(2, F)$ *and let Δ be a subset of $\mathbf{P}_1(\bar{F})$ of cardinality 2. Let K be a number field contained in F. The stabilizer of Δ inside Γ does not have many K-rational points.*

Proof. Let S be this stabilizer; we may assume that $S \neq \{\mathrm{id}\}$. Let $g \in S \setminus \{\mathrm{id}\}$. Then g is hyperbolic (see [Gerritzen and van der Put 1980, p. 7, line 2]), and hence has exactly two rational fixed points in $\mathbf{P}_1(F)$. Up to a change of projective coordinates, we may thus assume that $\Delta = \{0, \infty\}$. Then every element h of S is of the form $z \mapsto \lambda(h) z$, for some unique element $\lambda(h) \in K^\times$; moreover, unless $h = \mathrm{id}$, any such h is hyperbolic and thus is represented by a matrix having two eigenvalues with distinct absolute values, so that $|\lambda(h)| \neq 1$. Let us choose $h \in S \setminus \{\mathrm{id}\}$ such that $|\lambda(h)|$ is > 1 and minimal. By euclidean division, one has $S = \langle h \rangle$.

Then $S \cap \mathrm{PGL}(2, K)$ is generated by an element of the form h^a for some $a \in \mathbf{Z}$. Since h^a is semisimple, we have $H(h^a)^n \ll H(h^{an}) \ll H(h^a)^n$, for every $n \in \mathbf{Z}$ (see Section 4.4). This shows that $S \cap \mathrm{PGL}(2, K)$ does not have many rational points. □

In Section 8, we will need the following lemma.

Lemma 6.11. *Let r be a positive real number, $f \in \mathbf{C}_p[[z]]$ a power series which converges on the closed disk $E(0, r)$, and L_1 and L_2 closed subsets of $\mathbf{C}_p$ such that $f^{-1}(L_2) \subset L_1$. For every $x \in \mathbf{C}_p$, let $\delta(x; L_1)$ and $\delta(x; L_2)$ be the distances of x to L_1 and L_2, respectively. Then there exist real numbers $m \ge 0$, $c > 0$ and s such that $0 < s < r$ and such that $\delta(f(x); L_2) \ge c\delta(x; L_1)^m$ for every $x \in E(0, s)$.*

Proof. Write $f = \sum c_n z^n$. We may assume that there exists $a \in \mathbf{C}_p^\times$ such that $r = |a|$; composing f with homographies which map $E(0, r)$ to $E(0, 1)$ and $f(E(0, r))$ into the disk $E(0, 1)$, we assume that $r = 1$ and that $|c_n| \leq 1$ for all n. (Recall from Section 6.2 that homographies are Lipschitz for the distance δ.)

Let us first treat the case where $f(0) \notin L_2$. Then there exists a real number $s > 0$ such that $E(f(0), s) \cap L_2 = \varnothing$. For every $x \in E(0, 1)$ such that $|x| < s$, one has $|f(x) - f(0)| < s$; hence, $\delta(f(x); L_2) > s$. It suffices to set $m = 0$ and $c = s$.

We now assume that $f(0) \in L_2$, and hence $0 \in L_1$. Let $m = \operatorname{ord}_0(f - f(0))$. Since $f'(z) = \sum_{n \geq m} n c_n z^{n-1}$, there exists a real number s such that $0 < s \leq 1$ and such that $|f'(z)| = |mc_m||z|^{m-1}$ provided $|z| \leq s$. Moreover, $|f^{(n)}(z)/n!| \leq 1$ for every $n \geq 0$ and any $z \in E(0, 1)$. Considering the Taylor expansion

$$f(y) = \sum_{n \geq 0} \frac{1}{n!} f^{(n)}(x)(y - x)^n,$$

we then see that there exists a real number s' such that

$$f(E(x, u)) = E(f(x), |f'(x)|u)$$

for every real number u such that $0 < u \leq s'$ and $x \in E(0, 1)$ such that $0 < |x| \leq s$. If $u < \delta(x; L_1)$, then $E(x, u) \cap L_1 = \varnothing$; hence, $E(f(x), |f'(x)|u) \cap L_2 = \varnothing$. Consequently, $\delta(f(x); L_2) \geq |f'(x)|\, \delta(x; L_1)$. Since $0 \in L_1$, one has $|x| \geq \delta(x; L_1)$. Consequently,

$$\delta(f(x); L_2) \geq |mc_m||x|^{m-1} \delta(x; L_1) \geq |mc_m|\, \delta(x; L_1)^m.$$

This concludes the proof. □

7. Automorphisms of curves

The following result is already present in [Pila 2013]. For the clarity of exposition, we isolate it as a lemma.

Lemma 7.1. *Let k be an algebraically closed field of characteristic zero, B a smooth connected projective k-curve and $f : B \to \mathbf{P}_1$ a nonconstant morphism. Let $R_f \subset B$ be the ramification locus of f (the set of points of B at which f is not étale) and let $\Delta_f = f(R_f)$ be its discriminant locus.*

Assume that there exist automorphisms $g \in \operatorname{Aut}(\mathbf{P}_1)$ and $h \in \operatorname{Aut}(B)$ such that $f \circ h = g \circ f$, and that g has infinite order. Then B is isomorphic to $\mathbf{P}_1$, and one of the following cases holds:

- *The morphism f is an isomorphism (and $\Delta_f = \varnothing$).*
- *One has $\operatorname{Card}(R_f) = 2$ and $g(\Delta_f) = \Delta_f$.*

Proof. By construction, f induces a finite étale covering of $\mathbf{P}_1 - \Delta_f$.

Let $b \in R_f$. One has $df(b) = 0$; hence, $d(f \circ h)(b) = d(g \circ f)(b) = 0$. Since h is an automorphism of B, this implies that $df(h(b)) = 0$; hence, $h(b) \in R_f$. We thus have $h(R_f) \subset R_f$; hence, $h(R_f) = R_f$, because h is an isomorphism. Consequently, $g(\Delta_f) = \Delta_f$, so that some power of g fixes Δ_f pointwise. Since the identity is the only homography that fixes 3 points and g has infinite order, this implies that $\mathrm{Card}(\Delta_f) \leq 2$.

If $\mathrm{Card}(\Delta_f) \leq 1$, then $\mathbf{P}_1 - \Delta_f$ is simply connected. Hence, f is an isomorphism (and $\Delta_f = \varnothing$).

Otherwise, one has $\mathrm{Card}(\Delta_f) = 2$. Let $n = \deg(f)$. Up to a change of projective coordinates in $\mathbf{P}_1$, we may assume that $\Delta_f = \{0, \infty\}$. Then g is a homothety, because it leaves Δ_f invariant and has infinite order (otherwise, it would be of the form $g(z) = a/z$). Since all finite étale coverings of $\mathbf{P}_1 - \Delta_f$ are of Kummer type (equivalently, $\pi_1(\mathbf{P}_1 - \Delta_f) = \mathbf{Z}$), one has $B \simeq \mathbf{P}_1$ and the morphism f is conjugate to the morphism $z \mapsto z^n$ from $\mathbf{P}_1$ to itself.

We then remark that h is a homography of infinite order. Indeed, if $h^e = \mathrm{id}_B$, then $f = g^e \circ f$. Hence, $g^e = \mathrm{id}$ since f is surjective. Hence $e = 0$, since g has infinite order. As above, the formula $h(R_f) = R_f$ then implies that $\mathrm{Card}(R_f) \leq 2$. On the other hand, $\mathrm{Card}(R_f) \geq \mathrm{Card}(\Delta_f) = 2$. Hence, $\mathrm{Card}(R_f) = 2$. □

Proposition 7.2. *Let k be a field of characteristic zero. Let B be an integral k-curve in $\mathbf{P}_1^n$ possessing a smooth k-rational point. Let Γ_B be the stabilizer of B in $(\mathrm{Aut}(\mathbf{P}_1))^n$ and let $\Gamma_1 \subset \mathrm{Aut}(\mathbf{P}_1)$ be its image under the first projection. Assume that Γ_1 contains an element of infinite order. Then one of the following cases holds*:

(1) *The morphism $p_1|_B$ is constant.*

(2) *The morphism $p_1|_B$ is an isomorphism and the components of its inverse are either constant or homographies.*

(3) *There is a subset of $\mathbf{P}_1(\bar{k})$ of cardinality 2 which is invariant under every element of Γ_1.*

Proof. Assume that $p_1|_B$ is not constant. Let $\nu : B' \to B$ be the normalization of B and let $p_1' = p_1 \circ \nu : B' \to \mathbf{P}_1$. Let $g = (g_1, \dots, g_n)$ be an element of Γ_B. There exists a unique automorphism h of B' that lifts g, so $p_1' \circ h = g_1 \circ p_1'$. Since the curve B has smooth rational points, the curve B' is geometrically integral. Choosing g such that g_1 has infinite order, the preceding lemma implies that $\mathrm{Card}(R_{p_1'}) \in \{0, 2\}$.

Let us first assume that $\mathrm{Card}(R_{p_1'}) = 2$. Then $\mathrm{Card}(\Delta_{p_1'}) = 2$ as well. Moreover, the relation $p_1' \circ h = g_1 \circ p_1'$ implies that $g_1(\Delta_{p_1'}) \subset \Delta_{p_1'}$, so that case (3) holds.

Let us now assume that $\mathrm{Card}(R_{p_1'}) = 0$ and fix g such that g_1 has infinite order. By the preceding lemma, p_1' is an isomorphism; this implies that $p_1|_B$ is an isomorphism as well. Let f be its inverse and let $f_1, \dots, f_n$ be its components. Assume that case (2) does not hold, that is, for some j, the rational map f_j is neither constant, nor a homography; its ramification locus R_j is nonempty. Since g_1 has infinite order, the relation $g_j \circ f_j = f_j \circ g_1$ implies that g_j has infinite order as well. By the preceding lemma, one has $\mathrm{Card}(R_j) = 2$. Let then $g' = (g'_1, \dots, g'_n)$ be any element of Γ_B. The relation $g'_j \circ f_j = f_j \circ g'_1$ implies that $g'_1(R_j) \subset R_j$, so that case (3) holds. □

8. Proof of Theorem 2.7

We will reduce the proof of Theorem 2.7 to the following variant:

Proposition 8.1. *Let F be a finite extension of $\mathbf{Q}_p$ and let $(\Gamma_i)_{1\leq i\leq n}$ be a finite family of* arithmetic *Schottky subgroups of* PGL(2, F) *of ranks ≥ 2. As above, let us set $\Omega = \prod_{i=1}^n \Omega_{\Gamma_i}$ and $X = \prod_{i=1}^n X_{\Gamma_i}$, and let $p : \Omega \to X^{\mathrm{an}}$ be the morphism deduced from the morphisms $p_{\Gamma_i} : \Omega_{\Gamma_i} \to X^{\mathrm{an}}_{\Gamma_i}$.*

Let V be an irreducible algebraic subvariety of X and let W be an irreducible algebraic subvariety of Ω, maximal among those contained in $p^{-1}(V^{\mathrm{an}})$. If W is geometrically irreducible, then it is flat.

Lemma 8.2. *Proposition 8.1 implies Theorem 2.7.*

Proof. Let Y be the Zariski closure of W in $\mathbf{P}_1^n$; by assumption, W is an irreducible component of $Y^{\mathrm{an}} \cap \Omega$. Let W_0 be an irreducible component of $W_{\mathbf{C}_p}$. By [Ducros 2009, Théorème 7.16(v)], there exists a finite extension F' of F, contained in $\mathbf{C}_p$, and an irreducible component W' of $W_{F'}$ such that $W_0 = W'_{\mathbf{C}_p}$. Then W' is geometrically irreducible, as well as its Zariski closure Y'. By Proposition 5.5, $\Omega \cap Y'$ is geometrically irreducible. The inclusion $W' \subset \Omega \cap Y'$ and the inequality $\dim(W') = \dim(W_0) = \dim(W) = \dim(Y) \geq \dim(Y')$ imply that $W' = \Omega \cap Y'$. In particular, W' is irreducible algebraic and is contained in $p^{-1}(V^{\mathrm{an}}_{F'})$. Let us show that it is maximal. Let $W'_1 \subset \Omega_{F'}$ be an irreducible algebraic subvariety contained in $p^{-1}(V^{\mathrm{an}}_{F'})$ such that $W' \subsetneq W'_1$, and let $Y'_1 \subset (\mathbf{P}_1^n)_{F'}$ be the Zariski closure of W'_1. The image Y_1 of Y'_1 in $(\mathbf{P}_1^n)_F$ is Zariski closed, because F' is a finite extension of F, and $Y'_1 \subset (Y_1)_{F'}$. Moreover, $Y \subset Y_1$. There exists a unique irreducible component W_1 of $\Omega \cap Y_1$ that contains W, and W'_1 is an irreducible component of $W_{1,F'}$. Necessarily, W_1 is contained in $p^{-1}(V^{\mathrm{an}})$, because $W'_1 \subset p^{-1}(V^{\mathrm{an}}_{F'})$; this contradicts the maximality of W.

Applying Proposition 8.1 to W', we conclude that W' is flat. Consequently, $W_0 = W'_{\mathbf{C}_p}$ is flat, as was to be shown. □

8.3. To prove Proposition 8.1, we argue by induction and assume that it holds if there are less that n factors. Let W be an irreducible algebraic subvariety of Ω, maximal among those contained in $p^{-1}(V^{\mathrm{an}})$ and geometrically irreducible. Let Y be an irreducible subvariety of $\mathbf{P}_1^n$ such that W is an irreducible component of $Y^{\mathrm{an}} \cap \Omega$. By Corollary 5.7, Y is geometrically irreducible, $W = Y^{\mathrm{an}} \cap \Omega$ and W is topologically dense in Y.

The proof that W is flat requires intermediate steps and will be concluded in Proposition 8.11.

A crucial step will consist in proving that the stabilizer of W inside Γ has many points of bounded heights (Proposition 8.10). To that aim, we define in Section 8.7 an F-subanalytic subset R of $\mathrm{PGL}(2, F)^n$. The definition, close to that of a similar set in [Pila 2011; 2015], guarantees the following important property (Lemma 8.8): if B is a small enough subset of R then, for every $g \in B$, the translate $(g \cdot Y^{\mathrm{an}}) \cap \Omega$ is contained in $p^{-1}(V^{\mathrm{an}})$, and is independent of g. At this point, the maximality of W is invoked.

The existence of such blocks is established by applying the p-adic Pila–Wilkie theorem of [Cluckers et al. 2015]. We thus prove that R has many rational points (Lemma 8.9); these points are constructed using the action of the Schottky groups in a neighborhood of a boundary point ξ, applying material recalled in Section 6. The construction of such a point ξ, performed in Lemma 8.5, is actually the starting point of the proof.

The actual statement of Proposition 8.10 furnishes elements in Γ of a precise form. Using Proposition 7.2, we will finally conclude the proof of Proposition 8.1.

8.4. By assumption, $W = Y^{\mathrm{an}} \cap \Omega$; consequently, the j-th projection $q_j : (\mathbf{P}_1)^n \to \mathbf{P}_1$ is constant on Y if and only if it is constant on W, if and only if the j-th projection from X to X_j is constant on V, and in this case, its image is an F-rational point of $\mathbf{P}_1$, because W is geometrically irreducible. Deleting these constant factors, we thus assume that there does not exist $j \in \{1, \dots, n\}$ such that the j-th projection $q_j : (\mathbf{P}_1)^n \to \mathbf{P}_1$ is constant on Y. Consequently, $q_j|_Y : Y \to \mathbf{P}_1$ is surjective for every j; in particular, Y^{an} meets $q_j^{-1}(\mathscr{L}_{\Gamma_j})$.

Let $m = \dim(Y)$; by what precedes, we have $m > 0$, and $Y^{\mathrm{an}} \not\subset \Omega$.

Lemma 8.5. *Up to reordering the coordinates, there exists a smooth rigid point $\xi \in Y^{\mathrm{an}}$ and a connected open neighborhood U of ξ in $(\mathbf{P}_1^n)^{\mathrm{an}}$ such that the following properties hold*:

(1) *The first component $q_1(\xi)$ of ξ belongs to the limit set $\mathscr{L}_{\Gamma_1}$ of Γ_1.*

(2) *Letting $J = \{1, \dots, m\}$, the projection $q_J : \mathbf{P}_1^n \to \mathbf{P}_1^J$ induces a finite étale morphism from $U \cap Y^{\mathrm{an}}$ to its image in $(\mathbf{P}_1^J)^{\mathrm{an}}$.*

(3) *For every $j \in \{1, \dots, n\}$ and every point $y \in U \cap Y^{\mathrm{an}}$ such that $q_j(y) \in \mathscr{L}_{\Gamma_j}$, one has $q_1(y) \in \mathscr{L}_{\Gamma_1}$.*

Proof. For every subset V of Y^{an}, let us define a relation $\preceq_V$ on $\{1, \dots, n\}$ as follows: $i \preceq_V j$ if and only if, for every $y \in V$ such that $q_i(y) \in \mathscr{L}_{\Gamma_i}$, one has $q_j(y) \in \mathscr{L}_{\Gamma_j}$. This is a preordering relation. If $U \subset V \subset Y^{\mathrm{an}}$ and $i \preceq_V j$, then $i \preceq_U j$.

We define a decreasing sequence $(V_0, V_1, \dots, V_n)$ of nonempty open subsets of Y^{an} and a sequence $(j_0, j_1, \dots, j_n)$ of elements of $\{1, \dots, n\}$, such that for every k, $q_{j_k}(V_k)$ meets $\mathscr{L}_{\Gamma_{j_k}}$ and $1, \dots, k \preceq_{V_k} j_k$.

We start with $V_0 = Y^{\mathrm{an}}$. We have reduced to the case where $q_j(Y^{\mathrm{an}}) = \mathbf{P}_1$ for every j. In particular, $q_j(Y^{\mathrm{an}})$ meets $\mathscr{L}_{\Gamma_j}$. We may take $j_0 = 1$.

Let $k \geq 0$ be such that $V_0, V_1, \dots, V_k$ and $j_0, j_1, \dots, j_k$ are defined. If $k+1 \preceq_{V_k} j_k$, we set $V_{k+1} = V_k$ and $j_{k+1} = j_k$. Otherwise, one has $k+1 \not\preceq_{V_k} j_k$. Hence, there exists $y \in V_k$ such that $q_{k+1}(y) \in \mathscr{L}_{\Gamma_{k+1}}$ and $q_{j_k}(y) \notin \mathscr{L}_{\Gamma_{j_k}}$. Let $V_{k+1} = V_k \cap (q_{j_k})^{-1}(\Omega_{\Gamma_{j_k}})$; this is an open neighborhood of y in V_k such that $q_{j_{k+1}}(V_{k+1})$ meets $\mathscr{L}_{\Gamma_{j_{k+1}}}$. By construction, no element z of V_{k+1} satisfies $q_{j_k}(z) \in \mathscr{L}_{\Gamma_{j_k}}$, so that $j_k \preceq_{V_{k+1}} k+1$. We then set $j_{k+1} = k+1$.

Let $V = V_n$ and $i = j_n$, and let $y \in V$ be such that $q_i(y) \in \mathscr{L}_{\Gamma_i}$. Let Z be the dense open subscheme of Y consisting of smooth points at which dq_i does not vanish. Then Z^{an} is open and dense in Y^{an}, and $V \cap Z^{\mathrm{an}}$ is open and dense in V; hence, $q_i(V \cap Z^{\mathrm{an}})$ is dense in $q_i(V)$. Since $\mathscr{L}_{\Gamma_i}$ has no isolated points, we may assume that $y \in Z^{\mathrm{an}}$. Rigid points are dense in $q_i^{-1}(q_i(y)) \cap V \cap Z^{\mathrm{an}}$; there exists a rigid point ξ in $(q_i)^{-1}(q_i(y)) \cap V \cap Z^{\mathrm{an}}$. Since $q_i(y)$ is a rigid point, the point ξ is a rigid point of $V \cap Z^{\mathrm{an}}$ (and not only of its fiber of q_i). Moreover, $q_i(\xi) = q_i(y) \in \mathscr{L}_{\Gamma_i}$.

Since dq_i does not vanish at ξ, there exists a subset J of $\{1, \dots, n\}$ containing i such that the projection q_J from V to $(\mathbf{P}_1^J)^{\mathrm{an}}$ is finite étale at ξ. One has $\mathrm{Card}(J) = \dim(V) = m$. Consequently, there exists an open neighborhood U of ξ in $(\mathbf{P}_1^n)^{\mathrm{an}}$ such that q_J induces a finite étale morphism from $U \cap Y^{\mathrm{an}}$ to its image in $(\mathbf{P}_1^J)^{\mathrm{an}}$.

Reordering the coordinates, we may assume that $i = 1$ and $J = \{1, \dots, m\}$, hence the lemma. □

8.6. Choose ξ, $J = \{1, \dots, m\}$ and U as in the previous lemma; we may even assume that U is of the form $U_1 \times \cdots \times U_n$, where, for each i, U_i is an open neighborhood of $q_i(\xi)$ in $\mathbf{P}_1^{\mathrm{an}}$.

Let F' be a finite extension of F such that $\xi \in Y(F')$. Since W is geometrically irreducible, $W_{F'}$ is an irreducible algebraic subvariety of Ω. It is also maximal. Note that the flatness of $W_{F'}$ implies the flatness of W. Replacing F by F', we thus may assume that $\xi \in Y(F)$; then q_J induces a local isomorphism at ξ.

Let $\varphi = (\varphi_1, \dots, \varphi_n) : O \to Y^{\mathrm{an}} \cap U$ be an analytic section of $q_J|_{Y^{\mathrm{an}} \cap U}$, defined on an open neighborhood O of $q_J(\xi)$; we may assume that $O = U_1 \times \cdots \times U_m$.

By condition (3) of Lemma 8.5, $q_1(\varphi_j^{-1}(\mathscr{L}_{\Gamma_j})) \subset \mathscr{L}_{\Gamma_1}$ for every $j \in \{1, \dots, n\}$.

8.7. Let G be the $\mathbf{Q}$-algebraic group $\mathrm{PGL}(2)^n$, and let G_0 be the algebraic subgroup of G defined by

$$(g_1, \dots, g_n) \in G_0 \quad \Leftrightarrow \quad g_2 = \cdots = g_m = 1. \tag{8.7.1}$$

We denote by $q_1, \dots, q_n$ the projections of G to $\mathrm{PGL}(2)$. For every compact analytic domain $\mathfrak{F}$ of Ω, we define a subset $R_{\mathfrak{F}}$ of $G_0(F)$ by

$$g \in R_{\mathfrak{F}} \quad \Leftrightarrow \quad \dim(g \cdot Y^{\mathrm{an}} \cap \mathfrak{F} \cap p^{-1}(V^{\mathrm{an}})) = m. \tag{8.7.2}$$

Lemma 8.8. *Let $\mathfrak{F}$ be an affinoid domain of Ω.*

(1) *The set $R_{\mathfrak{F}}$ is an F-subanalytic subset of $G_0(F)$.*

(2) *For every $g \in R_{\mathfrak{F}}$, one has $(g \cdot Y^{\mathrm{an}}) \cap \Omega \subset p^{-1}(V^{\mathrm{an}})$.*

(3) *Let $M \subset R_{\mathfrak{F}}$ be a subset whose Zariski closure is irreducible; for every $g, h \in M$, one has $g \cdot Y = h \cdot Y$.*

Proof. (1) The sets V and Y are algebraic over F; hence, $V(\mathbf{C}_p)$ and $Y(\mathbf{C}_p)$ are rigid F-subanalytic. Since $\mathfrak{F}$ is affinoid, the morphism $p|_{\mathfrak{F}}$ defines a rigid F-subanalytic map from $\mathfrak{F}(\mathbf{C}_p)$ to $V(\mathbf{C}_p)$, so that $(\mathfrak{F} \cap p^{-1}(V^{\mathrm{an}}))(\mathbf{C}_p)$ is a rigid F-subanalytic set. Consequently, taking $\mathbf{C}_p$-points, $(g \cdot Y^{\mathrm{an}} \cap \mathfrak{F} \cap p^{-1}(V^{\mathrm{an}}))_g$ furnishes a rigid F-subanalytic family of rigid F-subanalytic subsets of $\Omega(\mathbf{C}_p)$, parameterized by $G_0(\mathbf{C}_p)$. By b-minimality, the set of points $g \in G_0(\mathbf{C}_p)$ such that $\dim(g \cdot Y^{\mathrm{an}} \cap \mathfrak{F} \cap p^{-1}(V^{\mathrm{an}})) = m$ is a rigid F-subanalytic subset of $G_0(\mathbf{C}_p)$. It then follows from Lemma 4.2 that $R_{\mathfrak{F}}$ is an F-subanalytic subset of $G_0(F)$.

(2) Let $g \in R_{\mathfrak{F}}$ and let us prove that $(g \cdot Y^{\mathrm{an}}) \cap \Omega \subset p^{-1}(V^{\mathrm{an}})$. Since $g \cdot Y^{\mathrm{an}}$ is irreducible and $g \cdot Y^{\mathrm{an}} \cap \mathfrak{F}$ has dimension $m = \dim(g \cdot Y^{\mathrm{an}})$, this intersection is Zariski dense in $g \cdot Y^{\mathrm{an}}$. Moreover, there exists a finite extension F' of F such that $g \cdot Y^{\mathrm{an}}_{F'} \cap \mathfrak{F}(F')$ is Zariski dense in $Y_{F'}$ (it suffices that $g \cdot Y^{\mathrm{an}} \cap \mathfrak{F}$ admits a smooth F'-point), so that the Zariski closure of $g \cdot Y^{\mathrm{an}} \cap \mathfrak{F}(F')$ in $(\mathbf{P}^n_1)_{F'}$ is equal to $g \cdot Y_{F'}$. Moreover, $g \cdot Y(F') \cap \mathfrak{F}(F')$ is F'-semialgebraic. Hence, Proposition 5.8 implies that $g \cdot Y^{\mathrm{an}}_{F'} \cap \Omega_{F'} \subset p^{-1}_{F'}(V^{\mathrm{an}}_{F'})$. Since p is defined over F and $g \in G(F)$, this implies that $(g \cdot Y^{\mathrm{an}}) \cap \Omega \subset p^{-1}(V^{\mathrm{an}})$.

(3) As a subset, $(M \cdot Y^{\mathrm{an}}) \cap \Omega$ is contained in $p^{-1}(V^{\mathrm{an}})$. By Proposition 5.8, its Zariski closure Y' satisfies $(Y')^{\mathrm{an}} \cap \Omega \subset p^{-1}(V^{\mathrm{an}})$ as well. Since Y and the Zariski closure of M are geometrically irreducible, Y' is geometrically irreducible.

Let $g \in M$; then $Y^{\mathrm{an}} \subset g^{-1} M \cdot Y^{\mathrm{an}} \subset g^{-1} \cdot (Y')^{\mathrm{an}}$, and hence $W \subset g^{-1} \cdot (Y')^{\mathrm{an}} \cap \Omega$. By maximality of W, one has $W = g^{-1} \cdot (Y')^{\mathrm{an}} \cap \Omega$. This implies $g \cdot Y = Y'$. Thus $g \cdot Y = h \cdot Y$ for every $g, h \in M$. □

We return to the context of Section 8.6. In particular, ξ is a point of $Y(F)$ such that $q_1(\xi) \in \mathscr{L}_{\Gamma_1}$, and the restriction to Y of the projection to the first m coordinates is étale at ξ, with a local analytic section φ defined on $U_1 \times \cdots \times U_m$.

Lemma 8.9. *There exist a real number $c > 0$, fundamental sets $\mathfrak{F}_i \subset \Omega_{\Gamma_i}$ and a subset Υ of $R_{\mathfrak{F}} \cap \Gamma$, where $\mathfrak{F} = \prod \mathfrak{F}_i$, such that the following hold:*

(1) *For all T large enough, one has $\operatorname{Card}(\Upsilon_T) \geq T^c$, where Υ_T denotes the set of all $\gamma \in \Upsilon$ such that $H(\gamma) \leq T$.*

(2) *The projection* q_1 *is injective on* Υ.

(3) *For all* $j \in \{1, \dots, n\}$ *such that* $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$, *one has* $\operatorname{Card}(q_j(\Upsilon)) = 1$.

Recall that there exists a number field K contained in F such that $\Gamma \subset \mathrm{PGL}(2, K)^n$, and H is induced by a fixed height function on $\mathrm{PGL}(2, \overline{\mathbf{Q}})^n$. In particular, Lemma 8.9 implies that the subset $R_{\mathfrak{F}}$ of $\mathrm{PGL}(2, F)^n$ has many K-rational points, in the sense of Section 4.5.

Proof. Let q be the genus of X_{Γ_1}; by Proposition 6.8, there exists a basis $\alpha_1, \dots, \alpha_q$ of Γ_1, an open neighborhood U_1' of $q_1(\xi)$ which is contained in U_1 and stable under the action of $\alpha_1, \dots, \alpha_q$, and a fundamental set $\mathfrak{F}_1$ for Γ_1 contained in U_1'. For simplicity of notation, we now assume that $U_1 = U_1'$.

We have introduced in Section 8.6 a local analytic section

$$\varphi = (\varphi_1, \dots, \varphi_n) : U_1 \times \cdots \times U_m \to Y^{\mathrm{an}} \cap U_1 \times \cdots \times U_n$$

of the projection $q_J : Y \to \mathbf{P}_1^J$, where $J = \{1, \dots, m\}$. Let $j \in \{1, \dots, n\}$ be such that $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$. Then $q_j(\xi)$ has a compact analytic neighborhood U_j' contained in Ω_{Γ_j}. Shrinking $U_1, \dots, U_m$ if necessary, we assume that the image of φ_j is contained in U_j' for every such j.

Let $a' = (a_1, \dots, a_n) \in W$ be a rigid point that belongs to the image of φ and such that $a_1 \in \mathfrak{F}_1$. Let $a = (a_1, \dots, a_m)$; we have $a' = \varphi(a)$. For $j \in \{2, \dots, n\}$, we also choose a fundamental set $\mathfrak{F}_j$ that contains a_j.

We claim that we can complete any element $\gamma_1 \in F_1$ which is a positive word γ_1 in $\alpha_1, \dots, \alpha_q$ to an element $\gamma \in \Gamma$ such that $\gamma^{-1} \in R_{\mathfrak{F}}$ and $H(\gamma) \ll c^{\ell_{\Gamma_1}(\gamma_1)}$, for some real number c.

Let us now prove the asserted claim. For any positive word γ_1 in $\alpha_1, \dots, \alpha_q$, one has $\gamma_1 \cdot a_1 \in U_1$; in particular, we can consider the point $a(\gamma_1) = (\gamma_1 \cdot a_1, a_2, \dots, a_m)$ of $U_1 \times \cdots \times U_m$ and its image $\varphi(a(\gamma_1))$ under the section φ.

By Section 6.3, there exists a real number $c_1 \geq 1$ such that $\delta(\alpha_j \cdot a_1; \mathscr{L}_{\Gamma_1}) \geq c_1^{-1} \delta(a_1; \mathscr{L}_{\Gamma_1})$, uniformly in a_1. By induction on the length $\ell_{\Gamma_1}(\gamma_1)$ of the positive word γ_1, this implies the inequality

$$\delta(\gamma_1 \cdot a_1; \mathscr{L}_{\Gamma_1}) \geq c_1^{-\ell_{\Gamma_1}(\gamma_1)}. \tag{8.9.1}$$

We first set $\gamma_2 = \cdots = \gamma_m = 1$.

Let $j > m$. Let $\psi_j : U_1 \to U_j$ be the analytic map with $\psi_j(x) = \varphi_j(x, a_2, \dots, a_m)$. By construction (Lemma 8.5), if $\psi_j(x) = \varphi_j(x, a_2, \dots, a_m) \in \mathscr{L}_{\Gamma_j}$, one has $x = q_1(x, a_2, \dots, a_m) \in \mathscr{L}_{\Gamma_1}$. In other words, one has $\psi_j^{-1}(\mathscr{L}_{\Gamma_j}) \subset \mathscr{L}_{\Gamma_1}$. Applying Lemma 6.11 to ψ_j, we obtain an inequality of the form

$$\delta(\varphi_j(x, a_2, \dots, a_m); \mathscr{L}_{\Gamma_j}) \gg \delta(x; \mathscr{L}_{\Gamma_1})^k,$$

for some integer $k \geq 0$ and all $x \in U_1$. In particular,

$$\delta(\varphi_j(a(\gamma_1)); \mathscr{L}_{\Gamma_j}) \gg \delta(\gamma_1 \cdot a_1; \mathscr{L}_{\Gamma_1})^k. \tag{8.9.2}$$

By Proposition 6.8, there exists $\gamma_j \in \Gamma_j$ such that $\varphi_j(a(\gamma_1)) \in \gamma_j \cdot \mathfrak{F}_j$. By Proposition 6.6 and Lemma 6.9, one has

$$H(\gamma_j) \ll \delta(\varphi_j(a(\gamma_1)); \mathscr{L}_{\Gamma_j})^{-\kappa}, \tag{8.9.3}$$

where κ is a positive real number, independent of γ_1. By equations (8.9.1), (8.9.2) and (8.9.3), we thus have

$$H(\gamma_j) \ll \delta(\gamma_1 \cdot a_1; \mathscr{L}_{\Gamma_j})^{-k\kappa} \ll c_1^{\ell_{\Gamma_1}(\gamma_1)k\kappa}. \tag{8.9.4}$$

Let $c = c_1^{k\kappa}$.

Let $\gamma = (\gamma_1, \dots, \gamma_n) \in \Gamma$. By what precedes, $H(\gamma) \ll c^{\ell_{\Gamma_1}(\gamma_1)}$. Moreover, $\varphi_j(a(\gamma_1)) \in \gamma_j \cdot \mathfrak{F}_j$ for every j; this follows from the fact that $a_j \in \mathfrak{F}_j$ if $j \leq m$, and from the construction of γ_j if $j > m$.

Let us prove $\gamma^{-1} \in R_{\mathfrak{F}}$. One has $W \subset p^{-1}(V^{\mathrm{an}})$ by assumption; since $\gamma \in \Gamma$, this implies $\gamma^{-1} \cdot W \subset p^{-1}(V^{\mathrm{an}})$. Consequently,

$$\gamma^{-1} \cdot Y^{\mathrm{an}} \cap \mathfrak{F} \cap p^{-1}(V^{\mathrm{an}}) \supset \gamma^{-1} \cdot W \cap \mathfrak{F} \cap p^{-1}(V^{\mathrm{an}}) = \gamma^{-1} \cdot W \cap \mathfrak{F}.$$

The analytic morphism

$$U_1 \times \cdots \times U_m \to W, \qquad (x_1, \dots, x_m) \mapsto \varphi(\gamma_1 \cdot x_1, x_2, \dots, x_m)$$

is an immersion and maps the point $a = (a_1, \dots, a_m)$ to the point $\varphi(a(\gamma_1)) \in \gamma \cdot \mathfrak{F}$. Since a is a rigid point, this morphism maps a neighborhood of a into $\gamma \cdot \mathfrak{F}$, so that $\dim(W \cap \gamma \cdot \mathfrak{F}) \geq m$. This proves $\gamma^{-1} \in R_{\mathfrak{F}}$.

Applying Lemma 6.9 to estimate $H(\gamma_1)$, we thus have shown the existence of a positive real number c such that for every positive word γ_1 in $\alpha_1, \dots, \alpha_q$, there exists an element $\gamma = (\gamma_1, \dots, \gamma_n)$ completing γ_1 such that $H(\gamma) \ll c^{\ell_{\Gamma_1}(\gamma_1)}$ and $\gamma^{-1} \in R_{\mathfrak{F}} \cap \Gamma$.

Let Υ' be the set of all such elements γ^{-1}, where γ_1 ranges over positive words in $\alpha_1, \dots, \alpha_q$. It is a subset of $R_{\mathfrak{F}} \cap \Gamma$. By construction, the projection q_1 is injective on Υ'. Moreover, since the number of positive words of length ℓ in $\alpha_1, \dots, \alpha_q$ is q^ℓ, the cardinality of Υ'_T is bounded from below by $q^{\log(T)/\log(c)} = T^{\log(q)/\log(c)}$, and the exponent of T is strictly positive, since $q \geq 2$. Finally, let j be such that $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$. By construction, $\varphi_j(a(\gamma_1)) \in \gamma_j \mathfrak{F}_j$; hence $\gamma_j \mathfrak{F}_j$ meets U'_j. By Corollary 6.7, the set S_j of such elements γ_j in Γ_j is finite. It follows that there is a subset Υ of Υ' that satisfies the conclusion of the proposition. □

Proposition 8.10. *Let G'_0 be the subgroup of G_0 consisting of elements (g_j) such that $g_j = \mathrm{id}$ if $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$. Both the stabilizer of W inside $G'_0 \cap \Gamma$ and its image in Γ_1 under the first projection have many rational points.*

Proof. Let c, Υ, $\mathfrak{F}_i$, $\mathfrak{F} = \prod \mathfrak{F}_i$ and $R = R_{\mathfrak{F}}$ be as given by Lemma 8.9; let $T_0 > 1$ be such that $\operatorname{Card}(\Upsilon_T) \geq T^c$ for $T \geq T_0$.

Let K be a number field contained in F such that all groups Γ_j are contained in $\mathrm{PGL}(2, K)$; the points of $R \cap \Gamma$ are K-rational points. Recall that for every real number T, we denote by $R(K; T)$ the set of K-rational points of R of height $\leq T$. One has $\Upsilon_T = \Upsilon \cap R(K; T)$.

Since R is F-subanalytic (Lemma 8.8), it is also $\mathbf{Q}_p$-subanalytic and we may apply the p-adic Pila–Wilkie theorem of [Cluckers et al. 2015], as stated in Theorem 4.7. Thus let $s \in \mathbf{N}$, $d \in \mathbf{R}$, $\varepsilon > 0$ and $B \subset R \times \mathbf{Q}_p^s$ be a family of blocks such that for every $T > 1$, there exists a subset $\Sigma_T \in \mathbf{Q}_p^s$ of cardinality $< dT^\varepsilon$ such that $R(K; T) \subset \bigcup_{\sigma \in \Sigma_T} B_\sigma$. Let also $t \in \mathbf{N}$ and $Z \subset G_0(F) \times \mathbf{Q}_p^t$ be a semialgebraic subset such that for every $\sigma \in \mathbf{Q}_p^s$, there exists $\tau \in \mathbf{Q}_p^t$ such that $B_\sigma \subset Z_\tau$ and $\dim(B_\sigma) = \dim(Z_\tau)$. Let finally r be an upper bound for the number of irreducible components of the Zariski closure of the sets Z_τ, for $\tau \in \mathbf{Q}_p^t$.

Let $T > T_0$. Since $\Upsilon_T \subset R(K; T)$, by the pigeonhole principle, there exists $\sigma \in \Sigma_T$ such that

$$\operatorname{Card}(\Upsilon_T \cap B_\sigma) \geq \frac{\operatorname{Card}(\Upsilon_T)}{\operatorname{Card}(\Sigma_T)} \geq \frac{1}{d} T^{c-\varepsilon}.$$

Moreover, the Zariski closure of B_σ in $\mathrm{PGL}(2)_F^n$ has at most r irreducible components. Consequently, we may choose such an irreducible component $\overline{M}$ whose trace M on B_σ satisfies

$$\operatorname{Card}(\Upsilon_T \cap M) \geq \frac{1}{dr} T^{c-\varepsilon}.$$

(Observe that $\overline{M}$ is indeed the Zariski closure of M.)

Let $g \in \Upsilon_T \cap M$. Since the Zariski closure of M is irreducible and $M \subset R_{\mathfrak{F}}$, it follows from Lemma 8.8 that the stabilizer of W inside $G_0 \cap \Gamma$ contains $g^{-1}M$; hence $g^{-1}(\Upsilon_T \cap M)$. By construction, the image of $g^{-1}(\Upsilon_T \cap M)$ under the projection of index j is $\{\mathrm{id}\}$ if $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$. This shows in particular that the stabilizer of W inside $G_0' \cap \Gamma$ contains $g^{-1}(\Upsilon_T \cap M)$. This set contains $\geq T^{c-\varepsilon}/dr$ points, and their heights are $\ll T^2$; the same holds for its image by the first projection, since this projection is injective on $g^{-1}(\Upsilon \cap M)$.

We thus have shown that the stabilizer of W inside $G_0' \cap \Gamma$ has many rational points, as well as its image under the first projection, concluding the proof. □

Proposition 8.11. *The subvariety W is flat.*

Proof. We have constructed in Section 8.6 an analytic map $\varphi : U_1 \times \cdots \times U_m \to Y$, which is a local section of the projection to the m first coordinates.

Let $a \in \prod_{i=2}^m (\Omega_{\Gamma_i} \cap U_i)$; let us denote by W_a the fiber of W over a under the projection to $\prod_{i=2}^m \mathbf{P}_1^{\mathrm{an}}$, and Y_a similarly. When a varies, the number of irreducible components of Y_a is uniformly bounded.

Let $\psi_a : (U_1)_{\mathscr{H}(a)} \to Y_a^{\mathrm{an}}$ be the analytic morphism deduced from φ. We claim that the components of ψ_a are either constant or homographies.

Let $g \in G_0 \cap \Gamma$ be an element such that $g \cdot W = W$, $g_1 \neq \mathrm{id}$ and $g_j = \mathrm{id}$ if $q_j(\xi) \notin \mathscr{L}_{\Gamma_j}$ (Proposition 8.10). Since $g \cdot W = W$, one has $g \cdot Y = Y$. Hence $g \cdot W_a = W_a$ and $g \cdot Y_a = Y_a$. The element g induces a commutative diagram

$$\begin{tikzcd} Y_a \arrow[r, "g"] \arrow[d] & Y_a \arrow[d] \\ (\mathbf{P}_1)_{\mathscr{H}(a)} \arrow[r, "g_1"] \arrow[u, dashed, bend left, "\psi_a"] & (\mathbf{P}_1)_{\mathscr{H}(a)} \arrow[u, dashed, bend right, "\psi_a"'] \end{tikzcd}$$

where the section ψ_a is analytic and defined over the open subset $(U_1)_{\mathscr{H}(a)}$ of $(\mathbf{P}_1)^{\mathrm{an}}_{\mathscr{H}(a)}$. Let Y'_a be the irreducible component of Y_a that contains $\psi_a(\xi_1)$; it is geometrically irreducible. Recall that g_1 has infinite order; replacing g_1 and g by some fixed power, we may thus assume that $g \cdot Y'_a = Y'_a$.

By Proposition 7.2, either $Y'_a \to (\mathbf{P}_1)_{\mathscr{H}(a)}$ is an isomorphism and the components of its inverse are constant or homographies, or there exists a subset Δ of $\mathbf{P}_1(\overline{\mathscr{H}(a)})$ such that $\mathrm{Card}(\Delta) = 2$ and $g_1(\Delta) = \Delta$ for every element $g = (g_1, \dots, g_n) \in G'_0 \cap \Gamma$ such that $g \cdot W = W$ and $g \cdot Y'_a = Y'_a$. Let us assume that we are in the latter case. Using that $\Gamma_1 \subset \mathrm{PGL}(2, F)$, we see that $\Delta \subset \mathbf{P}_1(\bar{F})$. By Lemma 6.10, the projection to Γ_1 of the stabilizer of W inside $G'_0 \cap \Gamma$ has few rational points, contradicting Proposition 8.10.

We thus have shown that the components of the analytic map ψ_a are either constant or given by homographies.

Let $j \in \{m+1, \dots, n\}$.

First assume that $q_j(\xi) \in \Omega_{\Gamma_j}$. Then $g_j = \mathrm{id}$, whence the relation $\psi_{a,j} = \psi_{a,j} \circ g_1$. Since $g_1 \neq \mathrm{id}$, this implies that $\psi_{a,j}$ is constant, i.e., φ_j does not depend on the coordinate x_1. Since U is reduced, the morphism φ_j is deduced by pull-back of an analytic map $\theta_j : \prod_{i=2}^m U_i \to \mathbf{P}_1^{\mathrm{an}}$.

Let us then assume that $q_j(\xi) \in \mathscr{L}_{\Gamma_j}$. Since the j-th component of φ takes the value $q_j(\xi)$, the section $\psi_{a,j}$ cannot be constant. It is thus a homography $\tau_{j,a}$.

A priori, one has $\tau_{j,a} \in \mathrm{PGL}(2, \mathscr{H}(a))$ for every a. However, by condition (3) of Lemma 8.5, one has $\varphi_j^{-1}(\mathscr{L}_{\Gamma_j}) \subset \mathscr{L}_{\Gamma_1}$. The limit sets $\mathscr{L}_{\Gamma_1}$ and $\mathscr{L}_{\Gamma_j}$ are contained in $\mathbf{P}_1(F)$ and have no isolated points, so that $\tau_{j,a}^{-1}$ maps an infinite subset of $\mathbf{P}_1(F)$ into $\mathbf{P}_1(F)$; this implies that $\tau_{j,a} \in \mathrm{PGL}(2, F)$.

Observe that for $x \in U_1 \cap \mathbf{P}_1(F)$, one has $\tau_{j,a} \cdot x = \psi_{a,j}(x) = \varphi(x, a)$. In particular, the assignment $a \mapsto \tau_{j,a}$ is induced by an analytic morphism. Since it takes its values in $\mathrm{PGL}(2, F)$, it is constant.

Let J' and J'' be the set of all $j \in \{m+1, \dots, n\}$ such that $q_j(\xi)$ belongs to $\mathscr{L}_{\Gamma_j}$ and Ω_{Γ_j}, respectively. Let $\Omega' = \Omega_{\Gamma_1} \times \prod_{j \in J'} \Omega_{\Gamma_j}$ and $\Omega'' = \prod_{i=2}^m \Omega_{\Gamma_i} \times \prod_{j \in J''} \Omega_{\Gamma_j}$; similarly, write $X' = X_1 \times \prod_{j \in J'} X_j$ and $X'' = \prod_{i=2}^m X_i \times \prod_{j \in J''} X_j$, and decompose the projection $p : \Omega \to X$ as (p', p''), where $p' : \Omega' \to X'$ and $p'' : \Omega'' \to X''$ are the natural projections.

Let Z' be the graph in $\left(\mathbf{P}_1 \times \prod_{j \in J'} \mathbf{P}_1\right)^{\mathrm{an}}$ of $(\tau_j)_{j \in J'}$ and Z'' the graph in $\left(\prod_{i=2}^{m} \mathbf{P}_1 \times \prod_{j \in J''} \mathbf{P}_1\right)^{\mathrm{an}}$ of $(\theta_j)_{j \in J''}$. Let Y' and Y'' be the Zariski closure of Z' and Z'', let W' and W'' be their traces in Ω' and Ω'', and let V' and V'' be the Zariski closures of $p'(Z')$ and $p''(Z'')$. It is clear that $Y' = Z'$ is the curve in $\mathbf{P}_1 \times \prod_{j \in J'} \mathbf{P}_1$ (with coordinates x_1 and x_j for $j \in J'$) given by the equations $x_j = \tau_j(x_1)$, and W' is its trace on Ω'. In particular, W' is flat.

By construction, $Z' \times Z''$ is a subspace of Y^{an} which meets W in a Zariski dense subset of itself; hence $Y = Y' \times Y''$ and $W = \Omega \cap Y^{\mathrm{an}} = W' \times W''$. Moreover, $p(W) = p'(W') \times p''(W'') \subset V$; hence $V' \times V'' \subset V$. Consequently, W'' is a maximal algebraic irreducible subset of $(p'')^{-1}((V'')^{\mathrm{an}})$. By induction, W'' is flat.

Consequently, $W = W' \times W''$ is flat, as was to be shown. □

9. A characterization of geodesic subvarieties

9.1. Let F be a finite extension of $\mathbf{Q}_p$ and let $(\Gamma_i)_{1 \le i \le n}$ be a finite family of arithmetic Schottky subgroups of ranks ≥ 2 in $\mathrm{PGL}(2, F)$ Let us set $\Omega = \prod_{i=1}^{n} \Omega_{\Gamma_i}$, $X = \prod_{i=1}^{n} X_{\Gamma_i}$, and let $p : \Omega \to X^{\mathrm{an}}$ be the morphism deduced from the morphisms $p_{\Gamma_i} : \Omega_{\Gamma_i} \to X_{\Gamma_i}^{\mathrm{an}}$.

Theorem 9.2. *Let W be a Zariski closed subvariety of Ω, geometrically irreducible. Then the following properties are equivalent*:

(i) *The variety W is geodesic.*

(ii) *Its projection $p(W)$ is algebraic.*

(iii) *The dimension of the Zariski closure of $p(W)$ in X is equal to* $\dim(W)$.

Proof. Let us assume that W is geodesic and show that $p(W)$ is algebraic.

We may assume that no projection p_{Γ_i} is constant on W. Define a relation $\sim$ on $\{1, \dots, n\}$ given by $i \sim j$ if there exists $g \in \mathrm{PGL}(2, F)$ (necessarily unique) such that $g\Gamma_i g^{-1}$ and Γ_j are commensurable and $z_j = g \cdot z_i$ for every $z \in W$. This is an equivalence relation. Fix an element j in each equivalence class; for i such that $i \sim j$, we may replace Γ_i by its conjugate $g\Gamma_i g^{-1}$ and assume that $z_j = z_i$ on W. This shows that W and Ω decompose as a product indexed by the set of equivalence classes of the following particular situation: all the subgroups Γ_i are commensurable, and W is the diagonal of Ω. It thus suffices to treat this particular case.

Let $\Gamma_0 = \bigcap_i \Gamma_i$ and X_0 be the algebraic curve associated with $\Omega_{\Gamma_0}/\Gamma_0$. Then, for every i, the morphism $f_i : W \to X_i^{\mathrm{an}}$ deduced from $f = p|_W$ factors as the composition of the uniformization $p_0 : \Omega_{\Gamma_0} \to X_0^{\mathrm{an}}$ and of a finite morphism $X_0^{\mathrm{an}} \to X_i^{\mathrm{an}}$. By GAGA [Berkovich 1990, Corollary 3.5.2; Poineau 2010, Appendix], a finite analytic morphism of algebraic curves is algebraic; consequently, there exists a finite morphism $q_i : X_0 \to X_i$ such that $f_i = q_i^{\mathrm{an}} \circ p_0$. Then $p(W)$ is the image of X_0 by the finite morphism $q = (q_1, \dots, q_n) : X_0 \to X$, hence is algebraic. This shows that (i) implies (ii). Since it is clear that (ii) implies (iii), it remains to prove that (iii) implies (i).

Let us assume now that the dimension of the Zariski closure V of $p(W)$ in X is equal to the dimension of W. By construction, W is a maximal irreducible algebraic subvariety of $p^{-1}(V^{\mathrm{an}})$. By Proposition 8.1, W is flat. A similar analysis as in the proof of the first implication shows that there is a partition of the indices $\{1, \dots, n\}$ under which W decomposes as a product of flat curves and points. Since it suffices to prove that each of these curves is geodesic, we may assume that W is a flat curve of the form

$$W = \{(z, g_2 \cdot z, \dots, g_n \cdot z)\} \cap \Omega,$$

where $g_2, \dots, g_n \in \mathrm{PGL}(2, F)$.

First assume that $n = 2$. Let then $g \in \mathrm{PGL}(2, F)$ be such that $W = \{(z, g \cdot z)\} \cap \Omega$ and let us prove that Γ_2 and $g\Gamma_1 g^{-1}$ are commensurable, a property which is equivalent to the finiteness of both orbit sets $\Gamma_2\backslash\Gamma_2 g\Gamma_1$ and $\Gamma_1\backslash\Gamma_1 g^{-1}\Gamma_2$.

Let us argue by contradiction and assume that $\Gamma_2\backslash\Gamma_2 g\Gamma_1$ is infinite. (The other finiteness is analogous, or follows by symmetry.) Fix a rigid point $z \in \Omega_{\Gamma_1}$. Let $A \subset \Gamma_1$ be a set such that gA is a set of representatives of $\Gamma_2\backslash\Gamma_2 g\Gamma_1$; by assumption, A is infinite. Since $\Gamma\backslash W \subset V^{\mathrm{an}}$, the algebraic variety V contains the infinite set of points $p(a \cdot z, g \cdot az) = (p_1(z), p_2(ga \cdot z))$, for $a \in A$; hence it contains its Zariski closure $\{p_1(z)\} \times X_2$. Since this holds for every $z \in W$, we deduce that V contains $X_1 \times X_2$, contradicting the assumption that $\dim(W) = 1$.

Let us now return to the general case. To prove that W is geodesic, it suffices to establish that the subgroups Γ_j and $g_j\Gamma_1 g_j^{-1}$ are commensurable for every $j \in \{2, \dots, n\}$. Up to renumbering the indices, it suffices to treat the case $j = 2$. Let $\Omega' = \Omega_{\Gamma_1} \times \Omega_{\Gamma_2}$, let $p' : \Omega' \to X' = X_1 \times X_2$ be the uniformization map, and denote by π the projections from Ω to Ω' and from X to X'. Let $W' = \pi(W)$ and $V' = \pi(V)$. By Chevalley's theorem, V' is an algebraic curve in X'. Obviously, W' is a flat curve contained in $(p')^{-1}((V')^{\mathrm{an}})$, and hence is a maximal irreducible algebraic subset of $(p')^{-1}((V')^{\mathrm{an}}) \cap \Omega'$. By the case $n = 2$, the Schottky groups Γ_2 and $g_2\Gamma_1 g_2^{-1}$ are commensurable, as was to be shown. This concludes the proof of Theorem 9.2. □

Corollary 9.3. *Let V be an irreducible* curve *in X. Then every irreducible algebraic subvariety of $\Omega_{\mathbf{C}_p}$ which is maximal among those contained in $p^{-1}(V_{\mathbf{C}_p}^{\mathrm{an}})$ is geodesic.*

Proof. Let W_0 be an irreducible algebraic subvariety of $\Omega_{\mathbf{C}_p}$, maximal among those contained in $p^{-1}(V_{\mathbf{C}_p}^{\mathrm{an}})$; let us prove that W_0 is geodesic. We may assume that $\dim(W_0) > 0$. Since p is surjective and has discrete fibers, one has $\dim(p^{-1}(V_{\mathbf{C}_p}^{\mathrm{an}})) = \dim(V_{\mathbf{C}_p}^{\mathrm{an}})$, hence $\dim(W_0) = 1$, so that W_0 is an irreducible component of $p^{-1}(V^{\mathrm{an}})_{\mathbf{C}_p}$. By Theorem 7.16 of [Ducros 2009], there exists a finite extension E of F and an irreducible component W of $p^{-1}(V^{\mathrm{an}})_E$ such that $W_0 = W_{\mathbf{C}_p}$.

By Theorem 9.2, W is geodesic. Consequently, W_0 is geodesic. □

Remark 9.4. This corollary suggests that the main results of the paper extend to maximal algebraic irreducible subvarieties of $p^{-1}(V^{\mathrm{an}})_{\mathbf{C}_p}$, without assuming that they are defined over a finite extension of F.

Acknowledgements

The research leading to this paper was initiated during the 2014 MSRI program "Model Theory, Arithmetic Geometry and Number Theory". We would like to thank the MSRI for its congenial atmosphere and hospitality. This research was partially supported by ANR-13-BS01-0006 (Valcomo) and by ANR-15-CE40-0008 (Défigéo). Loeser was partially supported by the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC Grant Agreement nr. 246903 NMNAG and the Institut Universitaire de France.

During the early stages of this project, Loeser benefited from stimulating discussions with Barry Mazur at MSRI, to whom we express our heartfelt thanks. We are also grateful to Daniel Bertrand for continuous support and encouragement.

The comments of Yves André, Jean-François Boutot, Zoé Chatzidakis, Antoine Ducros, Florent Martin and Jonathan Pila helped us to improve this paper, as well as the numerous remarks of a diligent referee; we thank them warmly.

References

[André 2003] Y. André, *Period mappings and differential equations: From $\mathbb{C}$ to $\mathbb{C}_p$*, MSJ Memoirs **12**, Mathematical Society of Japan, Tokyo, 2003. MR Zbl

[Ax 1971] J. Ax, "On Schanuel's conjectures", *Ann. of Math.* (2) **93** (1971), 252–268. MR Zbl

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

[Berkovich 1993] V. G. Berkovich, "Étale cohomology for non-Archimedean analytic spaces", *Inst. Hautes Études Sci. Publ. Math.* **78** (1993), 5–161. MR Zbl

[Bertrand 1976] D. Bertrand, "Séries d'Eisenstein et transcendance", *Bull. Soc. Math. France* **104**:3 (1976), 309–321. MR Zbl

[Boutot and Carayol 1992] J.-F. Boutot and H. Carayol, "Uniformisation p-adique des courbes de Shimura: les théorèmes de Čerednik et de Drinfel'd", pp. 45–158 in *Courbes modulaires et courbes de Shimura* (Orsay, 1987/1988), Astérisque **196-197**, Société Mathématique de France, Paris, 1992. MR Zbl

[Boutot and Zink 1995] J.-F. Boutot and T. Zink, "The p-adic uniformization of Shimura curves", preprint 95-107, University of Bielefeld SFB 343, 1995, http://sfb343.math.uni-bielefeld.de/sfb343/preprints/pr95107.ps.gz.

[Čerednik 1976] I. V. Čerednik, "Towers of algebraic curves that can be uniformized by discrete subgroups of $\mathrm{PGL}_2(k_w) \times E$", *Mat. Sb.* (*N.S.*) **99(141)**:2 (1976), 211–247, 296. In Russian; translated in *Math. USSR-Sb.* **28**:2 (1978), 187–215. MR

[Clark 2003] P. L. Clark, *Rational points on Atkin–Lehner quotients of Shimura curves*, Ph.D. thesis, Harvard University, 2003, https://search.proquest.com/docview/305334169. MR

[Cluckers and Loeser 2007] R. Cluckers and F. Loeser, "b-minimality", *J. Math. Log.* **7**:2 (2007), 195–227. MR Zbl

[Cluckers et al. 2006] R. Cluckers, L. Lipshitz, and Z. Robinson, "Analytic cell decomposition and analytic motivic integration", *Ann. Sci. École Norm. Sup.* (4) **39**:4 (2006), 535–568. MR Zbl

[Cluckers et al. 2015] R. Cluckers, G. Comte, and F. Loeser, "Non-Archimedean Yomdin–Gromov parametrizations and points of bounded height", *Forum Math. Pi* **3** (2015), e5, 60. MR Zbl

[Conrad 1999] B. Conrad, "Irreducible components of rigid spaces", *Ann. Inst. Fourier* (*Grenoble*) **49**:2 (1999), 473–541. MR Zbl

[Denef 1986] J. Denef, "p-adic semi-algebraic sets and cell decomposition", *J. Reine Angew. Math.* **369** (1986), 154–166. MR Zbl

[Denef and van den Dries 1988] J. Denef and L. van den Dries, "p-adic and real subanalytic sets", *Ann. of Math.* (2) **128**:1 (1988), 79–138. MR Zbl

[Drinfel'd 1976] V. G. Drinfel'd, "Coverings of p-adic symmetric domains", *Funkcional. Anal. i Priložen.* **10**:2 (1976), 29–40. In Russian; translated in *Funct. Anal. Appl.* **10**:2 (1976), 107–115. MR

[Ducros 2009] A. Ducros, "Les espaces de Berkovich sont excellents", *Ann. Inst. Fourier* (*Grenoble*) **59**:4 (2009), 1443–1552. MR Zbl

[Ducros 2016] A. Ducros, "Families of Berkovich spaces", preprint, 2016. To appear in *Astérisque*. arXiv

[Gerritzen and van der Put 1980] L. Gerritzen and M. van der Put, *Schottky groups and Mumford curves*, Lecture Notes in Mathematics **817**, Springer, 1980. MR Zbl

[Klingler et al. 2016] B. Klingler, E. Ullmo, and A. Yafaev, "The hyperbolic Ax–Lindemann–Weierstrass conjecture", *Publ. Math. Inst. Hautes Études Sci.* **123** (2016), 333–360. MR Zbl

[Lipshitz 1993] L. Lipshitz, "Rigid subanalytic sets", *Amer. J. Math.* **115**:1 (1993), 77–108. MR Zbl

[Lipshitz and Robinson 2000a] L. Lipshitz and Z. Robinson, "Model completeness and subanalytic sets", pp. 109–126 in *Rings of separated power series and quasi-affinoid geometry*, Astérisque **264**, Société Mathématique de France, Paris, 2000. MR Zbl

[Lipshitz and Robinson 2000b] L. Lipshitz and Z. Robinson, "Quasi-affinoid varieties", pp. 127–149 in *Rings of separated power series and quasi-affinoid geometry*, Astérisque **264**, Société Mathématique de France, Paris, 2000. MR Zbl

[Mumford 1970] D. Mumford, *Abelian varieties*, Tata Institute of Fundamental Research Studies in Math. **5**, Oxford University Press, 1970. MR Zbl

[Pila 2009] J. Pila, "On the algebraic points of a definable set", *Selecta Math.* (*N.S.*) **15**:1 (2009), 151–170. MR Zbl

[Pila 2011] J. Pila, "O-minimality and the André–Oort conjecture for $\mathbb{C}^n$", *Ann. of Math.* (2) **173**:3 (2011), 1779–1840. MR Zbl

[Pila 2013] J. Pila, "Modular Ax–Lindemann–Weierstrass with derivatives", *Notre Dame J. Form. Log.* **54**:3-4 (2013), 553–565. MR Zbl

[Pila 2015] J. Pila, "Functional transcendence via o-minimality", pp. 66–99 in *O-minimality and diophantine geometry* (Manchester, 2013), edited by G. O. Jones and A. J. Wilkie, London Mathematical Society Lecture Note Series **421**, Cambridge University Press, 2015. MR Zbl

[Pila and Tsimerman 2013] J. Pila and J. Tsimerman, "The André–Oort conjecture for the moduli space of abelian surfaces", *Compos. Math.* **149**:2 (2013), 204–216. MR Zbl

[Pila and Wilkie 2006] J. Pila and A. J. Wilkie, "The rational points of a definable set", *Duke Math. J.* **133**:3 (2006), 591–616. MR Zbl

[Pila and Zannier 2008] J. Pila and U. Zannier, "Rational points in periodic analytic sets and the Manin–Mumford conjecture", *Atti Accad. Naz. Lincei Rend. Lincei Mat. Appl.* **19**:2 (2008), 149–162. MR Zbl

[Poineau 2010] J. Poineau, "Raccord sur les espaces de Berkovich", *Algebra Number Theory* **4**:3 (2010), 297–334. MR Zbl

[Rumely 1989] R. S. Rumely, *Capacity theory on algebraic curves*, Lecture Notes in Mathematics **1378**, Springer, 1989. MR Zbl

[Seidenberg 1958] A. Seidenberg, "Abstract differential algebra and the analytic case", *Proc. Amer. Math. Soc.* **9** (1958), 159–164. MR Zbl

[Shimura 1961] G. Shimura, "On the zeta-functions of the algebraic curves uniformized by certain automorphic functions", *J. Math. Soc. Japan* **13** (1961), 275–331. MR Zbl

[Ullmo and Yafaev 2014] E. Ullmo and A. Yafaev, "Hyperbolic Ax–Lindemann theorem in the cocompact case", *Duke Math. J.* **163**:2 (2014), 433–463. MR Zbl

Communicated by Jonathan Pila
Received 2015-11-20 Revised 2017-09-02 Accepted 2017-09-03

antoine.chambert-loir@math.univ-paris-diderot.fr
Univ. Paris Diderot, Sorbonne Paris Cité, Institut de Mathématiques de Jussieu-Paris Rive Gauche, UMR 7586, F-75013, Paris, France

francois.loeser@upmc.fr
Sorbonne Universités, UPMC Univ Paris 06, UMR 7586 CNRS, Institut Mathématique de Jussieu-Paris Rive Gauche, F-75005, Paris, France

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2001

A modular description of $\mathscr{X}_0(n)$

Kęstutis Česnavičius

As we explain, when a positive integer n is not squarefree, even over $\mathbb{C}$ the moduli stack that parametrizes generalized elliptic curves equipped with an ample cyclic subgroup of order n does not agree at the cusps with the $\Gamma_0(n)$-level modular stack $\mathscr{X}_0(n)$ defined by Deligne and Rapoport via normalization. Following a suggestion of Deligne, we present a refined moduli stack of ample cyclic subgroups of order n that does recover $\mathscr{X}_0(n)$ over $\mathbb{Z}$ for all n. The resulting modular description enables us to extend the regularity theorem of Katz and Mazur: $\mathscr{X}_0(n)$ is also regular at the cusps. We also prove such regularity for $\mathscr{X}_1(n)$ and several other modular stacks, some of which have been treated by Conrad by a different method. For the proofs we introduce a tower of compactifications $\overline{\mathscr{E}\ell\ell}_m$ of the stack $\mathscr{E}\ell\ell$ that parametrizes elliptic curves—the ability to vary m in the tower permits robust reductions of the analysis of Drinfeld level structures on generalized elliptic curves to elliptic curve cases via congruences.

Chapter 1. Introduction

1.1. ***Algebraic stacks that refine*** $X_0(n)$**.** The study of the compactification $X_0(n)$ of the coarse moduli space of the algebraic stack $\mathscr{Y}_0(n)$ that parametrizes elliptic curves equipped with a cyclic subgroup of order n is key for many arithmetic problems, so one seeks to understand the arithmetic properties of $X_0(n)$, especially over $\mathbb{Z}$. For this, it is desirable to conceptualize the construction of $X_0(n)$ by realizing it as a coarse moduli space of an algebraic stack that compactifies $\mathscr{Y}_0(n)$.

MSC2010: primary 11G18; secondary 14D22, 14D23, 14G35.
Keywords: Elliptic curve, generalized elliptic curve, level structure, modular curve, moduli stack.

The sought compactifying stack $\mathscr{X}_0(n)$ was defined by Deligne and Rapoport [1973, IV.3.3] via a normalization procedure. However, $\mathscr{X}_0(n)$ lacks an *a priori* moduli interpretation, so instead one often considers the stack $\mathscr{X}_0(n)^{\mathrm{naive}}$ that parametrizes generalized elliptic curves whose smooth locus is equipped with a cyclic subgroup of order n that is ample, i.e., meets every irreducible component of every geometric fiber. Even though $\mathscr{X}_0(n)^{\mathrm{naive}}$ is algebraic, has $X_0(n)$ as its coarse moduli space, and agrees with $\mathscr{X}_0(n)$ on the elliptic curve locus, it seems to have been overlooked that

> If n is not squarefree, then $\mathscr{X}_0(n)$ and $\mathscr{X}_0(n)^{\mathrm{naive}}$ are genuinely different, even over $\mathbb{C}$.

1.2. ***Pathologies of*** $\mathscr{X}_0(p^2)^{\mathrm{naive}}$**.** To explain the difference, we set $n := p^2$ for some prime p, let $\mathscr{X}(1)$ denote the stack that parametrizes those generalized elliptic curves whose geometric fibers are integral, and consider the structure morphism

$$c : \mathscr{X}_0(p^2)^{\mathrm{naive}} \to \mathscr{X}(1)$$

which in terms of the moduli interpretation forgets the subgroup and contracts the generalized elliptic curve with respect to the identity section. We claim that the morphism c is not representable.

To see this, let E be the standard p-gon over $\mathbb{C}$ and let $\zeta_{p^2} \in \mathbb{C}^\times$ be a primitive root of unity of order p^2. Then $E^{\mathrm{sm}} = \mathbb{G}_m \times \mathbb{Z}/p\mathbb{Z}$ and each of the μ_p worth of automorphisms of E fixing $\mathbb{G}_m \times \{0\}$ stabilizes the cyclic subgroup $\langle(\zeta_{p^2}, 1)\rangle$ of order p^2. Each such automorphism contracts to the identity, so c is not representable.

In contrast, the morphism

$$\mathscr{X}_0(p^2) \to \mathscr{X}(1)$$

is representable by construction, so the $\mathscr{X}(1)$-stacks $\mathscr{X}_0(p^2)^{\mathrm{naive}}$ and $\mathscr{X}_0(p^2)$ are not isomorphic. The same p-gon example carried out over $\overline{\mathbb{F}}_p$ shows that $\mathscr{X}_0(p^2)^{\mathrm{naive}}$ is not even Deligne–Mumford (whereas $\mathscr{X}_0(p^2)$ is), a pathology that has already been pointed out in [Edixhoven 1990, 1.1.1.1; Conrad 2007].

1.3. ***A modular description of*** $\mathscr{X}_0(n)$**.** One of the main goals of this paper is to refine the definition of $\mathscr{X}_0(n)^{\mathrm{naive}}$ to obtain a moduli interpretation of $\mathscr{X}_0(n)$ even when n is not squarefree. The elliptic curve locus needs no refinement, so the issue is to incorporate the cusps in a way that avoids the nonrepresentability of c phenomenon. For this, we follow a suggestion of Deligne [2015]. To present Deligne's idea, we assume that $n = p^2$ for a prime p and work over $\mathbb{Z}[1/p]$.

In vague terms, the idea is to subsume the automorphisms causing the nonrepresentability of c into the moduli problem. To make this possible, the data being parametrized will involve algebraic stacks and not merely schemes. In precise terms, the moduli problem that in Chapter 5 will be proved to recover $\mathscr{X}_0(p^2)_{\mathbb{Z}[1/p]}$ assigns to every $\mathbb{Z}[1/p]$-scheme S the groupoid of tuples

$$\big(E \to S,\ G,\ S_{(1)},\ S_{(p)},\ S_{(p^2)},\ \mathcal{G}_{(1)},\ \mathcal{G}_{(p)},\ \mathcal{G}_{(p^2)}\big)$$

consisting of:

- a generalized elliptic curve $E \to S$;
- a cyclic subgroup $G \subset E_{S-S^\infty}$ of order p^2 over the elliptic curve locus $S - S^\infty$;
- open subschemes $S_{(1)}$, $S_{(p)}$, and $S_{(p^2)}$ of S that cover S, have $S - S^\infty$ as their pairwise intersections, and such that the degenerate geometric fibers of $E_{S_{(1)}}$ and $E_{S_{(p)}}$ are 1-gons and those of $E_{S_{(p^2)}}$ are p^2-gons;
- ample cyclic subgroups $\mathcal{G}_{(1)} \subset E^{\mathrm{sm}}_{S_{(1)}}$ and $\mathcal{G}_{(p^2)} \subset E^{\mathrm{sm}}_{S_{(p^2)}}$ of order p^2 that recover G over $S - S^\infty$;
- an ample cyclic subgroup $\mathcal{G}_{(p)} \subset \mathcal{E}^{\mathrm{sm}}_{(p)}$ of order p^2 of the universal generalized elliptic curve $\mathcal{E}_{(p)}$ whose degenerate geometric fibers are p-gons and whose contraction is $E_{S_{(p)}}$, subject to the requirement that $\mathcal{G}_{(p)}$ recovers G over $S - S^\infty$ (over which $\mathcal{E}_{(p)}$ is identified with E).

In essence, the moduli problem parametrizes generalized elliptic curves equipped with an ample cyclic subgroup of order p^2 with the caveat that over the part of the degeneracy locus prone to the nonrepresentability of c the subgroup has been upgraded to live inside a suitable universal "decontraction" $\mathcal{E}_{(p)}$ (which is an algebraic stack and not a scheme). The role of the $S_{(p^i)}$ is to remember the subdivision of the degeneracy locus S^∞ — without $S_{(1)}$ and $S_{(p)}$ we cannot single out those 1-gon degenerate geometric fibers of E that were "meant" to be p-gons but had to be "upgraded" in order to avoid the nonrepresentability of c.

1.4. ***Incorporating bad characteristics.*** After the work of Drinfeld and of Katz and Mazur, the extension of the above modular description of $\mathscr{X}_0(p^2)_{\mathbb{Z}[1/p]}$ to $\mathscr{X}_0(p^2)$ is a matter of technique. However, new difficulties at the cusps in characteristic p force us to impose an additional coherence requirement on $\mathcal{G}_{(p)}$, a requirement that holds automatically away from p and also on the elliptic curve locus (see Section 5.5 and Lemma 5.6) and that seems well suited for the analysis of $\mathcal{G}_{(p)}$ even over $\mathbb{Z}[1/p]$. With this proviso, we prove that for any n the analogue of the moduli problem described in Section 1.3 gives a moduli interpretation for $\mathscr{X}_0(n)$. We then use this moduli interpretation to prove the following extension of a regularity theorem of Katz and Mazur:

Theorem 1.5 (Theorem 5.13(a)). *The Deligne–Mumford stack $\mathscr{X}_0(n)$ is regular.*

In fact, $\mathscr{X}_0(n)_{\mathbb{Z}[1/n]}$ is even $\mathbb{Z}[1/n]$-smooth by [Deligne and Rapoport 1973, IV.6.7], whereas the elliptic curve locus $\mathscr{Y}_0(n)$ is regular by [Katz and Mazur 1985, 5.1.1], so Theorem 1.5 was known away from the closed substack of the cusps that lies in characteristics dividing n.

In the proof of Theorem 1.5, the eventual source of regularity is the combination of [Deligne and Rapoport 1973, V.4.13] and [Katz and Mazur 1985, 5.1.1] that proves the regularity of another modular stack $\mathscr{X}(n)$. The reduction to $\mathscr{X}(n)$ rests on the moduli interpretation of $\mathscr{X}_0(n)$ and on the regularity of $\mathscr{Y}_0(n)$. In particular, no stage of the argument requires any computations with universal deformation rings, other than what comes in from [Katz and Mazur 1985, Chapters 5–6] through our reliance on the regularity of $\mathscr{Y}(n)$ and $\mathscr{Y}_0(n)$.

We use Theorem 1.5 and the moduli interpretation of $\mathscr{X}_0(n)$ to prove that the coarse moduli space $X_0(n)$ is regular in a neighborhood of the cusps (see Theorem 6.7). This regularity is not new (see the introduction of Chapter 6) but our proof seems more conceptual.

1.6. ***The compactifications $\overline{\mathscr{E}\ell\ell}_m$.*** We have been vague about the base of the universal "decontraction" $\mathcal{E}_{(p)}$. For the construction of this base in general (beyond $n = p^2$), it is natural to fix an $m \in \mathbb{Z}_{\geq 1}$ and to consider the $\mathbb{Z}$-stack $\overline{\mathscr{E}\ell\ell}_m$ that parametrizes those generalized elliptic curves whose degenerate geometric fibers are m-gons. We prove in Theorem 3.1.6 that $\overline{\mathscr{E}\ell\ell}_m$ is algebraic, as well as proper and smooth over $\mathbb{Z}$, albeit is not Deligne–Mumford unless $m = 1$. Thus, each $\overline{\mathscr{E}\ell\ell}_m$ compactifies the stack $\mathscr{E}\ell\ell$ that parametrizes elliptic curves, and $\overline{\mathscr{E}\ell\ell}_1$ is the compactification that is sometimes called $\overline{\mathcal{M}}_{1,1}$.

As we describe in Section 3.2, the compactifications $\overline{\mathscr{E}\ell\ell}_m$ form an infinite tower, with transition maps given by contractions of generalized elliptic curves. This tower is the backbone of our study of $\mathscr{X}_0(n)$ and of several other "classical" modular curves. For these curves, the most important moduli-theoretic phenomenon that is not seen on the elliptic curve locus is the fact that "forgetful" contractions change generalized elliptic curves that underlie level structures. The ability to vary m in the tower $\{\overline{\mathscr{E}\ell\ell}_m\}_{m|m'}$ allows us to isolate the part of this phenomenon that has nothing to do with level structures. The remaining part that is specific to the level structure at hand may then be studied via "congruences" that reduce to the elliptic curve case.

1.7. ***Other modular curves.*** To illustrate the utility of $\overline{\mathscr{E}\ell\ell}_m$, let us consider the stack $\mathscr{X}(n)^{\mathrm{naive}}$ that parametrizes pairs consisting of a generalized elliptic curve $E \to S$ with n-gon degenerate geometric fibers and a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[n]$. (In the end, $\mathscr{X}(n)^{\mathrm{naive}}$ agrees with $\mathscr{X}(n)$ mentioned earlier and gives $\mathscr{X}(n)$ a moduli interpretation.) Using the work of Katz and Mazur, we prove via "mod n congruences with elliptic curves" that the forgetful map

$$\mathscr{X}(n)^{\mathrm{naive}} \to \overline{\mathscr{E}\ell\ell}_n$$

is representable and finite locally free of rank $\#\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$. It follows that $\mathscr{X}(n)^{\mathrm{naive}}$ is algebraic, proper and flat over $\mathbb{Z}$, and even Cohen–Macaulay. Other proofs of these properties of $\mathscr{X}(n)^{\mathrm{naive}}$ have been given by Conrad [2007]: the proof of the algebraicity used Hilbert schemes via tricanonical embeddings, whereas the Cohen–Macaulay property required a detailed analysis of the universal deformation rings at the cusps (in addition to the work of Katz and Mazur on the elliptic curve locus).

The relations with $\overline{\mathscr{E}\ell\ell}_m$ together with the "congruence method" that crucially uses the work of Katz and Mazur allow us to reprove the main results of [Conrad 2007] in Chapter 4. These include the moduli interpretations and the regularity of the modular stacks $\mathscr{X}(n)$ and $\mathscr{X}_1(n)$ (as well as some variants) and the construction of Hecke correspondences for $\mathscr{X}_1(n)$. The latter takes advantage of the theory of isogenies of generalized elliptic curves developed in Chapter 2. Away from the level, the moduli interpretations and the regularity have been proved by Deligne and Rapoport [1973, IV.3.5 and IV.4.14]; away from the cusps, they have been proved by Katz and Mazur [1985, 5.1.1]. Prior to the work of Conrad, [2007], the moduli interpretations and the regularity of $\mathscr{X}(n)$ and $\mathscr{X}_1(n)$ (among others) have been considered in an unfinished manuscript of Edixhoven [2001, especially 2.1.2].

1.8. ***Reliance on the literature.*** For what concerns generalized elliptic curves and Drinfeld level structures on them, we wish to explicate the logical dependence of our work on the three main references that we use: [Deligne and Rapoport 1973; Katz and Mazur 1985; Conrad 2007].

- We rely on [Deligne and Rapoport 1973] almost in its entirety; the sections of [op. cit.] that are logically independent from the work of this paper are II.§3, V.§2–3, VI.§2–6, and VII.§3–4.
- We make essential use of the results of [Katz and Mazur 1985, Chapters 1–6] and extend some of them to generalized elliptic curves (see, in particular, Section 4.2), but have no need for the results of [Katz and Mazur 1985, Chapters 7–14] (other than for comparison in Proposition 6.3 and Remarks 6.5 and 6.8).
- We use some auxiliary general results from sections 2.1 and 2.2 of [Conrad 2007] but the rest of [op. cit.] is logically independent from our work (as mentioned in Section 1.7, we give different proofs to the main results of [Conrad 2007]).

1.9. ***Notation and conventions.*** We let $\mathscr{E}\ell\ell$ denote the $\mathbb{Z}$-stack that, for variable schemes S, parametrizes elliptic curves $E \to S$. More precisely, for a scheme S, the objects (resp. the morphisms) of the groupoid $\mathscr{E}\ell\ell(S)$ are the elliptic curves $E \to S$ (resp. the isomorphisms between elliptic curves over S) and, for a scheme morphism $S' \to S$, the induced functor $\mathscr{E}\ell\ell(S) \to \mathscr{E}\ell\ell(S')$ is $E \mapsto E \times_S S'$. We use the analogous meaning of "parametrizes" when defining other stacks. Other than in the introduction, we use the notation $\mathscr{X}_{\Gamma_0(n)}$ (resp. $\mathscr{X}_{\Gamma_1(n)}$, etc.) introduced in Section 4.1.2 for stacky modular curves defined via normalization and the notation $\mathscr{X}_0(n)$ (resp. $\mathscr{X}_1(n)$, etc.) for stacks defined in terms of a moduli problem; once we prove that $\mathscr{X}_{\Gamma_0(n)} = \mathscr{X}_0(n)$ (and similarly in the other cases), we use the two notations interchangeably.

We use the definition of an fpqc cover for which all Zariski covers are fpqc; explicitly, $S' \to S$ is an fpqc cover if it is flat and every affine open $U \subset S$ is the union of images of finitely many affine opens of S'. An S-scheme S' is an fppf cover (or simply fppf) if $S' \to S$ is faithfully flat and locally of finite presentation. For a scheme S, we let S^{red} denote its associated reduced scheme. For an S-group algebraic space G, we let G^0 denote the subsheaf of sections that fiberwise factor through the identity component. We let $\mathscr{X}^{\mathrm{sm}}$ and $\Delta_{\mathscr{X}/S}$ denote the smooth locus and the diagonal of a morphism $\mathscr{X} \to S$. For a field k, we let $\bar{k}$ denote a choice of its algebraic closure. A geometric point is the spectrum of an algebraically closed field. For an $n \in \mathbb{Z}_{\geq 1}$, we set $\phi(n) := \#(\mathbb{Z}/n\mathbb{Z})^\times$.

For what concerns algebraic stack and algebraic space conventions, we follow [SP 2005–], except that "representable" stands for "representable by algebraic spaces." In particular, quasicompactness or separatedness of the diagonal are not part of the definition, but in practice end up being present (along with even stronger properties). An algebraic stack is Deligne–Mumford if its diagonal is unramified — for the equivalence with the étale atlas definition in the presence of quasicompactness and separatedness of the diagonal, see [Laumon and Moret-Bailly 2000, 8.1]. The relative dimension (at a point) of a smooth morphism of algebraic stacks is the difference of the relative dimensions (at a lift of the point) of the morphisms from a smooth atlas of the source, cf. [Laumon and Moret-Bailly 2000, bottom of p. 98].

Chapter 2. Isogenies of generalized elliptic curves

The main goal of this chapter is to expose a robust theory of isogenies of generalized elliptic curves. This theory is the subject of Section 2.2 and will be useful on several occasions, particularly, for algebraizing homomorphisms of formal generalized elliptic curves in Section 3.4 and for constructing Hecke correspondences for $\mathscr{X}_1(n)$ in Section 4.7. In order to prepare for the study of isogenies, in Section 2.1 we review several basic concepts, such as that of a homomorphism of generalized elliptic curves, and record some general results that will be useful throughout the paper.

2.1. Homomorphisms between generalized elliptic curves

In this section, we review basic definitions and properties of generalized elliptic curves, building up to the notion of a homomorphism, which will be studied in

Section 2.2. We assume that the reviewed concepts are familiar, so we concentrate on those aspects that will be used later. We begin with the notion of an n-gon, which is needed in order to define generalized elliptic curves. Informally, an n-gon is the curve obtained by gluing n-copies of $\mathbb{P}^1$ in a cyclic manner: the point 0 of the i-th copy gets identified with the point ∞ of the $(i+1)$-st copy.

Definition 2.1.1. For an $n \in \mathbb{Z}_{\geq 1}$ and an scheme S, the *standard n-gon over S* is the coequalizer of

$$\bigsqcup_{\mathbb{Z}/n\mathbb{Z}} S \rightrightarrows \bigsqcup_{\mathbb{Z}/n\mathbb{Z}} \mathbb{P}^1_S,$$

where the top (resp. the bottom) closed immersion includes the i-th copy of S as the 0 (resp. the ∞) section of the i-th (resp. $(i+1)$-st) copy of $\mathbb{P}^1_S$. A *Néron n-gon over S* (or an *n-gon over S*) is an S-scheme isomorphic to the standard n-gon over S. (We often omit "over S" if the base is implicit.)

Remark 2.1.2. Even though colimits usually do not exist in the category of schemes, the ones used in Definition 2.1.1 do exist and their formation commutes with base change in S. To see this, one checks directly (or with the help of [Ferrand 2003, 4.3]) that for $n \geq 2$ the sought coequalizer is the base change to S of the gluing of

$$\bigsqcup_{i\in\mathbb{Z}/n\mathbb{Z}} \mathrm{Spec}\big(\mathbb{Z}[X_i, Y_i]/(X_iY_i)\big)$$

obtained by identifying the opens

$$\mathrm{Spec}\big(\mathbb{Z}\big[Y_i, \tfrac{1}{Y_i}\big]\big) \quad \text{and} \quad \mathrm{Spec}\big(\mathbb{Z}\big[X_{i+1}, \tfrac{1}{X_{i+1}}\big]\big)$$

via $Y_i = 1/X_{i+1}$ for every $i \in \mathbb{Z}/n\mathbb{Z}$, and one treats the $n = 1$ case by realizing the standard 1-gon as the $\mathbb{Z}/n\mathbb{Z}$-quotient of the standard n-gon, cf. [Conrad 2007, top of p. 215].

We recall the definition of a generalized elliptic curve, which is a central notion for this paper.

Definition 2.1.3. A *generalized elliptic curve* over a scheme S is the data of

- a proper, flat, finitely presented morphism $E \to S$ each of whose geometric fibers is either a smooth connected curve of genus 1 or a Néron n-gon for some $n \geq 1$, and
- an S-morphism $E^{\mathrm{sm}} \times_S E \xrightarrow{+} E$ that restricts to a commutative S-group scheme structure on E^{sm} for which $+$ becomes an S-group action,

such that via pullback of line bundles the action $+$ induces the trivial action of E^{sm} on $\mathrm{Pic}^0_{E/S}$.

Remark 2.1.4. Our definition of a generalized elliptic curve is equivalent to the one given in [Deligne and Rapoport 1973, II.1.12]: the difference is that we have imposed the requirement that E^{sm} acts trivially on $\mathrm{Pic}^0_{E/S}$ at the outset. In [loc. cit.] this is replaced with the *a priori* milder requirement that on degenerate geometric fibers every translation by a smooth point induces a rotation on the underlying n-gon, which ends up being equivalent due to [Deligne and Rapoport 1973, II.1.7(ii) and II.1.13].

The requirement about the triviality of the induced action on $\mathrm{Pic}^0_{C/S}$ holds automatically on a large part of E^{sm}, namely, it always holds on the relative identity component $(E^{\mathrm{sm}})^0$ — to see this, we apply [Deligne and Rapoport 1973, II.1.14][1] to $\mathrm{Pic}^0_{E/S} \times_S E^{\mathrm{sm}}$ to get the openness of the locus of E^{sm} where the induced action on $\mathrm{Pic}^0_{E/S}$ is trivial, note that this locus is closed under the group law of E^{sm}, and conclude by noting that it contains the zero section. In particular, every elliptic curve is a generalized elliptic curve, and a generalized elliptic curve $E \to S$ is an elliptic curve over the open of S over which E is smooth.

Remark 2.1.5. The standard n-gon is canonically a generalized elliptic curve: due to its description recalled in Remark 2.1.2, its smooth locus is $\mathbb{G}_m \times \mathbb{Z}/n\mathbb{Z}$ and the translation action of this group scheme on itself extends to an action on the n-gon. By the previous remark, the triviality of the induced action on Pic^0 may be checked on the geometric fibers using [Deligne and Rapoport 1973, II.1.7(ii)]. For later use, we now describe the automorphism functor of this generalized elliptic curve.

Lemma 2.1.6. *For a fixed $n \in \mathbb{Z}_{\geq 1}$, let $E \to \operatorname{Spec} \mathbb{Z}$ be the standard n-gon generalized elliptic curve. There is the following identification of the automorphism functor of E:*

$$\mathrm{Aut}(E) \cong \mu_n \times \mathbb{Z}/2\mathbb{Z},$$

where the generator of $\mathbb{Z}/2\mathbb{Z}$ acts as inversion on E^{sm} and, for a scheme S and an index $i \in \mathbb{Z}/n\mathbb{Z}$, a section $\zeta \in \mu_n(S)$ acts on the i-th component of

$$E^{\mathrm{sm}}_S \cong (\mathbb{G}_m)_S \times \mathbb{Z}/n\mathbb{Z}$$

as scaling by ζ^i.

Proof. By [Deligne and Rapoport 1973, II.1.10], we have

$$\mathrm{Aut}(E) \cong \mu_n \rtimes \mathbb{Z}/2\mathbb{Z}$$

---

[1] We could also apply [Conrad 2007, 2.2.1] to avoid using the representability of $\mathrm{Pic}^0_{E/S}$ by a scheme. On the other hand, such representability may be proved as follows: by [Artin 1969, 7.3], the functor $\mathrm{Pic}^0_{E/S}$ is an algebraic space, so [Deligne and Rapoport 1973, II.2.6(i)] proves that the map

$$(E^{\mathrm{sm}})^0 \to \mathrm{Pic}^0_{E/S} \quad \text{defined by} \quad t \mapsto \mathscr{O}_E(t) \otimes \mathscr{O}_E(e)^{-1}$$

is an open immersion (where $e \in E(S)$ denotes the identity section), and the representability of $\mathrm{Pic}^0_{E/S}$ by a scheme follows from [BLR90 1990, 6.6/2(b)] applied to $\mathrm{Pic}^0_{E/S}$ acting on itself by translation (see also Remark 2.1.16).

with μ_n and $\mathbb{Z}/2\mathbb{Z}$ acting as described above, so we need to argue that $\mathbb{Z}/2\mathbb{Z}$ is central in $\mathrm{Aut}(E)$. For this, due to the $\mathbb{Z}$-universal schematic density of E^{sm} in E supplied by [EGA IV$_3$ 1966, 11.10.10], it suffices to note that every generalized elliptic curve automorphism of a base change of E must commute with inversion on E^{sm}. □

We turn to the closed subschemes $E^{\mathrm{sing}} \subset E$ and $S^{\infty,\pi} \subset S$ that measure the degeneration of E.

Definition 2.1.7. The *subscheme of nonsmoothness* of a generalized elliptic curve $E \xrightarrow{\pi} S$ is the closed subscheme $E^{\mathrm{sing}} \subset E$ defined by the first Fitting ideal sheaf $\mathrm{Fitt}_1(\Omega^1_{E/S}) \subset \mathscr{O}_E$. The *degeneracy locus* of $E \xrightarrow{\pi} S$ is the schematic image $S^{\infty,\pi} \subset S$ of E^{sing}.

Remark 2.1.8. The closed subscheme E^{sing} is supported at those points of E at which π is not smooth and its formation commutes with arbitrary base change in S, see [SGA 7$_{\mathrm{I}}$ 1972, VI, 5.3 and 5.4]. Even though the formation of schematic images often does not commute with nonflat base change, the formation of $S^{\infty,\pi}$ does commute with arbitrary base change, see [Conrad 2007, 2.1.12].

Remark 2.1.9. By [Deligne and Rapoport 1973, II.1.15], we have

$$S^{\infty,\pi} = \textstyle\bigsqcup_{n\geq 1} S^{\infty,\pi,n}$$

for closed subschemes $S^{\infty,\pi,n} \subset S$ such that only finitely many of the $S^{\infty,\pi,n}$ meet a given affine open of S and such that $E_{S^{\infty,\pi,n}}$ is fppf locally on $S^{\infty,\pi,n}$ isomorphic to the standard n-gon (which was discussed in Remark 2.1.5). In particular, every generalized elliptic curve $E \xrightarrow{\pi} S$ is, Zariski locally on S, projective because, by [Deligne and Rapoport 1973, II.1.20; Katz and Mazur 1985, 1.2.3], over the open

$$S - \textstyle\bigsqcup_{n\neq n'} S^{\infty,\pi,n}$$

the n'-torsion subscheme $E^{\mathrm{sm}}[n'] \subset E$ is a π-ample relative effective Cartier divisor.

We record a basic relationship between E^{sing} and its schematic image $S^{\infty,\pi}$ in the following lemma:

Lemma 2.1.10. *For a generalized elliptic curve $E \to S$, the map*

$$E^{\mathrm{sing}} \to S^{\infty,\pi}$$

is finite étale; *it has degree n over $S^{\infty,\pi,n}$.*

Proof. The map in question exists by the definition of $S^{\infty,\pi}$ and its formation commutes with base change in S by Remark 2.1.8. We may therefore assume that $S = S^{\infty,\pi,n}$ and that E is the standard n-gon. But in this case E^{sing} is a disjoint union of n copies of S and there is nothing to prove. □

Degenerate generalized curves possess canonical finite subgroups of multiplicative type and their torsion subgroups are amenable to scrutiny. We make this precise in the following lemma:

Lemma 2.1.11. *For every generalized elliptic curve $E \xrightarrow{\pi} S$ with $S^{\mathrm{red}} = (S^{\infty,\pi})^{\mathrm{red}}$ and every $d \in \mathbb{Z}_{\geq 1}$, the d-torsion $(E^{\mathrm{sm}})^0[d]$ is a finite locally free S-group scheme of order d that is étale locally on S isomorphic to μ_d. The S-group scheme*

$$E^{\mathrm{sm}}[d]/(E^{\mathrm{sm}})^0[d]$$

is étale and if $m \in \mathbb{Z}_{\geq 1}$ divides both d and the number of irreducible components of each geometric fiber of E, then $(E^{\mathrm{sm}}[d]/(E^{\mathrm{sm}})^0[d])[m]$ is étale locally on S isomorphic to $\mathbb{Z}/m\mathbb{Z}$.

Proof. Due to the fibral criterion for flatness [EGA IV$_3$ 1966, 11.3.11], the quasifinite, finitely presented, separated S-groups $(E^{\mathrm{sm}})^0[d]$ and $E^{\mathrm{sm}}[d]$ are flat. The fibers of $(E^{\mathrm{sm}})^0[d] \to S$ have degree d, so, due to [Deligne and Rapoport 1973, II.1.19], the S-group $(E^{\mathrm{sm}})^0[d]$ is finite locally free of rank d. Due to [Conrad 2014, B.4.1 and B.3.4], the claim about the étale local structure of $(E^{\mathrm{sm}})^0[d]$ reduces to case of geometric fibers.

Thanks to the settled claims about $(E^{\mathrm{sm}})^0[d]$, [EGA IV$_3$ 1966, 8.11.2] and [SGA 3$_{\mathrm{I\,(new)}}$ 2011, V, 4.1] imply that $E^{\mathrm{sm}}[d]/(E^{\mathrm{sm}})^0[d]$ is a separated, quasifinite, finitely presented, flat S-scheme. By inspecting geometric fibers we see that $E^{\mathrm{sm}}[d]/(E^{\mathrm{sm}})^0[d]$ is étale. The étale local structure of

$$(E^{\mathrm{sm}}[d]/(E^{\mathrm{sm}})^0[d])[m]$$

may be seen over the strict Henselizations of S, and hence even on geometric fibers. □

The focus of Chapter 2 is generalized elliptic curve homomorphisms. We recall their definition.

Definition 2.1.12. A *homomorphism* between generalized elliptic curves $E \to S$ and $E' \to S$ is an S-morphism

$$f : E \to E' \quad \text{with} \quad f(E^{\mathrm{sm}}) \subset E'^{\mathrm{sm}}$$

that intertwines the group laws of E^{sm} and E'^{sm}. Its *kernel* is the S-subscheme $\operatorname{Ker} f := E \times_{f,\, E',\, e'} S$ of E, where $\times_{f,\, E',\, e'}$ denotes the base change along f of the identity section $e' : S \to E'$.

Remark 2.1.13. Due to the S-universal schematic density of E^{sm} in E supplied by [EGA IV$_3$ 1966, 11.10.10] and the separatedness of $E' \to S$, a homomorphism f necessarily also intertwines the group actions $E^{\mathrm{sm}} \times E \to E$ and $E'^{\mathrm{sm}} \times E' \to E'$.

Remark 2.1.14. If a homomorphism f is surjective, then $f|_{E^{\mathrm{sm}}}$ is flat and $\operatorname{Ker} f$ is contained in E^{sm}, as may be checked on geometric fibers using the fibral criterion for flatness [EGA IV$_3$ 1966, 11.3.11]. In this case, $\operatorname{Ker} f$ is a finite locally free S-subgroup scheme of E^{sm}.

Example 2.1.15. The constant morphism that factors through e' is a homomorphism, the "zero homomorphism." Any elliptic curve isogeny is also a homomorphism. For a $d \in \mathbb{Z}_{\geq 1}$, the map

$$\mathbb{P}^1_S \to \mathbb{P}^1_S \quad \text{given on homogeneous coordinates by} \quad [x:y] \mapsto [x^d : y^d]$$

respects 0 and ∞, so it induces an S-morphism from the standard 1-gon over S to itself. This morphism restricts to the d-th power map on the $(\mathbb{G}_m)_S$ of the smooth locus of the 1-gon, so it is a homomorphism with kernel $(\mu_d)_S$.

Remark 2.1.16. Generalized elliptic curves are susceptible to limit arguments that reduce to a Noetherian base. More precisely, by [EGA IV$_2$ 1965, 8.8.2(ii), 8.10.5(xii), 11.2.6(ii)], Zariski locally on S, the underlying relative curve $E \to S$ is the base change of a proper and flat relative curve $E_0 \to S_0$ for which S_0 is of finite type over $\mathbb{Z}$. Thus, since the formation of E_0^{sm} commutes with base change, E^{sm} is necessarily of finite presentation. Moreover, by [EGA IV$_2$ 1965, 8.8.2(i)], after enlarging S_0, the commutative S-group action

$$E^{\mathrm{sm}} \times_S E \xrightarrow{+} E \quad \text{descends to a commutative } S_0\text{-group action} \quad E_0^{\mathrm{sm}} \times_{S_0} E_0 \xrightarrow{+} E_0.$$

The degenerate geometric fibers of $E_0 \to S_0$ are Néron n-gons: indeed, [Deligne and Rapoport 1973, II.1.3] applies because the condition of having only ordinary double points as singularities is equivalent to the unramifiedness of E_0^{sing}, whose formation commutes with base change (see Remark 2.1.8), whereas the triviality of the relative dualizing sheaf may be descended from an overfield using specialization techniques. Using Remark 2.1.4 to infer the triviality of the induced action of E_0^{sm} on $\operatorname{Pic}^0_{E_0/S_0}$, we conclude that $E_0 \to S_0$ is a generalized elliptic curve that descends $E \to S$ to a Noetherian base. Similarly, Zariski locally on S, elliptic curve homomorphisms are defined over a base that is of finite type over $\mathbb{Z}$.

By the limit arguments above, the open immersion $S - S^{\infty,\pi} \hookrightarrow S$ is always quasicompact.

2.2. Quotients of generalized elliptic curves by finite locally free subgroups

Even though homomorphisms between generalized elliptic curves are useful in practice, their structural properties are not immediately apparent. Moreover, guided by the theory of isogenies of elliptic curves, one suspects that for any finite locally free S-subgroup scheme $G \subset E^{\mathrm{sm}}$ with $E \to S$ a generalized elliptic curve, there should be an essentially unique homomorphism $E \to E'$ with kernel G. If G

intersects the identity components of the degenerate geometric fibers of $E \to S$ trivially, then the translation action of G on E is free, the fppf sheaf quotient E/G is a generalized elliptic curve, and

$$E \to E/G$$

is the sought “isogeny.” This special case is already useful — for instance, such isogenies are discussed in [Conrad 2007, 2.1.6] and exploited in several key proofs of [op. cit.].

The goal of this section is to explain how to make sense of isogenies of generalized elliptic curves in general. Theorem 2.2.4 and its proof explain how to build the desired “quotient by G” homomorphism $E \to E/G$, and we arrive at the concept of an isogeny in Definition 2.2.8. With Theorem 2.2.4 in hand, structural properties of arbitrary homomorphisms are susceptible to scrutiny and are detailed in Propositions 2.2.9 and 2.2.10.

We begin with an example that illustrates what E/G should be in a certain degenerate situation.

Example 2.2.1. Let E be the standard n-gon over $\mathbb{Z}$, and consider the subgroup $\mu_d \subset (E^{\mathrm{sm}})^0$ for some $d \in \mathbb{Z}_{\geq 1}$. We would like to build a generalized elliptic curve homomorphism

$$f_d : E \to E' \quad \text{with kernel } \mu_d.$$

By Remark 2.1.13, any such f_d is μ_d-equivariant, so it factors through the categorical quotient E/μ_d, which exists because E is projective and μ_d is finite. We claim that

$$E \to E/\mu_d$$

is already the desired $f_d : E \to E'$.

This claim follows from the description of E recalled in Remark 2.1.2. More precisely, if $n \geq 2$, then on $\operatorname{Spec}(\mathbb{Z}[X_i, Y_i]/(X_i Y_i))$ the action of

$$\mu_d = \operatorname{Spec}(\mathbb{Z}[T]/(T^d - 1))$$

is determined by

$$X_i \mapsto X_i \otimes T \quad \text{and} \quad Y_i \mapsto Y_i \otimes T,$$

so the ring of invariants is the $\mathbb{Z}$-subalgebra of $\mathbb{Z}[X_i, Y_i]/(X_i Y_i)$ generated by X_i^d and Y_i^d, and hence E/μ_d is the standard n-gon with the quotient map $E \to E/\mu_d$ induced by the d-th power map on each $\mathbb{P}^1_{\mathbb{Z}}$. The same description holds if $n = 1$, as the same computation performed $\mathbb{Z}/m\mathbb{Z}$-equivariantly on the m-gon cover for some $m \geq 2$ proves. Thus, the map $E \to E/\mu_d$ is a homomorphism whose kernel is μ_d, and it is initial among such homomorphisms, so it is the desired f_d.

Remark 2.2.2. Example 2.2.1 may be carried out over any base scheme S, which shows that the formation of f_d commutes with arbitrary base change. In particular, the formation of the categorical quotient E/μ_d commutes with arbitrary (possibly nonflat) base change.

Remark 2.2.3. For $d > 1$, the "isogeny" $E \to E/\mu_d$ constructed in Example 2.2.1 is not flat at the singular points, as the formal criterion for flatness [Bourbaki 1965, III, §5, n° 2, Theorem 1] reveals. In contrast, every isogeny between elliptic curves is flat.

Example 2.2.1 suggests that over an arbitrary base S, the desired quotient of a generalized elliptic curve $E \to S$ by a finite locally free S-subgroup $G \subset E^{\mathrm{sm}}$ may simply be the categorical quotient E/G. In Theorem 2.2.4 we prove that this indeed the case. The main issue that needs to be addressed is that the formation of categorical quotients does not in general commute with nonflat base change (as in the special case of forming the ring of invariants under the action of a finite group). Such phenomena do not occur for generalized elliptic curves because the analysis of E/G may be reduced to the cases when G is either diagonalizable or acts freely on E.

Theorem 2.2.4. *Let S be a scheme, $E \xrightarrow{\pi} S$ a generalized elliptic curve, and $G \subset E^{\mathrm{sm}}$ an S-subgroup scheme that is finite locally free over S. There is an S-scheme morphism*

$$q : E \to E/G$$

that is initial among G-equivariant S-morphisms from E to an S-scheme equipped with the trivial G-action (E is equipped with the translation action of G). *Moreover, q has the following properties.*

(i) *The formation of q commutes with arbitrary base change in S, and E/G is S-flat.*

(ii) *The map $q : E \to E/G$ is surjective, finite, and universally open.*

(iii) *There is a unique structure of a generalized elliptic curve on*

$$E/G \to S$$

for which q is a homomorphism. For this structure, q induces an S-group isomorphism

$$E^{\mathrm{sm}}/G \cong (E/G)^{\mathrm{sm}},$$

where E^{sm}/G is the fppf sheaf quotient; in particular, $E^{\mathrm{sm}} \xrightarrow{q} (E/G)^{\mathrm{sm}}$ is finite locally free.

(iv) *If E is an elliptic curve, then $q : E \to E/G$ is an isogeny with kernel G.*

Proof. Zariski locally on S the map π is projective (see Remark 2.1.9), so every finite set of points of any π-fiber is contained in an affine open of E (see [EGA II 1961, 4.5.4]). Therefore, by [SGA $3_{\mathrm{I\,(new)}}$ 2011, V, 4.1(i)] and its proof, E is covered by G-invariant affine opens and the initial q is nothing but the categorical quotient that is glued together from the rings of invariants of such G-invariant affines; moreover, this q is automatically a quotient map on the underlying topological spaces.

Since G acts freely on E^{sm}, by [SGA $3_{\mathrm{I\,(new)}}$ 2011, V, 4.1(iv)], the open S-subscheme

$$E^{\mathrm{sm}}/G \subset E/G$$

that results from the G-invariance of E^{sm} is identified with the fppf sheaf quotient of E^{sm} by G, the map $E^{\mathrm{sm}} \xrightarrow{q} E^{\mathrm{sm}}/G$ is finite locally free, and the formation of E^{sm}/G commutes with base change.

(i) The formation of E/G commutes with flat base change, so we may first assume that S is affine and then use Remark 2.1.16 to assume that $S = \operatorname{Spec} R$ for some Noetherian R. Moreover, by the previous paragraph, the claim is clear on the elliptic curve locus, so we may replace R by its completion along the ideal $I \subset R$ that cuts out the degeneracy locus $S^{\infty,\pi} \subset S$ to assume that R is I-adically complete and separated.

For such R, the intersections

$$G_{R/I^j} \cap (E^{\mathrm{sm}}_{R/I^j})^0 \quad \text{for } j \geq 1$$

are finite locally free R/I^j-subgroup schemes of G. By Grothendieck's existence theorem [Illusie 2005, 8.4.5, 8.4.7], these subgroups algebraize to a finite locally free R-subgroup

$$H \subset G \quad \text{with } H \subset (E^{\mathrm{sm}})^0.$$

The R/I-fibers of H are of multiplicative type, so H itself is of multiplicative type. At the cost of replacing R by a finite locally free cover we may assume that H is diagonalizable.

By [SGA $3_{\mathrm{I\,(new)}}$ 2011, I, 4.7.3], any R-module M equipped with an action of a diagonalizable H is a direct sum of χ-isotypic submodules for characters χ of H, so the submodule M^H of H-invariants is of formation compatible with arbitrary base change and is R-flat if M is. In particular, the categorical quotient E/H is R-flat and of formation compatible with base change. As may be checked on geometric R-fibers, G/H acts freely on E/H, so the further quotient $E/G = (E/H)/(G/H)$ is also R-flat and of formation compatible with base change.

(ii) The surjectivity of q follows from the first paragraph of the proof. By [SGA $3_{\mathrm{I\,(new)}}$ 2011, V, 4.1(ii)], the morphism q is integral, and hence even finite because it inherits the property of being of finite type from $E \to S$. In particular, q is universally closed, so it is also universally open by [Rydh 2013, 2.4] (which applies due to the bottom of p. 636 there and [SGA $3_{\mathrm{I\,(new)}}$ 2011, V, 4.1(iii)]).

(iii) We begin by arguing that E/G possesses the S-scheme properties required in Definition 2.1.3.

Due to [Atiyah and Macdonald 1969, 7.8], the morphism $E/G \to S$ inherits finite presentation from $E \to S$ thanks to the finiteness of $E \to E/G$ (and an initial reduction to Noetherian S based on (i)). By (ii),

$$E \to E/G, \quad \text{and hence also} \quad E \times_S E \to E/G \times_S E/G,$$

is a finite surjection, so the image of $\Delta_{E/S}(E)$ in $E/G \times_S E/G$, i.e., $\Delta_{(E/G)/S}(E/G)$, is closed. In other words, the finite type morphism $E/G \to S$ inherits separatedness from $E \to S$, so it also inherits properness by [EGA II 1961, 5.4.3 (ii)]. Finally, $E/G \to S$ is flat by (i). For the fibral properties, due to (i), we may assume that S is a geometric point.

If S is a geometric point and E is an elliptic curve, then E/G is its isogenous quotient. If S is a geometric point and E is the standard N-gon, then we set

$$H := G \cap (E^{\mathrm{sm}})^0, \quad \text{so } H \cong \mu_d \text{ for some } d \geq 1.$$

By Example 2.2.1, $E \to E/H$ is a "self-isogeny" of the standard N-gon, and, by construction, G/H acts freely on E/H. Therefore, E/G, which is identified with $(E/H)/(G/H)$, is the standard n-gon with $n = N/\#(G/H)$. This analysis also shows that $q(E^{\mathrm{sm}}) = (E/G)^{\mathrm{sm}}$.

Due to the paragraph preceding the proof of (i), all that remains to be shown is that the S-group scheme structure of $(E/G)^{\mathrm{sm}} \cong E^{\mathrm{sm}}/G$ extends to a unique action of $(E/G)^{\mathrm{sm}}$ on E/G; indeed, the induced action on $\mathrm{Pic}^0_{(E/G)/S}$ will automatically be trivial due to the fibral analysis of the previous paragraph and Remark 2.1.4. The uniqueness follows from the separatedness of E/G and the universal schematic density of $(E/G)^{\mathrm{sm}}$ in E/G supplied by [EGA IV_3 1966, 11.10.10]. For the same reason, for the existence we only need to produce a morphism

$$(E/G)^{\mathrm{sm}} \times_S E/G \to E/G$$

that extends the group law of $(E/G)^{\mathrm{sm}}$ — the relevant diagrams that encode the property of being a group scheme will automatically commute. To build this morphism from the one for E, it suffices to prove that

$$E^{\mathrm{sm}}/G \times_S E/G \cong (E^{\mathrm{sm}} \times_S E)/(G \times_S G),$$

where the quotients are categorical. For this isomorphism, it suffices to form the quotient on the right in stages and to note that the formation of E^{sm}/G commutes with base change along $E \to S$ whereas the formation of E/G commutes with base change along $E^{\mathrm{sm}}/G \to S$.

(iv) By (iii), $q : E \to E/G$ is a finite locally free homomorphism between elliptic curves over S and its kernel is G, i.e., q is an isogeny with kernel G. □

Remark 2.2.5. The categorical quotient E/G may also be analyzed with the tame stack formalism of Abramovich, Olsson, and Vistoli [AOV08 2008]. For this, the key point is that the quotient stack $[E/G]$ is tame by [AOV08 2008, Theorem 3.2] because the automorphism functors of its geometric points are of multiplicative type. Then, since E/G is the coarse moduli space of $[E/G]$ (see [Conrad 2005, Theorem 3.1]), E/G is S-flat and of formation compatible with arbitrary base change by [AOV08 2008, Corollary 3.3].

2.2.6. The quotient notation. In the sequel, whenever $E \to S$ is a generalized elliptic curve and $G \subset E^{\mathrm{sm}}$ is a finite locally free S-subgroup, we write E/G for the generalized elliptic curve constructed in Theorem 2.2.4. In the following corollary, we record some further properties of this quotient construction that follow from Theorem 2.2.4 and its proof.

Corollary 2.2.7. *Let $E \to S$ (resp. $E' \to S$) be a fixed (resp. variable) generalized elliptic curve over a scheme S.*

(a) *If $G \subset E^{\mathrm{sm}}$ is finite locally free S-subgroup, then the homomorphism $E \to E/G$ is initial among homomorphisms $f : E \to E'$ with $G \subset \operatorname{Ker} f$.*

(b) *If $f : E \to E'$ is a surjective homomorphism, then $\operatorname{Ker} f$ is a finite locally free S-subgroup of E^{sm}, and $\operatorname{Ker} f$ determines f up to an isomorphism in the sense that f induces an isomorphism*

$$E/(\operatorname{Ker} f) \cong E'.$$

(c) *If $G_1 \subset G_2 \subset E^{\mathrm{sm}}$ are finite locally free S-subgroups, then*

$$(E/G_1)/(G_2/G_1) \cong E/G_2.$$

Proof. (a) The map f is G-equivariant for the trivial G-action on E', so it uniquely factors through the categorical quotient $E \to E/G$. It remains to note that the induced map $(E/G)^{\mathrm{sm}} \to (E')^{\mathrm{sm}}$ intertwines the group laws, as may be checked on the fppf cover $E^{\mathrm{sm}} \to (E/G)^{\mathrm{sm}}$.

(b) The first claim was proved in Remark 2.1.14. Due to (a), f induces a homomorphism $E/(\operatorname{Ker} f) \to E'$ that is an isomorphism on the smooth loci. Due to [EGA IV$_4$ 1967, 17.9.5] and the S-flatness of $E/(\operatorname{Ker} f)$, checking that $E/(\operatorname{Ker} f) \to E'$ is an isomorphism may be done on geometric fibers, where it follows from the fact that an endomorphism of the standard n-gon that is an automorphism on the smooth locus must be an automorphism.

(c) The claim follows from the universal property of $E \to E/G_2$ recorded in (a). $\square$

Corollary 2.2.7(b) and the analogy with elliptic curves justify the following definition:

Definition 2.2.8. An *isogeny* between generalized elliptic curves $E \to S$ and $E' \to S$ is a surjective homomorphism $f : E \to E'$ (so, by Corollary 2.2.7(b), it induces an isomorphism $E' \cong E/(\operatorname{Ker} f)$). The *degree* of an isogeny f is the locally constant function on S given by the order of $\operatorname{Ker} f$.

The principal difference with the elliptic curve case is that an isogeny between generalized elliptic curves is not necessarily flat (see Remark 2.2.3). As we explain in Proposition 2.2.9 (whose elliptic curve case is [Katz and Mazur 1985, 2.4.2]), the structure of an arbitrary homomorphism may be completely understood in terms of isogenies (in turn, by Corollary 2.2.7(b), the structure of an isogeny is completely determined by its kernel).

Proposition 2.2.9. *Every homomorphism $f : E \to E'$ between generalized elliptic curves $E \to S$ and $E' \to S$ is Zariski locally on S either an isogeny or the zero homomorphism.*

Proof. Limit arguments described in Remark 2.1.16 allow us to reduce to the case when S is Noetherian, so the claim follows from [MFK94 1994, Proposition 6.1], which proves that on each connected component of S the map f is either surjective (i.e., an isogeny) or the zero homomorphism. $\square$

Due to Proposition 2.2.9, the following result describes how homomorphisms interact with the degeneracy loci of Definition 2.1.7 and the subschemes of nonsmoothness:

Proposition 2.2.10. *If $f : E \to E'$ is an isogeny between generalized elliptic curves $E \xrightarrow{\pi} S$ and $E' \xrightarrow{\pi'} S$, then $f|_{E^{\mathrm{sing}}}$ factors through E'^{sing} and $S^{\infty,\pi} \subset S^{\infty,\pi'}$.*

Proof. The second claim follows from the first because $S^{\infty,\pi}$ (resp. $S^{\infty,\pi'}$) is the schematic image of $E^{\mathrm{sing}} \to S$ (resp. of $E'^{\mathrm{sing}} \to S$). Moreover, since the formation of all the subschemes in question commutes with base change in S (see Remark 2.1.8), we may use Remark 2.1.9 to assume that $S = S^{\infty,\pi,n}$ and that E is the standard n-gon.

The intersection G of $\operatorname{Ker} f$ with the relative identity component $(E^{\mathrm{sm}})^0 = \mathbb{G}_m$ is a finite locally free S-subgroup scheme of both $\operatorname{Ker} f$ and $\mathbb{G}_m$. By parts (b) and (c) of Corollary 2.2.7, f is identified with the composite

$$E \to E/G \to (E/G)/((\operatorname{Ker} f)/G)$$

of isogenies. Therefore, since the assertion about $f|_{E^{\mathrm{sing}}}$ is compatible with composition, it suffices to treat the cases $G = \operatorname{Ker} f$ and $G = 0$ separately.

Since $\mathbb{G}_m$ has a unique finite locally free S-subgroup of a given order, Zariski locally on S we have $G = \mu_d$ for some $d \in \mathbb{Z}_{\geq 1}$. Thus, if $G = \operatorname{Ker} f$, then we may assume that f is the f_d described in Example 2.2.1 (see also Remark 2.2.2). For this f_d, the claim is clear:

$$E^{\text{sing}} \quad \text{is identified with} \quad \bigsqcup_{\mathbb{Z}/n\mathbb{Z}} S \text{ used in Definition 2.1.1}$$

and f_d is induced by the d-th power map on every $\mathbb{P}^1_S$ so maps E^{sing} to itself.

If $G = 0$, then f is étale, so that $\Omega^1_{E/S} \cong f^*\Omega^1_{E'/S}$. By [SGA 7_{I} 1972, VI, 5.1(a)], the formation of the closed subscheme cut out by a Fitting ideal of a finite type quasicoherent module commutes with pullback to another scheme, so this relation between the sheaves of differentials gives $E^{\text{sing}} = f^{-1}(E'^{\text{sing}})$. □

The inclusion $S^{\infty,\pi} \subset S^{\infty,\pi'}$ of Proposition 2.2.10 may be sharpened to a precise relation between the corresponding ideal sheaves. We record this in Proposition 2.2.11 and Remark 2.2.12.

Proposition 2.2.11. *If $f : E \to E'$ is an isogeny between generalized elliptic curves and if there is a $d \in \mathbb{Z}_{\geq 1}$ such that for every degenerate geometric fiber $E_{\bar{s}}$ the intersection* $(\operatorname{Ker} f)_{\bar{s}} \cap (E^{\text{sm}}_{\bar{s}})^0$ *has rank d, then the ideal sheaves in $\mathscr{O}_S$ of the degeneracy loci $S^{\infty,\pi}$ and $S^{\infty,\pi'}$ of E and E' are related by*

$$\mathscr{I}_{S^{\infty,\pi'}} = \mathscr{I}^d_{S^{\infty,\pi}}.$$

Remark 2.2.12. For any f, Zariski locally on S there exists a required d. In order to prove this, we may assume that $S = S^{\infty,\pi}$ and may work fppf locally on S, so Remark 2.1.9 reduces to the case when E is the standard n-gon. In this case $\operatorname{Ker} f \cap (E^{\text{sm}})^0$ is an open and closed S-subgroup of $\operatorname{Ker} f$, and the claim follows from the local constancy of its rank over S.

Proof of Proposition 2.2.11. It suffices to treat the case when $S = \operatorname{Spec} R$ for some Artinian local ring $(R, \mathfrak{m})$ that has a separably closed residue field $R/\mathfrak{m}$. The elliptic curve case is clear, so we assume that $E_{R/\mathfrak{m}}$ is degenerate. Moreover, by Corollary 2.2.7(c), quotients may be taken in stages, so we assume that either

$$\operatorname{Ker} f \subset (E^{\text{sm}})^0 \quad \text{or} \quad \operatorname{Ker} f \cap (E^{\text{sm}})^0 = 0.$$

We begin with the case $\operatorname{Ker} f \cap (E^{\text{sm}})^0 = 0$, when f is finite étale of rank $\#(\operatorname{Ker} f)$, so that $E^{\text{sing}} = f^{-1}(E'^{\text{sing}})$ by [SGA 7_{I} 1972, VI, 5.1(a)]. Lemma 2.1.10 then gives the desired $S^{\infty,\pi} = S^{\infty,\pi'}$.

In the remaining case when $\operatorname{Ker} f \subset (E^{\text{sm}})^0$, we first replace S by a flat cover to be able to assume that there is a finite étale S-subgroup $G \subset E^{\text{sm}}$ such that $G_{R/\mathfrak{m}}$ maps isomorphically to the component group of $E^{\text{sm}}_{R/\mathfrak{m}}$. Due to the settled $\operatorname{Ker} f \cap (E^{\text{sm}})^0 = 0$ case, passage to E/G and $E'/f(G)$ does not affect the degeneracy loci. Therefore, we may replace

$$E \text{ by } E/G \quad \text{and} \quad E' \text{ by } E'/f(G)$$

to reduce to the case when E is irreducible.

In this situation, since S is Artinian local and strictly Henselian, [Deligne and Rapoport 1973, VII.2.1] ensures that E is a base change of the Tate curve

$$\underline{\mathrm{Tate}}_1 \to \operatorname{Spec} \mathbb{Z}[\![q]\!]$$

[loc. cit.] proves that $\underline{\mathrm{Tate}}_1$ realizes $\operatorname{Spec} \mathbb{Z}[\![q]\!]$ as an étale double cover of the formal completion of $\overline{\mathscr{E}\ell\ell}_1$ along $\overline{\mathscr{E}\ell\ell}_1^\infty$; in the notation of [loc. cit.], $\underline{\mathrm{Tate}}_1 = \overline{\mathscr{G}}_m^q/q^{\mathbb{Z}}$). If, moreover, $\operatorname{Ker} f \subset (E^{\mathrm{sm}})^0$, then $\operatorname{Ker} f = \mu_{\#(\operatorname{Ker} f)}$ inside $(E^{\mathrm{sm}})^0$ (see Lemma 2.1.11), so that we are reduced to the case when

$$E \to S \quad \text{is} \quad \underline{\mathrm{Tate}}_1 \to \operatorname{Spec} \mathbb{Z}[\![q]\!] \qquad \text{and} \qquad \operatorname{Ker} f = \mu_d.$$

However, in this case the quotient map[2] $\underline{\mathrm{Tate}}_1 \to \underline{\mathrm{Tate}}_1/\mu_d$ is identified with the map

$$\underline{\mathrm{Tate}}_1 \to \underline{\mathrm{Tate}}_1(q^d) \quad \text{induced by “raising the coordinates to the } d\text{-th power,”}$$

as in Example 2.2.1 (compare with [Conrad 2007, 2.5.1]). It remains to recall from [Deligne and Rapoport 1973, VII.1.11] that the degeneracy locus of $\underline{\mathrm{Tate}}_1$ (resp. of $\underline{\mathrm{Tate}}_1(q^d)$) is cut out by the principal ideal $(q) \subset \mathbb{Z}[\![q]\!]$ (resp. $(q^d) \subset \mathbb{Z}[\![q]\!]$). □

Chapter 3. Compactifications of the stack of elliptic curves

Our approach to the study of level structures on generalized elliptic curves makes essential use of the tower $\{\overline{\mathscr{E}\ell\ell}_n\}_{n|n'}$ of compactifications of the stack $\mathscr{E}\ell\ell$ that parametrizes elliptic curves. The purpose of this chapter is to construct this tower and to detail its properties. We begin with the construction of the individual compactifications $\overline{\mathscr{E}\ell\ell}_n$ in Section 3.1, and proceed to expose the transition morphisms $\overline{\mathscr{E}\ell\ell}_{nm} \to \overline{\mathscr{E}\ell\ell}_n$ in Section 3.2. Section 3.3 proves that the coarse moduli space of $(\overline{\mathscr{E}\ell\ell}_n)_S$ is the “j-line” $\mathbb{P}^1_S$ for every n and every scheme S, whereas Section 3.4 uses the global structure of the stacks $\overline{\mathscr{E}\ell\ell}_n$ to algebraize formal generalized elliptic curves and their homomorphisms.

3.1. The compactification $\overline{\mathscr{E}\ell\ell}_n$ obtained by allowing n-gons for a fixed n

The goal of this section is to detail algebro-geometric properties of the $\mathbb{Z}$-stack $\overline{\mathscr{E}\ell\ell}_n$ obtained from the stack of elliptic curves $\mathscr{E}\ell\ell$ by “adjoining Néron n-gons” (see Definition 3.1.1). We prove in Theorem 3.1.6 that $\overline{\mathscr{E}\ell\ell}_n$ is a proper and smooth

[2] In the notation of [Deligne and Rapoport 1973, VII.1.10], we have $\underline{\mathrm{Tate}}_1(q^d) = \overline{\mathscr{G}}_m^{q^d}/(q^d)^{\mathbb{Z}}$ over $A = \mathbb{Z}[\![q]\!]$.

compactification of $\mathscr{E}\ell\ell$. This result has already been proved over $\mathbb{Z}[1/n]$ in [Deligne and Rapoport 1973, IV.2.2], which uses deformation-theoretic methods through its reliance on [Deligne and Rapoport 1973, III.1.2]. These methods require the number of the irreducible components of each geometric fiber of the generalized elliptic curve in question to be prime to the characteristic, so they do not seem to work without inverting n. A related difficulty is that even though the stack $\overline{\mathscr{E}\ell\ell}_n$ is algebraic, outside the elliptic curve locus it is not Deligne–Mumford in characteristics dividing n (see Theorem 3.1.6(b)), so $\overline{\mathscr{E}\ell\ell}_n$ may not possess universal deformation rings at some of its geometric points. To overcome these difficulties, we proceed indirectly by exploiting a convenient auxiliary algebraic stack $\mathscr{B}_n$ whose relationship to $\overline{\mathscr{E}\ell\ell}_n$ is described in Proposition 3.1.5.

We begin by defining the stack $\overline{\mathscr{E}\ell\ell}_n$ that we are going to study and later use.

Definition 3.1.1. For an $n \in \mathbb{Z}_{\geq 1}$, let $\overline{\mathscr{E}\ell\ell}_n$ denote the $\mathbb{Z}$-stack parametrizing those generalized elliptic curves $E \xrightarrow{\pi} S$ whose degenerate geometric fibers are n-gons. Let $\overline{\mathscr{E}\ell\ell}_n^{\infty}$ denote the closed substack of $\overline{\mathscr{E}\ell\ell}_n$ cut out by the degeneracy loci $S^{\infty,\pi}$ (defined in Definition 2.1.7).

Remark 3.1.2. The effectivity of descent data that is needed for $\overline{\mathscr{E}\ell\ell}_n$ to be a $\mathbb{Z}$-stack (for the fpqc topology) results from the S-ampleness of the relative effective Cartier divisor $E^{\mathrm{sm}}[n] \subset E$.

Remark 3.1.3. The well-definedness of the closed substack $\overline{\mathscr{E}\ell\ell}_n^{\infty}$ rests on the compatibility (recalled in Remark 2.1.8) of the formation of the degeneracy locus $S^{\infty,\pi}$ with base change.

We turn to the auxiliary stack $\mathscr{B}_n$ and to its relation to $\overline{\mathscr{E}\ell\ell}_n$.

3.1.4. The stack $\mathscr{B}_n$. Following [Deligne and Rapoport 1973, V.1.3], for an $n \in \mathbb{Z}_{\geq 1}$ we let $\mathscr{B}_n$ be the $\mathbb{Z}$-stack that, for variable schemes S, parametrizes the pairs (E, G) consisting of a generalized elliptic curve $E \to S$ whose degenerate geometric fibers are n-gons and a finite étale subgroup $G \subset E^{\mathrm{sm}}$ that is étale locally on S isomorphic to $\mathbb{Z}/n\mathbb{Z}$ and meets every irreducible component of every geometric fiber of $E \to S$. If $n = 1$, then G is the zero subgroup, so $\mathscr{B}_1 = \overline{\mathscr{E}\ell\ell}_1$.

Proposition 3.1.5. *Fix an $n \in \mathbb{Z}_{\geq 1}$.*

(a) *The $\mathbb{Z}$-stack $\mathscr{B}_n$ is Deligne–Mumford and $\mathbb{Z}$-smooth of relative dimension* 1.

(b) *The morphism*
$$\mathscr{B}_n \to \overline{\mathscr{E}\ell\ell}_n$$
that forgets G factors through the open substack $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}} \subset \overline{\mathscr{E}\ell\ell}_n$ obtained by removing the supersingular elliptic curves in characteristics dividing n. The induced morphism
$$\mathscr{B}_n \to \overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$$

is representable by schemes, separated, quasifinite, faithfully flat, and of finite presentation.

(c) *The stack* $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$ *is algebraic and* $\mathbb{Z}$*-smooth of relative dimension* 1.

Proof. (a) Both claims follow from [Deligne and Rapoport 1973, V.1.4].

(b) The morphism

$$q : \overline{\mathscr{E}\ell\ell}_n \to \overline{\mathscr{E}\ell\ell}_1 \quad \text{is well defined by} \quad q(E) = E/E^{\mathrm{sm}}[n]$$

(see Section 2.2.6), and, as in [Deligne and Rapoport 1973, VI.1.1], the j-invariant gives the morphism $j : \overline{\mathscr{E}\ell\ell}_1 \to \mathbb{P}^1_{\mathbb{Z}}$. Since $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$ is the preimage under $j \circ q$ of the open subscheme of $\mathbb{P}^1_{\mathbb{Z}}$ obtained by removing the supersingular j-invariants in characteristics dividing n, it is indeed an open substack of $\overline{\mathscr{E}\ell\ell}_n$.

The morphism $\mathscr{B}_n \to \overline{\mathscr{E}\ell\ell}_n$ factors through $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$ because a supersingular elliptic curve over an algebraically closed field of positive characteristic p cannot have $\mathbb{Z}/p\mathbb{Z}$ as a subgroup. Therefore, our task is to prove that for any generalized elliptic curve $E \to S$ whose geometric fibers are n-gons, ordinary elliptic curves in characteristic dividing n, or arbitrary elliptic curves in characteristic not dividing n, the functor

$$F_0 : S' \mapsto \big\{S'\text{-ample subgroups } G \subset E^{\mathrm{sm}}_{S'} \text{ that are étale locally on } S' \text{ isomorphic to } \mathbb{Z}/n\mathbb{Z}\big\}$$

on the category of S-schemes is representable by a separated, quasifinite, faithfully flat S-scheme B of finite presentation (the S'-ampleness of G as a relative effective Cartier divisor on $E_{S'}$ is equivalent to the condition that G meets every irreducible component of every geometric fiber of $E_{S'} \to S'$). In fact, it suffices to prove the same statement with “faithfully flat” replaced by “flat” and for the functor F'_0 obtained by dropping the S'-ampleness requirement from the definition of F_0: indeed, the surjectivity of $B \to S$ will follow from the imposed fibral assumptions on $E \to S$, whereas [EGA IV$_3$ 1966, 9.6.4] together with limit arguments ensures that the inclusion $F_0 \subset F'_0$ is representable by quasicompact open immersions.

We analyze F'_0 by studying the related functor

$$F_1 : S' \mapsto \big\{P \in E^{\mathrm{sm}}[n](S') \text{ that define a closed immersion } \mathbb{Z}/n\mathbb{Z} \hookrightarrow E^{\mathrm{sm}}_{S'}[n] \text{ by } 1 \mapsto P\big\}.$$

The map $F_1 \to F'_0$ that sends P to the copy of $\mathbb{Z}/n\mathbb{Z}$ that P generates is representable by schemes and finite étale of rank $\phi(n)$. Therefore, once we prove that F_1 is representable by a finitely presented, separated, quasifinite (and hence also quasiaffine, see [EGA IV$_3$ 1966, 8.11.2]), flat S-scheme, the desired claim about F'_0 will follow from [SGA 3$_{\mathrm{I\,(new)}}$ 2011, V, 4.1] (combined with [EGA IV$_2$ 1965, 2.2.11(iii); EGA IV$_4$ 1967, 17.7.5]).

The S-scheme $E^{\mathrm{sm}}[n]$ represents the functor of S'-homomorphisms

$$\mathbb{Z}/n\mathbb{Z} \to E^{\mathrm{sm}}_{S'}[n].$$

Such a homomorphism is a closed immersion if and only if its corresponding map f of finite locally free $\mathscr{O}_{S'}$-algebras is surjective, which is an open condition on S' because $\mathrm{Coker}(f)$ is a finitely generated $\mathscr{O}_{S'}$-module. Therefore, the inclusion $F_1 \subset E^{\mathrm{sm}}[n]$ is representable by open immersions, and is quasicompact by limit arguments, so the claims about F_1 follow.

(c) Both claims follow from (b). More precisely, if $X \to \mathscr{B}_n$ is a smooth atlas, then the composed morphism

$$X \to \overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$$

is representable by algebraic spaces, faithfully flat, and locally of finite presentation, so $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$ is algebraic by [SP 2005–, 06DC] (see also [Laumon and Moret-Bailly 2000, 10.6] for a related result), whereas, due to [EGA IV$_4$ 1967, 17.7.7], the $\mathbb{Z}$-smoothness of $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$ follows from that of $\mathscr{B}_n$ (for the relative dimension aspect, one may use [EGA IV$_2$ 1965, 6.1.2]). □

With Proposition 3.1.5 in hand, we are ready to address algebro-geometric properties of $\overline{\mathscr{E}\ell\ell}_n$ (see Proposition 3.3.2 for some further properties).

Theorem 3.1.6. *Fix an $n \in \mathbb{Z}_{\geq 1}$.*

(a) *The $\mathbb{Z}$-stack $\overline{\mathscr{E}\ell\ell}_n$ is algebraic with finite diagonal, proper, and smooth of relative dimension* 1.

(b) *The largest open substack of $\overline{\mathscr{E}\ell\ell}_n$ that is Deligne–Mumford is*

$$\overline{\mathscr{E}\ell\ell}_n - (\overline{\mathscr{E}\ell\ell}_n^{\infty})_{\mathbb{Z}/n\mathbb{Z}}.$$

(c) *The morphism* $\operatorname{Spec}\mathbb{Z} \to \overline{\mathscr{E}\ell\ell}_n^{\infty}$ *that corresponds to the standard n-gon is surjective, representable, and finite locally free of rank $2n$. In particular, the proper $\mathbb{Z}$-algebraic stack $\overline{\mathscr{E}\ell\ell}_n^{\infty}$ is irreducible, has geometrically irreducible $\mathbb{Z}$-fibers, and is $\mathbb{Z}$-smooth of relative dimension* 0.

(d) *The closed substack $\overline{\mathscr{E}\ell\ell}_n^{\infty} \subset \overline{\mathscr{E}\ell\ell}_n$ is a reduced relative effective Cartier divisor over* $\operatorname{Spec}\mathbb{Z}$.

Remark 3.1.7. In (b), the largest Deligne–Mumford open substack of the separated $\mathbb{Z}$-algebraic stack $\overline{\mathscr{E}\ell\ell}_n$ does make sense *a priori*. Indeed, the proof of [Conrad 2007, 2.2.5(2)] shows that if S is a scheme and $\mathscr{X}$ is an S-algebraic stack that is covered by S-separated open substacks, then there is a unique open substack

$$\mathscr{U} \subset \mathscr{X}$$

containing exactly those geometric points of $\mathscr{X}$ that have an unramified automorphism functor. (Equivalently, $\mathscr{U}$ contains those S-scheme valued points of $\mathscr{X}$ whose automorphism functors are unramified.) By Nakayama's lemma (or simply by [SP 2005–, 02GF (1)$\Leftrightarrow$(2)]), the diagonal $\Delta_{\mathscr{U}/S}$ is unramified, so $\mathscr{U}$ is Deligne–Mumford, and, by construction, $\mathscr{U}$ contains every Deligne–Mumford open substack of $\mathscr{X}$. Even though we take the unramifiedness of the diagonal as our definition of being Deligne–Mumford (see Section 1.9), in the case in hand $\mathscr{U}$ inherits separatedness from $\overline{\mathscr{E}\ell\ell}_n$, so, by [Laumon and Moret-Bailly 2000, 8.1], it also satisfies the étale atlas definition of a Deligne–Mumford stack.

Proof of Theorem 3.1.6. (a) The stack $\overline{\mathscr{E}\ell\ell}_n$ is a union of open substacks $\mathscr{E}\ell\ell$ and $\overline{\mathscr{E}\ell\ell}_n^{n\text{-}\mathrm{ord}}$, both of which are algebraic and $\mathbb{Z}$-smooth of relative dimension 1 by Proposition 3.1.5. Therefore, $\overline{\mathscr{E}\ell\ell}_n$ is also algebraic and $\mathbb{Z}$-smooth of relative dimension 1.

By [Conrad 2007, 3.2.4], the isomorphism functor of two generalized elliptic curves $E \to S$ and $E' \to S$ whose degenerate geometric fibers are n-gons is representable by a finite S-scheme,[3] so $\Delta_{\overline{\mathscr{E}\ell\ell}_n/\mathbb{Z}}$ is finite and, in particular, $\overline{\mathscr{E}\ell\ell}_n$ is separated. The morphism

$$\mathscr{E}\ell\ell \sqcup \operatorname{Spec}\mathbb{Z} \to \overline{\mathscr{E}\ell\ell}_n$$

whose restriction to $\operatorname{Spec}\mathbb{Z}$ corresponds to the standard n-gon is surjective on underlying topological spaces, so $\overline{\mathscr{E}\ell\ell}_n$ is quasicompact, and hence is of finite type over $\mathbb{Z}$. Its properness therefore results from the valuative criterion [Laumon and Moret-Bailly 2000, 7.10], which is satisfied due to the semistable reduction theorem for elliptic curves (and the availability of contractions, which are reviewed in Section 3.2.1).

(b) In the view of Remark 3.1.7, we only need to show that

$$\overline{\mathscr{E}\ell\ell}_n - (\overline{\mathscr{E}\ell\ell}_n^{\infty})_{\mathbb{Z}/n\mathbb{Z}}$$

contains those geometric points x of $\overline{\mathscr{E}\ell\ell}_n$ whose automorphism functor is unramified. If x lies in $\mathscr{E}\ell\ell = \overline{\mathscr{E}\ell\ell}_n - \overline{\mathscr{E}\ell\ell}_n^{\infty}$, then $\mathrm{Aut}(x)$ is unramified by [Deligne 1975, 5.3(I)] (or by [MFK94 1994, Corollary 6.2]). If x lies in $\overline{\mathscr{E}\ell\ell}_n^{\infty}$, then, by

[3] Here is a sketch for a proof of this representability that bypasses blowups used in [Conrad 2007, 3.2.2 and 3.2.4]: as in the proof of [Deligne and Rapoport 1973, III.2.5], one uses Hilbert schemes to get representability by a quasifinite, separated S-scheme; then, due to the valuative criterion, the key point is to check that if S is the spectrum of a strictly Henselian discrete valuation ring and E and E' are degenerating elliptic curves with identified generic fibers: $E_\eta = E'_\eta$, then $E = E'$; for this, the theory of Néron models (especially, [BLR90 1990, 7.4/3]) identifies $(E^{\mathrm{sm}})^0$ with $(E'^{\mathrm{sm}})^0$ and, since the reductions of η-rational points are dense in the special fibers, also E^{sm} with E'^{sm}; then Zariski's main theorem [BLR90 1990, 2.3/2$'$] produces the graph of the sought identification $E = E'$.

Lemma 2.1.6, $\mathrm{Aut}(x)$ is unramified if and only if x lies in

$$\overline{\mathscr{E}\ell\ell}_n^\infty - (\overline{\mathscr{E}\ell\ell}_n^\infty)_{\mathbb{Z}/n\mathbb{Z}}.$$

(c) For the asserted properties of the morphism, it suffices to note that for a generalized elliptic curve $E \xrightarrow{\pi} S$ with $S^{\infty,\pi,n} = S$, the functor of isomorphisms between E and the standard n-gon is representable by a finite locally free S-scheme of rank $2n$, as may be checked fppf locally on S with the help of Remark 2.1.9 and Lemma 2.1.6. The asserted properties of $\overline{\mathscr{E}\ell\ell}_n^\infty$ then follow by using [EGA IV$_4$ 1967, 17.7.7; EGA IV$_2$ 1965, 6.1.2] for the smoothness aspect.

(d) By (c), the stack $\overline{\mathscr{E}\ell\ell}_n^\infty$ is $\mathbb{Z}$-smooth, so it is also reduced. For the Cartier divisor claim, we may work over a smooth finite type scheme cover

$$X \to \overline{\mathscr{E}\ell\ell}_n, \quad \text{with } X^\infty \subset X \text{ being the preimage of } \overline{\mathscr{E}\ell\ell}_n^\infty.$$

By [Katz and Mazur 1985, 1.1.5.2], we may also base change from $\mathbb{Z}$ to an algebraically closed field. Then, for a point $x \in X^\infty$, by (a) and (c), both X and X^∞ are smooth at x and

$$\dim_x X^\infty = \dim_x X - 1.$$

Thus, $X^\infty \subset X$ is a Weil divisor and, since X is regular, also a desired Cartier divisor. $\square$

For later use we record the following proposition from [Conrad 2007, 3.2.4].

Proposition 3.1.8. *Let $E \xrightarrow{\pi} S$ and $E' \xrightarrow{\pi'} S$ be generalized elliptic curves such that*

$$S^{\infty,\pi,n} \cap S^{\infty,\pi',m} = \varnothing \quad \textit{whenever } n \neq m.$$

(a) *The fppf sheaf* $\mathrm{Isom}(E, E')$ *that parametrizes generalized elliptic curve isomorphisms is representable by a finite S-scheme of finite presentation.*

(b) *If S is integral and normal and η is its generic point, then any η-isomorphism*

$$E_\eta \simeq E'_\eta \quad \textit{extends to a unique } S\textit{-isomorphism} \quad E \simeq E'.$$

Proof. Part (a) has essentially been proved in footnote 3. Alternatively, Zariski locally on S there is an $n \in \mathbb{Z}_{\geq 1}$ such that E and E' correspond to objects of $\overline{\mathscr{E}\ell\ell}_n$, so (a) is a reformulation of the finiteness of the diagonal of $\overline{\mathscr{E}\ell\ell}_n$ proved in Theorem 3.1.6(a). To obtain (b) one combines (a) with the following useful lemma. $\square$

Lemma 3.1.9. *If S is an integral normal scheme, η is its generic point, and F is a finite S-scheme, then the pullback map $F(S) \to F(\eta)$ is bijective.*

Proof. The injectivity follows from the schematic dominance of $\eta \to S$ and the separatedness of $F \to S$. For the surjectivity, we may work Zariski locally on S to assume that $S = \operatorname{Spec} A$. Then the schematic image in F of an $x \in F(\eta)$ is $\operatorname{Spec} B$ for some finite A-subalgebra $B \subset \operatorname{Frac} A$. Since A is normal, $A = B$, so the schematic image is the sought extension of x to an element of $F(S)$. $\square$

3.2. The tower of compactifications

The compactifications $\overline{\mathscr{E}\ell\ell}_n$ introduced in the previous section are related to each other: they form an infinite tower in which the transition morphisms

$$\overline{\mathscr{E}\ell\ell}_{nm} \to \overline{\mathscr{E}\ell\ell}_n$$

encode contractions of generalized elliptic curves. The goal of this section is to use the already established results about $\overline{\mathscr{E}\ell\ell}_n$ to prove several basic properties, such as flatness, of these transition morphisms (see Theorem 3.2.4) and to deduce some concrete results about the generalized elliptic curves themselves (see Corollaries 3.2.5 and 3.2.6). We begin with a brief review of contractions.

3.2.1. Contraction with respect to a finite locally free subgroup. As is justified in [Conrad 2007, top of p. 218] (which is based on [Deligne and Rapoport 1973, IV.1.2]), if $E \to S$ is a generalized elliptic curve and $G \subset E^{\mathrm{sm}}$ is a finite locally free S-subgroup, then there is a generalized elliptic curve

$$c_G(E) \to S \quad \text{equipped with a surjective } S\text{-scheme morphism} \quad E \to c_G(E) \tag{3.2.1.1}$$

such that:

- the image under $E \to c_G(E)$ of each disjoint from G irreducible component of a geometric fiber of $E \to S$ is a single point, and
- the map $E \to c_G(E)$ restricts to a group isomorphism between the open complement of the union of such components and $(c_G(E))^{\mathrm{sm}}$.

In particular, if E is an elliptic curve, then $E = c_G(E)$.

These conditions ensure that G is identified with an S-subgroup of $c_G(E)^{\mathrm{sm}}$ that meets every irreducible component of every geometric fiber of $c_G(E) \to S$. Due to [Deligne and Rapoport 1973, IV.1.2], they also determine the data (3.2.1.1) uniquely up to a unique isomorphism. In particular, whenever $G' \subset E^{\mathrm{sm}}$ is another finite locally free S-subgroup that meets the same irreducible components of the geometric fibers of $E \to S$ as G, one gets a canonical identification

$$c_G(E) = c_{G'}(E). \tag{3.2.1.2}$$

For the same reason, the formation of $E \to c_G(E)$ commutes with arbitrary base change in S.

We call this $c_G(E)$ *the contraction of E with respect to G.* The compatibility of the formation of $c_G(E)$ with base change shows that for every $n, m \in \mathbb{Z}_{\geq 1}$, the identity map on $\mathscr{E}\ell\ell$ extends to the "contraction" $\mathbb{Z}$-morphism

$$\overline{\mathscr{E}\ell\ell}_{nm} \to \overline{\mathscr{E}\ell\ell}_{n} \quad \text{defined by} \quad E \mapsto c_{E^{\mathrm{sm}}[n]}(E).$$

Also, if (E, G) is classified by the stack $\mathscr{B}_{nm}$ of Section 3.1.4, then $(c_{G[n]}(E), G[n])$ is classified by the stack $\mathscr{B}_n$, so there is the "contraction" $\mathbb{Z}$-morphism

$$\mathscr{B}_{nm} \to \mathscr{B}_n \quad \text{defined by} \quad (E, G) \mapsto (c_{G[n]}(E), G[n]).$$

These and similar morphisms will be called *contractions* or *contraction morphisms* in the sequel (a slight abuse of terminology because it is not substacks of $\overline{\mathscr{E}\ell\ell}_{nm}$ or $\mathscr{B}_{nm}$ that are getting contracted).

In many situations, we will need a robust criterion for recognizing algebraic spaces and morphisms that are representable by algebraic spaces. The following lemma, which paraphrases [Conrad 2007, 2.2.5(1) and 2.2.7] and may be traced back to [Deligne and Rapoport 1973, IV.2.6], is well suited for this task.

Lemma 3.2.2. *Let S be a scheme and let $\mathscr{X}$ and $\mathscr{Y}$ be S-algebraic stacks whose diagonals $\Delta_{\mathscr{X}/S}$ and $\Delta_{\mathscr{Y}/S}$ are quasicompact and separated.*

(a) *The stack $\mathscr{X}$ is an algebraic space if and only if for every algebraically closed field $\bar{k}$ whose spectrum is equipped with a morphism to S, every object ξ of $\mathscr{X}(\bar{k})$, and every Artinian local $\bar{k}$-algebra A, the pullback of ξ to the groupoid $\mathscr{X}(A)$ has no nonidentity automorphism; if $\mathscr{X}$ is Deligne–Mumford, then $A = \bar{k}$ suffices.*

(b) *An S-morphism*

$$f : \mathscr{X} \to \mathscr{Y}$$

is representable by algebraic spaces if and only if for every algebraically closed field $\bar{k}$ whose spectrum is equipped with a morphism to S, every object ξ of $\mathscr{X}(\bar{k})$, and every Artinian local $\bar{k}$-algebra A, no nonidentity automorphism of the pullback of ξ to $\mathscr{X}(A)$ is sent to an identity automorphism in $\mathscr{Y}(A)$; if $\mathscr{X}$ is Deligne–Mumford, then $A = \bar{k}$ suffices.

Proof. (a) The necessity is clear. For the sufficiency, due to [Conrad 2007, 2.2.5(1)], it is enough to argue that the assumed condition implies the triviality of the automorphism functor of every ξ. This functor is a separated $\bar{k}$-group algebraic space G of finite type, so is necessarily a scheme due to [Artin 1969, 4.2], and is even $\bar{k}$-étale if $\mathscr{X}$ is Deligne–Mumford. The triviality of G is therefore equivalent to that of all the $G(A)$, with $A = \bar{k}$ being sufficient if $\mathscr{X}$ is Deligne–Mumford.

(b) The failure of the condition on ξ implies that the groupoid of A-points of some A-fiber of f has a nonidentity automorphism, and the necessity follows. For the sufficiency, due to [Conrad 2007, 2.2.7], it is enough to argue that the assumed condition implies that each $\bar{k}$-fiber X of f is an algebraic space, so it remains to observe that this condition ensures that X meets the criterion of (a). □

To infer further representability by schemes, we will often use the following well-known lemma:

Lemma 3.2.3. *For stacks $\mathscr{X}$ and $\mathscr{Y}$ over a scheme S, an S-morphism $f : \mathscr{X} \to \mathscr{Y}$ that is representable by algebraic spaces, separated, and locally quasifinite is representable by schemes; if, in addition, f is proper, then f is finite.*

Proof. This follows from [Laumon and Moret-Bailly 2000, A.2] (see also [Conrad 2007, 2.2.6]) and [EGA IV_4 1967, 18.12.4]. □

We are ready to exploit the relationship between the two contractions introduced in Section 3.2.1 to extract further information about the stacks $\overline{\mathscr{E}\ell\ell}_n$.

Theorem 3.2.4. *For $\mathscr{B}_n$ as in* Section 3.1.4 *and any $n, m \in \mathbb{Z}_{\geq 1}$, consider the commutative diagram*

$$\begin{array}{ccc} \mathscr{B}_{nm} & \xrightarrow{a} & \overline{\mathscr{E}\ell\ell}_{nm} \\ \downarrow{\scriptstyle c'} & & \downarrow{\scriptstyle c} \\ \mathscr{B}_n & \xrightarrow{b} & \overline{\mathscr{E}\ell\ell}_n \end{array}$$

in which c and c' are the contraction morphisms of Section 3.2.1 *and a and b forget the subgroup G.*

(a) *The contractions c and c' are flat and of finite presentation. Moreover, c is proper, with finite diagonal, and surjective, whereas c' is representable by schemes, separated, and quasifinite.*

(b) *The closed substack*

$$\overline{\mathscr{E}\ell\ell}_n^{\infty} \times_{\overline{\mathscr{E}\ell\ell}_n, c} \overline{\mathscr{E}\ell\ell}_{nm} \subset \overline{\mathscr{E}\ell\ell}_{nm}$$

is a relative effective Cartier divisor over $\operatorname{Spec} \mathbb{Z}$ *that is a positive integer multiple of $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$.*

(c) *The multiple needed in* (b) *is m, i.e.,*

$$[\overline{\mathscr{E}\ell\ell}_n^{\infty} \times_{\overline{\mathscr{E}\ell\ell}_n, c} \overline{\mathscr{E}\ell\ell}_{nm}] = m \cdot [\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}]$$

as Cartier divisors on $\overline{\mathscr{E}\ell\ell}_{nm}$.

Proof. The commutativity of the diagram follows from the identification discussed in Section 3.2.1.

By Proposition 3.1.5(b), the maps a and b are representable by schemes, separated, quasifinite, of finite presentation, flat, and faithfully flat onto $\overline{\mathscr{E}\ell\ell}_{nm}^{nm\text{-ord}}$ and $\overline{\mathscr{E}\ell\ell}_n^{n\text{-ord}}$, respectively.

(a) By Theorem 3.1.6(a), the stacks $\overline{\mathscr{E}\ell\ell}_{nm}$ and $\overline{\mathscr{E}\ell\ell}_{n}$ are $\mathbb{Z}$-proper with finite diagonal, so c is also proper, with finite diagonal, and of finite presentation. Since the contraction of the standard nm-gon with respect to its n-torsion is the standard n-gon, c is surjective. Moreover, $c|_{\mathscr{E}\ell\ell}$ is the identity, $\mathscr{E}\ell\ell$ and $\overline{\mathscr{E}\ell\ell}_{nm}^{nm\text{-ord}}$ cover $\overline{\mathscr{E}\ell\ell}_{nm}$, and, by Proposition 3.1.5(b), a is faithfully flat onto $\overline{\mathscr{E}\ell\ell}_{nm}^{nm\text{-ord}}$, so the flatness of c will follow once we establish that of c'.

It remains to establish the claims about c'. For the representability of c' by algebraic spaces, due to Lemma 3.2.2(b), it suffices to observe that if E is the standard nm-gon over an algebraically closed field and

$$G \simeq \mathbb{Z}/nm\mathbb{Z}$$

is a subgroup of E^{sm} that meets every irreducible component of E, then, by Lemma 2.1.6, no nonidentity automorphism of (E, G) restricts to the identity map on $(E^{\mathrm{sm}})^0$. The separatedness of c' follows from the separatedness of $b \circ c' = c \circ a$ and of b, and similarly for the finite presentation of c'. For the quasifiniteness of c' it therefore suffices to observe that a generalized elliptic curve over an algebraically closed field has finitely many subgroups of order nm. The representability of c' by schemes follows from Lemma 3.2.3.

Finally, since c' is a quasifinite map between separated Deligne–Mumford stacks that are smooth of relative dimension 1 over $\mathbb{Z}$, it is flat by [EGA IV$_2$ 1965, 6.1.5].

(b) Since c is flat by (a) and $\overline{\mathscr{E}\ell\ell}_{n}^{\infty} \subset \overline{\mathscr{E}\ell\ell}_{n}$ is a relative effective Cartier divisor over $\operatorname{Spec}\mathbb{Z}$ by Theorem 3.1.6(d), the pullback in question is also a relative effective Cartier divisor over $\operatorname{Spec}\mathbb{Z}$. Both

$$\overline{\mathscr{E}\ell\ell}_{n}^{\infty} \times_{\overline{\mathscr{E}\ell\ell}_{n},c} \overline{\mathscr{E}\ell\ell}_{nm} \quad \text{and} \quad \overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$$

are supported on the same closed subset of the underlying topological space of $\overline{\mathscr{E}\ell\ell}_{nm}$ and, by Theorem 3.1.6(c)–(d), this subset is irreducible and has $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$ as its associated reduced closed substack (see [Laumon and Moret-Bailly 2000, 5.6.1(ii)]). Moreover, $\overline{\mathscr{E}\ell\ell}_{nm}$ is regular, so on any of its scheme atlases Cartier divisors identify with Weil divisors. Passage to such an atlas then shows that $\overline{\mathscr{E}\ell\ell}_{n}^{\infty} \times_{\overline{\mathscr{E}\ell\ell}_{n},c} \overline{\mathscr{E}\ell\ell}_{nm}$ is a positive integer multiple of $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$ — the global constancy of the needed factor across the irreducible components of the pullback of $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$ to the atlas follows from the irreducibility of $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$ (to check that the generic points of such irreducible components map to the generic point of $\overline{\mathscr{E}\ell\ell}_{nm}^{\infty}$, one uses the fact that generizations lift along a flat morphism; see [Laumon and Moret-Bailly 2000, 5.8]).

(c) Due to (b) and the moduli interpretation, it suffices to find a single generalized elliptic curve $E \xrightarrow{\pi} S$ with nm-gon degenerate geometric fibers such that its contraction $E' \xrightarrow{\pi'} S$ with respect to $E^{\mathrm{sm}}[n]$ satisfies the equality

$$\mathscr{I}_{S^{\infty,\pi'}} = \mathscr{I}^{d}_{S^{\infty,\pi}} \quad \text{of } \mathscr{O}_S\text{-ideal sheaves for } d = m,$$

but does not satisfy this equality for any other $d \in \mathbb{Z}_{\geq 1}$ (here $\mathscr{I}_{S^{\infty,\pi}} \subset \mathscr{O}_S$ is the ideal sheaf that cuts out the degeneracy locus $S^{\infty,\pi} \subset S$, and likewise for $\mathscr{I}_{S^{\infty,\pi'}}$). Tate curves supply such E, see [Deligne and Rapoport 1973, VII.1.11 and VII.1.14]. □

We now record some concrete consequences of our analysis of the contraction

$$c : \overline{\mathscr{E}\ell\ell}_{nm} \to \overline{\mathscr{E}\ell\ell}_n.$$

Corollary 3.2.5. *For a generalized elliptic curve $E \xrightarrow{\pi} S$, let $\mathscr{I}_{S^{\infty,\pi}} \subset \mathscr{O}_S$ be the ideal sheaf that cuts out the degeneracy locus $S^{\infty,\pi} \subset S$. If the degenerate geometric fibers of $E \xrightarrow{\pi} S$ are nm-gons and $c_{E^{\mathrm{sm}}[n]}(E) \xrightarrow{\pi'} S$ is the contraction of $E \xrightarrow{\pi} S$ with respect to $E^{\mathrm{sm}}[n]$, then*

$$\mathscr{I}_{S^{\infty,\pi'}} = \mathscr{I}^m_{S^{\infty,\pi}}.$$

Proof. This is a reformulation of Theorem 3.2.4(c). □

Corollary 3.2.6. *For each $n \in \mathbb{Z}_{\geq 1}$, every generalized elliptic curve $E \to S$ is fppf locally on S the contraction* (with respect to some subgroup) *of a generalized elliptic curve $E' \to S$ for which the number of irreducible components of each degenerate geometric fiber is divisible by n. An fppf cover of S over which such an E' exists may be chosen to be locally quasifinite over S.*

Proof. We may assume that there is a $d \in \mathbb{Z}_{\geq 1}$ such that the degenerate geometric fibers of E are d-gons (see Remark 2.1.9). The first claim then follows from flatness, surjectivity, and finite presentation of $\overline{\mathscr{E}\ell\ell}_{nd} \xrightarrow{c} \overline{\mathscr{E}\ell\ell}_d$. The second claim follows from the first and [EGA IV$_4$ 1967, 17.16.2]. □

We conclude the section by using Corollary 3.2.6 to analyze the torsion subgroups of a generalized elliptic curve in a formal neighborhood of the degeneracy locus. Similar analysis in the case of Tate curves has been carried out in [Deligne and Rapoport 1973, VII.1.13–VII.1.15].

Proposition 3.2.7. *Let $E \xrightarrow{\pi} S$ be a generalized elliptic curve with $S = \operatorname{Spec} R$ for a Noetherian R that is complete and separated with respect to the ideal $I \subset R$ that cuts out $S^{\infty,\pi} \subset S$.*

(a) *For every $n \in \mathbb{Z}_{\geq 1}$, the S-group $(E^{\mathrm{sm}})^0$ has a unique finite locally free S-subgroup $B_n \subset (E^{\mathrm{sm}})^0$ of order n, and $B_n \simeq \mu_n$ étale locally on S. If an $m \in \mathbb{Z}_{\geq 1}$ divides both n and the number of irreducible components of each degenerate geometric fiber of E, then $E^{\mathrm{sm}}[n]$ has a unique finite locally free S-subgroup $A_{n,m}$ that meets precisely m irreducible components of every degenerate geometric fiber of E, contains every other finite locally free S-subgroup of $E^{\mathrm{sm}}[n]$ with this property, is of order nm, and fits into a short exact sequence*

$$0 \to B_n \to A_{n,m} \to C_m \to 0$$

with C_m isomorphic to $\mathbb{Z}/m\mathbb{Z}$ étale locally on S.

(b) *For every $n \in \mathbb{Z}_{\geq 1}$, over $S - S^{\infty,\pi}$ the group B_n from* (a) *fits into a short exact sequence*

$$0 \to (B_n)_{S-S^{\infty,\pi}} \to E_{S-S^{\infty,\pi}}[n] \to C'_n \to 0$$

with C'_n an $(S - S^{\infty,\pi})$-group scheme that is isomorphic to $\mathbb{Z}/n\mathbb{Z}$ étale locally on $S - S^{\infty,\pi}$.

Proof. (a) If S is an infinitesimal thickening of $S^{\infty,\pi}$, then Lemma 2.1.11 gives the claim. Therefore, the uniqueness and the existence of B_n and $A_{n,m}$ follow from [EGA III$_1$ 1961, 5.1.4 and 5.4.1] (the S-group structure of B_n may be read off from the action morphism $B_n \times_S E \to E$, so the nonproperness of E^{sm} does not intervene, and likewise for $A_{n,m}$). The étale local structure of B_n translates into that of its Cartier dual, so it may be read off on geometric fibers at points in $S^{\infty,\pi}$, and likewise for the étale local structure of C_m.

(b) In the case when n divides the number of irreducible components of each degenerate geometric fiber of E, the claim follows from (a). In general, C'_n is a finite locally free $(S - S^{\infty,\pi})$-group scheme of order n and it suffices to check that its geometric fibers are isomorphic to $\mathbb{Z}/n\mathbb{Z}$. In order to check this at a point $\eta \in S - S^{\infty,\pi}$, we choose a specialization $s \in S^{\infty,\pi}$ of η and use [EGA II 1961, 7.1.9] to find an S-scheme T that is the spectrum of a complete discrete valuation ring whose generic (resp. closed) point maps to η (resp. s). Due to the uniqueness of B_n, the formation of C'_n commutes with base change of E to T, so we are reduced to the case when $S = \operatorname{Spec} R$ for some complete discrete valuation ring R and $I \subset R$ is a nonzero power of the maximal ideal. In this case, Corollary 3.2.6 and [EGA IV$_4$ 1967, 18.5.11 (a)$\Leftrightarrow$(c)] supply a finite faithfully flat R-algebra R' such that $E_{R'}$ is the contraction of a generalized elliptic curve $E' \to \operatorname{Spec} R'$ for which n divides the number of irreducible components of each degenerate geometric fiber. Base change to R' reduces the claim to the settled case of E'. □

3.3. The coarse moduli space of $\overline{\mathscr{E}\ell\ell}_n$

We seek to prove in Proposition 3.3.2 that for any scheme S and any $n \in \mathbb{Z}_{\geq 1}$ the coarse moduli space of $(\overline{\mathscr{E}\ell\ell}_n)_S$ is isomorphic to $\mathbb{P}^1_S$, the "j-line." Of course, this is hardly surprising, but even in the $n = 1$ case we are not aware of a reference that would treat arbitrary S — for $n = 1$, [Deligne and Rapoport 1973, VI.1.1] settles the basic case $S = \operatorname{Spec} \mathbb{Z}$, whereas [Fulton and Olsson 2010, 2.1] handles general locally Noetherian S (the formation of the coarse moduli space need not commute with nonflat base change, so the $S = \operatorname{Spec} \mathbb{Z}$ case does not automatically imply the general case). We will build on the above result of Deligne and Rapoport through the following lemma.

The existence of all the coarse moduli spaces that we will consider in this section is guaranteed by [Keel and Mori 1997, 1.3(1)] (see also [Conrad 2005, 1.1; Rydh 2013, 6.12]).

Lemma 3.3.1. *Let $\mathscr{X}$ be a Deligne–Mumford stack that is separated, flat, and locally of finite type over $\mathbb{Z}$, and let*

$$f : \mathscr{X} \to X$$

be its coarse moduli space map. If $f_{\mathbb{F}_p} : \mathscr{X}_{\mathbb{F}_p} \to X_{\mathbb{F}_p}$ is the coarse moduli space map of $\mathscr{X}_{\mathbb{F}_p}$ for every prime p, then $f_S : \mathscr{X}_S \to X_S$ is the coarse moduli space map of $\mathscr{X}_S$ for every scheme S.

Proof. The formation of the coarse moduli space $f : \mathscr{X} \to X$ commutes with flat base change in X, and we may work fppf locally on X_S when checking that $f_S : \mathscr{X}_S \to X_S$ is the coarse moduli space of $\mathscr{X}_S$. We may therefore assume that $S = \operatorname{Spec} R$ for some ring R and, by [Abramovich and Vistoli 2002, 2.2.3 and its proof], that

$$X = \operatorname{Spec} A \quad \text{and} \quad \mathscr{X} = [(\operatorname{Spec} B)/G]$$

for some finite A-algebra B equipped with an action of a finite group G. In this situation, as is explained in [Conrad 2005, 3.1], we have $A = B^G$, the coarse moduli space of $\mathscr{X}_S$ is $\operatorname{Spec}((B \otimes_{\mathbb{Z}} R)^G)$, and we seek to prove that the map

$$j_R : B^G \otimes_{\mathbb{Z}} R \to (B \otimes_{\mathbb{Z}} R)^G$$

is an isomorphism granted that it is an isomorphism whenever $R = \mathbb{F}_p$ for any p.

The $\mathbb{Z}$-flatness of $\mathscr{X}$ ensures that B is torsion-free, so the abelian group B/B^G is also torsion-free. Therefore, $B^G \otimes_{\mathbb{Z}} R \to B \otimes_{\mathbb{Z}} R$, and hence also j_R, is injective for every $\mathbb{Z}$-module R. In order to conclude, we will prove that j_R is also surjective for every $\mathbb{Z}$-module R.

By passage to a filtered direct limit, we may assume that the $\mathbb{Z}$-module R is finitely generated. Thus, since the case $R = \mathbb{Z}$ is clear, we may assume that $R = \mathbb{Z}/n\mathbb{Z}$ for some $n \in \mathbb{Z}_{\geq 1}$. To then finally reduce to the assumed $R = \mathbb{Z}/p\mathbb{Z}$ case by dévissage, it remains to use the commutative diagram

$$\begin{tikzcd}
0 \arrow[r] & B^G \otimes_{\mathbb{Z}} R' \arrow[r] \arrow[d, hook, "j_{R'}"'] & B^G \otimes_{\mathbb{Z}} R \arrow[r] \arrow[d, hook, "j_R"'] & B^G \otimes_{\mathbb{Z}} R'' \arrow[r] \arrow[d, hook, "j_{R''}"'] & 0 \\
0 \arrow[r] & (B \otimes_{\mathbb{Z}} R')^G \arrow[r] & (B \otimes_{\mathbb{Z}} R)^G \arrow[r] & (B \otimes_{\mathbb{Z}} R'')^G &
\end{tikzcd}$$

that is in place whenever one has a short exact sequence $0 \to R' \to R \to R'' \to 0$ of $\mathbb{Z}$-modules. $\square$

We are ready for the promised conclusion about the coarse moduli space of $(\overline{\mathscr{E}\ell\ell}_n)_S$.

Proposition 3.3.2. *For any $n \in \mathbb{Z}_{\geq 1}$, the coarse moduli space of $\overline{\mathscr{E}\ell\ell}_n$ (resp. of the open substack $\mathscr{E}\ell\ell \subset \overline{\mathscr{E}\ell\ell}_n$) is isomorphic to $\mathbb{P}^1_{\mathbb{Z}}$ (resp. to $\mathbb{A}^1_{\mathbb{Z}} \subset \mathbb{P}^1_{\mathbb{Z}}$, with the map $\mathscr{E}\ell\ell \to \mathbb{A}^1_{\mathbb{Z}}$ being given by the j-invariant) and its formation commutes with base change to an arbitrary scheme S. In particular, $\overline{\mathscr{E}\ell\ell}_n$ is irreducible and has geometrically irreducible $\mathbb{Z}$-fibers.*

Proof. The last assertion follows from the rest because the map to the coarse moduli space induces a homeomorphism on topological spaces.

We begin with the $n = 1$ case, for which the base $S = \operatorname{Spec} \mathbb{Z}$ has been treated in [Deligne and Rapoport 1973, VI.1.1 and VI.1.3] and we only need to prove that the formation of the coarse moduli space of $\overline{\mathscr{E}\ell\ell}_1$ commutes with arbitrary base change. Let

$$\mathscr{C} \subset \overline{\mathscr{E}\ell\ell}_1$$

be the preimage of the open subscheme of $\mathbb{P}^1_{\mathbb{Z}}$ obtained by removing the sections $j = 0$ and $j = 1728$. By [Deligne 1975, 5.3(III)], the automorphism functor of every generalized elliptic curve classified by $\mathscr{C}$ is the constant group $\{\pm 1\}$. Therefore, as is explained in [ACV03 2003, §5.1], [Romagny 2005, §5], or [AOV08 2008, Appendix A], we may "quotient out" this constant group from the automorphism functors to obtain the algebraic stack $\mathscr{C} /\!\!/ \{\pm 1\}$ that is a "rigidification" of $\mathscr{C}$. By, for instance, [AOV08 2008, A.1], the rigidification map

$$\mathscr{C} \to \mathscr{C} /\!\!/ \{\pm 1\}$$

induces an isomorphism on coarse moduli spaces. However, by [Laumon and Moret-Bailly 2000, 8.1.1], the algebraic stack $\mathscr{C} /\!\!/ \{\pm 1\}$ is its own coarse moduli space. Thus, since the formation of $\mathscr{C} /\!\!/ \{\pm 1\}$ commutes with arbitrary base change, so does that of the coarse moduli space of $\mathscr{C}$. In particular, for every prime p, the map from the coarse moduli space of $(\overline{\mathscr{E}\ell\ell}_1)_{\mathbb{F}_p}$ to $\mathbb{P}^1_{\mathbb{F}_p}$ is an isomorphism on a dense open subscheme. However, this map is finite locally free due to the normality of its source inherited from the $\mathbb{F}_p$-smooth $(\overline{\mathscr{E}\ell\ell}_1)_{\mathbb{F}_p}$, so it is an isomorphism globally. This settles the $n = 1$ case for $S = \operatorname{Spec} \mathbb{F}_p$, and the general $n = 1$ case then follows from Lemma 3.3.1.

For general n, we begin by arguing that the coarse moduli space Y of $\overline{\mathscr{E}\ell\ell}_n$ is $\mathbb{Z}$-flat and that its formation commutes with arbitrary base change. By the settled $n = 1$ case, this is true on the elliptic curve locus, so we may focus on the open substack $\mathscr{C}_n \subset \overline{\mathscr{E}\ell\ell}_n$ that is the preimage of $\mathscr{C}$. By [Deligne and Rapoport 1973, II.2.8], every generalized elliptic curve has the automorphism -1 that restricts to inversion on the smooth locus. In particular, the constant group scheme $\{\pm 1\}$ is a canonical subgroup functor of the automorphism functor of every generalized elliptic curve classified by $\mathscr{C}_n$, so we may pass to the rigidification $\mathscr{C}_n /\!\!/ \{\pm 1\}$ and need to argue that its coarse moduli space is $\mathbb{Z}$-flat and of formation compatible with base change. This follows from [AOV08 2008, 3.3] because the algebraic stack $\mathscr{C}_n /\!\!/ \{\pm 1\}$ is tame by Lemma 2.1.6 and [Deligne 1975, 5.3(III)].

It remains to prove that the map $f : Y \to \mathbb{P}^1_{\mathbb{Z}}$ between the coarse moduli spaces of $\overline{\mathscr{E}\ell\ell}_n$ and $\overline{\mathscr{E}\ell\ell}_1$ is an isomorphism. By [Rydh 2013, 6.12], the coarse moduli space Y is $\mathbb{Z}$-proper, so the map in question is proper and quasifinite, and hence also finite by Lemma 3.2.3. Once we prove its flatness, and hence also local freeness, it will remain to inspect the elliptic curve locus to see that it is an isomorphism. Due to the $\mathbb{Z}$-flatness of Y and [EGA IV$_3$ 1966, 11.3.11], for the remaining flatness of f we may work $\mathbb{Z}$-fiberwise, and hence conclude with the help of [EGA IV$_2$ 1965, 6.1.5] after observing that for every field k, the reducedness of the k-smooth $(\overline{\mathscr{E}\ell\ell}_n)_k$ ensures the reducedness, and hence also the Cohen–Macaulay property, of its 1-dimensional coarse moduli space Y_k. □

3.4. Algebraization of formal generalized elliptic curves and of their homomorphisms

The goal of this section is to prove that a formal generalized elliptic curve that is adic over an affine Noetherian formal scheme and whose number of irreducible components of a degenerate geometric fiber is constant may be uniquely algebraized, and likewise for generalized elliptic curve homomorphisms — see Theorem 3.4.2 for a precise statement. Such algebraizability does not immediately follow from Grothendieck's formal GAGA formalism because the loci of smoothness may not be proper over the base, but it nevertheless is not surprising: if this formalism applied to the $\mathbb{Z}$-proper stack $\overline{\mathscr{E}\ell\ell}_n$ as it does in the scheme case, then the pullback map

$$\overline{\mathscr{E}\ell\ell}_n(R) \to \varprojlim_m \overline{\mathscr{E}\ell\ell}_n(R/I^m)$$

would be an equivalence for every adic Noetherian ring R with an ideal of definition I, and Theorem 3.4.2(a) would follow. The key difference from the scheme case is that a section of $(\overline{\mathscr{E}\ell\ell}_n)_R \to \operatorname{Spec} R$ is not a closed immersion. Nevertheless, an argument that we have extracted from [Olsson 2006, 5.4] proves a suitable formal GAGA statement recorded in Lemma 3.4.1 (see also [Aoki 2006b, §3.4; Aoki 2006a] for a similar argument).

Lemma 3.4.1. *Let R be a Noetherian ring that is complete and separated with respect to an ideal $I \subset R$. For every proper R-algebraic stack $\mathscr{X}$ with finite diagonal $\Delta_{\mathscr{X}/R}$ (for instance, for every proper Deligne–Mumford R-stack $\mathscr{X}$), the functor*

$$\mathscr{X}(R) \to \varprojlim_m \mathscr{X}(R/I^m) \tag{3.4.1.1}$$

is an equivalence of categories.

Proof. If $x, x' \in \mathscr{X}(R)$, then the isomorphism functor $\mathrm{Isom}(x, x')$ is a finite R-scheme, so

$$\mathrm{Isom}(x, x')(R) \to \varprojlim_m \mathrm{Isom}(x, x')(R/I^m)$$

is bijective by formal GAGA for schemes [EGA III$_1$ 1961, 5.1.6]. In other words, the functor (3.4.1.1) is fully faithful. For its essential surjectivity, suppose that

$$\{x_m \in \mathscr{X}(R/I^m)\}_{m \geq 1}$$

is a compatible sequence of objects. Due to the finiteness of $\Delta_{\mathscr{X}/R}$, each map

$$\mathrm{Spec}(R/I^m) \xrightarrow{x_m} \mathscr{X}_{R/I^m}$$

is representable by schemes and finite. Therefore, x_m corresponds to a coherent $\mathscr{O}_{\mathscr{X}_{R/I^m}}$-algebra $\mathscr{A}_m$. By formal GAGA for Artin stacks, i.e., by [Olsson 2006, A.1], the compatible system $\{\mathscr{A}_m\}_{m \geq 1}$ comes via base change from a unique coherent $\mathscr{O}_{\mathscr{X}}$-algebra $\mathscr{A}$. It remains to argue that the composition of the finite morphism $X \xrightarrow{x} \mathscr{X}$ corresponding to $\mathscr{A}$ and the structure morphism $\mathscr{X} \to \mathrm{Spec}\, R$ is an isomorphism. By construction, $x_{R/I^m} = x_m$ for every $m \geq 1$, so the claim will follow from [EGA III$_1$ 1961, 5.1.6] once we prove that the proper R-algebraic stack X is a finite R-scheme.

By [Conrad 2007, 2.2.5(2)], the algebraic space locus of X is open and contains $X_{R/I}$, so it must coincide with X. Since the relative dimension of X over R may be computed étale locally on X, [EGA IV$_3$ 1966, 13.1.3] proves that the relative dimension 0 locus of X is open, and hence must equal X because it contains $X_{R/I}$. To conclude that $X \to \mathrm{Spec}\, R$ is finite one then applies Lemma 3.2.3. □

The algebraization Theorem 3.4.2(a) has already been proved in [Conrad 2007, 2.2.4] by a different argument that does not use formal GAGA for Artin stacks (a similar argument had previously been used in [Deligne and Rapoport 1973, VII.1.10] to construct Tate curves), but it seems worthwhile to put this result in the context of Lemma 3.4.1. In contrast, the method of [Conrad 2007, 2.2.4] does not seem to suffice for the proof of the algebraizability of homomorphisms (beyond the case of isomorphisms), i.e., for Theorem 3.4.2(b). To algebraize homomorphisms we exploit their structure detailed in Section 2.2.

Theorem 3.4.2. *Let R be a Noetherian ring, complete and separated with respect to an ideal $I \subset R$.*

(a) *For each $n \in \mathbb{Z}_{\geq 1}$, every compatible under pullback sequence*

$$\{E_m \to \mathrm{Spec}(R/I^m)\}_{m \geq 1}$$

of generalized elliptic curves whose degenerate geometric fibers are n-gons is isomorphic to the sequence obtained via base change from a unique generalized elliptic curve $E \to \operatorname{Spec} R$.

(b) *For generalized elliptic curves* $E \to \operatorname{Spec} R$ *and* $E' \to \operatorname{Spec} R$, *every compatible sequence*

$$\{f_m : E_{R/I^m} \to E'_{R/I^m}\}_{m\geq 1}$$

of generalized elliptic curve homomorphisms (defined in Definition 2.1.12) *comes via base change from a unique generalized elliptic curve homomorphism*

$$f : E \to E'.$$

Proof. (a) Lemma 3.4.1 applied to $\overline{\mathscr{E}\ell\ell}_n$ proves the claim (for the uniqueness, Remark 2.1.9 ensures that the degenerate geometric fibers of E are n-gons).

(b) We begin with the case when all the f_m are isomorphisms (Lemma 3.4.1 does not apply because E need not correspond to an object of $\overline{\mathscr{E}\ell\ell}_n$ for any n). Due to Remark 2.1.9, there is no geometric point $\bar{s}$ of $\operatorname{Spec} R$ for which $E_{\bar{s}}$ and $E'_{\bar{s}}$ are both degenerate but have distinct numbers of irreducible components, so Proposition 3.1.8(a) shows that the isomorphism functor $\operatorname{Isom}(E, E')$ is a finite R-scheme. Therefore, by [EGA III$_1$ 1961, 5.1.6], the sequence

$$(f_m) \in \varprojlim_m \operatorname{Isom}(E, E')(R/I^m)$$

is induced by a desired unique

$$f \in \operatorname{Isom}(E, E')(R).$$

In the general case, by [EGA III$_1$ 1961, 5.4.1], the f_m algebraize to a unique R-morphism

$$f : E \to E',$$

and our task is to show that f is a generalized elliptic curve homomorphism. Since idempotents of R/I lift uniquely to R (see [EGA IV$_4$ 1967, 18.5.16(ii)]), we may use Proposition 2.2.9 to write

$$R = R' \times R'' \quad \text{and} \quad I = I' \times I''$$

in such a way that $(f_1)_{R'/I'}$ is the zero homomorphism and $(f_1)_{R''/I''}$ is an isogeny. Then R' (resp. R'') is complete and separated with respect to I' (resp. I'') and each $(f_m)_{R'/I'^m}$ (resp. $(f_m)_{R''/I''^m}$) is the zero homomorphism (resp. an isogeny). Thus, $f_{R'}$ must be the zero homomorphism, and we are reduced to the case when all the f_m are isogenies.

Let $K_m \subset E_{R/I^m}$ be the kernel of the isogeny f_m. The group law of K_m is the restriction of the action morphism

$$K_m \times E_{R/I^m} \to E_{R/I^m},$$

so [EGA III_1 1961, 5.1.4 and 5.4.1] supply a finite locally free R-subgroup $K \subset E^{\mathrm{sm}}$ that algebraizes all the K_m. Corollary 2.2.7(b) and the settled case when the f_m are isomorphisms then provide the identification $E/K \cong E'$, so f is identified with the isogeny $E \to E/K$ and hence is a homomorphism. □

Chapter 4. Modular descriptions of modular curves

With the compactifications $\overline{\mathscr{E}\ell\ell}_n$ at our disposal, we are ready to exhibit the moduli interpretations and the regularity of several classical modular curves, such as $\mathscr{X}(n)$ or $\mathscr{X}_1(n)$ (see Section 1.7 for an overview of our method and of previous work). We begin in Section 4.1 by reviewing the construction and the properties of modular curves of arbitrary congruence level. The moduli interpretations of $\mathscr{X}(n)$ and $\mathscr{X}_1(n)$ given in Sections 4.3 and 4.4 use Drinfeld structures on generalized elliptic curves, so in Section 4.2 we extend a number of properties of such structures from the elliptic curve case studied by Katz and Mazur. In Section 4.5, we synthesize the arguments used for $\mathscr{X}(n)$ and $\mathscr{X}_1(n)$ in the form of an axiomatic result, which we use in Section 4.6 to treat further modular curves $\widetilde{\mathscr{X}_1}(n; n')$, $\mathscr{X}_1(n; n')$, and $\mathscr{X}_0(n; n')$ for suitable n and n'. The analysis of $\mathscr{X}_1(n; n')$ is used in Section 4.7 to give a modular construction of some Hecke correspondences for $\mathscr{X}_1(n)$.

4.1. Modular curves of congruence level

The main goal of this section is to review the definition given by Deligne and Rapoport [1973, IV.3.3] of (stacky) modular curves over $\mathbb{Z}$ of congruence level. The definition is via a normalization procedure, and for general levels there is no known description of these $\mathbb{Z}$-curves as moduli spaces of generalized elliptic curves equipped with additional structure (one of the principal goals of this paper is to give such a description in the case of $\Gamma_0(n)$ level). The normalization procedure rests on the case of "no level," with which we begin.

4.1.1. The case of no additional level. In this case, the modular curve in question is the $\mathbb{Z}$-stack $\overline{\mathscr{E}\ell\ell}_1$ that parametrizes generalized elliptic curves with integral geometric fibers (see Definition 3.1.1). In the context of level structures, we will denote $\overline{\mathscr{E}\ell\ell}_1$ by $\mathscr{X}_{\mathrm{GL}_2(\widehat{\mathbb{Z}})}$, by $\mathscr{X}_{\Gamma(1)}$, or simply by $\mathscr{X}(1)$, and we will denote its elliptic curve locus $\mathscr{E}\ell\ell$ by similar notation with $\mathscr{X}$ replaced by $\mathscr{Y}$, e.g., by

$$\mathscr{Y}(1) \subset \mathscr{X}(1).$$

By Theorem 3.1.6(a)–(b) (i.e., by [Deligne and Rapoport 1973, III.2.5(i), III.1.2(iii), and IV.2.2]), the stack $\mathscr{X}(1)$ is Deligne–Mumford and the morphism

$$\mathscr{X}(1) \to \operatorname{Spec} \mathbb{Z}$$

is proper and smooth of relative dimension 1.

4.1.2. The case of an arbitrary congruence level H. The *level* is an open (and hence finite index) subgroup H of $\mathrm{GL}_2(\widehat{\mathbb{Z}})$. Its associated modular curve $\mathscr{X}_H$ is a Deligne–Mumford $\mathbb{Z}$-stack that, loosely speaking, compactifies the stack $\mathscr{Y}_H[1/\text{level}]$ which represents the "level H moduli problem" on elliptic curves over schemes on which bad primes that depend on the level are invertible. More precisely, given H, one fixes an $n \in \mathbb{Z}_{\geq 1}$ for which

$$\operatorname{Ker}(\mathrm{GL}_2(\widehat{\mathbb{Z}}) \twoheadrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})) \subset H \quad \text{and sets} \quad \bar{H} := \operatorname{Im}(H \to \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})).$$

One then lets $\mathscr{Y}_H[1/n]$ be the $\mathbb{Z}[1/n]$-stack that, for variable $\mathbb{Z}[1/n]$-schemes S, parametrizes elliptic curves $E \to S$ equipped with an S-point of the finite étale S-scheme

$$\bar{H} \backslash \operatorname{Isom}(E[n], (\mathbb{Z}/n\mathbb{Z})^2).$$

Finally, one defines $\mathscr{X}_H$ to be the Deligne–Mumford $\mathscr{X}(1)$-stack obtained by normalizing $\mathscr{X}(1)$ with respect to the "forgetful" finite étale morphism

$$\mathscr{Y}_H\big[\tfrac{1}{n}\big] \to \mathscr{Y}(1)_{\mathbb{Z}[\frac{1}{n}]}.$$

One lets $\mathscr{Y}_H$ be the preimage of $\mathscr{Y}(1)$ in $\mathscr{X}_H$. It is proved in [Deligne and Rapoport 1973, IV.3.6] that different choices of n lead to canonically isomorphic $\mathscr{X}_H$.

The map

$$\mathscr{X}_H \to \mathscr{X}(1) \tag{4.1.2.1}$$

is representable, finite, and also surjective because $\mathscr{X}(1)$ is irreducible. Moreover, by [EGA IV$_2$ 1965, 6.1.5] (which applies because of "going down" and the normality of $\mathscr{X}_H$), the map (4.1.2.1) is flat, so it is locally free of rank $[\mathrm{GL}_2(\widehat{\mathbb{Z}}) : H]$ and $\mathscr{X}_H$ is of relative dimension 1 over $\mathbb{Z}$ at every point. By [Deligne and Rapoport 1973, IV.6.7], the proper and flat structure morphism $\mathscr{X}_H \to \operatorname{Spec} \mathbb{Z}$ is even smooth over $\mathbb{Z}[1/n]$. If $H' \subset H$, then the finite étale $\mathscr{Y}(1)$-morphism

$$\mathscr{Y}_{H'}\big[\tfrac{1}{n}\big] \to \mathscr{Y}_H\big[\tfrac{1}{n}\big]$$

obtained from the S-morphisms

$$\overline{H'} \backslash \operatorname{Isom}\big(E[n], (\mathbb{Z}/n\mathbb{Z})^2\big) \to \bar{H} \backslash \operatorname{Isom}\big(E[n], (\mathbb{Z}/n\mathbb{Z})^2\big)$$

gives rise to the finite $\mathscr{X}(1)$-morphism

$$\mathscr{X}_{H'} \to \mathscr{X}_H.$$

Thus, due to the following lemma and Proposition 4.3.6, all the $\mathscr{X}_H$ are schemes for small enough H.

Lemma 4.1.3. *If the modular curve $\mathscr{X}_H$ has an open substack $U \subset \mathscr{X}_H$ whose geometric points have no nontrivial automorphisms, then U is a scheme that is quasiprojective over* $\operatorname{Spec}\mathbb{Z}$.

Proof. By Lemma 3.2.2(a), U is an algebraic space. Moreover, the coarse moduli space morphism $\mathscr{X}(1) \to \mathbb{P}^1_{\mathbb{Z}}$ is separated and quasifinite, so $U \to \mathbb{P}^1_{\mathbb{Z}}$ is also separated and quasifinite, and hence U is a scheme by Lemma 3.2.3. Finally, the morphism $U \to \mathbb{P}^1_{\mathbb{Z}}$ is quasiprojective by [EGA IV$_3$ 1966, 8.11.2] or by Zariski's main theorem [EGA IV$_3$ 1966, 8.12.6], so $U \to \operatorname{Spec}\mathbb{Z}$ is also quasiprojective. □

Remark 4.1.4. Due to Lemma 4.1.3 and [Conrad 2007, 2.2.5(2)], each $\mathscr{X}_H$ has a unique largest open subscheme. This subscheme contains exactly those geometric points of $\mathscr{X}_H$ whose automorphism functors are trivial.

One suspects that $\mathscr{X}_H$ is the "correct" modular curve of level H, in part because there is no other choice granted that one believes that such a modular curve should be representable and finite over $\mathscr{X}(1)$, normal, and agree with $\mathscr{Y}_H[1/n]$ over $\mathscr{Y}(1)_{\mathbb{Z}[1/n]}$. One of the bottlenecks limiting practical usefulness of the stacks $\mathscr{X}_H$ is the lack of descriptions of their functors of points (without inverting the level) in terms of generalized elliptic curves equipped with additional data. In the cases where such descriptions have been found, one has been able to analyze $\mathscr{X}_H$ more thoroughly, e.g., to prove that $\mathscr{X}_H$ is regular (and not just normal). Such regularity is useful in practice (but is not known in general) — for instance, through [EGA IV$_2$ 1965, 6.1.5] it would ensure flatness of the maps $\mathscr{X}_H \to \mathscr{X}_{H'}$ mentioned above. Similarly, the proof of the $\mathbb{Z}[1/n]$-smoothness of $(\mathscr{X}_H)_{\mathbb{Z}[1/n]}$ given in [Deligne and Rapoport 1973, IV.6.7] rests on the modular description of $(\mathscr{X}_H)_{\mathbb{Z}[1/n]}$ presented in [loc. cit.] for *any* H (however, this description is not explicit enough to *a priori* recover the "obvious" candidate descriptions for classical choices of H).

Modular descriptions of $\mathscr{X}_H$ are known for most "classical" H, and we will reprove some of them in Sections 4.3–4.6 below.

4.2. Drinfeld level structures on generalized elliptic curves via congruences

In order to efficiently handle all residue characteristics, the modular descriptions of various $\mathscr{X}_H$ that will be discussed in subsequent sections will use Drinfeld level structures on generalized elliptic curves. In the elliptic curve case, the necessary properties of such structures follow from the work of Katz and Mazur [1985], and the goal of this section is to extend them to the generalized elliptic curve case. Some such extensions have already been obtained in [Conrad 2007], but our method seems simpler, more direct, and applies in a wider range of situations. The key idea is to exploit "mod n congruences" with elliptic curves: the properties of various “mod n Drinfeld level structures” tend to be fppf local and to depend solely on the n-torsion $E^{\mathrm{sm}}[n]$, so for many purposes we may first use Corollary 3.2.6 to reduce to the case when $E^{\mathrm{sm}}[n]$ is finite locally free of rank n^2 and then apply the following lemma to further reduce to the elliptic curve case.

Lemma 4.2.1. *For every $n \in \mathbb{Z}_{\geq 1}$ and every generalized elliptic curve $E \to S$ for which n divides the number of irreducible components of each degenerate geometric fiber, there is an fppf cover $S' \to S$ and an elliptic curve $E' \to S'$ for which*

$$E^{\mathrm{sm}}_{S'}[n] \simeq E'[n].$$

Proof. We may work étale locally on S, so limit arguments allow us to assume that S is local and strictly Henselian. We may then also assume that the special fiber of E is degenerate, so the connected-étale sequence (together with Lemma 2.1.11) shows that $E^{\mathrm{sm}}[n]$ is an extension of $\mathbb{Z}/n\mathbb{Z}$ by μ_n. After passage to an fppf cover this extension splits and our task reduces to showing that fppf locally on $\operatorname{Spec}\mathbb{Z}$ there is an elliptic curve E' with $E'[n] \cong \mu_n \times \mathbb{Z}/n\mathbb{Z}$.

Via limit arguments, it suffices to find such an E' over each strict Henselization $(R, \mathfrak{m})$ of $\operatorname{Spec}\mathbb{Z}$ at every closed point. The conclusion then follows from choosing an ordinary elliptic curve over $R/\mathfrak{m}$, lifting its Weierstrass equation to R, and using the connected-étale sequence again. □

To make sense of Drinfeld level structures as alluded to above, we recall the following key definition:

Definition 4.2.2. For a finite abelian group A and a generalized elliptic curve $E \to S$, a *Drinfeld A-structure on E* is a homomorphism $\alpha : A \to E^{\mathrm{sm}}(S)$ for which the relative effective Cartier divisor

$$D_\alpha := \textstyle\sum_{a \in A} [\alpha(a)] \subset E^{\mathrm{sm}}$$

is an S-subgroup scheme. If this S-subgroup $G \subset E^{\mathrm{sm}}$ is given in advance, then we say that α is a *Drinfeld A-structure on G*.

Remark 4.2.3. By [Katz and Mazur 1985, 1.5.3], if $\#A$ is invertible on S, then a Drinfeld A-structure α on E amounts to an isomorphism induced by α between the constant S-group $\underline{A}_S$ and some S-subgroup of E^{sm}.

Convention 4.2.4. In the sequel we will sometimes deal with Drinfeld $\mathbb{Z}/nm\mathbb{Z}$- or $(\mathbb{Z}/nm\mathbb{Z})^2$-structures for fixed $n, m \in \mathbb{Z}_{\geq 1}$ and will want to obtain $\mathbb{Z}/n\mathbb{Z}$- or $(\mathbb{Z}/n\mathbb{Z})^2$-structures by restricting to the n-torsion subgroups. To make sense of this we need to choose noncanonical isomorphisms

$$\mathbb{Z}/n\mathbb{Z} \simeq (\mathbb{Z}/nm\mathbb{Z})[n] \quad \text{and} \quad (\mathbb{Z}/n\mathbb{Z})^2 \simeq (\mathbb{Z}/nm\mathbb{Z})^2[n].$$

The particular choices will never matter for the results, but for definiteness we always choose the isomorphisms induced by multiplication by m on $\mathbb{Z}$ or on $\mathbb{Z}^2$.

In the results below, the "compare with" references point to the elliptic curve cases treated by Katz and Mazur. We begin by detailing the properties of restrictions to subgroups of various Drinfeld structures on generalized elliptic curves. Parts (a) and (c) of Proposition 4.2.5 have been proved in [Conrad 2007, 2.3.2] by a different method that also eventually reduces to the elliptic curve case.

Proposition 4.2.5. *Let $n, m \in \mathbb{Z}_{\geq 1}$, and let $E \to S$ be a generalized elliptic curve.*

(a) (Compare with [Katz and Mazur 1985, 5.5.2(1) and 5.5.7(1)]). *If α is a Drinfeld $(\mathbb{Z}/nm\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[nm]$, then $\alpha|_{(\mathbb{Z}/nm\mathbb{Z})^2[n]}$ is a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[n]$ and $\alpha|_{\mathbb{Z}/nm\mathbb{Z}\times\{0\}}$ is a Drinfeld $\mathbb{Z}/nm\mathbb{Z}$-structure on E.*

(b) (Compare with [Katz and Mazur 1985, 5.5.8(1)]). *If $\alpha : (\mathbb{Z}/nm\mathbb{Z})^2 \to E^{\mathrm{sm}}(S)$ is a group homomorphism, every prime divisor of m divides n, and $\alpha|_{(\mathbb{Z}/nm\mathbb{Z})^2[n]}$ is a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[n]$, then α is a Drinfeld $(\mathbb{Z}/nm\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[nm]$ (so, in particular, the number of irreducible components of each degenerate geometric fiber of E is divisible by nm).*

(c) (Compare with [Katz and Mazur 1985, 5.5.7(2)]). *If α is a Drinfeld $\mathbb{Z}/nm\mathbb{Z}$-structure on E, then $\alpha|_{(\mathbb{Z}/nm\mathbb{Z})[n]}$ is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure on E.*

(d) (Compare with [Katz and Mazur 1985, 5.5.8(2)]). *If $\alpha : \mathbb{Z}/nm\mathbb{Z} \to E^{\mathrm{sm}}(S)$ is a group homomorphism, every prime divisor of m divides n, and $\alpha|_{(\mathbb{Z}/nm\mathbb{Z})[n]}$ is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure on E, then α is a Drinfeld $\mathbb{Z}/nm\mathbb{Z}$-structure on E.*

(e) (Compare with [Katz and Mazur 1985, 5.5.2(2)]). *For brevity, set $N := nm$. If α is a Drinfeld $(\mathbb{Z}/N\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[N]$ and $G \subset E^{\mathrm{sm}}$ is the subgroup $\sum_{i\in\mathbb{Z}/N\mathbb{Z}\times\{0\}}[\alpha(i)]$ supplied by* (a), *then*

$$\alpha|_{\{0\}\times\mathbb{Z}/N\mathbb{Z}} : \{0\} \times \mathbb{Z}/N\mathbb{Z} \to (E/G)^{\mathrm{sm}}(S)$$

is a Drinfeld $\mathbb{Z}/N\mathbb{Z}$-structure on $E^{\mathrm{sm}}[N]/G \subset (E/G)^{\mathrm{sm}}$.

Proof. It suffices to work fppf locally on S, so we may use Corollary 3.2.6 to reduce to the case when the number of irreducible components of each degenerate geometric fiber of E is divisible by nm (in parts (a) and (e) we are in this case at the outset). We may then apply Lemma 4.2.1 to assume further that there is an elliptic curve $E' \to S$ with $E'[nm] \simeq E^{\mathrm{sm}}[nm]$. By [Katz and Mazur 1985, 1.10.6 and 1.10.11], the properties under consideration depend solely on the S-group scheme $E^{\mathrm{sm}}[nm]$ equipped with the homomorphism α and not on the embedding of $E^{\mathrm{sm}}[nm]$ into a smooth S-group scheme of relative dimension 1 (such as E^{sm} or E'). Thus, the claims result from their elliptic curve cases. □

Cyclic subgroups of generalized elliptic curves will be important for us, so we recall their definition.

Definition 4.2.6. For a generalized elliptic curve $E \to S$, a finite locally free S-subgroup $G \subset E^{\mathrm{sm}}$ is *cyclic of order n* if fppf locally on S there is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure on G. For a G that is cyclic of order n, a section $g \in G(S)$ is a *generator of G* (or *generates G*) if the homomorphism

$$\alpha : \mathbb{Z}/n\mathbb{Z} \to E^{\mathrm{sm}}(S)$$

defined by $\alpha(1) = g$ is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure on G. An isogeny of constant degree n between generalized elliptic curves over S is *cyclic* if its kernel is cyclic of order n.

We turn to the properties of cyclic subgroups of generalized elliptic curves. Parts (a), (d), and (f) of Proposition 4.2.7 have also been reduced to the elliptic curve case in [Conrad 2007, 2.3.7, 2.3.8, and 2.3.5] by a different method.

Proposition 4.2.7. *Let $E \to S$ be a generalized elliptic curve, $G \subset E^{\mathrm{sm}}$ an S-subgroup that is finite locally free of rank n over S, and $G^\times \subset G$ the S-subsheaf of generators of G* (by [Katz and Mazur 1985, 1.6.5], the S-subsheaf $G^\times$ is a closed S-subscheme of G of finite presentation).

(a) (The Katz–Mazur cyclicity criterion; compare with [Katz and Mazur 1985, 6.1.1(1)]). *The subgroup G is cyclic of order n if and only if $G^\times$ is finite locally free of rank $\phi(n)$ over S. In particular, G is cyclic of order n if and only if it becomes cyclic of order n over an fpqc cover of S. If n is invertible on S and G is cyclic of order n, then $G^\times \to S$ is étale.*

(b) (Compare with [Katz and Mazur 1985, 6.1.1(2)]). *If $g \in G(S)$ is a generator of G, then*

$$G^\times = \textstyle\sum_{i \in (\mathbb{Z}/n\mathbb{Z})^\times} [i \cdot g] \quad \textit{as effective Cartier divisors on } E^{\mathrm{sm}}.$$

(c) (Compare with [Katz and Mazur 1985, 6.4.1]). *There is a finitely presented closed subscheme $T \subset S$ such that the base change $G_{S'}$ to an S-scheme S' is cyclic if and only if $S' \to S$ factors through T.*

(d) (Compare with [Katz and Mazur 1985, 6.8.7]). *If n is squarefree, then G is cyclic.*

(e) (Compare with [Katz and Mazur 1985, 5.5.4(3)]). *If G is cyclic of order n and the number of irreducible components of each degenerate geometric fiber of $E \to S$ is divisible by n, then the subgroup $E^{\mathrm{sm}}[n]/G$ of E/G is cyclic of order n.*

(f) (Compare with [Katz and Mazur 1985, 6.7.2]). *If G is cyclic and $g, g' \in G(S)$ are generators of G, then for every positive divisor d of n both $\frac{n}{d} \cdot g$ and $\frac{n}{d} \cdot g'$ are generators of the same S-subgroup*

$$G_d \subset G$$

that is cyclic of order d. In particular, if G is cyclic, then the fppf local on S subgroup of G defined in this way descends to a canonical cyclic S-subgroup $G_d \subset G$ of order d.

Proof. Cyclicity is an fppf local condition, so we may work fppf locally on S. We may therefore use Corollary 3.2.6 and Lemma 4.2.1 to assume that the number of irreducible components of each degenerate geometric fiber of $E \to S$ is divisible by n and that there is an elliptic curve $E' \to S$ such that $E^{\mathrm{sm}}[n] \simeq E'[n]$. Thus, since, by [Katz and Mazur 1985, 1.10.6 and its generalization 1.10.1], the properties under consideration depend solely on the S-group scheme $E^{\mathrm{sm}}[n]$ and its subgroup G, the claims follow from their elliptic curve cases (in (a), if n is invertible on S, then a cyclic G of order n becomes isomorphic to $\mathbb{Z}/n\mathbb{Z}$ over an étale cover of S, so that $G^\times$ becomes isomorphic to the constant subscheme $(\mathbb{Z}/n\mathbb{Z})^\times \subset \mathbb{Z}/n\mathbb{Z}$). □

Definition 4.2.8. For a generalized elliptic curve $E \to S$ and a cyclic S-subgroup $G \subset E^{\mathrm{sm}}$ of order n, the S-subgroup G_d defined in Proposition 4.2.7(f) is *the standard cyclic subgroup of G of order d*. Isogenies $f_1 : E \to E'$ and $f_2 : E' \to E''$ of constant degrees between generalized elliptic curves over S are *cyclic in standard order* if $\mathrm{Ker}(f_2 \circ f_1)$ is cyclic and $\mathrm{Ker}\, f_1$ is its standard cyclic subgroup (so that, in particular, f_1 and f_2 are both cyclic by Proposition 4.2.9(e) below).

In Propositions 4.2.9 and 4.2.10 we extend various results of [Katz and Mazur 1985, §6.7] about standard cyclic subgroups and standard order factorizations of cyclic isogenies to the case of generalized elliptic curves (Chapter 2 provides a robust extension of the notion of an isogeny). Some of these extensions will be important for the analysis of $\mathscr{X}_{\Gamma_0(n)}$ carried out in Chapter 5.

Proposition 4.2.9. *Let $E \to S$ be a generalized elliptic curve, let $G \subset E^{\mathrm{sm}}$ be a cyclic S-subgroup of order n, let d and d' be positive divisors n, and let*

$$G_d \subset G$$

denote the standard cyclic subgroup of order d.

(a) (Compare with [Katz and Mazur 1985, 6.7.4]). *If $d \mid d'$, then G_d is identified with the standard cyclic subgroup of $G_{d'}$ of order d.*

(b) *Interpreting the intersection as that of fppf subsheaves of G, we have*

$$G_d \cap G_{d'} = G_{\gcd(d,d')}.$$

(c) *If G meets precisely m irreducible components of every degenerate geometric fiber of E, then G_d meets precisely $m/\gcd\big(m, \frac{n}{d}\big)$ irreducible components of every degenerate geometric fiber of E.*

(d) (Compare with [Katz and Mazur 1985, 6.7.5]). *Letting $G_d^\times$ denote the S-scheme parametrizing the generators of G_d* (so that, by Proposition 4.2.7(a),

$G_d^\times$ is a closed subscheme of G_d and is finite locally free of rank $\phi(d)$ over S), *we have*

$$G = \textstyle\sum_{d|n} G_d^\times \quad \textit{as effective Cartier divisors on } E^{\mathrm{sm}}.$$

(e) (Compare with [Katz and Mazur 1985, 6.7.4]). *The quotient*

$$G/G_d \subset (E/G_d)^{\mathrm{sm}}$$

is a cyclic S-subgroup of order $\frac{n}{d}$, the image of any generator of G generates G/G_d, and if $d \mid d'$, then the standard cyclic subgroup of G/G_d of order $\frac{d'}{d}$ is identified with $G_{d'}/G_d$.

(f) (Compare with [Katz and Mazur 1985, 6.7.11 (2)]). *If n and $\frac{n}{d}$ have the same prime divisors, then $g \in G(S)$ generates G if and only if its image generates G/G_d, and, in particular, g generates G if and only if $g+h$ generates G for some (equivalently, for any) $h \in G_d(S)$.*

Proof. Part (a) follows from the definitions because we may work fppf locally to assume that G has a generator. Part (b) follows from (a): since $G_{\gcd(d,d')}$ lies inside both G_d and $G_{d'}$, it suffices to observe that $G_d/G_{\gcd(d,d')}$ and $G_{d'}/G_{\gcd(d,d')}$ have coprime orders and hence intersect trivially inside $G/G_{\gcd(d,d')}$. Part (c) follows from the definition of G_d. To prove part (d), we pass to an fppf cover of S over which G admits a generator and apply Proposition 4.2.7(b).

For the remaining (e) and (f), we work fppf locally on S and use Corollary 3.2.6 and Lemma 4.2.1 to assume that G has a generator, that the number of irreducible components of each degenerate geometric fiber of E is divisible by n, and that there is an elliptic curve $E' \to S$ with $E^{\mathrm{sm}}[n] \simeq E'[n]$. By [Katz and Mazur 1985, 1.10.6], the properties under consideration in (e) and (f) depend solely on the S-group G and not on its embedding into E^{sm} or E', so (e) and (f) follow from their elliptic curve cases. □

Proposition 4.2.10. *Let*

$$f_1 : E_0 \to E_1, \quad f_2 : E_1 \to E_2, \quad \textit{and} \quad f := f_2 \circ f_1 : E_0 \to E_2$$

be isogenies of constant degrees d_1, d_2, and $d_1 d_2$ between generalized elliptic curves over S.

(a) (Compare with [Katz and Mazur 1985, 6.7.8]). *If f is cyclic and* Ker f_2 *is étale over S, then f_1 and f_2 are cyclic in standard order.*

(b) (Compare with [Katz and Mazur 1985, 6.7.10]). *If d_1 and d_2 are coprime, then f is cyclic if and only if both f_1 and f_2 are cyclic, in which case f_1 and f_2 are cyclic in standard order.*

(c) (Compare with [Katz and Mazur 1985, 6.7.12]). *If f_1 and f_2 are cyclic*, d_1 *and* d_2 *have the same prime divisors*, *and* $g \in (\mathrm{Ker}\, f)(S)$ *is such that* $d_2 \cdot g$ *generates* $\mathrm{Ker}\, f_1$ *and* $f_1(g)$ *generates* $\mathrm{Ker}\, f_2$, *then* f_1 *and* f_2 *are cyclic in standard order and* g *generates* $\mathrm{Ker}\, f$.

(d) (Compare with [Katz and Mazur 1985, 6.7.15]). *If* $\{E_{i-1} \xrightarrow{f_i} E_i\}_{i=3}^n$ *are further isogenies of constant degrees* d_i *between generalized elliptic curves over* S *such that* $d_1, \dots, d_n$ *all have the same prime divisors and such that for each* $1 \le i \le n-1$ *the isogenies* f_i *and* f_{i+1} *are cyclic in standard order*, *then* $\mathrm{Ker}(f_n \circ \cdots \circ f_1)$ *is cyclic and each* $\mathrm{Ker}(f_i \circ \cdots \circ f_1)$ *is its standard cyclic subgroup.*

Proof. For notational convenience, we set $n := 2$ in (a)–(c). By Corollary 2.2.7 and [Katz and Mazur 1985, 1.10.6], the properties under consideration may be expressed in terms of the S-group scheme $\mathrm{Ker}(f_n \circ \dots \circ f_1)$ equipped with its S-subgroups $\mathrm{Ker}(f_i \circ \dots \circ f_1)$. Thus, since the claims are fppf local on S, Corollary 3.2.6 and Lemma 4.2.1 allow us to assume that the number of irreducible components of each degenerate geometric fiber of E_0 is divisible by $\prod_{i=1}^n d_i$ and that there is an elliptic curve $E' \to S$ with

$$E_0^{\mathrm{sm}}\big[\textstyle\prod_{i=1}^n d_i\big] \simeq E'\big[\textstyle\prod_{i=1}^n d_i\big].$$

This reduces to the elliptic curve cases treated by Katz–Mazur in [op. cit.]. □

We wish to prove in Proposition 4.2.11(b) a generalization of the claim of [Conrad 2007, 2.4.5] that is important for the definition of $\Gamma_1(N; n)$-structures given there. The argument given in [loc. cit.] seems to require further input: the "universal deformation technique" invoked towards the end of the proof does not seem to apply directly because it is based on [Deligne and Rapoport 1973, III.1.2(iii)] that requires the number of irreducible components of the closed fiber to be prime to the residue characteristic and the $\mathbb{Z}/N\mathbb{Z}$-structure P may interfere with this requirement.

Proposition 4.2.11. *Let* $E \to S$ *be a generalized elliptic curve*, *and let* $n, m \in \mathbb{Z}_{\ge 1}$.

(a) *If* $G \subset E^{\mathrm{sm}}$ *and* $H \subset E^{\mathrm{sm}}$ *are* S*-subgroups that are cyclic of orders* n *and* m, *respectively*, *and* α *and* β *are fppf local on* S *Drinfeld* $\mathbb{Z}/n\mathbb{Z}$*- and* $\mathbb{Z}/m\mathbb{Z}$*-structures on* G *and* H, *then*

$$\sum_{\substack{i \in \mathbb{Z}/n\mathbb{Z} \\ j \in \mathbb{Z}/m\mathbb{Z}}} [\alpha(i) + \beta(j)]$$

is an effective Cartier divisor on E^{sm} *that does not depend on the choices of* α *and* β *and descends to a well-defined relative effective Cartier divisor on* E^{sm} *over* S *denoted by* $[G + H]$.

(b) *Set* $d := \gcd(n, m)$ *and suppose that the number of irreducible components of each degenerate geometric fiber of* $E \to S$ *is divisible by* d. *If* $G \subset E^{\mathrm{sm}}$ *and* $H \subset E^{\mathrm{sm}}$ *are* S*-subgroups that are cyclic of orders* n *and* m, *respectively, and* $[G_d + H_d] = E^{\mathrm{sm}}[d]$, *then* $[G + H]$ *is a finite locally free* S*-subgroup scheme of* E^{sm} *of order* nm *and killed by* $\mathrm{lcm}(n, m)$, *and any Drinfeld* $\mathbb{Z}/n\mathbb{Z}$*-structure on* G *induces a Drinfeld* $\mathbb{Z}/n\mathbb{Z}$*-structure on* $[G + H]/H \subset (E/H)^{\mathrm{sm}}$.

Proof. For (a), the cases when either α or β is fixed suffice, so one only needs to observe that translation by an S-point is an automorphism of the S-scheme E^{sm} and hence commutes with the formation of the sum of effective Cartier divisors — for example, the left hand side of

$$\alpha(i) + H = \textstyle\sum_{j \in \mathbb{Z}/m\mathbb{Z}} [\alpha(i) + \beta(j)]$$

does not depend on β.

For (b), we work fppf locally on S and use Corollary 3.2.6 to assume that the number of irreducible components of each degenerate geometric fiber of $E \to S$ is divisible by nm and that there are Drinfeld $\mathbb{Z}/n\mathbb{Z}$- and $\mathbb{Z}/m\mathbb{Z}$-structures α and β on G and H. We then imitate the argument of [Conrad 2007, top of p. 231] given in the elliptic curve case. Namely, we use [Katz and Mazur 1985, 1.7.2 and 1.10.6] to "factor into prime powers" to reduce to the case when $n = p^r$ and $m = p^s$ for some prime p and $r \le s$ (the $r \ge s$ case of the last aspect of the claim will be argued separately in the last paragraph of this proof). We assume that $r \ge 1$ (otherwise $[G + H] = H$) and, after replacing S by an fppf cover, we choose a homomorphism $\widetilde{\alpha} : \mathbb{Z}/p^s\mathbb{Z} \to E(S)$ with $p^{s-r}\widetilde{\alpha}(1) = \alpha(1)$. By Proposition 4.2.5(b),

$$\widetilde{\alpha} + \beta : (\mathbb{Z}/p^s\mathbb{Z})^2 \to E^{\mathrm{sm}}(S)$$

is a Drinfeld $(\mathbb{Z}/p^s\mathbb{Z})^2$-structure on $E[p^s]$, so, by Proposition 4.2.5(e),

$$\widetilde{\alpha} : \mathbb{Z}/p^s\mathbb{Z} \to (E/H)^{\mathrm{sm}}(S)$$

is a Drinfeld $\mathbb{Z}/p^s\mathbb{Z}$-structure on E/H. Then, by Proposition 4.2.5(c),

$$\alpha : \mathbb{Z}/p^r\mathbb{Z} \to (E/H)^{\mathrm{sm}}(S)$$

is a Drinfeld $\mathbb{Z}/p^r\mathbb{Z}$-structure on a subgroup $K \subset (E/H)^{\mathrm{sm}}$. Finally, by [Katz and Mazur 1985, 1.11.3], the scheme $[G + H]$ is the preimage of K in E, so is a subgroup, as desired. Moreover, $[G + H]$ is killed by p^s because the quotient $[G + H]/E[p^r]$ is killed by its order, i.e., by p^{s-r}, whereas $E[p^r]$ is killed by p^r. By construction, α, whose particular choice is irrelevant for the argument, induces a Drinfeld $\mathbb{Z}/p^r\mathbb{Z}$-structure on $[G + H]/H$.

It remains to prove that any α also induces a Drinfeld $\mathbb{Z}/p^r\mathbb{Z}$-structure on $[G + H]/H \subset (E/H)^{\mathrm{sm}}$ when $r \ge s$ and $s \ge 1$. For this, by Proposition 4.2.5(e),

$\alpha|_{(\mathbb{Z}/p^r\mathbb{Z})[p^s]}$ induces a Drinfeld $\mathbb{Z}/p^s\mathbb{Z}$-structure on E/H, so, by Proposition 4.2.5(d), α induces a Drinfeld $\mathbb{Z}/p^r\mathbb{Z}$-structure on some S-subgroup $K' \subset (E/H)^{\mathrm{sm}}$, and it remains to apply [Katz and Mazur 1985, 1.11.3] again to deduce that the preimage of K' in E must equal $[G+H]$. □

One of the cornerstones of our approach to the study of various moduli stacks of Drinfeld A-structures on generalized elliptic curves is a direct reduction of many questions to the $A = (\mathbb{Z}/n\mathbb{Z})^2$ case. To make reductions of this sort feasible we will need the following result:

Proposition 4.2.12. *Let $E \to S$ be a generalized elliptic curve, let $n, m \in \mathbb{Z}_{\geq 1}$, let S' be a variable S-scheme, and recall* Convention 4.2.4.

(a) *If the number of irreducible components of each degenerate geometric fiber of $E \to S$ is divisible by nm and α is a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[n]$, then the functor*

$$S' \mapsto \big\{\textit{Drinfeld } (\mathbb{Z}/nm\mathbb{Z})^2\textit{-structures } \beta \textit{ on } E^{\mathrm{sm}}_{S'}[nm] \quad \textit{such that } \beta|_{(\mathbb{Z}/nm\mathbb{Z})[n]} = \alpha_{S'}\big\}$$

is representable by a finite locally free S-scheme of rank

$$\frac{\#\,\mathrm{GL}_2(\mathbb{Z}/nm\mathbb{Z})}{\#\,\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})}$$

that is étale if nm is invertible on S.

(b) (Compare with [Katz and Mazur 1985, 5.5.3]). *If $E \to S$ is a generalized elliptic curve for which n divides the number of irreducible components of each degenerate geometric fiber and α is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure on E, then the functor*

$$S' \mapsto \big\{\textit{Drinfeld } (\mathbb{Z}/n\mathbb{Z})^2\textit{-structures } \beta \textit{on } E^{\mathrm{sm}}_{S'}[n] \quad \textit{such that } \beta|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}} = \alpha_{S'}\big\}$$

is representable by a finite locally free S-scheme of rank $n \cdot \phi(n)$.

(c) (Compare with [Katz and Mazur 1985, 5.5.3]). *If the number of irreducible components of each degenerate geometric fiber of $E \to S$ is divisible by n and, for some S-subgroup $G \subset E$,*

$$\alpha : \mathbb{Z}/n\mathbb{Z} \to E^{\mathrm{sm}}(S) \quad \textit{and} \quad \beta : \mathbb{Z}/n\mathbb{Z} \to (E/G)^{\mathrm{sm}}(S)$$

are Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structures on G and on $E^{\mathrm{sm}}[n]/G$, respectively, then the functor

$$S' \mapsto \big\{\textit{Drinfeld } (\mathbb{Z}/n\mathbb{Z})^2\textit{-structures } \gamma \textit{ on } E^{\mathrm{sm}}_{S'}[n] \textit{ such that } \alpha_{S'} = \gamma|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}} \quad \textit{and} \quad \beta_{S'} = \gamma|_{\{0\}\times\mathbb{Z}/n\mathbb{Z}} : \mathbb{Z}/n\mathbb{Z} \to (E/G)^{\mathrm{sm}}(S')\big\}$$

is representable by a finite locally free S-scheme of rank n.

(d) *Set* $d := \gcd(n, m)$ *and* $N := \operatorname{lcm}(n, m)$. *If the number of irreducible components of each degenerate geometric fiber of* $E \to S$ *is divisible by* N *and* α *and* β *are, respectively, Drinfeld* $\mathbb{Z}/n\mathbb{Z}$*- and* $\mathbb{Z}/m\mathbb{Z}$*-structures on* E *such that*

$$\alpha|_{(\mathbb{Z}/n\mathbb{Z})[d]} + \beta|_{(\mathbb{Z}/m\mathbb{Z})[d]} : (\mathbb{Z}/d\mathbb{Z})^2 \to E^{\mathrm{sm}}(S)$$

is a Drinfeld $(\mathbb{Z}/d\mathbb{Z})^2$*-structure on* $E^{\mathrm{sm}}[d]$, *then the functor*

$$S' \mapsto \big\{\textit{Drinfeld } (\mathbb{Z}/N\mathbb{Z})^2\textit{-structures } \gamma \textit{ on } E^{\mathrm{sm}}_{S'}[N] \textit{ such that}$$
$$\alpha_{S'} = \gamma|_{(\mathbb{Z}/N\mathbb{Z}\times\{0\})[n]} \quad \textit{and} \quad \beta_{S'} = \gamma|_{(\{0\}\times\mathbb{Z}/N\mathbb{Z})[m]}\big\}$$

is representable by a finite locally free S*-scheme of rank* $N \cdot \phi(N)/(d \cdot \phi(d))$.

Proof. All the functors in question are fppf sheaves, so we may work fppf locally on S. Setting $N := nm$ (resp. $N := n$) in part (a) (resp. in parts (b) and (c)) for notational convenience, we may therefore apply Lemma 4.2.1 to assume that there is an elliptic curve $E' \to S$ with

$$E'[N] \simeq E^{\mathrm{sm}}[N].$$

By [Katz and Mazur 1985, 1.10.6], all the properties and functors under consideration depend solely on the S-scheme $E^{\mathrm{sm}}[N]$ (and its subgroup G in (c)), so we may pass to E' to reduce to the elliptic curve case. This already settles (b) and (c), and in order to also obtain (a) it remains to combine [EGA IV$_2$ 1965, 6.1.5] with [Katz and Mazur 1985, 5.1.1], which ensures that for every $\ell \in \mathbb{Z}_{\geq 1}$, the moduli stack parametrizing Drinfeld $(\mathbb{Z}/\ell\mathbb{Z})^2$-structures on elliptic curves is finite locally free of rank $\#\,\mathrm{GL}_2(\mathbb{Z}/\ell\mathbb{Z})$ over $\mathscr{E}\ell\ell$, étale over $\mathscr{E}\ell\ell_{\mathbb{Z}[1/\ell]}$, and regular.

For the remaining elliptic curve case of (d), we use [Katz and Mazur 1985, 1.7.2] to "factor into prime powers" and reduce to the case when

$$n = p^r \quad \text{and} \quad m = p^s \qquad \text{for some prime } p.$$

Without loss of generality $r \geq s$, so the case $s = 0$ is settled by (b). In the case $s \geq 1$, by Proposition 4.2.5(b) (i.e., by [Katz and Mazur 1985, 5.5.8(1)]), the functor in question is identified with the functor parametrizing $Q \in E(S')$ such that $p^{r-s}Q = \beta_{S'}(1)$. This functor is an $E[p^{r-s}]$-torsor, so it is representable by a finite locally free S-scheme of rank $p^{2(r-s)} = p^r \cdot \phi(p^r)/(p^s \cdot \phi(p^s))$. □

When proving the algebraicity of moduli stacks of Drinfeld structures on generalized elliptic curves we will sometimes rely on the representability of functors parametrizing various such structures on a fixed curve. The key case of this representability is Proposition 4.2.15(a) recorded below — further cases may be deduced from it with the help of Proposition 4.2.7(a). It will be important to have such representability when the structures being parametrized are assumed to be ample, so we first review the notion of ampleness.

Definition 4.2.13. A finite locally free S-subgroup $G \subset E^{\mathrm{sm}}$ of a generalized elliptic curve $E \to S$ is *ample* if G is S-ample as a relative effective Cartier divisor on E, equivalently, if G meets every irreducible component of every geometric fiber of $E \to S$. For a finite abelian group A, a Drinfeld A-structure α on E is *ample* if the S-subgroup $D_\alpha := \sum_{a \in A} [\alpha(a)] \subset E^{\mathrm{sm}}$ is ample.

Remark 4.2.14. The role of ampleness of α in the study of various stacks that classify Drinfeld A-structures on generalized elliptic curves is twofold: it facilitates descent considerations (e.g., the ones in the definition of a stack) by endowing $E \to S$ with a canonical S-ample line bundle $\mathscr{O}_E(D_\alpha)$, and it also kills undesirable automorphisms that would hinder the representability of various "forget the level" contraction morphisms (e.g., if α is ample and S is a geometric point, then one sees from Lemma 2.1.6 that only the identity automorphism of (E, α) fixes $(E^{\mathrm{sm}})^0$).

Proposition 4.2.15. *Let $E \to S$ be a generalized elliptic curve, let S' be a variable S-scheme, and recall the notation G_d and $[G + H]$ introduced in* Definition 4.2.8 *and* Proposition 4.2.11(a).

(a) *Fix $n, m \in \mathbb{Z}_{\geq 1}$, and set $d := \gcd(n, m)$ and $N := \operatorname{lcm}(n, m)$. The functor*

$$\mathcal{F} : S' \mapsto \big\{\text{cyclic } S'\text{-subgroups } G, H \subset E^{\mathrm{sm}}_{S'} \\ \text{of orders } n \text{ and } m \text{ with} [G_d + H_d] = E^{\mathrm{sm}}_{S'}[d]\big\}$$

(*resp. its analogue which, in addition, requires $[G + H]$ to be ample*) *is representable by a finitely presented, separated, quasifinite, flat S-scheme F that is étale if nm is invertible on S. If N divides the number of irreducible components of each degenerate geometric fiber of $E \to S$, then F* (defined without the ampleness requirement) *is finite locally free of rank*

$$\# \mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z}) \cdot \frac{d \cdot \phi(d)}{N \cdot \phi(N) \cdot \phi(n) \cdot \phi(m)}$$

over S.

(b) (Compare with [Katz and Mazur 1985, 6.8.1]). *For every $n \in \mathbb{Z}_{\geq 1}$, the functor*

$$\mathcal{I} : S' \mapsto \big\{\text{finite locally free } S'\text{-subgroups } G \subset E^{\mathrm{sm}}_{S'} \text{ of rank } n\big\}$$

(*resp. its analogue which, in addition, requires G to be ample*) *is representable by a finitely presented, separated, quasifinite, flat S-scheme I that is étale if n is invertible on S. If n divides the number of irreducible components of each degenerate geometric fiber of $E \to S$, then I* (defined without the ampleness requirement) *is finite locally free over S and its rank is constant and equals the number of subgroups of $(\mathbb{Z}/n\mathbb{Z})^2$ of order n.*

Remark 4.2.16. In (a), an important special case is $m = 1$, when $\mathcal{F}$ parametrizes cyclic subgroups of order n. In (b), due to Corollary 2.2.7(b), $\mathcal{I}$ parametrizes n-isogenies with source E.

Proof of Proposition 4.2.15. Due to [EGA IV$_3$ 1966, 9.6.4] and limit arguments that reduce to a Noetherian base, the additional ampleness requirement cuts out quasicompact open subfunctors of $\mathcal{F}$ and $\mathcal{I}$, so the ampleness variant of the claims will follow once we establish the rest.

To ease notation, we set $N := n$ in (b). By [EGA IV$_4$ 1967, 18.12.12], quasifinite and separated morphisms are quasiaffine, so effectivity of fppf descent for relatively quasiaffine schemes enables us to work fppf locally on S. We may therefore apply Corollary 3.2.6 to assume that E^{sm} is an open S-subgroup of the smooth locus of another generalized elliptic curve $E' \to S$ for which N divides the number of irreducible components of each degenerate geometric fiber. The functor $\mathcal{F}$ (resp. $\mathcal{I}$) is an open subfunctor of the corresponding functor $\mathcal{F}'$ (resp. $\mathcal{I}'$) for E', and the open immersion $\mathcal{F} \subset \mathcal{F}'$ (resp. $\mathcal{I} \subset \mathcal{I}'$) is quasicompact due to limit arguments, so it suffices to settle the claims for E' in place of E. We may then use Lemma 4.2.1 to assume that there is an elliptic curve $E'' \to S$ with

$$E''[N] \simeq E'^{\mathrm{sm}}[N].$$

Since E' and E'' give isomorphic functors $\mathcal{I}$, this reduces (b) to its elliptic curve case [Katz and Mazur 1985, 6.8.1].

For (a), we let $\mathcal{F}'_N$ denote the functor that parametrizes Drinfeld $(\mathbb{Z}/N\mathbb{Z})^2$-structures α on $E'^{\mathrm{sm}}_{S'}[N]$. By Proposition 4.2.12(a), $\mathcal{F}'_N$ is representable by a finite locally free S-scheme of rank $\#\mathrm{GL}_2(\mathbb{Z}/N\mathbb{Z})$ that is étale if N is invertible on S. By Proposition 4.2.5(a) and (c), there is a well-defined morphism

$$\mathcal{F}'_N \to \mathcal{F}'$$

that sends α to the pair of subgroups on which $\alpha|_{(\mathbb{Z}/N\mathbb{Z}\times\{0\})[n]}$ and $\alpha|_{(\{0\}\times\mathbb{Z}/N\mathbb{Z})[m]}$ are Drinfeld $\mathbb{Z}/n\mathbb{Z}$- and $\mathbb{Z}/m\mathbb{Z}$-structures, respectively. By Proposition 4.2.7(a) and Proposition 4.2.12(d), $\mathcal{F}'_N \to \mathcal{F}'$ is representable by schemes and finite locally free of rank

$$\frac{N\cdot\phi(N)\cdot\phi(n)\cdot\phi(m)}{d\cdot\phi(d)}.$$

Therefore, the desired claim about $\mathcal{F}'$ follows from [SGA 3$_{\mathrm{I(new)}}$ 2011, V, 4.1] (combined with [EGA IV$_2$ 1965, 2.2.11(ii); EGA IV$_4$ 1967, 17.7.5 and 17.7.7]). □

4.3. A modular description of $\mathscr{X}_{\Gamma(n)}$

The main goal of this section is to give a modular description of $\mathscr{X}_{\Gamma(n)}$, where $n \in \mathbb{Z}_{\geq 1}$ and

$$\Gamma(n) := \mathrm{Ker}(\mathrm{GL}_2(\widehat{\mathbb{Z}}) \twoheadrightarrow \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z}))$$

(see Section 4.1.2 for the definition of $\mathscr{X}_{\Gamma(n)}$; see also Section 1.9). This description and the proof of its correctness follow already from the results of [Conrad 2007], which also show the regularity and other properties of $\mathscr{X}_{\Gamma(n)}$. We reprove both the description and some of the properties of $\mathscr{X}_{\Gamma(n)}$ by exploiting a direct relationship with the compactification $\overline{\mathscr{E}\ell\ell}_n$ studied in Chapter 3. The resulting proofs seem more direct and more versatile — for instance, we will see in Section 4.4 that virtually the same strategy also handles the $H = \Gamma_1(n)$ case, which is significantly more complex for the methods of [op. cit.]. Another pleasant feature of this approach is that it eliminates the crutch of analytic uniformizations — for instance, in the proof of the "ampleness" of $\mathscr{X}(n)^\infty \subset \mathscr{X}(n)$ given in Proposition 4.3.2(b), the only input that is needed from the theory over $\mathbb{C}$ is the fact that the coarse moduli space of $(\overline{\mathscr{E}\ell\ell}_1)_\mathbb{C}$ is $\mathbb{P}^1_\mathbb{C}$ (this comes in through our reliance on [Deligne and Rapoport 1973, VI.1.1] in the proof of Proposition 3.3.2).

We begin by giving the definition of the modular stack $\mathscr{X}(n)$ that classifies generalized elliptic curves endowed with an ample level n structure, and proceed to establish enough of its properties to arrive at the identification $\mathscr{X}(n) = \mathscr{X}_{\Gamma(n)}$.

4.3.1. The stack $\mathscr{X}(n)$. This is the $\mathbb{Z}$-stack that, for a fixed $n \in \mathbb{Z}_{\geq 1}$, and for variable schemes S, parametrizes the pairs

$$\big(E \xrightarrow{\pi} S,\ \alpha : (\mathbb{Z}/n\mathbb{Z})^2 \to E^{\mathrm{sm}}(S)\big)$$

consisting of a generalized elliptic curve $E \xrightarrow{\pi} S$ whose degenerate geometric fibers are n-gons and an (automatically ample) Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure α on $E^{\mathrm{sm}}[n]$. The notation agrees with that of Section 4.1.1 because $\mathscr{X}(1) = \overline{\mathscr{E}\ell\ell}_1$. We let

$$\mathscr{X}(n)^\infty \subset \mathscr{X}(n) \quad \text{and} \quad \mathscr{Y}(n) \subset \mathscr{X}(n)$$

be the closed substack cut out by the degeneracy loci $S^{\infty,\pi}$ and its open complement (the elliptic curve locus), respectively. Due to Remark 4.2.3, for variable $\mathbb{Z}[1/n]$-schemes S, the base change $\mathscr{Y}(n)_{\mathbb{Z}[1/n]}$ parametrizes elliptic curves $E \to S$ equipped with an S-isomorphism $\alpha : \underline{(\mathbb{Z}/n\mathbb{Z})}^2_S \xrightarrow{\sim} E[n]$.

The results of Section 4.2 lead to the following direct relationship between $\mathscr{X}(n)$ and $\overline{\mathscr{E}\ell\ell}_n$.

Proposition 4.3.2. *Consider the $\mathbb{Z}$-morphism $f : \mathscr{X}(n) \to \overline{\mathscr{E}\ell\ell}_n$ that forgets α.*

(a) *The morphism f is representable, finite, and locally free of degree equal to* $\#\,\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$; *moreover, f is étale over* $\mathbb{Z}[1/n]$. *In particular, $\mathscr{X}(n)$ is a Cohen–Macaulay, reduced algebraic $\mathbb{Z}$-stack that is proper, flat, and of relative dimension* 1 *over* $\operatorname{Spec}\mathbb{Z}$ *at every point; moreover, $\mathscr{X}(n)$ is smooth over* $\mathbb{Z}[1/n]$.

(b) *The closed substack $\mathscr{X}(n)^\infty \subset \mathscr{X}(n)$ is the preimage of the closed substack $\overline{\mathscr{E}\ell\ell}_n^\infty \subset \overline{\mathscr{E}\ell\ell}_n$ and is a reduced relative effective Cartier divisor over* $\operatorname{Spec}\mathbb{Z}$ *that* *meets every irreducible component of every geometric fiber of* $\mathscr{X}(n) \to \operatorname{Spec} \mathbb{Z}$ *and is smooth over* $\mathbb{Z}[1/n]$.

Proof. (a) The asserted properties of f follow from Proposition 4.2.12(a), and those of $\mathscr{X}(n)$, other than the reducedness, then result from Theorem 3.1.6(a) (and [EGA IV$_2$ 1965, 6.4.2] for the Cohen–Macaulay aspect). By [EGA IV$_2$ 1965, 5.8.5], the reducedness amounts to the combination of (R$_0$) and (S$_1$). The Cohen–Macaulay aspect implies (S$_1$), whereas (R$_0$) follows from the $\mathbb{Z}$-flatness and $\mathbb{Z}[1/n]$-smoothness.

(b) In the given moduli interpretation, the map $\mathscr{X}(n) \to \overline{\mathscr{E}\ell\ell}_n$ does not change the underlying generalized elliptic curves, so an S-point of $\mathscr{X}(n)$ factors through $\mathscr{X}(n)^\infty$ if and only if its image in $\overline{\mathscr{E}\ell\ell}_n$ factors through $\overline{\mathscr{E}\ell\ell}_n^\infty$. In other words,

$$\mathscr{X}(n)^\infty = \mathscr{X}(n) \times_{\overline{\mathscr{E}\ell\ell}_n} \overline{\mathscr{E}\ell\ell}_n^\infty,$$

as desired. All the remaining claims then follow from (a) and from their counterparts for $\overline{\mathscr{E}\ell\ell}_n$ supplied by Theorem 3.1.6(c)–(d) and Proposition 3.3.2 (for the reducedness of $\mathscr{X}(n)^\infty$ one uses the (R$_0$)+(S$_1$) criterion as in the proof of (a)). $\square$

4.3.3. The contraction morphisms. Due to Proposition 4.2.5(a), the contraction morphism

$$\mathscr{X}(nm) \xrightarrow{c} \mathscr{X}(n) \quad \text{is well defined by} \quad (E, \alpha) \mapsto \big(c_{E^{\mathrm{sm}}[n]}(E), \alpha|_{(\mathbb{Z}/nm\mathbb{Z})^2[n]}\big)$$

(see Convention 4.2.4) for every $n, m \in \mathbb{Z}_{\geq 1}$. This morphism is compatible with its analogue for $\overline{\mathscr{E}\ell\ell}_n$ discussed in Section 3.2.1 in the sense that there is the commutative diagram

$$\begin{tikzcd} \mathscr{X}(nm) \arrow[r, "f_{nm}"] \arrow[d, "c"'] & \overline{\mathscr{E}\ell\ell}_{nm} \arrow[d] \\ \mathscr{X}(n) \arrow[r, "f_n"] & \overline{\mathscr{E}\ell\ell}_n \end{tikzcd}$$

whose horizontal maps forget the level structures α.

Proposition 4.3.4. *For every* $n, m \in \mathbb{Z}_{\geq 1}$, *the contraction* $c : \mathscr{X}(nm) \to \mathscr{X}(n)$ *is representable, finite, and locally free of rank* $\#\mathrm{GL}_2(\mathbb{Z}/nm\mathbb{Z})/\#\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$. *In particular, each* $\mathscr{X}(n)$ *is Deligne–Mumford.*

Proof. Since $\mathscr{X}(1)$ is Deligne–Mumford, the last assertion follows from the rest (applied with $n = 1$). The representability of c by algebraic spaces follows from Lemma 3.2.2(b) and Lemma 2.1.6.

The contraction c inherits properness and finite presentation from

$$\mathscr{X}(nm) \to \operatorname{Spec} \mathbb{Z},$$

and so is quasifinite due to its moduli interpretation. Therefore, by Lemma 3.2.3, the map c is representable by schemes and finite. It remains to prove that c is flat — once this is done, the asserted rank may be read off on the elliptic curve locus by using Proposition 4.3.2(a).

The flatness of the base change

$$\overline{\mathscr{E}\ell\ell}_{nm} \times_{\overline{\mathscr{E}\ell\ell}_n} \mathscr{X}(n) \xrightarrow{a} \mathscr{X}(n)$$

follows from that of $\overline{\mathscr{E}\ell\ell}_{nm} \to \overline{\mathscr{E}\ell\ell}_n$ supplied by Theorem 3.2.4(a). On the other hand,

$$\overline{\mathscr{E}\ell\ell}_{nm} \times_{\overline{\mathscr{E}\ell\ell}_n} \mathscr{X}(n)$$

parametrizes generalized elliptic curves endowed with a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[n]$ subject to the constraint that the degenerate geometric fibers are nm-gons, so the map

$$\mathscr{X}(nm) \xrightarrow{b} \overline{\mathscr{E}\ell\ell}_{nm} \times_{\overline{\mathscr{E}\ell\ell}_n} \mathscr{X}(n)$$

is flat by Proposition 4.2.12(a). In conclusion, the composite $c = a \circ b$ is also flat. □

We are ready for the promised identification $\mathscr{X}(n) = \mathscr{X}_{\Gamma(n)}$.

Theorem 4.3.5. *The Deligne–Mumford stack $\mathscr{X}(n)$ is regular and is identified with the stack $\mathscr{X}_{\Gamma(n)}$ of* Section 4.1.2 (see the proof for the description of the identification).

Proof. By [Katz and Mazur 1985, 5.1.1], the open substack $\mathscr{Y}(n) \subset \mathscr{X}(n)$ is regular. By combining this with the conclusions of Proposition 4.3.2, we see that $\mathscr{X}(n)$ satisfies both (R_1) and (S_2), i.e., is normal. Therefore, due to the conclusions of Proposition 4.3.4, $\mathscr{X}(n)$ is identified with the normalization of $\mathscr{X}(1)$ in $\mathscr{Y}(n)_{\mathbb{Z}[1/n]}$. However, the moduli interpretations of the $\mathscr{Y}(1)$-stacks $\mathscr{Y}(n)_{\mathbb{Z}[1/n]}$ and $\mathscr{Y}_{\Gamma(n)}[1/n]$ coincide (see Sections 4.1.2 and 4.3.1), so $\mathscr{X}(n)$ is identified with the normalization of $\mathscr{X}(1)$ in $\mathscr{Y}_{\Gamma(n)}[1/n]$, i.e., with $\mathscr{X}_{\Gamma(n)}$. To then extend the regularity of $\mathscr{Y}(n)$ supplied by [Katz and Mazur 1985, 5.1.1] to the regularity of the entire $\mathscr{X}(n)$, we recall that it follows from [Deligne and Rapoport 1973, 4.13] that $\mathscr{X}_{\Gamma(n)}$ is regular away from the supersingular points in characteristics dividing n. □

In the sequel we will identify $\mathscr{X}(n)$ and $\mathscr{X}_{\Gamma(n)}$. We conclude the section by recording all the cases in which $\mathscr{X}(n)$ is a scheme (see [Deligne and Rapoport 1973, IV.2.9] for such a result over $\mathbb{Z}[1/n]$).

Proposition 4.3.6. *The stack $\mathscr{X}(n)$ is a* (necessarily projective) *scheme over $\mathbb{Z}$ unless $n = p^s$ or $n = 2p^s$ for some prime p and some $s \in \mathbb{Z}_{\geq 1}$.*

Proof. If $n = p^s$ or $n = 2p^s$, then every supersingular elliptic curve E over $\overline{\mathbb{F}}_p$ equipped with a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure on $E[n]$ has multiplication by -1 as an automorphism, so $\mathscr{X}(n)$ cannot be a scheme. Outside of these cases, $n = n'n''$

for relatively prime $n' \geq 3$ and $n'' \geq 3$, so, due to [Katz and Mazur 1985, 2.7.2(1)] and Lemma 2.1.6, the geometric points of $\mathscr{X}(n)$ have no nontrivial automorphisms, and hence $\mathscr{X}(n)$ is a projective $\mathbb{Z}$-scheme by Lemma 4.1.3. □

4.4. A modular description of $\mathscr{X}_{\Gamma_1(n)}$

The main goal of this section is to give a modular description of $\mathscr{X}_{\Gamma_1(n)}$, where $n \in \mathbb{Z}_{\geq 1}$ and

$$\Gamma_1(n) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix} \right) \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \quad \text{such that} \quad a \equiv 1 \bmod n \quad \text{and} \quad c \equiv 0 \bmod n \right\}$$

(see Section 4.1.2 for the definition of $\mathscr{X}_{\Gamma_1(n)}$; see also Section 1.9). The overall strategy is similar to the case of $\Gamma(n)$ treated in the previous section: through relations with the compactifications $\overline{\mathscr{E}\ell\ell}_m$ we infer enough properties of the stack $\mathscr{X}_1(n)$ that classifies generalized elliptic curves endowed with an ample Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure to arrive at the identification $\mathscr{X}_1(n) = \mathscr{X}_{\Gamma_1(n)}$. As in the case of $\Gamma(n)$, this identification and the finer properties of $\mathscr{X}_1(n)$, such as regularity, follow already from the results of [Conrad 2007], but the alternative proofs given below seem simpler. In particular, when proving the regularity of $\mathscr{X}_1(n)$ we do not use any computations with schemes of $\Gamma_1(n)$-structures on Tate curves or with universal deformation rings, but instead directly deduce such regularity from the regularity of $\mathscr{X}(n)$.

4.4.1. The stack $\mathscr{X}_1(n)$. This is the $\mathbb{Z}$-stack that, for a fixed $n \in \mathbb{Z}_{\geq 1}$ and for variable schemes S, parametrizes the pairs

$$\left(E \xrightarrow{\pi} S,\ \alpha : \mathbb{Z}/n\mathbb{Z} \to E^{\mathrm{sm}}(S)\right)$$

consisting of a generalized elliptic curve $E \xrightarrow{\pi} S$ and an ample Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure α on E. As before, we let

$$\mathscr{X}_1(n)^{\infty} \subset \mathscr{X}_1(n) \quad \text{and} \quad \mathscr{Y}_1(n) \subset \mathscr{X}_1(n)$$

be the closed substack cut out by the degeneracy loci $S^{\infty,\pi}$ and its open complement (the elliptic curve locus), respectively.

For a positive divisor m of n, we let

$$\mathscr{X}_1(n)_{(m)} \subset \mathscr{X}_1(n)$$

be the open substack that classifies those (E, α) for which the degenerate geometric fibers of $E \to S$ are m-gons (the openness follows from Remark 2.1.9), and we set

$$\mathscr{X}_1(n)^{\infty}_{(m)} := \mathscr{X}_1(n)_{(m)} \cap \mathscr{X}_1(n)^{\infty}.$$

When m varies, the open substacks $\mathscr{X}_1(n)_{(m)}$ cover $\mathscr{X}_1(n)$, and we will use them to prove the algebraicity of $\mathscr{X}_1(n)$.

Proposition 4.4.2. *Let $f_{(m)} : \mathscr{X}_1(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_m$ be the $\mathbb{Z}$-morphism that forgets α.*

(a) *The morphism $f_{(m)}$ is representable by schemes, quasifinite, separated, flat, and of finite presentation; moreover, $f_{(m)}$ is étale over $\mathbb{Z}[1/n]$. In particular, $\mathscr{X}_1(n)$ is an algebraic $\mathbb{Z}$-stack with a quasicompact and separated diagonal and is flat, of finite presentation, and of relative dimension* 1 *over* $\operatorname{Spec}\mathbb{Z}$ *at every point; moreover, $\mathscr{X}_1(n)$ is smooth over $\mathbb{Z}[1/n]$.*

(b) *The closed substack $\mathscr{X}_1(n)^{\infty}_{(m)} \subset \mathscr{X}_1(n)_{(m)}$ is the preimage of $\overline{\mathscr{E}\ell\ell}^{\infty}_m \subset \overline{\mathscr{E}\ell\ell}_m$. In particular, $\mathscr{X}_1(n)^{\infty} \subset \mathscr{X}_1(n)$ is a reduced relative effective Cartier divisor over* $\operatorname{Spec}\mathbb{Z}$ *that is smooth over $\mathbb{Z}[1/n]$.*

Proof. (a) The asserted properties of $f_{(m)}$ follow from Proposition 4.2.15(a) and Proposition 4.2.7(a). Since the $\mathscr{X}_1(n)_{(m)}$ cover $\mathscr{X}_1(n)$, the asserted properties of $\mathscr{X}_1(n)$ follow from those of $f_{(m)}$ and from Theorem 3.1.6(a).

(b) For the first assertion, it suffices to observe that in the given moduli interpretation, the map $f_{(m)}$ does not change the underlying generalized elliptic curve. The remaining assertions then follow from the first, (a), and Theorem 3.1.6(c)–(d), using the $(\mathrm{R}_0)+(\mathrm{S}_1)$ criterion together with [EGA IV_2 1965, 6.4.2] to establish the claimed reducedness. □

4.4.3. The relation to $\mathscr{X}(n)$. There is a forgetful contraction morphism

$$g : \mathscr{X}_1(n) \to \mathscr{X}(1),$$

and, due to Proposition 4.2.5(a), also an $\mathscr{X}(1)$-morphism

$$h : \mathscr{X}(n) \to \mathscr{X}_1(n), \quad (E, \alpha) \mapsto \big(c_{\alpha|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}}}(E), \alpha|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}}\big)$$

that contracts E with respect to the unique finite locally free subgroup of E^{sm} on which $\alpha|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}}$ is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure.

We will extract further information about $\mathscr{X}_1(n)$ by studying h. The main difficulty is that h changes E, which makes its key properties, such as flatness, less transparent. To overcome this, we will further exploit the compactifications $\overline{\mathscr{E}\ell\ell}_m$.

Theorem 4.4.4. (a) *The morphism $h : \mathscr{X}(n) \to \mathscr{X}_1(n)$ is representable, finite, and locally free of rank $n \cdot \phi(n)$. In particular, $\mathscr{X}_1(n) \to$* $\operatorname{Spec}\mathbb{Z}$ *is proper, $\mathscr{X}_1(n)$ is regular, and $\mathscr{X}_1(n)^{\infty}$ meets every irreducible component of every geometric $\mathbb{Z}$-fiber of $\mathscr{X}_1(n)$.*

(b) *The contraction $g : \mathscr{X}_1(n) \to \mathscr{X}(1)$ is representable, finite, and locally free of rank $\#\,\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})/(n \cdot \phi(n))$.*

(c) *The stack $\mathscr{X}_1(n)$ is Deligne–Mumford and is identified with the stack $\mathscr{X}_{\Gamma_1(n)}$ of* Section 4.1.2; *more precisely, both $\mathscr{X}_1(n)$ and $\mathscr{X}_{\Gamma_1(n)}$ are the normalizations of $\mathscr{X}(1)$ in $\mathscr{Y}_1(n)_{\mathbb{Z}[1/n]} \cong \mathscr{Y}_{\Gamma_1(n)}[1/n]$.*

Proof. (a) The representability of h by algebraic spaces follows from Lemma 3.2.2(b) and Lemma 2.1.6. Let $\mathscr{X}(n)_{(m)} \subset \mathscr{X}(n)$ be the h-preimage of $\mathscr{X}_1(n)_{(m)}$, let $h_{(m)} : \mathscr{X}(n)_{(m)} \to \mathscr{X}_1(n)_{(m)}$ be the restriction of h, and let $f_{(m)} : \mathscr{X}_1(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_m$ be the forgetful map studied in Proposition 4.4.2. By (3.2.1.2), the composition $f_{(m)} \circ h_{(m)}$ agrees with the composition

$$\mathscr{X}(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_n \xrightarrow{c} \overline{\mathscr{E}\ell\ell}_m$$

in which the first map forgets the Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure. Therefore, the universal property of the fiber product gives the commutative diagram

$$\begin{tikzcd}
\mathscr{X}(n)_{(m)} \arrow[r, "h'"] \arrow[dr, "h_{(m)}"'] & \mathscr{X}_1(n)_{(m)} \times_{\overline{\mathscr{E}\ell\ell}_m} \overline{\mathscr{E}\ell\ell}_n \arrow[r] \arrow[d, "h''"] & \overline{\mathscr{E}\ell\ell}_n \arrow[d, "c"'] \\
& \mathscr{X}_1(n)_{(m)} \arrow[r, "f_{(m)}"] & \overline{\mathscr{E}\ell\ell}_m
\end{tikzcd}$$

in which the square is Cartesian. By Proposition 4.2.12(b), the map h' is representable and finite locally free of rank $n \cdot \phi(n)$. By Theorem 3.2.4(a), the base change h'' of c is proper, flat, and surjective. The representable map $h_{(m)}$ is therefore proper, flat, surjective, and, due to its moduli interpretation, also quasifinite. Since h inherits these properties, we see from Lemma 3.2.3 that h is representable by schemes and finite locally free. Its rank is determined on the elliptic curve locus, so equals $n \cdot \phi(n)$.

The remaining claims follow from the combination of Proposition 4.3.2, Theorem 4.3.5, and [EGA IV$_2$ 1965, 6.5.3(i)], once we establish the $\mathbb{Z}$-separatedness of $\mathscr{X}_1(n)$. For this, since the diagonal map $\Delta_{\mathscr{X}_1(n)/\mathbb{Z}}$ is separated and of finite type by Proposition 4.4.2(a), its properness follows from the commutative diagram

$$\begin{tikzcd}
\mathscr{X}(n) \arrow[r, "\Delta_{\mathscr{X}(n)/\mathbb{Z}}"] \arrow[d, "h"] & \mathscr{X}(n) \times_{\mathbb{Z}} \mathscr{X}(n) \arrow[d, "h \times h"] \\
\mathscr{X}_1(n) \arrow[r, "\Delta_{\mathscr{X}_1(n)/\mathbb{Z}}"] & \mathscr{X}_1(n) \times_{\mathbb{Z}} \mathscr{X}_1(n)
\end{tikzcd}$$

and the properness of $(h \times h) \circ \Delta_{\mathscr{X}(n)/\mathbb{Z}}$.

(b) Since $\mathscr{X}_1(n) \to \operatorname{Spec} \mathbb{Z}$ is proper, g is also proper. Moreover, g is representable by algebraic spaces and quasifinite due to its moduli interpretation, Lemma 3.2.2(b), and Lemma 2.1.6. Thus, due to Lemma 3.2.3, g is representable by schemes and finite. The remaining assertions follow by considering the composite

$$\mathscr{X}(n) \xrightarrow{h} \mathscr{X}_1(n) \xrightarrow{g} \mathscr{X}(1)$$

and combining (a) with Proposition 4.3.4.

(c) Thanks to (b), the Deligne–Mumford property is inherited from $\mathscr{X}(1)$. For the rest, due to the regularity of $\mathscr{X}_1(n)$ and the finiteness of $\mathscr{X}_1(n) \to \mathscr{X}(1)$, we need to identify the stack $\mathscr{Y}_1(n)_{\mathbb{Z}[1/n]}$ with the stack $\mathscr{Y}_{\Gamma_1(n)}[1/n]$ that, for variable $\mathbb{Z}[1/n]$-schemes S, parametrizes pairs consisting of an elliptic curve $E \to S$ and an S-point of the finite étale S-scheme

$$\left\{\left(\begin{smallmatrix} 1 & * \\ 0 & * \end{smallmatrix}\right) \subset \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})\right\} \setminus \mathrm{Isom}\big(E[n], (\mathbb{Z}/n\mathbb{Z})^2\big).$$

The datum of such an S-point amounts to the datum of an isomorphism between $\mathbb{Z}/n\mathbb{Z}$ and a subgroup of E, so the sought identification results from Remark 4.2.3. $\square$

4.5. An axiomatic criterion for recognizing correctness of a modular description

The arguments of the preceding section that supplied the identification

$$\mathscr{X}_1(n) = \mathscr{X}_{\Gamma_1(n)}$$

and proved the regularity of $\mathscr{X}_{\Gamma_1(n)}$ illustrate a general method that will similarly handle more complicated cases in the sequel. Therefore, in order to avoid repetitiveness, we wish to present the following axiomatic result that ensures that for any open subgroup $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ any "good enough" candidate stack $\mathscr{X}'_H$ agrees with the $\mathscr{X}_H$ defined in Section 4.1.2 and that $\mathscr{X}_H$ is automatically regular whenever such a good candidate is present. Of course, the main difficulty of this approach to the regularity of $\mathscr{X}_H$ lies in finding a suitable $\mathscr{X}'_H$. In all the cases presented in the sequel, the candidate $\mathscr{X}'_H$ will be defined by a modular description of its functor of points and Theorem 4.5.1 will act as a criterion for recognizing that this modular description actually yields $\mathscr{X}_H$.

Theorem 4.5.1. *Let $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ be an open subgroup, let $n \in \mathbb{Z}_{\geq 1}$ be such that $\Gamma(n) \subset H$, and let $\mathscr{X}'_H$ be a $\mathbb{Z}$-stack.*

(a) *If there is a cover*

$$\mathscr{X}'_H = \textstyle\bigcup_{m|n} (\mathscr{X}'_H)_{(m)} \quad \text{by open substacks} \quad (\mathscr{X}'_H)_{(m)} \subset \mathscr{X}'_H$$

each of which admits a representable by algebraic spaces, separated, finite type morphism

$$(\mathscr{X}'_H)_{(m)} \to \overline{\mathscr{E}\ell\ell}_{d(m)}$$

for some $d(m) \in \mathbb{Z}_{\geq 1}$, then $\mathscr{X}'_H$ is algebraic, has a quasicompact and separated diagonal $\Delta_{\mathscr{X}'_H/\mathbb{Z}}$, and is of finite type over $\mathbb{Z}$.

(b) *If $\mathscr{X}'_H$ is algebraic, has a quasicompact and separated diagonal, is of finite type over $\mathbb{Z}$, and*

(1) *there is a proper, flat, and surjective* $\mathbb{Z}$*-morphism* $\mathscr{X}(n) \xrightarrow{h} \mathscr{X}'_H$,

then $\mathscr{X}'_H$ *is regular*, $\mathscr{X}'_H \to \operatorname{Spec} \mathbb{Z}$ *is a proper, flat surjection, and* $(\mathscr{X}'_H)_{\mathbb{Z}[1/n]}$ *is* $\mathbb{Z}[1/n]$*-smooth.*

(c) *If* $\mathscr{X}'_H$ *is algebraic*, $\mathbb{Z}$*-proper, and satisfies* (1) *together with*

(2) *there is a representable by algebraic spaces* $\mathbb{Z}$*-morphism* $\mathscr{X}'_H \xrightarrow{g} \mathscr{X}(1)$ *that over* $\mathbb{Z}[1/n]$ *is identified with the morphism* $\mathscr{Y}_H[1/n] \to \mathscr{Y}(1)_{\mathbb{Z}[1/n]}$ *of* Section 4.1.2, *and*

(3) *the composition* $g \circ h : \mathscr{X}(n) \to \mathscr{X}(1)$ *is identified with the contraction of* Section 4.3.3,

then $\mathscr{X}'_H$ *is Deligne–Mumford and the morphism g induces the identification*

$$\mathscr{X}_H = \mathscr{X}'_H;$$

more precisely, then both $\mathscr{X}_H$ *and* $\mathscr{X}'_H$ *are the normalizations of* $\mathscr{X}(1)$ *in* $\mathscr{Y}_H[1/n]$.

Remark 4.5.2. The flatness of h is one of the most stringent requirements. For the $\mathscr{X}'_H$ that we will construct this flatness will be supplied by the results of Katz and Mazur through congruences with elliptic curves (see Proposition 4.2.12(b) and the proof of Theorem 4.4.4(a) for an example).

Proof of Theorem 4.5.1. (a) The algebraicity of each $(\mathscr{X}'_H)_{(m)}$ follows from that of $\overline{\mathscr{E}\ell\ell}_{d(m)}$ supplied by Theorem 3.1.6(a) (see [Laumon and Moret-Bailly 2000, 4.5(ii)]). This suffices for the algebraicity of $\mathscr{X}'_H$ because the diagonal $\Delta_{\mathscr{X}'_H/\mathbb{Z}}$ factors as the composition

$$\mathscr{X}'_H = \textstyle\bigcup_{m|n} (\mathscr{X}'_H)_{(m)} \to \bigcup_{m|n} (\mathscr{X}'_H)_{(m)} \times_{\mathbb{Z}} (\mathscr{X}'_H)_{(m)} \subset \mathscr{X}'_H \times_{\mathbb{Z}} \mathscr{X}'_H$$

in which the inclusion is representable by open immersions. Since the inclusion is also quasicompact and each $(\mathscr{X}'_H)_{(m)}$ is separated over $\mathbb{Z}$, i.e., each $\Delta_{(\mathscr{X}'_H)_{(m)}/\mathbb{Z}}$ is proper, it also follows that $\Delta_{\mathscr{X}'_H/\mathbb{Z}}$ is quasicompact and separated.

(b) In the commutative diagram

$$\begin{tikzcd} \mathscr{X}(n) \arrow[r, "\Delta_{\mathscr{X}(n)/\mathbb{Z}}"] \arrow[d, "h"] & \mathscr{X}(n) \times_{\mathbb{Z}} \mathscr{X}(n) \arrow[d, "h \times h"] \\ \mathscr{X}'_H \arrow[r, "\Delta_{\mathscr{X}'_H/\mathbb{Z}}"] & \mathscr{X}'_H \times_{\mathbb{Z}} \mathscr{X}'_H \end{tikzcd}$$

the composite $(h \times h) \circ \Delta_{\mathscr{X}(n)/\mathbb{Z}}$ is proper, $\Delta_{\mathscr{X}'_H/\mathbb{Z}}$ is separated and of finite type, and h is surjective, so $\Delta_{\mathscr{X}'_H/\mathbb{Z}}$ is proper. In other words, $\mathscr{X}'_H \to \operatorname{Spec} \mathbb{Z}$ is separated, so $\mathscr{X}'_H$ inherits $\mathbb{Z}$-properness from $\mathscr{X}(n)$. Due to the flatness and surjectivity of h, the flatness, regularity, and smoothness aspects for $\mathscr{X}'_H$ follow from the corresponding aspects for $\mathscr{X}(n)$ supplied by Proposition 4.3.2(a) and Theorem 4.3.5.

(c) The Deligne–Mumford property follows from the representability of g. The map g inherits properness from $\mathscr{X}'_H \to \operatorname{Spec}\mathbb{Z}$ and quasifiniteness from $g \circ h$, so g is finite by Lemma 3.2.3. Moreover, $\mathscr{X}'_H$ is normal by (b), so, due to the requirement (2), g identifies $\mathscr{X}'_H$ with the normalization of $\mathscr{X}(1)$ with respect to $\mathscr{Y}_H[1/n] \to \mathscr{Y}(1)_{\mathbb{Z}[1/n]}$. On the other hand, by definition, this normalization is $\mathscr{X}_H$ (see Section 4.1.2). □

Example 4.5.3. Theorem 4.5.1 is useful for proving that "obvious" candidate modular descriptions for various mixtures of standard moduli problems are correct. When treating "mixture situations," one cannot simply "reduce to individual constituents" via fiber products (unlike on the elliptic curve locus): such "reductions" fail already in situations where no mixtures are involved, for instance,

$$\mathscr{X}(15) \ncong \mathscr{X}(3) \times_{\mathscr{X}(1)} \mathscr{X}(5), \quad \text{even though} \quad \mathscr{Y}(15) \cong \mathscr{Y}(3) \times_{\mathscr{Y}(1)} \mathscr{Y}(5),$$

as one sees by inspecting the ramification at the cusps

$$\big(\text{e.g.,} \quad \mathbb{C}[\![q^{\frac{1}{15}}]\!] \ncong \mathbb{C}[\![q^{\frac{1}{3}}]\!] \otimes_{\mathbb{C}[\![q]\!]} \mathbb{C}[\![q^{\frac{1}{5}}]\!]\big).$$

The concrete example of a "mixture situation" for which we wish to illustrate Theorem 4.5.1 has

$$H = \Gamma(d) \cap \Gamma_1(\ell) \quad \text{with coprime } d, \ell \in \mathbb{Z}_{\geq 1}.$$

For this H, due to the factorizations of Drinfeld structures discussed in [Katz and Mazur 1985, 1.7.2], the "obvious" candidate $\mathscr{X}'_H$ is the stack that, for variable schemes S, parametrizes ample Drinfeld $((\mathbb{Z}/d\mathbb{Z})^2 \times \mathbb{Z}/\ell\mathbb{Z})$-structures α on generalized elliptic curves $E \to S$ subject to the requirement that $\alpha|_{(\mathbb{Z}/d\mathbb{Z})^2 \times \{0\}}$ is a Drinfeld $(\mathbb{Z}/d\mathbb{Z})^2$-structure on $E^{\mathrm{sm}}[d]$ (so d divides the number of irreducible components of each degenerate geometric fiber of $E \to S$).

For this $\mathscr{X}'_H$, we let the maps h and g in Theorem 4.5.1 be the forgetful contractions with $n = d\ell$ and let

$$(\mathscr{X}'_H)_{(m)} \subset \mathscr{X}'_H$$

be the open substack parametrizing those $E \to S$ whose degenerate geometric fibers are m-gons. The requirements of Theorem 4.5.1(a) are met due to [Katz and Mazur 1985, 1.7.2] and Propositions 4.2.5(a), 4.2.7(a), and 4.2.15(a) (with $(n, m) = (d\ell, d)$ in the latter). The requirement (b)(1) is checked with the help of a diagram analogous to the one in the proof of Theorem 4.4.4(a), the key point being that the induced map

$$\mathscr{X}(n)_{(m)} \to (\mathscr{X}'_H)_{(m)} \times_{\overline{\mathscr{E}\ell\ell}_m} \overline{\mathscr{E}\ell\ell}_n$$

from the h-preimage $\mathscr{X}(n)_{(m)}$ of $(\mathscr{X}'_H)_{(m)}$ is finite locally free of rank $\ell \cdot \phi(\ell)$ due to Proposition 4.2.12(b). The requirement (c)(2) is checked as in the proof of Theorem 4.4.4(c) by using the fact that the image of H in $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ is the pointwise stabilizer of $(\mathbb{Z}/d\mathbb{Z})^2 \times \mathbb{Z}/\ell\mathbb{Z}$ in $(\mathbb{Z}/n\mathbb{Z})^2$. Finally, the requirement (c)(3) follows from the definitions of g and h.

In conclusion,

$$\mathscr{X}'_H = \mathscr{X}_{\Gamma(d)\cap\Gamma_1(\ell)}$$

and $\mathscr{X}_{\Gamma(d)\cap\Gamma_1(\ell)}$ is regular (such regularity at the cusps is not an automatic consequence of the regularity of $\mathscr{X}_{\Gamma(d)}$ and $\mathscr{X}_{\Gamma_1(\ell)}$).

4.6. A modular description of $\mathscr{X}_{\Gamma_1(n;\,n')}$ and $\mathscr{X}_{\Gamma_0(n;\,n')}$ for suitable n and n'

Let n and n' be positive integers, and let

$$\Gamma_1(n;\,n') \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$$

be the preimage of the subgroup of $\mathrm{GL}_2(\mathbb{Z}/nn'\mathbb{Z})$ that stabilizes the subgroup $\{0\} \times (\mathbb{Z}/nn'\mathbb{Z})[n']$ in $(\mathbb{Z}/nn'\mathbb{Z})^2$ and that fixes $(\mathbb{Z}/nn'\mathbb{Z})[n] \times \{0\}$ pointwise. Our goal is to prove that the "obvious" candidate modular description for $\mathscr{X}_{\Gamma_1(n;n')}$ presented in Section 4.6.1 is correct under the assumption that

$$\mathrm{ord}_p(n') \leq \mathrm{ord}_p(n) + 1$$

for every prime p. The importance of $\mathscr{X}_{\Gamma_1(n;\,n')}$ stems from its role in defining Hecke correspondences for $\mathscr{X}_1(n)$ (see Section 4.7), but there also are the following reasons for treating $H = \Gamma_1(n;\,n')$.

- The techniques used below to study $\mathscr{X}_{\Gamma_1(n;\,n')}$ simultaneously expose properties of the stack $\mathscr{X}_0(n)^{\mathrm{naive}}$ that parametrizes generalized elliptic curves equipped with an ample cyclic subgroup of order n. Although in general $\mathscr{X}_0(n)^{\mathrm{naive}}$ does not agree with $\mathscr{X}_{\Gamma_0(n)}$, its properties will nevertheless be crucial for the study of $\mathscr{X}_{\Gamma_0(n)}$ in Chapter 5.
- Under the additional assumption that $\mathrm{ord}_p(n') \leq \mathrm{ord}_p(n)$ for all $p \mid \gcd(n, n')$, the correctness of the candidate modular description of $\mathscr{X}_{\Gamma_1(n;\,n')}$ also follows from the results of [Conrad 2007] but it seems worthwhile to simplify the proofs of [op. cit.] with the help of the general Theorem 4.5.1. In fact, Conrad does not assume that $\mathrm{ord}_p(n') \leq 1$ for $p \nmid n$, but outside this case the forgetful contraction morphism from the algebraic stack $\mathcal{M}_{\Gamma_1(n;\,n')}$ that he constructs in *op. cit.* to $\mathscr{X}(1)$ is not representable (even over $\mathbb{C}$), so $\mathcal{M}_{\Gamma_1(n;\,n')}$ cannot agree with $\mathscr{X}_{\Gamma_1(n;\,n')}$ (a related pathology is that $\mathcal{M}_{\Gamma_1(n;\,n')}$ is not Deligne–Mumford in characteristics $p \nmid n$ with $p^2 \mid n'$).

In order to also recover and generalize the results of [Conrad 2007] in the cases when $\mathrm{ord}_p(n') > 1$ for some prime $p \nmid n$, we initially drop *all* requirements on n and n', define a certain stack $\mathscr{X}_1(n;\,n')$ that agrees with the stack $\mathcal{M}_{\Gamma_1(n;\,n')}$ considered in *op. cit.*(in the cases in which $\mathcal{M}_{\Gamma_1(n;\,n')}$ was defined), prove that $\mathscr{X}_1(n; n')$ is algebraic, $\mathbb{Z}$-proper, and regular (among other properties), and only then impose assumptions on n and n' in order to arrive at the agreement with $\mathscr{X}_{\Gamma_1(n;\,n')}$.

4.6.1. The stack $\mathscr{X}_1(n; n')$. This is the $\mathbb{Z}$-stack that, for fixed $n, n' \in \mathbb{Z}_{\geq 1}$ with $d := \gcd(n, n')$ and for variable schemes S, parametrizes the triples

$$\big(E \xrightarrow{\pi} S,\ \alpha : \mathbb{Z}/n\mathbb{Z} \to E^{\mathrm{sm}}(S),\ H\big)$$

consisting of a generalized elliptic curve $E \xrightarrow{\pi} S$, a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure α on some S-subgroup $G \subset E^{\mathrm{sm}}$, and a cyclic S-subgroup $H \subset E^{\mathrm{sm}}$ of order n' subject to the requirements that

$$[G_d + H_d] = E^{\mathrm{sm}}[d] \quad \text{and} \quad [G + H] \text{ is ample} \tag{4.6.1.1}$$

(we implicitly use Definition 4.2.8 and Proposition 4.2.11(a) to make sense of $[G_d + H_d]$ and $[G + H]$). The effectivity of descent needed for $\mathscr{X}_1(n; n')$ to be a stack is ensured by the ampleness of $[G+H]$ as in Remark 4.2.14. The requirement $[G_d + H_d] = E^{\mathrm{sm}}[d]$ implies that the number of irreducible components of each degenerate geometric fiber of E is divisible by d, so Proposition 4.2.11(b) ensures that $[G + H]$ is a finite locally free S-subgroup of E^{sm} of rank nn' that is killed by $\mathrm{lcm}(n, n')$.

We let

$$\mathscr{X}_1(n; n')^{\infty} \subset \mathscr{X}_1(n; n') \quad \text{and} \quad \mathscr{Y}_1(n; n') \subset \mathscr{X}_1(n; n')$$

be the closed substack cut out by the degeneracy loci $S^{\infty,\pi}$ and its open complement (the elliptic curve locus), respectively. Similarly to the case of $\mathscr{X}_1(n)$ (discussed in Section 4.4.1), for every positive divisor m of $\mathrm{lcm}(n, n')$, we let

$$\mathscr{X}_1(n; n')_{(m)} \subset \mathscr{X}_1(n; n')$$

be the open substack over which the degenerate geometric fibers of E are m-gons.

4.6.2. Variants $\widetilde{\mathscr{X}_1}(n; n')$ and $\mathscr{X}_0(n; n')$. Slight modifications of the definition of $\mathscr{X}_1(n; n')$ give the following related stacks:

- the stack $\widetilde{\mathscr{X}_1}(n; n')$ obtained by replacing the datum H by the datum of a Drinfeld $\mathbb{Z}/n'\mathbb{Z}$-structure β on some S-subgroup $H \subset E^{\mathrm{sm}}$ subject to (4.6.1.1);
- the stack $\mathscr{X}_0(n; n')$ obtained by replacing the datum α by the datum of a cyclic S-subgroup $G \subset E^{\mathrm{sm}}$ of order n subject to (4.6.1.1).

Due to Proposition 4.2.7(a), the forgetful maps

$$\widetilde{\mathscr{X}_1}(n; n') \to \mathscr{X}_1(n; n') \quad \text{and} \quad \mathscr{X}_1(n; n') \to \mathscr{X}_0(n; n') \tag{4.6.2.1}$$

are representable by schemes, finite locally free of ranks $\phi(n')$ and $\phi(n)$, respectively, and, over $\mathbb{Z}[1/n']$ and $\mathbb{Z}[1/n]$, respectively, étale. As before, for every positive divisor m of $\mathrm{lcm}(n, n')$ we let

$$\widetilde{\mathscr{X}_1}(n; n')_{(m)} \subset \widetilde{\mathscr{X}_1}(n; n') \quad \text{and} \quad \mathscr{X}_0(n; n')_{(m)} \subset \mathscr{X}_0(n; n')$$

be the open substacks over which the degenerate geometric fibers of E are m-gons, let

$$\widetilde{\mathscr{X}_1}(n; n')^{\infty} \subset \widetilde{\mathscr{X}_1}(n; n') \quad \text{and} \quad \mathscr{X}_0(n; n')^{\infty} \subset \mathscr{X}_0(n; n')$$

be the degeneracy loci, and let

$$\widetilde{\mathscr{Y}_1}(n; n') \subset \widetilde{\mathscr{X}_1}(n; n') \quad \text{and} \quad \mathscr{Y}_0(n; n') \subset \mathscr{X}_0(n; n')$$

be the elliptic curve loci.

For suitably constrained n and n', the stacks $\widetilde{\mathscr{X}_1}(n; n')$ and $\mathscr{X}_0(n; n')$ were also considered in [Conrad 2007] (in the notation $\mathcal{M}_{\widetilde{\Gamma}_1(N;\, n)}$ and $\mathcal{M}_{\Gamma_0(N;\, n)}$). There $\widetilde{\mathscr{X}_1}(n; n')$ was often used as an intermediary in the proofs of the properties of $\mathscr{X}_1(n; n')$, whereas $\mathscr{X}_0(n; n')$ was mentioned on page 273 in relation to modifications that one needs to make to the method of [op. cit.] to also construct Hecke correspondences for $\mathscr{X}_0(n)$. We will see below that the proofs of the properties of $\mathscr{X}_1(n; n')$ will also prove the corresponding properties of $\widetilde{\mathscr{X}_1}(n; n')$ and $\mathscr{X}_0(n; n')$.

4.6.3. Contraction maps from $\mathscr{X}(nn')$. There is a forgetful contraction map

$$\mathscr{X}(nn') \to \widetilde{\mathscr{X}_1}(n; n') \tag{4.6.3.1}$$

that sends a Drinfeld $(\mathbb{Z}/nn'\mathbb{Z})^2$-structure γ to

$$\alpha := \gamma|_{(\mathbb{Z}/nn'\mathbb{Z})[n] \times \{0\}} \quad \text{and} \quad \beta := \gamma|_{\{0\} \times (\mathbb{Z}/nn'\mathbb{Z})[n']}$$

(see Proposition 4.2.5(a) and (c) and Convention 4.2.4) and contracts the underlying generalized elliptic curve accordingly. Similar forgetful contraction maps

$$\mathscr{X}(nn') \to \mathscr{X}_1(n; n') \quad \text{and} \quad \mathscr{X}(nn') \to \mathscr{X}_0(n; n')$$

are the compositions of (4.6.3.1) with the forgetful maps from (4.6.2.1).

We are ready to address the basic properties of the stack $\mathscr{X}_1(n; n')$ and its variants.

Theorem 4.6.4. *Fix $n, n' \in \mathbb{Z}_{\geq 1}$ and let $\mathscr{X} \in \{\widetilde{\mathscr{X}_1}(n; n'), \mathscr{X}_1(n; n'), \mathscr{X}_0(n; n')\}$.*

(a) *The $\mathbb{Z}$-stack $\mathscr{X}$ is algebraic, regular, proper, flat, and of relative dimension* 1 *over* $\operatorname{Spec} \mathbb{Z}$ *at every point; moreover, $\mathscr{X}$ is smooth over $\mathbb{Z}\big[\frac{1}{nn'}\big]$. The diagonal $\Delta_{\mathscr{X}/\mathbb{Z}}$ is finite.*

(b) *The forgetful contraction map $\mathscr{X}(nn') \to \mathscr{X}$ is representable by schemes and is finite locally free of constant positive rank.*

(c) *The closed substack* $\mathscr{X}^\infty \subset \mathscr{X}$ *is a reduced relative effective Cartier divisor over* $\operatorname{Spec} \mathbb{Z}$ *that meets every irreducible component of every geometric* $\mathbb{Z}$*-fiber of* $\mathscr{X}$ *and is smooth over* $\mathbb{Z}[\frac{1}{nn'}]$.

Proof. (a) By Proposition 4.2.15(a) and the finiteness of the maps (4.6.2.1), for every positive divisor m of $\operatorname{lcm}(n, n')$ the forgetful map $\mathscr{X}_{(m)} \to \overline{\mathscr{E}\ell\ell}_m$ is representable, separated, and of finite type, so, by Theorem 4.5.1(a), $\mathscr{X}$ is algebraic and has a quasicompact and separated diagonal.

Except for the relative dimension and the diagonal aspects, the rest of the claim follows from Theorem 4.5.1(b) once we prove that the forgetful contraction $\mathscr{X}(nn') \to \widetilde{\mathscr{X}_1}(n; n')$ is proper, flat, and surjective. For this, we first let $\mathscr{X}(nn')_{(m)}$ for every positive divisor m of $\operatorname{lcm}(n, n')$ be the preimage of $\widetilde{\mathscr{X}_1}(n; n')_{(m)}$. Due to Theorem 3.2.4(a), it then suffices to note that, by Proposition 4.2.12(a) and (d), the induced map

$$\mathscr{X}(nn')_{(m)} \to \widetilde{\mathscr{X}_1}(n; n')_{(m)} \times_{\overline{\mathscr{E}\ell\ell}_m} \overline{\mathscr{E}\ell\ell}_{nn'},$$

both components of which are forgetful, is finite locally free of constant positive rank.

The relative dimension aspect will follow from the corresponding aspect for $\mathscr{X}(nn')$ once we prove that the surjective map $\mathscr{X}(nn') \to \widetilde{\mathscr{X}_1}(n; n')$ is finite locally free. In fact, due to Lemma 3.2.3 and the previous paragraph, representability by algebraic spaces and quasifiniteness would suffice. The representability is inherited from $\mathscr{X}(nn') \to \mathscr{X}(1)$ and the quasifiniteness follows from the moduli interpretation.

The diagonal $\Delta_{\mathscr{X}/\mathbb{Z}}$ is proper due to the $\mathbb{Z}$-separatedness of $\mathscr{X}$ and is quasifinite due to Theorem 3.1.6(a), so its finiteness follows from Lemma 3.2.3.

(b) Due to the proof of (a) and the fact that the forgetful contractions (4.6.2.1) are representable and finite locally free, only the constancy of the rank requires attention and we may focus on $\mathscr{X}_0(n; n')$. Moreover, since $\mathscr{Y}_0(n; n')$ is dense in $\mathscr{X}_0(n; n')$, we may work on the elliptic curve locus. Therefore, since the rank of $\mathscr{Y}(nn') \to \mathscr{Y}(1)$ is constant, the conclusion follows from Proposition 4.2.15(a) which proves that $\mathscr{Y}_0(n; n') \to \mathscr{Y}(1)$ is finite locally free of constant positive rank.

(c) The assertion about the geometric fibers follows from the corresponding assertion for $\mathscr{X}(nn')^\infty \subset \mathscr{X}(nn')$ supplied by Proposition 4.3.2(b), so it suffices to prove that for each positive divisor m of $\operatorname{lcm}(n, n')$ the restriction $\mathscr{X}^\infty_{(m)} \subset \mathscr{X}_{(m)}$ of $\mathscr{X}^\infty \subset \mathscr{X}$ is a reduced relative effective Cartier divisor over $\operatorname{Spec} \mathbb{Z}$ that is smooth over $\mathbb{Z}[\frac{1}{nn'}]$. To do so, it suffices to note that $\mathscr{X}^\infty_{(m)}$ is the pullback of $\overline{\mathscr{E}\ell\ell}^\infty_m$, to apply Theorem 3.1.6(c)–(d) and Proposition 4.2.15(a), to use the properties of the forgetful maps (4.6.2.1), and to use the (R_0)+(S_1) criterion for reducedness. □

In principle it is possible to determine the largest Deligne–Mumford open substacks of $\widetilde{\mathscr{X}_1}(n; n')$, $\mathscr{X}_1(n; n')$, and $\mathscr{X}_0(n; n')$ (such open substacks make sense *a priori* due to Remark 3.1.7): one needs to inspect the defining modular descriptions to determine those geometric points whose automorphism functors are not étale. To illustrate the procedure, in Proposition 4.6.5 we exhibit large Deligne–Mumford open substacks of $\widetilde{\mathscr{X}_1}(n; n')$, $\mathscr{X}_1(n; n')$, and $\mathscr{X}_0(n; n')$ (the actual Deligne–Mumford loci of $\mathscr{X}_1(n; n')$ and $\mathscr{X}_0(n; n')$ may be larger). For the stack $\mathcal{M}_{\Gamma_1(N;n)}$ considered in [Conrad 2007], Proposition 4.6.5(b) improves on [Conrad 2007, 3.1.7] by proving that the Deligne–Mumford locus includes all the cusps in characteristics $p \mid N$ (even when $p^2 \mid n$).

Proposition 4.6.5. *Fix $n, n' \in \mathbb{Z}_{\geq 1}$ and set $d := \gcd(n, n')$.*

(a) *The stack $\widetilde{\mathscr{X}_1}(n; n')$ is Deligne–Mumford. In fact, the forgetful contraction morphism*
$$\widetilde{\mathscr{X}_1}(n; n') \to \mathscr{X}(1)$$
is representable by algebraic spaces.

(b) *The open substack of $\mathscr{X}_1(n; n')$ obtained by removing the closed substacks $\mathscr{X}_1(n; n')^{\infty}_{\mathbb{F}_p}$ for the primes p with $\operatorname{ord}_p(n') \geq \operatorname{ord}_p(n) + 2$ is Deligne–Mumford. If $\operatorname{ord}_p(n') \leq \operatorname{ord}_p(n) + 1$ for every prime p, then the forgetful contraction morphism*
$$\mathscr{X}_1(n; n') \to \mathscr{X}(1)$$
is representable by algebraic spaces.

(c) *The open substack of $\mathscr{X}_0(n; n')$ obtained by removing the closed substacks $\mathscr{X}_0(n; n')^{\infty}_{\mathbb{F}_p}$ for the primes p with $|\operatorname{ord}_p(n) - \operatorname{ord}_p(n')| \geq 2$ is Deligne–Mumford. If $|\operatorname{ord}_p(n) - \operatorname{ord}_p(n')| \leq 1$ for every prime p, then the forgetful contraction morphism*
$$\mathscr{X}_0(n; n') \to \mathscr{X}(1)$$
is representable by algebraic spaces.

Proof. We recall from Lemma 2.1.6 that the automorphism functor of the standard m-gon generalized elliptic curve is $\mu_m \times \mathbb{Z}/2\mathbb{Z}$. To test the Deligne–Mumford property of an open substack of $\widetilde{\mathscr{X}_1}(n; n')$, $\mathscr{X}_1(n; n')$, or $\mathscr{X}_0(n; n')$, we will use the criterion of having unramified automorphism functors at geometric points (see Remark 3.1.7). To test the representability of contraction morphisms, we will use Lemma 3.2.2(b). These preliminary remarks already settle part (a).

(b) Our task is to show that if p is a prime, E is the standard m-gon with $p \mid m$ over an algebraically closed field $\bar{k}$, and (E, α, H) is an object of $\mathscr{X}_1(n; n')(\bar{k})$ with $\operatorname{ord}_p(n') \leq \operatorname{ord}_p(n) + 1$, then $\mu_p \subset \operatorname{Aut}(E)$ does not fix both α and H. By decomposing into primary parts with the help of [Katz and Mazur 1985, 1.7.2] and by contracting away from the p-primary part of $[G + H]$, we loose no generality by assuming that n, n', and m are powers of p and $m > 1$.

Suppose that μ_p fixes both α and H. Then α cannot be ample, so H is ample, $H \cap (E^{\mathrm{sm}})^0$ contains $\mu_p \subset (E^{\mathrm{sm}})^0$, and $\mathrm{ord}_p(n') \geq 2$. Therefore, the standard cyclic subgroup $H_p \subset H$ of order p is contained in $(E^{\mathrm{sm}})^0$ and hence equals μ_p. Moreover, due to the requirement $\mathrm{ord}_p(n') \leq \mathrm{ord}_p(n) + 1$, we have $n > 1$, so, by Proposition 4.2.5(a), the requirement $[G_d + H_d] = E^{\mathrm{sm}}[d]$ implies that $[G_p + H_p] = E^{\mathrm{sm}}[p]$. The latter forces G_p to project isomorphically onto the p-torsion subgroup of the component group of E^{sm}, so G injects into this component group. Since H is ample and $H \cap (E^{\mathrm{sm}})^0 \neq 0$, this violates the requirement $\mathrm{ord}_p(n') \leq \mathrm{ord}_p(n) + 1$ unless G is ample, that is, unless α is ample, which is a contradiction.

(c) Our task is to show that if p is a prime, E is the standard m-gon with $p \mid m$ over an algebraically closed field $\bar{k}$, and (E, G, H) is an object of $\mathscr{X}_0(n; n')(\bar{k})$ with $|\mathrm{ord}_p(n) - \mathrm{ord}_p(n')| \leq 1$, then $\mu_p \subset \mathrm{Aut}(E)$ does not fix both G and H. As in the proof of (b), we assume that n, n', and m are powers of p and $m > 1$.

Suppose that μ_p fixes both G and H. By the conclusion of (b), μ_p cannot fix any Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure (resp. $\mathbb{Z}/n'\mathbb{Z}$-structure) on G (resp. H), so G and H must both be ample, and hence must both contain $\mu_p \subset (E^{\mathrm{sm}})^0$. Then $G_p = H_p = \mu_p$ inside $(E^{\mathrm{sm}})^0$, which is a contradiction to the requirement $[G_p + H_p] = E^{\mathrm{sm}}[p]$ inherited from $[G_d + H_d] = E^{\mathrm{sm}}[d]$. $\square$

With Proposition 4.6.5 in hand, we are ready for identifications with suitable modular curves $\mathscr{X}_H$.

Theorem 4.6.6. *Fix $n, n' \in \mathbb{Z}_{\geq 1}$.*

(a) *Let $\widetilde{\Gamma}_1(n; n')$ be the preimage in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ of the subgroup of $\mathrm{GL}_2(\mathbb{Z}/nn'\mathbb{Z})$ that fixes the subgroups $(\mathbb{Z}/nn'\mathbb{Z})[n] \times \{0\}$ and $\{0\} \times (\mathbb{Z}/nn'\mathbb{Z})[n']$ pointwise in $(\mathbb{Z}/nn'\mathbb{Z})^2$. The forgetful contraction $\widetilde{\mathscr{X}_1}(n; n') \to \mathscr{X}(1)$ induces the identification*

$$\widetilde{\mathscr{X}_1}(n; n') = \mathscr{X}_{\widetilde{\Gamma}_1(n;\, n')}.$$

(b) *Let $\Gamma_1(n; n')$ be the preimage in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ of the subgroup of $\mathrm{GL}_2(\mathbb{Z}/nn'\mathbb{Z})$ that fixes the subgroup $(\mathbb{Z}/nn'\mathbb{Z})[n] \times \{0\}$ pointwise and stabilizes the subgroup $\{0\} \times (\mathbb{Z}/nn'\mathbb{Z})[n']$ in $(\mathbb{Z}/nn'\mathbb{Z})^2$. If $\mathrm{ord}_p(n') \leq \mathrm{ord}_p(n) + 1$ for every prime p, then the forgetful contraction $\mathscr{X}_1(n; n') \to \mathscr{X}(1)$ induces the identification*

$$\mathscr{X}_1(n; n') = \mathscr{X}_{\Gamma_1(n;\, n')}.$$

(c) *Let $\Gamma_0(n; n')$ be the preimage in $\mathrm{GL}_2(\widehat{\mathbb{Z}})$ of the subgroup of $\mathrm{GL}_2(\mathbb{Z}/nn'\mathbb{Z})$ that stabilizes the subgroups $(\mathbb{Z}/nn'\mathbb{Z})[n] \times \{0\}$ and $\{0\} \times (\mathbb{Z}/nn'\mathbb{Z})[n']$ in*

$(\mathbb{Z}/nn'\mathbb{Z})^2$. *If* $|\mathrm{ord}_p(n') - \mathrm{ord}_p(n)| \le 1$ *for every prime* p, *then the forgetful contraction* $\mathscr{X}_0(n;\, n') \to \mathscr{X}(1)$ *induces the identification*

$$\mathscr{X}_0(n;\, n') = \mathscr{X}_{\Gamma_0(n;\, n')}.$$

Proof. By Proposition 4.6.5, the imposed assumptions on n and n' ensure that the forgetful contraction morphisms to $\mathscr{X}(1)$ are representable by algebraic spaces. Therefore, due to Theorem 4.6.4 and Theorem 4.5.1(c), we only need to show that, for variable $\mathbb{Z}[\frac{1}{nn'}]$-schemes S, the $\mathscr{Y}(1)_{\mathbb{Z}[\frac{1}{nn'}]}$-stacks

$$\widetilde{\mathscr{Y}_1}(n;\, n')_{\mathbb{Z}[\frac{1}{nn'}]}, \quad \mathscr{Y}_1(n;\, n')_{\mathbb{Z}[\frac{1}{nn'}]}, \quad \text{and} \quad \mathscr{Y}_0(n;\, n')_{\mathbb{Z}[\frac{1}{nn'}]}$$

parametrize elliptic curves $E \to S$ equipped with an S-point of

$$\overline{\widetilde{\Gamma}_1(n;\, n')} \backslash \mathrm{Isom}(E[nn'], (\mathbb{Z}/nn'\mathbb{Z})^2), \quad \overline{\Gamma_1(n;\, n')} \backslash \mathrm{Isom}(E[nn'], (\mathbb{Z}/nn'\mathbb{Z})^2),$$
$$\text{and} \quad \overline{\Gamma_0(n;\, n')} \backslash \mathrm{Isom}(E[nn'], (\mathbb{Z}/nn'\mathbb{Z})^2),$$

respectively, where overlines denote images in $\mathrm{GL}_2(\mathbb{Z}/nn'\mathbb{Z})$. For this, it suffices to inspect the defining modular descriptions of $\widetilde{\mathscr{X}_1}(n;\, n')$, $\mathscr{X}_1(n;\, n')$, and $\mathscr{X}_0(n;\, n')$ and to use the definitions of $\widetilde{\Gamma}_1(n;\, n')$, $\Gamma_1(n;\, n')$, and $\Gamma_0(n;\, n')$ given in the statements of (a), (b), and (c). □

4.7. A modular construction of Hecke correspondences for $\mathscr{X}_1(n)$

We wish to explain how the results of Sections 2.2, 4.4, and 4.6 give rise to a Hecke correspondence

$$\pi_1, \pi_2 : \mathscr{X}_{\Gamma_1(n;\, p)} \rightrightarrows \mathscr{X}_{\Gamma_1(n)}$$

for every $n \in \mathbb{Z}_{\ge 1}$ and every squarefree $p \in \mathbb{Z}_{\ge 1}$ that may or may not be coprime with n.

In terms of the moduli interpretations given in Sections 4.4.1 and 4.6.1 and proved in Theorems 4.4.4(c) and 4.6.6(b), the maps are given by

$$\pi_1((E, \alpha, H)) = (c_\alpha(E), \alpha) \quad \text{and} \quad \pi_2((E, \alpha, H)) = (E/H, \alpha),$$

and are well defined due to the last aspect of Proposition 4.2.11(b) (we let $c_\alpha(E)$ denote the contraction of E with respect to the unique subgroup on which α is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure). To argue that we have exhibited a correspondence, it suffices to prove the following lemma:

Lemma 4.7.1. *The maps π_1 and π_2 are representable, finite locally free, and surjective.*

Proof. Since π_1 is the $\mathscr{X}(1)$-morphism induced by the inclusion $\Gamma_1(n;\, p) \subset \Gamma_1(n)$, its finiteness follows from the finiteness of $\mathscr{X}_H \to \mathscr{X}_{H'}$ observed in the last paragraph of Section 4.1.2. By Theorem 4.4.4(a), $\mathscr{X}_{\Gamma_1(n)}$ is regular, so the flatness of π_1 follows from [EGA IV$_2$ 1965, 6.1.5]. The surjectivity of π_1 may be checked over $(\mathscr{Y}_{\Gamma_1(n)})_{\mathbb{Q}}$.

For the representability of π_2, due to Lemma 3.2.2(b) and the representability of $\mathscr{X}_{\Gamma_1(n;\, p)} \to \mathscr{X}(1)$, it suffices to observe that if E is a generalized elliptic curve over an algebraically closed field and $H \subset E^{\mathrm{sm}}$ is a finite subgroup, then every automorphism i of E that stabilizes H and induces the identity map on E/H must fix $(E^{\mathrm{sm}})^0$ because the endomorphism $\mathrm{id}_{E^{\mathrm{sm}}} - i|_{E^{\mathrm{sm}}}$ of E^{sm} factors through H. The properness of π_2 follows from the $\mathbb{Z}$-properness of $\mathscr{X}_{\Gamma_1(n;\, p)}$ and $\mathscr{X}_{\Gamma_1(n)}$, so its quasifiniteness may be checked on geometric fibers. Finiteness of π_2 is then supplied by Lemma 3.2.3, and its flatness follows from [EGA IV$_2$ 1965, 6.1.5]. Finally, the surjectivity of π_2 may be checked over $(\mathscr{Y}_{\Gamma_1(n)})_{\mathbb{Q}}$. □

In the case when p is a prime, the Hecke correspondence above has already been constructed in [Conrad 2007, 4.4.3] by a different method: due to the lack of the theory of quotients of generalized elliptic curves by arbitrary finite locally free subgroups, [loc. cit.] first defines π_2 by the same formula on the elliptic curve locus and then argues that the resulting map extends uniquely to the entire $\mathscr{X}_{\Gamma_1(n;p)}$. The construction above seems simpler and more direct, and it also produces the map ξ of [Conrad 2007, 4.4.3]: if e and e' are the identity sections of $E \to S$ and $E/H \to S$, then there is a map

$$(e')^*(\Omega^1_{(E/H)/S}) \to e^*(\Omega^1_{E/S})$$

whose formation is compatible with base change in S.

Chapter 5. A modular description of $\mathscr{X}_{\Gamma_0(n)}$

For an integer $n \in \mathbb{Z}_{\geq 1}$ and the subgroup

$$\Gamma_0(n) := \left\{ \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathrm{GL}_2(\widehat{\mathbb{Z}}) \mid c \equiv 0 \bmod n \right\},$$

the goal of this chapter is to exhibit the modular curve $\mathscr{X}_{\Gamma_0(n)}$ defined via normalization (see Section 4.1.2) as a moduli stack parametrizing generalized elliptic curves equipped with a "$\Gamma_0(n)$-structure," which on the elliptic curve locus is the datum of a subgroup that is cyclic of order n in the sense of Definition 4.2.6. The proof of the correctness of this moduli interpretation in Theorem 5.13 will simultaneously deduce the regularity of $\mathscr{X}_{\Gamma_0(n)}$ from that of $\mathscr{Y}_{\Gamma_0(n)}$ proved by Katz and Mazur. We begin with a naive modular description that recovers $\mathscr{X}_{\Gamma_0(n)}$ only for squarefree n and then proceed to refine the naive description to a description that works for any n.

Throughout Chapter 5 we fix an integer $n \in \mathbb{Z}_{\geq 1}$.

5.1. ***The stack*** $\mathscr{X}_0(n)^{\mathrm{naive}}$**.** This is the $\mathbb{Z}$-stack that, for variable schemes S, parametrizes the pairs

$$(E \xrightarrow{\pi} S,\ G)$$

consisting of a generalized elliptic curve $E \xrightarrow{\pi} S$ and an ample S-subgroup $G \subset E^{\mathrm{sm}}$ that is cyclic of order n (in the sense of Definition 4.2.6). We call such a G a *naive* $\Gamma_0(n)$*-structure on* E.

We let

$$\mathscr{Y}_0(n)^{\mathrm{naive}} \subset \mathscr{X}_0(n)^{\mathrm{naive}}$$

be the open substack that parametrizes those pairs for which E is an elliptic curve. For each positive divisor m of n, we let

$$\mathscr{X}_0(n)^{\mathrm{naive}}_{(m)} \subset \mathscr{X}_0(n)^{\mathrm{naive}}$$

be the open substack that parametrizes those pairs for which the degenerate geometric fibers of E are m-gons.

In the notation of Section 4.6.2, one has

$$\mathscr{X}_0(n)^{\mathrm{naive}} = \mathscr{X}_0(n;\, 1),$$

so, by Theorem 4.6.4(a), the stack $\mathscr{X}_0(n)^{\mathrm{naive}}$ is algebraic, proper and flat over $\operatorname{Spec}\mathbb{Z}$, and regular with finite diagonal $\Delta_{\mathscr{X}_0(n)^{\mathrm{naive}}/\mathbb{Z}}$. By Theorem 4.6.4(b) (and its proof), the morphism

$$\mathscr{X}(n) \to \mathscr{X}_0(n)^{\mathrm{naive}}$$

that sends a Drinfeld $(\mathbb{Z}/n\mathbb{Z})^2$-structure α to the subgroup on which $\alpha|_{\mathbb{Z}/n\mathbb{Z}\times\{0\}}$ is a Drinfeld $\mathbb{Z}/n\mathbb{Z}$-structure and contracts the underlying generalized elliptic curve with respect to this subgroup is finite locally free of rank $n\cdot\phi(n)^2$.

If n is squarefree, then Theorem 4.6.6(c) proves that the contraction

$$\mathscr{X}_0(n)^{\mathrm{naive}} \to \mathscr{X}(1) \quad \text{is identified with the structure morphism} \quad \mathscr{X}_{\Gamma_0(n)} \to \mathscr{X}_0(1).$$

This identification fails when n is divisible by p^2 for some prime p: variants of the example given in Section 1.2 show that for such n the contraction

$$\mathscr{X}_0(n)^{\mathrm{naive}} \to \mathscr{X}(1)$$

is not representable.

5.2. ***The notation*** $d(m)$**.** For a positive divisor m of n, we set

$$d(m) := \frac{m}{\gcd\bigl(m, \frac{n}{m}\bigr)},$$

so that $d(m)$ depends *both* on m and on the integer n that is fixed throughout.

To explain the role of the function $m \mapsto d(m)$ in the context of $\Gamma_0(n)$-structures on generalized elliptic curves, let E be the standard m-gon over an algebraically closed field and suppose that E is equipped with an ample cyclic subgroup $G \subset E^{\mathrm{sm}}$ of order n. Then $G \cap (E^{\mathrm{sm}})^0 = \mu_{n/m}$ and $\mu_m \subset \operatorname{Aut}(E)$ is the subgroup of those automorphisms that induce the identity map on the contraction of E with respect to the zero section (see Lemma 2.1.6). The further subgroup of $\mathrm{Aut}(E)$ that in addition stabilizes G is therefore $\mu_m \cap \mu_{n/m} = \mu_{\gcd(m,n/m)}$ (intersection in $(E^{\mathrm{sm}})^0$), and this subgroup acts trivially on precisely $d(m)$ of the m irreducible components of E.

When refining G to a $\Gamma_0(n)$-structure on such an E, we will only remember the contraction $c_{E^{\mathrm{sm}}[d(m)]}(E)$ that is a $d(m)$-gon together with the standard cyclic subgroup $G_{(n/m)\cdot d(m)}$ of order $\frac{n}{m} \cdot d(m)$. In addition, we will require the datum of a compatible ample cyclic G' of order n on every E' that contracts to (a base change) of $c_{E^{\mathrm{sm}}[d(m)]}(E)$ and that has m-gon degenerate geometric fibers. Different m may give the same $d(m)$, so there is no way to recover m from $c_{E^{\mathrm{sm}}[d(m)]}(E)$ alone; to overcome this, we will incorporate m into the data that comprises a $\Gamma_0(n)$-structure.

For the precise definition of a $\Gamma_0(n)$-structure given in Section 5.10, we need the following preparations.

5.3. ***The stack of "decontractions".*** Fix a positive divisor m of n and suppose that we have a generalized elliptic curve $E \xrightarrow{\pi} S$ and an open subscheme $S_{\pi,(m)} \subset S$ that contains the elliptic curve locus $S - S^{\infty,\pi}$ and such that the degenerate geometric fibers of $E_{S_{\pi,(m)}}$ are $d(m)$-gons. (Such an $S_{\pi,(m)}$ will be part of the data of a $\Gamma_0(n)$-structure on E.) The base change $E_{S_{\pi,(m)}}$ determines a map $S_{\pi,(m)} \to \overline{\mathscr{E}\ell\ell}_{d(m)}$, so we may consider the fiber product algebraic stack

$$S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m,$$

which parametrizes "decontractions" of $E_{S_{\pi,(m)}}$, or, more precisely, which, for variable $S_{\pi,(m)}$-schemes S', parametrizes the pairs

$$\big(E' \xrightarrow{\pi'} S',\ \iota' : E_{S'} \xrightarrow{\sim} c_{E'^{\mathrm{sm}}[d(m)]}(E')\big)$$

consisting of a generalized elliptic curve $E' \xrightarrow{\pi'} S'$ whose degenerate geometric fibers are m-gons and a specified S'-isomorphism ι'. We denote the universal object of $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$ by

$$(\mathcal{E}_{\pi,(m)}, \iota_{\pi,(m)}).$$

The base change of $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$ (resp. of $\mathcal{E}_{\pi,(m)}$) to $S - S^{\infty,\pi}$ is identified with $S - S^{\infty,\pi}$ (resp. with $E_{S-S^{\infty,\pi}}$), and the same holds over the entire $S_{\pi,(m)}$ if $d(m) = m$.

We will endow the universal "decontraction" $\mathcal{E}_{\pi,(m)}$ with additional structures. The algebraic stack $\mathcal{E}_{\pi,(m)}$ is typically not a scheme, but there are two ways to think about such structures concretely:

- As compatible with isomorphisms and base change structures on E' for each

$$(E' \xrightarrow{\pi'} S', \iota');$$

- As compatible under the pullbacks

$$S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} X_1 \rightrightarrows S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} X_0$$

structures on the "decontractions" over the indicated bases, where

$$X_1 \rightrightarrows X_0 \to \overline{\mathscr{E}\ell\ell}_m$$

is a once and for all fixed scheme presentation of the algebraic stack $\overline{\mathscr{E}\ell\ell}_m$, so that

$$S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} X_1 \rightrightarrows S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} X_0 \to S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$$

is a scheme presentation of the algebraic stack $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$ (by Theorem 3.1.6(a), the algebraic stacks $\overline{\mathscr{E}\ell\ell}_m$ and $\overline{\mathscr{E}\ell\ell}_{d(m)}$ have finite diagonals, so $X_0 \times_{\overline{\mathscr{E}\ell\ell}_m} X_0$ and similar fiber products that would *a priori* be algebraic spaces are schemes).

The second way has the advantage of avoiding set-theoretic difficulties that would need to be addressed in order to make the first way completely rigorous.

The contractions of the generalized elliptic curves parametrized by the stack $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$ are identified. In particular, the degenerate geometric fibers of these curves have canonically isomorphic component groups because the identity component of such a fiber may be used to fix the "direction" of the m-gon. This observation lies behind the following lemma:

Lemma 5.4. *Let $E \xrightarrow{\pi} S$ and $E' \xrightarrow{\pi'} S$ be generalized elliptic curves whose degenerate geometric fibers are m-gons and let $\iota : c(E) \xrightarrow{\sim} c(E')$ be an S-isomorphism between their contractions with respect to the identity sections.*

(a) *If S is a geometric point, then there is a unique identification*

$$E^{\mathrm{sm}}/(E^{\mathrm{sm}})^0 = E'^{\mathrm{sm}}/(E'^{\mathrm{sm}})^0$$

of the component groups that is induced by any isomorphism $E \simeq E'$ that is compatible with ι.

(b) *If $S^{\mathrm{red}} = (S^{\infty,\pi})^{\mathrm{red}}$* (so that also $S^{\mathrm{red}} = (S^{\infty,\pi'})^{\mathrm{red}}$), *then there is a unique S-identification*

$$(E^{\mathrm{sm}})[m]/(E^{\mathrm{sm}})^0[m] = (E'^{\mathrm{sm}})[m]/(E'^{\mathrm{sm}})^0[m]$$

whose base change to any geometric S-point $\bar{s}$ is induced by any $\bar{s}$-isomorphism $E_{\bar{s}} \simeq E'_{\bar{s}}$ compatible with $\iota_{\bar{s}}$. Any S-isomorphism $i : E \simeq E'$ compatible with ι induces this identification.

(c) *For $g \in E^{\mathrm{sm}}(S)$ and $g' \in E'^{\mathrm{sm}}(S)$, the set of $s \in S$ for which g and g' meet the same* (in the sense of (a)) *irreducible components of $E_{\bar{s}}$ and $E'_{\bar{s}}$ forms an open subscheme of S that is also closed if $S^{\mathrm{red}} = (S^{\infty,\pi})^{\mathrm{red}}$.*

Proof. (a) If either E or E' is smooth, then ι itself induces the desired identification. We may therefore assume that both E and E' are degenerate. Then, by Remark 2.1.9, both E and E' are isomorphic to the standard m-gon discussed in Remark 2.1.5. Moreover, any two isomorphisms $E \simeq E'$ that are compatible with ι differ by an automorphism of E' that is the identity map on $(E'^{\mathrm{sm}})^0$. It remains to observe that, by Lemma 2.1.6, any automorphism of E' that is the identity map on $(E'^{\mathrm{sm}})^0$ induces the identity map on $E'^{\mathrm{sm}}/(E'^{\mathrm{sm}})^0$.

(b) If S is a geometric point, then

$$(E^{\mathrm{sm}})[m]/(E^{\mathrm{sm}})^0[m] = E^{\mathrm{sm}}/(E^{\mathrm{sm}})^0,$$

and likewise for E', so the claim follows from (a). In general, by Lemma 2.1.11, both

$$(E^{\mathrm{sm}})[m]/(E^{\mathrm{sm}})^0[m] \quad \text{and} \quad (E'^{\mathrm{sm}})[m]/(E'^{\mathrm{sm}})^0[m]$$

are étale, so we may and do assume that $S = S^{\mathrm{red}}$. In this case, by Remark 2.1.9, i exists fppf locally on S. Moreover, any i satisfies the defining property, so we only need to check that two different i induce the same identification. For this, the case of a local strictly Henselian S suffices and reduces to the settled case of a geometric point.

(c) We may assume that $S = S^{\infty,\pi} = S^{\infty,\pi'}$ and S is reduced and may work fppf locally on S. We therefore use Remark 2.1.9 to fix an S-isomorphism $i : E \xrightarrow{\sim} E'$ that is compatible with ι and to assume that E is the standard m-gon. In this case, the label of the component of E^{sm} that meets g is locally constant on S, and likewise for $\iota^{-1}(g')$. □

5.5. ***Coherence of a cyclic subgroup of the universal "decontraction".*** In the notation of Section 5.3, part of the data of a $\Gamma_0(n)$-structure will be an ample cyclic $(S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m)$-subgroup

$$\mathcal{G}_{(m)} \subset \mathcal{E}^{\mathrm{sm}}_{\pi,(m)}$$

of order n, or, in more concrete terms, for every $(E' \xrightarrow{\pi'} S', \iota')$ an ample cyclic S'-subgroup $G' \subset E'^{\mathrm{sm}}$ of order n that is compatible with base change and with isomorphisms of pairs (E', ι') (for the notion of cyclicity, see Definition 4.2.6).

In order to isolate a well-behaved class of such $\mathcal{G}_{(m)}$, we say that $\mathcal{G}_{(m)}$ is *coherent* if:

For every $S_{\pi,(m)}$-scheme S' and every pair of objects

$$\big(E'_1 \xrightarrow{\pi'_1} S', \iota'_1\big) \quad \text{and} \quad \big(E'_2 \xrightarrow{\pi'_2} S', \iota'_2\big)$$

of $(S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m)(S')$, the pullbacks $G'_1 \subset E'^{\mathrm{sm}}_1$ and $G'_2 \subset E'^{\mathrm{sm}}_2$ of $\mathcal{G}_{(m)}$ fpqc locally on S' have generators g'_1 and g'_2 that meet the same (in the sense of Lemma 5.4(a)) irreducible components of the geometric fibers of E'_1 and E'_2 and satisfy

$$(\iota'_1)^{-1}\big(\tfrac{n}{m} \cdot g'_1\big) = (\iota'_2)^{-1}\big(\tfrac{n}{m} \cdot g'_2\big).$$

(The last equality takes place in E and makes sense because $\frac{n}{m} \cdot g'_1$ lies in the contraction $c_{E'^{\mathrm{sm}}_1[d(m)]}(E'_1)$ by Proposition 4.2.9(c), and likewise for $\frac{n}{m} \cdot g'_2$.) Equivalently, the coherence of $\mathcal{G}_{(m)}$ is a condition of the existence of compatible fpqc local generators of the pullbacks of $\mathcal{G}_{(m)}$ along the two projections

$$(S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m) \times_{S_{\pi,(m)}} (S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m) \rightrightarrows S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m,$$

where compatibility amounts to the conditions imposed on g'_1 and g'_2 above.

In what follows, the purpose of the coherence condition is to ensure that $\mathcal{G}_{(m)}$ is uniquely determined by its pullback to *any* $(E' \xrightarrow{\pi'} S', \iota')$ with $S' = S_{\pi,(m)}$, provided that such an (E', ι') exists. Lemma 5.7 will justify this, and its aspect (iii) will show that no generality is lost if one strengthens the coherence condition by fixing an fpqc local generator g'_1 of G'_1 in advance.

Any $\mathcal{G}_{(m)}$ is coherent if $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m = S_{\pi,(m)}$, and also if n is a unit on $S_{\pi,(m)}$ as we now show.

Lemma 5.6. *If n is invertible on $S_{\pi,(m)}$, then every ample cyclic $(S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m)$-subgroup $\mathcal{G}_{(m)} \subset \mathcal{E}^{\mathrm{sm}}_{\pi,(m)}$ of order n is coherent.*

Proof. We will show that for every pair $(E'_1 \xrightarrow{\pi'_1} S', \iota'_1)$ and $(E'_2 \xrightarrow{\pi'_2} S', \iota'_2)$ as in the definition of coherence, desired generators g'_1 and g'_2 of G'_1 and G'_2 exist even étale locally on S'. For this, due to Lemma 5.4(c), we may assume that S' is local strictly Henselian and that the special fibers $(E'_1)_{s'}$ and $(E'_2)_{s'}$ are degenerate. Moreover, since $(E'_1)^{\mathrm{sm}}[n]$ and $(E'_2)^{\mathrm{sm}}[n]$ are étale and G'_1 and G'_2 are constant, we may assume further that S' is a geometric point. In the case of a geometric point, it suffices to transport any choice of a g'_1 across any S'-isomorphism $(E'_1, \iota'_1) \simeq (E'_2, \iota'_2)$. $\square$

The following key lemma analyses the coherence condition beyond the case when n is a unit by exhibiting a universal property satisfied by pullbacks of a coherent $\mathcal{G}_{(m)}$. This property compensates for the loss of a direct reduction to geometric points that governed the case of an invertible n.

Lemma 5.7. *Let m be a positive divisor of n, let $d \in \mathbb{Z}_{\geq 1}$ be a multiple of m, let $E \xrightarrow{\pi} S$ and $E' \xrightarrow{\pi'} S$ be generalized elliptic curves whose degenerate geometric* *fibers are d-gons, and let*

$$\iota : c_{E^{\mathrm{sm}}[d(m)]}(E) \xrightarrow{\sim} c_{E'^{\mathrm{sm}}[d(m)]}(E')$$

be an S-isomorphism. For every cyclic S-subgroup $G \subset E^{\mathrm{sm}}$ of order n that meets precisely m irreducible components of every degenerate geometric fiber of E, there is a unique cyclic S-subgroup $G' \subset E'^{\mathrm{sm}}$ of order n such that:

(i) *Over $S - S^{\infty,\pi} = S - S^{\infty,\pi'}$ there is an equality $\iota(G_{S-S^{\infty,\pi}}) = G'_{S-S^{\infty,\pi'}}$.*

(ii) *fpqc locally on S there exist generators g of G and g' of G' that meet the same irreducible components of the geometric fibers of E and E'* (in the sense of Lemma 5.4(a)) *and satisfy*

$$\iota\big(\tfrac{n}{m} \cdot g\big) = \tfrac{n}{m} \cdot g'.$$

(So G' meets precisely m irreducible components of every degenerate geometric fiber of E'.)

Moreover, this unique G' is such that:

(iii) *For every S-scheme T and every generator $\tilde{g}$ of G_T, fpqc locally on T there exists a generator $\tilde{g}'$ of G'_T such that $\tilde{g}$ and $\tilde{g}'$ meet the same irreducible components of the geometric fibers of E and E' and satisfy*

$$\iota\big(\tfrac{n}{m} \cdot \tilde{g}\big) = \tfrac{n}{m} \cdot \tilde{g}'.$$

(iv) *The standard cyclic subgroups $G_{(n/m)\cdot d(m)} \subset G$ and $G'_{(n/m)\cdot d(m)} \subset G'$ of order $\frac{n}{m} \cdot d(m)$ satisfy*

$$\iota(G_{(n/m)\cdot d(m)}) = G'_{(n/m)\cdot d(m)}.$$

Remark 5.8. Due to Proposition 4.2.9(c), the equalities displayed in (ii)–(iv) make sense.

Proof of Lemma 5.7. We have broken the argument up into six steps.

Step 1: *The claim of* (iv) *follows from the rest.* The subgroups $\iota(G_{(n/m)\cdot d(m)})$ and $G'_{(n/m)\cdot d(m)}$ of E'^{sm} are cyclic of order $\frac{n}{m} \cdot d(m)$, agree with $\iota((G_{(n/m)\cdot d(m)})_{S-S^{\infty,\pi}})$ over $S - S^{\infty,\pi'}$, and fpqc locally on S have generators $\iota\big(\frac{m}{d(m)} \cdot g\big)$ and $\frac{m}{d(m)} \cdot g'$ whose $\frac{n}{m}$-multiples equal $\iota\big(\frac{n}{m} \cdot \big(\frac{m}{d(m)} \cdot g\big)\big)$. Therefore, $\iota(G_{(n/m)\cdot d(m)})$ and $G'_{(n/m)\cdot d(m)}$ must be equal because they satisfy (i) and (ii) when n, m, and G are replaced by $\frac{n}{m} \cdot d(m)$, $d(m)$, and $G_{(n/m)\cdot d(m)}$, respectively ($G_{(n/m)\cdot d(m)}$ meets precisely $d(m)$ irreducible components of every degenerate geometric fiber of E due to Proposition 4.2.9(c)).

Step 2: *The claim of* (iii). We may assume that $T = S$ and may work fpqc locally on S, so we fix g, g', and $\tilde{g}$ over S. In order to find a desired fpqc local $\tilde{g}'$, we work Zariski locally on S and use limit arguments together with Lemma 5.4(c) to reduce to the case when $S = \operatorname{Spec} R$ for some Noetherian R. Then we pass to an fpqc cover to assume that R is complete and separated with respect to the ideal I that cuts out $S^{\infty,\pi}$ (equivalently, with respect to the ideal that cuts out $S^{\infty,\pi'}$; see Corollary 3.2.5).

By Proposition 3.2.7(a), $E^{\mathrm{sm}}[n]$ (resp. $E'^{\mathrm{sm}}[n]$) has the largest finite locally free S-subgroup $A_{n,m}$ (resp. $A'_{n,m}$) that meets precisely m irreducible components of every degenerate geometric fiber of E (resp. E'), so $G \subset A_{n,m}$ and $G' \subset A'_{n,m}$. Moreover, Proposition 3.2.7(a) supplies extensions

$$\begin{tikzcd} 0 \arrow[r] & B_n \arrow[r] \arrow[d, equal] & A_{n,m} \arrow[r] & C_m \arrow[r] \arrow[d, equal] & 0 \\ 0 \arrow[r] & B_n \arrow[r] & A'_{n,m} \arrow[r] & C_m \arrow[r] & 0 \end{tikzcd}$$

of S-group schemes, where the identification of B_n is via ι and the identification of C_m is via Lemma 5.4(b) (applied over R/I^j for every $j \geq 1$ to the contractions of E_{R/I^j} and E'_{R/I^j} with respect to the m-torsion). As may be checked on degenerate geometric fibers, the generators $g \in G(S)$ and $g' \in G'(S)$ project to the same section of C_m that gives an isomorphism $C_m \simeq \mathbb{Z}/m\mathbb{Z}$.

The homomorphism $G \to C_m$ is finite locally free and, by Proposition 4.2.10(a), its kernel is the standard cyclic subgroup $G_{n/m} \subset G$ of order $\frac{n}{m}$. By replacing g and g' by $u \cdot g$ and $u \cdot g'$ for a suitable $u \in \underline{(\mathbb{Z}/n\mathbb{Z})^\times}(S)$, we reduce to the case when g and $\tilde{g}$ have the same image in C_m. Then $g - \tilde{g} \in G_{n/m}$, so $\frac{n}{m} \cdot g = \frac{n}{m} \cdot \tilde{g}$, which means that we may choose $\tilde{g}'$ to be g'.

For the rest of the proof, we focus on the remaining claim about the existence and uniqueness of G'.

Step 3: *Reduction to the case when n is a prime power.* The group G, as well as any candidate G', decomposes as a product of its p-primary parts for various primes p dividing n. By [Katz and Mazur 1985, 1.7.2], cyclicity of G or of G' is equivalent to the cyclicity of the primary factors, and the datum of a generator of G or of G' corresponds to the datum of a generator of each primary factor. Therefore, for the existence and the uniqueness of the sought G' we may assume that n is a prime power.

For the rest of the proof, we assume that $n = p^r$ and $m = p^s$ for some prime p and $r, s \in \mathbb{Z}_{\geq 0}$.

Step 4: *The case $s = 0$.* For the existence, $\iota(G)$ fulfills the requirements (i)–(ii). The uniqueness reduces to the case of an Artinian local S and then follows from Proposition 3.2.7(a).

For the rest of the proof, we assume that $s \geq 1$, so that $\frac{n}{m} \neq n$.

Step 5: *Uniqueness of G'.* Due to the claim concerning (iii) (i.e., due to Step 2), we may assume that the two candidates $G'_1, G'_2 \subset E'^{\mathrm{sm}}$ have generators g'_1 and g'_2 that meet the same irreducible components of the geometric fibers of E' and satisfy $\frac{n}{m} \cdot g'_1 = \frac{n}{m} \cdot g'_2$. Furthermore, we may assume that the base S is Noetherian, then local, then complete, and finally Artinian, and that E' is nonsmooth over S. Then, since $g'_1 - g'_2 \in (E'^{\mathrm{sm}})^0(S)$ and $\frac{n}{m} \cdot g'_1 = \frac{n}{m} \cdot g'_2$, we have

$$g'_2 = g'_1 + h \quad \text{for some } h \in (E'^{\mathrm{sm}})^0\big[\tfrac{n}{m}\big](S).$$

By Lemma 2.1.11 and Proposition 4.2.10(a), the S-group $(E'^{\mathrm{sm}})^0[n/m]$ is the standard cyclic subgroup of G'_1 of order $\frac{n}{m}$, so Proposition 4.2.9(f) ensures that $g'_1 + h$ generates G'_1, which means that $G'_1 = G'_2$.

Step 6: *Existence of G'.* Due to the uniqueness of G', for its existence we may work fpqc locally on S, so we fix a generator g of G. Moreover, as in Step 2 we reduce to the case when $S = \operatorname{Spec} R$ for a Noetherian R that is complete and separated with respect to the ideal $I \subset R$ that cuts out $S^{\infty,\pi}$ and use Proposition 3.2.7(a) to obtain the diagram of extensions displayed in Step 2.

By Proposition 3.2.7(a), $E'^{\mathrm{sm}}[m] \subset A'_{n,m}$, so $E'^{\mathrm{sm}}[m/d(m)] \subset A'_{n,m}$, too, and hence the image of $A'_{n,m}$ under the multiplication by $m/d(m)$ map of E'^{sm} is a finite locally free S-subgroup of $A'_{(n/m)\cdot d(m),d(m)}$ of order $\big(\frac{n}{m} \cdot d(m)\big) \cdot d(m)$. This image therefore equals $A'_{(n/m)\cdot d(m),d(m)}$, so, since $\iota(m/d(m) \cdot g)$ lies in $A'_{(n/m)\cdot d(m),d(m)}$, after replacing S by a finite locally free cover we may choose a $g' \in A'_{n,m}(S)$ with

$$\frac{m}{d(m)} \cdot g' = \iota\Big(\frac{m}{d(m)} \cdot g\Big).$$

Since $E'^{\mathrm{sm}}[m/d(m)]$ is an extension of $(C_m)[m/d(m)]$ by $(B_n)[m/d(m)]$, after a further finite locally free cover of S we may adjust g' by a lift to $(E'^{\mathrm{sm}}[m/d(m)])(S)$ of the difference of the images of g and g' in C_m to arrange that g and g' have the same image in C_m and hence meet the same irreducible components of the geometric fibers of E and E'.

By Proposition 4.2.5(d), g' generates a cyclic S-subgroup $G' \subset E'^{\mathrm{sm}}$ of order n. Since $(m/d(m)) \mid (n/m)$, the group G' satisfies (ii). Thus, to complete Step 6, and hence also the proof of Lemma 5.7, it suffices to show that

$$\iota(G_{S-S^{\infty,\pi}}) = G'_{S-S^{\infty,\pi'}}.$$

We have $G \subset A_{n,m}$ and $G' \subset A'_{n,m}$ with g and g' projecting to the same section of C_m. Moreover, by Proposition 3.2.7(b) and the diagram displayed in Step 2, both $\iota((A_{n,m})_{S-S^{\infty,\pi'}})$ and $(A'_{n,m})_{S-S^{\infty,\pi'}}$ are the preimages in $E'_{S-S^{\infty,\pi'}}[n]$ of the unique $(S - S^{\infty,\pi'})$-subgroup of $(E'^{\mathrm{sm}}[n]/B_n)_{S-S^{\infty,\pi'}}$ of order m, so

$$\iota \text{ identifies } A_{n,m} \text{ and } A'_{n,m} \quad \text{over } S - S^{\infty,\pi'}.$$

We claim that under this identification via ι, the image of $g_{S-S^{\infty,\pi}}$ in $A_{n,m}/B_n$ agrees with the image of $g'_{S-S^{\infty,\pi'}}$ in $A'_{n,m}/B_n$. Since $A'_{n,m}/B_n$ is finite étale, it suffices to check the claimed agreement on the geometric fibers at the points in $S - S^{\infty,\pi'}$, so the technique used in the proof of Proposition 3.2.7(b) reduces the proof of the claimed agreement to the case when R is a discrete valuation ring and E and E' have smooth generic fibers but nonsmooth closed fibers. In this case, by Proposition 3.1.8(b), ι extends to a unique isomorphism $E \simeq E'$, which then must induce the identification of the groups C_m for E and E'. Thus, in this case the claimed agreement follows from the agreement of the images of g and g' in C_m.

Returning to the proof of $\iota(G_{S-S^{\infty,\pi}}) = G'_{S-S^{\infty,\pi'}}$, via the above reasoning, we conclude that $g'_{S-S^{\infty,\pi'}} - \iota(g_{S-S^{\infty,\pi}})$ lies in B_n. Moreover, since $(m/d(m)) \mid (n/m)$, the construction of g' ensures that

$$\tfrac{n}{m} \cdot g'_{S-S^{\infty,\pi'}} = \tfrac{n}{m} \cdot \iota(g_{S-S^{\infty,\pi}}).$$

Therefore, there is an $h \in ((B_n)[n/m])(S - S^{\infty,\pi'})$ such that

$$g'_{S-S^{\infty,\pi'}} = \iota(g_{S-S^{\infty,\pi}}) + h.$$

By the uniqueness aspect of the first assertion of Proposition 3.2.7(a) and by Proposition 4.2.9(c), $(B_n)[n/m]$ is the standard cyclic subgroup of G of order $\frac{n}{m}$, so $\iota(g_{S-S^{\infty,\pi}}) + h$ generates $\iota(G_{S-S^{\infty,\pi}})$ by Proposition 4.2.9(f). The sought equality $\iota(G_{S-S^{\infty,\pi}}) = G'_{S-S^{\infty,\pi'}}$ follows. □

We are ready for the definition of a $\Gamma_0(n)$-structure on a generalized elliptic curve.

5.9. ***$\Gamma_0(n)$-structures.*** For a generalized elliptic curve $E \xrightarrow{\pi} S$, a $\Gamma_0(n)$-*structure on* E is a tuple

$$\big(G,\ \{S_{\pi,(m)}\}_{m|n},\ \{\mathcal{G}_{(m)}\}_{m|n}\big)$$

consisting of the following data:

(1) a cyclic $(S - S^{\infty,\pi})$-subgroup $G \subset E_{S-S^{\infty,\pi}}$ of order n (in the sense of Definition 4.2.6);

(2) for each positive divisor m of n, an open subscheme $S_{\pi,(m)} \subset S$ such that

  (2.1) $S = \bigcup_m S_{\pi,(m)}$;

  (2.2) if $m \neq m'$, then $S_{\pi,(m)} \cap S_{\pi,(m')} = S - S^{\infty,\pi}$;

  (2.3) the degenerate geometric fibers of $E_{S_{\pi,(m)}}$ are $d(m)$-gons, where

$$d(m) = \frac{m}{\gcd\big(m, \frac{n}{m}\big)};$$

(3) for each positive divisor m of n, in the notation of Section 5.3, an ample cyclic $(S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m)$-subgroup

$$\mathcal{G}_{(m)} \subset \mathcal{E}^{\mathrm{sm}}_{\pi,(m)}$$

of order n such that

(3.1) on the elliptic curve locus,

$$(\mathcal{G}_{(m)})_{S-S^{\infty,\pi}} = \iota_{\pi,(m)}(G);$$

(3.2) the cyclic subgroup $\mathcal{G}_{(m)}$ is coherent in the sense of Section 5.5.

Remark 5.9.1. If $E \to S$ is smooth, then the data (2)–(3) are uniquely determined by (1) and a $\Gamma_0(n)$-structure on E is nothing more than a cyclic S-subgroup of order n.

Remark 5.9.2. If n is invertible on S, then, by Lemma 5.6, the requirement (3.2) is superfluous.

Remark 5.9.3. If n is squarefree, then $d(m) = m$ for every m, so that $S_{\pi,(m)}$ is the open subscheme of S obtained by removing all the $S^{\infty,\pi,m'}$ with $m' \neq m$, the "decontraction" $\mathcal{E}_{\pi,(m)}$ is $E_{S_{\pi,(m)}}$ itself, and a $\Gamma_0(n)$-structure on E is nothing else than an ample cyclic S-subgroup of E^{sm} order n.

In general, the datum $\{S_{\pi,(m)}\}_{m|n}$ of (2) is equivalent to a subdivision

$$S^{\infty,\pi} = \bigsqcup\nolimits_{m|n} S^{\infty}_{\pi,(m)},$$

subject to the requirement that $S^{\infty}_{\pi,(m)} \subset S^{\infty,\pi,d(m)}$ for every m. In this notation,

$$S_{\pi,(m)} = S - \left(\bigcup\nolimits_{m' \neq m} S^{\infty}_{\pi,(m')}\right).$$

Remark 5.9.4. The subgroup $\mathcal{G}_{(m)}$ determines an ample cyclic $S_{\pi,(m)}$-subgroup

$$G_{(m)} \subset E^{\mathrm{sm}}_{S_{\pi,(m)}}$$

of order $\frac{n}{m} \cdot d(m)$ such that $(G_{(m)})_{S-S^{\infty,\pi}}$ is a standard cyclic subgroup of G. To build $G_{(m)}$, we choose an fppf cover S' of $S_{\pi,(m)}$ for which there is an object $(E' \to S', \iota')$ of $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$, let $G' \subset E'^{\mathrm{sm}}$ be the pullback of $\mathcal{G}_{(m)}$, and use Proposition 4.2.9(c) to set

$$(G_{(m)})_{S'} := (\iota')^{-1}(G'_{(n/m)\cdot d(m)}).$$

Lemma 5.7(iv) shows the agreement of the two pullbacks of $(G_{(m)})_{S'}$ to $S' \times_{S_{\pi,(m)}} S'$, and hence also the effectivity of descent to the sought $G_{(m)}$ over $S_{\pi,(m)}$, as well as the independence of the resulting $G_{(m)}$ on the choice of S' and (E', ι').

By construction and Lemma 5.7(iv), $\iota_{\pi,(m)}(G_{(m)})$ is a standard cyclic subgroup of $\mathcal{G}_{(m)}$.

The principal reason why the stack $\mathscr{X}_0(n)$ that we are about to introduce is practical to work with even when n is not squarefree is Lemma 5.12(a) below.

5.10. ***The stack*** $\mathscr{X}_0(n)$**.** In order to construct this $\mathbb{Z}$-stack, we begin by letting S be a variable scheme and by defining the categories $\mathscr{X}_0(n)(S)$.

The objects of $\mathscr{X}_0(n)(S)$ are the tuples

$$\big(E \xrightarrow{\pi} S,\ G,\ \{S_{\pi,(m)}\}_{m|n},\ \{\mathcal{G}_{(m)}\}_{m|n}\big)$$

consisting of a generalized elliptic curve $E \xrightarrow{\pi} S$ and a $\Gamma_0(n)$-structure on E.

In $\mathscr{X}_0(n)(S)$, a morphism

$$\big(E_1 \xrightarrow{\pi_1} S,\ G_1,\ \{S_{\pi_1,(m)}\},\ \{\mathcal{G}_{(m),1}\}\big) \to \big(E_2 \xrightarrow{\pi_2} S,\ G_2,\ \{S_{\pi_2,(m)}\},\ \{\mathcal{G}_{(m),2}\}\big)$$

between two tuples such that $S_{\pi_1,(m)} = S_{\pi_2,(m)}$ for every positive divisor m of n consists of:

(I) an S-isomorphism $i_E : E_1 \xrightarrow{\sim} E_2$ of generalized elliptic curves such that

$$(i_E)_{S-S^{\infty,\pi_1}}(G_1) = G_2;$$

(II) for each positive divisor m of n, an isomorphisms $i_{(m)}$ of stacks over

$$S_{\pi_1,(m)} = S_{\pi_2,(m)}$$

and an isomorphism $i_{\mathcal{E}_{(m)}}$ of generalized elliptic curves that fit into the commutative diagram

$$\begin{tikzcd}
\mathcal{E}_{\pi_1,(m)} \arrow[r, "\sim", "i_{\mathcal{E}_{(m)}}"'] \arrow[d] & \mathcal{E}_{\pi_2,(m)} \arrow[d] \\
S_{\pi_1,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m \arrow[r, "\sim", "i_{(m)}"'] & S_{\pi_2,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m
\end{tikzcd}$$

and such that $i_{\mathcal{E}_{(m)}}$ induces the isomorphism $(i_E)_{S_{\pi_1,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m}$ between the contractions of $\mathcal{E}_{\pi_1,(m)}$ and $\mathcal{E}_{\pi_2,(m)}$ with respect to

$$\mathcal{E}^{\mathrm{sm}}_{\pi_1,(m)}[d(m)] \quad \text{and} \quad \mathcal{E}^{\mathrm{sm}}_{\pi_2,(m)}[d(m)],$$

respectively, and satisfies

$$i_{\mathcal{E}_{(m)}}(\mathcal{G}_{(m),1}) = \mathcal{G}_{(m),2}.$$

There are no morphisms between tuples for which $S_{\pi_1,(m)} \neq S_{\pi_2,(m)}$ for some m.

In concrete terms, the datum $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ of (II) amounts to

(II′) an $S_{\pi_1,(m)}$-isomorphism

$$i_{(m)} : S_{\pi_1,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m \xrightarrow{\sim} S_{\pi_2,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$$

together with: for every object $(E'_1 \to S',\, \iota'_1)$ of $S_{\pi_1,(m)} \times_{\overline{\mathcal{E}\ell\ell}_{d(m)}} \overline{\mathcal{E}\ell\ell}_m$ with $i_{(m)}$-image $(E'_2 \to S',\, \iota'_2)$, a generalized elliptic curve isomorphism

$$i_{E'_1,E'_2} : E'_1 \xrightarrow{\sim} E'_2$$

that is compatible with $(i_E)_{S'}$ (via ι'_1 and ι'_2), brings the pullback of $\mathcal{G}_{(m),1}$ to the pullback of $\mathcal{G}_{(m),2}$, and whose formation commutes with isomorphisms and base change of pairs (E'_1, ι'_1).

A compatible with i_E pair of isomorphisms $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ always exists (send (E'_1, ι'_1) to $(E'_1, \iota'_1 \circ (i_E)^{-1}_{S'})$) and, thanks to $i_{\mathcal{E}_{(m)}}$, is unique up to a unique isomorphism. However, this unique $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ may not automatically respect $\mathcal{G}_{(m),1}$ and $\mathcal{G}_{(m),2}$. In practice, the uniqueness up to a unique isomorphism means that the lack of canonicity in the choice of $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ does not matter and that the construction of $\mathscr{X}_0(n)$ stays in the realm of 2-categories.

The existence of a unique $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ compatible with i_E ensures that:

- $\mathscr{X}_0(n)(S)$ is a groupoid; and
- the base change functor $\mathscr{X}_0(n)(S) \to \mathscr{X}_0(n)(S')$ along variable scheme morphisms $S' \to S$ turns $\mathscr{X}_0(n)$ into a $\mathbb{Z}$-stack for the fppf topology (see [SP 2005–, 026F] for stack axioms).

We let

$$\mathscr{X}_0(n)^{\infty} \subset \mathscr{X}_0(n) \quad \text{and} \quad \mathscr{Y}_0(n) \subset \mathscr{X}_0(n)$$

be the closed substack cut out by the degeneracy loci $S^{\infty,\pi}$ and its open complement (the elliptic curve locus), respectively. By Remark 5.9.1, there is an identification

$$\mathscr{Y}_0(n) = \mathscr{Y}_0(n)^{\text{naive}}.$$

By Remark 5.9.3, if n is squarefree, then $\mathscr{X}_0(n)$ is identified with $\mathscr{X}_0(n)^{\text{naive}}$.

For a positive divisor m of n, we let

$$\mathscr{X}_0(n)_{(m)} \subset \mathscr{X}_0(n)$$

be the open substack cut out by the subschemes $S_{\pi,(m)}$. For every tuple classified by $\mathscr{X}_0(n)_{(m)}$, the degenerate geometric fibers of E are $d(m)$-gons.

5.11. ***The contraction*** $\mathscr{X}_0(n)^{\text{naive}} \to \mathscr{X}_0(n)$**.** Let $E \xrightarrow{\pi} S$ be a generalized elliptic curve equipped with a naive $\Gamma_0(n)$-structure, i.e., with an ample cyclic S-subgroup $G \subset E^{\text{sm}}$ of order n. To build a $\Gamma_0(n)$-structure on a generalized elliptic curve $\widetilde{E} \xrightarrow{\widetilde{\pi}} S$ out of (E, G), we first construct $\widetilde{E}$ by letting $S_{\widetilde{\pi},(m)}$, for a positive divisor m of n, be the largest open subscheme of S over which the degenerate geometric fibers of E are m-gons and by letting $\widetilde{E}$ be the gluing of the contractions $c_{E^{\text{sm}}[d(m)]}(E_{S_{\widetilde{\pi},(m)}})$ along $E_{S-S^{\infty,\pi}}$. We endow $\widetilde{E}_{S-S^{\infty,\widetilde{\pi}}}$ with the cyclic subgroup $G_{S-S^{\infty,\pi}}$ of order n. This produces the data (1) and (2), so it remains to explain how to get (3).

For a fixed positive divisor m of n, each $S_{\widetilde{\pi},(m)}$-scheme S', and each generalized elliptic curve $E' \to S'$ whose degenerate geometric fibers are m-gons and that is equipped with an S'-isomorphism

$$\iota' : \widetilde{E}_{S'} = c_{E^{\mathrm{sm}}[d(m)]}(E_{S'}) \xrightarrow{\sim} c_{E'^{\mathrm{sm}}[d(m)]}(E'),$$

we endow E' with the unique cyclic S'-subgroup G' of order n supplied by Lemma 5.7. Due to the uniqueness, the formation of G' commutes with base change and with isomorphisms of pairs (E', ι'). In other words, the subgroups G' give rise to a cyclic subgroup $\mathcal{G}_{(m)} \subset \mathcal{E}^{\mathrm{sm}}_{\pi,(m)}$ of order n, which agrees with G on the elliptic curve locus due to Lemma 5.7(i), is ample due to Lemma 5.7(ii), and is coherent due to Lemma 5.7(iii). This gives the sought datum (3).

The construction of $\widetilde{E}$ and of its $\Gamma_0(n)$-structure respects isomorphisms and base change of pairs (E, G), so we obtain the sought contraction morphism

$$\mathscr{X}_0(n)^{\mathrm{naive}} \to \mathscr{X}_0(n),$$

which for each positive divisor m of n restricts to a morphism

$$\mathscr{X}_0(n)^{\mathrm{naive}}_{(m)} \to \mathscr{X}_0(n)_{(m)}.$$

The following lemma together with Lemma 5.7 is the driving force of our analysis of $\mathscr{X}_0(n)$.

Lemma 5.12. *Let m be a positive divisor of n.*

(a) *The square*

$$\begin{array}{ccc} \mathscr{X}_0(n)^{\mathrm{naive}}_{(m)} & \longrightarrow & \overline{\mathscr{E}\ell\ell}_m \\ \big\downarrow & & \big\downarrow \\ \mathscr{X}_0(n)_{(m)} & \longrightarrow & \overline{\mathscr{E}\ell\ell}_{d(m)} \end{array}$$

is Cartesian.

(b) *The map $\mathscr{X}_0(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_{d(m)}$ is representable by schemes, of finite presentation, separated, quasifinite, and flat; moreover, it is étale over $\mathbb{Z}[1/n]$.*

Proof. (a) For a generalized elliptic curve $E \xrightarrow{\pi} S$, part of the data of a $\Gamma_0(n)$-structure α on E with $S_{\pi,(m)} = S$ is the datum of a naive $\Gamma_0(n)$-structure G' on E' for every $(E' \xrightarrow{\pi'} S, \iota')$ classified by $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$. The assignment of this naive $\Gamma_0(n)$-structure gives the morphism

$$\mathscr{X}_0(n)_{(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m \to \mathscr{X}_0(n)^{\mathrm{naive}}_{(m)},$$

which, by construction of the contraction $\mathscr{X}_0(n)^{\mathrm{naive}}_{(m)} \to \mathscr{X}_0(n)_{(m)}$ in Section 5.11, is a left inverse to the induced morphism

$$\mathscr{X}_0(n)^{\mathrm{naive}}_{(m)} \to \mathscr{X}_0(n)_{(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m.$$

To prove that it is also a right inverse, we need to argue that α agrees with the $\Gamma_0(n)$-structure on E determined as in Section 5.11 by the naive $\Gamma_0(n)$-structure G' on E'. For this, the key point is the coherence requirement (3.2) on the $\mathcal{G}_{(m)}$ that is part of α: thanks to it and to the uniqueness aspect of Lemma 5.7, for every $(E'' \xrightarrow{\pi''} S, \iota'')$ classified by $S_{\pi,(m)} \times_{\overline{\mathscr{E}\ell\ell}_{d(m)}} \overline{\mathscr{E}\ell\ell}_m$, the naive $\Gamma_0(n)$-structure G'' on E'' that is part of α is also the one determined by G' through Lemma 5.7, and likewise over any S-scheme S'.

(b) We prove the asserted properties with the representability by schemes requirement replaced by representability by algebraic spaces — due to Lemma 3.2.3, this loses no generality.

By Proposition 4.2.15(a) (applied with $m = 1$ there), $\mathscr{X}_0(n)^{\text{naive}}_{(m)} \to \overline{\mathscr{E}\ell\ell}_m$ enjoys all the properties in question. Moreover, these properties are fppf local on the target (for the representability by algebraic spaces, see [SP 2005–, 04SK] or [Laumon and Moret-Bailly 2000, 10.4.2]) and, by Theorem 3.2.4(a), $\overline{\mathscr{E}\ell\ell}_m \to \overline{\mathscr{E}\ell\ell}_{d(m)}$ is surjective, flat, and of finite presentation. With the help of (a), we therefore conclude that $\mathscr{X}_0(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_{d(m)}$ inherits the properties in question. □

We are ready for the sought identification $\mathscr{X}_0(n) = \mathscr{X}_{\Gamma_0(n)}$ and for the regularity of $\mathscr{X}_{\Gamma_0(n)}$.

Theorem 5.13. (a) *The stack $\mathscr{X}_0(n)$ is Deligne–Mumford and regular. The map $\mathscr{X}_0(n) \to \mathscr{X}(1)$ that forgets the $\Gamma_0(n)$-structure and contracts with respect to the identity section induces the identification*

$$\mathscr{X}_0(n) = \mathscr{X}_{\Gamma_0(n)};$$

more precisely, $\mathscr{X}_0(n)$ and $\mathscr{X}_{\Gamma_0(n)}$ are the normalizations of $\mathscr{X}(1)$ in

$$\mathscr{Y}_0(n)_{\mathbb{Z}[\frac{1}{n}]} \cong \mathscr{Y}_{\Gamma_0(n)}\big[\tfrac{1}{n}\big].$$

(b) *The substack $\mathscr{X}_0(n)^\infty \subset \mathscr{X}_0(n)$ is a reduced relative effective Cartier divisor over* $\operatorname{Spec}\mathbb{Z}$ *that meets every irreducible component of every geometric fiber of $\mathscr{X}_0(n) \to \operatorname{Spec}\mathbb{Z}$ and is smooth over $\mathbb{Z}[1/n]$.*

Proof. (a) We will use the axiomatic Theorem 4.5.1. To apply its part (a), and hence to prove the algebraicity of $\mathscr{X}_0(n)$ and the quasicompactness and separatedness of $\Delta_{\mathscr{X}_0(n)/\mathbb{Z}}$, we use the open cover $\mathscr{X}_0(n) = \bigcup_{m|n} \mathscr{X}_0(n)_{(m)}$ and appeal to Lemma 5.12(b). To then apply Theorem 4.5.1(b), and hence to prove the regularity of $\mathscr{X}_0(n)$, we let $\mathscr{X}(n) \to \mathscr{X}_0(n)$ be the composition of the contractions

$$\mathscr{X}(n) \to \mathscr{X}_0(n)^{\text{naive}} \quad \text{and} \quad \mathscr{X}_0(n)^{\text{naive}} \to \mathscr{X}_0(n)$$

of Sections 5.1 and 5.11 and note that this composition is proper, flat, and surjective due to Section 5.1, Lemma 5.12(a), and Theorem 3.2.4(a). Finally, in order to prove that $\mathscr{X}_0(n)$ is Deligne–Mumford and $\mathscr{X}_0(n) = \mathscr{X}_{\Gamma_0(n)}$, by Theorem 4.5.1(c), we need to prove that the map

$$\mathscr{X}_0(n) \to \mathscr{X}(1)$$

is representable by algebraic spaces and that its base change to $\mathscr{Y}(1)_{\mathbb{Z}[1/n]}$ is identified with

$$\mathscr{Y}_{\Gamma_0(n)}\big[\tfrac{1}{n}\big] \to \mathscr{Y}(1)_{\mathbb{Z}[\frac{1}{n}]}.$$

Since $\mathscr{Y}_0(n) = \mathscr{Y}_0(n)^{\mathrm{naive}}$, the latter identification results from the fact that the image of $\Gamma_0(n)$ in $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$ is the stabilizer of the subgroup $\mathbb{Z}/n\mathbb{Z} \times \{0\}$ in $(\mathbb{Z}/n\mathbb{Z})^2$ (compare with the proof of Theorem 4.4.4(c)).

Due to Lemma 3.2.2(b), the representability of $\mathscr{X}_0(n) \to \mathscr{X}(1)$ will follow once we prove that, for every Artinian local algebra A over an algebraically closed field $\bar{k}$ and every $\xi \in \mathscr{X}_0(n)(\bar{k})$, no nonidentity automorphism of $\xi|_A$ maps to an identity automorphism in $\mathscr{X}(1)(A)$. More concretely, by Lemma 2.1.6, we need to prove that for every positive divisor d of n and every prime divisor p of d, there is no $\Gamma_0(n)$-structure α on the standard d-gon E over $\bar{k}$ such that some nonidentity automorphism $i \in \mu_p(A) \subset \mathrm{Aut}(E)(A)$ fixes the pullback α_A of α to A. For the sake of contradiction, we fix such α and i.

We let m be such that α has $S_{\pi,(m)} \neq \varnothing$, so, in particular, $d(m) = d$. We let $(\widetilde{E}, \iota)$ be the standard m-gon over $\bar{k}$ equipped with the canonical isomorphism $\iota : E \xrightarrow{\sim} c_{\widetilde{E}^{\mathrm{sm}}[d]}(\widetilde{E})$. Up to unique isomorphism, the pair of isomorphisms $(i_{(m)}, i_{\mathcal{E}_{(m)}})$ that extends i as in Section 5.10 sends $(\widetilde{E}_A, \iota_A)$ to $(\widetilde{E}_A, \iota_A \circ i^{-1})$, so the ample cyclic A-subgroups $\widetilde{G} \subset \widetilde{E}^{\mathrm{sm}}_A$ and $\widetilde{G}' \subset \widetilde{E}^{\mathrm{sm}}_A$ of order n that are the pullbacks of $\mathcal{G}_{(m)}$ corresponding to $(\widetilde{E}_A, \iota_A)$ and $(\widetilde{E}_A, \iota_A \circ i^{-1})$ must be equal:

$$\widetilde{G} = \widetilde{G}' \quad \text{inside } \widetilde{E}_A.$$

We replace A by an Artinian local fppf cover to assume that the automorphism $\iota_A \circ i \circ \iota_A^{-1}$ of $c_{\widetilde{E}^{\mathrm{sm}}_A[d]}(\widetilde{E}_A)$ is the contraction of an automorphism

$$\tilde{i} \in \mu_m(A) \subset \mathrm{Aut}(\widetilde{E})(A).$$

Then $\tilde{i}$ gives an isomorphism $(\widetilde{E}_A, \iota_A \circ i^{-1}) \xrightarrow{\sim} (\widetilde{E}_A, \iota_A)$, so must satisfy

$$\tilde{i}(\widetilde{G}') = \widetilde{G}, \quad \text{i.e.,} \quad \tilde{i}(\widetilde{G}) = \widetilde{G}.$$

The latter equality means that $\tilde{i}$ also lies in $\widetilde{G} \cap (\widetilde{E}^{\mathrm{sm}}_A)^0 = (\mu_{n/m})_A$, that is,

$$\tilde{i} \in \mu_{\gcd(m,n/m)}(A).$$

However, $\mu_{\gcd(m,n/m)}$ acts trivially on $c_{\widetilde{E}^{\mathrm{sm}}[d(m)]}(\widetilde{E})$ by the definition of $d(m)$ (see Section 5.2), which means that $\iota_A \circ i \circ \iota_A^{-1} = \mathrm{id}$ and contradicts the assumption that

$$i \neq \mathrm{id}.$$

(b) By the proof of (a), $\mathscr{X}(n) \to \mathscr{X}_0(n)$ is surjective, so the claim about the geometric fibers follows from the corresponding claim for $\mathscr{X}(n)^\infty \subset \mathscr{X}(n)$ proved in Proposition 4.3.2(b).

For the rest, we may work on $\mathscr{X}_0(n)_{(m)}$ and may focus on the corresponding claims for

$$\mathscr{X}_0(n)^\infty_{(m)} := \mathscr{X}_0(n)_{(m)} \cap \mathscr{X}_0(n)^\infty,$$

so it suffices to observe that $\mathscr{X}_0(n)^\infty_{(m)}$ is the preimage of $\overline{\mathscr{E}\ell\ell}^\infty_{d(m)}$ under the map

$$\mathscr{X}_0(n)_{(m)} \to \overline{\mathscr{E}\ell\ell}_{d(m)},$$

to apply Theorem 3.1.6(c)–(d) and Lemma 5.12(b), and to use the (R_0)+(S_1) criterion for reducedness. □

Chapter 6. Implications for coarse moduli spaces

The main goal of this chapter is to take advantage of the moduli interpretation of $\mathscr{X}_0(n)$ presented in Chapter 5 to prove that the coarse moduli space $X_0(n)$ is regular at the cusps (and, in fact, regular on a large open subscheme, see Theorem 6.7). This regularity is not new: [Edixhoven 1990, §1.2] uses the results of Katz and Mazur to verify via an explicit computation that the completion of $X_0(n)$ along the cusps is regular (such regularity is also a special case of an earlier assertion of Gross and Zagier [1986, Proposition III.1.4]). In contrast, the proof given below rests on Theorem 5.13(a), but requires no computation of completions.

We also exploit Lemma 3.3.1 to obtain a base change result for coarse moduli spaces X_H of arbitrary congruence level H (see Proposition 6.4). To prepare for it, we review general properties of X_H.

6.1. ***The coarse moduli space of $\mathscr{X}_H$.*** For an open subgroup $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$, the finite type Deligne–Mumford $\mathbb{Z}$-stack $\mathscr{X}_H$ of Section 4.1.2 is separated, so it has a coarse moduli space X_H (by [Keel and Mori 1997, 1.3(1)], for instance). We let

$$Y_H \subset X_H$$

be the open that is the coarse moduli space of the "elliptic curve locus"

$$\mathscr{Y}_H \subset \mathscr{X}_H.$$

We write $X(n)$, $Y_0(n)$, etc. for $X_{\Gamma(n)}$, $Y_{\Gamma_0(n)}$, etc.

Since $X(1) = \mathbb{P}^1_{\mathbb{Z}}$ (see Proposition 3.3.2) and X_H inherits $\mathbb{Z}$-properness from $\mathscr{X}_H$ (see [Rydh 2013, 6.12]), the induced map

$$X_H \to X(1)$$

is finite, so X_H is a projective $\mathbb{Z}$-scheme. Moreover, X_H inherits normality from $\mathscr{X}_H$ (see [Abramovich and Vistoli 2002, 2.2.3] and compare with the proof of Lemma 3.3.1), so $X_H \to X(1)$ is even locally free of constant rank by [EGA IV$_2$ 1965, 6.1.5]. In particular, X_H is flat and of relative dimension 1 over $\operatorname{Spec}\mathbb{Z}$ at every point.

Due to Lemma 4.1.3 (and the sentence preceding it), $\mathscr{X}_H = X_H$ whenever H is small enough. The analysis of the case of arbitrary H is facilitated by the following lemma:

Lemma 6.2 [Deligne and Rapoport 1973, IV.3.10(iii)]. *For an open subgroup $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ and an $n \geq 1$, if*

$$\Gamma(n) \subset H \quad \textit{and} \quad \bar{H} := \mathrm{Im}(H \to \mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})),$$

then X_H is identified with the categorical quotient $X(n)/\bar{H}$. □

The coarse moduli spaces Y_H and X_H have been studied extensively in [Katz and Mazur 1985], albeit with somewhat different terminology, notation, and setup. In order to put the results below in the context of the work of [Katz and Mazur 1985], we explicate the relationship between the terminology of [op. cit.] and that of the approach based on the systematic use of the theory of algebraic stacks.

Proposition 6.3. *Let $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ be an open subgroup, let $n \in \mathbb{Z}_{\geq 1}$ be such that $\Gamma(n) \subset H$, and let $\bar{H}$ be the image of H in $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$.*

(a) *The "quotient moduli problem"* $[\Gamma(n)]/\bar{H}$ (in the sense of [Katz and Mazur 1985, §7.1]) *is identified with $\mathscr{Y}_H$.*
(b) *The "coarse moduli scheme"* $\mathrm{M}([\Gamma(n)]/\bar{H})$ (in the sense of [Katz and Mazur 1985, §8.1]) *is identified with Y_H.*
(c) *The "compactified coarse moduli scheme"* $\overline{\mathrm{M}}([\Gamma(n)]/\bar{H})$ (in the sense of [Katz and Mazur 1985, §8.6]) *is identified with X_H.*

Proof. (a) In the case $H = \Gamma(n)$, the identification $[\Gamma(n)] = \mathscr{Y}(n)$ over $\mathscr{E}\ell\ell$ amounts to the definitions given in [Katz and Mazur 1985, §5.1 and §3.1] and Section 4.3.1, so the identification $[\Gamma(n)] = \mathscr{Y}_{\Gamma(n)}$ is part of Theorem 4.3.5. Therefore, in general, the desired identification over $\operatorname{Spec}\mathbb{Z}[1/n]$ results by [Katz and Mazur 1985, 7.1.3(2)], and hence also over all of $\operatorname{Spec}\mathbb{Z}$ by [Katz and Mazur 1985, 7.1.3 (5)–(6)].

(b) If $\mathscr{Y}_H$ is representable, then the claim follows from (a) and the definition of [Katz and Mazur 1985, 8.1.1]. Therefore, in general, the claim follows from Lemma 6.2.

(c) By (b), it suffices to observe that X_H is the normalization of $X(1)$ in Y_H, since $\overline{\mathrm{M}}([\Gamma(n)]/\bar{H})$ is defined as the normalization of $\mathbb{P}^1_{\mathbb{Z}} = X(1)$ in $\mathrm{M}([\Gamma(n)]/\bar{H})$. □

Before turning to the case $H = \Gamma_0(n)$, we record the following general result that holds for every H. Its part (a) has been proved in [Deligne and Rapoport 1973, VI.6.7] by a different method, and the proof given below is in essence due to Katz and Mazur. Its part (b) complements [Katz and Mazur 1985, 8.5.3].

Proposition 6.4. *Let $H \subset \mathrm{GL}_2(\widehat{\mathbb{Z}})$ be an open subgroup, and let $n \in \mathbb{Z}_{\geq 1}$ be such that $\Gamma(n) \subset H$.*

(a) *The coarse moduli space $(X_H)_{\mathbb{Z}[1/n]}$ of $(\mathscr{X}_H)_{\mathbb{Z}[1/n]}$ is $\mathbb{Z}[1/n]$-smooth.*

(b) *For any $\mathbb{Z}[1/\gcd(6, n)]$-scheme S, the canonical map from the coarse moduli space of $(\mathscr{X}_H)_S$ to $(X_H)_S$ is an isomorphism.*

Proof. Let $\bar{H}$ denote the image of H in $\mathrm{GL}_2(\mathbb{Z}/n\mathbb{Z})$.

(a) The coarse moduli space $X(n^2)$ may be covered by $\mathrm{GL}_2(\mathbb{Z}/n^2\mathbb{Z})$-invariant open subschemes that are affine over $\mathbb{Z}$ and are preimages of $\mathbb{Z}$-affine open subschemes of $X(1)$, so Lemma 6.2 and [Katz and Mazur 1985, Theorem on p. 508 in the section "Notes on Chapters 8 and 10"] reduce the proof to the case when $H = \Gamma(n^2)$. For this H, the $n = 1$ case is clear and if $n \geq 2$, then the geometric points of $\mathscr{X}(n^2)_{\mathbb{Z}[1/n]}$ have no nontrivial automorphisms by [Katz and Mazur 1985, 2.7.2(1)] and Lemma 2.1.6. Thus, if $n \geq 2$, then Lemma 3.2.2(a) ensures that

$$X(n^2)_{\mathbb{Z}[1/n]} = \mathscr{X}(n^2)_{\mathbb{Z}[1/n]}$$

and [Deligne and Rapoport 1973, IV.2.5] provides the sought $\mathbb{Z}[1/n]$-smoothness of $X(n^2)_{\mathbb{Z}[1/n]}$.

(b) We work locally on $\mathbb{Z}[1/\gcd(6, n)]$, so we assume that S is either a $\mathbb{Z}\left[\frac{1}{6}\right]$-scheme or a $\mathbb{Z}[1/n]$-scheme.

Since $\mathscr{X}_H \to \mathscr{X}(1)$ is representable, the automorphism group of every geometric point of $\mathscr{X}_H$ is of order dividing 24. Therefore, by [Olsson 2006, 2.12], étale locally on its coarse moduli space, $\mathscr{X}_H$ is the quotient of an affine scheme $\operatorname{Spec} A$ by an action of a finite group G whose order divides 24. Thus, the case when S is a $\mathbb{Z}\left[\frac{1}{6}\right]$-scheme follows from the fact that the formation of the ring of invariants A^G commutes with arbitrary base change if $\#G$ is invertible in A.

For the remainder of the proof we assume that S is a $\mathbb{Z}[1/n]$-scheme, so applying Lemma 3.3.1 with $\mathscr{X} = (\mathscr{X}_H)_{\mathbb{Z}[1/n]}$ reduces the proof to the case when $S = \operatorname{Spec} \mathbb{F}_p$ with $p \nmid n$. We therefore let X' be the coarse moduli space of $(\mathscr{X}_H)_{\mathbb{F}_p}$ and seek to prove that the finite map

$$f : X' \to (X_H)_{\mathbb{F}_p}$$

is an isomorphism. The source and the target curves of f are $\mathbb{F}_p$-smooth (equivalently, normal): the target due to (a) and the source due to the $\mathbb{F}_p$-smoothness of $(\mathscr{X}_H)_{\mathbb{F}_p}$ ensured by [Deligne and Rapoport 1973, IV.6.7]. Therefore, f is locally free by [EGA IV$_2$ 1965, 6.1.5]. To conclude that its rank is 1, it suffices to exhibit a fiberwise dense open substack $\mathscr{U} \subset \mathscr{Y}_H[1/n]$ whose coarse moduli space is of formation compatible with base change to $\mathbb{F}_p$.

We choose $\mathscr{U}$ to be the preimage of the complement of $j=0$ and $j=1728$ in $\mathbb{A}^1_{\mathbb{Z}[1/n]}$, let $\mathcal{E} \to \mathscr{U}$ denote the universal elliptic curve, and let

$$\mathcal{F} := \bar{H} \setminus \operatorname{Isom}(\mathcal{E}[n], (\mathbb{Z}/n\mathbb{Z})^2)$$

be the finite étale $\mathscr{U}$-stack of level H structures on $\mathcal{E}$ (compare with Section 4.1.2). The universal level H-structure is a section α of $\mathcal{F} \to \mathscr{U}$, as is $[-1]^*_{\mathcal{E}}(\alpha)$. Since $\mathcal{F} \to \mathscr{U}$ is finite étale, the substack $\mathscr{V} \subset \mathscr{U}$ over which $\alpha = [-1]^*_{\mathcal{E}}(\alpha)$ is both open and closed. By [Deligne 1975, 5.3(III)], the automorphism stack of $\mathcal{E}$ is the constant $\{\pm 1\}_{\mathscr{U}}$, so the open complement $\mathscr{U} \setminus \mathscr{V}$ is its own coarse moduli space, whereas the coarse moduli space of $\mathscr{V}$ is the rigidification $\mathscr{V} /\!\!/ \{\pm 1\}$ (in the notation of [AOV08 2008, Appendix]). Since the formation of $\mathscr{V} /\!\!/ \{\pm 1\}$ commutes with arbitrary base change, so does the formation of the coarse moduli space of $\mathscr{U}$. □

Remark 6.5. For a version of Proposition 6.4(a) in residue characteristics dividing n and suitable H, see [Katz and Mazur 1985, 10.10.3(5)].

Remark 6.6. In Proposition 6.4(b), for some subgroups H one cannot remove the requirement that $\gcd(6, n)$ be invertible on S. For instance, by [Česnavičius 2017, Theorem 3.2], the canonical map from the coarse moduli space of $(\mathscr{X}_{\Gamma_1(4)})_{\mathbb{F}_2}$ to $(X_{\Gamma_1(4)})_{\mathbb{F}_2}$ is not an isomorphism.

We are ready for the promised regularity of $X_0(n)$ at the cusps. Similar techniques may be used to prove analogous regularity results for $X(n)$ or $X_1(n)$ (or even for $\widetilde{X}_1(n; n')$, $X_1(n; n')$, or $X_0(n; n')$ with n and n' as in Theorem 4.6.6), but we do not explicate them because in many cases $X(n) = \mathscr{X}(n)$ and $X_1(n) = \mathscr{X}_1(n)$ (see Proposition 4.3.6 and Lemma 4.1.3), and in these cases the entire $X(n)$ or $X_1(n)$ is regular by Theorem 4.3.5 or Theorem 4.4.4(a).

Theorem 6.7. *For an $n \in \mathbb{Z}_{\geq 1}$, the open subscheme $U \subset X_0(n)$ obtained by removing the closed points corresponding to $j = 0$ or $j = 1728$ in residue characteristics dividing n is regular.*

Proof. The regularity of $X_0(n)_{\mathbb{Z}[1/n]}$ follows from Proposition 6.4(a), so it suffices to prove the regularity of the coarse moduli space of the preimage

$$\mathscr{U} \subset \mathscr{X}_0(n)$$

of the open subscheme of $\mathbb{P}^1_{\mathbb{Z}}$ obtained by removing the sections $j=0$ and $j=1728$.

By the moduli interpretation of $\mathscr{X}_0(n)$ given in Section 5.10 and Theorem 5.13(a), the constant group $\{\pm 1\}_{\mathscr{U}}$ is a subgroup of the automorphism group of the universal object of $\mathscr{U}$. In fact, due to [Deligne 1975, 5.3(III)] and the representability of $\mathscr{U} \to \mathscr{X}(1)$, this automorphism group equals $\{\pm 1\}_{\mathscr{U}}$. Therefore, the coarse moduli space of $\mathscr{U}$ is the rigidification $\mathscr{U} /\!\!/ \{\pm 1\}$. By [AOV08 2008, A.1], the map

$$\mathscr{U} \twoheadrightarrow \mathscr{U} /\!\!/ \{\pm 1\}$$

is étale, and, by Theorem 5.13(a), the stack $\mathscr{U}$ is regular, so $\mathscr{U}/\!\!/\{\pm 1\}$ is also regular, as desired. □

Remark 6.8. One may use the structure of the fibers $X_0(n)_{\mathbb{F}_p}$ with $p \mid n$ to sharpen Theorem 6.7. For instance, if n is squarefree, then, due to Proposition 6.4 and [Katz and Mazur 1985, 13.5.6 and Theorem on p. 508], in Theorem 6.7 one may require that the removed points are in addition supersingular (and likewise for general n and those removed points that lie on the reduced components of $X_0(n)_{\mathbb{F}_p}$). For a more thorough analysis of the coarse space $X_0(n)$, see [Edixhoven 1990].

We end by proving that $\mathscr{X}_0(n)^{\text{naive}}$ yields the same coarse moduli space $X_0(n)$, and hence suffices for many purposes (however, the proof of Theorem 6.7 does rely on the finer $\mathscr{X}_0(n)$ through the representability of $\mathscr{X}_0(n) \to \mathscr{X}_0(1)$).

Proposition 6.9. *For every $n \in \mathbb{Z}_{\geq 1}$, the contraction morphism*

$$\mathscr{X}_0(n)^{\text{naive}} \to \mathscr{X}_0(n)$$

defined in Section 5.11 *induces an isomorphism on coarse moduli spaces.*

Proof. The coarse moduli space $X_0(n)'$ of $\mathscr{X}_0(n)^{\text{naive}}$ exists due to the finiteness of the diagonal of $\mathscr{X}_0(n)^{\text{naive}}$ supplied by Theorem 4.6.4(a) (see [Rydh 2013, 6.12]). As in Section 6.1, the map

$$X_0(n)' \to \mathbb{P}^1_{\mathbb{Z}}$$

is finite, so, since $\mathscr{Y}_0(n)^{\text{naive}} = \mathscr{Y}_0(n)$, it suffices to prove that $X_0(n)'$ is normal.

For the normality, we work Zariski locally on $X_0(n)'$ and note that each open substack

$$\mathscr{U} \subset \mathscr{X}_0(n)^{\text{naive}}$$

that has an affine coarse moduli space $\operatorname{Spec} A$ satisfies $A = \Gamma(\mathscr{U}, \mathscr{O}_{\mathscr{U}})$ by the universal property for maps to $\mathbb{A}^1_{\mathbb{Z}}$. To then see that $\Gamma(\mathscr{U}, \mathscr{O}_{\mathscr{U}})$ is integrally closed in its total ring of fractions it suffices to use the normality of $\mathscr{U}$ supplied by Theorem 4.6.4(a) and the fact that generizations lift along smooth morphisms from algebraic spaces to $\mathscr{U}$ (see [Laumon and Moret-Bailly 2000, 5.7.1]). □

Remark 6.10. The same proof shows that, in the notation of Section 4.6, for every $n, n' \in \mathbb{Z}_{\geq 1}$ the coarse moduli spaces of $\mathscr{X}_1(n; n')$ and $\mathscr{X}_0(n; n')$ agree with those of $\mathscr{X}_{\Gamma_1(n;n')}$ and $\mathscr{X}_{\Gamma_0(n;n')}$.

Acknowledgements

I thank Pierre Deligne for correspondence about the moduli interpretation in the $\Gamma_0(n)$ case and for permitting me to make his letter [Deligne 2015] available. The modular description of $\mathscr{X}_0(n)$ presented in Chapter 5 is inspired by the ideas explained there. I thank the referee for a very careful reading of the manuscript and for numerous helpful suggestions. I thank the MathOverflow community — the reading of several anonymous discussions has been useful while working on some aspects of this paper. I thank Rebecca Bellovin, George Boxer, Brian Conrad, Bas Edixhoven, Benedict Gross, Dino Lorenzini, Martin Olsson, Ken Ribet, and Sug Woo Shin for helpful conversations or correspondence. I thank the Miller Institute for Basic Research in Science at the University of California Berkeley for its support.

References

[Abramovich and Vistoli 2002] D. Abramovich and A. Vistoli, "Compactifying the space of stable maps", *J. Amer. Math. Soc.* **15**:1 (2002), 27–75. MR Zbl

[ACV03 2003] D. Abramovich, A. Corti, and A. Vistoli, "Twisted bundles and admissible covers", *Comm. Algebra* **31**:8 (2003), 3547–3618. MR Zbl

[Aoki 2006a] M. Aoki, "Erratum: "Hom stacks" [Manuscripta Math. **119**:1 (2006), 37–56]", *Manuscripta Math.* **121**:1 (2006), 135. MR

[Aoki 2006b] M. Aoki, "Hom stacks", *Manuscripta Math.* **119**:1 (2006), 37–56. MR Zbl

[AOV08 2008] D. Abramovich, M. Olsson, and A. Vistoli, "Tame stacks in positive characteristic", *Ann. Inst. Fourier* (*Grenoble*) **58**:4 (2008), 1057–1091. MR Zbl

[Artin 1969] M. Artin, "Algebraization of formal moduli, I", pp. 21–71 in *Global Analysis (Papers in Honor of K. Kodaira)*, Univ. Tokyo Press, 1969. MR Zbl

[Atiyah and Macdonald 1969] M. F. Atiyah and I. G. Macdonald, *Introduction to commutative algebra*, Addison-Wesley, Reading, MA, 1969. MR Zbl

[BLR90 1990] S. Bosch, W. Lütkebohmert, and M. Raynaud, *Néron models*, Ergebnisse der Mathematik (3) **21**, Springer, 1990. MR Zbl

[Bourbaki 1965] N. Bourbaki, *Éléments de mathématique, Fascicule XXXI: Algèbre commutative, Chapitre 7: Diviseurs*, Actualités Scientifiques et Industrielles **1314**, Hermann, Paris, 1965. MR Zbl

[Česnavičius 2017] K. Česnavičius, "Coarse base change fails for some modular curves", *Algebr. Geom.* **4**:4 (2017), 444–451. MR

[Conrad 2005] B. Conrad, "The Keel–Mori theorem via stacks", preprint, Stanford, 2005, http://math.stanford.edu/~conrad/papers/coarsespace.pdf.

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

[Conrad 2014] B. Conrad, "Reductive group schemes", pp. 93–444 in *Autour des schémas en groupes, I*, Panor. Synthèses **42/43**, Soc. Math. France, Paris, 2014. MR Zbl

[Deligne 1975] P. Deligne, "Courbes elliptiques: formulaire d'après J. Tate", pp. 53–73 in *Modular functions of one variable, IV* (Antwerp, 1972), edited by B. J. Birch and W. Kuyk, Lecture Notes in Math. **476**, Springer, 1975. MR Zbl

[Deligne 2015] P. Deligne, "Letter to Česnavičius", 15th of July 2015, http://www.math.uni-bonn.de/people/kestutis/Deligne-2015-07-15.pdf.

[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, II* (Antwerp, 1972), edited by P. Deligne and W. Kuyk, Lecture Notes in Math. **349**, Springer, 1973. MR Zbl

[Edixhoven 1990] B. Edixhoven, "Minimal resolution and stable reduction of $X_0(N)$", *Ann. Inst. Fourier (Grenoble)* **40**:1 (1990), 31–67. MR Zbl

[Edixhoven 2001] B. Edixhoven, "Modular parametrizations at primes of bad reduction", unfinished manuscript, 2001. dated Sept. 7.

[EGA II 1961] A. Grothendieck, "Eléments de géométrie algébrique, II: Étude globale élémentaire de quelques classes de morphismes", *Inst. Hautes Études Sci. Publ. Math.* **8** (1961), 5–222. MR Zbl

[EGA III_1 1961] A. Grothendieck, "Eléments de géométrie algébrique, III: Étude cohomologique des faisceaux cohérents, I", *Inst. Hautes Études Sci. Publ. Math.* **11** (1961), 5–167. MR Zbl

[EGA IV_2 1965] A. Grothendieck, "Eléments 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 Zbl

[EGA IV_3 1966] A. Grothendieck, "Eléments 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 Zbl

[EGA IV_4 1967] A. Grothendieck, "Eléments 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 Zbl

[Ferrand 2003] D. Ferrand, "Conducteur, descente et pincement", *Bull. Soc. Math. France* **131**:4 (2003), 553–585. MR Zbl

[Fulton and Olsson 2010] W. Fulton and M. Olsson, "The Picard group of $\mathscr{M}_{1,1}$", *Algebra Number Theory* **4**:1 (2010), 87–104. MR Zbl

[Gross and Zagier 1986] B. H. Gross and D. B. Zagier, "Heegner points and derivatives of L-series", *Invent. Math.* **84**:2 (1986), 225–320. MR Zbl

[Illusie 2005] L. Illusie, "Grothendieck's existence theorem in formal geometry", pp. 179–233 in *Fundamental algebraic geometry*, Math. Surveys Monogr. **123**, Amer. Math. Soc., Providence, RI, 2005. MR Zbl

[Katz and Mazur 1985] N. M. Katz and B. Mazur, *Arithmetic moduli of elliptic curves*, Annals of Mathematics Studies **108**, Princeton University Press, 1985. MR Zbl

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

[Laumon and Moret-Bailly 2000] G. Laumon and L. Moret-Bailly, *Champs algébriques*, Ergebnisse der Mathematik (3) **39**, Springer, 2000. MR Zbl

[MFK94 1994] D. Mumford, J. Fogarty, and F. Kirwan, *Geometric invariant theory*, 3rd ed., Ergebnisse der Mathematik (2) **34**, Springer, 1994. MR Zbl

[Olsson 2006] M. C. Olsson, "$\underline{\mathrm{Hom}}$-stacks and restriction of scalars", *Duke Math. J.* **134**:1 (2006), 139–164. MR Zbl

[Romagny 2005] M. Romagny, "Group actions on stacks and applications", *Michigan Math. J.* **53**:1 (2005), 209–236. MR Zbl

[Rydh 2013] D. Rydh, "Existence and properties of geometric quotients", *J. Algebraic Geom.* **22**:4 (2013), 629–669. MR Zbl

[SGA $3_{\mathrm{I(new)}}$ 2011] M. Demazure and A. Grothendieck, *Schémas en groupes, Tome I: Propriétés générales des schémas en groupes, Exposés I–VII* (Séminaire de Géométrie Algébrique du Bois Marie 1962–1964), Société Mathématique de France, 2011. MR Zbl

[SGA 7_{I} 1972] A. Grothendieck, *Groupes de monodromie en géométrie algébrique, I: Exposés I–II, VI–IX* (Séminaire de Géométrie Algébrique du Bois Marie 1967–1969), Lecture Notes in Math. **288**, Springer, Berlin, 1972. MR Zbl

[SP 2005–] P. Belmans, A. J. de Jong, et al., "The Stacks project", electronic reference, 2005–, http://stacks.math.columbia.edu.

Communicated by Brian Conrad
Received 2016-01-06 Revised 2017-07-13 Accepted 2017-10-05

kestutis@math.uni-bonn.de *Mathematisches Institut, Universität Bonn, D-53115 Bonn, Germany*

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2091

Elementary equivalence versus isomorphism, II

Florian Pop

In this note we give sentences $\boldsymbol{\vartheta}_K$ in the language of fields which describe the isomorphy type of K among finitely generated fields, provided the Kronecker dimension $\dim(K)$ satisfies $\dim(K) < 3$. This extends results by Rumely (1980) concerning global fields; see also Scanlon (2008).

1. Introduction

We begin by recalling Rumely's result [1980] showing that for every global field k there exists a sentence $\boldsymbol{\vartheta}_k^{\mathrm{Ru}}$ which *characterizes the isomorphy type of k among global fields*, i.e., if l is any *global field,* then $\boldsymbol{\vartheta}_k^{\mathrm{Ru}}$ holds in l if and only if $l \cong k$ as fields.

It is one of the *main open questions* in the first-order theory of finitely generated fields whether a fact similar to Rumely's result mentioned above holds for all finitely generated fields K. We notice that the question above is related to, but much stronger than, the still open *elementary equivalence versus isomorphism problem,* which asks whether the isomorphism type of every finitely generated field K is encoded in the whole first-order theory $\mathfrak{Th}(K)$ of K; see, e.g., [Pop 2003] for details and literature about this, as well as [Scanlon 2008].[1]

In the present note we show that the answer to the above question is positive for finitely generated fields K having Kronecker dimension $\dim(K) < 3$, which are precisely the finite fields, the global fields, and the function fields of (algebraic) curves over global fields.

Supported by the John Templeton Foundation Grant ID 13394 and the NSF grant DMS-1265290.
MSC2010: primary 11G30, 14H25; secondary 03C62, 11G99, 12F20, 12G10, 12L12, 13F30.
Keywords: elementary equivalence versus isomorphism, first-order definability, finitely generated fields, Milnor K-groups, Galois étale cohomology, Kato's higher local-global principles.

[1] To the best of my knowledge, Bjorn Poonen was among the first to ask whether Rumely's result [1980] might hold in higher dimensions. It was claimed in [Scanlon 2008] that the answer to this question is positive for all finitely generated fields. Unfortunately, the proof has a gap (see *Erratum,* J. Amer. Math. Soc. **24**:3 (2011), p. 917). Nevertheless, Scanlon's work appears to reduce the problem to "definability of valuations".

Main Result. *For every finitely generated field K having* $\dim(K) < 3$, *there exists a first-order sentence in the language of fields $\boldsymbol{\vartheta}_K$ such that for finitely generated fields L, the sentence $\boldsymbol{\vartheta}_K$ holds in L if and only if $L \cong K$ as fields.*

The more precise form of the result is as follows: Recall that by one of the main results in [Pop 2002], for every $d \geq 0$, there exists a sentence φ_d (in the language of fields) such that for all finitely generated fields K, φ_d holds in K if and only if $\dim(K) = d$. In particular, if K is any finitely generated field, then φ_0 holds in K if and only if K is a finite field, φ_1 holds in K if and only if K is a global field, and finally φ_2 holds in K if and only if $\dim(K) = 2$.

In particular, given a global field K, consider the sentence $\boldsymbol{\vartheta}_K$ given by $\varphi_1 \wedge \boldsymbol{\vartheta}_K^{\mathrm{Ru}}$. Then if $\boldsymbol{\vartheta}_K$ holds in a finitely generated field L, one has the following: First, $\dim(L) = 1$, because φ_1 holds in L, and hence L is a global field. Second, $L \cong K$ because $\boldsymbol{\vartheta}_K^{\mathrm{Ru}}$ holds in the global field L.

In the case $\dim(K) = 2$, let $k_0 = K^{\mathrm{abs}}$ be the constant subfield of K, i.e., the set of elements of K which are algebraic over the prime field of K. Then k_0 is finite if and only if $\mathrm{char}(K) > 0$, and if so, K is the function field of a projective smooth geometrically integral surface over k_0. Letting (t_0, t_1) be a separable transcendence basis of K, there exists $t_2 \in K$ such that $K = k_0(t_0, t_1, t_2)$, with t_0, t_1, t_2 satisfying an absolutely irreducible polynomial $f(T_0, T_1, T_2)$ over k_0. And if $\mathrm{char}(K) = 0$, then K is the function field of a projective smooth k_0-curve, and for every nonconstant $t_1 \in K$ there exists $t_2 \in K$ such that $K = k_0(t_1, t_2)$, with t_1, t_2 satisfying an irreducible polynomial $f(T_1, T_2) \in k_0[T_1, T_2]$. The precise result proven will be the following; see Section 5 for proofs.

Theorem 1.1. *Let K be a finitely generated field. The following hold*:

(1) *For every finite field k_0 and absolutely irreducible polynomial $f = f(T_0, T_1, T_2)$ over k_0, there exists a formula $\psi_{k_0,f}(\mathrm{t}_0, \mathrm{t}_1, \mathrm{t}_2)$ with free variables $\mathrm{t}_0, \mathrm{t}_1, \mathrm{t}_2$ such that the following are equivalent*:
   - (i) *The sentence $\boldsymbol{\vartheta}_K$ defined by $\exists\, \mathrm{t}_0, \mathrm{t}_1, \mathrm{t}_2\ \psi_{k_0,f}(\mathrm{t}_0, \mathrm{t}_1, \mathrm{t}_2)$ holds in K.*
   - (ii) *There exist $t_0, t_1, t_2 \in K$ such that $K = k_0(t_0, t_1, t_2)$ and $f(t_0, t_1, t_2) = 0$.*
     - ($*$) *In particular, suppose that $\boldsymbol{\vartheta}_K$ holds in K. Then for all finitely generated fields L, $\boldsymbol{\vartheta}_K$ holds in L if and only if $L \cong K$ as abstract fields.*

(2) *For every number field k_0 and absolutely irreducible polynomial $f = f(T_1, T_2)$ over k_0, there exists a formula $\psi_{k_0,f}(\mathrm{t}_1, \mathrm{t}_2)$ with free variables $\mathrm{t}_1, \mathrm{t}_2$ such that the following are equivalent*:
   - (i) *The sentence $\boldsymbol{\vartheta}_K$ defined by $\exists\, \mathrm{t}_1, \mathrm{t}_2\ \psi_{k_0,f}(\mathrm{t}_1, \mathrm{t}_2)$ holds in K.*
   - (ii) *There exist $t_1, t_2 \in K$ such that $K = k_0(t_1, t_2)$ and $f(t_1, t_2) = 0$.*
     - ($*$) *In particular, suppose that $\boldsymbol{\vartheta}_K$ holds in K. Then for all finitely generated fields L, $\boldsymbol{\vartheta}_K$ holds in L if and only if $L \cong K$ as abstract fields.*

The result above is based on and uses in an essential way, among other things, previous results by Rumely, Poonen, and Pop. First, the above-mentioned sentences φ_d single out the finite fields, the global fields, and the fields of curves over global fields among all finitely generated fields K. Second, Poonen [2007] showed that there exists a predicate, i.e., formula $\psi^{\text{abs}}(\mathfrak{x})$ with one free variable $\mathfrak{x}$ such that for all finitely generated fields K, one has $k_0 := K^{\text{abs}} = \{x \in K \mid \psi^{\text{abs}}(x) \text{ is true in } K\}$. Further, techniques developed in [Poonen 2007] (using [Pop 2002] as well) give formulas $\psi_r(\mathfrak{x}_1, \dots, \mathfrak{x}_r, \mathfrak{x}_{r+1})$ with $r+1$ free variables such that for $x_1, \dots, x_{r+1} \in K$, one has that $\psi_r(x_1, \dots, x_r, x_{r+1})$ holds in K if and only if $x_1, \dots, x_r$ are algebraically independent over k_0, but $x_1, \dots, x_r, x_{r+1}$ are not. Hence, for $x_1, \dots, x_r \in K$ algebraically independent over k_0, the relative algebraic closure of $k_0(x_1, \dots, x_r)$ in K is given by $L := \{x_{r+1} \in K \mid \psi_r(x_1, \dots, x_r, x_{r+1}) \text{ holds in } K\}$. Finally, Poonen [2007] showed that there exists a sentence ψ_0 which holds in a finitely generated field K if and only if $\operatorname{char}(K) = 0$.

Hence, in the case $\dim(K) = 2$, one has the following: First, $k_0 = K^{\text{abs}}$ is finite if and only if $\operatorname{char}(K) > 0$ if and only if ψ_0 does not hold in K. If so, K is the function field of a projective smooth surface over k_0. Therefore, there exist separable transcendence bases t_0, t_1 of $K\,|\,k_0$ satisfying that the relative algebraic closure $k \subset K$ of $k_0(t_0)$ in K is a global function field, and furthermore that $K\,|\,k$ is the function field of a (projective smooth) geometrically integral k-curve X. Thus $K = k(t_1, t_2)$ for a properly chosen t_2. Second, if K has characteristic zero, then $k := k_0 = K^{\text{abs}}$ is a number field, and K is the function field of a projective smooth geometrically integral k-curve X. So for $t_1 \in K \setminus k$, and properly chosen $t_2 \in K$, one has $K = k(t_1, t_2)$. Hence, one can deduce Theorem 1.1 above from the following theorem; see Section 5 for detailed proofs.

Theorem 1.2. *The k-valuations of function fields $K = k(X)$ of projective smooth geometrically integral k-curves X over global fields k are* uniformly first-order definable. *In particular, there exist formulas* $\deg_N(\mathfrak{t})$, $\psi^R(\mathfrak{t},\mathfrak{t}')$, $\psi^0(\mathfrak{t},\mathfrak{t}')$, *with free variables* $\mathfrak{t}$, $\mathfrak{t}'$, *such that for every $K\,|\,k$ as above and $t \in K \setminus k$, the following hold*:

(a) $\deg_N(t)$ *is true in K if and only if t has degree N as a function of $K\,|\,k$, i.e.,* $[K : k(t)] = N$.

(b) $R := \{t' \in K \mid \psi^R(t, t') \text{ is true in } K\}$ *is the integral closure of $k[t]$ in K.*

(c) $k[t] = \{t' \in K \mid \psi^0(t, t') \text{ is true in } K\}$.

We mention that the formulas $\deg_N(\mathfrak{t})$, $\psi^R(\mathfrak{t},\mathfrak{t}')$, $\psi^0(\mathfrak{t},\mathfrak{t}')$ are quite explicit; see Section 5. In particular, so is Theorem 5.3, which is slightly more general than Theorem 1.2 above. The main technical tool in the proof is one of Kato's higher Hasse local-global principles (LGPs) for H^3; see Theorem 2.1 below. If similar LGPs would be available in higher dimensions, it would be possible to extend the methods of this paper to higher dimensions.

2. Reviewing well known facts

2A. ***The Hasse–Brauer–Noether local-global principle.*** We recall briefly the famous Hasse–Brauer–Noether LGP for the Brauer group of a global field k. Let $\mathbb{P}(k)$ be the set of nontrivial places of k. For $v \in \mathbb{P}(k)$, we denote by k_v the completion of k with respect to v. Then k_v is a locally compact (nondiscrete) field, and the Brauer group $\mathrm{Br}(k_v)$ of k_v admits a canonical embedding $\mathrm{inv}_v : \mathrm{Br}(k_v) \to \mathbb{Q}/\mathbb{Z}$, called the *invariant* (*isomorphism*), satisfying the following:

(a) If $k_v = \mathbb{C}$, then $\mathrm{Br}(k_v) = 0$ and inv_v is the trivial map.

(b) If $k_v = \mathbb{R}$, then $\mathrm{inv}_v : \mathrm{Br}(k_v) \to \frac{1}{2}\mathbb{Z}/\mathbb{Z} \subset \mathbb{Q}/\mathbb{Z}$ is an isomorphism.

(c) In the remaining cases, $\mathrm{inv}_v : \mathrm{Br}(k_v) \to \mathbb{Q}/\mathbb{Z}$ is an isomorphism.

The Hasse–Brauer–Noether LGP asserts that the canonical sequence

$$0 \to \mathrm{Br}(k) \to \bigoplus_v \mathrm{Br}(k_v) \to \mathbb{Q}/\mathbb{Z} \to 0$$

is exact. Here, the first map is the direct sum of all the canonical restriction maps $\mathrm{Br}(k) \to \mathrm{Br}(k_v)$; thus implicitly, for every division algebra D over k there exist only finitely many v such that $D \otimes_k k_v$ is not a matrix algebra. And the second map is the sum of the invariant morphisms.

Moreover, if ${}_n(\ )$ denotes the n-torsion, then identifying the n-torsion in $\mathbb{Q}/\mathbb{Z}$ canonically with $\mathbb{Z}/n$, the above exact sequence gives rise canonically to an exact sequence

$$0 \to {}_n\mathrm{Br}(k) \to \bigoplus_v {}_n\mathrm{Br}(k_v) \to \mathbb{Z}/n \to 0.$$

2B. ***Hasse higher LGPs (after Kato).*** It is a fundamental observation by Kato [1986] that the above local-global principle has higher dimensional variants as follows: First, following [Kato 1986], for every positive integer n, say $n = mp^r$ with p the characteristic, and an integer twist i, one sets $\mathbb{Z}/n(0) = \mathbb{Z}/n$, and defines in general $\mathbb{Z}/n(i) := \mu_m^{\otimes i} \oplus W_r\,\Omega^i_{\log}[-i]$, where $W_r\,\Omega_{\log}$ is the logarithmic part of the de Rham–Witt complex on the étale site; see [Illusie 1979] for details. In this notation, for every (finitely generated) field K one has

$$\mathrm{H}^1(K, \mathbb{Z}/n) = \mathrm{Hom}_{\mathrm{cont}}(G_K, \mathbb{Z}/n), \qquad \mathrm{H}^2(K, \mathbb{Z}/n(1)) = {}_n\mathrm{Br}(K),$$

where G_K is the absolute Galois group of K. Hence, the cohomology groups $\mathrm{H}^{i+1}(K, \mathbb{Z}/n(i))$ have a particular arithmetical significance for $i = 0, 1$. Further, in this notation, the Hasse–Brauer–Noether LGP is a local-global principle for the cohomology group $\mathrm{H}^2(K, \mathbb{Z}/n(1))$, and note that global fields have Kronecker dimension $d = \dim(K) = 1$.

This led Kato to the fundamental idea that for finitely generated fields K of Kronecker dimension d there should exist similar LGPs for $\mathrm{H}^{d+1}(K, \mathbb{Z}/n(d))$. And

Kato [1986] proved that such higher dimension LGPs do indeed hold for $d = 2$, i.e., for $\mathrm{H}^3(K, \mathbb{Z}/n(2))$, where K is the function field of an integral curve over some global field, or equivalently the function field of an integral two dimensional scheme of finite type.

We describe below one of Kato's local-global principles for $\mathrm{H}^3(K, \mathbb{Z}/n(2))$, and will use that LGP in the cases $n = 2$, $\mathrm{char}(K) \neq 2$ as well as $n = 3$, $\mathrm{char}(K) = 2$. The situation is as follows. Let k be a global field, and $K\,|\,k$ be the function field of a complete smooth geometrically integral k-curve X. Let S be the arithmetical complete normal curve with function field $\kappa(S) = k$; hence $S = \mathrm{Spec}\,\mathcal{O}_k$ if k is a number field, and S is the unique projective smooth curve with function field k if k is a global field of positive characteristic. Then by Abhyankar's regularization theorems of surfaces [1965], $X \to \mathrm{Spec}\,k$ is the generic fiber of a proper morphism $\mathcal{X} \to S$ of regular schemes (and having further properties, e.g., having NCD on $\mathcal{X}$ as reduced fibers, etc.). For $i \geq 0$, we denote by $\mathcal{X}_i \subset \mathcal{X}$ the points of dimension i in $\mathcal{X}$. Then for $x \in \mathcal{X}$ one has:

(a) $x \in \mathcal{X}_0 \Leftrightarrow \mathcal{O}_x$ is a two dimensional local ring $\Leftrightarrow \kappa(x)$ is a finite field.

(b) $x \in \mathcal{X}_1 \Leftrightarrow \mathcal{O}_x$ is a discrete valuation ring $\Leftrightarrow \kappa(x)$ is a global field.

For $s \in S_0$ we denote by v_s the canonical valuation of $\mathcal{O}_s$ and by k_s the completion of k at s. For $x_1 \in \mathcal{X}_1$ we denote by v_{x_1} the canonical valuation of $\mathcal{O}_{x_1}$, and by K_{x_1} the completion of K at x_1. Notice that $x_1 \mapsto s$ under $\mathcal{X} \to S$ if and only if $\mathcal{O}_s \prec \mathcal{O}_{x_1}$, that is, the local ring $\mathcal{O}_s$ is dominated by the local ring $\mathcal{O}_{x_1}$ under $k \hookrightarrow K$.

Next let L be an arbitrary field, and recall the canonical isomorphism (generalizing the classical Kummer theory isomorphism) $h^1 : L^\times/n \to \mathrm{H}^1(L, \mathbb{Z}/n(1))$.[2] As explained in [Kato 1986, §1], the isomorphism h^1 gives rise canonically for all $q \neq 0$ to morphisms[3]

$$\begin{aligned} h^q &: K_q^{\mathrm{M}}(L)/n \to \mathrm{H}^q(L, \mathbb{Z}/n(q)), \\ \{a_1, \dots, a_q\}/n &\mapsto h^1(a_1) \cup \cdots \cup h^1(a_q) =: a_1 \cup \cdots \cup a_q. \end{aligned}$$

Let $\mathfrak{v}$ be a discrete valuation of L. Then for every uniformizing parameter $\pi \in L$ at $\mathfrak{v}$, one defines the *boundary homomorphism*

$$\partial_{\mathfrak{v}} : \mathrm{H}^{q+1}(L, \mathbb{Z}/n(q+1)) \to \mathrm{H}^q(\lambda, \mathbb{Z}/n(q))$$

by $\pi \cup a_1 \cdots \cup a_q \mapsto a_1 \cup \cdots \cup a_q$ and $a_0 \cup a_1 \cdots \cup a_q \mapsto 0$, provided all $a_0, a_1, \dots, a_q$ are $\mathfrak{v}$-units. We notice that in general, this homomorphism depends on the uniformizing parameter π. Further, if the Galois action on $\mathbb{Z}/n(1)$ is trivial, then

[2] Recall that for every abelian group A, we denote $A/n := A/(nA)$.

[3] By the (now proven) Milnor–Bloch–Kato conjecture, h^q are isomorphisms. Nevertheless, that fact in its full generality is not needed here, because one could work as well with the subgroup generated by symbols $\mathrm{H}^q_{\cup}(L, \mathbb{Z}/n(q)) \subseteq \mathrm{H}^q(L, \mathbb{Z}/n(q))$.

all Galois modules $\mathbb{Z}/n(q)$ are actually isomorphic to $\mathbb{Z}/n$, and $\partial_{\mathfrak{v}}$ gives rise to morphisms

$$\partial_{\mathfrak{v}} : \mathrm{H}^{q+2}(L, \mathbb{Z}/n(q+1)) \to \mathrm{H}^{q+1}(\lambda, \mathbb{Z}/n(q)).$$

We will use two instances of these homomorphisms for $q \le 2$ and L a finitely generated field of Kronecker dimension equal to q containing μ_{2n} (thus having no orderings).[4] First, let k be a global field, $K\,|\,k$ be the function field of a complete smooth k-curve X, and $\mathcal{X} \to S$, etc., be as introduced above. For $x_1 \in \mathcal{X}_1$, let $\mathfrak{v} := v_{x_1}$ be the corresponding discrete valuation of K. The boundary homomorphisms we will consider are

$$\partial_{x_1} : \mathrm{H}^3(K, \mathbb{Z}/n(2)) \to \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)).$$

For later use we notice that for $f, g, h \in K$ such that g, h are v_{x_1}-units, one has

$$\partial_{x_1}(f \cup g \cup h) = v_{x_1}(f) \cdot \bar{g} \cup \bar{h} \quad \text{in } \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)),$$

where $u \mapsto \bar{u}$ is the residue map $\mathcal{O}_{x_1} \to \kappa(x_1)$. In particular, if the v_{x_1}-values of $f, g, h \in K$ are all divisible by n (for instance, if f, g, h are all v_{x_1}-units), then $\partial_{x_1}(f \cup g \cup h) = 0$.

For $q = 1$, $L = \kappa(x_1)$ is the residue field of K at some $x_1 \in \mathcal{X}_1$ and $\mathfrak{v}$ is some finite place $\mathfrak{p}_0$ of $\kappa(x_1)$. Thus the boundary homomorphisms we will consider are

$$\partial_{\mathfrak{p}_0} : \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(\mathfrak{p}_0), \mathbb{Z}/n(0)) = \mathbb{Z}/n.$$

Notice that $\partial_{\mathfrak{p}_0}$ is nothing but the local component of the Hasse–Brauer–Noether LGP for the global field $\kappa(x_1)$ defined by the place $\mathfrak{p}_0$.

Following [Kato 1986, §1], for all $x_1 \in \mathcal{X}_1$, $x_0 \in \mathcal{X}_0$, one defines *boundary homomorphisms*

$$\partial_{x_1 x_0} : \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$$

as follows. First, if $x_0 \notin \overline{\{x_1\}}$, set $\partial_{x_1 x_0} = 0$. Second, if $x_0 \in \overline{\{x_1\}}$, proceed as follows: Recall that $\mathcal{O}_{x_0}$ is a two dimensional regular local ring, and set $\mathcal{X}_{x_0} := \operatorname{Spec} \mathcal{O}_{x_0} \hookrightarrow \mathcal{X}$. Then $x_0 \in \overline{\{x_1\}}$ if and only if $x_1 \in \mathcal{X}_{x_0}$. If so, then $\mathcal{O}_{x_1}$ is some localization of $\mathcal{O}_{x_0}$ and the image $\overline{\mathcal{O}}_{x_0} \subset \kappa(x_1)$ of $\mathcal{O}_{x_0}$ under the projection $\mathcal{O}_{x_1} \to \kappa(x_1)$ is a local Noetherian ring of Krull dimension one. Thus its integral closure $\widetilde{\mathcal{O}}_{x_0}$ in $\kappa(x_1)$ is a Dedekind domain with finitely many maximal ideals $\mathfrak{p}_i$, and thus a principal ideal domain. Further, every completion $\kappa(x_1)_{\mathfrak{p}_i}$ is a localization of the global field $\kappa(x_1)$ and the residue fields $\kappa(\mathfrak{p}_i)\,|\,\kappa(x_0)$ are finite fields. Kato defined $\partial_{x_1 x_0}$ as follows, where the last map is the sum of the corestriction maps:

$$\partial_{x_1 x_0} : \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \bigoplus_{\mathfrak{p}_i} \mathrm{H}^1(\kappa(\mathfrak{p}_i), \mathbb{Z}/n) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n).$$

[4]Recall that in these cases, the equality $\mathrm{H}^3_{\cup} = \mathrm{H}^3$ has been known for a while already.

Finally, one of the local-global principles Kato gives — which is essential for the methods of this paper — is the following; see [Kato 1986, p. 145, Corollary].

Theorem 2.1. *With the above notation, suppose that K has no orderings, e.g., $\mu_{2n} \subset K$. Then via the obvious direct sums of the above boundary homomorphisms one gets a long exact sequence of the following form, where the last map is given by the sum*:

$$0 \to \mathrm{H}^3(K, \mathbb{Z}/n(2)) \to \bigoplus_{x_1 \in \mathcal{X}_1} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \bigoplus_{x_0 \in \mathcal{X}_0} \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n) \to \mathbb{Z}/n \to 0.$$

In particular, recalling that $\mathcal{X}_{x_0}$ is the set of all the $x_1 \in \mathcal{X}_1$ such that $x_0 \in \overline{\{x_1\}}$, the map $\mathrm{H}^3(K, \mathbb{Z}/n(2)) \to \bigoplus_{x_1 \in \mathcal{X}_{x_0}} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$ is trivial for each $x_0 \in \mathcal{X}_0$.

2C. ***An arithmetical application/interpretation.*** In the following discussion, suppose that the Galois action on $\mathbb{Z}/n(1)$ is trivial, so that $\mathbb{Z}/n(q)$ are isomorphic to $\mathbb{Z}/n$ as Galois modules.

(1) Recall ${}_n\mathrm{Br}(L) = H^2(L, \mathbb{Z}/n(2))$ and $H^3(L, \mathbb{Z}/n(3))$ are generated by symbols $a \cup b$ and $a \cup b \cup c$, respectively, with $a, b, c \in L^\times$.

(2) Now suppose that n is a prime number. For $a, b \in L^\times$ consider the field extension $L_a | L$ defined by $h^1(a) \in \mathrm{H}^1(L, \mathbb{Z}/n(1))$, the norm map $N_a : L_a^\times \to L^\times$, and the cyclic algebra $A_{a,b}$ with $[A_{a,b}] = a \cup b \in \mathrm{H}^2(L, \mathbb{Z}/n(1))$. Then $a \cup b \in \mathrm{H}^2(L, \mathbb{Z}/n(1))$ is trivial if and only if $b \in N_a(L_a^\times)$. Furthermore, if $A_{a,b}$ is a division algebra, let $N_{a,b} : A_{a,b}^\times \to L^\times$ be the reduced norm of $A_{a,b}$. Then by [Merkurjev and Suslin 1982], $N_{a,b}$ represents $c \in L^\times$ if and only if $a \cup b \cup c \in H^3(L, \mathbb{Z}/n(3))$ is trivial.

   Therefore, since the conditions $b \in \mathrm{im}(N_a)$ and/or $c \in \mathrm{im}(N_{a,b})$ are first-order expressible, we conclude that $a \cup b$ and/or $a \cup b \cup c$ being (non)trivial are first-order expressible. Hence, the following hold:

   $(*)$ *The subsets $\Sigma_2 \subset L^\times \times L^\times$ and $\Sigma_3 \subset L^\times \times L^\times \times L^\times$ defined by*

$$\Sigma_2 := \{(a, b) \mid a \cup b \text{ is nontrivial}\},$$
$$\Sigma_3 := \{(a, b, c) \mid a \cup b \cup c \text{ is nontrivial}\}$$

   are first-order definable subsets.

(3) Let $K | k$ be a function field in one variable over a global field k as above. Let $\tilde{k} | k$ be a finite extension with $\mu_n \subset \tilde{k}$, and $\widetilde{K} := K\tilde{k}$. Then $\widetilde{K} | \tilde{k}$ is the function field of the complete smooth geometrically integral $\tilde{k}$-curve $\widetilde{X} := X \times_k \tilde{k}$. As in the case of $K | k$, we consider proper regular models $\widetilde{\mathcal{X}} \to \widetilde{S}$ of $\widetilde{X} \to \operatorname{Spec} \tilde{k}$, and the sets $\widetilde{\mathcal{X}}_i \subset \widetilde{\mathcal{X}}$ for $i = 0, 1, 2$. In particular, for every $\tilde{x}_1 \in \widetilde{\mathcal{X}}_1$, the local rings $\mathcal{O}_{\tilde{x}_1}$ are discrete valuation rings of $\widetilde{K}$, and we denote by $\widetilde{K}_{\tilde{x}_1}$ the

corresponding completions of $\widetilde{K}$. Then if $A_{a,b}$ is a division algebra as in (2) above, the following holds:

$N_{a,b}$ represents $c \in \widetilde{K}^{\times}$ over $\widetilde{K}$ if and only if $N_{a,b}$ represents c over $\widetilde{K}_{\tilde{x}_1}$ for all $\tilde{x}_1 \in \widetilde{\mathcal{X}}_1$.

3. Consequences of Kato's local-global principles

3A. *General facts.* Let K be a finitely generated field of Kronecker dimension $\dim(K) = 2$, and $k_0 = K^{\text{abs}}$ be its absolute subfield. If $\text{char}(K) = 0$, then $k := k_0$ is a number field, and $S = \operatorname{Spec} \mathcal{O}_k$ is the "canonical global curve" with function field k. Further, K is the function field of a projective regular S-surface $\mathcal{X} \to S$, having as a generic fiber a smooth projective geometrically integral k-curve X. If $\text{char}(K) = p > 0$, there exist (many) global function subfields $k \subset K$ of K with $k = \bar{k} \cap K$ such that letting S be the unique projective smooth k_0-curve, there exist projective smooth S-surfaces $\mathcal{X} \to S$ having as generic fiber a projective smooth k-curve X.

In the above notation, we denote by $S_i \subset S$, $X_i \subset X$, $\mathcal{X}_i \subset \mathcal{X}$ the points of dimension i in the corresponding schemes. In particular, one has the following:

- $S_0 \subset S$, $\mathcal{X}_0 \subset \mathcal{X}$, $X_0 \subset X$ are the closed points in the corresponding schemes.
- $S = S_0 \cup \{\eta\}$ and $X = X_0 \cup \{\eta_X\}$, where $\eta \in S$, $\eta_X \in X$ are the generic points.
- $\mathcal{X} = \mathcal{X}_0 \cup \mathcal{X}_1 \cup \{\eta_{\mathcal{X}}\}$, and $\eta_{\mathcal{X}} = \eta_X$, $X_0 \subset \mathcal{X}_1$ under the canonical inclusion $X \hookrightarrow \mathcal{X}_1$.

Notation/Remarks 3.1. Let n be a fixed prime number such that the group of roots of unity μ_{2n} of order $2n$ is contained in K. We notice/define the following:

(1) The local rings $\mathcal{O}_s$ at the closed points $s \in S_0$ of S are exactly the valuation rings of the nonarchimedean places of k. Further, for $x \in \mathcal{X}$ one has $x \mapsto s$ if and only if the corresponding local rings dominate each other: $\mathcal{O}_s \prec \mathcal{O}_x$ under $k \hookrightarrow K$.

(2) For $x_1 \in \mathcal{X}_1$, let $C_{x_1} = \overline{\{x_1\}} \subset \mathcal{X}$ be the schematic closure of x_1 in $\mathcal{X}$. Then C_{x_1} is an *arithmetic curve* on $\mathcal{X}$ with generic point $x_1 \in \mathcal{X}_1$. For $s \in S$, let $\mathcal{X}_s \to \kappa(s)$ be the fiber of $\mathcal{X} \to S$ at s. Then $\mathcal{X}_s$ is a projective (maybe nonreduced) one dimensional $\kappa(v)$-scheme of finite type. In the above notation, one has the following:

(a) $x_1 \mapsto \eta \in S$ if and only if $x_1 \in X_0$ if and only if $C_{x_1} \to S$ is finite dominant. If so, C_{x_1} is called a *horizontal curve* on $\mathcal{X}$, and we denote

$$\mathcal{X}_{1,\eta} := \{x_1 \in \mathcal{X}_1 \mid C_{x_1} \text{ is a horizontal curve}\}.$$

(b) $x_1 \mapsto s \in S_0$ if and only if C_{x_1} is a reduced irreducible component of $\mathcal{X}_s \to \kappa(s)$. If so, C_{x_1} is called a *vertical curve* on $\mathcal{X}$, and we denote

$$\mathcal{X}_{1,0} := \{x_1 \in \mathcal{X}_1 \mid C_{x_1} \text{ is a vertical curve}\}.$$

(3) One obviously has $\mathcal{X}_1 = \mathcal{X}_{1,\eta} \cup \mathcal{X}_{1,0}$, and the map $\mathcal{X}_{1,0} \to S_0$ has finite fibers.

(4) Since the generic fiber $X \to k$ of $\mathcal{X} \to S$ is a projective smooth geometrically integral k-curve, there exists a (unique) nonempty maximal open subset $U = U_{\mathcal{X}}$ of S such that $\mathcal{X}_U := \mathcal{X} \times_S U \to U$ is a family of projective smooth curves with geometrically integral fibers. For $s \in U_0 := U \cap S_0$, letting $x_s \in \mathcal{X}_s \subset \mathcal{X}$ be the generic point, one has:

(a) x_s is the unique preimage of s in $\mathcal{X}_1$, so $x_s \in \mathcal{X}_{1,0}$ and $C_{x_s} = \mathcal{X}_s$.

(b) $\kappa(s)$ is relatively algebraically closed in $\kappa(x_s)$.

(5) For $f \in K^\times$, let $|\mathrm{div}(f)| := \{P \in X_0 \mid v_P(f) \neq 0\} \subset X_0$ be the support of the divisor (f) of f viewed as a function on $X \to k$. Then $C_P = \overline{\{P\}}$ with $P \in |\mathrm{div}(f)|$ are distinct horizontal curves and $\bigcup_{P \in |\mathrm{div}(f)|} C_P$ is the closure of $|\mathrm{div}(f)|$ in $\mathcal{X}$. Therefore, there exists a unique maximal open subset $U_f = U_{\mathcal{X}f} \subset U$ satisfying:

(a) $\bigcup_{P \in |\mathrm{div}(f)|} C_P \to S$ is étale above U_f. Hence, $C_P \cap \mathcal{X}_s \cap C_{P'} = \varnothing$ for $P' \neq P$.

(b) For $s \in U_f$ and its unique preimage $x_s \in \mathcal{X}_{1,0}$ under $\mathcal{X}_{1,0} \to S$, the following hold:
- n is invertible in $\kappa(s)$.
- f is a v_{x_s}-unit, and its residue $\bar{f} \in \kappa(x_s)$ nonconstant: $\bar{f} \in \kappa(x_s) \setminus \kappa(s)$.

(6) Finally, we notice that $U_f \subseteq U_0$ has the following permanence property: Let $\tilde{k}\,|\,k$ be a finite extension, and set $\widetilde{K} := K\tilde{k}$. Let $\widetilde{S} \to S$ be the normalization of S in $k \hookrightarrow \tilde{k}$, and $\widetilde{\mathcal{X}} \to \widetilde{S}$ be a minimal proper regular model of $\widetilde{K}\,|\,\tilde{k}$ which dominates $\mathcal{X} \to S$. In particular, the generic fiber $\widetilde{X} \to \tilde{k}$ of $\widetilde{\mathcal{X}} \to \widetilde{S}$ is the normalization $\widetilde{X} \to X$ of X in $K \hookrightarrow \widetilde{K}$. Let $U_{\widetilde{\mathcal{X}}} \subset \widetilde{S}$ be the maximal open subset such that $\widetilde{\mathcal{X}}_{U_{\widetilde{\mathcal{X}}}} := \widetilde{\mathcal{X}} \times_{\widetilde{S}} U_{\widetilde{\mathcal{X}}} \to U_{\widetilde{\mathcal{X}}}$ is smooth and has reduced geometrically integral fibers, and define the subsets $U_{\widetilde{\mathcal{X}}f} \subseteq U_{\widetilde{\mathcal{X}}}$ for the model $\widetilde{\mathcal{X}} \to \widetilde{S}$ of $\widetilde{K}|\tilde{k}$ and $f \in \widetilde{K}$ in the way the subsets $U_{\mathcal{X}f} \subseteq U_{\mathcal{X}}$ of S were defined above for the model $\mathcal{X} \to S$ of $K\,|\,k$ and $f \in K$. Then one has:

Lemma 3.2. *In the above notation, let $\widetilde{U}_{\mathcal{X}f} \subseteq \widetilde{U}_{\mathcal{X}}$ be the preimages of $U_{\mathcal{X}f} \subseteq U_{\mathcal{X}}$ under the map $\widetilde{S} \to S$. Then $\widetilde{\mathcal{X}} \times_{\widetilde{S}} \widetilde{U}_{\mathcal{X}} \to \widetilde{U}_{\mathcal{X}}$ is smooth, whence $\widetilde{U}_{\mathcal{X}} \subseteq U_{\widetilde{\mathcal{X}}}$ and $\widetilde{\mathcal{X}} \times_{\widetilde{S}} \widetilde{U}_{\mathcal{X}} = \mathcal{X} \times_S \widetilde{U}_{\mathcal{X}}$. Further, $\widetilde{U}_{\mathcal{X}f} \subseteq U_{\widetilde{\mathcal{X}}f}$, and the morphism $\widetilde{\mathcal{X}} \to \mathcal{X}$ is finite above $\mathcal{X} \times_S U_{\mathcal{X}}$.*

Proof. For the first inclusion, let $\mathcal{X}^{\mathrm{n}} \to \widetilde{S}$ denote the normalization of $\mathcal{X} \to S$ in the field extension $K \hookrightarrow \widetilde{K}$. Then $\widetilde{\mathcal{X}}$ being regular, it is also normal. Thus $\widetilde{\mathcal{X}} \to \widetilde{S}$ dominates $\mathcal{X}^{\mathrm{n}} \to \widetilde{S}$. Moreover, since the base change $\mathcal{X} \times_S \widetilde{S} \to \widetilde{S}$ is dominant and finite over $\mathcal{X} \to S$, it follows that $\mathcal{X}^{\mathrm{n}} \to \widetilde{S}$ dominates $\mathcal{X} \times_S \widetilde{S} \to \widetilde{S}$. Thus finally $\widetilde{\mathcal{X}} \to \widetilde{S}$ dominates $\mathcal{X} \times_S \widetilde{S} \to \widetilde{S}$. To simplify notation, set $U := U_{\mathcal{X}}$ and $\widetilde{U} := \widetilde{U}_{\mathcal{X}}$. Since $\mathcal{X}_U := \mathcal{X} \times_S U \to U$ is smooth and has reduced geometrically integral fibers, so is the base change $\mathcal{X}_U \times_S \widetilde{U} \to \widetilde{U}$, and in particular $\mathcal{X}_U \times_S \widetilde{U}$ is regular. Hence, by the minimality of $\widetilde{\mathcal{X}} \to \widetilde{S}$, one has $\widetilde{\mathcal{X}} \times_{\widetilde{S}} \widetilde{U} = \mathcal{X}^{\mathrm{n}} \times_{\widetilde{S}} \widetilde{U} = \mathcal{X}_U \times_S \widetilde{U}$. Therefore, $\widetilde{\mathcal{X}}_{\widetilde{U}} := \widetilde{\mathcal{X}} \times_{\widetilde{S}} \widetilde{U} \to \mathcal{X}_U$ is finite because $\mathcal{X}^{\mathrm{n}} \to \mathcal{X}$ is. Since $\widetilde{\mathcal{X}}_{\widetilde{U}} = \mathcal{X} \times_S \widetilde{U} \to \widetilde{U}$ is also smooth, one has $\widetilde{U}_{\mathcal{X}} \subseteq U_{\widetilde{\mathcal{X}}}$ by the maximality of the latter. The other inclusion follows immediately from the fact that being étale is preserved under base change, and the fact that $\widetilde{\mathcal{X}} \times_{\widetilde{S}} \widetilde{U}_{\mathcal{X}} = \mathcal{X} \times_S \widetilde{U}_{\mathcal{X}}$. □

3B. *A local-global principle for* $\mathrm{H}^3(\mathcal{X}, f)$. We work in the context and the notation of the previous subsection. Let $\mathcal{X}_{1f} \subset \mathcal{X}_1$ be the preimage of U_f under $\mathcal{X}_1 \to S$. We notice that $\mathcal{X}_1 \setminus \mathcal{X}_{1f}$ is the finite closed subset of $\mathcal{X}_{1,0}$ consisting of all $x_1 \in \mathcal{X}_1$ which map into the (finite) closed set $S_0 \setminus U_f$.

Notation. Let $\mathrm{H}^3(\mathcal{X}, f) \subset \mathrm{H}^3(K, \mathbb{Z}/n(2))$ denote the set of all the symbols $f \cup a \cup b$ with $a, b \in k^\times$ which are nontrivial over some completion K_{x_1} with $x_1 \in \mathcal{X}_{1f}$.

Lemma 3.3. *Let $D_f \subseteq |\mathrm{div}(f)|$ be the set of all P such that $v_P(f)$ is not divisible by n in $v_P(K)$. Suppose that K has no orderings. Then for every $f \cup a \cup b \in \mathrm{H}^3(\mathcal{X}, f)$ there exists $P \in D_f$ such that $f \cup a \cup b$ is nontrivial over K_P.*

Proof. Let $z_1 \in \mathcal{X}_1$ be a given point. Then by the concrete description of the boundary homomorphism $\hat{\partial}_{z_1}$ as given before Theorem 2.1, one has that if $v_{z_1}(f)$, $v_{z_1}(a)$, $v_{z_1}(b)$ are all divisible by n, then $f \cup a \cup b$ is trivial over the completion K_{z_1} at v_{z_1}. In particular, if $z_1 \in \mathcal{X}_{1,\eta}$, then $a, b \in k^\times$ are v_{z_1}-units. Thus $v_{z_1}(a) = 0 = v_{z_1}(b)$ are divisible by n. Hence, if $v_{z_1}(f)$ is divisible by n, then $f \cup a \cup b$ is trivial over K_{z_1}.

Returning to the proof of the lemma, let $f \cup a \cup b \in \mathrm{H}^3(\mathcal{X}, f)$ be a given element, and let $x_1 \in \mathcal{X}_{1f} \subset \mathcal{X}_1$ be such that $f \cup a \cup b$ is nontrivial over the completion K_{x_1}.

Case 1. $x_1 \in \mathcal{X}_{1,\eta} = X_0$. Then v_{x_1} is trivial on k, so $v_{x_1}(a) = 0 = v_{x_1}(b)$. Hence, since $f \cup a \cup b \in \mathrm{H}^3(\mathcal{X}, f)$ is nontrivial over the completion K_{x_1}, it follows by the discussion above that $v_{x_1}(f)$ is not divisible by n. Thus $P := x_1 \in D_f$, and we are done.

Case 2. $x_1 \in \mathcal{X}_{1,0}$. Let $s \in U_f$ be the image of x_1 under $\mathcal{X}_{1f} \to U_f \subset S$. Then by the definition of $\mathcal{X}_{1f}$, one has that $x_s := x_1$ is the unique preimage of s in $\mathcal{X}_1$, and the following hold:

- $\mathcal{X}_s$ is a projective smooth geometrically integral $\kappa(s)$-curve, and $C_{x_s} = \mathcal{X}_s$.
- For all $P \neq P'$ in $|\mathrm{div}(f)|$, if $x_0 \in \mathcal{X}_s \cap C_P$, then $x_0 \notin \mathcal{X}_s \cap C_{P'}$.
- n is invertible in $\kappa(s)$.
- f is a v_{x_s}-unit, and its residue $\bar{f} \in \kappa(x_s)$ is nonconstant, i.e., $\bar{f} \in \kappa(x_s) \setminus \kappa(s)$.

From this we reason as follows. Let $x_0 \in \mathcal{X}_0$ be a closed point with $x_0 \mapsto s \in U_f$, and let $z_1 \in \mathcal{X}_1$ satisfy that $x_0 \in C_{z_1}$ and not all $v_{z_1}(a)$, $v_{z_1}(b)$, $v_{z_1}(f)$ are zero. Then we have:

(a) If $z_1 \in \mathcal{X}_{1,0}$, i.e., z_1 maps to some $s' \in S_0$, then C_{z_1} is a vertical curve; and since $C_{z_1} \ni x_0 \mapsto s$, we must have $C_{z_1} = \mathcal{X}_s$, so that $z_1 = x_1$, etc.

(b) If $P := z_1 \in \mathcal{X}_{1,\eta} = X_0$, then a, b are v_P-units, and therefore $v_P(f) \neq 0$. Thus $P \in |\mathrm{div}(f)|$, and $x_0 \in \mathcal{X}_s \cap C_P$. Moreover, P is the unique point in $|\mathrm{div}(f)|$ with the property $x_0 \in C_P \cap \mathcal{X}_s$.

From this we can conclude the following. Let $x_0 \in \mathcal{X}$ be a closed point above the point $s \in U_f$. Then there exist at most two points $z_1 \in \mathcal{X}_1$, each satisfying that $x_0 \in C_{z_1}$ and at least one of the values $v_{z_1}(a)$, $v_{z_1}(b)$, $v_{z_1}(f)$ is nonzero. Further, the two (potential) points are

- the given point $x_s = x_1 \in \mathcal{X}_{1,0}$ — note that $v_{x_s}(f) = 0$, $\bar{f} \in \kappa(x_s)$ is nonconstant, etc.;
- the unique $P_0 \in |\mathrm{div}(f)|$ such that $x_0 \in C_{P_0} \cap \mathcal{X}_s$ — note that a, b are v_{P_0}-units.

Therefore, the image of $f \cup a \cup b \in \mathrm{H}^3(K, \mathbb{Z}/n(2))$ in $\bigoplus_{x_1 \in \mathcal{X}_{x_0}} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1))$ under the homomorphism of Theorem 2.1 actually lies in

$$\mathrm{H}^2(\kappa(P_0), \mathbb{Z}/n(1)) \oplus \mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1)).$$

Let us compute $\partial_{x_s}(f \cup a \cup b)$. First, since $s \in U_f$, we have $f \in \mathcal{O}_{x_s}^\times$ and $\bar{f} \in \kappa(x_s)$ is nonconstant. Second, every uniformizing parameter $\pi \in k$ at s is also a uniformizing parameter π at x_s, because $\mathcal{X}_s$ is reduced by the fact that $s \in U_f$. For such a π, set $a = \pi^q a'$ and $b = \pi^r b'$ with $a', b' \in \mathcal{O}_s^\times$. Then setting $c = a'^r / b'^q \in \mathcal{O}_s^\times$, it follows by the definition of ∂_{x_s} that we have $0 \neq \partial_{x_s}(f \cup a \cup b) = \bar{f} \cup \bar{c} \in \mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1))$.

Next, recall that since $s \in U_f$, the special fiber $\mathcal{X}_s \to \kappa(s)$ is a complete smooth geometrically integral model of the global function field $\kappa(x_s) | \kappa(s)$. Since $\bar{f} \cup \bar{c}$ is nontrivial in $\mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1))$, by the Hasse–Brauer–Noether LGP, there exists a closed point $x_0 \in \mathcal{X}_{s,0} \subset \mathcal{X}_0$ such that $\bar{f} \cup \bar{c}$ is nontrivial over the completion $\kappa(x_s)_{x_0}$. Equivalently, the boundary homomorphism

$$\partial_{x_0} : \mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$$

maps $\bar{f} \cup \bar{\gamma}$ to some nontrivial element in $\mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$.

Let $\mathcal{O}_{x_0}$ be the local ring of $x_0 \in \mathcal{X}$ viewed as a closed point of $\mathcal{X}$, and let $\overline{\mathcal{O}}_{x_0} \subset \kappa(x_s)$ be the image of $\mathcal{O}_{x_0}$ under the canonical projection $\mathcal{O}_{x_s} \to \kappa(x_s)$. Then by scheme-theoretical nonsense, it follows that $\overline{\mathcal{O}}_{x_0}$ is the local ring of the point $x_0 \in \mathcal{X}_{s,0}$ viewed as a closed point of $\mathcal{X}_s$. Hence, since the latter is a smooth curve over $\kappa(s)$, and thus regular, it follows that $\overline{\mathcal{O}}_{x_0} = \mathcal{O}_{\mathcal{X}_s, x_0}$ is regular. Therefore, by the definition of ∂_{x_s, x_0} as described before Theorem 2.1, it follows that $\partial_{x_s x_0}(\bar{f} \cup \bar{c}) = \partial_{x_0}(\bar{f} \cup \bar{c})$. Hence we conclude that $\partial_{x_s x_0}(\bar{f} \cup \bar{c}) \in \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$ is nontrivial. Viewing x_0 as a closed point of $\mathcal{X}$, we conclude that the image of $f \cup a \cup b$ under $\mathrm{H}^3(K, \mathbb{Z}/n(2)) \to \mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$ is nontrivial.

On the other hand, the image of $f \cup a \cup b$ in $\bigoplus_{x_1 \in \mathcal{X}_{x_0}} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1))$ lies in $\mathrm{H}^2(\kappa(P_0), \mathbb{Z}/n(1)) \oplus \mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1))$, by the discussion above.

For a contradiction, suppose that the image of $f \cup a \cup b$ in $\mathrm{H}^2(\kappa(P_0), \mathbb{Z}/n(1))$ is trivial. Then the image of $f \cup a \cup b$ in $\bigoplus_{x_1 \in \mathcal{X}_{x_0}} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1))$ lies in $\mathrm{H}^2(\kappa(x_s), \mathbb{Z}/n(1))$, and this image is $\bar{f} \cup \bar{c}$. Thus, the image of $f \cup a \cup b$ under the canonical map

$$\mathrm{H}^3(K, \mathbb{Z}/n(2)) \to \bigoplus_{x_1 \in \mathcal{X}_{x_0}} \mathrm{H}^2(\kappa(x_1), \mathbb{Z}/n(1)) \to \mathrm{H}^1(\kappa(x_0), \mathbb{Z}/n)$$

is nontrivial. This contradicts Theorem 2.1.

Therefore, the image of $f \cup a \cup b$ in $\mathrm{H}^2(\kappa(P_0), \mathbb{Z}/n(1))$ must be nontrivial. Since that image is $v_{P_0}(f) \cdot a \cup b$, we conclude, first, that $v_{P_0}(f)$ is not divisible by n, so that $P_0 \in D_f$, and second, that $f \cup a \cup b$ is nontrivial over K_{P_0}. □

3C. ***The Chebotarev density theorem and the size of*** $\mathbf{H}^3(\mathcal{X}, f)$**.** Let $\lambda | k$ be a finite extension with $\mu_{2n} \subset \lambda$. Consider $\alpha \in k$ which is not an n-th power in λ, or equivalently, $\tilde{\lambda} := \lambda[\sqrt[n]{\alpha}\,]$ is a cyclic extension of degree n of λ. Let further $\hat{\lambda} | k$ be some finite Galois extension of k containing $\tilde{\lambda}$, and let $\hat{T} \to \widetilde{T} \to T \to S$ be the normalizations of S in the field extensions $k \hookrightarrow \lambda \hookrightarrow \tilde{\lambda} \hookrightarrow \hat{\lambda}$. For a generator $\sigma \in \mathrm{Gal}(\tilde{\lambda} | \lambda)$, consider a preimage $\tau \in \mathrm{Gal}(\hat{\lambda} | \lambda) \subseteq \mathrm{Gal}(\hat{\lambda} | \lambda)$. Let $\hat{T}_\alpha \to S_\alpha$ be the sets of all the points $\hat{z} \mapsto s$ such that α is a v_s-unit and τ is the Frobenius $\kappa(\hat{z}) | \kappa(s)$. Notice that by the Chebotarev density theorem, S_α has a positive Dirichlet density, and that for $\hat{z} \in \hat{T}_\alpha$ and its image $s \in S_\alpha$ one has $\kappa(\hat{z}) = \kappa(s)[\hat{\gamma}]$ with $\hat{\gamma}^m = \bar{\alpha}$ and m the order of τ.

Finally, let $\tilde{z}_\alpha \to z_\alpha$ be the images of $\hat{z}_\alpha$ in $\widetilde{T} \to T$. Then for $\widetilde{S}_\alpha \ni \tilde{z} \mapsto z \in T_\alpha$, one has that $\tilde{z} | z$ is unramified and has σ as Frobenius automorphism: $\kappa(z) = \kappa(s)$ and $\kappa(\tilde{z}) = \kappa(s)[\gamma]$, where $\gamma^n = \bar{\alpha}$. Thus we have showed the following:

Fact 3.4. *Let $\lambda | k$ be a finite extension of global fields, $\mu_{2n} \subset \lambda$, and $\alpha \in k$ not an n-th power in λ. Let $T \to S$ be the normalization of S in $k \hookrightarrow \lambda$. There exist subsets $S_\alpha \subset S$ of positive Dirichlet density and $T_\alpha \subset T$ mapping onto S_α such that for all $T_\alpha \ni z \mapsto s \in S_\alpha$ one has $k_s = \lambda_z$, and α and n are v_s-units, and α is not an n-th power in $k_s = \lambda_z$.*

Notation/Remarks 3.5. For $\alpha, \delta \in k^\times$ and $V_\delta := \{s \in S \mid v_s(\delta) = 0\} \subset S$ open and nonempty, and the subgroup $\mathrm{H}_\delta = \{\beta \in k^\times \mid v_s(\beta - 1) > 2v_s(2n) \text{ if } s \notin V_\delta\} \subset k^\times$, consider/define

(1) $\mathrm{H}_{\delta\alpha} := \alpha \cup \mathrm{H}_\delta \subset \mathrm{H}^2(k, \mathbb{Z}/n(1))$,

(2) $\mathrm{H}_{f\delta\alpha} := f \cup \alpha \cup \mathrm{H}_\delta = f \cup \mathrm{H}_{\delta\alpha} \subset \mathrm{H}^3(K, \mathbb{Z}/n(2))$.

Notice that by Hensel's lemma we have that if $\beta \in \mathrm{H}_\delta$ then β is an n-th power in k_s for all $s \notin V_\delta$.

Lemma 3.6. *Suppose that $V_\delta \subseteq U_f$ and that $\mathrm{H}_{f\delta\alpha} \neq 0$. Then every nonzero $f \cup \alpha \cup \beta \in \mathrm{H}_{f\delta\alpha}$ lies in $\mathrm{H}^3(\mathcal{X}, f)$, and for every such $f \cup \alpha \cup \beta$ the following hold*:

(1) *There exists $P \in X$ such that $f \cup \alpha \cup \beta$ is nontrivial over K_P. Hence, $P \in D_f$ and α is not an n-th power in K_P nor in the residue field $\kappa(P)$.*

(2) $\mathrm{H}_{f\delta^*\alpha}$ *is nontrivial over K_P for all P as in* (1) *above and all $\delta^* \in k^\times$.*

Proof. We first prove that every nonzero $f \cup \alpha \cup \beta \in \mathrm{H}_{f\delta\alpha}$ actually lies in $\mathrm{H}^3(\mathcal{X}, f)$. Indeed, by Theorem 2.1, there exists some $x_1 \in \mathcal{X}_1$ such that $f \cup \alpha \cup \beta$ is nontrivial over the completion K_{x_1}. Let $x_1 \mapsto s \in S$ be the image of x_1 in S. We claim that $s \in V_\delta$. By contradiction suppose that $s \notin V_\delta$. Then by Notation/Remarks 3.5, β is an n-th power in k_s and in particular, $\alpha \cup \beta$ is trivial over k_s. Further, since $x_1 \mapsto s$, we have $k_s \subseteq K_{x_1}$. Hence $f \cup \alpha \cup \beta$ is trivial over $\widetilde{K}_{\tilde{x}_1}$, a contradiction! Thus finally $s \in V_\delta$, and since $V_\delta \subseteq U_f$ we have $s \in U_f$.

By (1), since $f \cup \alpha \cup \beta \in \mathrm{H}^3(\mathcal{X}, f)$, by Lemma 3.3 it follows that there exists $P \in D_f$ such that $f \cup \alpha \cup \beta$ is nontrivial over K_P. In particular, α is not an n-th power in K_P, etc.

For (2), by the discussion above, α is not an n-th power in $\lambda := \kappa(P)$. In the notation from Fact 3.4, for some fixed $s^* \in V_{\delta^*} \cap S_\alpha$, let β^* be a uniformizing parameter at s^* such that β^* is an n-th power in k_s for all $s \notin V_{\delta^*}$. Then $\alpha \cup \beta^*$ satisfies first, $\beta^* \in \mathrm{H}_{\delta^*}$, so $\alpha \cup \beta^* \in \mathrm{H}_{\delta^*\alpha}$ and $f \cup \alpha \cup \beta^* \in \mathrm{H}_{f\delta^*\alpha}$. Second, $\alpha \cup \beta^*$ is trivial over all k_s with $s \notin V_{\delta^*}$, because β^* is an n-th power in k_s. On the other hand, $\alpha \cup \beta^*$ is not trivial over λ_z for $z \mapsto s^*$ because β^* is a uniformizing parameter at s^* and at all $z \mapsto s^*$. Hence $\alpha \cup \beta^*$ is not trivial over $\kappa(P) \subset \kappa(P)_z$. But then, since $\hat{\partial}_P : \mathrm{H}^3(K_P, \mathbb{Z}/n(2)) \to \mathrm{H}^2(\kappa(P), \mathbb{Z}/n(1))$ is an isomorphism and $\hat{\partial}_P(f \cup \alpha \cup \beta^*) = v_P(f) \cdot \alpha \cup \beta^* \neq 0$, we get that $f \cup \alpha \cup \beta^*$ is nontrivial over K_P. □

4. Detecting the k-valuations of $K \,|\, k$

In this section we work in the context/notation of the previous sections: $n \neq \mathrm{char}(K)$ is a prime number and $\mu_{2n} \subset K$. So if $n = 2$, then $\mu_4 \subset K$, and if $n \neq 2$, then $\mu_n \subset K$.

4A. *The sets $\mathcal{U}_\bullet$.*

Notation/Remarks 4.1. In the usual context we have the following:

(1) For $u \in K$ and $\alpha, c \in k$, set $u_{\alpha,c} = 1 - c(1-u) + \alpha c^n (1-u)^n$, and further define $u_c := u_{0,c} = 1 + c(u-1)$ and $u_\alpha := u_{\alpha,1} = u + \alpha(1-u)^n$.

(2) For $u \in K^\times$ and $c \in k^\times$ we set $K_{u,\alpha,c} := K[\sqrt[n]{u_c}, \sqrt[n]{u_{\alpha,c}}\,]$, and notice that $K_{u,\alpha,c} \,|\, K$ is a $\mathbb{Z}/n$-elementary abelian extension of degree 1, n, or n^2.

(3) In Notation/Remarks 3.5, suppose that $\mathrm{H}_{f\delta\alpha} \neq 0$. Thus $\mathrm{H}_{f\delta^*\alpha} \neq 0$ for all $\delta^* \in k^\times$. We set $\mathcal{U}_{f\alpha} := \{u \in K^\times \mid \mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_{u,\alpha,c}$ for all $c, \delta^* \in k^\times\}$.

(4) For u, α, c as above, set $D_{u,\alpha,c} := \{P \in D_f \mid u_c, u_{\alpha,c} \in \mathcal{O}_P^\times\}$.

(5) Finally, let $D_{f\alpha} := \{P \in D_f \mid \alpha$ is not an n-th power in $\kappa(P)\}$. Note that by Lemma 3.6, if $\mathrm{H}_{f\delta\alpha}$ is nontrivial then $D_{f\alpha}$ is nonempty.

Lemma 4.2. *Let $Y \to X$, $Q \mapsto P$, be the normalization of X in $K \hookrightarrow L := K_{u,\alpha,c}$ and suppose that $\sqrt[n]{\alpha} \notin L_Q$. Then $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$ and hence $P \in D_{u,\alpha,c}$.*

Proof. Let us analyze what happens if either u_c or $u_{\alpha,c}$ is not a v_P-unit. We first claim that u is v_P-integral. Indeed, by contradiction, suppose that $v_P(u) < 0$. Then $v_P(1/u) > 0$, and $u_{\alpha,c} = \eta\alpha(-cu)^n$, where η is a principal v_P-unit. But then η is a principal v_Q-unit too, and hence η is an n-th power in L_Q. Conclude that $\sqrt[n]{\alpha} \in L_Q$: contradiction! Thus finally u must be v_P-integral. Further, if u is a principal v_P-unit, then so are $u_c, u_{\alpha,c}$, and thus $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$. Hence, it is left to analyze what happens if $v_P(u) \geq 0$ and u is not a principal v_P-unit. First we remark that $1-u$ is a v_P-unit, and hence so is $\alpha c^n(1-u)^n$. Second, both u_c and $u_{\alpha,c}$ are v_P-integral. Therefore, since $u_{\alpha,c} = u_c + \alpha c^n(1-u)^n$, it follows that at least one of the elements u_c and $u_{\alpha,c}$ is a v_P-unit. By contradiction, suppose that either u_c or $u_{\alpha,c}$ is not a v_P-unit. Then either $v_P(u_c) = 0$ and $v_P(u_{\alpha,c}) > 0$, or vice versa.

Case 1. $v_P(u_c) = 0$ and $v_P(u_{\alpha,c}) > 0$.

Then $\alpha = -u_c(1 - u_{\alpha,c}/u_c)/c^n(1-u)^n$. Since $\mu_{2n} \subset K$, it follows that -1 is an n-th power in K, and since $1 - u_{\alpha,c}/u_c$ is a principal v_P-unit, it is an n-th power in L_Q. Hence all the factors on the right-hand side are n-th powers in L_Q. Thus $\sqrt[n]{\alpha} \in L_Q$: contradiction!

Case 2. $v_P(u_{\alpha,c}) = 0$ and $v_P(u_c) > 0$.

Then $\alpha = u_{\alpha,c}(1 - u_c/u_{\alpha,c})/c^n(1-u)^n$ with $1 - u_c/u_{\alpha,c}$ a principal v_P-unit. But then all the factors on the right-hand side are n-th powers in $K_P \subset L_Q$. Hence $\sqrt[n]{\alpha} \in K_P \subseteq L_Q$: contradiction!

We thus conclude that $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$, as claimed. □

Lemma 4.3. *Suppose that $V_\delta \subseteq U_f$ and $\mathrm{H}_{f\delta\alpha}$ is nontrivial. Then the following hold*:

(1) *If $u \in \mathcal{U}_{f\alpha}$ then $u_c \in \mathcal{U}_{f\alpha}$ for all $c \in k$. And if $c \neq 0$ and $u_c \in \mathcal{U}_{f\alpha}$, then $u \in \mathcal{U}_{f\alpha}$.*

(2) $1 + \bigcup_{P \in D_{f\alpha}} \mathfrak{m}_P \subseteq \mathcal{U}_{f\alpha}$.

(3) *For every $u \in \mathcal{U}_{f\alpha}$ and each resulting $u_c, u_{\alpha,c}$ the following hold*:
  - (a) *There exists $P \in D_{f\alpha}$ with $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$ and $\mathrm{H}_{f\delta^*\alpha}$ nontrivial over $K_P K_{u,\alpha,c}$ for all δ^*.*
  - (b) *There exists δ^* such that if $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_P K_{u,\alpha,c}$, then $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$ and $P \in D_{f\alpha}$.*

Proof. (1): For all $a, c, c' \in k$, $(u_c)_{a,c'} = 1 - cc'(1-u) + a(cc')^n(1-u)^n = u_{a,cc'}$, and therefore $(u_c)_{c'} = u_{cc'}$ and $(u_c)_{\alpha,c'} = u_{\alpha,cc'}$. Hence $\{(u_c)_{c'}, (u_c)_{\alpha,c'}\} = \{u_{cc'}, u_{\alpha,cc'}\}$. Now suppose that $u \in \mathcal{U}_{f\alpha}$. Then by the definition of $\mathcal{U}_{f\alpha}$ it follows that $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_{u,\alpha,c''}$ for all $c'' \in k$ and all $\delta^* \in k^\times$. In particular, setting $c'' := cc'$, it follows that $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_{u_c,\alpha,c'}$, etc. The converse is clear, because given c' and u_c, by the discussion above one has that $\{u_{c'}, u_{\alpha,c'}\} = \{(u_c)_{c'/c}, (u_c)_{\alpha,c'/c}\}$, etc.

For the proof of assertions (2) and (3) we first set up notation as follows: For $u \in K^\times$ and $c \in k$, set as usual $L := K_{u,\alpha,c}$, and further, $l := L \cap \bar{k}$. Let $T \to S$ be the normalization of S in $k \hookrightarrow l$, and $\mathcal{Y} \to T$ be the minimal proper regular model of $L\,|\,l$ which dominates $\mathcal{X} \to S$. In particular, the generic fiber $Y \to l$ of $\mathcal{Y} \to T$ is the normalization $Y \to X$ of X in the field extension $K \hookrightarrow L$.

(2): Let $u \in 1 + \mathfrak{m}_P$ be a principal unit at some $P \in D_{f\alpha}$. Since $P \in D_{f\alpha}$, we have by definition that α is not an n-th power in $\kappa(P)$ nor in K_P, and $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over K_P for all $\delta^* \in k^\times$ by Lemma 3.6. On the other hand, since u is a principal v_P-unit, $u_c, u_{\alpha,c}$ are principal v_P-units too (by mere definitions). Therefore, P is totally split in the field extension $L\,|\,K$, and thus for every $Q \mapsto P$ one has $L_Q = K_P$. Hence, $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $L_Q = L_P$ (because it was nontrivial over K_P). But then $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $L \subset L_Q$ too.

(3): For the proper regular model $\mathcal{Y} \to T$ of $L\,|\,l$ and $f \in L$, we define the open nonempty subsets $U_{\mathcal{Y}f} \subseteq U_{\mathcal{Y}}$ of T, as we defined the sets $U_f \subseteq U_{\mathcal{X}}$ of S for the proper regular model $\mathcal{X} \to S$ of $K\,|\,k$ and $f \in K$ at Notation/Remarks 3.1(5). For both assertions (a) and (b), we consider δ^* which satisfy $V_{\delta^*} \subseteq U_f$, and the preimage of V_{δ^*} under $T \to S$ is contained in $U_{\mathcal{Y}f}$. For such a $\delta^* \in k^\times$ let $f \cup \alpha \cup \beta^* \in \mathrm{H}_{f\delta^*\alpha}$ be nontrivial over L.

Claim. The image of $f \cup \alpha \cup \beta^*$ in $\mathrm{H}^3(L, \mathbb{Z}/n(2))$ lies in $\mathrm{H}^3(\mathcal{Y}, f)$.

Indeed, since $f \cup \alpha \cup \beta^*$ is nontrivial over L, by Theorem 2.1, there exists some $y_1 \in \mathcal{Y}_1$ such that $f \cup \alpha \cup \beta^*$ is nontrivial over L_{y_1}. Let $y_1 \mapsto z \mapsto s^*$ be the images of y_1 in $T \to S$. We claim that $s \in V_{\delta^*}$, and thus $z \in U_{\mathcal{Y}f}$ by the definition of δ^*. Indeed, by contradiction, suppose that $s^* \notin V_{\delta^*}$. Then reasoning as in the proof of Lemma 3.6, taking into account that β^* is an n-th power in k_s for $s \notin V_{\delta^*}$, we conclude that $\alpha \cup \beta^*$ is trivial over k_{s^*} because β^* is in k_{s^*}. Hence $f \cup \alpha \cup \beta^*$ is trivial over L_{y_1}, because $k_{s^*} \subset L_{y_1}$. Contradiction! The claim is proved.

(a): For $u \in \mathcal{U}_{f\alpha}$ and $f \cup \alpha \cup \beta^* \in \mathrm{H}_{f\delta^*\alpha}$, which is nontrivial over L, by Lemma 3.3 applied to $f \cup \alpha \cup \beta^* \in \mathrm{H}^3(\mathcal{Y}, f)$, one found that there exists some $Q \in Y$ such that $v_Q(f)$ is not divisible by n in $v_Q(L)$, and α is not an n-th power in L_Q, nor in $\kappa(Q)$. Further, $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $L_Q = K_P K_{u,\alpha,c}$ for all $\delta^* \in k^\times$. Let $Q \mapsto P \in X$ be the image of Q in X, and consider the canonical embeddings $K_P \hookrightarrow L_Q$, $\kappa(P) \hookrightarrow \kappa(Q)$, and recall that $v_Q = e(Q\,|\,P)\,v_P$, where $e(Q\,|\,P)$ is the ramification index of $v_Q\,|\,v_P$. Hence the following hold:

- Since $v_Q(f) \notin n \cdot v_Q(L)$, one has that $v_P(f) \notin n \cdot v_P(K)$. Therefore, $P \in D_f$.
- Since $\sqrt[n]{\alpha} \notin \kappa(Q)$, one has that $\sqrt[n]{\alpha} \notin \kappa(P)$. Therefore, $P \in D_{f\alpha}$.

Hence $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_P K_{u,\alpha,c}$, and $P \in D_{f\alpha}$ and $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$ by Lemma 4.2.

(b): Clear from the discussion above. $\square$

4B. ***The k-rings $\mathfrak{R}_\bullet$ and $R_\bullet$.*** In the above notation and context, we introduce the ring stabilizer $\mathfrak{R}_{f\alpha}$ of $\mathcal{U}_{f\alpha}$, which will play an essential role in describing k-valuations of $K\,|\,k$.

Notation/Remarks 4.4. (1) Let $A, +, \cdot$ be a commutative ring with 1_A, and if $k \subset A$ is a subfield in A having identity equal to 1_A, we consider A as a k-algebra. It seems that the following are well known facts to (a nonempty set of) experts:

> *Let $X \subset A$ satisfy $X = -X$ and $0_A \in X$, and set $X_0 := \{x \in A \mid x + X \subseteq X\}$. Then $X_0 \subseteq X$, and $R_X := \{a \in X_0 \mid a \cdot X_0 \subseteq X_0\}$ is a subring of R, which contains 1_A if and only if $1_A \in X_0$. Moreover, if A is a k-algebra, and X is stable under multiplication with k, then R_X is a k vector subspace.*

(2) Given a commutative ring $A, +, \cdot$ with 1_A as above, let $*$ and $\circ$ be the transport of the usual addition and multiplication, respectively, on the underlying set A via $a \mapsto a + 1_A$. Hence $a * b = a + b - 1_R$ and $a \circ b = ab - a - b + 2$, and A endowed with $*, \circ$ is an isomorphic copy of $A, +, \cdot$ which we denote $\mathfrak{A}$.

If $X \subset A$ is a nonempty subset as in (1) above, we let $\mathfrak{R}_X \subset \mathfrak{A}$, or simply $\mathfrak{R}$ if no confusion is possible, be the corresponding subring of $\mathfrak{A}$.

(3) In the context and notation of the previous subsection, recall the nonempty set $X := \mathcal{U}_{f\alpha}$ of K. We notice that in Lemma 4.3(1) one has $u_c \in \mathcal{U}_{f\alpha}$ if (and only if) $u \in \mathcal{U}_{f\alpha}$ (provided $c \neq 0$). On the other hand, $c \circ u = u \circ c = (c-1)(u-1)+1 = u_{c-1}$. Hence for $u \in K$, $c \in k$, one has $c \circ u \in \mathcal{U}_{f\alpha}$ if (and only if) $u \in \mathcal{U}_{f\alpha}$ (provided $c \neq 1$).

In particular, $X := \mathcal{U}_{f\alpha}$ is closed with respect to multiplication $\circ$ by elements of k, and therefore, X is symmetric with respect to the addition $*$.

(4) For $X := \mathcal{U}_{f\alpha}$, we denote by $\mathfrak{R}_{f\alpha} := \mathfrak{R}_{\mathcal{U}_{f\alpha}}$ the corresponding subring of $K, *, \circ$ (the latter being an isomorphic copy of the field $K, +, \cdot$ as mentioned above).

Hence $R_{f\alpha} := \mathfrak{R}_{f\alpha} - 1$ is a subring of the field $K, +, \cdot$ with the usual addition and multiplication.

Lemma 4.5. *The ring $\mathfrak{R}_{f\alpha}, *$ is a $k, *, \circ$ vector space. Thus $R_{f\alpha}$ is a k-subspace of $K, +$.*

Proof. Clear by the discussion at (1) and (3) above. $\square$

Lemma 4.6. *In the above notation, let $X := \mathcal{U}_{f\alpha}$. Then one has $X_0 \subseteq \bigcap_{P \in D_{f\alpha}} \mathcal{O}_P^\times$.*

Proof. Indeed, by contradiction, suppose that there exists $u_0 \in X_0$ such that $v_{P_0}(u_0) \neq 0$ for some $P_0 \in D_{f\alpha}$. Then using the weak approximation lemma, we can choose $t \in K$ such that $v_{P_0}(t-1) > 0$, i.e., t is a principal v_{P_0}-unit, and $v_{P'}(u_0+t-1)<0$ for all $P'\neq P_0$, $P'\in D_{f\alpha}$. Then $v_{P''}(u_0+t-1)\neq 0$ for all $P''\in D_{f\alpha}$. On the other hand, since u is a v_{P_0} principal unit, it follows by Lemma 4.3(2) that $t\in\mathcal{U}_{f\alpha}=X$. Hence, since $u_0\in X_0$, we must have $t*u_0\in\mathcal{U}_{f\alpha}$ (by the definition of X_0), i.e, $u:=t*u_0=u_0+t-1\in\mathcal{U}_{f\alpha}$. But then by Lemma 4.3(3a), it follows that for every c, there must exist some $P\in D_{f\alpha}$ such that $u_c, u_{\alpha,c}\in\mathcal{O}_P^\times$. Hence, for $c=1$, one gets $u_0+t-1=u=u_c\in\mathcal{O}_P^\times$ for some $P\in D_{f\alpha}$, contradicting the fact that $v_{P''}(u_0+t-1)\neq 0$ for all $P''\in D_{f\alpha}$. □

Key Lemma 4.7. *One has* $\mathfrak{R}_{f\alpha}=1+\bigcap_{P\in D_{f\alpha}}\mathfrak{m}_P$, *and therefore*, $R_{f\alpha}=\bigcap_{P\in D_{f\alpha}}\mathfrak{m}_P$.

Proof. Let $X=\mathcal{U}_{f\alpha}$ and $X_0\subset X$ as in Notation/Remarks 4.4.

For the inclusion "$\subseteq$", consider the partition $D_{f\alpha}=D^2\cup D^1\cup D^0$, where

- $P\in D^2$ if and only if $v_P(\mathfrak{R}_{f\alpha})\neq 0$,
- $P\in D^1$ if and only if $v_P(\mathfrak{R}_{f\alpha})=0$ and $\mathfrak{R}_{f\alpha}$ is not contained in $1+\mathfrak{m}_P$,
- $P\in D^0$ if and only if $\mathfrak{R}_{f\alpha}\subseteq 1+\mathfrak{m}_P$.

Clearly, in order to show that $\mathfrak{R}_{f\alpha}\subseteq 1+\bigcap_{P\in D_{f\alpha}}\mathfrak{m}_P$, we have to show that D^2 and D^1 are empty. By contradiction, suppose that at least one of the sets D^2, D^1 is nonempty.

Case 1. *D^2 is nonempty.*

Let $P\in D^2$ and $t\in\mathfrak{R}_{f\alpha}$ be such that $v_P(t)\neq 0$. Using the weak approximation lemma, choose any principal v_P-unit u' such that $v_{P'}(t+u'-1)>0$ for all $P'\neq P$ from $D_{f\alpha}$. Since $u'\in 1+\mathfrak{m}_P$, it follows by Lemma 4.3 that $u'\in\mathcal{U}_{f\alpha}$. Since $t\in\mathfrak{R}_{f\alpha}\subseteq X_0$, and $u'\in X=\mathcal{U}_{f\alpha}$, we get (by the definition of X_0) that $t*u'\in\mathcal{U}_{f\alpha}$. On the other hand, one has $u:=t*u'=t+u'-1$, and therefore $v_{P''}(u)\neq 0$ for all $P''\in D_{f\alpha}$, thus contradicting Lemma 4.3(3).

Case 2. *D^2 is empty, and D^1 nonempty.*

For $P\in D^1$ we have $\mathfrak{R}_{f\alpha}\subset\mathcal{O}_P^\times$ and $\mathfrak{R}_{f\alpha}$ not contained in $1+\mathfrak{m}_P$. In particular, the image $\overline{R}_{f\alpha}$ of $R_{f\alpha}$ under the residue map $\mathcal{O}_P\to\kappa(P)$ is a nontrivial k-subring of $\kappa(P)$. Since $\kappa(P)$ is a finite field extension of k, it follows that $\overline{R}_{f\alpha}$ is a k-subfield of $\kappa(P)$. Hence there exists $t\in\mathfrak{R}_{f\alpha}$ whose image in $\kappa(P)$ is the given element $\alpha\in k$. In order to conclude, using the weak approximation lemma, choose $u'\in 1+\mathfrak{m}_P\subseteq\mathcal{U}_{f\alpha}$ such that $v_P(t+u'-1-\alpha)>0$ for all $P'\neq P$ from $D_{f\alpha}$. Then reasoning as above, it follows that $t*u'\in\mathcal{U}_{f\alpha}=X$. On the other hand, as above, $u:=t*u'=u+u'-1$ has the property that $v_{P''}(u-\alpha)>0$ for all $P''\in D_{f\alpha}$. Therefore, u is a $v_{P''}$-unit with residue $\bar{u}=\alpha$ in $\kappa(P'')$ for all $P''\in D_{f\alpha}$. Recalling that $u=t*u'\in\mathcal{U}_{f\alpha}$, for $c=1$ one has $\{u_c, u_{\alpha,c}\}=\{u, u+\alpha(1-u)^n\}$, and $\alpha=\bar{u}\in\kappa(P)$ for all $P\in D_{f\alpha}$. Thus α is an n-th power in $K_P K_{u,\alpha,c}$ for all $P\in D_{f\alpha}$, contradicting Lemma 4.3(3).

For the converse inclusion "$\supseteq$", we have to show that for every $u_1 = 1 + \tilde{u} \in 1 + \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P$ with $\tilde{u} \in \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P$, the following hold:

(a) $u_1 \in X_0$, or equivalently, $u_1 * u \in X$ for all $u \in X$.

(b) $u_1 \circ X_0 \subseteq X_0$, or equivalently, $u_1 \circ u_0 * u \in X$ for all $u_0 \in X_0$, $u \in X$.

For (a), we show that $u_1 * u \in X$ for all $u \in X$. Indeed, $t := u_1 * u = u_1 + u - 1 = u + \tilde{u}$. Since $u \in X$, by Lemma 4.3(3), there exists $P \in D_{f\alpha}$ such that $u_c, u_{\alpha,c} \subset \mathcal{O}_P^\times$, and for all δ^* one has that $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_P K_{u,\alpha,c}$. We now claim that $K_P K_{u,\alpha,c} = K_P K_{t,\alpha,c}$. Indeed, since $t = u + \tilde{u}$, $v_P(\tilde{u}) > 0$, and $u_c, u_{\alpha,c} \in \mathcal{O}_P^\times$, it follows that $t_c, t_{\alpha,c} \in \mathcal{O}_P$ and $\bar{t}_c = \bar{u}_c$ and $\bar{t}_{\alpha,c} = \bar{u}_{\alpha,c}$ in $\kappa(P)^\times$. Hence, by Hensel's lemma it follows that $\sqrt[n]{u_c}$, $\sqrt[n]{u_{\alpha,c}}$ and $\sqrt[n]{t_c}$, $\sqrt[n]{t_{\alpha,c}}$ generate the same extension of K_P. Therefore, if $f \cup \alpha \cup \beta^* \in \mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $L_Q = K_P K_{u,\alpha,c}$, it is nontrivial over $K_P K_{u,\alpha,c} = K_P K_{t,\alpha,c}$, and thus also over $K_{t,\alpha,c} \subset K_P K_{t,\alpha,c}$, etc.

For (b), we show that $u_0 * u \in X$ for all $u \in X$ implies that $(u_1 \circ u_0) * u \in X$ for all $u \in X$. First, recall that by Notation/Remarks 4.4(4), it follows that $u_0 \in \mathcal{O}_P^\times$ for all $P \in D_{f\alpha}$. Hence one gets

$$t := u_1 \circ u_0 * u = ((u_1 - 1)(u_0 - 1) + 1) + u - 1 = \tilde{u}(u_0 - 1) + u = u + \tilde{u}',$$

where $\tilde{u}' = \tilde{u}(u_0 - 1) \in \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P$ since $u_0 \in \bigcap_{P \in D_{f\alpha}} \mathcal{O}_P^\times$ and $\tilde{u} \in \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P$. On the other hand, $u \in \mathcal{U}_{f\alpha}$ and $P \in D_{f\alpha}$ are such that $\mathrm{H}_{f\delta^*\alpha}$ is nontrivial over $K_P K_{u,\alpha,c}$. Hence $K_P K_{u,\alpha,c} = K_P K_{t,\alpha,c}$, by the fact that $\bar{u} = \bar{t}$ in $\kappa(P)^\times$ (and Hensel's lemma). Finally it follows that $t \in \mathcal{U}_{f\alpha} = X$, as claimed.

This concludes the proof of Key Lemma 4.7. $\square$

5. Proof of Theorems 1.1 and 1.2

5A. ***Defining the k-valuation rings.*** In the notation and hypotheses of the previous sections, let $K|k$ be a smooth fibration of a finitely generated field K with $\dim(K) = 2$, and X the complete smooth k-curve with $K = k(X)$. By Riemann–Roch, if $P \in X$ is a closed point and $m \gg 0$, there exist functions $f \in K$ such that $(f)_\infty = mP$, and letting m be prime to n, we have $P \in D_f$. Further, setting $\lambda := \kappa(P)$, there exist "many" $\alpha \in k^\times$ such that α is not an n-th power in $\kappa(P)$. Hence there exists α such that $D_{f\alpha}$ is nonempty. Thus by Key Lemma 4.7, it follows that $\mathfrak{R}_{f\alpha} = 1 + \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P$.

For f and α as above, we set $g := f + 1$, and notice that $(g)_\infty = mP$, etc. We repeat the constructions above and we get $\mathfrak{R}_{g\alpha} = 1 + \bigcap_{Q \in D_{g\alpha}} \mathfrak{m}_Q$.

Since $|\mathrm{div}(f)| \cap |\mathrm{div}(g)| = \{P\}$, by the weak approximation lemma one has

$$\Big(1 + \bigcap_{P \in D_{f\alpha}} \mathfrak{m}_P\Big) \cdot \Big(1 + \bigcap_{Q \in D_{g\alpha}} \mathfrak{m}_Q\Big) = 1 + \mathfrak{m}_P.$$

Therefore, one can recover $\mathcal{O}_P$, $\mathfrak{m}_P$ from $\mathfrak{R}_{f\alpha}$ and $\mathfrak{R}_{g\alpha}$ as follows:

$$\mathfrak{m}_P = \mathfrak{R}_{f\alpha} \cdot \mathfrak{R}_{g\alpha} - 1 \quad \text{and} \quad \mathcal{O}_P = \{t \in K \mid t\,\mathfrak{m}_P \subseteq \mathfrak{m}_P\}.$$

We thus have a *first-order recipe* to define all the k-valuation rings of $K\,|\,k$:

Recipe 5.1. Let $K = k(X)$ with X a complete smooth k-curve X. Suppose that a predicate $\psi(\mathfrak{x})$ is given which defines k inside K, i.e., $k = \{x \in K \mid \psi(x) \text{ holds in } K\}$. Let $n \neq \operatorname{char}(K)$ be a prime number, and notice that describing the k-valuation rings of $K\,|\,k$ is equivalent to describing the $k[\mu_{2n}]$-valuation rings of $K[\mu_{2n}]$. Supposing that $\mu_{2n} \subset K$, consider the following steps:

(1) For every $\delta \in k^\times$ let $\mathrm{H}_\delta = \{\beta \in k^\times \mid v_s(\beta - 1) > 2v_s(2n) \text{ if } s \notin V_\delta\} \subset k^\times$.
Note that H_δ is a definable subset of K, provided a predicate $\psi(\mathfrak{x})$ is given which defines k inside K. Indeed, given the global field k, the valuation rings $\mathcal{O}_s$, $\mathfrak{m}_s$ of k are definable inside k by Rumely's recipe [1980] mentioned in Section 1. Thus, one has

$$\beta \in \mathrm{H}_\delta \quad \text{if and only if} \quad \forall \mathcal{O}_s, \mathfrak{m}_s\, (\delta \notin \mathcal{O}_s \Rightarrow \beta - 1 \in 4n^2 \cdot \mathfrak{m}_s).$$

(2) For every $f \in K$ and $\alpha \in k^\times$, set $\mathrm{H}_{f\delta\alpha} := f \cup \alpha \cup \mathrm{H}_\delta \subset \mathrm{H}^3(K, \mathbb{Z}/n(2))$.

(3) Let $\mathcal{U}_{f\alpha} := \{u \in K^\times \mid \mathrm{H}_{f\delta\alpha} \text{ is nontrivial over } K_{u,\alpha,c} \text{ for all } c \in k,\ \delta \in k^\times\}$, where $K_{u,\alpha,c} := K[\sqrt[n]{u_c}, \sqrt[n]{u_{\alpha,c}}\,]$ and $u_c := 1 - c(1-u)$, $u_{\alpha,c} =: u_c + \alpha c^n (1-u)^n$.
Note that the fact that $\mathrm{H}_{f\delta\alpha}$ is nontrivial over $K_{u,\alpha,c}$ is first-order expressible as follows (see Section 2C): $f \cup \alpha$ is nontrivial and there exists $\beta \in \mathrm{H}_\delta$ such that the reduced norm $N_{f,\alpha}$ of the division algebra $A_{f,\alpha}$ does not represent β over $K_{u,\alpha,c}$.

(4) Let $\mathfrak{R}_{f\alpha} := \{u \in \mathcal{U}_{f\alpha} \mid u + \mathcal{U}_{f\alpha} - 1 \subseteq \mathcal{U}_{f\alpha},\, (u-1)(\mathcal{U}_{f\alpha} - 1) + 1 \subseteq \mathcal{U}_{f\alpha}\}$.

(5) Repeat the process above for $g := f + 1$.

(6) Set $\mathfrak{m} := \mathfrak{R}_{f\alpha} \cdot \mathfrak{R}_{g\alpha} - 1$ and $\mathcal{O} := \{t \in K \mid t\,\mathfrak{m} \subseteq \mathfrak{m}\}$.

Conclusion 5.2. The k-valuation rings $\mathcal{O}_P$, $\mathfrak{m}_P$ of $K\,|\,k$ are among the definable sets $\mathcal{O}$, $\mathfrak{m}$. Precisely, for every $\mathcal{O}_P$, $\mathfrak{m}_P$ there exist f and α such that $\mathcal{O} = \mathcal{O}_P$, and $1/f \in \mathfrak{m} = \mathfrak{m}_P$.

5B. ***Concluding the proof of Theorem 1.2.*** First, the above Recipe 5.1 is a uniform first-order description of the k-valuation rings of function fields of complete smooth k-curves.

In order to give the formula $\deg_N(\mathfrak{t})$, we first recall that k is a Hilbertian field. Hence, for every nonconstant function $t \in K$ there exist (infinitely many) specializations $t \mapsto a \in k$ such that the point $P \in X$ with $\bar{t} = a$ in $\kappa(P)$ is unique. Hence $[\kappa(P) : k] = [K : k(t)]$ is the degree of t. Thus, one possibility for the formula

$\mathrm{def}_N(\mathfrak{t})$ would be

$$(\forall \mathcal{O}, \mathfrak{m} : \mathfrak{t} \in k + \mathfrak{m} \Rightarrow [\mathcal{O}/\mathfrak{m} : k] \leq N) \,\&\, (\exists \mathcal{O}, \mathfrak{m} : \mathfrak{t} \in k + \mathfrak{m} \,\&\, [\mathcal{O}/\mathfrak{m} : k] = N).$$

For the formula $\psi^R(\mathfrak{t}, \mathfrak{t}')$, let $R \subset K$ be the integral closure of $k[t]$ in K. Then for any k-valuation ring $\mathcal{O}$, one has $t \in \mathcal{O}$ if and only if $k[t] \subset \mathcal{O}$ if and only if $R \subset \mathcal{O}$. Further, R is actually the intersection of all the $\mathcal{O}$ which contain $k[t]$, or equivalently, which contain t. Thus the formula $\psi^R(\mathfrak{t}, \mathfrak{t}')$ could be

$$\forall \mathcal{O}, \mathfrak{m}\, (\mathfrak{t} \in \mathcal{O} \Rightarrow \mathfrak{t}' \in \mathcal{O}).$$

Finally, for the formula $\psi^0(\mathfrak{t}, \mathfrak{t}')$, let $R \subset K$ be the integral closure of $k[t]$ in K. Recall that $t' \in K$ lies in $k[t]$ if and only if $t' \in R$ and for all $\mathcal{O}, \mathfrak{m}$ one has that if the residue $\bar{t} \in \mathcal{O}/\mathfrak{m}$ lies in $k \subset \mathcal{O}/\mathfrak{m}$, then the residue $\bar{t}' \in \mathcal{O}/\mathfrak{m}$ lies in $k \subset \mathcal{O}/\mathfrak{m}$. (Indeed, this follows again from the fact that k is Hilbertian.) Thus one possibility for the formula $\psi^0(\mathfrak{t}, \mathfrak{t}')$ could be

$$\forall \mathcal{O}, \mathfrak{m}\, ((\mathfrak{t} \in \mathcal{O} \Rightarrow \mathfrak{t}' \in \mathcal{O}) \,\&\, (\mathfrak{t} \in \mathfrak{m} + k \Rightarrow \mathfrak{t}' \in \mathfrak{m} + k)).$$

This completes the proof of Theorem 1.2.

5C. *Proof of Theorem 1.1.* In order to prove Theorem 1.1, we recall that by the main results of [Pop 2002] combined with the description of the absolute constants in [Poonen 2007], one has the following:

(1) For every finitely generated field K over a number field k with $d := \mathrm{tr.deg}(K \,|\, k)$, there exists a formula with d free variables $\varphi_K^0(\mathfrak{t}_1, \ldots, \mathfrak{t}_d)$ such that the sentence
$$\exists t_1, \ldots, t_d\, \varphi_K^0(t_1, \ldots, t_d)$$
is true in K. Moreover, if L is any other finitely generated field such that $\varphi_K^0(u_1, \ldots, u_d)$ is true in L for some choice of $u_1, \ldots, u_d \in L$, then the map $(t_1, \ldots, t_d) \mapsto (u_1, \ldots, u_d)$ extends to an embedding of fields $K \hookrightarrow L$.

(2) Now suppose that $d = 1$. Then using the above $\mathrm{deg}_N(\mathfrak{t})$ we have proved the following theorem.

Theorem 5.3. *Let ϑ_K be the sentence $\exists \mathfrak{t}\, (\varphi_K^0(\mathfrak{t}) \,\&\, \mathrm{deg}_N(\mathfrak{t}))$. Then the following hold*:

(1) *The sentence ϑ_K is true in K if and only if there exists some $t \in K$ such that $\varphi_K^0(t)$ is true in K, and t has degree N in K.*

(2) *Suppose that ϑ_K is true in K. Then for every finitely generated field L, the sentence ϑ_K is true in L if and only if K and L are isomorphic as fields.*

Acknowledgements

I would like to thank the participants at several activities, e.g., AWS 2003, AIM Workshop 2004, INI Cambridge 2005, HIM Bonn 2009, ALANT III in 2014, for the debates on the topic and suggestions concerning this problem. Special thanks are due to Bjorn Poonen and Jakob Stix for discussing technical aspects of the proofs. The author would also like to thank the University of Heidelberg and University of Bonn for the excellent working conditions during his visit in 2015–16 as a *Humboldt Preisträger.*

References

[Abhyankar 1965] S. S. Abhyankar, "Resolution of singularities of arithmetical surfaces", pp. 111–152 in *Arithmetical Algebraic Geometry* (West Lafayette, IN, 1963), edited by O. F. G. Schilling, Harper & Row, New York, 1965. MR Zbl

[Illusie 1979] L. Illusie, "Complexe de de Rham–Witt et cohomologie cristalline", *Ann. Sci. École Norm. Sup.* (4) **12**:4 (1979), 501–661. MR Zbl

[Kato 1986] K. Kato, "A Hasse principle for two-dimensional global fields", *J. Reine Angew. Math.* **366** (1986), 142–183. MR Zbl

[Merkurjev and Suslin 1982] A. S. Merkurjev and A. A. Suslin, "K-cohomology of Severi–Brauer varieties and the norm residue homomorphism", *Izv. Akad. Nauk SSSR Ser. Mat.* **46**:5 (1982), 1011–1046. In Russian; translated in *Math. USSR-Izv.* **21**:2 (1983), 307–340. MR Zbl

[Poonen 2007] B. Poonen, "Uniform first-order definitions in finitely generated fields", *Duke Math. J.* **138**:1 (2007), 1–22. MR Zbl

[Pop 2002] F. Pop, "Elementary equivalence versus isomorphism", *Invent. Math.* **150**:2 (2002), 385–408. MR Zbl

[Pop 2003] F. Pop, "Elementary equivalence of finitely generated fields", course notes, Arizona Winter School, 2003, available at http://swc.math.arizona.edu/aws/2003/03PopNotes.pdf.

[Rumely 1980] R. S. Rumely, "Undecidability and definability for the theory of global fields", *Trans. Amer. Math. Soc.* **262**:1 (1980), 195–217. MR Zbl

[Scanlon 2008] T. Scanlon, "Infinite finitely generated fields are biinterpretable with $\mathbb{N}$", *J. Amer. Math. Soc.* **21**:3 (2008), 893–908. Correction in **24**:3 (2011), 917. MR Zbl

Communicated by Bjorn Poonen
Received 2016-04-23 Revised 2017-05-11 Accepted 2017-08-03

pop@math.upenn.edu *Department of Mathematics, University of Pennsylvania, Philadelphia, PA, United States*

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2113

On the algebraic structure of iterated integrals of quasimodular forms

Nils Matthes

We study the algebra $\mathcal{I}^{\mathrm{QM}}$ of iterated integrals of quasimodular forms for $\mathrm{SL}_2(\mathbb{Z})$, which is the smallest extension of the algebra QM_* of quasimodular forms which is closed under integration. We prove that $\mathcal{I}^{\mathrm{QM}}$ is a polynomial algebra in infinitely many variables, given by Lyndon words on certain monomials in Eisenstein series. We also prove an analogous result for the M_*-subalgebra $\mathcal{I}^{M}$ of $\mathcal{I}^{\mathrm{QM}}$ of iterated integrals of modular forms.

1. Introduction

Quasimodular forms, a generalization of modular forms, were first introduced in [Kaneko and Zagier 1995] in a context motivated by mathematical physics. The $\mathbb{C}$-algebra QM_* of quasimodular forms for the full modular group $\mathrm{SL}_2(\mathbb{Z})$ can be defined, in a slightly ad hoc fashion, as the polynomial ring $\mathbb{C}[E_2, E_4, E_6]$, where E_{2k} denotes the normalized Eisenstein series of weight $2k$:

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} n^{2k-1} \frac{q^n}{1-q^n}, \quad q = e^{2\pi i \tau},$$

where B_{2k} are the Bernoulli numbers. In particular, QM_* contains the algebra of modular forms $M_* \cong \mathbb{C}[E_4, E_6]$.

The derivative of a quasimodular form (of weight k) is again a quasimodular form (of weight $k+2$); this was essentially already known to Ramanujan (see [Zagier 2008, Proposition 15]). On the other hand, the integral of a quasimodular form is in general not quasimodular. For example, a primitive of E_2 would have to be of weight zero, but every quasimodular form of weight zero is constant.

The goal of this paper is to study the smallest algebra extension of QM_* which is closed under integration. For this, the idea is to iteratively adjoin primitives

MSC2010: primary 11F11; secondary 11F67.
Keywords: quasimodular forms, iterated integrals.

to QM_*, which eventually leads to adjoining all (indefinite) *iterated integrals*

$$I(f_1,\dots,f_n;\tau)=(2\pi i)^n \int_{\tau\le\tau_1\le\cdots\le\tau_n\le i\infty}\cdots\int f_1(\tau_1)\cdots f_n(\tau_n)\,\mathrm{d}\tau_1\cdots\mathrm{d}\tau_n, \qquad (1\text{-}1)$$

where $f_1,\dots,f_n$ are quasimodular forms (a precise definition will be given in Definition 2.6). The integrals (1-1) were first studied by Manin [2006], and later by Brown [2016] and Hain [2016], in the case where all the f_i are modular forms.[1] In all of these treatments, the focus lies rather on arithmetic aspects of these iterated integrals, for example their special values at cusps of the upper half-plane. By contrast, we study them solely as holomorphic functions of τ. It is also worth noting that even in the modular case, the iterated integrals we study in the present paper are slightly more general than the ones introduced in [Manin 2006; Brown 2016; Hain 2016]. For example, if $f(\tau)$ is a modular form of weight k, then the integral $\int_\tau^{i\infty} f(\tau_1)\tau_1^n\,\mathrm{d}\tau_1$ is an iterated integral of modular forms in the sense of the present paper *for every* $n\ge 0$, while [Manin 2006; Brown 2016; Hain 2016] also require $n\le k-2$.

Now let $\mathcal{I}^{\mathrm{QM}}$ be the QM_*-algebra generated by all the integrals (1-1), which is the smallest algebra extension of QM_* closed under integration. It turns out that $\mathcal{I}^{\mathrm{QM}}$ is not finitely generated, but still has a manageable structure, which is captured by the notion of shuffle algebra (which is just the graded dual of the tensor algebra with a certain commutative multiplication, the so-called shuffle product) [Reutenauer 1993]. More precisely, let $V=\mathbb{C}\cdot E_2\oplus M_*$ be the $\mathbb{C}$-vector space spanned by all modular forms and the Eisenstein series E_2, and let $\mathbb{C}\langle V\rangle$ be the shuffle algebra on V. Our main result is the following.

Theorem (Theorem 4.3). *The* QM_**-linear morphism*

$$\varphi^{\mathrm{QM}}:\mathrm{QM}_*\otimes_{\mathbb{C}}\mathbb{C}\langle V\rangle\to\mathcal{I}^{\mathrm{QM}}, \qquad [f_1\mid\cdots\mid f_n]\mapsto I(f_1,\dots,f_n;\tau)$$

is an isomorphism of QM_**-algebras.*

A similar result holds for the M_*-subalgebra $\mathcal{I}^M$ of $\mathcal{I}^{\mathrm{QM}}$ of iterated integrals of modular forms (see Theorem 4.5).[2] The surjectivity of φ^{QM} can be reduced to the fact that every quasimodular form can be written uniquely as a polynomial in n-th derivatives of modular forms and the Eisenstein series E_2; see [Zagier 2008, Proposition 20]. The proof of injectivity is more elaborate and amounts to showing that iterated integrals of modular forms and the Eisenstein series E_2 are linearly

[1] More precisely, Manin only defined iterated integrals of cusp forms, and the extension to all modular forms is due to Brown.

[2] After this paper was submitted for publication, the author learned that, in the case of iterated integrals of modular forms, a very similar result has also been proved by Brown [2017, Proposition 4.4] using a slightly different method.

independent over QM_*. It extends results of [Lochak et al. 2017], which dealt with iterated integrals of Eisenstein series. In both cases, the key is to use a general result on linear independence of iterated integrals [Deneufchâtel et al. 2011]. It would be interesting to prove similar results for quasimodular forms for congruence subgroups.

The Milnor–Moore theorem [Milnor and Moore 1965] states that if k has characteristic zero, then $k\langle V\rangle$ is isomorphic to a polynomial algebra (usually in infinitely many variables). Fixing a (totally ordered) basis $\mathcal{B}$ of V, Radford [1979] has given explicit generators of $k\langle V\rangle$ in terms of Lyndon words on $\mathcal{B}$ (see Section 4). Using this, we get the following theorem.

Theorem (Theorem 4.9). *Let $\mathcal{B}$ be a basis of $\mathbb{C}\cdot E_2 \oplus M_*$. We have a natural isomorphism*

$$\mathcal{I}^{\mathrm{QM}} \cong \mathrm{QM}_*[\mathrm{Lyn}(\mathcal{B}^*)], \tag{1-2}$$

where the right-hand side is the polynomial QM_**-algebra on the set* $\mathrm{Lyn}(\mathcal{B}^*)$ *of Lyndon words of* $\mathcal{B}$.

Again, a similar result holds for $\mathcal{I}^{M}$. Since QM_* has an explicit basis given by monomials in the Eisenstein series E_2, E_4 and E_6, the isomorphism (1-2) can be made completely explicit, and may be viewed as an analog of the isomorphism $\mathrm{QM}_* \cong \mathbb{C}[E_2, E_4, E_6]$ [Kaneko and Zagier 1995].

Finally, we note that classically, integrals of modular forms play an important role in Eichler–Shimura theory, where they give rise to group-cocycles (say for $\mathrm{SL}_2(\mathbb{Z})$ or more generally for some congruence subgroup thereof) with values in homogeneous polynomials. This has been generalized by Manin [2006], and later by Brown [2016] and Hain [2016], who attach certain nonabelian cocycles to iterated integrals of modular forms. Although it is not the main focus of this article, in the Appendix we show how one can attach cocycles to quasimodular forms (for $\mathrm{SL}_2(\mathbb{Z})$), partly since we found no mention of this in the literature. On the other hand, we leave the definition and study of cocycles attached to iterated integrals of quasimodular forms for future investigation.

The plan of the paper is as follows. In Section 2, we collect the necessary background on quasimodular forms and their iterated integrals. In Section 3, we prove a linear independence result for iterated integrals of quasimodular forms. This result is then put to use in Section 4, where the main results are proved. In the Appendix, we discuss the above-mentioned generalization of the classical Eichler–Shimura theory to quasimodular forms for $\mathrm{SL}_2(\mathbb{Z})$.

2. Preliminaries

Throughout the paper, all modular and quasimodular forms will be for $\mathrm{SL}_2(\mathbb{Z})$. We fix some notation. Let $\mathfrak{H} = \{z \in \mathbb{C} \mid \mathrm{Im}(z) > 0\}$ be the upper half-plane with canonical coordinate τ. For every $k \in \mathbb{Z}$, we have a group action of $\mathrm{SL}_2(\mathbb{Z})$ on the set of all functions $f : \mathfrak{H} \to \mathbb{C}$ (not necessarily holomorphic), defined by $(\gamma, f) \mapsto f|_k \gamma$, where

$$(f|_k \gamma)(\tau) := (c\tau + d)^{-k} f\Big(\frac{a\tau+b}{c\tau+d}\Big).$$

For fixed $\tau \in \mathfrak{H}$, we also define a map $X : \mathrm{SL}_2(\mathbb{Z}) \to \mathbb{C}$ by $X(\gamma) = \frac{1}{2\pi i}\frac{c}{c\tau+d}$. Note that X has infinite, and thus Zariski dense, image.

Recap of modular forms. Denote by M_k the space of modular forms of weight $k \in \mathbb{Z}$. By definition, these are the holomorphic functions $f : \mathfrak{H} \to \mathbb{C}$, which satisfy $f|_k \gamma = f$ for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$, and which are "holomorphic at the cusp". The latter condition means that in the Fourier expansion $f(\tau) = \sum_{n\in\mathbb{Z}} a_n q^n$ (which exists since for $\gamma = \left(\begin{smallmatrix}1&1\\0&1\end{smallmatrix}\right) \in \mathrm{SL}_2(\mathbb{Z})$, the condition $f|_k \gamma = f$ is just $f(\tau+1) = f(\tau)$ for all τ), $a_n = 0$ for all $n < 0$. Examples of modular forms include the Eisenstein series

$$E_{2k}(\tau) = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} n^{2k-1} \frac{q^n}{1-q^n} = 1 - \frac{4k}{B_{2k}} \sum_{n=1}^{\infty} \Big(\sum_{d|n} d^{2k-1}\Big) q^n,$$

which is a modular form of weight $2k$, for $k \geq 2$ (the B_{2k} are Bernoulli numbers). The $\mathbb{C}$-vector space of all modular forms M_* is a graded (for the weight) $\mathbb{C}$-algebra $M_* = \bigoplus_{k\in\mathbb{Z}} M_k$, which is well-known to be isomorphic to the polynomial algebra $\mathbb{C}[E_4, E_6]$. Proofs of all these facts and much more on modular forms can be found, for example, in [Zagier 2008].

Quasimodular forms. Quasimodular forms are a generalization of modular forms which was first introduced in [Kaneko and Zagier 1995]; see also [Bloch and Okounkov 2000, §3; Zagier 2008, §5.3]. The definition we give here is due to W. Nahm[3] and is also used for example in [Martin and Royer 2005].

Definition 2.1. Let $k, p \in \mathbb{Z}$ with $p \geq 0$. A *quasimodular form* of weight k and depth $\leq p$ is a function $f : \mathfrak{H} \to \mathbb{C}$ with the following property: there exist holomorphic functions $f_r : \mathfrak{H} \to \mathbb{C}$, for $0 \leq r \leq p$, which have Fourier expansions $\sum_{n=0}^{\infty} a_n q^n$ such that

$$(f|_k \gamma)(\tau) = \sum_{r=0}^{p} f_r(\tau) X(\gamma)^r, \quad \text{for all } \gamma \in \mathrm{SL}_2(\mathbb{Z}). \tag{2-1}$$

We denote by $\mathrm{QM}_k^{\leq p}$ the $\mathbb{C}$-vector space of quasimodular forms of weight k and depth $\leq p$, and set

$$\mathrm{QM}_k := \bigcup_{p\geq 0} \mathrm{QM}_k^{\leq p}, \qquad \mathrm{QM}_* := \bigoplus_{k\in\mathbb{Z}} \mathrm{QM}_k.$$

[3] See [Zagier 2008, §5.3].

Remark 2.2. (i) It is clear from the definition that, if $f_1 \in \mathrm{QM}_{k_1}^{\leq p_1}$, $f_2 \in \mathrm{QM}_{k_2}^{\leq p_2}$, then $f_1 f_2 \in \mathrm{QM}_{k_1+k_2}^{\leq p_1+p_2}$. In other words, QM_* is a graded (for the weight) and filtered (for the depth) $\mathbb{C}$-algebra.

(ii) Using the fact that X is Zariski dense, it is easy to see that the functions $f_r(\tau)$ are uniquely determined by $f(\tau)$. Also, applying (2-1) with $\gamma = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$, we see that $f_0(\tau) = f(\tau)$. In particular, every quasimodular form is holomorphic on $\mathfrak{H}$ and at the cusp.

Every modular form is a quasimodular form of depth zero; more precisely, $M_k = \mathrm{QM}_k^{\leq 0}$. An example of a quasimodular form which is not modular is the Eisenstein series of weight two $E_2(\tau) = 1 - 24\sum_{n=1}^{\infty} n \frac{q^n}{1-q^n}$, which transforms as

$$(E_2|_2\,\gamma)(\tau) = E_2(\tau) + 12X(\gamma) = E_2(\tau) - \frac{6i}{\pi}\frac{c}{c\tau+d} \tag{2-2}$$

for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. In particular, $E_2 \in \mathrm{QM}_2^{\leq 1} \setminus M_2$.

The following proposition recalls basic properties of QM_* that will be of use later.

Proposition 2.3. (i) *The $\mathbb{C}$-algebra QM_* is closed under the differential operator $D := \frac{1}{2\pi i}\frac{d}{d\tau} = q\frac{d}{dq}$. More precisely, for f quasimodular of weight k and depth $\leq p$, we have*

$$(D(f)|_{k+2}\,\gamma)(\tau) = \sum_{r=0}^{p+1}(D(f_r)(\tau) + (k-r+1)f_{r-1}(\tau))X(\gamma)^r.$$

In particular, $D(\mathrm{QM}_k^{\leq p}) \subset \mathrm{QM}_{k+2}^{\leq p+1}$ for all $k, p \in \mathbb{Z}$.

(ii) *We have*

$$\mathrm{QM}_k = \begin{cases} \{0\} & \text{if } k<0, \\ \mathbb{C}\cdot E_2 & \text{if } k=2, \\ D(\mathrm{QM}_{k-2}) \oplus M_k & \text{else.} \end{cases}$$

In particular, $\mathrm{QM}_* = \mathbb{C}\cdot E_2 \oplus D(\mathrm{QM}_*) \oplus M_*$, and

$$\mathrm{QM}_* \cong \mathbb{C}[E_2, E_4, E_6]$$

as graded $\mathbb{C}$-algebras.

Proof. For (i), simply apply D to both sides of (2-1). The first equality in (ii) follows from [Zagier 2008, Proposition 20(iii)], and the isomorphism $\mathrm{QM}_* \cong \mathbb{C}[E_2, E_4, E_6]$ is essentially a consequence of this, but can also be proved independently (see [Bloch and Okounkov 2000, Proposition 3.5(ii)]). □

Remark 2.4. Relaxing the condition in the definition of quasimodular forms that every f_r be a holomorphic function, one can define the notion of *weakly quasimodular form* of weight k and depth $\leq p$ as a meromorphic function $f : \mathfrak{H} \to \mathbb{C}$ satisfying (2-1), but where the functions $f_r(\tau)$ are only required to be meromorphic on $\mathfrak{H}$ and have Fourier series of the form $\sum_{n=-M}^{\infty} a_n q^n$ (f_r is "meromorphic at the cusp”). As in the case of quasimodular forms, one shows easily that the functions $f_r(\tau)$ are uniquely determined by $f(\tau)$ (see Remark 2.2). Moreover, Proposition 2.3(i) generalizes straightforwardly to weakly quasimodular forms.

We end this subsection with a short lemma, for which we couldn't find a suitable reference. Denote by $\Delta = \frac{1}{1728}(E_4^3 - E_6^2)$ Ramanujan's cusp form of weight 12.

Lemma 2.5. *Let $g \in \mathrm{QM}_* \setminus \{0\}$ and $\alpha \in \mathbb{C}$ be such that*

$$D(g) = (\alpha E_2)\cdot g. \tag{2-3}$$

Then α is a nonnegative integer, and $g = \beta\Delta^\alpha$ for some $\beta \in \mathbb{C}\setminus\{0\}$.

Proof. Let $g = \sum_{n=0}^{\infty} a_n q^n$, so that $D(g) = \sum_{n=0}^{\infty} n a_n q^n$. Comparing coefficients on both sides of (2-3) yields that α equals the smallest integer $m \geq 0$ such that $a_m \neq 0$. On the other hand, $D(\Delta)/\Delta = E_2$ [Zagier 2008, proof of Proposition 7], and from the chain rule, $D(\Delta^\alpha)/\Delta^\alpha = \alpha E_2$, which gives the result. □

Iterated integrals on the upper half-plane. Iterated integrals of modular forms were first considered by Manin [2006] (for cusp forms), and later by Brown [2016] (in general). They are generalizations of the classical Eichler integrals

$$\int_\tau^{i\infty} f(z) z^m \,\mathrm{d}z, \quad m = 0, \ldots, k-2, \tag{2-4}$$

where f is a cusp form of weight k [Eichler 1957; Lang 1976]. Extending (2-4) to a general modular form poses the problem of logarithmic divergences, which arise from the constant term in the Fourier series of f. A procedure for regularizing such integrals is described in [Brown 2016], and we borrow it to define iterated integrals of quasimodular forms. Since it is perhaps not so well-known, we give some details for the convenience of the reader.

Let $W \subset \mathcal{O}(\mathfrak{H})$ be the $\mathbb{C}$-subalgebra of holomorphic functions $f : \mathfrak{H} \to \mathbb{C}$, which have an everywhere convergent Fourier series $f(\tau) = \sum_{n=0}^{\infty} a_n q^n$ with $q = e^{2\pi i\tau}$. Note that $\mathrm{QM}_* \subset W$. For $f(\tau) \in W$, let $f^\infty = a_0$, and $f^0(\tau) = f(\tau) - f^\infty = \sum_{n=1}^{\infty} a_n q^n$. Let $\mathbb{C}\langle W\rangle$ (sometimes denoted by $T^c(W)$) be the shuffle algebra [Reutenauer 1993], i.e., the graded dual of the tensor algebra $T(W) = \bigoplus_{k\geq 0} W^{\otimes n}$ on W, where the grading is by the length of tensors. Elements of $(W^{\otimes n})^\vee$ will be written using bar notation $[f_1 \mid f_2 \mid \cdots \mid f_n]$, and a general element of $\mathbb{C}\langle W\rangle$ is a $\mathbb{C}$-linear combination of those. The product on $\mathbb{C}\langle W\rangle$ is the shuffle product $\sqcup\!\sqcup$, which is defined on the basic elements by

$$[f_1 \mid \cdots \mid f_r] \sqcup\!\sqcup [f_{r+1} \mid \cdots \mid f_{r+s}] = \sum_{\sigma\in\Sigma_{r,s}} [f_{\sigma(1)} \mid \cdots \mid f_{\sigma(r+s)}], \tag{2-5}$$

where $\Sigma_{r,s}$ denotes the set of all the permutations on the set $\{1, \ldots, r+s\}$ such that $\sigma^{-1}(1) < \cdots < \sigma^{-1}(r)$ and $\sigma^{-1}(r+1) < \cdots < \sigma^{-1}(r+s)$.

Define a $\mathbb{C}$-linear map $R : \mathbb{C}\langle W\rangle \to \mathbb{C}\langle W\rangle$ by the formula

$$R[f_1 \mid \cdots \mid f_n] = \sum_{i=0}^{n} (-1)^{n-i} [f_1 \mid \cdots \mid f_i] ⧢ [f_n^\infty \mid \cdots \mid f_{i+1}^\infty].$$

Following [Brown 2016, §4], we make the following definition.

Definition 2.6. For $f_1, \dots, f_n \in W$, define their regularized iterated integral

$$I(f_1, \dots, f_n; \tau) := (2\pi i)^n \sum_{i=0}^{n} (-1)^{n-i} \int_\tau^{i\infty} R[f_1 \mid \cdots \mid f_i] \int_0^\tau [f_n^\infty \mid \cdots \mid f_{i+1}^\infty], \quad (2\text{-}6)$$

where

$$\int_a^b [f_1 \mid \cdots \mid f_n] := \int_{0 \le t_1 \le \cdots \le t_n \le 1} (\gamma_a^b)^*(f_1(\tau_1)\,\mathrm{d}\tau_1) \cdots (\gamma_a^b)^*(f_n(\tau_n)\,\mathrm{d}\tau_n)$$

denotes the usual iterated integral along the straight line path γ_a^b from a to b.

Remark 2.7. Using the change of variables $\tau \mapsto q = e^{2\pi i \tau}$, it is easy to see that $I(f_1, \dots, f_n; \tau) \in W[\log(q)]$, where $\log(q) := 2\pi i \tau$. By the same token, if all of the f_i have rational Fourier coefficients, then $I(f_1, \dots, f_n; \tau)$ will also have rational coefficients, as a series in q and $\log(q)$.

Proposition 2.8. *The functions $I(f_1, \dots, f_n; \tau)$ satisfy the following properties.*

(i) *The product of any two of them is given by the shuffle product*

$$I(f_1, \dots, f_r; \tau) I(f_{r+1}, \dots, f_{r+s}; \tau) = \sum_{\sigma \in \Sigma_{r,s}} I(f_{\sigma(1)}, \dots, f_{\sigma(r+s)}; \tau). \quad (2\text{-}7)$$

(ii) *They satisfy the differential equation*

$$\frac{1}{2\pi i} \frac{d}{d\tau}\Big|_{\tau=\tau_0} I(f_1, \dots, f_n; \tau) = -f_1(\tau_0) I(f_2, \dots, f_n; \tau_0). \quad (2\text{-}8)$$

(iii) *We have the integration by parts formulas*

$$I(f_1, \dots, f_i, D(g), f_{i+1}, \dots, f_n; \tau) = I(f_1, \dots, f_i, g f_{i+1}, \dots, f_n; \tau) - I(f_1, \dots, f_i g, f_{i+1}, \dots, f_n; \tau), \quad (2\text{-}9)$$

as well as

$$I(D(g), f_2, \dots, f_n; \tau) = I(g f_2, f_3, \dots, f_n; \tau) - g(\tau) I(f_2, \dots, f_n; \tau),$$

and

$$I(f_1, \dots, f_{n-1}, D(g); \tau) = g(i\infty) I(f_1, \dots, f_{n-1}; \tau) - I(f_1, \dots, f_{n-1} g; \tau).$$

Proof. Using the definition (2-6), all of these follow from the analogous properties for usual iterated integrals; see, e.g., [Hain 1987]. $\square$

A criterion for linear independence of iterated integrals. Let $\mathrm{Frac}(W)$ be the field of fractions of the $\mathbb{C}$-algebra W introduced in the last subsection. By the quotient rule, it is easy to see that $\mathrm{Frac}(W)$ is closed under $D = \frac{1}{2\pi i}\frac{d}{d\tau}$.

The following theorem is a special case of the main result of [Deneufchâtel et al. 2011].

Theorem 2.9. *Let $\mathcal{F} = (f_i)_{i\in I}$ be a family of elements of W, and let $\mathcal{C} \subset \mathrm{Frac}(W)$ be a subfield which is closed under D and contains $\mathcal{F}$. The following are equivalent*:

(i) *The family of iterated integrals $(I(f_1, \dots, f_n; \tau) \mid f_i \in I,\ n \geq 0)$ is linearly independent over $\mathcal{C}$.*

(ii) *The family $\mathcal{F}$ is linearly independent over $\mathbb{C}$, and we have*

$$D(\mathcal{C}) \cap \mathrm{Span}_{\mathbb{C}}(\mathcal{F}) = \{0\}.$$

Proof. This is the special case of Theorem 2.1 in [Deneufchâtel et al. 2011], with the notation $k = \mathbb{C}$, $(\mathcal{A}, \mathrm{d}) = (\mathrm{Frac}(\mathcal{O}(\mathfrak{H})), D)$, $X = \{A_{f_i} \mid f_i \in \mathcal{F}\}$, $M = -\sum_{i\in I} f_i A_{f_i}$ and $S = \sum_{n\geq 0} \sum_{f_{i_1},\dots,f_{i_n}\in S} I(f_1, \dots, f_n; \tau)\cdot A_{f_1}\cdots A_{f_n}$. Note that it follows from (2-8) that

$$D(S) = M \cdot S,$$

as required in [loc. cit.]. □

Remark 2.10. Variants of Theorem 2.9 have been known before; see [Brown 2009, Lemma 3.6].

3. Linear independence of iterated integrals of quasimodular forms

In this section, we apply Theorem 2.9 to deduce linear independence of a large family of iterated integrals of quasimodular forms. More precisely, our main result is the following theorem.

Theorem 3.1. *Let $\mathcal{B}$ be a $\mathbb{C}$-linearly independent family of elements of $\mathbb{C}\cdot E_2 \oplus M_*$. Then the family of iterated integrals*

$$(I(f_1, \dots, f_n; \tau) \mid f_i \in \mathcal{B})$$

is linearly independent over $\mathrm{Frac}(\mathrm{QM}_*) \cong \mathbb{C}(E_2, E_4, E_6)$.

Two auxiliary lemmas. For the proof of Theorem 3.1, we need two lemmas.

Lemma 3.2. *Let $f, g \in \mathbb{C}[E_2, E_4, E_6]$ be such that $g \neq 0$ and such that f and g are coprime. Assume that $D(f/g) \in \mathbb{C}[E_2, E_4, E_6]$. Then $g = \beta\Delta^\alpha$ for some $\alpha \in \mathbb{Z}_{\geq 0}$ and some $\beta \in \mathbb{C}\setminus\{0\}$, where $\Delta := \frac{1}{1728}(E_4^3 - E_6^2)$ is Ramanujan's cusp form of weight* 12.

Proof. By the quotient rule, we have

$$D\Big(\frac{f}{g}\Big) = \frac{D(f)g - fD(g)}{g^2} = \frac{D(f) - fD(g)/g}{g}.$$

The left-hand side is contained in $\mathbb{C}[E_2, E_4, E_6]$ by assumption, and since also $D(f)$ and g are in $\mathbb{C}[E_2, E_4, E_6]$, we have $fD(g)/g \in \mathbb{C}[E_2, E_4, E_6]$. But then, as f and g have no common factor, g must divide $D(g)$, i.e., there exists $h \in \mathbb{C}[E_2, E_4, E_6]$ such that

$$D(g) = gh.$$

Since $D : \mathrm{QM}_* \to \mathrm{QM}_*$ is homogeneous of weight 2 (see Proposition 2.3(i)), we have $h \in \mathrm{QM}_2$, i.e., $h = \alpha E_2$ with $\alpha \in \mathbb{C}$. In other words, g solves the differential equation $D(g) = (\alpha E_2) \cdot g$. But by Lemma 2.5, α must be a nonnegative integer and $g = \beta \Delta^\alpha$ for some $\beta \in \mathbb{C} \setminus \{0\}$. □

Lemma 3.3. *Let f be a weakly quasimodular form such that its derivative $D(f)$ is a quasimodular form. Then f is a quasimodular form.*

Proof. It is no loss of generality to assume that f is of weight $k \in \mathbb{Z}$ and depth $\leq p$, where $p \geq 0$. By the definition of weakly quasimodular forms (see also Remark 2.2), there exist uniquely determined meromorphic functions $f_r(\tau)$, for $0 \leq r \leq p$, such that

$$(f|_k \gamma)(\tau) = \sum_{r=0}^{p} f_r(\tau) X(\gamma)^r$$

for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Therefore, we only need to show that every $f_r(\tau)$ is holomorphic, including at the cusp.

To this end, by Proposition 2.3(i), we know that

$$(D(f)|_{k+2}\gamma)(\tau) = \sum_{r=0}^{p+1} (D(f_r)(\tau) + (k - r + 1) f_{r-1}(\tau)) X(\gamma)^r, \tag{3-1}$$

and since $D(f)$ is a quasimodular form by assumption, every coefficient of (3-1) is holomorphic, including at the cusp.

The constant term, with respect to $X(\gamma)$, in (3-1) equals $D(f_0)(\tau)$, which is holomorphic by assumption. But a meromorphic function whose derivative is holomorphic everywhere is itself holomorphic everywhere. An easy induction argument, using the fact that the coefficients of (3-1) are holomorphic, now shows that in fact every $f_r(\tau)$ is holomorphic. □

Proof of Theorem 3.1. We use the criterion of Theorem 2.9 in the case where $\mathcal{C} = \mathrm{Frac}(\mathrm{QM}_*)$ and $\mathcal{F} = \mathcal{B}$. Since $\mathcal{B}$ is linearly independent over $\mathbb{C}$ by assumption, it is enough to prove that if $h \in \mathrm{Frac}(\mathrm{QM}_*)$ then

$$D(h) = \sum_{f \in \mathcal{B}} \alpha_f f \ \text{ and } \ \alpha_f \in \mathbb{C} \quad \Rightarrow \quad \alpha_f = 0, \text{ for all } f \in \mathcal{B}.$$

Also, since $\mathcal{B}$ spans a subspace of $\mathbb{C} \cdot E_2 \oplus M_*$, it clearly suffices to prove that $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$ implies that $D(h) = 0$, or equivalently, that h is constant. Thus, the following proposition completes the proof of Theorem 3.1.

Proposition 3.4. *Suppose that $h \in \mathrm{Frac}(\mathrm{QM}_*) \cong \mathbb{C}(E_2, E_4, E_6)$ is such that $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$. Then h is constant.*

Proof. Write $h = f/g$ with $f, g \in \mathbb{C}[E_2, E_4, E_6]$ such that $g \neq 0$ and f and g are coprime. Writing f as a $\mathbb{C}$-linear combination of its homogeneous components, it is enough to show the proposition for f homogeneous of weight k_f.

First, we know from Lemma 3.2 that $g = \beta\Delta^\alpha$ for some $\alpha \in \mathbb{Z}_{\geq 0}$ and $\beta \in \mathbb{C} \setminus \{0\}$, where Δ is Ramanujan's cusp form of weight 12. In particular, g is a cusp form of weight $k_g = 12\alpha$.

Since f is quasimodular of weight k_f and depth $\leq p$, there exist holomorphic (including at the cusp) functions $f_r(\tau)$, for $0 \leq r \leq p$, such that

$$(f|_{k_f}\gamma)(\tau) = \sum_{r=0}^{p} f_r(\tau)X(\gamma)^r$$

for all $\gamma \in \mathrm{SL}_2(\mathbb{Z})$. Setting $h_r(\tau) := \frac{f_r}{g}(\tau)$, we also have, for $k := k_f - k_g$,

$$(h|_k\gamma)(\tau) = \sum_{r=0}^{p} h_r(\tau)X(\gamma)^r.$$

Moreover, the functions $h_r(\tau)$ are meromorphic; thus, h is a weakly quasimodular form (of weight k and depth $\leq p$). By assumption, $D(h)$ is a quasimodular form (necessarily of weight $k+2$ and depth $\leq p+1$), and using Lemma 3.3, this implies that $h \in \mathrm{QM}_k^{\leq p}$. Therefore, every $h_r(\tau)$ is holomorphic, including at the cusp.

Summarizing, we have seen that $h \in \mathrm{Frac}(\mathrm{QM}_*)$ such that $D(h) \in \mathrm{QM}_*$ implies that $h \in \mathrm{QM}_*$. But we even have $D(h) \in \mathbb{C} \cdot E_2 \oplus M_*$ by assumption, and therefore Proposition 2.3(ii) now implies that h is constant, as was to be shown. □

4. Iterated integrals of quasimodular forms and shuffle algebras

We describe the QM_*-algebra of iterated integrals of quasimodular forms, which is the smallest algebra which contains QM_* and is closed under integration. Using the results of the last section, we show that it is canonically isomorphic to an explicit shuffle algebra. A similar result holds for the M_*-subalgebra of iterated integrals of modular forms.

The algebra of iterated integrals of quasimodular forms.

Definition 4.1. Define $\mathcal{I}^{\mathrm{QM}}$ to be the QM_*-module generated by all iterated integrals of quasimodular forms:

$$\mathcal{I}^{\mathrm{QM}} = \mathrm{Span}_{\mathrm{QM}_*}\{I(f_1, \dots, f_n; \tau) \mid f_i \in \mathrm{QM}_*\}.$$

We also denote by $\mathcal{I}_n^{\mathrm{QM}}$ the QM_*-linear submodule, which is spanned by all of the $I(f_1, \dots, f_r; \tau)$ with $r \leq n$.

The subspaces $\mathcal{I}_n^{\mathrm{QM}}$ define an ascending filtration $\mathcal{I}_\bullet^{\mathrm{QM}}$ on $\mathcal{I}^{\mathrm{QM}}$, called the length filtration (in analogy with the length filtration on iterated integrals [Hain 1987]). It follows from (2-7) that $\mathcal{I}^{\mathrm{QM}}$ is a filtered QM_*-algebra. However, the length is not a grading, as shown by the next result.

Proposition 4.2. *Let $f_1, \dots, f_n$ be quasimodular forms. Then*

$$I(f_1, \dots, f_{i-1}, D(f_i), f_{i+1}, \dots, f_n; \tau) \in \mathcal{I}_{n-1}^{\mathrm{QM}}.$$

Proof. This follows immediately from the integration by parts formula (2-9). □

$\mathcal{I}^{\mathrm{QM}}$ as a shuffle algebra. We let V be the $\mathbb{C}$-vector space $\mathbb{C} \cdot E_2 \oplus M_*$, and denote by $\mathbb{C}\langle V\rangle$ the shuffle algebra on V (see Section 2). Recall that this is the graded dual of the tensor algebra $T(V)$, whose grading is given by the length of tensors. Elements of $\mathbb{C}\langle V\rangle$ are $\mathbb{C}$-linear combination of the basic elements $[f_1 \mid \cdots \mid f_n]$, and the product on $\mathbb{C}\langle V\rangle$ is the shuffle product (2-5).

The following theorem is the main result of this paper.

Theorem 4.3. *The QM_*-linear map*

$$\varphi^{\mathrm{QM}} : \mathrm{QM}_* \otimes_{\mathbb{C}} \mathbb{C}\langle V\rangle \to \mathcal{I}^{\mathrm{QM}}, \qquad [f_1 \mid \cdots \mid f_n] \mapsto I(f_1, \dots, f_n; \tau) \tag{4-1}$$

is an isomorphism of QM_*-algebras.

Proof. Let $\mathcal{B}$ be a basis of V, so that the family $([f_1 \mid \cdots \mid f_n] \mid f_i \in \mathcal{B})$ is a basis of $\mathbb{C}\langle V\rangle$. The injectivity of φ^{QM} follows from the $\mathrm{Frac}(\mathrm{QM}_*)$-linear independence of the family

$$\mathcal{F} = (I(f_1, \dots, f_n; \tau) \mid f_i \in \mathcal{B}), \tag{4-2}$$

which is a consequence of Theorem 3.1.

To obtain the surjectivity, we need to prove that the family (4-2) generates $\mathcal{I}^{\mathrm{QM}}$. To this end, we prove inductively that for every $n \geq 0$, we have $\mathcal{I}_n^{\mathrm{QM}} \subset \mathrm{Span}_{\mathrm{QM}_*} \mathcal{F}$. The case $n = 0$ is trivial. Now let $n \geq 1$ and assume that for every $r \leq n - 1$, we have $\mathcal{I}_r^{\mathrm{QM}} \subset \mathrm{Span}_{\mathrm{QM}_*} \mathcal{F}$. Given quasimodular forms $f_1, \dots, f_n$, we can write $f_i = g_i + D(h_i)$, where $g_i \in \mathbb{C} \cdot E_2 \oplus M_*$ and $h_i \in D(\mathrm{QM}_*)$ by Proposition 2.3(ii). Then by linearity,

$$I(f_1, \dots, f_n; \tau) = I(g_1, \dots, g_n; \tau) + \sum_{i=1}^{n} I(g_1, \dots, g_{i-1}, D(h_i), g_{i+1}, \dots, g_n) + \cdots, \tag{4-3}$$

where the $\cdots$ above signifies iterated integrals which have at least two $D(h_i)$ as integrands. The first term on the right is contained in $\mathrm{Span}_{\mathrm{QM}_*}\mathcal{F}$, since $g_i \in \mathbb{C}\cdot E_2 \oplus M_*$ for every i and $\mathcal{B}$ is a basis. On the other hand, all other terms in the sum (4-3) are iterated integrals which contain at least one $D(h_i)$. By Proposition 4.2, it thus follows that

$$I(f_1,\dots,f_n;\tau) \equiv I(g_1,\dots,g_n;\tau) \mod \mathcal{I}^{\mathrm{QM}}_{n-1},$$

and we conclude using the induction hypothesis. Finally, it is clear that φ^{QM} is a homomorphism of algebras, since both sides of (4-1) are endowed with the shuffle product. □

The algebra of iterated integrals of modular forms. In this section, we study the subalgebra $\mathcal{I}^M$ of $\mathcal{I}^{\mathrm{QM}}$, generated by iterated integrals of modular forms.

Definition 4.4. Define $\mathcal{I}^M$ to be the M_*-module generated by all iterated integrals of modular forms:

$$\mathcal{I}^M = \mathrm{Span}_{M_*}\{I(f_1,\dots,f_n;\tau) \mid f_i \in M_*\}.$$

As in the case of $\mathcal{I}^{\mathrm{QM}}$, the length of iterated integrals defines the length filtration $\mathcal{I}^M_\bullet$ on $\mathcal{I}^M$, and $\mathcal{I}^M$ is a filtered M_*-subalgebra of $\mathcal{I}^{\mathrm{QM}}$. We let $\mathbb{C}\langle M_*\rangle$ be the shuffle algebra on the $\mathbb{C}$-vector space M_*.

Theorem 4.5. *The M_*-linear map*

$$\varphi^M : M_* \otimes_{\mathbb{C}} \mathbb{C}\langle M_*\rangle \to \mathcal{I}^M, \qquad [f_1 \mid \cdots \mid f_n] \mapsto I(f_1,\dots,f_n;\tau)$$

is an isomorphism of M_*-algebras.

Proof. The morphism φ^M is surjective by definition. It is also injective, since for a basis $\mathcal{B}_M$ of M_*, the iterated integrals $I(f_1,\dots,f_n;\tau)$ with $f_i \in \mathcal{B}_M$ are linearly independent over M_* by Theorem 3.1, as $M_* \subset \mathrm{Frac}(\mathrm{QM}_*)$. □

A polynomial basis for $\mathcal{I}^{\mathrm{QM}}$. Recall from Proposition 2.3(ii) that QM_* is isomorphic to the polynomial algebra $\mathbb{C}[E_2, E_4, E_6]$. A similar, but slightly more involved statement holds for the QM_*-algebra $\mathcal{I}^{\mathrm{QM}}$ of iterated integrals of quasimodular forms. Namely, $\mathcal{I}^{\mathrm{QM}}$ is a polynomial algebra over QM_* in infinitely many variables, which are given by certain Lyndon words.

In the following, if $(S,<)$ is a totally ordered set, we will endow the free monoid S^* on S with the lexicographical order induced by $<$. Also, the *length* of w is simply the number of letters of w.

Definition 4.6. A *Lyndon word* on S^* is a nontrivial word $w \in S^* \setminus \{1\}$ such that for all factorizations $w = uv$ with $u, v \neq 1$, we have $w < v$. We denote by $\mathrm{Lyn}(S^*)$ the set of all Lyndon words on S^*.

Example 4.7. Let $S = \{a, b\}$ with total order $a < b$. Then the Lyndon words on S^* of length at most four are

$$a,\ b,\ ab,\ aab,\ abb,\ aaab,\ aabb,\ abbb.$$

Now for a field k and any set S, define $k\langle S\rangle$ to be the shuffle algebra on the free k-vector space generated by S. If k is of characteristic zero, then by the Milnor–Moore theorem [Milnor and Moore 1965], $k\langle S\rangle$ is isomorphic to a polynomial algebra (in possibly infinitely many variables). The following refinement is due to Radford.

Theorem 4.8 [Radford 1979]. *If k has characteristic zero, then $k\langle S\rangle$ is freely generated, as a k-algebra, by the set of Lyndon words* $\mathrm{Lyn}(S^*)$. *Equivalently, $k\langle S\rangle \cong k[\mathrm{Lyn}(S^*)]$, the polynomial algebra on* $\mathrm{Lyn}(S^*)$.

Returning to quasimodular forms, consider again the $\mathbb{C}$-vector space

$$V = \mathbb{C}\cdot E_2 \oplus M_*,$$

and let $\mathcal{B} = \bigcup_{k\geq 0} \mathcal{B}_k$ be the homogeneous basis of V given by $\mathcal{B}_k = \{E_4^a E_6^b \mid 4a + 6b = k\}$ for $k \neq 2$, and $\mathcal{B}_2 = \{E_2\}$. The basis $\mathcal{B}$ can be ordered for the lexicographical order as follows: if $E_4^a E_6^b, E_4^{a'} E_6^{b'} \in \mathcal{B}_k$, then

$$E_4^a E_6^b < E_4^{a'} E_6^{b'} :\Leftrightarrow a < a',\ \text{or } a = a' \text{ and } b < b',$$

and if $f \in \mathcal{B}_k$, $g \in \mathcal{B}_{k'}$ with $k < k'$, then $f < g$.

Now, since for $f_1, \dots, f_n \in \mathcal{B}$, the iterated integrals $I(f_1, \dots, f_n; \tau)$ are linearly independent over QM_* (by Theorem 3.1), we can canonically identify the set of all $I(f_1, \dots, f_n; \tau)$ with the free monoid $\mathcal{B}^*$ and order $\mathcal{B}^*$ for the lexicographical ordering induced from the order on $\mathcal{B}$ above. The next result is a formal consequence of Theorems 4.3, 4.5 and 4.8.

Theorem 4.9. *The elements of* $\mathrm{Lyn}(\mathcal{B}^*)$ *are algebraically independent over* QM_* *and we have a natural isomorphism of* QM_**-algebras*

$$\mathrm{QM}_*[\mathrm{Lyn}(\mathcal{B}^*)] \cong \mathcal{I}^{\mathrm{QM}},$$

which is filtered for the length, where the left-hand side is the polynomial QM_**-algebra on* $\mathrm{Lyn}(\mathcal{B}^*)$. *Explicitly, the isomorphism maps an element*

$$w = f_1 \cdots f_n \in \mathrm{Lyn}(\mathcal{B}^*)$$

to the iterated integral $I(f_1, \dots, f_n; \tau)$. Similarly, we have a natural isomorphism of M_*-algebras

$$M_*[\mathrm{Lyn}(\mathcal{B}_M^*)] \cong \mathcal{I}^M,$$

where $\mathcal{B}_M = \mathcal{B} \setminus \{E_2\}$.

Example 4.10. The following table gives all elements of $\mathrm{Lyn}(\mathcal{B}^*)$ involving iterated integrals of length at most two of quasimodular forms of total weight at most 12. For ease of notation, we have dropped the τ from $I(f_1,\dots,f_n;\tau)$.

	Length		
Weight	0	1	2
0	—	$I(1)$	—
2	—	$I(E_2)$	—
4	—	$I(E_4)$	$I(1,E_4)$
6	—	$I(E_6)$	$I(1,E_6), I(E_2,E_4)$
8	—	$I(E_4^2)$	$I(1,E_4^2), I(E_2,E_6)$
10	—	$I(E_4E_6)$	$I(1,E_4E_6), I(E_2,E_4^2), I(E_4,E_6)$
12	—	$I(E_4^3), I(E_6^2)$	$I(1,E_4^3), I(1,E_6^2), I(E_2,E_4E_6), I(E_4,E_4^2)$

Also, the list of all elements of $\mathrm{Lyn}(\mathcal{B}^*)$ consisting of iterated integrals of length at most three of quasimodular forms of total weight 12 is given by

$$\begin{aligned}&\{I(E_4^3),\ I(E_6^2),\ I(1,E_4^3),\ I(1,E_6^2),\ I(E_2,E_4E_6),\ I(E_4,E_4^2),\\ &\ I(1,1,E_4^3),\ I(1,1,E_6^2),\ I(1,E_2,E_4E_6),\ I(1,E_4,E_4^2),\ I(1,E_6,E_6),\\ &\ I(1,E_4^2,E_4),\ I(1,E_4E_6,E_2),\ I(E_2,E_2,E_4^2),\ I(E_2,E_4,E_6),\ I(E_2,E_6,E_4)\}.\end{aligned}$$

Appendix: Eichler–Shimura for quasimodular forms

In this appendix, we show how one can attach one-cocycles to quasimodular forms. This extends the classical Eichler–Shimura theory of the cocycles attached to modular forms, and is probably well-known to the experts, but the author does not know of a suitable reference for the precise statements.

Throughout this appendix, we will freely use some elementary concepts from the cohomology of groups, for which we refer to [Weibel 1994, Chapter 6].

Cocycles attached to modular forms. We first briefly recall how modular forms give rise to cocycles for $\mathrm{SL}_2(\mathbb{Z})$. A standard reference is [Lang 1976, Chapter VI].

For $d \geq 0$, let $\mathbb{Q}[X,Y]_d$ be the $\mathbb{Q}$-vector space of homogeneous polynomials in X and Y of degree d. It is a right $\mathrm{SL}_2(\mathbb{Z})$-module by defining

$$P(X,Y)|_\gamma = P(aX+bY, cX+dY) \quad \text{for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),\ P \in \mathbb{Q}[X,Y]_d.$$

With this action, given a modular form f of weight $k \geq 2$, it is straightforward to verify that the holomorphic differential one-form

$$\underline{f}(\tau) := (2\pi i)^{k-1} f(\tau)(X-\tau Y)^{k-2}\,\mathrm{d}\tau \in \Omega^1(\mathfrak{H}) \otimes_{\mathbb{Q}} \mathbb{Q}[X,Y]_{k-2}$$

is $\mathrm{SL}_2(\mathbb{Z})$-invariant, where $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{H}$ in the usual way via fractional linear transformations. Fixing a base point τ_0 of $\mathfrak{H}$ (possibly $i\infty$), it follows from the $\mathrm{SL}_2(\mathbb{Z})$-invariance that the function

$$r_{f,\tau_0} : \mathrm{SL}_2(\mathbb{Z}) \to \mathbb{C}[X,Y]_{k-2}, \qquad \gamma \mapsto \int_{\tau}^{\tau_0} \underline{f}(\tau) - \left(\int_{\gamma.\tau}^{\tau_0} \underline{f}(\tau)\right)\Big|_{\gamma}$$

(regularized as in Section 2 if $\tau_0 = i\infty$) is a one-cocycle, i.e., it satisfies

$$r_{f,\tau_0}(\gamma_1\gamma_2) = r_{f,\tau_0}(\gamma_1)|_{\gamma_2} + r_{f,\tau_0}(\gamma_2)$$

for all $\gamma_1, \gamma_2 \in \mathrm{SL}_2(\mathbb{Z})$. Its cohomology class does not depend on τ_0, and we denote this class simply by $[r_f]$.

The same construction can also be applied to the complex conjugate

$$\overline{\underline{f}(\tau)} := (-2\pi i)^{k-1}\,\overline{f(\tau)}(X - \bar{\tau}Y)^{k-2}\,\mathrm{d}\bar{\tau}$$

of the one-form $\underline{f}(\tau)$, and we denote by $[r_{\bar{f}}]$ the resulting cohomology class.

Theorem A.1 (Eichler–Shimura). *For every $k \geq 2$, the morphism*

$$M_k \oplus \overline{S}_k \to H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X,Y]_{k-2}) \otimes_{\mathbb{Q}} \mathbb{C}, \qquad (f, \bar{g}) \mapsto [r_f] + [r_{\bar{g}}]$$

is an isomorphism of $\mathbb{C}$-vector spaces. Here, $\overline{S}_k$ denotes the complex conjugate of the $\mathbb{C}$-vector space of cusp forms of weight k.

Cocycles for the braid group. The fact that r_f is a cocycle hinges on the modularity of f. In order to incorporate quasimodular forms into the picture, we need to consider instead of $\mathrm{SL}_2(\mathbb{Z})$ the braid group $B_3 = \langle \sigma_1, \sigma_2 : \sigma_1\sigma_2\sigma_1 = \sigma_2\sigma_1\sigma_2 \rangle$ on three strands. It is a central extension

$$1 \to \mathbb{Z} \to B_3 \to \mathrm{SL}_2(\mathbb{Z}) \to 1, \tag{A-1}$$

and also the fundamental group of the quotient of $\mathbb{C}^\times \times \mathfrak{H}$ by the $\mathrm{SL}_2(\mathbb{Z})$-action

$$\gamma.(z,\tau) = ((c\tau + d)z, \gamma.\tau) \quad \text{for } \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(\mathbb{Z}),$$

where $\mathrm{SL}_2(\mathbb{Z})$ acts on $\mathfrak{H}$ as before. We refer to [Hain 2011, §8] for more details and further equivalent descriptions of B_3.

Next, we compute the cohomology groups $H^1(B_3, \mathbb{Q}[X,Y]_d)$, where B_3 acts on $\mathbb{Q}[X,Y]_d$ via the projection $B_3 \to \mathrm{SL}_2(\mathbb{Z})$.

Proposition A.2. *We have canonical isomorphisms*

$$H^1(B_3, \mathbb{Q}[X,Y]_d) \cong \begin{cases} H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X,Y]_d) & \text{for } d \geq 1, \\ \mathbb{Q} & \text{for } d = 0. \end{cases}$$

Proof. The Hochschild–Serre spectral sequence (see [Weibel 1994, §6.8.3]) associated to the extension (A-1) yields an exact sequence

$$0 \to H^1(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X,Y]_d) \to H^1(B_3, \mathbb{Q}[X,Y]_d) \to H^1(\mathbb{Z}, \mathbb{Q}[X,Y]_d)^{\mathrm{SL}_2(\mathbb{Z})} \to 0,$$

where we have used the fact that $H^2(\mathrm{SL}_2(\mathbb{Z}), \mathbb{Q}[X,Y]_d) = \{0\}$, as $\mathrm{SL}_2(\mathbb{Z})$ has virtual cohomological dimension equal to one. The proposition now follows easily from this. $\square$

Quasimodular forms and braid group cocycles. In light of Theorem A.1, the isomorphisms of Proposition A.2 suggest attaching a one-cocycle $B_3 \to \mathbb{C}$ to the Eisenstein series E_2. Indeed, this can be done as follows.

First, the modular transformation property of E_2 (2-2) implies that the differential one-form

$$2\pi i E_2(\tau)\,\mathrm{d}\tau - 12\frac{\mathrm{d}z}{z} \in \Omega^1(\mathbb{C}^\times \times \mathfrak{H}) \tag{A-2}$$

is $\mathrm{SL}_2(\mathbb{Z})$-invariant, i.e., it descends to the quotient $\mathrm{SL}_2(\mathbb{Z})\backslash(\mathbb{C}^\times \times \mathfrak{H})$. Denote by

$$\underline{E_2}(\xi,\tau) := \varphi^*\Bigl(2\pi i E_2(\tau)\,\mathrm{d}\tau - 12\frac{\mathrm{d}z}{z}\Bigr) = 2\pi i E_2(\tau)\,\mathrm{d}\tau - 12\,\mathrm{d}\xi \in \Omega^1(\mathbb{C} \times \mathfrak{H})$$

the pull-back of (A-2) along the universal covering map $\varphi : \mathbb{C}\times\mathfrak{H} \to \mathrm{SL}_2(\mathbb{Z})\backslash(\mathbb{C}^\times\times\mathfrak{H})$. Clearly, $\underline{E_2}(\xi,\tau)$ is B_3-invariant and it follows that for any base point (ξ_0,τ_0) (for example, $(\xi_0,\tau_0) = (0, i\infty)$), the function

$$r_{E_2,(\xi_0,\tau_0)} : B_3 \to \mathbb{C}, \qquad \gamma \mapsto \int_{(\xi,\tau)}^{(\xi_0,\tau_0)} \underline{E_2}(\xi,\tau) - \biggl(\int_{\gamma.(\xi,\tau)}^{(\xi_0,\tau_0)} \underline{E_2}(\xi,\tau)\biggr)\bigg|_\gamma$$

is a well-defined cocycle (again, regularization is needed if $\tau_0 = i\infty$).

Remark A.3. The integral $I(E_2;\tau)$ introduced in Section 2 is actually equal to $\int_\tau^{i\infty} \underline{E_2}(\xi,\tau)$, where we embed $\mathfrak{H}$ into $\mathbb{C}\times\mathfrak{H}$ by $\tau \mapsto (0,\tau)$. However, that embedding is not B_3-equivariant, and indeed the integral $I(E_2;\tau)$ does not give rise to a cocycle for B_3; for this, one really needs to lift the form $2\pi i E_2(\tau)\,\mathrm{d}\tau$ to the form $\underline{E_2}(\xi,\tau)$.

Now, since the cocycle $r_{E_2,(\xi_0,\tau_0)}$ is nonzero, its cohomology class (which is again independent of the choice of base point (ξ_0,τ_0)) is nontrivial. The Eichler–Shimura theorem (Theorem A.1) together with Proposition A.2 then implies the next result.

Corollary A.4. *For every $k \geq 2$, the morphism*

$$V_k \oplus \overline{S}_k \to H^1(B_3, \mathbb{Q}[X,Y]_{k-2}) \otimes_{\mathbb{Q}} \mathbb{C}, \qquad (f,\bar{g}) \mapsto [r_f] + [r_{\bar{g}}],$$

where $V := M_* \oplus \mathbb{C}\cdot E_2$, is an isomorphism of $\mathbb{C}$-vector spaces.

One can also attach a cocycle r_{f,τ_0} to a general quasimodular form $f \in \mathrm{QM}_k$ of weight k as follows. By Proposition 2.3(ii), we know that f can be written uniquely as a $\mathbb{C}$-linear combination of derivatives of modular forms and of derivatives of E_2. Thus, we can write

$$f = \sum \lambda_g \cdot D^{p_g}(g), \quad \lambda_g \in \mathbb{C},\ p_g \geq 0,$$

where either g is a modular form of weight $k-2p_g$ or $g = E_2$. Therefore, we may define $r_{f,\tau_0} : B_3 \to \mathbb{C}[X, Y]_{\leq k-2} := \bigoplus_{0\leq d\leq k-2} \mathbb{C}[X, Y]_d$ by

$$r_{f,\tau_0} := \sum \lambda_g \cdot r_{g,\tau_0}.$$

Using this definition, one sees in particular that the cocycles of quasimodular forms can be expressed in terms of the cocycles attached to modular forms and to E_2. This is of course in line with Corollary A.4.

Remark A.5. In [Manin 2006; Brown 2016; Hain 2016], certain nonabelian $\mathrm{SL}_2(\mathbb{Z})$-cocycles given in terms of iterated integrals of modular forms are studied. It would be natural to try and extend this theory to nonabelian B_3-cocycles attached to iterated integrals of quasimodular forms (perhaps along the lines suggested in [Hain 2016, §14]), but this is beyond the scope of the present paper.

Acknowledgments

Very many thanks to Pierre Lochak for bringing the article [Deneufchâtel et al. 2011] to the author's attention. Also, many thanks to Francis Brown, Erik Panzer and the referees for corrections as well as very helpful suggestions and to Don Zagier for inspiring discussions on the Appendix. The results of this paper were found while the author was a Ph.D. student at Universität Hamburg under the supervision of Ulf Kühn.

References

[Bloch and Okounkov 2000] S. Bloch and A. Okounkov, "The character of the infinite wedge representation", *Adv. Math.* **149**:1 (2000), 1–60. MR Zbl

[Brown 2009] F. C. S. Brown, "Multiple zeta values and periods of moduli spaces $\overline{\mathfrak{M}}_{0,n}$", *Ann. Sci. Éc. Norm. Supér.* (4) **42**:3 (2009), 371–489. MR Zbl

[Brown 2016] F. Brown, "Multiple modular values and the relative completion of the fundamental group of $M_{1,1}$", preprint, 2016. arXiv

[Brown 2017] F. Brown, "A class of non-holomorphic modular forms, II: Equivariant iterated Eisenstein integrals", preprint, 2017. arXiv

[Deneufchâtel et al. 2011] M. Deneufchâtel, G. H. E. Duchamp, V. H. N. Minh, and A. I. Solomon, "Independence of hyperlogarithms over function fields via algebraic combinatorics", pp. 127–139 in *Algebraic informatics* (Linz, Austria, 2011), edited by F. Winkler, Lecture Notes in Comput. Sci. **6742**, Springer, 2011. MR Zbl

[Eichler 1957] M. Eichler, "Eine Verallgemeinerung der Abelschen Integrale", *Math. Z.* **67** (1957), 267–298. MR Zbl

[Hain 1987] R. M. Hain, "The geometry of the mixed Hodge structure on the fundamental group", pp. 247–282 in *Algebraic geometry* (Brunswick, ME, 1985), edited by S. J. Bloch, Proceedings of Symposia in Pure Math. **46**, American Mathematical Society, Providence, RI, 1987. MR Zbl

[Hain 2011] R. Hain, "Lectures on moduli spaces of elliptic curves", pp. 95–166 in *Transformation groups and moduli spaces of curves* (Hangzhou, China), edited by L. Ji and S.-T. Yau, Advanced Lectures in Math. **16**, International Press, Somerville, MA, 2011. MR Zbl

[Hain 2016] R. Hain, "The Hodge–de Rham theory of modular groups", pp. 422–514 in *Recent advances in Hodge theory* (Vancouver, 2013), edited by M. Kerr and G. Pearlstein, London Math. Society Lecture Note Series **427**, Cambridge University Press, 2016. MR Zbl

[Kaneko and Zagier 1995] M. Kaneko and D. Zagier, "A generalized Jacobi theta function and quasimodular forms", pp. 165–172 in *The moduli space of curves* (Texel, Netherlands, 1994), edited by R. Dijkgraaf et al., Progress in Math. **129**, Birkhäuser, Boston, 1995. MR Zbl

[Lang 1976] S. Lang, *Introduction to modular forms*, Grundlehren der math. Wissenschaften **222**, Springer, 1976. MR Zbl

[Lochak et al. 2017] P. Lochak, N. Matthes, and L. Schneps, "Elliptic multiple zeta values and the elliptic double shuffle relations", preprint, 2017. arXiv

[Manin 2006] Y. I. Manin, "Iterated integrals of modular forms and noncommutative modular symbols", pp. 565–597 in *Algebraic geometry and number theory*, edited by V. Ginzburg, Progress in Math. **253**, Birkhäuser, Boston, 2006. MR Zbl

[Martin and Royer 2005] F. Martin and E. Royer, "Formes modulaires et périodes", pp. 1–117 in *Formes modulaires et transcendance* (Marseille, 2003), edited by S. Fischler et al., Sémin. Congr. **12**, Société Mathématique de France, Paris, 2005. MR Zbl

[Milnor and Moore 1965] J. W. Milnor and J. C. Moore, "On the structure of Hopf algebras", *Ann. of Math.* (2) **81** (1965), 211–264. MR Zbl

[Radford 1979] D. E. Radford, "A natural ring basis for the shuffle algebra and an application to group schemes", *J. Algebra* **58**:2 (1979), 432–454. MR Zbl

[Reutenauer 1993] C. Reutenauer, *Free Lie algebras*, London Math. Society Monographs New Series **7**, Oxford University Press, 1993. MR Zbl

[Weibel 1994] C. A. Weibel, *An introduction to homological algebra*, Cambridge Studies in Advanced Math. **38**, Cambridge University Press, 1994. MR Zbl

[Zagier 2008] D. Zagier, "Elliptic modular forms and their applications", pp. 1–103 in *The 1-2-3 of modular forms* (Nordfjordeid, Norway, 2004), edited by K. Ranestad, Springer, 2008. MR Zbl

Communicated by Yuri Manin
Received 2016-11-08 Revised 2017-06-15 Accepted 2017-09-08

nilsmath@mpim-bonn.mpg.de *Max-Planck-Institut für Mathematik, Bonn, Germany*

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2131

On the density of zeros of linear combinations of Euler products for $\sigma > 1$

Mattia Righetti

It has been conjectured by Bombieri and Ghosh that the real parts of the zeros of a linear combination of two or more L-functions should be dense in the interval $[1, \sigma^*]$, where σ^* is the least upper bound of the real parts of such zeros. In this paper we show that this is not true in general. Moreover, we describe the optimal configuration of the zeros of linear combinations of orthogonal Euler products by showing that the real parts of such zeros are dense in subintervals of $[1, \sigma^*]$ whenever $\sigma^* > 1$.

1. Introduction

Let $L(s)$ be a Dirichlet series and let $\sigma^* = \sigma^*(L)$ be the least upper bound of the real parts of the zeros of $L(s)$. It is well known that σ^* is finite (see, e.g., Titchmarsh [1975, §9.41]). For the Riemann zeta function we know that $\sigma^* \leq 1$, and it is expected that the Riemann hypothesis holds, i.e., $\sigma^* = \frac{1}{2}$. A similar situation is expected for many Euler products (see, e.g., Selberg [1992]).

On the other hand, we have recently proved [Righetti 2016a], for a large class of L-functions with a polynomial Euler product, that nontrivial linear combinations have zeros for $\sigma > 1$. This is not surprising since many examples of such linear combinations were already known to have zeros for $\sigma > 1$ from work of Davenport and Heilbronn [1936a; 1936b] on the Hurwitz and Epstein zeta functions. We also refer to later important works of Cassels [1961], Conrey and Ghosh [1994], Saias and Weingartner [2009], and Booker and Thorne [2014].

Since for this type of Dirichlet series we know that there are zeros in the region of absolute convergence, which we may always suppose to be $\sigma > 1$, it is of interest to know the distribution of such zeros in this half-plane. With respect to the distribution of the imaginary parts the problem was completely solved by Jessen and Tornehave [1945]. Indeed it is known that the number of zeros in any rectangle

MSC2010: primary 11M41; secondary 11M26.
Keywords: zeros of Dirichlet series, value distribution, asymptotic distribution functions, convexity.

$[\sigma_1, \sigma_2] \times [T_1, T_2]$, with $1 < \sigma_1 < \sigma_2$, satisfies (cf. Theorem 31 of [Jessen and Tornehave 1945])

$$N(\sigma_1, \sigma_2, T_1, T_2) = c(T_2 - T_1) + o(|T_2 - T_1|), \quad \text{when } |T_2 - T_1| \to \infty, \tag{1-1}$$

for some nonnegative constant $c = c(\sigma_1, \sigma_2)$. Note that by a classical application of the Bohr almost periodicity of Dirichlet series and Rouché's theorem we easily have that $c > 0$ whenever $N(\sigma_1, \sigma_2, T_1, T_2) > 0$.

On the other hand the situation regarding the distribution of the real parts of the zeros is much more complicated. In fact some Epstein zeta functions studied by Bombieri and Mueller [2008] are known to have the property that the real parts of their zeros are dense in the interval $[1, \sigma^*]$. Note that these functions may be written as a linear combination of two Hecke L-functions. Other examples of linear combinations with this property may be found in Bombieri and Ghosh [2011], although not explicitly stated. Moreover, we remarked in [Righetti 2016a] that, as a consequence of the technique used to prove the main result there, the real parts of the zeros of nontrivial combinations of orthogonal L-functions are dense in a small interval $[1, 1 + \eta]$, for some $\eta > 0$ (cf. Corollary 1 of [Righetti 2016a]). Hence one might expect, as conjectured by Bombieri and Ghosh [2011, p. 230], that the real parts of the zeros of linear combinations of two or more L-functions should be dense in the whole interval $[1, \sigma^*]$. However this is too much to hope for as one can see from the following general counterexample.

Theorem 1.1. *Let $N \geq 2$ be an integer and let $F_j(s) = \sum_{n=1}^{\infty} a_j(n) n^{-s}$ be distinct nonidentically zero Dirichlet series absolutely convergent for $\sigma > 1$, $j = 1, \dots, N$, with $\sum_{j=1}^{N} |a_j(1)| \neq 0$. Then, for any $\boldsymbol{x} = (x_1, \dots, x_N) \in \mathbb{C}^N$ such that $\sum_{j=1}^{N} x_j a_j(1) = 0$ but the Dirichlet series $L_{\boldsymbol{x}}(s) = \sum_{j=1}^{N} x_j F_j(s)$ is not identically zero, there exist infinitely many projectively inequivalent vectors $\boldsymbol{c} \in \mathbb{C}^N$ such that $L_{\boldsymbol{c}}(s)$ has no zeros in some vertical strip $\sigma_1 < \sigma < \sigma_2$ with $1 < \sigma_1 < \sigma_2 < \sigma^*(L_{\boldsymbol{c}})$.*

Remark. The above statement is very general, but in particular may be applied to linear combinations of linearly independent L-functions. Moreover, it is easy to show that the same argument works also for a-values with $a \neq 0$.

This has to be compared with what happens for $\frac{1}{2} < \sigma < 1$. There it is known that *joint universality* of L-functions implies that the real parts of the zeros of any linear combination of these L-functions are dense in $\left[\frac{1}{2}, 1\right]$ (see, e.g., [Bombieri and Gosh 2011, p. 230]). Furthermore joint universality is known to hold for many families of L-functions and recently Lee, Nakamura and Pańkowski [Lee et al. 2017] have shown that this property holds in an axiomatic setting such as the Selberg class under a strong Selberg orthonormality conjecture.

We can actually prove more, i.e., it is in general possible to construct Dirichlet series, given by a linear combination of L-functions, which have many *distinct* vertical strips without zeros, i.e., such that between every two vertical strips without zeros there is at least one zero.

Theorem 1.2. *Let $k \geq 1$ be an integer and, for $j = 1, \dots, k+1$, let $F_j(s) = \sum_{n=1}^{\infty} a_j(n) n^{-s}$ be a Dirichlet series absolutely convergent for $\sigma > 1$ with $a_j(1) \neq 0$. Suppose that*

$$\det \begin{pmatrix} a_1(1) & a_1(2) & \cdots & a_1(k+1) \\ a_2(1) & a_2(2) & \cdots & a_2(k+1) \\ \vdots & \vdots & \ddots & \vdots \\ a_{k+1}(1) & a_{k+1}(2) & \cdots & a_{k+1}(k+1) \end{pmatrix} \neq 0. \tag{1-2}$$

Then there exists at least one $\boldsymbol{c} \in \mathbb{C}^{k+1}$ such that the Dirichlet series $L_{\boldsymbol{c}}(s) = \sum_{j=1}^{k+1} c_j F_j(s)$ has at least k distinct vertical strips without zeros in the region $1 < \sigma < \sigma^*(L_{\boldsymbol{c}})$.

Remark. Note that trivially every nonzero scalar multiple of a vector $\boldsymbol{c}$ of Theorems 1.1 or 1.2 has the same property. On the other hand, in Theorem 1.1, for every $\boldsymbol{x}$ the vectors $\boldsymbol{c}$ are given by the intersection of a ball and a hyperplane in $\mathbb{C}^N$, hence there are clearly infinitely many projectively inequivalent such vectors; see Section 6 for details. Besides, the proof of Theorem 1.2 seems to suggest that there may actually be infinitely many projectively inequivalent vectors $\boldsymbol{c}$ with the same property in this case too.

The proof of Theorem 1.2 is actually constructive and may be used to explicitly obtain coefficients $\boldsymbol{c}$. As a concrete example we apply it to $\zeta(s)$, $L(s, \chi_1)$ and $L(s, \overline{\chi_1})$, where χ_1 is the unique Dirichlet character mod 5 such that $\chi_1(2) = i$, which satisfy the hypotheses of Theorem 1.2. We thus obtain the Dirichlet series

$$L(s) = c_1 L(s, \chi_1) + c_2 L(s, \overline{\chi_1}) + c_3 \zeta(s),$$

where

$$\begin{aligned} c_1 &= -\frac{1}{L(8, \overline{\chi_1})} \frac{L(16, \overline{\chi_1})\zeta(8) - L(8, \overline{\chi_1})\zeta(16)}{L(16, \chi_1)\zeta(8) - L(8, \chi_1)\zeta(16)} \\ &\qquad = -0.08260584\ldots - i0.99658995\ldots, \\ c_2 &= \frac{1}{L(8, \overline{\chi_1})} \qquad = \ \ 1.00000059\ldots + i0.00375400\ldots, \\ c_3 &= \frac{1}{L(8, \overline{\chi_1})} \frac{L(8, \chi_1)L(16, \overline{\chi_1}) - L(8, \overline{\chi_1})L(16, \chi_1)}{\zeta(8)L(16, \chi_1) - L(8, \chi_1)\zeta(16)} \\ &\qquad = -0.91739597\ldots + i0.99283727\ldots. \end{aligned}$$

In Figure 1 we see part of two distinct vertical strips without zeros of $L(s)$ within

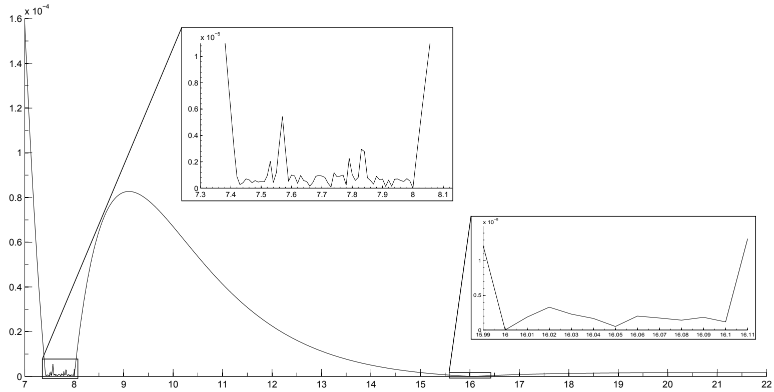

Figure 1. Approximate plot of

$$\min_t \left| c_1 L(\sigma + it, \chi_1) + c_2 L(\sigma + it, \overline{\chi_1}) + c_3 \zeta(\sigma + it) \right|$$

for $\sigma \in [7, 22]$ and $t \in [0, 2000]$ with step 0.01.

the vertical strip $1 < \sigma < \sigma^*$. We recall that, by [Saias and Weingartner 2009], there are zeros in the vertical strip $1 < \sigma < 1 + \eta$ for some $\eta > 0$.

Actually Figure 1 shows that another interesting phenomenon happens for linear combinations of *orthogonal* (see (1-3)) L-functions: it looks like that whenever there is one zero then there should be a small closed interval, either around or beside its real part, where the real parts of the zeros are dense. The bulk of this paper is devoted to showing that this is indeed true.

We first recall that, as a consequence of the work of Jessen and Tornehave [1945] on the asymptotic number of zeros mentioned above, we have the following general result. We denote by $\sigma_u(L)$ the abscissa of uniform convergence of $L(s)$.

Theorem 1.3. *Suppose $L(s) = \sum_{n=n_0}^{\infty} a(n)/n^s$ has $a(n_0) \neq 0$ and $\sigma^*(L) > \sigma_u(L)$. Then in any vertical strip $\sigma_u(L) < \alpha \le \sigma \le \sigma^*(L)$, $L(s)$ has only a finite number of zero-free vertical strips and a finite number of isolated vertical lines containing zeros. In particular, if $\rho_0 = \beta_0 + i\gamma_0$ is a zero of $L(s)$ with $\beta_0 > \sigma_u(L)$, then either $\sigma = \beta_0$ is an isolated vertical line as above or there exist $\sigma_1 \le \beta_0 \le \sigma_2$, with $\sigma_1 < \sigma_2$, such that the set*

$$\left\{ \beta \in [\sigma_1, \sigma_2] \mid \exists \gamma \in \mathbb{R} \text{ such that } L(\beta + i\gamma) = 0 \right\}$$

is dense in $[\sigma_1, \sigma_2]$.

The first part of Theorem 1.3 is a reinterpretation of Theorem 31 of [Jessen and Tornehave 1945] in view of Theorem 8 of the same paper. The second part follows from the first one by a simple set-theoretic argument.

Therefore we just need to prove that a linear combination of orthogonal Euler products has no isolated vertical lines containing zeros. As in [Righetti 2016a] we work in an axiomatic setting, and at the end of the introduction we briefly mention some important families of L-functions satisfying the required properties. Given a complex function $F(s)$ we consider the following properties:

(I) $F(s) = \sum_{n=1}^{\infty} a_F(n)/n^s$ is absolutely convergent for $\sigma > 1$;

(II) $\log F(s) = \sum_p \sum_{k=1}^{\infty} b_F(p^k)/p^{ks}$ is absolutely convergent for $\sigma > 1$, with $|b_F(p^k)| \ll p^{k\theta}$ for every prime p and every $k \geq 1$, for some $\theta < \frac{1}{2}$;

(III) for any $\varepsilon > 0$, $|a_F(n)| \ll n^{\varepsilon}$ for every $n \geq 1$.

Definition. For any integer $N \geq 1$, we say that $F_1(s), \ldots, F_N(s)$ satisfying (I) and (II) are *orthogonal* if

$$\sum_{p \leq x} \frac{a_{F_i}(p)\overline{a_{F_j}(p)}}{p} = (m_{i,j} + o(1)) \log\log x, \quad x \to \infty, \tag{1-3}$$

with $m_{i,i} > 0$ and $m_{i,j} = 0$ if $i \neq j$.

Remark. There are some differences between the axioms that in [Righetti 2016a] define the class $\mathcal{E}$ and the above axioms (I)–(III), so that in principle we cannot say that the results that we obtained in [Righetti 2016a] may be applied here or *vice versa*. However most of the known families of L-functions satisfy, or are supposed to satisfy, both the axioms of $\mathcal{E}$ and (I)–(III).

We can now state the main theorems. We consider separately the cases $N = 2$ and $N \geq 3$ since they are handled in different ways and yield different results, although the underling idea is the same.

Theorem 1.4. *Let $F_1(s)$, $F_2(s)$ be orthogonal functions satisfying* (I) *and* (II), $c_1, c_2 \in \mathbb{C} \setminus \{0\}$, *and $L(s) = c_1F_1(s) + c_2F_2(s)$. Then $L(s)$ has no isolated vertical lines containing zeros in the half-plane $\sigma > 1$.*

Theorem 1.5. *Suppose $N \geq 3$ is an integer*, $c_1, \ldots, c_N \in \mathbb{C} \setminus \{0\}$, $c \in \mathbb{C}$, *and $F_1(s), \ldots, F_N(s)$ are orthogonal functions satisfying* (I)–(III). *If we write $L(s) = \sum_{j=1}^{N} c_j F_j(s) - c$, then $L(s)$ has no isolated vertical lines containing zeros in the half-plane $\sigma > 1$.*

Theorems 1.4 and 1.5 are obtained by suitably adapting the works of Bohr and Jessen [1930; 1932], Jessen and Wintner [1935], Jessen and Tornehave [1945], Borchsenius and Jessen [1948], and Lee [2014] on the value distribution of Dirichlet series. Note that, however, most of these papers refer to results on particular Dirichlet series in the strip $\frac{1}{2} < \sigma < 1$, while we work in the half-plane $\sigma > 1$ with far more general Dirichlet series. Hence, although the ideas are similar, the results are quite different in nature and technical difficulty. The proofs will be given in Sections 4 and 5 respectively.

Remark. Note that orthogonality is necessary in Theorems 1.4 and 1.5 as is shown by the following simple example

$$(1-2^{-s})\zeta(s)-\tfrac{3}{4}\zeta(s)=\Big(\frac{1}{4}-\frac{1}{2^s}\Big)\zeta(s),$$

which clearly vanishes, in the half-plane of absolute convergence $\sigma>1$, only on the vertical line $\sigma=2$. We mention here that in the proof of Theorems 1.4 and 1.5, roughly speaking, orthogonality is just used to bound particular oscillatory integrals (see the end of Section 2) and therefore to show that certain distribution functions behave "nicely" (see Section 3).

From Theorems 1.4 and 1.5 we obtain the following interesting consequence, which should be compared with Corollary 1 of [Righetti 2016a].

Corollary 1.6. *Let $L(s)$ be as in Theorems 1.4 or 1.5. If $\sigma^*(L)>1$, then there exists $\eta>0$ such that the set*

$$\big\{\beta\in[\sigma^*(L)-\eta,\sigma^*(L)]\mid \exists\gamma \text{ such that } L(\beta+i\gamma)=0\big\}$$

is dense in $[\sigma^*(L)-\eta,\sigma^*(L)]$.

Proof. If $\sigma^*=\sigma^*(L)$ is itself the real part of a zero, the result follows immediately from the second part of Theorem 1.3 and Theorems 1.4 and 1.5, choosing $\eta=\sigma^*-\sigma_1>0$ and $\sigma_2=\sigma^*$. Suppose otherwise that σ^* is not the real part of a zero. Then by definition σ^* is the limit point of the real part of certain zeros of $L(s)$. Note that in general if $L(\sigma+it)\neq 0$, then either for any $\varepsilon>0$ there exist β_ε with $|\sigma-\beta_\varepsilon|<\varepsilon$ and $\gamma_\varepsilon\in\mathbb{R}$ such that $L(\beta_\varepsilon+i\gamma_\varepsilon)=0$, i.e., σ is the limit point of the real part of certain zeros of $L(s)$, or there exists an open interval $(\sigma-\delta,\sigma+\delta)$, for some $\delta>0$, which does not contain any real part of the zeros. Since by Theorem 1.3 the number of zero-free vertical strips in $\sigma^*-\varepsilon<\sigma<\sigma^*$ is finite for every small $\varepsilon>0$, we can take $\eta=\varepsilon$ small enough so that there are none. □

By Theorems 1.1 and 1.2 we see that Theorems 1.4 and 1.5 are optimal, in the sense that without conditions on the coefficients $\boldsymbol{c}$ we cannot expect stronger results on the density of the real parts of the zeros. On the other hand it may be true that one could provide necessary and sufficient conditions on the coefficients of a linear combination of L-functions to guarantee Bombieri and Ghosh's conjecture to hold, but this seems out of reach at the moment. Here we just mention the following example with the Davenport–Heilbronn type L-functions studied by Bombieri and Ghosh [2011]. As we already remarked, Bombieri and Ghosh do not say whether these functions do have the property that the real parts of their zeros are dense in $[1,\sigma^*]$. However, in our Ph.D. thesis [Righetti 2016b] we gave necessary and

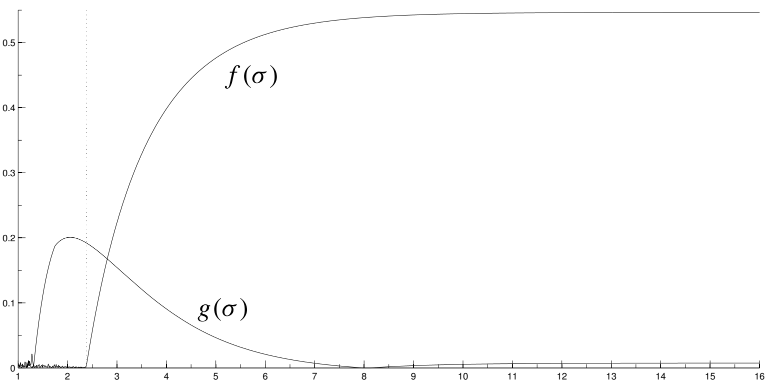

Figure 2. Approximate plot of

$$f(\sigma) = \min_t \left| \frac{L(\sigma+it, \chi_1)}{L(\sigma+it, \overline{\chi_1})} + \frac{1+i\tau}{1-i\tau} \right|,$$

$$g(\sigma) = \min_t \left| \frac{L(\sigma+it, \chi_1)}{L(\sigma+it, \overline{\chi_1})} - \frac{L(8, \chi_1)}{L(8, \overline{\chi_1})} \right|,$$

where $\sigma \in [1.01, 16.01]$ and $t \in [0, 2000]$ with step 0.01.

sufficient conditions on the coefficients of these Dirichlet series for this to happen, namely:

Theorem 1.7. *Let $\xi \in \mathbb{R}$, χ_1 be the unique Dirichlet character mod 5 such that $\chi_1(2) = i$, q be a positive integer and χ_0 be the principal character mod q. Then there exists $\xi_{\max}(q)$, such that the real parts of the zeros for $\sigma > 1$ of*

$$f(s, \xi, q) = \tfrac{1}{2}\big[(1-i\xi)L(s, \chi_1\chi_0) + (1+i\xi)L(s, \overline{\chi_1}\chi_0)\big]$$

are dense in the interval $[1, \sigma^*(\xi, q)]$ if and only if $|\xi| \le \xi_{\max}(q)$. In particular, if $6 \nmid q$ it is sufficient to take $|\xi| \le 6.5851599$.

Proof. The proof is a continuation of the proof of Theorem 7 of [Bombieri and Gosh 2011] using results of Kershner [1936, Theorems II–III] on the support function of the inner border of the sum of convex curves. We refer to Theorem 4.1.3 of [Righetti 2016b] for details. □

As an example we see in Figure 2 that the real parts of the zeros of Davenport–Heilbronn type L-function

$$f(s,\tau) = \tfrac{1}{2}\big[(1-i\tau)L(s, \chi_1) + (1+i\tau)L(s, \overline{\chi_1})\big], \quad \tau = -\frac{1+\sqrt{5}}{2} - \sqrt{1+\Big(\frac{1+\sqrt{5}}{2}\Big)^2},$$

are dense up to $\sigma^* = 2.3822861089\ldots$. On the other hand, we see that the real parts of the zeros of $L(s, \chi_1) - cL(s, \overline{\chi_1})$, where

$$c = \frac{L(8, \chi_1)}{L(8, \overline{\chi_1})} = 0.99997181\ldots + i0.00750790\ldots,$$

are dense close to $\sigma = 1$ (cf. Corollary 1 of [Righetti 2016a]), there are no zeros with real part in the interval [2, 7], but $s = 8$ is clearly a zero.

Note that in the previous results we don't ask for a functional equation or meromorphic continuation to the whole complex plane. However, in many concrete cases these are known to hold, so one might ask what happens if one adds these conditions. On account of this we show that Theorem 1.1 may be modified so that the resulting Dirichlet series is an L-function with functional equation and, of course, without Euler product. We therefore consider functions $F(s)$ satisfying (I) and

(IV) $(s-1)^m F(s)$ is an analytic continuation as an entire function of finite order for some $m \geq 0$,

(V) $F(s)$ satisfies a functional equations of the form $\Phi(s) = \omega\overline{\Phi(1-\bar{s})}$, where $|\omega| = 1$ and

$$\Phi(s) = Q^s \prod_{j=1}^{r} \Gamma(\lambda_j s + \mu_j) F(s) = \gamma(s) F(s),$$

say, with $r \geq 0$, $Q > 0$, $\lambda_j > 0$ and $\operatorname{Re} \mu_j \geq 0$,

although such requirements can actually be relaxed.

Theorem 1.8. *Let $N \geq 3$ be an integer, $(r, Q, \boldsymbol{\lambda}, \boldsymbol{\mu})$ fixed parameters, and let $F_1(s), \ldots, F_N(s)$ be functions satisfying* (I), (II), (IV) *and* (V) *for some $|\omega_j| = 1$, $j = 1, \ldots, N$. Suppose furthermore that $\omega_h \neq \omega_k$ for some $h, k \in \{1, \ldots, N\}$. Then there exist infinitely many $\boldsymbol{c} \in \mathbb{C}^N$ such that $L_{\boldsymbol{c}}(s) = \sum_{j=1}^{N} c_j F_j(s)$ satisfies* (IV), (V) *and has no zeros in some vertical strip $\sigma_1 < \sigma < \sigma_2$ with $1 < \sigma_1 < \sigma_2 < \sigma^*(L_{\boldsymbol{c}})$.*

To give a concrete example of the above result, we fix an integer $q \geq 7$, square-free, $(q, 6) = 1$ and $q \not\equiv 2 \pmod 4$, and consider the Dirichlet L-functions associated with primitive characters $\chi \pmod q$. Their number is $\varphi^*(q) = \prod_{p|q}(p-2)$ and at least half of them have the same *parity*. We denote by $\mathcal{W}(q)$ the set of such characters and we have that $|\mathcal{W}(q)| \geq 3$. As a consequence of Theorem 1 of Kaczorowski, Molteni and Perelli [Kaczorowski et al. 2010], we have that $\omega_{\chi_1} \neq \omega_{\chi_2}$ if $\chi_1 \neq \chi_2$ for $\chi_1, \chi_2 \in \mathcal{W}(q)$, so we may apply Theorem 1.8 to the Dirichlet L-functions associated with distinct characters of $\mathcal{W}(q)$.

On the other hand, we mention that Bombieri and Hejhal [1995] have shown that, under the generalized Riemann hypothesis and a weak pair correlation of the zeros, linear combinations with real coefficients of Euler products with the same functional equation have asymptotically almost all of their zeros on the line $\sigma = \frac{1}{2}$.

As concrete examples of families of L-functions satisfying the properties required by Theorems 1.4 and 1.5 we refer to [Righetti 2016a] for Artin L-functions, automorphic L-functions and the Selberg class. Here we only recall that the relevant analytic properties of the automorphic L-functions and their orthogonality can be found in the papers of Rudnick and Sarnak [1996], Iwaniec and Sarnak [2000], Bombieri and Hejhal [1995], Kaczorowski and Perelli [2000], Kaczorowski, Molteni and Perelli [Kaczorowski et al. 2007], Liu and Ye [2005], and Avdispahić and Smajlović [2010]. Moreover, we refer to Selberg [1992] and the surveys of Kaczorowski [2006], Kaczorowski and Perelli [1999], and Perelli [2005] for a thorough discussion on the Selberg class.

For the computations we have used the software packages PARI/GP [2016] and MATLAB®. These were made by truncating the Dirichlet series to the first 70 000 terms, which guarantees accuracy to eight decimal places for the values given above.

2. Radii of convexity of power series

Let $F(s)$ be a function satisfying (I) and (II). Then we can write $F(s)$ as an absolutely convergent Euler product $F(s) = \prod_p F_p(s)$ for $\sigma > 1$, where the local factor $F_p(s)$ is determined by $\log F_p(s) = \sum_{k=1}^{\infty} b_F(p^k) p^{-ks}$. Then, in most of the results on the value distribution of $F(s)$ for some fixed σ, a fundamental ingredient is the convexity of the curves $\log F_p(\sigma + it)$, $t \in \mathbb{R}$, at least for infinitely many primes p. In this section we collect and prove some results on this matter which will be needed later.

Let $\mathcal{A}$ be the class of functions $f(z) = z + \sum_{n=2}^{\infty} b(n) z^n$ which are regular on $D = \{|z| < 1\}$. Let $\mathcal{F}$ be any subclass of $\mathcal{A}$, then we write $r_c(\mathcal{F})$ for the largest r, with $0 < r \leq 1$, such that $f(\{|z| < r\})$ is convex.

Proposition 2.1 [Yamashita 1982, Theorem 2]. *Let $\mathcal{B} = \{f \in \mathcal{A} \mid |b(n)| \leq n,\ n \geq 2\}$. Then $r_c(\mathcal{B}) \geq R_1$, where R_1 is the smallest root in* $(0, 1)$ *of* $2(1-X)^4 = 1 + 4X + X^2$. *Let $K > 0$ and $\mathcal{G}(K) = \{f \in \mathcal{A} \mid |b(n)| \leq K,\ n \geq 2\}$. Then $r_c(\mathcal{G}(K)) \geq R_2(K)$, where $R_2(K)$ is the smallest root in* $(0, 1)$ *of* $X^3 - 3X^2 + 4X = (1 - X)^3/K$.

The proof of the above proposition is actually a simple consequence of the following result of Alexander and Remak (see Theorem 1 of [Goodman 1957]).

Theorem 2.2 (Alexander–Remak). *If $f(z) = z + \sum_{n=2}^{\infty} b(n) z^n \in \mathcal{A}$ and*

$$\sum_{n=2}^{\infty} n^2 |b(n)| \leq 1,$$

then $f(D)$ is convex.

Adapting Yamashita's proof [1982, §2] we obtain the following:

Proposition 2.3. *Let* $K > 0$ *and* $\mathcal{H}(K) = \{f \in \mathcal{A} \mid |b(n)| \le Kn^2,\ n \ge 2\}$. *Then* $r_c(\mathcal{H}(K)) \ge R_3(K)$, *where* $R_3(K)$ *is the smallest root in* $(0, 1)$ *of*

$$X^5 - 5X^4 + 11X^3 + X^2 + 16X = (1 - X)^5/K.$$

Remark 2.4. Note that $R_3(K)$ is a strictly decreasing function of K, with

$$\sup_{K>0} R_3(K) = \lim_{K\to 0^+} R_3(K) = 1 \quad \text{and} \quad \inf_{K>0} R_3(K) = \lim_{K\to+\infty} R_3(K) = 0.$$

Moreover, for any $K > 0$ we have $R_3(K) \le R_2(K)$.

Proof of Proposition 2.3. For $f(z) = z + \sum_{n=2}^{\infty} b(n)z^n \in \mathcal{H}(K)$ and any $r \le R_3 = R_3(K)$ we have

$$\sum_{n=2}^{\infty} n^2|b(n)|r^{n-1} \le K \sum_{n=2}^{\infty} n^4 R_3^{n-1} = K\,\frac{R_3^5 - 5R_3^4 + 11R_3^3 + R_3^2 + 16R_3}{(1 - R_3)^5} = 1,$$

where the last equality follows from the fact that R_3 is chosen as the smallest real root in $(0, 1)$ of $X^5 - 5X^4 + 11X^3 + X^2 + 16X = (1 - X)^5/K$. Therefore we can apply Theorem 2.2 to $h(z) = r^{-1} f(rz)$, which is thus convex on $|z| < 1$. Hence $f(\{|z| < r\})$ is convex for any $r \le R_3$ and thus $R_3 \le r_c(\mathcal{H}(K))$. □

From this we obtain an explicit version of Theorem 13 of [Jessen and Wintner 1935] and Lemma 2.5 of [Lee 2014].

Proposition 2.5. *Let* N *be a fixed positive integer*,

$$G_j(z) = \sum_{n=1}^{\infty} a_j(n)z^n, \quad j = 1, \ldots, N,$$

and suppose there exist positive real numbers ρ_j *and* K_j *such that* $|a(n)| \le K_j \rho_j^{1-n}$ *for every* $n \ge 2$. *For any* $\boldsymbol{y} = (y_1, \ldots, y_J) \in \mathbb{C}^N$, *define*

$$g(r, \theta, \boldsymbol{y}) = \sum_{j=1}^{N} \operatorname{Re}\bigl(G_j(re^{2\pi i\theta})\overline{y_j}\bigr),$$

where $0 < r < \min_j \rho_j$ *and* $\theta \in [0, 1]$. *If* $\sum_{j=1}^{N} \overline{y_j} a_j(1) \ne 0$, *then there exists a positive constant* C *such that for any* $\delta > 0$ *we have*

$$\left|\int_0^1 e^{ig(r,\theta,\boldsymbol{y})}\, d\theta\right| \le \frac{24}{\sqrt{C\delta r\|\boldsymbol{y}\|}} \tag{2-1}$$

for every $0 < r \le R_3\bigl(\frac{1}{\delta}\sqrt{\sum_j |K_j|^2}\bigr) \min_j \rho_j$ *and every* $\boldsymbol{y}$ *such that* $\bigl|\sum_{j=1}^{N} \overline{y_j} a_j(1)\bigr| \ge \delta\|\boldsymbol{y}\| > 0$.

Proof. The proof is a combination of Theorems 12 and 13 of [Jessen and Wintner 1935] and Lemma 2.5 of [Lee 2014], and we use the aforementioned results to obtain explicit constants. Consider the power series

$$f(z) = \sum_{n=1}^{\infty}\Bigl(\sum_{j=1}^{N} \overline{y}_j a_j(n)\Bigr) z^n \quad \text{and} \quad h(z) = \sum_{n=1}^{\infty} n^2 \Bigl(\sum_{j=1}^{N} \overline{y}_j a_j(n)\Bigr) z^n.$$

Since, by hypothesis and the Cauchy–Schwarz inequality, we have

$$\Bigl|\sum_{j=1}^{N} \overline{y}_j a_j(n)\Bigr| \le \frac{\|\boldsymbol{y}\|\sqrt{\sum_j |K_j|^2}}{(\min_j \rho_j)^{n-1}} \quad \forall n \ge 2, \tag{2-2}$$

$f(z)$ and $h(z)$ are both holomorphic for $|z| < \min_j \rho_j$ and, by definition, we have

$$g(r, \theta, \boldsymbol{y}) = \operatorname{Re} f(re^{2\pi i\theta}) \quad \text{and} \quad g''(r, \theta, \boldsymbol{y}) = \frac{\partial^2}{\partial \theta^2} g(r, \theta, \boldsymbol{y}) = -4\pi^2 \operatorname{Re} h(re^{2\pi i\theta}).$$

By Proposition 2.1 we have that $f(re^{2\pi i\theta})$ is a parametric representation of a convex curve if

$$r \le R_2\biggl(\frac{\|\boldsymbol{y}\|\sqrt{\sum_j |K_j|^2}}{\bigl|\sum_{j=1}^{N} \overline{y}_j a_j(1)\bigr|}\biggr) \min_j \rho_j.$$

Indeed, substituting $w = z/\min_j \rho_j$, we have

$$\tilde{f}(w) = \frac{f(z/\min_j \rho_j)}{(\min_j \rho_j)\bigl(\sum_{j=1}^{N} \overline{y}_j a_j(1)\bigr)} = w + \sum_{n=2}^{\infty} (\min_j \rho_j)^{n-1} \biggl(\frac{\sum_{j=1}^{N} \overline{y}_j a_j(n)}{\sum_{j=1}^{J} \overline{y}_j a_j(1)}\biggr) w^n$$

and, by (2-2),

$$\tilde{f}(w) \in \mathcal{G}\biggl(\frac{\|\boldsymbol{y}\|\sqrt{\sum_j |K_j|^2}}{\bigl|\sum_{j=1}^{J} \overline{y}_j a_j(1)\bigr|}\biggr).$$

Analogously, by Proposition 2.3 we have that $h(re^{2\pi i\theta})$ is a parametric representation of a convex curve if

$$r \le R_3\biggl(\frac{\|\boldsymbol{y}\|\sqrt{\sum_j |K_j|^2}}{\bigl|\sum_{j=1}^{N} \overline{y}_j a_j(1)\bigr|}\biggr) \min_j \rho_j. \tag{2-3}$$

Therefore, by Remark 2.4, both $f(re^{2\pi i\theta})$ and $h(re^{2\pi i\theta})$ are parametric representations of convex curves for any fixed r satisfying (2-3). This implies that both $g(r, \theta, \boldsymbol{y})$ and $g''(r, \theta, \boldsymbol{y})$ have exactly two zeros mod 1. By the mean value theorem, we have that also $g'(r, \theta, \boldsymbol{y})$ has exactly two zeros mod 1, which separate those of $g''(r, \theta, \boldsymbol{y})$. Note that the zeros of $g'(r, \theta, \boldsymbol{y})$ and $g''(r, \theta, \boldsymbol{y})$ depend continuously on r and $\boldsymbol{y}$ since $g'(r, \theta, \boldsymbol{y})$ and $g''(r, \theta, \boldsymbol{y})$ are continuous functions in each variable.

We now consider the midpoints of the four arcs mod 1 determined by the zeros of $g'(r,\theta,\boldsymbol{y})$ and $g''(r,\theta,\boldsymbol{y})$. These midpoints clearly depend continuously on r and $\boldsymbol{y}$, and divide [0, 1] into four arcs, namely I_1, I_2, I_3 and I_4, such that I_1 and I_3 each contain one zero of $g'(r,\theta,\boldsymbol{y})$, while I_2 and I_4 each contain one zero of $g''(r,\theta,\boldsymbol{y})$. By van der Corput's lemma for oscillatory integrals (see [Titchmarsh 1986, Lemmas 4.2 and 4.4]) we have

$$\left|\int_{I_2\cup I_4} e^{ig(r,\theta,y)}\,d\theta\right| \le \frac{8}{\min\limits_{I_2\cup I_4}|g'(r,\theta,\boldsymbol{y})|}$$

and

$$\left|\int_{I_1\cup I_3} e^{ig(r,\theta,y)}\,d\theta\right| \le \frac{16}{\sqrt{\min\limits_{I_1\cup I_3}|g''(r,\theta,\boldsymbol{y})|}}.$$

Writing

$$g(r,\theta,\boldsymbol{y}) = r\left|\sum_{j=1}^{N}\overline{y}_j a_j(1)\right|\cos(2\pi(\theta-\xi)) + r^2 O(\|\boldsymbol{y}\|)$$

for some ξ, we see that by continuity there exists a positive constant C such that

$$\frac{g'(r,\theta,\boldsymbol{y})}{r\left|\sum_{j=1}^{N}\overline{y}_j a_j(1)\right|} \ge C \text{ on } I_2 \text{ and } I_4, \quad \text{and} \quad \frac{g''(r,\theta,\boldsymbol{y})}{r\left|\sum_{j=1}^{N}\overline{y}_j a_j(1)\right|} \ge C \text{ on } I_1 \text{ and } I_3$$

for every r satisfying (2-3) and $\boldsymbol{y}\in\mathbb{C}^N$.

We fix $\delta>0$, $\boldsymbol{y}\neq\boldsymbol{0}$ such that

$$\left|\sum_{j=1}^{J}\overline{y}_j a_j(1)\right| \ge \delta\|\boldsymbol{y}\|, \quad r \le R_3\left(\frac{1}{\delta}\sqrt{\sum_j |K_j|^2}\right)\min_j \rho_j,$$

and we obtain

$$\left|\int_0^1 e^{ig(r,\theta,y)}\,d\theta\right| \le \frac{8}{C\delta r\|\boldsymbol{y}\|} + \frac{16}{\sqrt{C\delta r\|\boldsymbol{y}\|}}.$$

Since $1/(C\delta r\|\boldsymbol{y}\|) \le 1/\sqrt{C\delta r\|\boldsymbol{y}\|}$ when $C\delta r\|\boldsymbol{y}\|\ge 1$,

$$\left|\int_0^1 e^{ig(r,\theta,y)}\,d\theta\right| \le \frac{24}{\sqrt{C\delta r\|\boldsymbol{y}\|}} \quad \text{for } \|\boldsymbol{y}\| \ge \frac{1}{C\delta r}.$$

On the other hand, we clearly have that $\left|\int_0^1 e^{ig(r,\theta,y)}\,d\theta\right| \le 1$, hence (2-1) holds whenever the RHS is ≥ 1. Therefore the result follows from the simple fact that the RHS of (2-1) is >24 when $0<\|\boldsymbol{y}\|<1/(C\delta r)$. □

Theorem 2.6. *Let $F_1(s), \dots, F_N(s)$ be orthogonal functions satisfying* (I) *and* (II). *Then there exists a positive constant A and infinitely many primes p such that*

$$\left| \int_0^1 \exp\Bigl(i \sum_{j=1}^N \operatorname{Re}\Bigl(\overline{y_j} \log F_{j,p}\Bigl(\sigma + i\frac{2\pi\theta}{\log p}\Bigr)\Bigr)\Bigr)\, d\theta \right| \le \frac{A}{\sqrt{\|\boldsymbol{y}\|}} p^{\sigma/2} \tag{2-4}$$

for every $\sigma \ge 1$ and every $\boldsymbol{y} = (y_1, \dots, y_N) \in \mathbb{C}^N \setminus \{\mathbf{0}\}$.

Proof. We want to apply Proposition 2.5 to

$$G_j(z) = \sum_{n=1}^\infty \frac{b_{F_j}(p^n)}{\sqrt{m_{j,j}}} z^n, \quad j = 1, \dots, N,$$

where the $m_{j,j}$ are as in (1-3). By (II) there exist K_{F_j} and $\theta_j < \frac{1}{2}$ such that for every prime p and every $n \ge 2$ we have $|b_{F_j}(p^n)| \le K_{F_j} p^{n\theta_j} \le K_{F_j} p^{2(n-1)\theta_j}$, $j = 1, \dots, N$. Thus, for $j = 1, \dots, N$ and every prime p we may take $K_j = K_{F_j}/\sqrt{m_{j,j}}$ and $\rho_j = p^{-2\theta_j}$.

On the other hand, by orthogonality we have that for any $\boldsymbol{y} \ne \mathbf{0}$

$$\sum_{p \le x} \left| \frac{\overline{y_1} b_{F_1}(p)}{\sqrt{m_{1,1}}} + \dots + \frac{\overline{y_N} b_{F_N}(p)}{\sqrt{m_{N,N}}} \right|^2 \Big/ p \sim \|\boldsymbol{y}\|^2 \log\log x, \quad \text{as } x \to \infty.$$

In particular this implies that there are infinitely many primes p such that

$$\left| \frac{\overline{y_1} b_{F_1}(p)}{\sqrt{m_{1,1}}} + \dots + \frac{\overline{y_N} b_{F_N}(p)}{\sqrt{m_{N,N}}} \right| \ge \frac{\|\boldsymbol{y}\|}{4}$$

for every $\boldsymbol{y} \ne \mathbf{0}$. For each such prime p we take $r = p^{-\sigma}$ and $\delta = \frac{1}{4}$. Then Proposition 2.5 yields

$$\left| \int_0^1 \exp\Bigl(i \sum_{j=1}^N \operatorname{Re}\Bigl(\frac{\overline{y_j}}{\sqrt{m_{j,j}}} \log F_{j,p}\Bigl(\sigma + i\frac{2\pi\theta}{\log p}\Bigr)\Bigr)\Bigr)\, d\theta \right| \le \frac{48}{\sqrt{C\|\boldsymbol{y}\|}} p^{\sigma/2} \tag{2-5}$$

when

$$p^{-\sigma} \le R_3\Biggl(4\sqrt{\sum_j \frac{|K_{F_j}|^2}{m_{j,j}}}\Biggr) p^{-2\max_j \theta_j} \tag{2-6}$$

and $\boldsymbol{y} \ne \mathbf{0}$. Note that (2-6) holds for every $\sigma \ge 1$ if p is sufficiently large since $\max_j \theta_j < \frac{1}{2}$. Now, substituting

$$\boldsymbol{y}' = (y_1', \dots, y_N') = \Bigl(\frac{y_1}{\sqrt{m_{1,1}}}, \dots, \frac{y_N}{\sqrt{m_{N,N}}} \Bigr)$$

in (2-5) we obtain that there are infinitely many primes p such that

$$\left|\int_0^1 \exp\left(i\sum_{j=1}^{N} \mathrm{Re}\left(\overline{y_j'} \log F_{j,p}\left(\sigma + i\frac{2\pi\theta}{\log p}\right)\right)\right) d\theta\right| \\ \leq \frac{48}{\sqrt{C}\sqrt[4]{m_{1,1}|y_1'|^2+\cdots+m_{N,N}|y_N'|^2}} p^{\sigma/2}$$

for every $\sigma \geq 1$ and every $\boldsymbol{y'} \in \mathbb{C}^N \setminus \{\boldsymbol{0}\}$. Since clearly there exists a positive constant D such that $\sqrt{m_{1,1}|y_1'|^2+\cdots+m_{N,N}|y_N'|^2} \geq D\|\boldsymbol{y'}\|$, the result follows immediately with $A = 48/\sqrt{DC}$. □

Remark 2.7. From the proof we have that (2-4) holds for $\sigma \geq 1$ because $\max_j \theta_j < \frac{1}{2}$ by (II). Therefore if we had that $\max_j \theta_j < \frac{\kappa}{2}$ for some $0 < \kappa < 1$, we would immediately have that (2-4) holds for every $\sigma \geq \kappa$.

3. On some distribution functions

This section is an adaptation of Chapter II of [Borchsenius and Jessen 1948]. We will also use Theorem 2.6 similarly to how Borchsenius and Jessen use Theorem 13 of [Jessen and Wintner 1935]. The particular distribution functions under investigation in this section may be found in [Lee 2014] and they will be used in Sections 4 and 5 for the proofs of Theorems 1.4 and 1.5. We refer to [Lee 2014] for a brief introduction to the theory developed by Jessen and Tornehave [1945] and Borchsenius and Jessen [1948] and how it may be applied to linear combinations of Euler products.

Given a function $F(s)$ satisfying (I) and (II), and a positive integer n we write

$$F_n(s) = \prod_{m=1}^{n} F_{p_m}(s) \quad \text{and} \quad F_n(\sigma, \boldsymbol{\theta}) = F_n(\sigma, \theta_1, \ldots, \theta_n) = \prod_{m=1}^{n} F_{p_m}\left(\sigma + i\frac{2\pi\theta_m}{\log p_m}\right),$$

where p_m is the m-th prime and $F_p(s)$ is determined by

$$\log F_p(s) = \sum_{k=1}^{\infty} b_F(p^k)p^{-ks}.$$

Remark 3.1. For any $n \geq 1$, $F_n(s)$ is well defined as a Dirichlet series (and Euler product) absolutely convergent for $\sigma > \theta = \theta_F$ by (II). Moreover, $F_n(s)$ and $\log F_n(s)$ converge uniformly for $\sigma \geq \sigma_0 > 1$ to $F(s)$ and $\log F(s)$, respectively.

Let $F_1(s), \ldots, F_N(s)$ be orthogonal functions satisfying (I) and (II). For $\boldsymbol{\theta} \in [0,1]^n$, we define

$$\boldsymbol{F}_n(\sigma, \boldsymbol{\theta}) = \big(F_{1,n}(\sigma, \boldsymbol{\theta}), \ldots, F_{N,n}(\sigma, \boldsymbol{\theta})\big)$$

and

$$\mathbf{log}\, \boldsymbol{F}_n(\sigma, \boldsymbol{\theta}) = \big(\log F_{1,n}(\sigma, \boldsymbol{\theta}), \dots, \log F_{N,n}(\sigma, \boldsymbol{\theta})\big).$$

To these functions we attach some distribution functions, namely for any Borel set $\boldsymbol{E} \subseteq \mathbb{C}^N$, $j, l \in \{1, \dots, N\}$, $j \neq l$ and $\sigma > 1$, we set

$$\boldsymbol{\lambda}_{\sigma,n;j}(\boldsymbol{E}) = \int_{\boldsymbol{W}_{\mathbf{log}\,\boldsymbol{F}_n}(\sigma,\boldsymbol{E})} \left| \frac{F'_{j,n}}{F_{j,n}}(\sigma, \boldsymbol{\theta}) \right|^2 d\boldsymbol{\theta} \tag{3-1}$$

and

$$\boldsymbol{\lambda}_{\sigma,n;j,l;\tau}(\boldsymbol{E}) = \int_{\boldsymbol{W}_{\mathbf{log}\,\boldsymbol{F}_n}(\sigma,\boldsymbol{E})} \left| \frac{F'_{j,n}}{F_{j,n}}(\sigma, \boldsymbol{\theta}) + \tau \frac{F'_{l,n}}{F_{l,n}}(\sigma, \boldsymbol{\theta}) \right|^2 d\boldsymbol{\theta}, \tag{3-2}$$

where $\boldsymbol{W}_{\mathbf{log}\,\boldsymbol{F}_n}(\sigma, \boldsymbol{E}) = \{\boldsymbol{\theta} \in [0, 1)^n \mid \mathbf{log}\,\boldsymbol{F}_n(\sigma, \boldsymbol{\theta}) \in \boldsymbol{E}\}$, and $\tau = \pm 1, \pm i$.

A distribution function μ on $\mathbb{C}^n$ is *absolutely continuous* (with respect to the Lebesgue measure, meas) if for every Borel set $\boldsymbol{E} \subseteq \mathbb{C}^n$, $\operatorname{meas}(\boldsymbol{E}) = 0$ implies $\mu(\boldsymbol{E}) = 0$ (cf. [Bogachev 2007, Definition 3.2.1]). By the Radon–Nikodym theorem (see, e.g., Theorem 3.2.2 in [Bogachev 2007]) this holds if and only if there exists a Lebesgue integrable function $G_\mu : \mathbb{C}^n \to \mathbb{R}_{\geq 0}$ such that

$$\mu(\boldsymbol{E}) = \int_{\boldsymbol{E}} G_\mu(\boldsymbol{x})\, d\boldsymbol{x}$$

for any Borel set $\boldsymbol{E} \subseteq \mathbb{C}^n$; $G_\mu(\boldsymbol{x})$ is the *density* of μ.

As a sufficient condition for absolute continuity we recall here the following result (cf. [Borchsenius and Jessen 1948, §6; Bogachev 2007, §3.8]).

Lemma 3.2. *Let μ be a distribution function on $\mathbb{C}^n$ and let $\hat{\mu}$ be its Fourier transform. If $\int_{\mathbb{C}^n} \|\boldsymbol{y}\|^q |\hat{\mu}(\boldsymbol{y})|\, d\boldsymbol{y} < \infty$ for some integer $q \geq 0$, then μ is absolutely continuous with density $G_\mu(\boldsymbol{x}) \in \mathcal{C}^q(\mathbb{C}^n)$ determined by the Fourier inversion formula*

$$G_\mu(\boldsymbol{x}) = \frac{1}{(2\pi)^{2n}} \int_{\mathbb{C}^n} e^{-i\langle \boldsymbol{x}, \boldsymbol{y}\rangle} \hat{\mu}(\boldsymbol{y})\, d\boldsymbol{y}.$$

We have the following result on the distribution functions defined above.

Theorem 3.3. *Let $F_1(s), \dots, F_N(s)$ be orthogonal functions satisfying* (I) *and* (II). *Then there exists $n_0 \geq 1$ such that the distribution functions $\boldsymbol{\lambda}_{\sigma,n;j}$ and $\boldsymbol{\lambda}_{\sigma,n;j,l;\tau}$ are absolutely continuous with continuous densities $G_{\sigma,n;j}(\boldsymbol{x})$ and $G_{\sigma,n;j,l;\tau}(\boldsymbol{x})$ for every $n \geq n_0$, $\sigma \geq 1$, $j, l \in \{1, \dots, N\}$, $j \neq l$ and $\tau = \pm 1, \pm i$. More generally for any $k \geq 0$ there exists $n_k \geq 1$ such that $G_{\sigma,n;j}(\boldsymbol{x})$, $G_{\sigma,n;j,l;\tau}(\boldsymbol{x}) \in \mathcal{C}^k(\mathbb{C}^N)$ for every $n \geq n_k$, $\sigma \geq 1$.*
Moreover, $\boldsymbol{\lambda}_{\sigma,n;j}$ and $\boldsymbol{\lambda}_{\sigma,n;j,l;\tau}$ converge weakly to some distribution functions $\boldsymbol{\lambda}_{\sigma;j}$ and $\boldsymbol{\lambda}_{\sigma;j,l;\tau}$ as $n \to \infty$, which are absolutely continuous with densities $G_{\sigma;j}(\boldsymbol{x})$, $G_{\sigma;j,l;\tau}(\boldsymbol{x}) \in \mathcal{C}^\infty(\mathbb{C}^N)$ for every $\sigma \geq 1$, $j, l \in \{1, \dots, N\}$, $j \neq l$ and $\tau = \pm 1, \pm i$. The functions $G_{\sigma,n;j}(\boldsymbol{x})$ and $G_{\sigma,n;j,l;\tau}(\boldsymbol{x})$ and their partial derivatives

converge uniformly for $\boldsymbol{x} \in \mathbb{C}^n$ and $1 \le \sigma \le M$ to $G_{\sigma;j}(\boldsymbol{x})$ and $G_{\sigma;j,l;\tau}(\boldsymbol{x})$ and their partial derivatives as $n \to \infty$ for every $M > 1$.

Proof. The proof is an adaptation of Theorem 5 of Borchsenius and Jessen [1948] (see also [Lee 2014, pp. 1827–1830]). We prove it just for $\boldsymbol{\lambda}_{\sigma,n;j}$ since the proof for the other distributions is completely similar.

We compute the Fourier transform of the functions $\boldsymbol{\lambda}_{\sigma,n;j}$ and get

$$\widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y}) = \int_{[0,1]^n} \exp\biggl(i \sum_{h=1}^{N} \mathrm{Re}(\log F_{h,n}(\sigma,\boldsymbol{\theta})\overline{y_h})\biggr) \left| \frac{F'_{j,n}}{F_{j,n}}(\sigma,\boldsymbol{\theta}) \right|^2 d\boldsymbol{\theta}, \tag{3-3}$$

for any $\boldsymbol{y} = (y_1, \dots, y_N) \in \mathbb{C}^N$. By Lemma 3.2, to prove the first part it is sufficient to show that for every $k \ge 0$ there exists n_k such that, for any $M > 1$, $\|\boldsymbol{y}\|^k \widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})$ is Lebesgue integrable for every $n \ge n_k$ and $1 \le \sigma \le M$. We recall that by (II) there exist K_{F_j} and $\theta_{F_j} < \frac{1}{2}$ such that

$$|b_{F_j}(p^n)| \le K_{F_j} p^{n\theta_{F_j}}$$

for every prime p and $k \ge 1$, $j = 1, \dots, N$. Then we have

$$\begin{aligned} |\widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})| \le \sup_{\sigma>1} \left| \frac{F'_{j,n}}{F_{j,n}}(\sigma,\boldsymbol{\theta}) \right|^2 &\le \sum_{m=1}^{n} \log^2 p_m \sum_{k=1}^{\infty} \frac{|b_{F_j}(p_m^k)|^2}{p_m^{2k\sigma}} \\ &\le K_{F_j}^2 \sum_p \frac{\log^2 p}{p^{2(\sigma-\theta_{F_j})}} < \infty \end{aligned} \tag{3-4}$$

for every $n \ge 1$ and $\sigma \ge 1$. Hence it is sufficient to show that there exist constants $C_k > 0$ and $n_k \ge 1$ such that for any $M > 1$ we have

$$|\widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})| \le C_k \|\boldsymbol{y}\|^{-\frac{5}{2}-k} \quad \text{as } \|\boldsymbol{y}\| \to \infty$$

for every $n \ge n_k$ and $1 \le \sigma \le M$. To prove this, note that we can write (cf. [Borchsenius and Jessen 1948, (47); Lee 2014, (3.24)])

$$\begin{aligned} \widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y}) = \sum_{m=1}^{n} K_{2,j}(p_m,\boldsymbol{y}) \prod_{\substack{\ell=1\\ \ell\ne m}}^{n} K_{0,j}(p_\ell,\boldsymbol{y}) \\ + \sum_{\substack{m,k=1\\ m\ne k}}^{n} K_{1,j}(p_m,\boldsymbol{y}) \overline{K_{1,j}(p_k,-\boldsymbol{y})} \prod_{\substack{\ell=1\\ \ell\ne m,k}}^{n} K_{0,j}(p_\ell,\boldsymbol{y}), \end{aligned} \tag{3-5}$$

where, for any prime p and $j \in \{1, \dots, N\}$, we take

$$K_{0,j}(p, \boldsymbol{y}) = \int_0^1 \exp\biggl(i \sum_{h=1}^N \operatorname{Re}\Bigl(\log F_{h,p}\Bigl(\sigma + i \frac{2\pi\theta}{\log p}\Bigr)\overline{y_h}\Bigr)\biggr)\, d\theta,$$

$$\begin{aligned} &K_{1,j}(p, \boldsymbol{y}) \\ &\quad = \int_0^1 \exp\biggl(i \sum_{h=1}^N \operatorname{Re}\Bigl(\log F_{h,p}\Bigl(\sigma + i \frac{2\pi\theta}{\log p}\Bigr)\overline{y_h}\Bigr)\biggr) \frac{F'_{j,p}}{F_{j,p}}\Bigl(\sigma + i \frac{2\pi\theta}{\log p}\Bigr)\, d\theta, \end{aligned} \tag{3-6}$$

$$\begin{aligned} &K_{2,j}(p, \boldsymbol{y}) \\ &\quad = \int_0^1 \exp\biggl(i \sum_{h=1}^N \operatorname{Re}\Bigl(\log F_{h,p}\Bigl(\sigma + i \frac{2\pi\theta}{\log p}\Bigr)\overline{y_h}\Bigr)\biggr) \biggl|\frac{F'_{j,p}}{F_{j,p}}\Bigl(\sigma + i \frac{2\pi\theta}{\log p}\Bigr)\biggr|^2 d\theta. \end{aligned}$$

Hence, we just need to estimate the functions defined in (3-6).

For all primes p and $j \in \{1, \dots, N\}$ we clearly have

$$|K_{0,j}(p, \boldsymbol{y})| \le 1. \tag{3-7}$$

On the other hand, by the hypotheses on $F_1(s), \dots, F_N(s)$ we can apply Theorem 2.6 and obtain a positive constant A and infinitely many primes p such that

$$|K_{0,j}(p, \boldsymbol{y})| \le \frac{A}{\sqrt{\|\boldsymbol{y}\|}} p^{\sigma/2} \tag{3-8}$$

for every $\sigma \ge 1$, $\boldsymbol{y} \ne \boldsymbol{0}$ and $j \in \{1, \dots, N\}$. Thus, putting together (3-7) and (3-8) we obtain that for any fixed integer $q \ge 1$ there exists m_q such that

$$\prod_{\substack{\ell=1 \\ \ell \ne m,k}}^n |K_{0,j}(p_\ell, \boldsymbol{y})| \le \biggl[\frac{A}{\sqrt{\|\boldsymbol{y}\|}} p_{m_q}^{\sigma/2}\biggr]^q \tag{3-9}$$

for every $m, k \le n$, $n \ge m_q$, $\sigma \ge 1$, $\boldsymbol{y} \ne \boldsymbol{0}$ and $j \in \{1, \dots, N\}$. Since we shall need it later, we also note that from the fact that $|e^{it} - 1 - it| \le t^2/2$ and by (II), for every prime p we get (cf. [Borchsenius and Jessen 1948, (50); Lee 2014, p. 1830])

$$|K_{0,j}(p, \boldsymbol{y}) - 1| \le \frac{\|\boldsymbol{y}\|^2}{2} \biggl(\sum_{h=1}^N K_{F_j}^2\biggr) \frac{1}{p^{2(\sigma - \max_h \theta_{F_h})}}. \tag{3-10}$$

For $K_{1,j}(p, \boldsymbol{y})$, using the fact that $|e^{it} - 1| \le |t|$ and (II), we obtain for any $\sigma \ge 1$ and any prime p (cf. [Borchsenius and Jessen 1948, (52); Lee 2014, (3.27)])

$$|K_{1,j}(p, \boldsymbol{y})| \le \|\boldsymbol{y}\| K_{F_j} \sqrt{\sum_{h=1}^N K_{F_h}^2} \frac{\log p}{p^{2(\sigma - \max_h \theta_{F_h})}}. \tag{3-11}$$

Finally, for any prime p, $\sigma \geq 1$ and $j \in \{1, \dots, N\}$, we simply have (cf. [Borchsenius and Jessen 1948, (53); Lee 2014, (3.26)])

$$|K_{2,j}(p, \boldsymbol{y})| \leq \int_0^1 \left| \frac{F'_{j,p}}{F_{j,p}}\left(\sigma + i\frac{2\pi\theta}{\log p}\right)\right|^2 d\theta \overset{\text{(II)}}{\leq} K_{F_j}^2 \frac{\log^2 p}{p^{2(\sigma-\theta_{F_j})}}. \tag{3-12}$$

Putting (3-7), (3-9), (3-11) and (3-12) into (3-5), for any fixed $M > 1$, $j \in \{1, \dots, N\}$ and $q \geq 0$, we get

$$|\widehat{\lambda_{\sigma,n;j}}(\boldsymbol{y})| \leq K_{F_j}^2 A^q \|\boldsymbol{y}\|^{-q/2} p_{m_q}^{q\sigma/2} \sum_{m=1}^{n} \frac{\log^2 p_m}{p_m^{2(\sigma-\theta_{F_j})}}$$
$$+ K_{F_j}^2 \left(\sum_{h=1}^{N} K_{F_h}^2\right) A^q \|\boldsymbol{y}\|^{2-q/2} p_{m_q}^{q\sigma/2} \left(\sum_{m=1}^{n} \frac{\log p_m}{p_m^{2(\sigma-\max_h \theta_{F_h})}}\right)^2$$

for any $n \geq m_q$, $\sigma \geq 1$ and $\boldsymbol{y} \neq \boldsymbol{0}$. Choosing $q = 9 + 2k$, $n_k = m_{9+2k}$ and

$$C_k = \left(\sum_{h=1}^{N} K_{F_h}^2\right) A^{9+2k} p_{n_k}^{(9+2k)M/2} \left(1 + \left(\sum_{h=1}^{N} K_{F_h}^2\right)^2 \sum_p \frac{\log p}{p^{2(\sigma-\max_h \theta_{F_h})}}\right)$$
$$\times \sum_p \frac{\log p}{p^{2(\sigma-\max_h \theta_{F_h})}}$$

we have

$$|\widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})| \leq C_k \|\boldsymbol{y}\|^{-\frac{5}{2}-k} \quad \text{when } \|\boldsymbol{y}\| \geq 1, \tag{3-13}$$

for every $n \geq n_k = m_{9+2k}$, $1 \leq \sigma \leq M$ and $j \in \{1, \dots, N\}$. Therefore, by Lemma 3.2, since n_k doesn't depend on M and since M is arbitrary, it follows that $\boldsymbol{\lambda}_{\sigma,n;j}$, $j = 1, \dots, N$, are absolutely continuous with continuous density for every $n \geq n_0$ and every $\sigma \geq 1$, while $G_{\sigma,n;j}(\boldsymbol{x}) \in \mathcal{C}^k(\mathbb{C}^N)$ for every $j \in \{1, \dots, N\}$, $n \geq n_k$ and $\sigma \geq 1$.

On the other hand, by (3-4), (3-5), (3-7), (3-10), (3-11), and (3-12), we have (cf. [Borchsenius and Jessen 1948, (60); Lee 2014, p. 1830])

$$|\widehat{\boldsymbol{\lambda}_{\sigma,n+1;j}}(\boldsymbol{y}) - \widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})| \ll \|\boldsymbol{y}\|^2 \frac{\log p_{n+1}}{p_{n+1}^{2(\sigma-\max_h \theta_{F_h})}}$$

for every $n \geq 1$, $\sigma \geq 1$ and $j \in \{1, \dots, N\}$. By the triangle inequality we thus get

$$|\widehat{\boldsymbol{\lambda}_{\sigma,n+k;j}}(\boldsymbol{y}) - \widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y})| \ll \|\boldsymbol{y}\|^2 \sum_{m=n+1}^{n+k} \frac{\log p_m}{p_m^{2(\sigma-\max_h \theta_{F_h})}}$$
$$\leq \|\boldsymbol{y}\|^2 \sum_{m=n+1}^{\infty} \frac{\log p_m}{p_m^{2(\sigma-\max_h \theta_{F_h})}} \tag{3-14}$$

for every $n,\ k \geq 1$ and $\sigma \geq 1$. Hence, by Cauchy's criterion, there exist the limit functions

$$\widehat{\boldsymbol{\lambda}_{\sigma;j}}(\boldsymbol{y}) = \lim_{n\to\infty} \widehat{\boldsymbol{\lambda}_{\sigma,n;j}}(\boldsymbol{y}), \qquad j = 1, \dots, N,$$

and by (3-14) it is clear that the convergence is uniform in $\|\boldsymbol{y}\| \leq a$, for every $a > 0$. Therefore, by Lévy's convergence theorem (see, e.g., Theorem 8.8.1 in [Bogachev 2007]), we have that $\widehat{\boldsymbol{\lambda}_{\sigma;j}}(\boldsymbol{y})$ is the Fourier transform of some distribution function $\boldsymbol{\lambda}_{\sigma;j}$ and $\boldsymbol{\lambda}_{\sigma,n;j} \to \boldsymbol{\lambda}_{\sigma;j}$ weakly as $n \to \infty$, for $j = 1, \dots, N$. Moreover by (3-13) we have that we may apply the dominated convergence theorem and thus $\boldsymbol{\lambda}_{\sigma;j}$ are absolutely continuous for every $\sigma \geq 1$ and $j \in \{1, \dots, N\}$, with density $G_{\sigma;j}(\boldsymbol{x}) \in \mathcal{C}^\infty(\mathbb{C})$ (for the arbitrariness of M and k). Moreover, since $G_{\sigma,n;j}(\boldsymbol{x})$ and $G_{\sigma;j}(\boldsymbol{x})$ are determined by the inverse Fourier transform (see Lemma 3.2), the dominated convergence theorem yields that $G_{\sigma,n;j}(\boldsymbol{x})$ and their partial derivatives converge uniformly for $\boldsymbol{x} \in \mathbb{C}^n$ and $1 \leq \sigma \leq M$ toward $G_{\sigma;j}(\boldsymbol{x})$ and their partial derivatives for every $j \in \{1, \dots, N\}$. □

Theorem 3.4. *For any $\alpha > 0$ and $q \geq 0$ the densities $G_{\sigma;j}(\boldsymbol{x})$ and $G_{\sigma,n;j}(\boldsymbol{x})$, $n \geq n_q$, together with their partial derivatives of order $\leq q$, have a majorant of the form $K_q e^{-\alpha\|\boldsymbol{x}\|^2}$ for every $\sigma \geq 1$, $j, l \in \{1, \dots, N\}$, $j \neq l$ and $\tau = \pm 1, \pm i$.*

Proof. This is a straightforward adaptation of Theorems 6 and 9 of [Borchsenius and Jessen 1948]. □

Theorem 3.5. *The distribution functions $\boldsymbol{\lambda}_{\sigma;j}$, $\boldsymbol{\lambda}_{\sigma;j,l;\tau}$, $\boldsymbol{\lambda}_{\sigma,n;j}$ and $\boldsymbol{\lambda}_{\sigma,n;j,l;\tau}$, for $n \geq n_0$, depend continuously on σ, and their densities $G_{\sigma;j}(\boldsymbol{x})$, $G_{\sigma;j,l;\tau}(\boldsymbol{x})$, $G_{\sigma,n;j}(\boldsymbol{x})$ and $G_{\sigma,n;j,l;\tau}(\boldsymbol{x})$, together with their partial derivatives of order $\leq q$ if $n \geq n_q$, are continuous in σ for every $\sigma \geq 1$, $j, l \in \{1, \dots, N\}$, $j \neq l$ and $\tau = \pm 1, \pm i$.*

Proof. As in Theorem 9 of [Borchsenius and Jessen 1948] the result follows from (3-13), (3-14) and the Fourier inversion formula. □

Remark 3.6. As for Remark 2.7, note that Theorems 3.3, 3.4 and 3.5 hold for $\sigma > 1$ because $\max_j \theta_{F_j} < \frac{1}{2}$ by (II). Therefore if we had that $\max_j \theta_{F_j} < \kappa/2$ for some $0 < \kappa < 1$ we would immediately have that (2-4) holds for every $\sigma > \kappa$.

4. Zeros of sums of two Euler products

Let $F_1(s)$ and $F_2(s)$ be functions satisfying (I) and (II), and $c_1, c_2 \in \mathbb{C} \setminus \{0\}$. We then set

$$L(s) = c_1 F_1(s) + c_2 F_2(s).$$

To study the distribution of the zeros of $L(s)$ for $\sigma > 1$, we note that, since $F_1(s)F_2(s) \neq 0$ for $\sigma > 1$,

$$L(s) = 0 \quad \Leftrightarrow \quad \log\biggl(\frac{F_1(s)}{F_2(s)}\biggr) = \log\biggl(-\frac{c_2}{c_1}\biggr).$$

This idea was used by Gonek [1981], and later by Bombieri and Mueller [2008] and Bombieri and Ghosh [2011]. Moreover, if $F_1(s)$ and $F_2(s)$ are orthogonal, then it is easy to show that $\frac{F_1}{F_2}(s)$ satisfies (I), (II) and, if we write $\frac{F_1}{F_2}(s) = \sum_{n=1}^{\infty} a(n)n^{-s}$,

$$\sum_{p\le x} \frac{|a(p)|^2}{p} = (\kappa + o(1)) \log\log x, \quad x \to \infty, \tag{4-1}$$

for some constant $\kappa > 0$. Therefore Theorem 1.4 follows immediately from the following more general result on the value distribution of the logarithm of an Euler product.

Theorem 4.1. *Let $F(s)$ be a function satisfying* (I), (II) *and* (4-1), *and $c \in \mathbb{C}$. Then the Dirichlet series* $\log F(s) - c$ *has no isolated vertical lines containing zeros in the half-plane $\sigma > 1$.*

Proof. The first part of the proof is similar to Borchsenius and Jessen's application [1948, Theorems 11 and 13] of their Theorems 5–9 to the Riemann zeta function.

For every $n \ge 1$ consider the Dirichlet series $\log F_n(s)$, which are absolutely convergent for $\sigma > \theta_F$ by Remark 3.1. Let $\nu_{\sigma,n}$ be, for every $\sigma > \theta_F$, the *asymptotic distribution function* of $\log F_n(s)$ with respect to $|(F_n'/F_n)(s)|^2$, defined for any Borel set $E \subseteq \mathbb{C}$ by (cf. [Borchsenius and Jessen 1948, §7])

$$\nu_{\sigma,n}(E) = \lim_{T_2 - T_1 \to \infty} \frac{1}{T_2 - T_1} \int_{V_{\log F_n}(\sigma, T_1, T_2, E)} \left| \frac{F_n'}{F_n}(s) \right|^2 dt,$$

where $V_{\log F_n}(\sigma, T_1, T_2, E) = \{t \in (T_1, T_2) \mid \log F_n(\sigma + it) \in E\}$. For $\sigma \ge 1$, we compute its Fourier transform and, by the Kronecker–Weyl theorem (see, e.g., [Karatsuba and Voronin 1992, §A.8]) we get (cf. [Borchsenius and Jessen 1948, p. 160] or [Lee 2014, p. 1819])

$$\widehat{\nu_{\sigma,n}}(y) = \int_{[0,1]^n} \exp\bigl(i \operatorname{Re}(\log F_n(\sigma, \boldsymbol{\theta}) \bar{y})\bigr) \left| \frac{F_n'}{F_n}(\sigma, \boldsymbol{\theta}) \right|^2 d\boldsymbol{\theta} \overset{(3\text{-}3)}{=} \widehat{\boldsymbol{\lambda}_{\sigma,n;1}}(y),$$

with $N = 1$. For simplicity we write $\lambda_{\sigma,n} = \boldsymbol{\lambda}_{\sigma,n;1}$. By the uniqueness of the Fourier transform (see, e.g., Proposition 3.8.6 in [Bogachev 2007]) we have that $\nu_{\sigma,n} = \lambda_{\sigma,n}$ as distribution functions for every $\sigma \ge 1$ and $n \ge 1$.

By Theorem 3.3 we know that $\nu_{\sigma,n} = \lambda_{\sigma,n}$ is absolutely continuous for $n \ge n_0$ with density $G_{\sigma,n}(x)$ which is a continuous function of both σ and x (see Theorem 3.5). Hence for any $n \ge n_0$, $x \in \mathbb{C}$ and $\sigma > \theta_F$ we have that the *Jensen function* $\varphi_{\log F_n - x}(\sigma)$ (see, e.g., Theorem 5 of [Jessen and Tornehave 1945]) is twice differentiable with continuous second derivative (cf. [Borchsenius and Jessen 1948, §9])

$$\varphi''_{\log F_n - x}(\sigma) = 2\pi G_{\sigma,n}(x). \tag{4-2}$$

Note that in order to apply Theorems 3.3 and 3.5 we have implicitly made use of the orthogonality hypothesis.

On the other hand, for any $1 < \sigma_1 < \sigma_2$, by the uniform convergence of $\log F_n(s)$ of Remark 3.1 and by Theorem 6 of [Jessen and Tornehave 1945], we have that

$$\varphi_{\log F_n - x}(\sigma) \to \varphi_{\log F - x}(\sigma) \quad \text{as } n \to \infty \tag{4-3}$$

uniformly for $\sigma_1 \le \sigma \le \sigma_2$. Moreover, by Theorem 3.3, $G_{\sigma,n}(x)$ converges uniformly for $\sigma_1 \le \sigma \le \sigma_2$ toward $G_\sigma(x)$, which is continuous in both σ and x. Then, by (4-2), (4-3), the convexity of $\varphi_{\log F_n - x}$ and Theorem 7.17 in [Rudin 1976] we obtain that for any $x \in \mathbb{C}$ the Jensen function $\varphi_{\log F - x}(\sigma)$ is twice differentiable with continuous second derivative

$$\varphi''_{\log F - x}(\sigma) = 2\pi G_\sigma(x).$$

We fix an arbitrary $c \in \mathbb{C}$ and we note the following: Suppose that $\varphi''_{\log F - c}(\sigma_0) > 0$ for some $\sigma_0 > 1$. Then, by continuity, there exists $\varepsilon_0 > 0$ such that $\varphi''_{\log F - c}(\sigma) > 0$ for every $\sigma \in (\sigma_0 - \varepsilon_0, \sigma_0 + \varepsilon_0)$. Then, for any $0 < \varepsilon < \varepsilon_0$, by Theorem 31 of [Jessen and Tornehave 1945] and the mean value theorem, we have

$$\begin{aligned}\lim_{T_2 - T_1 \to \infty} & \frac{N_{\log F - c}(\sigma_0 - \varepsilon, \sigma_0 + \varepsilon, T_1, T_2)}{T_2 - T_1} \\ &= \frac{1}{2\pi}\big(\varphi'_{\log F - c}(\sigma_0 + \varepsilon) - \varphi'_{\log F - c}(\sigma_0 - \varepsilon)\big) = \frac{\varepsilon}{2\pi}\varphi''_{\log F - c}(\sigma_\varepsilon) > 0,\end{aligned}$$

for some $\sigma_\varepsilon \in (\sigma_0 - \varepsilon, \sigma_0 + \varepsilon)$, i.e., there are infinitely many zeros with real part $\sigma \in (\sigma_0 - \varepsilon, \sigma_0 + \varepsilon)$. This means, by letting $\varepsilon \to 0^+$, that σ_0 is the limit point of the real parts of some zeros of $\log F(s) - c$ (or σ_0 is itself a zero).

Now, suppose there exists $\rho_0 = \beta_0 + i\gamma_0$ with $\beta_0 > 1$ such that $\log F(\rho_0) - c = 0$. If we suppose that $\varphi''_{\log F - c}(\beta_0) > 0$, then $\sigma = \beta_0$ cannot be an isolated vertical line containing zeros since β_0 is the limit point of the real parts of some zeros. Suppose otherwise that $\varphi''_{\log F - c}(\tilde{\sigma}) = 0$, and for any $\delta > 0$ consider the intervals $I_\delta^+ = (\tilde{\sigma}, \tilde{\sigma} + \delta)$ and $I_\delta^- = (\tilde{\sigma} - \delta, \tilde{\sigma})$. Note that in general, if $\varphi''_{\log F - c}(\sigma) = 0$ for every $\sigma \in (\sigma_1, \sigma_2)$, for some $1 < \sigma_1 < \sigma_2$, then Theorem 31 of [Jessen and Tornehave 1945] and the mean value theorem imply that $\log F(s) - c$ has no zeros for $\sigma_1 < \sigma < \sigma_2$. Therefore, in at least one of I_δ^+ or I_δ^- there are infinitely many σ such that $\varphi''_{\log F - c}(\sigma) > 0$, for any $\delta > 0$, by almost periodicity. Hence, letting $\delta \to 0$, we see that there exists a sequence $\{\sigma_\delta\}_\delta$ such that $\varphi''_{\log F - c}(\sigma_\delta) > 0$ and $\sigma_\delta \to \beta_0$. Since every σ_δ is the limit point of the real parts of some zeros, we conclude that also β_0 is the limit point of the real parts of some zeros. □

5. *c*-values of sums of at least three Euler products

We first state the following simple result which is a generalization of Lemma 2.4 of [Lee 2014].

Lemma 5.1. *Let $F(s)$ be a function satisfying* (I), (II) *and* (III), $\sigma_0 > \frac{1}{2}$ *and k be a fixed positive integer. Then there exists a positive constant $A_k(\sigma_0)$ such that*

$$\int_{[0,1]^n} |F_n(\sigma, \boldsymbol{\theta})|^{2k}\, d\boldsymbol{\theta} \le A_k(\sigma_0) \quad \text{and} \quad \int_{[0,1]^n} |F_n'(\sigma, \boldsymbol{\theta})|^{2k}\, d\boldsymbol{\theta} \le A_k(\sigma_0)$$

for every $n \ge 1$ and $\sigma \ge \sigma_0$.

Proof. As in Lemma 2.4 of [Lee 2014] the proof follows from a bound of

$$\mathcal{J}_k(z_1, \dots, z_n, w_1, \dots, w_n) = \int_{[0,1]^n} \prod_{j=1}^{k} F_n(\sigma + z_j, \boldsymbol{\theta}) \overline{F_n(\sigma + \overline{w_j}, \boldsymbol{\theta})}\, d\boldsymbol{\theta}$$

and Cauchy's integral formula on polydiscs. This bound may be obtained with the same computations as in Lemma 2.5 of [Lee 2014] by replacing the Ramanujan bound $|a(n)| \le 1$ with the weaker Ramanujan conjecture $|a(n)| \ll_\varepsilon n^\varepsilon$, where we take $0 < \varepsilon < (2\sigma_0 - 1)/4$. □

Proof of Theorem 1.5. To handle this case we follow an idea of Lee [2014, §3.2] and we use the distribution functions studied in Section 3, similarly to what we have done in the previous section for $N = 2$. We give only a sketch of the proof.

For every $n \ge 1$ we write

$$L_n(s) = \sum_{j=1}^{N} c_j F_{j,n}(s),$$
$$L_n(\sigma, \boldsymbol{\theta}) = L_n(\sigma, \theta_1, \dots, \theta_n) = \sum_{j=1}^{N} c_j F_{j,n}(s, \theta_1, \dots, \theta_n).$$

Let $\nu_{\sigma,n}$ be the asymptotic distribution function of $L_n(s)$ with respect to $|L_n'(s)|^2$ defined for any Borel set $E \subseteq \mathbb{C}$ by (cf. [Borchsenius and Jessen 1948, §7])

$$\nu_{\sigma,n}(E) = \lim_{T_2 - T_1 \to \infty} \frac{1}{T_2 - T_1} \int_{V_{L_n}(\sigma, T_1, T_2, E)} |L_n'(s)|^2\, dt,$$

where $V_{L_n}(\sigma, T_1, T_2, E) = \{t \in (T_1, T_2) \mid L_n(\sigma + it) \in E\}$. As in Theorem 4.1, by the Kronecker–Weyl theorem and the uniqueness of the Fourier transform, we have that $\nu_{\sigma,n} = \lambda_{\sigma,n}$, for any $n \ge 1$ and $\sigma \ge 1$, where $\lambda_{\sigma,n}$ is the distribution function of $L_n(s, \boldsymbol{\theta})$ with respect to $|L_n'(s, \boldsymbol{\theta})|^2$, defined for every Borel set $E \subseteq \mathbb{C}$ by

$$\lambda_{\sigma,n}(E) = \int_{W_{L_n}(\sigma, E)} |L_n'(\sigma, \boldsymbol{\theta})|^2\, d\boldsymbol{\theta},$$

with $W_{L_n}(\sigma, E) = \{\boldsymbol{\theta} = (\theta_1, \dots, \theta_n) \in [0,1)^n \mid L_n(\sigma, \boldsymbol{\theta}) \in E\}$. We want to show that there exists $\tilde{n} \ge 1$ such that $\lambda_{\sigma,n}$, and hence $\nu_{\sigma,n}$, is absolutely continuous with continuous density, which we call $H_{\sigma,n}(x)$, for every $n \ge \tilde{n}$ and $\sigma \ge 1$.

As in [Lee 2014, pp. 1830–1831], we compute the Fourier transform of $\lambda_{\sigma,n}$ and, for $\sigma \geq 1$ and $n \geq n_0$, we get

$$\widehat{\lambda_{\sigma,n}}(y) = \sum_{j,l=1}^{N} \overline{c_j} c_l (2\pi)^N \int_{\mathbb{R}_+^N} \int_{\mathbb{R}^N} \exp\biggl(i \sum_{h=1}^{N} |c_h \bar{y}| r_h \sin(2\pi(\theta_h - \alpha_h)) - 2\pi i \theta_j + 2\pi i \theta_l\biggr) \times r_j r_l G_{\sigma,n;j,l}(\boldsymbol{r}) \frac{dr_1}{r_1} \cdots \frac{dr_N}{r_N} \, d\theta_1 \cdots d\theta_N,$$

where $\boldsymbol{r} = (\log r_1 + 2\pi i \theta_1, \ldots, \log r_N + 2\pi i \theta_N)$, α_h is determined by the argument of $c_h \bar{y}$, for $h = 1, \ldots, N$, and

$$G_{\sigma,n;j,l}(\boldsymbol{x}) = \begin{cases} G_{\sigma,n;j}(\boldsymbol{x}), & j = l, \\ \sum_{\tau=\pm1,\pm i} \bar{\tau} G_{\sigma,n;j,l;\tau}(\boldsymbol{x}), & j \neq l \end{cases}$$

is defined from the densities of the distribution functions $\boldsymbol{\lambda}_{\sigma,n;j}$ and $\boldsymbol{\lambda}_{\sigma,n;j,l;\tau}$ of Section 3.

For any $h \in \{1, \ldots, N\}$ and any $\varepsilon > 0$ let

$$A_{h,\varepsilon} = \bigl\{\theta \in \mathbb{R} \mid |\theta - \alpha_h - m\pi| < \varepsilon \text{ for some } m \in \mathbb{Z}\bigr\}.$$

Then we note that integrating by parts with respect to r_h, $h = 1, \ldots, N$, and using the majorant $K_N \exp\bigl(-\bigl[\sum_{h=1}^{N} \log^2 r_h + \theta_h^2\bigr]\bigr)$ of Theorem 3.4 for the partial derivatives up to order N of the density $G_{\sigma,n;j,l}(\boldsymbol{r})$, for $n \geq n_N$ and $\sigma \geq 1$, we obtain (cf. [Lee 2014, p. 1832])

$$\begin{aligned} &\int_{\mathbb{R}\setminus A_{1,\varepsilon}} \cdots \int_{\mathbb{R}\setminus A_{N,\varepsilon}} \int_{\mathbb{R}_+^N} \exp\biggl(i \operatorname{Re}\biggl(\sum_{h=1}^{N} r_h c_h \bar{y} e^{2\pi i \theta_h}\biggr) - 2\pi i \theta_j + 2\pi i \theta_l\biggr) \\ &\qquad\qquad \times r_j r_l G_{\sigma,n;j,l}(\boldsymbol{r}) \frac{dr_1}{r_1} \cdots \frac{dr_N}{r_N} \, d\theta_1 \cdots d\theta_N \\ &\quad \ll \prod_{h=1}^{N} \int_{\mathbb{R}\setminus A_{h,\varepsilon}} \frac{1}{|c_h \bar{y}| \sin(2\pi(\theta_h - \alpha_h))} e^{-\theta_h^2} d\theta_h \\ &\quad \ll \frac{1}{(\varepsilon |y|)^N} \end{aligned} \tag{5-1}$$

for every $n \geq n_N$, $\sigma \geq 1$ and $y \neq 0$. Analogously, integrating by parts with respect to θ_h, $h = 1, \ldots, N$, using van der Corput's lemma for oscillatory integrals (see, e.g., Lemma 4.2 in [Titchmarsh 1986]) on each interval $[\alpha_h + m_h/2 - \varepsilon, \alpha_h + m_h/2 + \varepsilon]$ with $\varepsilon < \frac{1}{2}$, and the majorant $K_N \exp\bigl(-\bigl[\sum_{h=1}^{N} \log^2 r_h + \theta_h^2\bigr]\bigr)$ of Theorem 3.4 for the partial derivatives up to order N of the density $G_{\sigma,n;j,l}(\boldsymbol{r})$, $n \geq n_N$ and $\sigma \geq 1$, we obtain (cf. [Lee 2014, p. 1832])

$$
\begin{aligned}
\int_{\mathbb{R}_+^N}\int_{A_{1,\varepsilon}}\cdots\int_{A_{N,\varepsilon}} \exp\biggl(i\,\mathrm{Re}\biggl(\sum_{h=1}^N r_h c_h \bar{y} e^{2\pi i\theta_h}\biggr) - 2\pi i\theta_j + 2\pi i\theta_l\biggr)& \\
\times r_j r_l G_{\sigma,n;j,l}(\boldsymbol{r})\frac{dr_1}{r_1}\cdots\frac{dr_N}{r_N}\,d\theta_1\cdots d\theta_N& \\
\ll \prod_{h=1}^N \int_{\mathbb{R}_+} \frac{1}{|c_h\bar{y}|} e^{-\log^2 r_h} dr_h& \\
\ll \frac{1}{|y|^N},& \qquad (5\text{-}2)
\end{aligned}
$$

for every $n \ge n_N$, $\sigma \ge 1$, $|y| \ge \max_h 1/|c_h|$ and $\varepsilon > 0$ sufficiently small. Note that to apply Theorem 3.4 we have implicitly made use of the orthogonality hypothesis. Fixing $\varepsilon > 0$ sufficiently small so that (5-2) holds and putting together (5-1) and (5-2), we obtain

$$
|\widehat{\nu_{\sigma,n}}(y)| = |\widehat{\lambda_{\sigma,n}}(y)| \ll |y|^{-N} \ll |y|^{-3} \qquad (5\text{-}3)
$$

since $N \ge 3$, for every $n \ge n_N$, $\sigma \ge 1$ and $|y| \ge \max(1, \max_h |c_h|^{-1})$. By Lemma 3.2 we have thus proved that $\nu_{\sigma,n}$ is absolutely continuous for every $n \ge \tilde{n} = n_N$ and $\sigma \ge 1$. Moreover, since $\nu_{\sigma,n}$ depends continuously on σ (cf. [Borchsenius and Jessen 1948, §7]), we have that $\widehat{\nu_{\sigma,n}}$ is continuous in σ. Therefore (5-3) and the Fourier inversion formula imply that $H_{\sigma,n}(x)$ is continuous in both σ and x. Note that all implied constants in (5-3) are independent of n.

Now we prove that the absolutely continuous distribution functions $\lambda_{\sigma,n}$ converge weakly as $n \to \infty$ toward the absolutely continuous distribution function λ_σ with density $H_\sigma(x)$ which is continuous in both σ and x. Moreover, we want to show that, for any $1 < \sigma_1 < \sigma_2$, $H_{\sigma,n}(x)$ converges uniformly for $\sigma_1 \le \sigma \le \sigma_2$ toward $H_\sigma(x)$ as $n \to \infty$.

For this, note that

$$
\begin{aligned}
&L_{n+1}(\sigma,\boldsymbol{\theta},\theta_{n+1}) \\
&\quad= \sum_{j=1}^N c_j F_{j,n}(\sigma,\boldsymbol{\theta}) F_{j,p_{n+1}}\biggl(\sigma + i\frac{2\pi\theta_{n+1}}{\log p_{n+1}}\biggr) \\
&\quad\overset{\text{(III)}}{=} \sum_{j=1}^N c_j F_{j,n}(\sigma,\boldsymbol{\theta})\biggl(1 + \frac{a_{F_j}(p_{n+1})}{p_{n+1}^\sigma} e^{2\pi i\theta_{n+1}} + O_\varepsilon\biggl(\frac{1}{p_{n+1}^{2(\sigma-\varepsilon)}}\biggr)\biggr) \\
&\quad= L_n(\sigma,\boldsymbol{\theta}) + \frac{e^{2\pi i\theta_{n+1}}}{p_{n+1}^\sigma}\sum_{j=1}^N c_j a_{F_j}(p_{n+1}) F_{j,n}(\sigma,\boldsymbol{\theta}) + O_\varepsilon\biggl(\frac{\sum_j |F_{j,n}|}{p_{n+1}^{2(\sigma-\varepsilon)}}\biggr) \qquad (5\text{-}4)
\end{aligned}
$$

for every $\sigma \geq 1$ and $0 < \varepsilon < \frac{1}{2}$. Similarly

$$\begin{aligned} L'_{n+1}&(\sigma, \boldsymbol{\theta}, \theta_{n+1}) \\ &= L'_n(\sigma, \boldsymbol{\theta}) + \frac{e^{2\pi i \theta_{n+1}}}{p_{n+1}^{\sigma}} \sum_{j=1}^{N} c_j a_{F_j}(p_{n+1}) \big[F'_{j,n}(\sigma, \boldsymbol{\theta}) - \log p_{n+1} F_{j,n}(\sigma, \boldsymbol{\theta}) \big] \\ &\qquad\qquad + O_\varepsilon \biggl(\frac{\log p_{n+1} \sum_j |F_{j,n}| + |F'_{j,n}|}{p_{n+1}^{2(\sigma-\varepsilon)}} \biggr) \end{aligned}$$

for every $\sigma \geq 1$ and $0 < \varepsilon < \frac{1}{2}$. Hence we have (cf. [Lee 2014, (3.20)])

$$\begin{aligned} \widehat{\lambda_{\sigma,n+1}}&(y) - \widehat{\lambda_{\sigma,n}}(y) \\ &= \int_{[0,1]^{n+1}} \big[e^{i \operatorname{Re}(L_{n+1}(\sigma,\boldsymbol{\theta},\theta_{n+1})\bar{y})} - e^{i \operatorname{Re}(L_n(\sigma,\boldsymbol{\theta})\bar{y})} \big] |L'_n(\sigma, \boldsymbol{\theta})|^2 \, d\boldsymbol{\theta} \, d\theta_{n+1} \\ &\quad + \frac{2}{p_{n+1}^{\sigma}} \int_{[0,1]^{n+1}} e^{i \operatorname{Re}(L_{n+1}(\sigma,\boldsymbol{\theta},\theta_{n+1})\bar{y})} \operatorname{Re} \biggl(\overline{L'_n(\sigma, \boldsymbol{\theta})} e^{2\pi i \theta_{n+1}} \\ &\qquad\qquad \times \sum_{j=1}^{N} c_j a_{F_j}(p_{n+1}) \big(F'_{j,n}(\sigma, \boldsymbol{\theta}) - \log p_{n+1} F_{j,n}(\sigma, \boldsymbol{\theta}) \big) \biggr) \\ &\qquad\qquad\qquad \times d\boldsymbol{\theta} \, d\theta_{n+1} \\ &\quad + O_\varepsilon \biggl(\frac{\log p_{n+1}}{p_{n+1}^{2(\sigma-\varepsilon)}} \int_{[0,1]^{n+1}} \Bigl(1 + \sum_j |F'_{j,n}| \Bigr) \Bigl(\sum_j |F_{j,n}| + |F'_{j,n}| \Bigr) \, d\boldsymbol{\theta} \, d\theta_{n+1} \biggr) \\ &\quad + O_\varepsilon \biggl(\frac{\log^2 p_{n+1}}{p_{n+1}^{4(\sigma-\varepsilon)}} \int_{[0,1]^{n+1}} \Bigl(\sum_j |F_{j,n}| + |F'_{j,n}| \Bigr)^2 \, d\boldsymbol{\theta} \, d\theta_{n+1} \biggr). \end{aligned} \tag{5-5}$$

for every $\sigma \geq 1$ and $0 < \varepsilon < \frac{1}{2}$.

For the first term, using again $|e^{it} - 1 - it| \leq t^2/2$, we obtain (cf. [Lee 2014, (3.22)])

$$\biggl| \int_0^1 \big[e^{i \operatorname{Re}(L_{n+1}(\sigma,\boldsymbol{\theta},\theta_{n+1})\bar{y})} - e^{i \operatorname{Re}(L_n(\sigma,\boldsymbol{\theta})\bar{y})} \big] \, d\theta_{n+1} \biggr| \ll_{\varepsilon,a} \frac{\sum_j |F_{j,n}| + |F_{j,n}|^2}{p_{n+1}^{2(\sigma-\varepsilon)}}$$

for $|y| \leq a$, $a > 0$, $\sigma \geq 1$ and $0 < \varepsilon < \frac{1}{2}$. For the second term we get directly from (5-4) and $|e^{it} - 1| \leq |t|$ that

$$\biggl| \int_0^1 e^{i \operatorname{Re}(L_{n+1}(\sigma,\boldsymbol{\theta},\theta_{n+1})\bar{y})} e^{\pm 2\pi i \theta_{n+1}} \, d\theta_{n+1} \biggr| \ll_{\varepsilon,a} \frac{\sum_j |F_{j,n}|}{p_{n+1}^{(\sigma-\varepsilon)}}$$

for $|y| \leq a$, $a > 0$, $\sigma \geq 1$ and $0 < \varepsilon < \frac{1}{2}$. We fix $0 < \varepsilon < \frac{1}{2}$, then putting these together, by triangle inequality and Lemma 5.1 with $\sigma_0 = 1$, we get (cf. [Lee 2014, p. 1826])

$$|\widehat{\lambda_{\sigma,n+1}}(y) - \widehat{\lambda_{\sigma,n}}(y)| \ll_{a,\varepsilon} \frac{\log p_{n+1}}{p_{n+1}^{2(\sigma-\varepsilon)}}$$

uniformly for $|y| \le a$, $a > 0$, and for every $\sigma \ge 1$. It follows that for any $k > 0$

$$|\widehat{\lambda_{\sigma,n+k}}(y) - \widehat{\lambda_{\sigma,n}}(y)| \ll_{a,\varepsilon} \sum_{m=n+1}^{n+k} \frac{\log p_m}{p_m^{2(\sigma-\varepsilon)}} \le \sum_{m=n+1}^{\infty} \frac{\log p_m}{p_m^{2(\sigma-\max_h \theta_{F_h})}} \tag{5-6}$$

for every $n,\ k \ge 1$ and $\sigma \ge 1$, uniformly for $|y| \le a,\ a > 0$. Hence, by Cauchy's criterion, there exists the limit function

$$\widehat{\lambda_\sigma}(y) = \lim_{n\to\infty} \widehat{\lambda_{\sigma,n}}(y)$$

and by (3-14) the convergence is uniform in $|y| \le a$ for every $a > 0$. Therefore, by Lévy's convergence theorem, we have that $\widehat{\lambda_\sigma}(y)$ is the Fourier transform of some distribution function λ_σ, and $\lambda_{\sigma,n} \to \lambda_\sigma$ weakly as $n \to \infty$. Moreover, since the constants in (5-3) are independent of n, we may apply the dominated convergence theorem and thus λ_σ is absolutely continuous for every $\sigma \ge 1$, with continuous (both in σ and x) density $H_\sigma(x)$. Furthermore, since $H_{\sigma,n}(x)$ and $H_\sigma(x)$ are determined by the Fourier inversion formula (see Lemma 3.2), the uniform convergence of $\widehat{\lambda_{\sigma,n}}(y) \to \widehat{\lambda_\sigma}(y)$ and (5-3) imply that $H_{\sigma,n}(x)$ converges, uniformly with respect to both $1 \le \sigma \le M$, $M > 1$, and $x \in \mathbb{C}$, toward $H_\sigma(x)$.

Now, similarly to Theorem 4.1, for $n \ge \tilde{n}$ and $c \in \mathbb{C}$ we have that the Jensen function $\varphi_{L_n-c}(\sigma)$ is twice differentiable with continuous second derivative (cf. [Borchsenius and Jessen 1948, §9])

$$\varphi''_{L_n-c}(\sigma) = 2\pi H_{\sigma,n}(c). \tag{5-7}$$

On the other hand, for any $1 < \sigma_1 < \sigma_2$, by the uniform convergence of $F_{j,n}(s)$, $j = 1, \dots, N$, of Remark 3.1 and by Theorem 6 of [Jessen and Tornehave 1945], we have that

$$\varphi_{L_n-c}(\sigma) \to \varphi_{L-c}(\sigma) \quad \text{as } n \to \infty \tag{5-8}$$

uniformly for $\sigma_1 \le \sigma \le \sigma_2$. By (5-7), (5-8), the convexity of $\varphi_{L_n-c}(\sigma)$ and Theorem 7.17 in [Rudin 1976] we obtain that the Jensen function $\varphi_L(\sigma)$ is twice differentiable with continuous second derivative

$$\varphi''_{L-c}(\sigma) = 2\pi H_\sigma(c).$$

At this point, the same final argument of Theorem 4.1 yields the result. □

6. Dirichlet series with vertical strips without zeros

In this section we collect the proofs of Theorems 1.1, 1.2 and 1.8.

Proof of Theorem 1.1. Since $L_{\boldsymbol{x}}(s)$ is not identically zero, then $\sigma^*(L_{\boldsymbol{x}}) < +\infty$ and hence we fix

$$\sigma_2 > \sigma_1 > \max\Big(\sigma^*(L_{\boldsymbol{x}}),\ \max_{1\le j\le N} \sigma^*(F_j)\Big).$$

Then, by definition of $\sigma^*(L_{\boldsymbol{x}})$ and Theorem 8 of [Jessen and Tornehave 1945], there exists $\varepsilon > 0$ such that $|L_{\boldsymbol{x}}(s)| > \varepsilon$ for $\sigma_1 \le \sigma \le \sigma_2$. Moreover, there exists $M > 0$ such that $|F_j(s)| \le M$ for $\sigma_1 \le \sigma \le \sigma_2$. On the other hand, if we consider the hyperplanes $H(\sigma) = \{\boldsymbol{z} \in \mathbb{C}^N \mid L_{\boldsymbol{z}}(\sigma) = 0\}$ we have

$$\lim_{\sigma\to+\infty} \operatorname{dist}(\boldsymbol{x}, H(\sigma)) = \lim_{\sigma\to+\infty} \frac{|L_{\boldsymbol{x}}(\sigma)|}{\sqrt{\sum_j |F_j(\sigma)|^2}} = 0.$$

Therefore there exists $\beta > \sigma_2$ such that $\operatorname{dist}(\boldsymbol{x}, H(\beta)) < \varepsilon/(4\sqrt{N}M)$. Then for any $\boldsymbol{0} \ne \boldsymbol{c} \in B_{\varepsilon/(2\sqrt{N}M)}(\boldsymbol{x}) \cap H(\beta)$ we have $L_{\boldsymbol{c}}(\beta) = 0$ and, by the triangle and Cauchy–Schwartz inequalities,

$$|L_{\boldsymbol{c}}(s)| \ge |L_{\boldsymbol{x}}(s)| - |L_{\boldsymbol{c}-\boldsymbol{x}}(s)| > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}$$

for $1 \le \sigma^*(L_{\boldsymbol{x}}) < \sigma_1 \le \sigma \le \sigma_2 < \beta \le \sigma^*(L_{\boldsymbol{c}})$. This concludes the proof since $B_{\varepsilon/(2\sqrt{N}M)}(\boldsymbol{x}) \cap H(\beta)$ clearly contains infinitely many projectively inequivalent vectors $\boldsymbol{c}$.

Proof of Theorem 1.2. We write $N = k + 1 \ge 2$. If $N = 2$ then the result follows from Theorem 1.1; so we suppose that $N \ge 3$.
Note that $\boldsymbol{x} \in \mathbb{C}^N$ is such that $L_{\boldsymbol{x}}(\sigma) = 0$ for some $\sigma > 1$ if and only if $\boldsymbol{x} = (x_1, \dots, x_N)$ belongs to the hyperplane

$$F_1(\sigma)x_1 + \cdots + F_N(\sigma)x_N = 0. \tag{6-1}$$

If $\sigma > \max_{1\le j\le N} \sigma^*(F_j) = \tilde{\sigma}_0$, then the space of solutions of (6-1) has dimension $N - 1 \ge 2$ and is generated by

$$v_j^{(1)}(\sigma) = \Big(-\frac{1}{F_1(\sigma)}, 0, \dots, \frac{1}{F_j(\sigma)}, \dots, 0\Big), \quad j = 2, \dots, N.$$

Moreover we define inductively for $h = 2, \dots, N-1$ the vectors

$$v_j^{(h)}(\sigma_1, \dots, \sigma_h) = v_j^{(h-1)}(\sigma_1, \dots, \sigma_{h-1}) - \frac{L_{v_j^{(h-1)}(\sigma_1,\dots,\sigma_{h-1})}(\sigma_h)}{L_{v_h^{(h-1)}(\sigma_1,\dots,\sigma_{h-1})}(\sigma_h)} v_h^{(h-1)}(\sigma_1, \dots, \sigma_{h-1}),$$

$j = h+1, \dots, N$. Note that these are well defined linear combinations of $v_j^{(1)}(\sigma_1)$, $j = 2, \dots, N$, hence solutions of (6-1), if $\sigma_1 > \tilde{\sigma}_0$ and $\sigma_h > \sigma^*(L_{v_h^{(h-1)}(\sigma_1,\dots,\sigma_{h-1})})$,

$h = 2, \ldots, N-1$. Actually, by definition it is clear that, under these conditions, $v_j^{(h)}(\sigma_1, \ldots, \sigma_h)$ is a solution of

$$\begin{gathered} F_1(\sigma_1)x_1 + \cdots + F_N(\sigma_1)x_N = 0, \\ \vdots \\ F_1(\sigma_h)x_1 + \cdots + F_N(\sigma_h)x_N = 0. \end{gathered}$$

Moreover, for any $1 \le m \le N-1$ we consider the vector

$$v_m(\sigma_1, \ldots, \sigma_{m-1}, \infty, \ldots, \infty) = \lim_{\sigma_m \to \infty} \cdots \lim_{\sigma_{N-1} \to \infty} v_N^{(N-1)}(\sigma_1, \ldots, \sigma_{N-1}) \quad (6\text{-}2)$$

and for simplicity we write $v_N(\sigma_1, \ldots, \sigma_{N-1}) = v_N^{(N-1)}(\sigma_1, \ldots, \sigma_{N-1})$. Note that there exists a finite set of explicit conditions on $\sigma_1, \ldots, \sigma_{N-1}$ for which these limits exist, i.e., there exist $\tilde{\sigma}_j$, $j = 1, \ldots, N-1$, which depend only on the Dirichlet series $F_1, \ldots, F_N$, such that $v_m(\sigma_1, \ldots, \sigma_{m-1}, \infty, \ldots, \infty)$ exists for every $1 \le m \le N-1$ if $\sigma_l > \tilde{\sigma}_l$ for every $l = 1, \ldots, N-1$. These conditions actually correspond to the fact that the vector $v_m(\sigma_1, \ldots, \sigma_{m-1}, \infty, \ldots, \infty)$ is a generator of the one-dimensional vector space (by (1-2), reordering the functions if needed) defined by the system

$$\begin{aligned} F_1(\sigma_1)x_1 + \cdots + F_N(\sigma_1)x_N &= 0, \\ &\vdots \\ F_1(\sigma_{m-1})x_1 + \cdots + F_N(\sigma_{m-1})x_N &= 0, \\ a_1(1)x_1 + \cdots + a_N(1)x_N &= 0, \\ &\vdots \\ a_1(N-m)x_1 + \cdots + a_N(N-m)x_N &= 0. \end{aligned}$$

Hence, in particular, this implies that the definition of $v_m(\sigma_1, \ldots, \sigma_{m-1}, \infty, \ldots, \infty)$ is independent from the order of the limits and that $L_{v_m(\sigma_1, \ldots, \sigma_{m-1}, \infty, \ldots, \infty)}(\sigma_l) = 0$, $l = 1, \ldots, m-1$.

We work by induction on $h \in [1, N-2]$. For $h = 1$ we fix

$$\sigma_{1,2} > \sigma_{1,1} > \max\bigl(\sigma^*(L_{v_1(\infty, \ldots, \infty)}), \tilde{\sigma}_0\bigr),$$

and take

$$\varepsilon_1 = \min_{\sigma_{1,1} \le \sigma \le \sigma_{1,2}, t \in \mathbb{R}} |L_{v_1(\infty, \ldots, \infty)}(\sigma + it)| > 0$$

and

$$M_1 = \max_{1 \le j \le N} \max_{\sigma_{1,1} \le \sigma \le \sigma_{1,2}, t \in \mathbb{R}} |F_j(\sigma + it)| < \infty.$$

Note that $M_1 > 0$ by the choice of $\sigma_{1,1}$ and $\sigma_{1,2}$. By (6-2), we can choose $\beta_1 > \sigma_{1,2}$ such that

$$\left\| v_1(\infty, \dots, \infty) - v_2(\beta_1, \infty, \dots, \infty) \right\| < \frac{\varepsilon_1}{2\sqrt{N} M_1}.$$

Then, since $v_2(\beta_1, \infty, \dots, \infty)$ is a solution of (6-1) with $\sigma = \beta_1$, we have that $L_{v_2(\beta_1,\infty,\dots,\infty)}(\beta_1) = 0$. Moreover for $\sigma_{1,1} \le \sigma \le \sigma_{1,2}$ we have, by the triangle and Cauchy–Schwartz inequalities,

$$\begin{aligned} |L_{v_2(\beta_1,\infty,\dots,\infty)}(s)| &\ge |L_{v_1(\infty,\dots,\infty)}(s)| - |L_{v_1(\infty,\dots,\infty) - v_2(\beta_1,\infty,\dots,\infty)}(s)| \\ &\ge \varepsilon_1 - \frac{\varepsilon_1}{2} = \frac{\varepsilon_1}{2} = \delta_1 > 0. \end{aligned}$$

By induction we suppose that for any fixed $1 < h \le N - 2$ there exist

$$\sigma_{1,1} < \sigma_{1,2} < \beta_1 < \cdots < \sigma_{h,1} < \sigma_{h,2} < \beta_h$$

and $\delta_h > 0$ such that

$$\min_{1 \le l \le h} \min_{\sigma_{l,1} < \sigma < \sigma_{l,2}, t \in \mathbb{R}} |L_{v_{h+1}(\beta_1,\dots,\beta_h,\infty,\dots,\infty)}(\sigma + it)| > \delta_h.$$

These hypotheses mean that the Dirichlet series $L_{v_{h+1}(\beta_1,\dots,\beta_h,\infty,\dots,\infty)}(s)$, which vanishes for $s = \beta_1, \dots, \beta_h$, has at least h distinct vertical strips without zeros in the region $1 < \sigma < \sigma^*(L_{v_{h+1}(\beta_1,\dots,\beta_h,\infty,\dots,\infty)})$.

For the inductive step $h \mapsto h + 1$, we take

$$\sigma_{h+1,2} > \sigma_{h+1,1} > \max\left(\sigma^*(L_{v_{h+1}(\beta_1,\dots,\beta_h,\infty,\dots,\infty)}), \max_{h+1 \le j \le N} \sigma^*(L_{v_j^{(h)}(\beta_1,\dots,\beta_h)}), \tilde{\sigma}_h \right),$$

$$\varepsilon_{h+1} = \min\left(\delta_h, \min_{\sigma_{h+1,1} \le \sigma \le \sigma_{h+1,2}, t \in \mathbb{R}} |L_{v_{h+1}(\beta_1,\dots,\beta_h,\infty,\dots,\infty)}(\sigma + it)| \right) > 0$$

and

$$M_{h+1} = \max_{1 \le j \le N} \max_{\sigma_{1,1} \le \sigma \le \sigma_{h+1,2}, t \in \mathbb{R}} |F_j(\sigma + it)| < \infty.$$

Note that since $\sigma_{h+1,1} > \sigma_{1,2}$ we have $M_{h+1} > 0$. Then we choose $\beta_{h+1} > \sigma_{h+1,2}$ such that

$$\begin{aligned} &\left\| v_{h+1}(\beta_1, \dots, \beta_h, \infty, \dots, \infty) - v_{h+2}(\beta_1, \dots, \beta_h, \beta_{h+1}, \infty, \dots, \infty) \right\| \\ &\qquad < \frac{\varepsilon_{h+1}}{2\sqrt{N} M_{h+1}}, \end{aligned}$$

which exists by definition. Moreover, by the triangle and Cauchy–Schwartz inequalities, we have that

$$
\begin{aligned}
&\left|L_{v_{h+2}(\beta_1,\ldots,\beta_{h+1},\infty,\ldots,\infty)}(s)\right| \\
&\quad\geq \left|L_{v_{h+1}(\beta_1,\ldots,\beta_h,\infty,\ldots,\infty)}(s)\right| - \left|L_{v_{h+2}(\beta_1,\ldots,\beta_{h+1},\infty,\ldots,\infty)-v_{h+2}(\beta_1,\ldots,\beta_{h+1},\infty,\ldots,\infty)}(s)\right| \\
&\quad\geq \delta_h - \frac{\varepsilon_{h+1}}{2} \geq \frac{\varepsilon_{h+1}}{2} = \delta_{h+1}
\end{aligned}
$$

for any $\sigma_{l,1} \leq \sigma \leq \sigma_{l,2}$, $l = 1, \ldots, h+1$.

When $h+1 = N-2+1 = N-1$ we have just one vector

$$\boldsymbol{c} = v_N(\beta_1, \ldots, \beta_{N-1}) \in \mathbb{C}^N \setminus \{\boldsymbol{0}\}$$

and the corresponding Dirichlet series $L_{\boldsymbol{c}}(s)$ has, as noted above, at least $N-1$ distinct vertical strips without zeros in the region $1 < \sigma < \sigma^*(L_{\boldsymbol{c}})$.

Proof of Theorem 1.8. For any $j = 1, \ldots, N$, let α_j be a square root of ω_j. Without loss of generality we may suppose that $h = 1$ and $k = 2$. Note that, since $|\omega_j| = 1$ and $\omega_1 \neq \omega_2$ then $\alpha_1 \neq \pm\alpha_2$ and we may suppose $\alpha_1 \notin \mathbb{R}$. It follows that the system of equations

$$
\begin{aligned}
\operatorname{Re}(\alpha_1)x_1 + \cdots + \operatorname{Re}(\alpha_N)x_N &= 0, \\
\operatorname{Im}(\alpha_1)x_1 + \cdots + \operatorname{Im}(\alpha_N)x_N &= 0
\end{aligned}
\tag{6-3}
$$

defines a real vector space V_∞ of dimension $N-2 \geq 1$ which may be written as

$$
V_\infty = \left\{ \left(\sum_{j=3}^{\infty} \left(\frac{\operatorname{Im}(\alpha_2)\operatorname{Im}(\alpha_1\overline{\alpha_j})}{\operatorname{Im}(\alpha_1)\operatorname{Im}(\alpha_1\overline{\alpha_2})} - \frac{\operatorname{Im}(\alpha_j)}{\operatorname{Im}(\alpha_1)} \right) t_j, -\sum_{j=3}^{\infty} \frac{\operatorname{Im}(\alpha_1\overline{\alpha_j})}{\operatorname{Im}(\alpha_1\overline{\alpha_2})} t_j, t_3, \ldots, t_N \right) \;\middle|\; t_3, \ldots, t_N \in \mathbb{R} \right\}.
$$

Let $v_\infty \in V_\infty$ be the vector corresponding to a fixed choice $(\tau_1, \ldots, \tau_N) \in \mathbb{R}^{N-2} \setminus \{\boldsymbol{0}\}$ and $\boldsymbol{c}_0 = (\overline{\alpha_1} v_{\infty,1}, \ldots, \overline{\alpha_N} v_{\infty,N})$. We take $\sigma_2 > \sigma_1 > \max(\sigma^*(L_{\boldsymbol{c}_0}))$, then, by Theorem 8 of [Jessen and Tornehave 1945], there exists $\varepsilon > 0$ such that $|L_{\boldsymbol{c}_0}(s)| > \varepsilon$ for $\sigma_1 \leq \sigma \leq \sigma_2$. Moreover, there exists $M > 0$ such that $|F_j(s)| \leq M$ for $\sigma_1 \leq \sigma \leq \sigma_2$. On the other hand, for any fixed $\sigma > \sigma_2$, the system of equations

$$
\begin{aligned}
\operatorname{Re}(\alpha_1 F_1(\sigma))x_1 + \cdots + \operatorname{Re}(\alpha_N F_N(\sigma))x_N &= 0, \\
\operatorname{Im}(\alpha_1 F_1(\sigma))x_1 + \cdots + \operatorname{Im}(\alpha_N F_N(\sigma))x_N &= 0
\end{aligned}
\tag{6-4}
$$

defines a real vector space V_σ of dimension at least $N-2$. However, since $F_j(\sigma) \to a_j(1) = 1$ as $\sigma \to \infty$, $j = 1, 2$, there exists $\sigma_0 > \sigma_2$ such that V_σ has dimension

$N-2$ for every $\sigma > \sigma_0$ and

$$V_\sigma = \Biggl\{ \Biggl(\sum_{j=3}^{\infty} \biggl(\frac{\operatorname{Im}(\alpha_2 F_2(\sigma)) \operatorname{Im}(\alpha_1 \overline{\alpha_j} F_1(\sigma) \overline{F_j(\sigma)})}{\operatorname{Im}(\alpha_1 F_1(\sigma)) \operatorname{Im}(\alpha_1 \overline{\alpha_2} F_1(\sigma) \overline{F_2(\sigma)})} - \frac{\operatorname{Im}(\alpha_j) F_j(\sigma)}{\operatorname{Im}(\alpha_1) F_1(\sigma)} \biggr) t_j, \\ - \sum_{j=3}^{\infty} \frac{\operatorname{Im}(\alpha_1 \overline{\alpha_j} F_1(\sigma) \overline{F_j(\sigma)})}{\operatorname{Im}(\alpha_1 \overline{\alpha_2} F_1(\sigma) \overline{F_2(\sigma)})} t_j, t_3, \dots, t_N \Biggr) \Biggm| t_3, \dots, t_N \in \mathbb{R} \Biggr\}.$$

Let $v_\sigma \in V_\sigma$ be the vector corresponding to $(\tau_1, \dots, \tau_N)$, then $\|v_\infty - v_\sigma\| \to 0$ as $\sigma \to \infty$. Therefore there exists $\beta > \sigma_0$ such that, taking $\boldsymbol{c} = (\overline{\alpha_1} v_{\beta,1}, \dots, \overline{\alpha_N} v_{\beta,N})$, we have $\|\boldsymbol{c}_0 - \boldsymbol{c}\| < \varepsilon/(2\sqrt{N}M)$. Then by (6-4) we have that $L_{\boldsymbol{c}}(\beta) = 0$ and, by the triangle and Cauchy–Schwartz inequalities, that

$$|L_{\boldsymbol{c}}(s)| \geq |L_{\boldsymbol{c}_0}(s)| - |L_{\boldsymbol{c}-\boldsymbol{c}_0}(s)| > \varepsilon - \frac{\varepsilon}{2} = \frac{\varepsilon}{2}$$

for $1 \leq \sigma^*(L_{\boldsymbol{c}_0}) < \sigma_1 \leq \sigma \leq \sigma_2 < \sigma_0 < \beta \leq \sigma^*(L_{\boldsymbol{c}})$. Moreover

$$\begin{aligned} \Phi(s) &= \sum_{j=1}^{N} \overline{\alpha_j} v_{\beta,j} \Phi_j(s) = \sum_{j=1}^{N} \overline{\alpha_j} v_{\beta,j} \omega_j \overline{\Phi_j(1-\bar{s})} \\ &= \sum_{j=1}^{N} \alpha_j v_{\beta,j} \overline{\Phi_j(1-\bar{s})} = \overline{\Phi(1-\bar{s})}. \end{aligned}$$

Acknowledgments

This paper is part of my Ph.D. thesis at the Department of Mathematics of the University of Genova. I express my sincere gratitude to my supervisor Professor Alberto Perelli for his support and advice. I thank Professor Giuseppe Molteni for carefully reading the manuscript and suggesting several improvements. I also wish to thank Professor Enrico Bombieri for an enlightening discussion on this topic and for helpful comments concerning the paper. Finally, I would like to thank the referees for pointing out some inaccuracies and for the many suggestions which greatly improved the exposition.

References

[Avdispahić and Smajlović 2010] M. Avdispahić and L. Smajlović, "On the Selberg orthogonality for automorphic L-functions", *Arch. Math.* (*Basel*) **94**:2 (2010), 147–154. MR Zbl

[Bogachev 2007] V. I. Bogachev, *Measure theory*, vols. I and II, Springer, 2007. MR Zbl

[Bohr and Jessen 1930] H. Bohr and B. Jessen, "Über die Werteverteilung der Riemannschen Zetafunktion, I", *Acta Math.* **54**:1 (1930), 1–35. MR Zbl

[Bohr and Jessen 1932] H. Bohr and B. Jessen, "Über die Werteverteilung der Riemannschen Zetafunktion, II", *Acta Math.* **58**:1 (1932), 1–55. MR Zbl

[Bombieri and Gosh 2011] E. Bombieri and A. Gosh, "Around the Davenport–Heilbronn function", *Uspekhi Mat. Nauk* **66**:2(398) (2011), 15–66. In Russian; translated in *Russian Math. Surveys* **66**:2 (2011), 221–270. MR Zbl

[Bombieri and Hejhal 1995] E. Bombieri and D. A. Hejhal, "On the distribution of zeros of linear combinations of Euler products", *Duke Math. J.* **80**:3 (1995), 821–862. MR Zbl

[Bombieri and Mueller 2008] E. Bombieri and J. Mueller, "On the zeros of certain Epstein zeta functions", *Forum Math.* **20**:2 (2008), 359–385. MR Zbl

[Booker and Thorne 2014] A. R. Booker and F. Thorne, "Zeros of L-functions outside the critical strip", *Algebra Number Theory* **8**:9 (2014), 2027–2042. MR Zbl

[Borchsenius and Jessen 1948] V. Borchsenius and B. Jessen, "Mean motions and values of the Riemann zeta function", *Acta Math.* **80** (1948), 97–166. MR Zbl

[Cassels 1961] J. W. S. Cassels, "Footnote to a note of Davenport and Heilbronn", *J. London Math. Soc.* **36** (1961), 177–184. MR Zbl

[Conrey and Ghosh 1994] J. B. Conrey and A. Ghosh, "Turán inequalities and zeros of Dirichlet series associated with certain cusp forms", *Trans. Amer. Math. Soc.* **342**:1 (1994), 407–419. MR Zbl

[Davenport and Heilbronn 1936a] H. Davenport and H. Heilbronn, "On the Zeros of Certain Dirichlet Series", *J. London Math. Soc.* **S1-11**:3 (1936), 181–185. MR Zbl

[Davenport and Heilbronn 1936b] H. Davenport and H. Heilbronn, "On the Zeros of Certain Dirichlet Series, II", *J. London Math. Soc.* **S1-11**:4 (1936), 307–312. MR Zbl

[Gonek 1981] S. M. Gonek, "The zeros of Hurwitz's zeta function on $\sigma = \frac{1}{2}$", pp. 129–140 in *Analytic number theory* (Philadelphia, PA, 1980), Lecture Notes in Math. **899**, Springer, 1981. MR Zbl

[Goodman 1957] A. W. Goodman, "Univalent functions and nonanalytic curves", *Proc. Amer. Math. Soc.* **8** (1957), 598–601. MR Zbl

[Iwaniec and Sarnak 2000] H. Iwaniec and P. Sarnak, "Perspectives on the analytic theory of L-functions", *Geom. Funct. Anal.* Special Volume, Part II (2000), 705–741. MR Zbl

[Jessen and Tornehave 1945] B. Jessen and H. Tornehave, "Mean motions and zeros of almost periodic functions", *Acta Math.* **77** (1945), 137–279. MR Zbl

[Jessen and Wintner 1935] B. Jessen and A. Wintner, "Distribution functions and the Riemann zeta function", *Trans. Amer. Math. Soc.* **38**:1 (1935), 48–88. MR Zbl

[Kaczorowski 2006] J. Kaczorowski, "Axiomatic theory of L-functions: the Selberg class", pp. 133–209 in *Analytic number theory*, Lecture Notes in Math. **1891**, Springer, 2006. MR Zbl

[Kaczorowski and Perelli 1999] J. Kaczorowski and A. Perelli, "The Selberg class: a survey", pp. 953–992 in *Number theory in progress, II* (Zakopane Kościelisko, Poland, 1997), de Gruyter, Berlin, 1999. MR Zbl

[Kaczorowski and Perelli 2000] J. Kaczorowski and A. Perelli, "On the structure of the Selberg class, III: Sarnak's rigidity conjecture", *Duke Math. J.* **101**:3 (2000), 529–554. MR Zbl

[Kaczorowski et al. 2007] J. Kaczorowski, G. Molteni, and A. Perelli, "Some remarks on the unique factorization in certain semigroups of classical L-functions", *Funct. Approx. Comment. Math.* **37**:Part 2 (2007), 263–275. MR Zbl

[Kaczorowski et al. 2010] J. Kaczorowski, G. Molteni, and A. Perelli, "A converse theorem for Dirichlet L-functions", *Comment. Math. Helv.* **85**:2 (2010), 463–483. MR Zbl

[Karatsuba and Voronin 1992] A. A. Karatsuba and S. M. Voronin, *The Riemann zeta-function*, De Gruyter Expositions in Mathematics **5**, de Gruyter, Berlin, 1992. MR Zbl

[Kershner 1936] R. Kershner, "On the Addition of Convex Curves", *Amer. J. Math.* **58**:4 (1936), 737–746. MR Zbl

[Lee 2014] Y. Lee, "On the zeros of Epstein zeta functions", *Forum Math.* **26**:6 (2014), 1807–1836. MR Zbl

[Lee et al. 2017] Y. Lee, T. Nakamura, and Ł. Pańkowski, "Selberg's orthonormality conjecture and joint universality of L-functions", *Math. Z.* **286**:1-2 (2017), 1–18. MR

[Liu and Ye 2005] J. Liu and Y. Ye, "Selberg's orthogonality conjecture for automorphic L-functions", *Amer. J. Math.* **127**:4 (2005), 837–849. MR Zbl

[PARI/GP 2016] PARI/GP *version* `2.7.0`, The PARI Group, Univ. Bordeaux, 2016, Available at http://pari.math.u-bordeaux.fr/.

[Perelli 2005] A. Perelli, "A survey of the Selberg class of L-functions, I", *Milan J. Math.* **73** (2005), 19–52. MR Zbl

[Righetti 2016a] M. Righetti, "Zeros of combinations of Euler products for $\sigma > 1$", *Monatsh. Math.* **180**:2 (2016), 337–356. MR Zbl

[Righetti 2016b] M. Righetti, *Zeros of combinations of Euler products for* $\sigma > 1$, Ph.D. thesis, Università di Genova, 2016.

[Rudin 1976] W. Rudin, *Principles of mathematical analysis*, 3rd ed., McGraw–Hill, New York, 1976. MR Zbl

[Rudnick and Sarnak 1996] Z. Rudnick and P. Sarnak, "Zeros of principal L-functions and random matrix theory", *Duke Math. J.* **81**:2 (1996), 269–322. MR Zbl

[Saias and Weingartner 2009] E. Saias and A. Weingartner, "Zeros of Dirichlet series with periodic coefficients", *Acta Arith.* **140**:4 (2009), 335–344. MR Zbl

[Selberg 1992] A. Selberg, "Old and new conjectures and results about a class of Dirichlet series", pp. 367–385 in *Proceedings of the Amalfi Conference on Analytic Number Theory* (Maiori, Italy, 1989), Univ. Salerno, 1992. MR Zbl

[Titchmarsh 1975] E. C. Titchmarsh, *The theory of functions*, Reprint of the 2nd (1939) ed., Oxford University Press, 1975. MR Zbl

[Titchmarsh 1986] E. C. Titchmarsh, *The theory of the Riemann zeta-function*, 2nd ed., Clarendon Press, New York, 1986. MR Zbl

[Yamashita 1982] S. Yamashita, "Radii of univalence, starlikeness, and convexity", *Bull. Austral. Math. Soc.* **25**:3 (1982), 453–457. MR Zbl

Communicated by Peter Sarnak
Received 2016-11-21 Revised 2017-07-18 Accepted 2017-08-14

Dipartimento di Matematica, Università di Genova, 16146, Genova, Italy
righetti@crm.umontreal.ca *Current address: Centre de Recherches Mathématiques, Université de Montréal, P.O. Box 6128, Centre-Ville Station, Montréal QC H3C 3J7, Canada*

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2165

Adams operations on matrix factorizations

Michael K. Brown, Claudia Miller, Peder Thompson and Mark E. Walker

We define Adams operations on matrix factorizations, and we show these operations enjoy analogues of several key properties of the Adams operations on perfect complexes with support developed by Gillet and Soulé. As an application, we give a proof of a conjecture of Dao and Kurano concerning the vanishing of Hochster's θ pairing.

1. Introduction

We establish a theory of Adams operations on the Grothendieck group of matrix factorizations and use these operations to prove a conjecture of Dao and Kurano [2014, Conjecture 3.1(2)] concerning the vanishing of Hochster's θ pairing for a pair of modules defined on an isolated hypersurface singularity.

Let Q be a commutative Noetherian ring and let $f \in Q$. A *matrix factorization* of f in Q is a $\mathbb{Z}/2$-graded, finitely generated projective Q-module $P = P_0 \oplus P_1$, equipped with an odd degree Q-linear endomorphism d satisfying $d^2 = f\,\mathrm{id}_P$. In other words, a matrix factorization is a pair of maps of finitely generated projective Q-modules, $(\alpha : P_1 \to P_0, \beta : P_0 \to P_1)$, satisfying $\alpha\beta = f\,\mathrm{id}_{P_0}$ and $\beta\alpha = f\,\mathrm{id}_{P_1}$.

When $f = 0$, a matrix factorization of f is the same thing as a $\mathbb{Z}/2$-graded complex of finitely generated projective Q-modules. In this case, we have the evident $\mathbb{Z}/2$-graded analogues of chain maps and homotopies of such. These, in fact, generalize to an arbitrary f. The matrix factorizations of $f \in Q$ form the objects of a category $\mathrm{mf}(Q, f)$, in which a morphism between objects P and P' of $\mathrm{mf}(Q, f)$ is a degree zero Q-linear map $g : P \to P'$ such that $d_{P'} \circ g = g \circ d_P$. In other words, a morphism is a pair of maps $g_0 : P_0 \to P_0'$ and $g_1 : P_1 \to P_1'$ causing the evident pair of squares to commute. A *homotopy* joining morphisms $g_1, g_2 : P \to P'$ in $\mathrm{mf}(Q, f)$ is a Q-linear map $h : P \to P'$ of odd degree such that $d_{P'}h + hd_P = g_1 - g_2$. The *homotopy category* of $\mathrm{mf}(Q, f)$ is the category $[\mathrm{mf}(Q, f)]$ obtained from $\mathrm{mf}(Q, f)$ by identifying homotopic morphisms. It is well-known that, when Q is

This work was partially supported by a grant from the Simons Foundation (#318705 for Mark Walker) and grants from the National Science Foundation (NSF Award DMS-0838463 for Michael Brown and Peder Thompson and NSF Award DMS-1003384 for Claudia Miller).
MSC2010: primary 13D15; secondary 13D02, 13D09, 13D22.
Keywords: Adams operations, matrix factorizations, Hochster's theta pairing.

regular and f is not a zero divisor, $[\mathrm{mf}(Q, f)]$ may be equipped with a canonical triangulated structure (see, for instance, [Orlov 2004] Section 3.1).

Much of the interest in matrix factorizations arises from the following result. For a Noetherian ring R, let $D^b(R)$ denote the bounded derived category of R. Objects of $D^b(R)$ are bounded complexes of finitely generated R-modules, and morphisms are obtained from chain maps by inverting the collection of quasiisomorphisms. Let $\mathrm{Perf}(R)$ denote the full triangulated subcategory of $D^b(R)$ consisting of bounded complexes of finitely generated and projective R-modules, and let $D_{\mathrm{sing}}(R)$ denote the Verdier quotient $D^b(R)/\mathrm{Perf}(R)$, called the *singularity category* of R. The following theorem is essentially due to work of Buchweitz [1986] and Eisenbud [1980]; this particular formulation of the result is proven by Orlov.

Theorem 1 [Orlov 2004, Theorem 3.9] . *If Q is regular and f is not a zero divisor, there is an equivalence of triangulated categories*

$$[\mathrm{mf}(Q, f)] \xrightarrow{\sim} D_{\mathrm{sing}}(Q/(f))$$

determined by sending a matrix factorization $(\alpha : P_1 \to P_0, \beta : P_0 \to P_1)$ *to* $\mathrm{coker}(\alpha)$.

Remark 1.1. In [Orlov 2004], Orlov assumes Q contains a field and has finite Krull dimension, but these assumptions are in fact not needed for this theorem to hold.

Let $R := Q/(f)$. Under the assumptions of Theorem 1, the Grothendieck group $K_0(\mathrm{mf}(Q, f))$ of the triangulated category $[\mathrm{mf}(Q, f)]$ is isomorphic to the quotient $G_0(R)/(\mathrm{im}(K_0(R) \to G_0(R)))$. So, defining a notion of Adams operations on $K_0(\mathrm{mf}(Q, f))$, in this setting, amounts to defining such operations on this quotient.

For a closed subset Z of $\mathrm{Spec}(Q)$, define $\mathcal{P}^Z(Q)$ to be the category of bounded complexes of finitely generated and projective Q-modules whose homology is supported on Z. Gillet–Soulé define lambda and Adams operations on the Grothendieck group $K_0^Z(Q) := K_0(\mathcal{P}^Z(Q))$ [Gillet and Soulé 1987, Sections 3 and 4]. It is tempting to mimic their approach to define Adams operations on $K_0(\mathrm{mf}(Q, f))$, since $\mathrm{mf}(Q, f)$ is somewhat analogous to $\mathcal{P}^{V(f)}(Q)$. But their construction relies on the Dold–Kan correspondence relating $\mathbb{N}$-graded complexes to simplicial modules; since matrix factorizations are $\mathbb{Z}/2$-graded, such an approach is not available for $K_0(\mathrm{mf}(Q, f))$.

Instead, we model our approach after the construction of the *cyclic Adams operations* ψ^p_{cyc} on $K_0^Z(Q)$ developed by the authors in [BMTW 2017] (see also [Atiyah 1966; Haution 2009; Köck 1997]). Let us give a brief summary of the construction of the operations ψ^p_{cyc} and some of their properties.

Fix a prime p. We assume that p is invertible in Q and that Q contains all p-th roots of unity (when Q is local, the case of primary interest to us, we can find such a prime p, at least after passing to a faithfully flat extension of Q). For a perfect complex of Q-modules X, let $T^p(X)$ denote the p-th tensor power of X, which comes equipped with a canonical left action by the symmetric group Σ_p. For a p-th root of unity $w \in Q$, set $T^p(X)^{(w)}$ to be the eigenspace of eigenvalue w for the action of the p-cycle $(1\,2\,\cdots\,p)$ on $T^p(X)$. We define

$$\psi^p_{\mathrm{cyc}}(X) = [T^p(X)^{(1)}] - [T^p(X)^{(\zeta)}]$$

where ζ is a primitive p-th root of unity.

In Sections 2 and 3 of [BMTW 2017], it is established that this formula induces a well-defined operation on $K_0^Z(Q)$ (see also [Haution 2009]). In fact, by Corollary 6.14 of [BMTW 2017], if $p!$ is invertible in Q, then ψ^p_{cyc} agrees with the p-th Adams operation on $K_0^Z(Q)$ defined by Gillet–Soulé. More generally, we have:

Theorem 2 [BMTW 2017, Theorem 3.7] . *If p is a prime, and Q contains $1/p$ and all the p-th roots of unity, then the action of ψ^p_{cyc} on $K_0^Z(Q)$ satisfies the four Gillet–Soulé axioms defining a degree p Adams operation.*

We refer the reader to Theorem 3.7 of [BMTW 2017] for a precise statement of the four Gillet–Soulé axioms. A consequence of Theorem 2 is that the action of ψ^p_{cyc} on $K_0^Z(Q)_{\mathbb{Q}} := K_0^Z(Q) \otimes \mathbb{Q}$ is diagonalizable: there is a “weight decomposition”

$$K_0^Z(Q)_{\mathbb{Q}} = \bigoplus_{i=c}^{d} K_0^Z(Q)^{(i)}_{\mathbb{Q}},$$

where $K_0^Z(Q)^{(i)}_{\mathbb{Q}}$ is the eigenspace of ψ^p_{cyc} of eigenvalue p^i, and c is the codimension of Z [loc. cit., Corollary 3.12].

In Section 2, we use the operations ψ^p_{cyc} as a model to construct cyclic Adams operations ψ^p_{cyc} on the Grothendieck group $K_0(\mathrm{mf}(Q, f))$, as well as more general versions for matrix factorizations with a support condition. In Theorem 2.10 and Proposition 2.13, we prove:

Theorem 3. *If p is prime, and Q contains $1/p$ and all the p-th roots of unity, the operator ψ^p_{cyc} on $K_0(\mathrm{mf}(Q, f))$ satisfies the evident analogues of the four Gillet–Soulé axioms for a p-th Adams operation.*

Moreover, if Q is regular and $f \in Q$ is not a zero divisor, the canonical surjection

$$K_0^{V(f)}(Q) \twoheadrightarrow K_0(\mathrm{mf}(Q, f))$$

is compatible with the action of ψ^p_{cyc}.

For Q regular, f not a zero divisor, and $R = Q/(f)$, given a finitely generated R-module M, let $[M]_{\mathrm{stable}} \in K_0(\mathrm{mf}(Q, f))$ denote the image of $[M] \in G_0(R)$ under the canonical surjection $G_0(R) \twoheadrightarrow K_0(\mathrm{mf}(Q, f))$ given by Theorem 1.

Corollary 4. *Assume Q is a regular ring containing $1/p$ and all the p-th roots of unity for some prime p, and suppose $f \in Q$ is not a zero divisor. The action of ψ^p_{cyc} induces an eigenspace decomposition*

$$K_0(\mathrm{mf}(Q, f))_{\mathbb{Q}} = \bigoplus_{i=1}^{d} K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}}.$$

Moreover, if M is a finitely generated R-module, then

$$[M]_{\mathrm{stable}} \in \bigoplus_{i=\operatorname{codim}_R M+1}^{d} K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}}.$$

In Section 3, we give an application of the above results. For the rest of this introduction, assume Q is a regular local ring with maximal ideal $\mathfrak{m}$, and assume f is a nonzero element of $\mathfrak{m}$. Assume also that $R = Q/(f)$ is an isolated singularity; that is, $R_{\mathfrak{p}}$ is regular for all $\mathfrak{p} \in \operatorname{Spec}(R) \setminus \{\mathfrak{m}\}$. Then for any pair of finitely generated R-modules (M, N), we have

$$\operatorname{Tor}^R_i(M, N) \cong \operatorname{Tor}^R_{i+2}(M, N) \quad \text{and} \quad \operatorname{length} \operatorname{Tor}^R_i(M, N) < \infty$$

for $i \gg 0$. This motivates the following definition.

Definition 1.2. With Q, f, R as above, for a pair of finitely generated R-modules (M, N), set

$$\theta_R(M, N) = \operatorname{length}(\operatorname{Tor}^R_{2i}(M, N)) - \operatorname{length}(\operatorname{Tor}^R_{2i+1}(M, N))$$

for $i \gg 0$.

The pairing $\theta_R(-,-)$ is called *Hochster's theta pairing*, since it first appeared in work of Hochster [1981]. The theta pairing should be regarded as the analogue, for the singularity category $D_{\mathrm{sing}}(R)$, of the intersection multiplicity pairing that occurs, for example, in Serre's multiplicity conjectures. There has been much recent work on better understanding the theta pairing, including when it vanishes and how it relates to more classical invariants. Buchweitz and van Straten [2012] show that, for complex isolated hypersurface singularities, the theta pairing can be recovered from the linking form on the link of an isolated singularity. In the same setting, Polishchuk and Vaintrob [2012] relate it to the classical residue pairing using the boundary bulk map. It was conjectured by Dao that θ vanishes for all isolated hypersurface singularities R such that $\dim(R)$ is even, and this has now been proven in almost all cases; see [Moore et al. 2011; Buchweitz and Van Straten 2012; Polishchuk and Vaintrob 2012; Walker 2017]. We refer the reader to Section 3 of [Dao and Kurano 2014] for additional history of the theta pairing and a list of several other conjectures.

One such conjecture, [Dao and Kurano 2014, Conjecture 3.1(2)], is an analogue of Serre's vanishing conjecture (see the remark on page 111 of [Serre 2000]). This conjecture was proven by Dao in the case where R is excellent and contains a field, using a geometric approach [Dao 2013, Theorem 3.5]. As an application of the properties of Adams operations on matrix factorizations that we establish in Section 2, we prove this conjecture in full generality:

Theorem 5 (see Theorem 3.19). *Let $(Q, \mathfrak{m})$ be a regular local ring and $f \in \mathfrak{m}$ with $f \neq 0$. Suppose that $R = Q/(f)$ is an isolated singularity. If M and N are finitely generated R-modules such that*

$$\dim M + \dim N \leq \dim R$$

then $\theta_R(M, N) = 0$.

We close this introduction with a sketch of our proof of Theorem 5. We easily reduce to the case where there is a prime p such that Q contains $1/p$ and all p-th roots of unity. Given a matrix factorization $P = (\alpha : P_1 \to P_0, \beta : P_0 \to P_1)$ of f, one may obtain a matrix factorization P° of $-f$ by negating β. In Proposition 3.18, we show

$$\theta_R(M, N) = \chi([M]_{\text{stable}} \cup [N]^\circ_{\text{stable}}),$$

where $- \cup -$ is the pairing induced by tensor product of matrix factorizations, and χ denotes the Euler characteristic. The assumptions ensure that $[M]_{\text{stable}} \cup [N]^\circ_{\text{stable}}$ is a class in $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))$, the Grothendieck group of $\mathbb{Z}/2$-graded complexes of finitely generated projective Q-modules with finite length homology, so that χ is well-defined. By Corollary 4 and the linearity of χ, we may assume that the classes $[M]_{\text{stable}}$ and $[N]_{\text{stable}}$ lie in eigenspaces $K_0(\text{mf}(Q, 0))^{(i)}_{\mathbb{Q}}$ and $K_0(\text{mf}(Q, 0))^{(j)}_{\mathbb{Q}}$, respectively, where $i + j > d = \dim Q$. By properties of the operations ψ^p_{cyc} established in Theorem 3, $[M]_{\text{stable}} \cup [N]^\circ_{\text{stable}} \in K_0(\text{mf}^{\mathfrak{m}}(Q, 0))^{(i+j)}_{\mathbb{Q}}$.

At this point, one would like to argue that $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))_{\mathbb{Q}} = K_0(\text{mf}^{\mathfrak{m}}(Q, 0))^{(d)}_{\mathbb{Q}}$, which would force $[M]_{\text{stable}} \cup [N]^\circ_{\text{stable}} = 0$. Indeed, one might expect $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))$ to be generated by the $\mathbb{Z}/2$-folding of the class of the Koszul complex on a regular sequence of generators of $\mathfrak{m}$, which lies in $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))^{(d)}$ by the axioms in Theorem 3; this would be parallel to what occurs for bounded $\mathbb{Z}$-graded complexes. The proof of Theorem 5 sketched here would then be almost exactly the same as Gillet and Soulé's proof of Serre's vanishing conjecture.

We are not able to prove $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))$ is generated by the Koszul complex, and indeed we have come to suspect this might be false (see Example 3.6). Fortunately, for the proof of Dao and Kurano's conjecture, one needs only the weaker property that there is an equality of maps $\chi \circ \psi^p_{\text{cyc}} = p^d \chi$ from $K_0(\text{mf}^{\mathfrak{m}}(Q, 0))$ to $\mathbb{Z}$; we prove this in Theorem 3.8.

2. Adams operations on matrix factorizations

In this section, we define cyclic Adams operations on matrix factorizations, closely following the construction of cyclic Adams operations on perfect complexes with support found in Sections 2 and 3 of [BMTW 2017]. We prove these operations enjoy analogues of many of the key properties of the operations on perfect complexes with support constructed in [loc. cit.].

2A. ***Construction.*** Let Q be a Noetherian commutative ring, $f \in Q$ any element (including possibly $f = 0$), and G a finite group. Let $\mathrm{mf}(Q, f; G)$ be the category of *G-equivariant matrix factorizations*. When G is the trivial group, this is the category described in the introduction. More generally, an object of $\mathrm{mf}(Q, f; G)$ is an object P of $\mathrm{mf}(Q, f)$ equipped with a G-action (i.e., a group homomorphism $G \to \mathrm{Aut}_{\mathrm{mf}(Q,f)}(P)$), and a morphism is a G-equivariant morphism of matrix factorizations.

The category $\mathrm{mf}(Q, f; G)$ is an exact category, with the notion of exactness given degree-wise in the evident manner.

Remark 2.1. We could equivalently define an object of $\mathrm{mf}(Q, f; G)$ to consist of a pair of $Q[G]$-modules P_0 and P_1 that are finitely generated and projective as Q-modules, together with a pair of morphisms of $Q[G]$-modules, $(\alpha : P_1 \to P_0, \beta : P_0 \to P_1)$, such that $\alpha\beta$ and $\beta\alpha$ are each multiplication by f (which is central in $Q[G]$). Moreover, if $|G|$ is invertible in Q, we have $\mathrm{mf}(Q, f; G) = \mathrm{mf}(Q[G], f)$.

Example 2.2. If $f = 0$ (and G is trivial), $\mathrm{mf}(Q, 0)$ is the category of $\mathbb{Z}/2$-graded complexes of finitely generated projective Q-modules, with morphisms being chain maps.

A *homotopy* joining morphisms $g_1, g_2 : P \to P'$ in $\mathrm{mf}(Q, f; G)$ is defined just as in the introduction, with the added condition that it be G-equivariant. In detail, it is a Q-linear, G-equivariant map $h : P \to P'$ of degree 1 such that $d_{P'}h + hd_P = g_1 - g_2$. The *homotopy category* of $\mathrm{mf}(Q, f; G)$ is the category $[\mathrm{mf}(Q, f; G)]$ obtained from $\mathrm{mf}(Q, f; G)$ by identifying homotopic morphisms.

Given a ring homomorphism $Q \to Q'$ sending f to f', there is an evident functor $\mathrm{mf}(Q, f; G) \to \mathrm{mf}(Q', f'; G)$ given by extension of scalars along $Q \to Q'$. When $Q' = Q_{\mathfrak{p}}$ for $\mathfrak{p} \in \mathrm{Spec}(Q)$, we write this functor as $P \mapsto P_{\mathfrak{p}}$.

For an object $P \in \mathrm{mf}(Q, f; G)$, define the *support* of P to be

$$\mathrm{supp}(P) = \{\mathfrak{p} \in \mathrm{Spec}(Q) \mid P_{\mathfrak{p}} \text{ is not homotopy equivalent to 0 in } \mathrm{mf}(Q_{\mathfrak{p}}, f; G)\}.$$

Given a closed subset Z of $\mathrm{Spec}(Q)$, define $\mathrm{mf}^Z(Q, f; G)$ to be the full subcategory of $\mathrm{mf}(Q, f)$ consisting of objects P satisfying $\mathrm{supp}(P) \subseteq Z$. Note that $\mathrm{mf}^Z(Q, f; G)$ is a full, exact subcategory of $\mathrm{mf}(Q, f; G)$, and $[\mathrm{mf}^Z(Q, f; G)]$ is a full subcategory of $[\mathrm{mf}(Q, f; G)]$.

We will mainly use the notion of supports for matrix factorizations when $f = 0$ and G is trivial, in which case objects of $\mathrm{mf}(Q, 0)$ are ($\mathbb{Z}/2$-graded) complexes. One must be careful in this situation not to conflate the notion of being homotopy equivalent to 0 with being acyclic. The former implies the latter, but the latter does not imply the former in general. These conditions are equivalent, however, in the following case:

Lemma 2.3. *If Q is a regular ring, an object $P \in \mathrm{mf}(Q, 0)$ is contractible if and only if $H_0(P) = H_1(P) = 0$.*

Proof. Suppose $P = (\alpha_0 : P_0 \to P_1, \alpha_1 : P_1 \to P_0)$ is acyclic, and set $M = \ker(\alpha_1) = \mathrm{im}(\alpha_0)$ and $N = \ker(\alpha_0) = \mathrm{im}(\alpha_1)$. We claim that M and N are projective. It suffices to prove $M_{\mathfrak{p}}$ and $N_{\mathfrak{p}}$ are free for all primes $\mathfrak{p}$. Since

$$0 \to M_{\mathfrak{p}} \to (P_1)_{\mathfrak{p}} \to (P_0)_{\mathfrak{p}} \to (P_1)_{\mathfrak{p}} \to \cdots$$

is exact, we see that, for any d, $M_{\mathfrak{p}}$ is a d-th syzygy of some other $Q_{\mathfrak{p}}$-module. Taking $d > \dim(Q_{\mathfrak{p}})$ gives that $M_{\mathfrak{p}}$ is free. Similarly, N is projective.

Choose splittings $\pi_0 : P_0 \to N$ and $\pi_1 : P_1 \to M$ of the inclusions $N \hookrightarrow P_0$ and $M \hookrightarrow P_1$. Define $A : P_0 \to N \oplus M$ and $B : P_1 \to N \oplus M$ to be given by $\binom{\pi_0}{\alpha_0}$ and $\binom{\alpha_1}{\pi_1}$, respectively. Set $E := \left(\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right)$ and $F := \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$.

We have the following isomorphism of matrix factorizations

$$\begin{tikzcd} P_0 \arrow[r, "\alpha_0"] \arrow[d, "A"'] & P_1 \arrow[r, "\alpha_1"] \arrow[d, "B"'] & P_0 \arrow[d, "A"'] \\ N \oplus M \arrow[r, "E"] & N \oplus M \arrow[r, "F"] & N \oplus M \end{tikzcd}$$

and the bottom matrix factorization is clearly contractible. □

Remark 2.4. When Q is regular, f is not a zero divisor, and G is trivial, the support of any object of $\mathrm{mf}(Q, f)$ is a subset of

$$\mathrm{Sing}(R) := \{\mathfrak{p} \in \mathrm{Spec}(R) \mid R_{\mathfrak{p}} \text{ is not regular}\}$$

where $R = Q/(f)$, and where we identify $\mathrm{Spec}\, R$ with its image in $\mathrm{Spec}\, Q$. Thus, in this case, we have

$$\mathrm{mf}(Q, f) = \mathrm{mf}^{\mathrm{Sing}(R)}(Q, f).$$

Eventually, we will be making the additional assumption that R is an isolated singularity, meaning Q, and hence R, is local, and $\mathrm{Sing}(R) = \{\mathfrak{m}\}$.

Define the Grothendieck group $K_0(\mathrm{mf}^Z(Q, f; G))$ to be the abelian monoid given by isomorphism classes of objects of $\mathrm{mf}^Z(Q, f; G)$ under the operation of direct sum, modulo the relations $[P] = [P'] + [P'']$ if there exists a short exact sequence $0 \to P' \to P \to P'' \to 0$ and $[P] = [P']$ if P and P' are homotopy equivalent. As with the K-theory of complexes, $K_0(\mathrm{mf}^Z(Q, f; G))$ is an abelian group, since $[P]+[\Sigma(P)]=0$, where $\Sigma(P)$ denotes the suspension of P.

For $P \in \mathrm{mf}(Q, f; G)$ and $P' \in \mathrm{mf}(Q, f'; G')$, the tensor product $P \otimes_Q P'$ is the usual tensor product of Q-modules, with grading determined by $|p \otimes p'| = |p|+|p'|$ and differential $\partial(p\otimes p') = d_P(p)\otimes p' + (-1)^{|p|} p \otimes d_{P'}(p')$. The group $G\times G'$ acts in the evident manner, and the resulting object belongs to $\mathrm{mf}(Q, f+f'; G\times G')$, since ∂^2 is multiplication by $f+f'$. Note, in particular, that the n-th tensor power of an object of $\mathrm{mf}(Q, f)$ belongs to $\mathrm{mf}(Q, nf)$.

We proceed to define cyclic Adams operations on $K_0(\mathrm{mf}^Z(Q, f))$. The construction is closely parallel to that for $K_0^Z(Q)$ given in [BMTW 2017], with one minor exception: the need to "divide by p".

For an integer $n \geq 1$, we define a functor

$$T^n : \mathrm{mf}^Z(Q, f) \to \mathrm{mf}^Z(Q, nf; \Sigma_n)$$

given, on objects, by sending $P \in \mathrm{mf}^Z(Q, f)$ to the matrix factorization

$$T^n(P) = \overbrace{P \otimes_Q \cdots \otimes_Q P}^{n \text{ times}}$$

equipped with the left action of Σ_n given by

$$\sigma(p_1 \otimes \cdots \otimes p_n) = \pm p_{\sigma^{-1}(1)} \otimes \cdots \otimes p_{\sigma^{-1}(n)}.$$

The sign is uniquely determined by the following rule: if σ is the transposition $(i\ i+1)$ for some $1 \leq i \leq n-1$ and $p_1, \ldots, p_n$ are homogenous elements of P, then

$$\sigma(p_1 \otimes \cdots \otimes p_n) = (-1)^{|p_i||p_{i+1}|} k p_1 \otimes \cdots p_{i-1} \otimes p_{i+1} \otimes p_i \otimes p_{i+2} \otimes \cdots \otimes p_n.$$

The rule for morphisms is the evident one.

Following Section 2 of [BMTW 2017], for any i and j, let $\Sigma_{i,j}$ be the image of the canonical homomorphism $\Sigma_i \times \Sigma_j \hookrightarrow \Sigma_{i+j}$, and define a pairing

$$\star_{i,j} : K_0(\mathrm{mf}^Z(Q, if); \Sigma_i) \times K_0(\mathrm{mf}^Z(Q, jf); \Sigma_j) \to K_0(\mathrm{mf}^Z(Q, (i+j)f); \Sigma_{i+j})$$

induced by the bifunctor $(P, P') \mapsto Q[\Sigma_{i+j}] \otimes_{Q[\Sigma_{i,j}]} P \otimes_Q P'$. This pairing is well-defined, commutative, and associative, by an argument identical to the proof of Lemma 2.4 in [loc. cit.].

The proof of Theorem 2.2 in [loc. cit.] also holds nearly verbatim for matrix factorizations and leads to a proof of:

Theorem 2.5. *For a commutative Noetherian ring Q, closed subset Z of* $\mathrm{Spec}(Q)$, *element $f \in Q$, and integer $n \geq 1$, there is a function*

$$t_\Sigma^n : K_0(\mathrm{mf}^Z(Q, f)) \to K_0(\mathrm{mf}^Z(Q, nf; \Sigma_n))$$

such that, for an object $P \in \mathrm{mf}^Z(Q, f)$, *we have*

$$t^n_\Sigma([P]) = [T^n(P)].$$

Remark 2.6. As in [BMTW 2017, §5], if k is a positive integer such that $k!$ is invertible in Q, then one can use Theorem 2.5 to establish an operation λ^k on $K_0(\mathrm{mf}^Z(Q, f))$ that is induced from the k-th exterior power functor. Since we won't use such operations in this paper, we omit the details.

We now assume p is a prime that is invertible in Q, and we define C_p to be the subgroup of Σ_p generated by the p-cycle $(1\,2\,\cdots\,p)$. For any p-th root of unity ζ belonging to Q (including the case $\zeta = 1$), let Q_ζ denote the $Q[C_p]$-module Q equipped with the C_p-action $\sigma q = \zeta q$. For $P \in \mathrm{mf}^Z(Q, pf; C_p)$, we define

$$P^{(\zeta)} := \mathrm{Hom}_{Q[C_p]}(Q_\zeta, P) = \ker(\sigma - \zeta : P \to P).$$

Since p is invertible and ζ belongs to Q, the module Q_ζ is a direct summand of $Q[C_p]$, and so $P \mapsto P^{(\zeta)}$ is an exact functor. It therefore induces a map

$$\phi^p_\zeta : K_0(\mathrm{mf}^Z(Q, pf; C_p)) \xrightarrow{[P]\mapsto[P^{(\zeta)}]} K_0(\mathrm{mf}^Z(Q, pf)),$$

and so we may form the composition

$$K_0(\mathrm{mf}^Z(Q, f)) \xrightarrow{t^p_\Sigma} K_0(\mathrm{mf}^Z(Q, pf; \Sigma_p)) \xrightarrow{\mathrm{res}} K_0(\mathrm{mf}^Z(Q, pf; C_p)) \xrightarrow{\phi^p_\zeta} K_0(\mathrm{mf}^Z(Q, pf)).$$

We come upon the need to "divide by p". In general, if $u \in Q$ is a unit, we define an autoequivalence

$$\mathrm{mult}_u : \mathrm{mf}^Z(Q, f) \to \mathrm{mf}^Z(Q, uf)$$

by sending a matrix factorization (α, β) to $(\alpha, u\beta)$. (Its inverse is given by $\mathrm{mult}_{u^{-1}}$.) For example, in Section 3C, we will employ the functor mult_{-1}, which we will write as $\mathrm{mult}_{-1}(P) = P^\circ$. Here, we use $\mathrm{mult}_{1/p}$, and we define t^p_ζ to be the composition

$$K_0(\mathrm{mf}^Z(Q, f)) \xrightarrow{\phi^p_\zeta \circ \mathrm{res} \circ t^p_\Sigma} K_0(\mathrm{mf}^Z(Q, pf)) \xrightarrow{\mathrm{mult}_{1/p}} K_0(\mathrm{mf}^Z(Q, f)).$$

Let A_p denote the subring of $\mathbb{C}$ given by $\mathbb{Z}[1/p, e^{2\pi i/p}]$.

Definition 2.7. Assume p is a prime, Q is a (commutative, Noetherian) A_p-algebra, f is any element of Q, and Z is a closed subset of $\mathrm{Spec}(Q)$. Define

$$\psi^p_{\mathrm{cyc}} = \sum_\zeta \zeta t^p_\zeta : K_0(\mathrm{mf}^Z(Q, f)) \to K_0(\mathrm{mf}^Z(Q, f)),$$

where the sum ranges over all p-th roots of unity. (In this formula, the ζ occurring as a coefficient is interpreted as belonging to $\mathbb{Z}[e^{2\pi i/p}]$ whereas the ζ occurring as a subscript denotes its image in Q under the map $A_p \to Q$.)

Remark 2.8. The image of ψ^p_{cyc} is contained in the group $K_0(\mathrm{mf}^Z(Q, f)) \otimes_{\mathbb{Z}} \mathbb{Z}[e^{2\pi i/p}]$. But, by an argument identical to the proof of Corollary 3.5 in [BMTW 2017], we have

$$\sum_{\zeta} \zeta t^p_{\zeta} = t^p_1 - t^p_{\zeta'}$$

for any fixed primitive p-th root of unity ζ', and thus the image of ψ^p_{cyc} can be taken to be $K_0(\mathrm{mf}^Z(Q, f))$.

Remark 2.9. Setting $\phi^p = \sum_{\zeta} \zeta \phi^p_{\zeta}$, one gets another formulation

$$\psi^p_{\mathrm{cyc}} = \mathrm{mult}_{1/p} \circ \phi^p \circ \mathrm{res} \circ t^p_{\Sigma}.$$

2B. ***Axioms for Adams operations on matrix factorizations à la Gillet–Soulé.*** In this subsection, we show the operations ψ^p_{cyc} satisfy the following analogues of the axioms of Gillet and Soulé (see Theorem 3.7 in [BMTW 2017]).

Theorem 2.10. *Assume p is a prime, Q is a (commutative, Noetherian) A_p-algebra, f, f_1, f_2 are any elements of Q, and Z is a closed subset of* $\mathrm{Spec}(Q)$:

(1) ψ^p_{cyc} *is a group endomorphism of* $K_0(\mathrm{mf}^Z(Q, f))$.

(2) *For $\alpha \in K_0(\mathrm{mf}^Z(Q, f_1))$ and $\beta \in K_0(\mathrm{mf}^W(Q, f_2))$,*

$$\psi^p_{\mathrm{cyc}}(\alpha \cup \beta) = \psi^p_{\mathrm{cyc}}(\alpha) \cup \psi^p_{\mathrm{cyc}}(\beta) \in K_0(\mathrm{mf}^{Z\cap W}(Q, f_1 + f_2)),$$

where $\cup$ is the multiplication rule on Grothendieck groups induced by tensor product. The three operators ψ^p_{cyc} in the equation are, from left to right, acting on $K_0(\mathrm{mf}^{Z\cap W}(Q, f_1 + f_2))$, $K_0(\mathrm{mf}^Z(Q, f_1))$, and $K_0(\mathrm{mf}^W(Q, f_2))$.

(3) ψ^p_{cyc} *is functorial in the following sense: Suppose $\rho : Q \to Q'$ is map of A_p-algebras, $f' = \rho(f)$, and $\tilde{\rho}^{-1}(Z) \subseteq Z'$ where $\tilde{\rho}$:* $\mathrm{Spec}\, Q' \to \mathrm{Spec}\, Q$ *is the induced map on spectra. Then extension of scalars along ρ induces a map $K_0(\mathrm{mf}^Z(Q, f)) \to K_0(\mathrm{mf}^{Z'}(Q', f'))$ that commutes with the actions of ψ^p_{cyc}.*

(4) *If $f = gh$, so that $(g, h) := (Q \xrightarrow{g} Q, Q \xrightarrow{h} Q)$ is an object of* $\mathrm{mf}^{V(g,h)}(Q, f)$, *we have*

$$\psi^p_{\mathrm{cyc}}[(g, h)] = p[(g, h)].$$

Proof. The proofs of (1)–(3) are essentially identical to the proofs of parts (1)–(3) of Theorem 3.7 in [BMTW 2017]. As for (4), let $(0, 0)$ denote the matrix factorization $(Q \xrightarrow{0} Q, Q \xrightarrow{0} Q)$ of 0, and let X denote the tensor product

$$(g, ph) \otimes_Q (0, 0) \otimes_Q \cdots \otimes_Q (0, 0).$$

Set $\zeta := e^{2\pi i/p}$ and $\sigma := (1\,2\,\cdots\,p) \in C_p$. We equip X with a C_p action by letting σ act on the i-th factor of X in the following way: If x has odd degree, $\sigma \cdot x = \zeta^{i-1} x$. If x has even degree, $\sigma \cdot x = x$.

We claim that there is an isomorphism

$$T^p([g,h]) \cong (g, ph) \otimes_Q (0,0) \otimes_Q \cdots \otimes_Q (0,0)$$

in $\mathrm{mf}^{V(g,h)}(Q, pf; C_p)$. To prove the claim, let V be a free Q-module of rank p with a fixed basis $\{e_0, \dots, e_{p-1}\}$. We identify the underlying Q-modules of $T^p((g,h))$ and X with the exterior algebra $\bigwedge V$ of V; under this identification, the action of C_p on $T^p((g,h))$ is given by

$$\sigma(e_{i_1} \wedge \cdots \wedge e_{i_n}) = e_{\sigma^{-1}(i_1)} \wedge \cdots \wedge e_{\sigma^{-1}(i_n)},$$

and the action of C_p on X is given by

$$\sigma(e_{i_1} \wedge \cdots \wedge e_{i_n}) = \zeta^{i_1+\cdots+i_n} e_{i_1} \wedge \cdots \wedge e_{i_n}.$$

For $0 \le i \le p-1$, define $v_i := 1/p \sum_j \zeta^{ij} e_j$. Then $v_0, \dots, v_{p-1}$ form a basis of V. Let $\alpha : \bigwedge V \to \bigwedge V$ denote the Q-algebra automorphism given by $e_i \mapsto v_i$. Then α yields an isomorphism $T^p((g,h)) \xrightarrow{\sim} X$ of C_p-equivariant matrix factorizations; this proves the claim.

(In checking the details here, it is useful to note the following: The "differential" on $T^p((g,h))$ is given by s_0+s_1, where s_0 is left-multiplication by $h(e_0+\cdots+e_{p-1})$, and s_1 is given by the Koszul differential on the sequence $(g, g, \dots, g)$. Similarly, the "differential" on X is given by $t_0 + t_1$, where t_0 is left-multiplication by phe_0 and t_1 is given by the Koszul differential on the sequence $(g, 0, \dots, 0)$.)

By Remark 2.9, and the result analogous to Lemma 3.11 of [BMTW 2017] for matrix factorizations (with essentially the same proof), we have

$$\psi^p_{\mathrm{cyc}}([(g,h)]) = \mathrm{mult}_{1/p}\big(\phi^p([(g,ph)]) \cup \phi^p([(0,0)]) \cup \cdots \cup \phi^p([(0,0)])\big).$$

Here, ϕ^p acts as the identity on the first factor, which is equipped with the trivial action of C_p. Furthermore, direct calculation on the $(i+1)$-st factor yields

$$\phi^p([(0,0]) = [I] + \zeta^i[\Sigma I] = (1-\zeta^i)[I]$$

where I denotes the unit matrix factorization $(0 \xrightarrow{0} Q, Q \xrightarrow{0} 0)$. Thus, one obtains

$$\psi^p_{\mathrm{cyc}}([(g,h)]) = \mathrm{mult}_{1/p}([(g,ph)] \cup [I] \cup \cdots \cup [I]) \prod_{i=1}^{p-1}(1-\zeta^i) = p[(g,h)],$$

since $\prod_{i=1}^{p-1}(1-\zeta^i) = p$. □

Corollary 2.11. *If $a = (a_1, \dots, a_n)$ is a sequence of elements in an A_p-algebra Q and $K(a)$ is the associated $\mathbb{Z}/2$-folded Koszul complex, regarded as an object of $\mathrm{mf}^{V(a_1,\dots,a_n)}(Q,0)$, then*

$$\psi^p_{\mathrm{cyc}}([K(a)]) = p^n[K(a)] \in K_0(\mathrm{mf}^{V(a_1,\dots,a_n)}(Q,0)).$$

Proof. This follows from parts (2) and (4) of the theorem, because $K(a)$ is the tensor product of the matrix factorizations $(a_i, 0)$ and $\mathbb{Z}/2$-folding commutes with tensor product. $\square$

2C. *Diagonalizability.* Suppose Q is a regular ring and $f \in Q$ is a not a zero divisor. Recall, from the introduction, that $\mathcal{P}^{V(f)}(Q)$ denotes the category of bounded complexes of finitely generated and projective Q-modules whose homology is supported on $V(f)$, and $K_0^{V(f)}(Q)$ denotes its Grothendieck group. In this subsection, we construct a surjection

$$\rho_f : K_0^{V(f)}(Q) \twoheadrightarrow K_0(\mathrm{mf}(Q, f))$$

that commutes with the actions of ψ^p_{cyc}. Using this, and Corollary 3.12 of [BMTW 2017] (the proof of which is really due to Gillet–Soulé), we deduce that the action of ψ^p_{cyc} on $K_0(\mathrm{mf}(Q, f))_{\mathbb{Q}}$ decomposes the latter into eigenspaces of the expected weights.

Let K_f denote the Koszul dga associated to f, so that, as a Q-algebra, $K_f = Q[\epsilon]/(\epsilon^2)$ with $|\epsilon| = 1$, and it is equipped with the Q-linear differential d satisfying $d(\epsilon) = f$. Let $P(K_f/Q)$ denote the full subcategory of the category of dg-K_f-modules consisting of those that are finitely generated and projective as Q-modules. An object of $P(K_f/Q)$ is thus a bounded complex P of finitely generated projective Q-modules equipped with a degree one Q-linear map $s : P_\cdot \to P_{\cdot+1}$ satisfying $d_P s + s d_P = f$ and $s^2 = 0$. (The map s is given by multiplication by ϵ.) A morphism from (P, d_P, s) to $(P', d_{P'}, s')$ is a chain map g such that $gs = s'g$. A homotopy from g_1 to g_2 is a degree one map h such that $d_{P'}h + hd_P = g_1 - g_2$ and $hs = s'h$.

There are functors

$$\mathcal{P}^{V(f)}(Q) \xleftarrow{F} P(K_f/Q) \xrightarrow{\text{Fold}} \mathrm{mf}(Q, f),$$

where F is the forgetful functor that sends (P, d_P, s) to (P, d_P), and Fold sends (P, d, s) to the following matrix factorization: the even degree part is $\bigoplus_i P_{2i}$, the odd degree part is $\bigoplus_i P_{2i+1}$ and the degree one endomorphism is $\partial := d + s$.

Define $K_0(P(K_f/Q))$ to be the Grothendieck group of objects modulo relations coming from short exact sequences and homotopy equivalences as usual.

Lemma 2.12. *If f is not a zero divisorin a regular ring Q, the functor F induces an isomorphism*

$$K_0(P(K_f/Q)) \xrightarrow{\sim} K_0^{V(f)}(Q).$$

Proof. Let $R = Q/(f)$. One has an evident quasiisomorphism $K_f \xrightarrow{\sim} R$ of dga's, and hence an equivalence of triangulated categories $D^b(R) \xrightarrow{\sim} D^b(K_f)$ induced by restriction of scalars. Thus, one has an isomorphism

$$G_0(R) = K_0(D^b(R)) \xrightarrow{\sim} K_0(D^b(K_f)).$$

We may model $D^b(K_f)$ by semiprojective K_f-modules with finitely generated homology. Since Q is regular, the good truncation of such a complex in sufficiently high degree is a complex of projective Q-modules. It thus follows from Quillen's resolution theorem that the inclusion map determines an isomorphism

$$K_0(P(K_f/Q)) \xrightarrow{\sim} K_0(D^b(K_f)).$$

We thus obtain an isomorphism $G_0(R) \xrightarrow{\sim} K_0(P(K_f/Q))$, which we can describe explicitly as follows: If M is a finitely generated R-module, form a (possibly infinite) K_f-semiprojective resolution $P \xrightarrow{\sim} M$ of M. Then the map sends $[M]$ to $[P']$ where P' is a good truncation of P in sufficiently high degree.

We also have the more classical isomorphism $G_0(R) \xrightarrow{\sim} K_0^{V(f)}(Q)$, sending $[M]$ to the class of a Q-projective resolution of M. Since the complex P' constructed above is an example of such a resolution, it is clear that the triangle

$$\begin{tikzcd} K_0(P(K_f/Q)) \arrow[rr, "F"] & & K_0^{V(f)}(Q) \\ & G_0(R) \arrow[ul, "\cong"] \arrow[ur, "\cong"'] & \end{tikzcd}$$

commutes. $\square$

The functor Fold induces a map from $K_0(P(K_f/Q))$ to $K_0(\mathrm{mf}(Q, f))$, and thus, using the lemma, we obtain the desired map $\rho_f : K_0^{V(f)}(Q) \to K_0(\mathrm{mf}(Q, f))$. Explicitly, the construction shows that if an object $P \in \mathcal{P}^{V(f)}(Q)$ admits a degree one map s satisfying $ds + sd = f$ and $s^2 = 0$, then $\rho_f([P]) = [\mathrm{Fold}(P, d, s)]$. In particular, the map ρ_f is surjective, since for a matrix factorization $(\alpha : P_1 \to P_0, \beta : P_0 \to P_1) \in \mathrm{mf}(Q, f)$, we have $(\alpha, \beta) = \mathrm{Fold}(P, \alpha, \beta)$.

Since there exists an isomorphism $G_0(Q/(f)) \xrightarrow{\sim} K_0^{V(f)}(Q)$ which sends the class of a finitely generated $Q/(f)$-module to the class of a chosen Q-projective resolution of it, we obtain a surjective map

$$G_0(Q/(f)) \twoheadrightarrow K_0(\mathrm{mf}(Q, f)).$$

Note that this surjection agrees with the one induced by the inverse of the equivalence $[\mathrm{mf}(Q, f)] \xrightarrow{\sim} D_{\mathrm{sing}}(Q/(f))$ from Theorem 1 of the introduction.

Given a finitely generated $Q/(f)$-module M, let $[M]_{\mathrm{stable}} \in K_0(\mathrm{mf}(Q, f))$ denote the image of $[M]$ under the above surjection $G_0(Q/(f)) \twoheadrightarrow K_0(\mathrm{mf}(Q, f))$. Explicitly, for such an M, one may find a Q-projective resolution (P, d) of it for which there exists a degree one endomorphism s of P satisfying $ds + sd = f$ and $s^2 = 0$ (by taking, for instance as above, a good truncation in sufficiently high degree of a K_f-semiprojective resolution $P \xrightarrow{\sim} M$). Then $[M]_{\mathrm{stable}} = [\mathrm{Fold}(P, d, s)]$.

We will use the following result to deduce the diagonalizability of ψ^p_{cyc} on the Grothendieck group of matrix factorizations from the corresponding result for complexes.

Proposition 2.13. *Assume Q is a regular A_p-algebra and $f \in Q$ is a not a zero divisor. The map ρ_f commutes with the Adams operations ψ^p_{cyc}.*

Proof. We need to show the diagram

$$\begin{tikzcd}
K_0^{V(f)}(Q) \arrow[r, "\rho_f"] \arrow[d, "\psi^p_{\mathrm{cyc}}"] & K_0(\mathrm{mf}(Q, f)) \arrow[d, "{[Y]\mapsto[T^p(Y)^{(1)}]-[T^p(Y)^{(\zeta)}]}"] \\
K_0^{V(f)}(Q) \arrow[r, "\rho_{pf}"] \arrow[d, "="] & K_0(\mathrm{mf}(Q, pf)) \arrow[d, "\mathrm{mult}_{1/p}"] \\
K_0^{V(f)}(Q) \arrow[r, "\rho_f"] & K_0(\mathrm{mf}(Q, f))
\end{tikzcd}$$

commutes.

It suffices to check the commutativity of the top square on classes $[P]$ for which there exists an s with $ds + sd = f$ and $s^2 = 0$. Recall that the induced differential $T^p(d)$ on $T^p(P)$ is given by

$$T^p(d)(x_1 \otimes \cdots \otimes x_p) = \sum_{i=1}^{p} (-1)^{|x_1|+\cdots+|x_{i-1}|} x_1 \otimes \cdots \otimes d(x_i) \otimes \cdots \otimes x_p,$$

and we define $T^p(s)$ to be the degree one map given by the same formula with s in place of d. Then $T^p(d)T^p(s) + T^p(s)T^p(d) = pf$ and $T^p(s)^2 = 0$. Moreover, it follows from the definitions that there is a canonical isomorphism

$$T^p(\mathrm{Fold}(P, d, s)) \cong \mathrm{Fold}(T^p(P), T^p(d), T^p(s)) \in \mathrm{mf}(Q, pf),$$

and this isomorphism is equivariant for the action of Σ_p. The commutativity of the top square in the diagram follows.

The bottom square commutes by the more general lemma below. $\square$

Lemma 2.14. *If Q is a regular, $f \in Q$ is not a zero divisor, and $u \in Q$ is a unit, the triangle*

$$\begin{tikzcd}
& K_0(\mathrm{mf}(Q, f)) \arrow[dd, "\mathrm{mult}_u"] \\
K_0^{V(f)}(Q) \arrow[ur, two heads, "\rho_f"] \arrow[dr, two heads, "\rho_{uf}"'] & \\
& K_0(\mathrm{mf}(Q, uf))
\end{tikzcd}$$

commutes.

Proof. Again, it suffices to check the commutativity of the diagram on classes $[P]$ such that P is a complex with differential d for which there exists an s with $ds + sd = f$ and $s^2 = 0$. If $[P]$ is such a class, $\rho_f([P]) = [\mathrm{Fold}(P, d, s)]$.

Before applying ρ_{uf}, first replace (P, d) by the isomorphic complex (P', d') with $P'_i = P_i$ for all i and with $d'_i = d_i$ for i odd and $d'_i = ud_i$ for i even. Defining s' as $s'_i = s_i$ for i odd and $s'_i = us_i$ for i even, one has $d's' + s'd' = uf$. Then $\rho_{uf}([P]) = [\mathrm{Fold}(P,' d', s')] = \mathrm{mult}_u([\mathrm{Fold}(P, d, s)]) = (\mathrm{mult}_u \circ \rho_f)([P])$. □

Theorem 2.15. *Assume Q is a regular A_p-algebra of dimension d and $f \in Q$ is not a zero divisor. There is a decomposition*

$$K_0(\mathrm{mf}(Q, f))_{\mathbb{Q}} = \bigoplus_{i=1}^{d} K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}},$$

which is independent of p, such that ψ^p_{cyc} acts on $K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}}$ as multiplication by p^i. Moreover, for a finitely generated $Q/(f)$-module M, we have

$$[M]_{\mathrm{stable}} \in \bigoplus_{i=\mathrm{codim}_{Q/(f)} M+1}^{d} K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}}.$$

Proof. This follows from Corollary 3.12 of [BMTW 2017] and Proposition 2.13 by defining $K_0(\mathrm{mf}(Q, f))^{(i)}_{\mathbb{Q}}$ to be the image of $K_0^{V(f)}(Q)^{(i)}_{\mathbb{Q}}$ under $\rho_f \otimes \mathbb{Q}$. □

We close this subsection with a technical result needed below.

Corollary 2.16. *If Q is a regular A_p-algebra for a prime p, $f \in Q$ is not a zero divisor, and $u \in Q$ is a unit, we have an equality of maps $\psi^p_{\mathrm{cyc}} \circ \mathrm{mult}_u = \mathrm{mult}_u \circ \psi^p_{\mathrm{cyc}}$ from $K_0(\mathrm{mf}(Q, f))$ to $K_0(\mathrm{mf}(Q, uf))$.*

Proof. By Proposition 2.13, the diagonal maps in the commutative diagram of Lemma 2.14 commute with the action of ψ^p_{cyc}, and these maps are surjective. □

3. Dao and Kurano's Conjecture

In this section, we apply the results of Section 2 to give a proof of Theorem 5 from the introduction.

3A. ***Some properties of $\mathbb{Z}/2$-graded complexes.*** We will need some general results about $\mathbb{Z}/2$-graded complexes. Much of what we need holds in great generality, and so we start by working over a Noetherian commutative ring B.

Let $\mathrm{LF}(B, 0)$ denote the abelian category of all $\mathbb{Z}/2$-graded complexes of B-modules ("LF" stands for "linear factorization"), and let $\mathrm{lf}(B, 0)$ denote the full subcategory of $\mathrm{LF}(B, 0)$ consisting of complexes whose components are finitely generated B-modules. An object of $\mathrm{LF}(B, 0)$ consists of a pair of B-modules, M^0 and M^1, together with maps $d^0 : M^0 \to M^1$ and $d^1 : M^1 \to M^0$ such that

$d^1 \circ d^0 = 0 = d^0 \circ d^1$. Morphisms are given by the evident $\mathbb{Z}/2$-graded analogues of chain maps. We also have the evident $\mathbb{Z}/2$-versions of quasiisomorphisms and homotopies of chain maps. For objects $X, Y \in \mathrm{LF}(B, 0)$, let $\mathrm{Hom}_{\mathrm{LF}}(X, Y)$ denote the $\mathbb{Z}/2$-analogue of the mapping complex construction. So $\mathrm{Hom}_{\mathrm{LF}}(X, Y) \in \mathrm{LF}(B, 0)$ with $\mathrm{Hom}_{\mathrm{LF}}(X, Y)^{\epsilon} = \bigoplus_{\epsilon'+\epsilon''=\epsilon} \mathrm{Hom}_B(X^{\epsilon'}, Y^{\epsilon''})$. Note that the zero cycles in $\mathrm{Hom}_{\mathrm{LF}}(X, Y)$ are, by definition, the set of morphisms from X to Y in $\mathrm{LF}(B, 0)$, and $H^0 \mathrm{Hom}_{\mathrm{LF}}(X, Y)$ is the set of morphisms modulo homotopy.

We write $X \otimes_{\mathrm{LF}} Y \in \mathrm{LF}(B, 0)$ for the evident $\mathbb{Z}/2$-graded analogue of the tensor product of complexes, so that

$$(X \otimes_{\mathrm{LF}} Y)^{\epsilon} = \bigoplus_{\epsilon=\epsilon'+\epsilon''} X^{\epsilon'} \otimes_B Y^{\epsilon''}.$$

We will also need the notion of the totalization $\mathrm{Tot}(X.)$ of a bounded complex

$$X. := (0 \to X_m \to \cdots \to X_0 \to 0)$$

of objects of $\mathrm{LF}(B, 0)$, defined in a manner similar to the $\mathbb{Z}$-graded setting. In more detail, we have

$$\mathrm{Tot}(X.)^{\epsilon} = \bigoplus_{i=0}^{m} X_i^{i+\epsilon},$$

with superscripts taken modulo 2. Moreover, if

$$0 \to X_m \to \cdots \to X_0 \to M \to 0$$

is an exact sequence in $\mathrm{LF}(B, 0)$, then there is a natural quasiisomorphism

$$\mathrm{Tot}(X.) \xrightarrow{\sim} M$$

in $\mathrm{LF}(B, 0)$.

For $M \in \mathrm{LF}(B, 0)$, define $Z(M)$ to be the $\mathbb{Z}/2$-graded module consisting of the kernels of the two maps comprising the complex M, and define $B(M)$ to be the $\mathbb{Z}/2$-graded module given by the images of the two maps comprising M. Let $\mathrm{H}(M)$ denote the $\mathbb{Z}/2$-graded module consisting of the homology modules of M. Each of B, Z, and H can be interpreted as a functor from $\mathrm{LF}(B, 0)$ to itself, and they restrict to functors from $\mathrm{lf}(B, 0)$ to itself. Note that $B(M) \subseteq Z(M)$ and $\mathrm{H}(M) = Z(M)/B(M)$.

Recall that $\mathrm{mf}(B, 0)$ is the full subcategory of $\mathrm{lf}(B, 0)$ consisting of complexes whose components are projective B-modules.

Definition 3.1. An object $X \in \mathrm{mf}(B, 0)$ is called *proper* if $Z(X)$, $B(X)$ and $\mathrm{H}(X)$ are all projective R-modules.

For $M \in \mathrm{lf}(B, 0)$, an exact sequence of the form

$$\cdots \to X_m \to \cdots \to X_1 \to X_0 \to M \to 0$$

such that $X_i \in \mathrm{mf}(B, 0)$ is proper for all i and each of the induced sequences

$$\begin{aligned}
&\cdots \to B(X_m) \to \cdots \to B(X_1) \to B(X_0) \to B(M) \to 0,\\
&\cdots \to Z(X_m) \to \cdots \to Z(X_1) \to Z(X_0) \to Z(M) \to 0,\\
&\cdots \to \mathrm{H}(X_m) \to \cdots \to \mathrm{H}(X_1) \to \mathrm{H}(X_0) \to \mathrm{H}(M) \to 0
\end{aligned}$$

is also exact is called a *Cartan–Eilenberg resolution* of M. Such a resolution is *bounded* if $X_j = 0$ for all $j \gg 0$.

Lemma 3.2. *If B is a Noetherian commutative ring, and at least one of $X, Y \in$ $\mathrm{mf}(B, 0)$ is proper, then there is a natural isomorphism*

$$\mathrm{H}(X) \otimes_{\mathrm{LF}} \mathrm{H}(Y) \xrightarrow{\sim} \mathrm{H}(X \otimes_{\mathrm{LF}} Y).$$

Proof. The proof is the same as for the classical Künneth Theorem. □

Lemma 3.3. *If B is a Noetherian commutative ring, then every $M \in \mathrm{lf}(B, 0)$ admits a Cartan–Eilenberg resolution. If B is regular, every $M \in \mathrm{lf}(B, 0)$ admits a bounded Cartan–Eilenberg resolution.*

Proof. Choose projective resolutions of $B^0(M)$, $B^1(M)$, $\mathrm{H}^0(M)$ and $\mathrm{H}^1(M)$, and make repeated use of the horseshoe lemma, just as in the proof of the classical version of this result. If B is regular, all of the chosen projective resolutions in the proof may be chosen to be bounded. □

Recall that $[\mathrm{mf}(B, 0)]$ denotes the category with the same objects as $\mathrm{mf}(B, 0)$ and with morphism sets given by $\mathrm{Hom}_{[\mathrm{mf}(B,0)]}(X, Y) := H^0(\mathrm{Hom}_{\mathrm{LF}}(X, Y))$. We write $\mathfrak{D}(\mathrm{lf}(B, 0))$ for the category obtained from $\mathrm{lf}(B, 0)$ by inverting all quasiisomorphisms.

Proposition 3.4. *If B is regular, the canonical functor*

$$[\mathrm{mf}(B, 0)] \xrightarrow{\sim} \mathfrak{D}(\mathrm{lf}(B, 0))$$

is an equivalence.

Proof. Let M be an object in $\mathfrak{D}(\mathrm{lf}(B, 0))$. Applying Lemma 3.3, choose a bounded Cartan–Eilenberg resolution $X_\cdot$ of M. Then the canonical map $\mathrm{Tot}(X_\cdot) \to M$ is a quasiisomorphism, and $\mathrm{Tot}(X_\cdot)$ is an object of $\mathrm{mf}(B, 0)$; thus, the functor is essentially surjective. It is fully faithful by Lemma 2.3. □

We are especially interested in complexes with finite length homology. Let $\mathrm{lf}^{\mathrm{fl}}(B, 0)$ and $\mathrm{mf}^{\mathrm{fl}}(B, 0)$ denote the full subcategories of $\mathrm{lf}(B, 0)$ and $\mathrm{mf}(B, 0)$ consisting of those complexes M such that $H^0(M)$ and $H^1(M)$ are finite length B-modules. Since this condition is preserved by quasiisomorphism, we may form $[\mathrm{mf}^{\mathrm{fl}}(B, 0)]$ and $\mathfrak{D}(\mathrm{lf}^{\mathrm{fl}}(B, 0))$, and they may be identified as full subcategories of

$[\mathrm{mf}(B,0)]$ and $\mathscr{D}(\mathrm{lf}(B,0))$. Moreover, it follows from Proposition 3.4 that the canonical functor induces an equivalence

$$[\mathrm{mf}^{\mathrm{fl}}(B,0)] \xrightarrow{\simeq} \mathscr{D}(\mathrm{lf}^{\mathrm{fl}}(B,0)),$$

provided B is regular.

It will be convenient to give an alternative description of the category $\mathrm{LF}(B,0)$ and of the constructions just described. Fix a degree two indeterminate t and form the $\mathbb{Z}$-graded algebra $\tilde{B} := B[t,t^{-1}]$, which we regard as a dg-ring with trivial differential. Recall that a dg-$\tilde{B}$-module is a graded $\tilde{B}$-module M equipped with a degree one $\tilde{B}$-linear map $d : M \to M$ such that $d^2 = 0$. Since t is a degree two invertible element, a dg-$\tilde{B}$-module is the same things as a $\mathbb{Z}$-graded complex of B-modules M together with a specified isomorphism $t : M \xrightarrow{\simeq} M[2]$ of complexes. A morphism between two such pairs, say from (M,t) to (M',t'), is a chain map from M to M' that commutes with t and t'. There is an evident equivalence of abelian categories

$$\mathrm{dg}\text{-}\tilde{B}\text{-}\mathrm{Mod} \xrightarrow{\simeq} \mathrm{LF}(B,0)$$

that sends a dg-$\tilde{B}$-module M to the object

$$(M^0 \xrightarrow{d} M^1 \xrightarrow{t^{-1}d} M^0)$$

of $\mathrm{LF}(B,0)$. Moreover, the notions of mapping complex, tensor product, quasi-isomorphism, homotopy equivalence and totalization defined above for $\mathrm{LF}(B,0)$ correspond to the standard notions for dg-modules. This equivalence thus allows us to employ standard results from differential graded algebra.

3B. ***Adams operations on $\mathbb{Z}/2$-graded complexes with finite length homology.*** Let Q be a regular local ring with maximal ideal $\mathfrak{m}$. Recall that $\mathrm{mf}^{\mathfrak{m}}(Q,0)$ is the category of $\mathbb{Z}/2$-graded complexes of finite rank free Q-modules whose homology has support in $\{\mathfrak{m}\}$; notice that $\mathrm{mf}^{\mathfrak{m}}(Q,0) = \mathrm{mf}^{\mathrm{fl}}(Q,0)$, where the right-hand side is as defined in Section 3A.

Recall that $K_0^{\mathfrak{m}}(Q)$ is the Grothendieck group of the category of bounded $\mathbb{Z}$-graded complexes of projective Q-modules whose homology has support in $\{\mathfrak{m}\}$. It is easy to prove that $K_0^{\mathfrak{m}}(Q)$ is a free abelian group of rank one, generated by the class of the Koszul complex on a regular system of generators of $\mathfrak{m}$. One might thus expect the answer to the following question to be positive:

Question 3.5. *For a regular local ring $(Q,\mathfrak{m})$, is $K_0(\mathrm{mf}^{\mathfrak{m}}(Q,0))$ a free abelian group of rank one, generated by the $\mathbb{Z}/2$-folded Koszul complex?*

We know the answer to be "yes" if $\dim(Q) \le 2$, but the general situation remains unknown. The following example illustrates the difficulty:

Example 3.6. Let $(Q, \mathfrak{m})$ be a regular local ring of dimension three, and suppose x, y, z form a regular sequence of generators for the maximal ideal $\mathfrak{m}$. Let

$$0 \to Q \xrightarrow{i} Q^3 \xrightarrow{A} Q^3 \xrightarrow{p} Q \to 0$$

be the usual Koszul complex on x, y, z (so that, for example, p is given by the row matrix (x, y, z)). The $\mathbb{Z}/2$-folding of this Koszul complex,

$$K := \left(Q^3 \oplus Q \xrightarrow{\left[\begin{smallmatrix} A & 0 \\ 0 & 0 \end{smallmatrix}\right]} Q^3 \oplus Q \xrightarrow{\left[\begin{smallmatrix} 0 & i \\ p & 0 \end{smallmatrix}\right]} Q^3 \oplus Q \right),$$

determines a class $[K]$ in $K_0(\mathrm{mf}^{\mathfrak{m}}(Q, 0))$.

Now define $B : Q^3 \to Q^3$ to be the map $i \circ p$. Then $AB = 0 = BA$, so that $X = (Q^3 \xrightarrow{A} Q^3 \xrightarrow{B} Q^3)$ is a $\mathbb{Z}/2$-graded complex. Moreover, $\ker(B) = \mathrm{im}(A)$ and $\ker(A)/\mathrm{im}(B) \cong Q/\mathfrak{m}$, so that $X \in \mathrm{mf}^{\mathfrak{m}}(Q, 0)$. We do not know whether $[X]$ is a multiple of $[K]$ in $K_0(\mathrm{mf}^{\mathfrak{m}}(Q, 0))$.

To explain the relevance of Question 3.5, let us define the *Euler characteristic* of an object $X \in \mathrm{mf}^{\mathfrak{m}}(Q, 0)$ to be

$$\chi(X) = \text{length } H^0(X) - \text{length } H^1(X).$$

Then χ determines a group homomorphism

$$\chi : K_0(\mathrm{mf}^{\mathfrak{m}}(Q, 0)) \to \mathbb{Z}.$$

For example, if K is the $\mathbb{Z}/2$-folded Koszul complex on a regular system of generators for $\mathfrak{m}$, then $\chi(K) = 1$. Assume now that Q is a regular local A_p-algebra for a prime p (that is, assume p is invertible in Q and that Q contains a primitive p-th root of unity), so that the cyclic Adams operation ψ^p_{cyc} acts on $K_0(\mathrm{mf}^{\mathfrak{m}}(Q, 0))$. We have $\psi^p_{\mathrm{cyc}}([K]) = p^d[K]$, where $d = \dim(Q)$, by Corollary 2.11. If the answer to Question 3.5 were affirmative, we would obtain as an immediate consequence the identity

$$\chi \circ \psi^p_{\mathrm{cyc}} = p^d \chi \tag{3.7}$$

of maps from $K_0(\mathrm{mf}^{\mathfrak{m}}(Q, 0))$ to $\mathbb{Z}$. Moreover, this equation plays a key role in the proof of Theorem 5.

Although we are unable to answer Question 3.5, we are nevertheless able to prove an analogue to [Gillet and Soulé 1987, Proposition 7.1].

Theorem 3.8. *For a regular local ring Q of dimension d that is an A_p-algebra for some prime p,* (3.7) *holds.*

The proof of this theorem occupies the remainder of this subsection.

Fix a prime p, and let B be a commutative Noetherian A_p-algebra. Recall the functor t^p_ζ defined on $\mathrm{mf}(B, 0)$ that sends X to $T^p(X)^{(\zeta)}$, where ζ is a p-th root of unity. It will be useful to interpret this functor as a composition

$$\mathrm{mf}(B, 0) \xrightarrow{T^p} \mathrm{mf}(B', 0) \xrightarrow{Y \mapsto Y^{(\zeta)}} \mathrm{mf}(B, 0)$$

where we set $B' = B[C_p] = B[\sigma]/(\sigma^p - 1)$. Since B is an A_p-algebra, B' is isomorphic to a product of p copies of B equipped with an action of C_p. So, an object of $\mathrm{mf}(B', 0)$ is the same thing as an object of $\mathrm{mf}(B, 0)$ equipped with an action of C_p, and if B is regular, then so is B'.

The functors above preserve the condition that homology has finite length, and they send homotopic maps to homotopic maps, so that we have an induced functor

$$t^p_\zeta : [\mathrm{mf}^{\mathrm{fl}}(B, 0)] \to [\mathrm{mf}^{\mathrm{fl}}(B, 0)]$$

given as the composition of functors

$$[\mathrm{mf}^{\mathrm{fl}}(B, 0)] \xrightarrow{T^p} [\mathrm{mf}^{\mathrm{fl}}(B', 0)] \xrightarrow{Y \mapsto Y^{(\zeta)}} [\mathrm{mf}^{\mathrm{fl}}(B, 0)].$$

We will need a "derived" version of the functor t^p_ζ. When B is regular, then we may use the equivalence of Proposition 3.4 to obtain a functor

$$\mathbf{t}^p_\zeta : \mathcal{D}(\mathrm{lf}^{\mathrm{fl}}(B, 0)) \to [\mathrm{mf}^{\mathrm{fl}}(B, 0)].$$

Explicitly, for $M \in \mathrm{lf}^{\mathrm{fl}}(B, 0)$, $\mathbf{t}^p_\zeta(M) = t^p_\zeta(P)$ where P is any object of $\mathrm{mf}^{\mathrm{fl}}(B, 0)$ for which there exists a quasiisomorphism $P \xrightarrow{\sim} M$.

Given $M \in \mathrm{lf}(B, 0)$, recall that $\mathrm{H}(M)$ denotes the object of $\mathrm{lf}(B, 0)$ given by the $\mathbb{Z}/2$-graded B-module with components $H^0(M)$ and $H^1(M)$, regarded as a complex with trivial differential. In terms of the dg-ring $\tilde{B}$, $\mathrm{H}(M)$ corresponds to the homology of a dg-$\tilde{B}$-module, which is naturally a dg-$\tilde{B}$-module with trivial differential (since $\tilde{B}$ has trivial differential). If $M \in \mathrm{lf}^{\mathrm{fl}}(B, 0)$, we define its Euler characteristic by

$$\chi(M) := \mathrm{length}\, H^0(M) - \mathrm{length}\, H^1(M),$$

as above.

Lemma 3.9. *If B is a regular A_p-algebra, then for any $M \in \mathrm{lf}^{\mathrm{fl}}(B, 0)$ and any p-th root of unity ζ, we have*

$$\chi(\mathbf{t}^p_\zeta(M)) = \chi(\mathbf{t}^p_\zeta(\mathrm{H}(M))).$$

Theorem 3.8 is a relatively easy consequence of Lemma 3.9. Before proving Lemma 3.9, we must introduce the following notation and establish one more preliminary result. For a bounded complex

$$X_\cdot = (0 \to X_m \to X_{m-1} \to \cdots \to X_1 \to X_0 \to 0)$$

of objects of $\mathrm{LF}(B, 0)$, we write $\mathcal{H}_q(X.) \in \mathrm{LF}(B, 0)$ for its homology taken in the abelian category $\mathrm{LF}(B, 0)$; that is,

$$\mathcal{H}_q(X.) = \ker(X_q \to X_{q-1}) / \operatorname{im}(X_{q+1} \to X_q).$$

We write $\mathrm{H}(X.)$ for the complex of objects of $\mathrm{LF}(B, 0)$ obtained by applying H term-wise

$$\mathrm{H}(X.) := (0 \to \mathrm{H}(X_d) \to \cdots \to \mathrm{H}(X_0) \to 0).$$

Note that $\mathrm{H}(X.)$ is a complex of $\mathbb{Z}/2$-graded modules, and we regard it as another complex of objects in $\mathrm{LF}(B, 0)$.

Lemma 3.10. *For a Noetherian commutative ring B, assume*

$$Y. := (0 \to Y_m \to \cdots \to Y_0 \to 0)$$

is a complex in $\mathrm{lf}(B, 0)$ *such that both* $\mathcal{H}_q\, \mathrm{H}(Y.)$ *and* $\mathrm{H}\, \mathcal{H}_q(Y.)$ *have finite length for all q. Then* $\mathrm{Tot}(Y.)$ *belongs to* $\mathrm{lf}^{\mathrm{fl}}(B, 0)$, *and we have*

$$\begin{aligned}\chi(\mathrm{Tot}(Y.)) &= \sum_{q \in \mathbb{Z}, \epsilon \in \mathbb{Z}/2} (-1)^{q+\epsilon} \operatorname{length} \mathcal{H}_q(\mathrm{H}^\epsilon(Y.)) \\ &= \sum_{q \in \mathbb{Z}, \epsilon \in \mathbb{Z}/2} (-1)^{q+\epsilon} \operatorname{length} \mathrm{H}^\epsilon(\mathcal{H}_q(Y.)).\end{aligned}$$

Proof. Our proof uses spectral sequences and is similar to the proof of the analogous fact concerning $\mathbb{Z}$-graded bicomplexes, but some care is needed to deal with the $\mathbb{Z}/2$-grading.

We find it most convenient to work in the setting of dg-$\tilde{B}$-modules. Recall that a dg-$\tilde{B}$-module is the same thing as pair consisting of a $\mathbb{Z}$-graded complex of B-modules and a degree 2 automorphism. A graded $\tilde{B}$-module is a dg-$\tilde{B}$-module with trivial differential.

Let us say that a graded $\tilde{B}$-module H has *finite length* if H^i has finite length as a B-module for each $i \in \mathbb{Z}$ (or, equivalently, for $i = 0, 1$). In this case, we define

$$\tilde{\chi}(H) = \operatorname{length}_B(H^0) - \operatorname{length}_B(H^1).$$

(Note that $\tilde{\chi}(H) = \operatorname{length}_B(H^{2m}) - \operatorname{length}_B(H^{2n+1})$ for any $m, n \in \mathbb{Z}$.) It is clear that if $Y \in \mathrm{lf}^{\mathrm{fl}}(B, 0)$, then

$$\chi(Y) = \tilde{\chi}(\tilde{H}(Y))$$

where χ is as defined before, and $\tilde{H}(Y)$ denotes the homology of Y regarded in the canonical way as a graded $\tilde{B}$-module.

We will need the following fact. If (M, d) is a dg-$\tilde{B}$-module such that the underlying graded $\tilde{B}$-module M has finite length, then $H(M, d)$ also has finite length, and $\tilde{\chi}(H(M, d)) = \tilde{\chi}(M)$. This is seen to hold by a straightforward calculation.

We view $Y_.$ as a bicomplex $Y_.^{\cdot}$ with $m+1$ rows, whose m-th row, for $0 \le j \le m$, is

$$\cdots \to Y_j^{-1} \to Y_j^0 \to Y_j^1 \to \cdots,$$

along with a degree $(2,0)$ isomorphism of bicomplexes $t : Y_.^{\cdot} \xrightarrow{\sim} Y_.^{\cdot+2}$. Since this bicomplex is uniformly bounded in the vertical direction, we have two strongly convergent spectral sequences of the form

$$\begin{aligned}{}'E_2^{p,-q} &= H_q(H^p(Y_.^{\cdot})) \Longrightarrow H^{p-q}(\mathrm{Tot}(Y_.^{\cdot})) \quad \text{and}\\ {}''E_2^{p,-q} &= H^p(H_q(Y_.^{\cdot})) \Longrightarrow H^{p-q}(\mathrm{Tot}(Y_.^{\cdot})).\end{aligned}$$

Let $E_r^{*,*}$, for $r \ge 2$, refer to either of these two spectral sequences. The isomorphism $t : Y_.^{\cdot} \xrightarrow{\sim} Y_.^{\cdot+2}$ induces isomorphisms

$$t : E_r^{p,-q} \xrightarrow{\sim} E_r^{p+2,-q}$$

for each $r \ge 2$, and similarly on the underlying D_r-terms, and these isomorphisms commute with all the maps of the exact couple.

For any r, define a $\mathbb{Z}$-graded B-module $\mathrm{Tot}(E_r)$ by

$$\mathrm{Tot}(E_r)^n := \bigoplus_{p+q=n} E_r^{p,q}.$$

The isomorphism t induces an isomorphism of degree 2 on $\mathrm{Tot}(E_r)$ making it into a graded $\tilde{B}$-module. For each r, the differential d_r on the E_r's induces a degree one map (which we will also write as d_r) on $\mathrm{Tot}(E_r)$, and since this map commutes with t, we have that $(\mathrm{Tot}(E_r), d_r)$ is a dg-$\tilde{B}$-module. Finally, we have an identity

$$\mathrm{Tot}(E_{r+1}) = H(\mathrm{Tot}(E_r), d_r)$$

of graded $\tilde{B}$-modules.

Returning to the two specific instances of this spectral sequence, the assumptions give that each of $\mathrm{Tot}('E_2)$ and $\mathrm{Tot}(''E_2)$ has finite length, and that we have

$$\begin{aligned}\tilde{\chi}(\mathrm{Tot}('E_2)) &= \sum_{q\in\mathbb{Z},\epsilon\in\mathbb{Z}/2} (-1)^{q+\epsilon} \operatorname{length} \mathscr{H}_q(\mathrm{H}^\epsilon(Y_.))\\ \tilde{\chi}(\mathrm{Tot}(''E_2)) &= \sum_{q\in\mathbb{Z},\epsilon\in\mathbb{Z}/2} (-1)^{q+\epsilon} \operatorname{length} \mathrm{H}^\epsilon(\mathscr{H}_q(Y_.)).\end{aligned} \tag{3.11}$$

By the general fact mentioned above, we get that each of $\mathrm{Tot}(E_3), \mathrm{Tot}(E_4), \ldots$ also has finite length, and, moreover,

$$\tilde{\chi}(\mathrm{Tot}(E_2)) = \tilde{\chi}(\mathrm{Tot}(E_3)) = \cdots = \tilde{\chi}(\mathrm{Tot}(E_\infty)).$$

(Note that the spectral sequence degenerates after at most $m+2$ steps, so that $E_{m+2} = E_{m+3} = \cdots = E_\infty$.)

Now, for $\epsilon = 0, 1$, the B-module $H^\epsilon \operatorname{Tot}(Y)$ admits a filtration by B-submodules whose subquotients are $E_\infty^{\epsilon,0}, E_\infty^{\epsilon-1,1}, \dots, E_\infty^{\epsilon-m,m}$, and hence

$$\begin{aligned}\chi(\operatorname{Tot}(Y)) &= \tilde{\chi}(H(\operatorname{Tot}(Y)) \\ &= \sum_q \operatorname{length} E_\infty^{-q,q} - \sum_q \operatorname{length} E_\infty^{1-q,q} \\ &= \tilde{\chi}(\operatorname{Tot}(E_\infty)) = \tilde{\chi}(\operatorname{Tot}(E_2)).\end{aligned}$$

By (3.11), the proof is complete. □

Proof of Lemma 3.9. We may assume, without loss of generality, that $M = P$ belongs to $\mathrm{mf}^{\mathrm{fl}}(B, 0)$. Let

$$\cdots \to 0 \to X_m \to X_{m-1} \to \cdots \to X_1 \to X_0 \to P \to 0$$

be a bounded Cartan–Eilenberg resolution of P. Since P is an object of $\mathrm{mf}(B, 0)$, the induced quasiisomorphism $\operatorname{Tot}(X.) \xrightarrow{\sim} P$ is a homotopy equivalence, a fact that will be used below.

Recall that X_i is proper. In particular, $H(X_i)$ is projective for all i, and the induced complex

$$\cdots \to 0 \to H(X_m) \to H(X_{m-1}) \to \cdots \to H(X_1) \to H(X_0) \to H(P) \to 0$$

is also exact. The latter gives, by definition,

$$\mathbf{t}_\zeta^p(\mathrm{H}(P)) = t_\zeta^p(\operatorname{Tot}(\mathrm{H}(X.))) = T^p(\operatorname{Tot}(\mathrm{H}(X.)))^{(\zeta)}. \tag{3.12}$$

For any bounded complex $Y.$ of objects of $\mathrm{mf}(B, 0)$, write $T^p(Y.)$ for the complex of objects in $\mathrm{mf}(B, 0)$ that, in degree j, is

$$T^p(Y.)_j = \bigoplus_{i_1+\cdots+i_p=j} Y_{i_1} \otimes_{\mathrm{LF}} \cdots \otimes_{\mathrm{LF}} Y_{i_p}.$$

For example, if $p = 2$, then $T^2(Y.)$ is the complex

$$\cdots \to (Y_2 \otimes Y_0 \oplus Y_1 \otimes Y_1 \oplus Y_0 \otimes Y_2) \to (Y_1 \otimes Y_0 \oplus Y_0 \otimes Y_1) \to Y_0 \otimes Y_0 \to 0.$$

Each term of the complex $T^p(Y)$ admits an evident signed action by C_p, and the maps of this complex respect these actions, so that we may regard $T^p(Y.)$ as a complex in $\mathrm{mf}(B', 0)$, where $B' := B[C_p]$. We have an identity

$$T^p(\operatorname{Tot}(Y.)) = \operatorname{Tot}(T^p(Y.)) \tag{3.13}$$

of objects of $\mathrm{mf}(B', 0)$.

Since B is an A_p-algebra, $(-)^{(\zeta)}$ is an exact functor from $\mathrm{lf}(B', 0)$ to $\mathrm{lf}(B, 0)$. In fact, B' is a product of copies of B, and this functor is given by extension of scalars along one of the canonical projections $B' \twoheadrightarrow B$. In particular, we have

$$\mathrm{Tot}(Y_.)^{(\zeta)} = \mathrm{Tot}(Y_.^{(\zeta)}) \tag{3.14}$$

for any bounded complex $Y_.$ of objects of $\mathrm{lf}(B', 0)$, and

$$\mathrm{H}(Y)^{(\zeta)} = \mathrm{H}(Y^{(\zeta)}) \tag{3.15}$$

for any object $Y \in \mathrm{lf}(B', 0)$.

Since each X_i is proper, Lemma 3.2 implies that we have canonical isomorphisms

$$\mathrm{H}(X_{i_1}) \otimes_{\mathrm{LF}} \cdots \otimes_{\mathrm{LF}} \mathrm{H}(X_{i_p}) \xrightarrow{\simeq} \mathrm{H}(X_{i_1} \otimes_{\mathrm{LF}} \cdots \otimes_{\mathrm{LF}} X_{i_p})$$

which combine to give an isomorphism

$$T^p(\mathrm{H}(X_.)) \xrightarrow{\simeq} \mathrm{H}(T^p(X_.)) \tag{3.16}$$

of complexes of objects of $\mathrm{mf}(B', 0)$.

Combining these facts gives

$$\begin{aligned}
\mathbf{t}_\zeta^p(\mathrm{H}(P)) &= T^p(\mathrm{Tot}(\mathrm{H}(X_.)))^{(\zeta)}, && \text{by (3.12)},\\
&= (\mathrm{Tot}(T^p(\mathrm{H}(X_.))))^{(\zeta)}, && \text{by (3.13)},\\
&= \mathrm{Tot}(T^p(\mathrm{H}(X_.))^{(\zeta)}), && \text{by (3.14)},\\
&= \mathrm{Tot}(\mathrm{H}(T^p(X_.))^{(\zeta)}), && \text{by (3.16)},\\
&= \mathrm{Tot}(\mathrm{H}(T^p(X_.)^{(\zeta)})), && \text{by (3.15)}.
\end{aligned}$$

We now apply Lemma 3.10 to the complex $Y_. := T^p(X_.)^{(\zeta)}$ of objects in $\mathrm{mf}(B, 0)$, which gives

$$\sum_{q,\epsilon} (-1)^{q+\epsilon} \operatorname{length} \mathcal{H}_q(\mathrm{H}^\epsilon(Y_.)) = \sum_{q,\epsilon} (-1)^{q+\epsilon} \operatorname{length} \mathrm{H}^\epsilon(\mathcal{H}_q(Y_.)). \tag{3.17}$$

Since we have shown that $\mathrm{Tot}(\mathrm{H}(Y_.)) \cong \mathbf{t}_\zeta^p(\mathrm{H}(P))$, the left-hand side of (3.17) is $\chi(\mathbf{t}_\zeta^p(\mathrm{H}(P)))$.

Recall that, since P belongs to $\mathrm{mf}(B, 0)$, the quasiisomorphism $\mathrm{Tot}(X_.) \xrightarrow{\simeq} P$ is a homotopy equivalence. It follows that the map

$$\mathrm{Tot}(Y_.) \cong T^p(\mathrm{Tot}(X_.))^{(\zeta)} \to T^p(P)^{(\zeta)}.$$

is also a homotopy equivalence. We get

$$\mathrm{H}^\epsilon(\mathcal{H}_q(Y_.)) \cong \begin{cases} \mathrm{H}^\epsilon(t_\zeta^p(P)) & \text{if } q = 0,\\ 0 & \text{otherwise,} \end{cases}$$

which shows that the right-hand side of (3.17) is $\chi(t_\zeta^p(P))$. □

Proof of Theorem 3.8. Let $P \in \mathrm{mf}^{\mathfrak{m}}(Q,0) = \mathrm{mf}^{\mathrm{fl}}(Q,0)$. By definition,

$$\chi(\psi^p_{\mathrm{cyc}}([P])) = \sum_{\zeta} \zeta\, \chi(t^p_\zeta(P)).$$

By Lemma 3.9, the value of the right-hand side of this equation coincides with $\sum_\zeta \zeta\, \chi(\mathbf{t}^p_\zeta(\mathrm{H}(P)))$. Since $\mathrm{H}(P)$ has trivial differential, the class

$$[\mathrm{H}(P)] \in K_0(\mathscr{D}(\mathrm{lf}^{\mathrm{fl}}(Q,0))) \cong K_0(\mathrm{mf}^{\mathfrak{m}}(Q,0))$$

is an integer multiple of the class of the residue field $k = Q/\mathfrak{m}$, which in turn coincides with the class of the folded Koszul complex $K \in \mathrm{mf}^{\mathfrak{m}}(Q,0)$. This proves that the equation of Theorem 3.8 holds in general provided it holds for the class $[K]$, and that special case is known to hold by Corollary 2.11. □

3C. *Proof of the conjecture.* Throughout this section, we assume $(Q,\mathfrak{m})$ is a regular local ring and f is a nonzero element of $\mathfrak{m}$, and we set $R = Q/(f)$. We also assume R is an isolated singularity; that is, we assume $R_{\mathfrak{p}}$ is regular for all $\mathfrak{p} \in \mathrm{Spec}(R) \setminus \{\mathfrak{m}\}$. Recall from the introduction that these conditions lead to a well-defined invariant for a pair (M,N) of finitely generated R-modules:

$$\theta_R(M,N) = \mathrm{length}(\mathrm{Tor}^R_{2n}(M,N)) - \mathrm{length}(\mathrm{Tor}^R_{2n+1}(M,N))$$

for $n \gg 0$.

For a finitely generated R-module M, $[M]_{\mathrm{stable}}$ denotes its associated class in $K_0(\mathrm{mf}(Q,f))$, given by the surjection $G_0(R) \twoheadrightarrow K_0(\mathrm{mf}(Q,f))$ described in Section 2C. Recall that $[M]_{\mathrm{stable}} = [\mathrm{Fold}(P,d,s)]$, where P is a Q-projective resolution of M admitting a degree one endomorphism s that satisfies $ds + sd = f$ and $s^2 = 0$, that is, a Koszul resolution.

For a matrix factorization $X \in \mathrm{mf}(Q,f)$, write X° for $\mathrm{mult}_{-1} X \in \mathrm{mf}(Q,-f)$. That is, if $X = (\alpha : P_1 \to P_0, \beta : P_0 \to P_1)$, then $X^\circ = (\alpha, -\beta)$. We also use the notation $(-)^\circ$ to denote the induced isomorphism $K_0(\mathrm{mf}(Q,f)) \xrightarrow{\sim} K_0(\mathrm{mf}(Q,-f))$. For a finitely generated R-module N, the class $[N]^\circ_{\mathrm{stable}}$ is the image of $[N]$ under $G_0(R) \twoheadrightarrow K_0(\mathrm{mf}(Q,-f))$, using that $Q/(f) = Q/(-f)$.

Proposition 3.18. *For Q, $\mathfrak{m}$, f, R, M and N as in Definition 1.2*,

$$\theta_R(M,N) = \chi([M]_{\mathrm{stable}} \cup [N]^\circ_{\mathrm{stable}}).$$

Proof. First note that, since f is an isolated singularity, one has

$$K_0(\mathrm{mf}(Q,\pm f)) = K_0(\mathrm{mf}^{\mathfrak{m}}(Q,\pm f))$$

and hence

$$[M]_{\mathrm{stable}} \cup [N]^\circ_{\mathrm{stable}} \in K_0(\mathrm{mf}^{\mathfrak{m}}(Q, f + (-f))) = K_0(\mathrm{mf}^{\mathfrak{m}}(Q,0)).$$

Choose matrix factorizations $X = (d_1 : X_1 \to X_0, d_0 : X_0 \to X_1)$ and $Y = (d_1' : Y_1 \to Y_0, d_0' : Y_0 \to Y_1)$ such that $[X] = [M]_{\text{stable}}$ and $[Y] = [N]^{\circ}_{\text{stable}}$. Assume, without loss of generality, that N is maximal Cohen–Macaulay, and $N = \operatorname{coker}(d_1')$.

Let Z denote the object $(0 \to N, N \to 0)$ of $\operatorname{lf}(Q, -f)$; here, 0 is in odd degree and N is in even degree. Let $\alpha : Y \to Z$ be the morphism in $\operatorname{lf}(Q, -f)$ given by the canonical surjection in even degree and, of course, the zero map in odd degree. Since $\theta(M, N)$ clearly coincides with the Euler characteristic of $X \otimes Z$, it suffices to show that the morphism

$$\operatorname{id} \otimes \alpha : X \otimes Y \to X \otimes Z$$

in $\operatorname{lf}(Q, 0)$ is a quasiisomorphism. The map $\operatorname{id} \otimes \alpha$ is clearly surjective, so it suffices to show that its kernel is acyclic. An easy calculation shows that $\ker(\operatorname{id} \otimes \alpha) \cong X \otimes T$, where T is the object $(Y_1 \xrightarrow{\text{id}} Y_1, Y_1 \xrightarrow{-f} Y_1) \in \operatorname{lf}(Q, -f)$. Since T is contractible, $X \otimes T$ is contractible; thus, $\operatorname{id} \otimes \alpha$ is a quasiisomorphism. $\square$

We now prove the conjecture of Dao and Kurano:

Theorem 3.19. *Let $(Q, \mathfrak{m})$ be a regular local ring and $f \in \mathfrak{m}$ a nonzero element, and assume $R := Q/(f)$ is an isolated singularity. If M and N are finitely generated R-modules such that*

$$\dim M + \dim N \le \dim R$$

then $\theta_R(M, N) = 0$.

Proof. Let p be any prime that is invertible in Q. We start by reducing to the case where Q contains a primitive p-th root of unity. If not, we form the faithfully flat extension $Q \subseteq Q'$ where Q' is the localization of $Q[x]/(x^p - 1)$ at any one of the maximal ideals lying over $\mathfrak{m}$, and set $R' = Q'/f \cong R \otimes_Q Q'$. Note that $R \subseteq R'$ is also faithfully flat, and thus

$$\operatorname{Tor}_i^R(M, N) \otimes_R R' \cong \operatorname{Tor}_i^{R'}(M \otimes_R R', N \otimes_R R').$$

It follows that

$$\theta_{R'}(M \otimes_R R', N \otimes_R R') = [R'/\mathfrak{m}' : R/\mathfrak{m}] \cdot \theta_R(M, N),$$

and so we may replace Q with Q'.

Set $d = \dim Q$, $c_M = \operatorname{codim}_Q M$ and $c_N = \operatorname{codim}_Q N$. The hypothesis that $\dim M + \dim N \le \dim R = d - 1$ yields $c_M + c_N \ge d + 1$. By Theorem 2.15, the classes $[M]_{\text{stable}}, [N]_{\text{stable}} \in K_0(\operatorname{mf}(Q, f)) \otimes \mathbb{Q}$ decompose uniquely as

$$[M]_{\text{stable}} = \sum_{i=c_M}^{d} X_i \quad \text{and} \quad [N]_{\text{stable}} = \sum_{j=c_N}^{d} Y_j,$$

where X_i and Y_j are such that $\psi^p_{\mathrm{cyc}}(X_i) = p^i X_i$ and $\psi^p_{\mathrm{cyc}}(Y_j) = p^j Y_j$. Then

$$[N]^\circ_{\mathrm{stable}} = \sum_{j=c_N}^{d} Y_j^\circ$$

and, by Corollary 2.16, $\psi^p_{\mathrm{cyc}}(Y_j^\circ) = p^j Y_j^\circ$ for all j.

By Proposition 3.18, we have

$$\theta_R(M, N) = \chi([M]_{\mathrm{stable}} \cup [N]^\circ_{\mathrm{stable}}) = \sum_{i,j} \chi(X_i \cup Y_j^\circ),$$

and so it suffices to prove $\chi(X_i \cup Y_j^\circ) = 0$ for all i and j. For any i and j,

$$\begin{aligned} p^d \chi(X_i \cup Y_j^\circ) &= \chi(\psi^p_{\mathrm{cyc}}(X_i \cup Y_j^\circ)) \\ &= \chi(\psi^p_{\mathrm{cyc}}(X_i) \cup \psi^p_{\mathrm{cyc}}(Y_j^\circ)) \\ &= \chi(p^i X_i \cup p^j Y_j^\circ) \\ &= p^{i+j} \chi(X_i \cup Y_j^\circ), \end{aligned}$$

where the first equality is by Theorem 3.8, the second is by Theorem 2.10, and the third is by definition of X_i and Y_j. Since Theorem 2.15 yields that $i + j \geq c_M + c_N > d$, we conclude that $\chi(X_i \cup Y_j^\circ) = 0$. $\square$

Acknowledgements

We thank Luchezar Avramov for helpful conversations in preparing this paper. We thank Dave Benson, Oliver Haution, and Bernhard Köck for leading us to relevant literature [Benson 1984; Haution 2009; Köck 1997] and Paul Roberts for making available to us his unpublished 1996 lecture notes on Adams operations.

References

[Atiyah 1966] M. F. Atiyah, "Power operations in K-theory", *Quart. J. Math. Oxford Ser.* (2) **17** (1966), 165–193. MR Zbl

[Benson 1984] D. J. Benson, "Lambda and psi operations on Green rings", *J. Algebra* **87**:2 (1984), 360–367. MR Zbl

[BMTW 2017] M. K. Brown, C. Miller, P. Thompson, and M. E. Walker, "Cyclic Adams operations", *J. Pure Appl. Algebra* **221**:7 (2017), 1589–1613. MR Zbl

[Buchweitz 1986] R.-O. Buchweitz, "Maximal Cohen–Macaulay modules and Tate-cohomology over Gorenstein rings", preprint, Univ. Hannover, 1986, http://tinyurl.com/ragnarolaf.

[Buchweitz and Van Straten 2012] R.-O. Buchweitz and D. Van Straten, "An index theorem for modules on a hypersurface singularity", *Mosc. Math. J.* **12**:2 (2012), 237–259. MR Zbl

[Dao 2013] H. Dao, "Decent intersection and Tor-rigidity for modules over local hypersurfaces", *Trans. Amer. Math. Soc.* **365**:6 (2013), 2803–2821. MR Zbl

[Dao and Kurano 2014] H. Dao and K. Kurano, "Hochster's theta pairing and numerical equivalence", *J. K-Theory* **14**:3 (2014), 495–525. MR Zbl

[Eisenbud 1980] D. Eisenbud, "Homological algebra on a complete intersection, with an application to group representations", *Trans. Amer. Math. Soc.* **260**:1 (1980), 35–64. MR Zbl

[Gillet and Soulé 1987] H. Gillet and C. Soulé, "Intersection theory using Adams operations", *Invent. Math.* **90**:2 (1987), 243–277. MR Zbl

[Haution 2009] O. Haution, *Steenrod operations and quadratic forms*, Ph.D. thesis, Université Paris VI, 2009, http://www.mathematik.uni-muenchen.de/~haution/these.pdf.

[Hochster 1981] M. Hochster, "The dimension of an intersection in an ambient hypersurface", pp. 93–106 in *Algebraic geometry* (Chicago, 1980), edited by A. Libgober and P. Wagreich, Lecture Notes in Math. **862**, Springer, Berlin, 1981. MR Zbl

[Köck 1997] B. Köck, "Adams operations for projective modules over group rings", *Math. Proc. Cambridge Philos. Soc.* **122**:1 (1997), 55–71. MR Zbl

[Moore et al. 2011] W. F. Moore, G. Piepmeyer, S. Spiroff, and M. E. Walker, "Hochster's theta invariant and the Hodge–Riemann bilinear relations", *Adv. Math.* **226**:2 (2011), 1692–1714. MR Zbl

[Orlov 2004] D. O. Orlov, "Triangulated categories of singularities and D-branes in Landau–Ginzburg models", *Tr. Mat. Inst. Steklova* **246** (2004), 240–262. In Russian; translated in *Proc. Steklov Inst. Math.* **246** (2004), 227–248. MR Zbl

[Polishchuk and Vaintrob 2012] A. Polishchuk and A. Vaintrob, "Chern characters and Hirzebruch–Riemann–Roch formula for matrix factorizations", *Duke Math. J.* **161**:10 (2012), 1863–1926. MR Zbl

[Serre 2000] J.-P. Serre, *Local algebra*, Springer, Berlin, 2000. MR Zbl

[Walker 2017] M. E. Walker, "On the vanishing of Hochster's θ invariant", *Ann. K-Theory* **2**:2 (2017), 131–174. MR Zbl

Communicated by Craig Huneke
Received 2016-12-31 Accepted 2017-08-09

mkbrown5@math.wisc.edu — *Department of Mathematics, University of Wisconsin-Madison, Madison, WI, United States*

clamille@syr.edu — *Mathematics Department, Syracuse University, Syracuse, NY, United States*

peder.thompson@ttu.edu — *Department of Mathematics and Statistics, Texas Tech University, Lubbock, TX, United States*

mark.walker@unl.edu — *Deptartment of Mathematics, University of Nebraska-Lincoln, Lincoln, NE, United States*

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2193

Rationality does not specialize among terminal fourfolds

Alexander Perry

We show that rationality does not specialize in flat projective families of complex fourfolds with terminal singularities. This answers a question of Totaro, who established the corresponding result in dimensions greater than 4.

1. Introduction

Rationality behaves subtly in families of complex algebraic varieties. In general, given a flat projective family, the locus of rational fibers forms a countable union of locally closed subsets of the base [de Fernex and Fusi 2013, Proposition 2.3]. Recently, Hassett, Pirutka, and Tschinkel [2016] produced a smooth projective family of fourfolds where none of these locally closed subsets is dense, but their union is dense (even in the Euclidean topology). In particular, rationality is neither an open nor closed condition in smooth families.

This paper concerns the question of whether the locally closed subsets parametrizing the rational fibers of a family are actually closed, i.e., whether rationality specializes.

Question 1. Given a flat projective family of complex varieties, does geometric rationality of the generic fiber imply the same of every fiber?

Without further restrictions, the answer is negative: specializations of rational varieties need not even be rationally connected, as shown by a family of smooth cubic surfaces degenerating to a cone over a smooth cubic curve. However, if the fibers of the family are required to be smooth of dimension at most 3, Timmerscheidt [1982] proved the answer is positive. In fact, as Totaro observed, it follows from the results of de Fernex and Fusi [2013] and Hacon and Mckernan [2007] that the answer remains positive if the fibers are allowed to have log terminal singularities and dimension at most 3.

In higher dimensions, however, Totaro [2016b] showed that rationality does not

This work was partially supported by an NSF postdoctoral fellowship, DMS-1606460.
MSC2010: 14E08.
Keywords: rationality, specialization, fourfold, terminal singularities.

specialize among varieties with mild singularities. Namely, specialization fails in every dimension greater than 4 if terminal singularities (the mildest type of singularity arising from the minimal model program) are allowed, and in dimension 4 if canonical singularities (the second mildest type of singularity) are allowed. This left open the possibility that rationality specializes among terminal fourfolds. The purpose of this paper is to show that this fails too.

Theorem 2. *There is a flat projective family of fourfolds over a Zariski open neighborhood U of the origin $0 \in \mathbb{A}^1$ in the complex affine line such that*:

(1) *All the fibers have terminal singularities.*

(2) *The fibers over $U \setminus \{0\}$ are rational.*

(3) *The fiber over 0 is stably irrational.*

Our proof of Theorem 2 closely follows [Totaro 2016b]. There, starting from a stably irrational smooth quartic fourfold $Y \subset \mathbb{P}^5$ (known to exist by [Totaro 2016a]), Totaro constructs a family of fivefolds satisfying conditions (1)–(3) in Theorem 2 by deforming the cone over Y to rational fivefolds. More generally, starting from any smooth hypersurface $Y \subset \mathbb{P}^n$ which is Fano of index at least 2, his construction produces a family of n-folds satisfying (1) and (2), whose fiber over 0 is birational to $Y \times \mathbb{P}^1$. It is thus tempting to take $Y \subset \mathbb{P}^4$. However, then the only potential candidate for Y is a cubic threefold such that $Y \times \mathbb{P}^1$ is irrational, the existence of which is a difficult open problem.

Our idea is to instead take Y to be a quartic double solid. Then Y is a Fano threefold of index 2, and can be chosen to be stably irrational by Voisin's seminal work [2015]. Although Y is not a hypersurface in projective space, it is a hypersurface in a *weighted* projective space, which we show is enough to run Totaro's argument.

The natural question left open by this paper is whether rationality specializes among smooth varieties of dimension greater than 3.

Conventions. We work over the field of complex numbers $\mathbb{C}$. For positive integers $a_0, \dots, a_n$, we denote by $\mathbb{P}(a_0, \dots, a_n)$ the weighted projective space with weights a_i. We use superscripts to denote that a weight is repeated with multiplicity, e.g., $\mathbb{P}(1^4, 2) = \mathbb{P}(1, 1, 1, 1, 2)$. For a vector bundle $\mathcal{E}$ on a scheme S, the associated projective bundle is $\mathbb{P}(\mathcal{E}) = \mathrm{Proj}_S(\mathrm{Sym}(\mathcal{E}^\vee))$.

2. Proof of Theorem 2

Let $Y \to \mathbb{P}^3$ be a quartic double solid, i.e., a double cover of $\mathbb{P}^3$ branched along a smooth quartic surface. We regard Y as a hypersurface in the weighted projective space $\mathbb{P}(1^4, 2)$, cut out by a polynomial of the form

$$f_4(x_0, \dots, x_4) = x_4^2 - h_4(x_0, \dots, x_3),$$

where $h_4(x_0, \dots, x_3)$ is a quartic. Let $X \subset \mathbb{P}(1^4, 2, 1)$ be the cone over Y defined by the same polynomial $f_4(x_0, \dots, x_4)$ in the bigger weighted projective space $\mathbb{P}(1^4, 2, 1)$. For a stably irrational choice of Y, the variety X will form the central fiber in the promised family of fourfolds.

Lemma 3. *X is birational to $Y \times \mathbb{P}^1$, and has terminal singularities.*

Proof. This can be deduced from a general result on cones (see [Kollár 2013, §3.1]), but we give a direct argument. Let H denote the pullback of the hyperplane class on $\mathbb{P}^3$ to Y. Define

$$\pi : \widetilde{X} = \mathbb{P}(\mathcal{O}_Y(-H) \oplus \mathcal{O}_Y) \to Y.$$

There is a natural morphism $\widetilde{X} \to \mathbb{P}(1^4, 2, 1)$ given as follows. Let ζ denote the divisor corresponding to the relative $\mathcal{O}(1)$ line bundle on $\widetilde{X}$. Then

$$\pi_*(\mathcal{O}_{\widetilde{X}}(\zeta)) = \mathcal{O}_Y(H) \oplus \mathcal{O}_Y \quad \text{and} \quad \pi_*(\mathcal{O}_{\widetilde{X}}(2\zeta)) = \mathcal{O}_Y(2H) \oplus \mathcal{O}_Y(H) \oplus \mathcal{O}_Y.$$

Hence $\mathrm{H}^0(\widetilde{X}, \mathcal{O}_{\widetilde{X}}(\zeta)) \cong \mathbb{C}^4 \oplus \mathbb{C}$, and $\mathrm{H}^0(\widetilde{X}, \mathcal{O}_{\widetilde{X}}(2\zeta))$ has a canonical 1-dimensional subspace corresponding to the canonical section of $\mathcal{O}_Y(2H)$. This data specifies the morphism $\widetilde{X} \to \mathbb{P}(1^4, 2, 1)$. In fact, this morphism factors through $X \subset \mathbb{P}(1^4, 2, 1)$ and gives a resolution of singularities $f : \widetilde{X} \to X$ with a single exceptional divisor

$$E = \mathbb{P}(\mathcal{O}_Y) \subset \widetilde{X},$$

which is contracted to $[0, 0, 0, 0, 0, 1] \in X$. Thus the first claim of the lemma holds.

Note that X is normal with $\mathbb{Q}$-Cartier canonical divisor. We show that the discrepancy of the exceptional divisor E above is 1, so that X has terminal singularities, completing the proof. Write $K_{\widetilde{X}} = f^*(K_X) + aE$. Then by adjunction

$$K_E = (K_{\widetilde{X}} + E)|_E = (a+1)E|_E.$$

Observe that $E \cong Y$, so $K_E = -2H$, and $E = \zeta - \pi^* H$, so $E|_E = -H$. We conclude $a = 1$. □

Next, choose a nonzero polynomial $g_3(x_0, \dots, x_4) \in \mathrm{H}^0(\mathbb{P}(1^4, 2), \mathcal{O}(3))$ of weighted degree 3. We consider the flat family $\mathcal{X} \to \mathbb{A}^1$ over the affine line whose fiber $\mathcal{X}_t \subset \mathbb{P}(1^4, 2, 1)$ over $t \in \mathbb{A}^1$ is given by

$$f_4(x_0, \dots, x_4) + t g_3(x_0, \dots, x_4) x_5 = 0.$$

Note that $X = \mathcal{X}_0$.

Lemma 4. *There is a Zariski open neighborhood U of $0 \in \mathbb{A}^1$ such that*:

(1) *$\mathcal{X}_t$ has terminal singularities for all $t \in U$.*

(2) *$\mathcal{X}_t$ is rational for $t \in U \setminus \{0\}$.*

Proof. The fiber $\mathcal{X}_0$ has terminal singularities by Lemma 3. Since this condition is Zariski open in families [Nakayama 2004, Corollary VI.5.3], there is a Zariski open neighborhood U of $0 \in \mathbb{A}^1$ such that all fibers of $\mathcal{X}_U \to U$ are terminal. Further, observe that for $t \neq 0$, projection away from the x_5-coordinate gives a birational map from $\mathcal{X}_t$ to $\mathbb{P}(1^4, 2)$. Indeed, this map is an isomorphism over the locus where $g_3(x_0, \dots x_4) \neq 0$ in $\mathbb{P}(1^4, 2)$. Hence $\mathcal{X}_t$ is rational for $t \neq 0$. □

Now we can prove Theorem 2. By [Voisin 2015, Theorem 1.1], a very general quartic double solid is stably irrational. Taking such a Y in the above construction and combining Lemmas 3 and 4, we conclude that $\mathcal{X}_U \to U$ is a family of fourfolds satisfying all of the required conditions. □

Acknowledgements

Theorem 2 was conceived during Burt Totaro's talk on [Totaro 2016b] at the Higher Dimensional Algebraic Geometry Conference at the University of Utah in July 2016. I thank Burt for his comments on a draft of this paper, and the organizers of the conference for a stimulating environment. I also thank János Kollár for bringing [Timmerscheidt 1982] to my attention, and the referee for useful comments.

References

[de Fernex and Fusi 2013] T. de Fernex and D. Fusi, "Rationality in families of threefolds", *Rend. Circ. Mat. Palermo* (2) **62**:1 (2013), 127–135. MR Zbl

[Hacon and Mckernan 2007] C. D. Hacon and J. Mckernan, "On Shokurov's rational connectedness conjecture", *Duke Math. J.* **138**:1 (2007), 119–136. MR Zbl

[Hassett et al. 2016] B. Hassett, A. Pirutka, and Y. Tschinkel, "Stable rationality of quadric surface bundles over surfaces", preprint, 2016. arXiv

[Kollár 2013] J. Kollár, *Singularities of the minimal model program*, Cambridge Tracts in Mathematics **200**, Cambridge University Press, 2013. MR Zbl

[Nakayama 2004] N. Nakayama, *Zariski-decomposition and abundance*, MSJ Memoirs **14**, Mathematical Society of Japan, Tokyo, 2004. MR Zbl

[Timmerscheidt 1982] K. Timmerscheidt, "On deformations of three-dimensional rational manifolds", *Math. Ann.* **258**:3 (1982), 267–275. MR Zbl

[Totaro 2016a] B. Totaro, "Hypersurfaces that are not stably rational", *J. Amer. Math. Soc.* **29**:3 (2016), 883–891. MR Zbl

[Totaro 2016b] B. Totaro, "Rationality does not specialise among terminal varieties", *Math. Proc. Cambridge Philos. Soc.* **161**:1 (2016), 13–15. MR Zbl

[Voisin 2015] C. Voisin, "Unirational threefolds with no universal codimension 2 cycle", *Invent. Math.* **201**:1 (2015), 207–237. MR Zbl

Communicated by János Kollár
Received 2017-01-11 Revised 2017-07-07 Accepted 2017-09-28

aperry@math.columbia.edu Department of Mathematics, Columbia University, New York, NY, United States

ALGEBRA AND NUMBER THEORY 11:9 (2017)
dx.doi.org/10.2140/ant.2017.11.2197

Topological noetherianity for cubic polynomials

Harm Derksen, Rob H. Eggermont and Andrew Snowden

Let $P_3(\boldsymbol{k}^\infty)$ be the space of cubic polynomials in infinitely many variables over the algebraically closed field $\boldsymbol{k}$ (of characteristic $\neq 2, 3$). We show that this space is GL_∞-noetherian, meaning that any GL_∞-stable Zariski closed subset is cut out by finitely many orbits of equations. Our method relies on a careful analysis of an invariant of cubics we introduce called q-rank. This result is motivated by recent work in representation stability, especially the theory of twisted commutative algebras. It is also connected to uniformity problems in commutative algebra in the vein of Stillman's conjecture.

1. Introduction

Let $P_d(\boldsymbol{k}^n)$ be the space of degree d polynomials in n variables over an algebraically closed field $\boldsymbol{k}$ of characteristic $\neq 2, 3$. Let $P_d(\boldsymbol{k}^\infty)$ be the inverse limit of the $P_d(\boldsymbol{k}^n)$, equipped with the Zariski topology and its natural GL_∞ action (see Section 1G). This paper is concerned with the following question:

Question 1.1. Is the space $P_d(\boldsymbol{k}^\infty)$ noetherian with respect to the GL_∞ action? That is, can every Zariski closed GL_∞-stable subspace be defined by finitely many orbits of equations?

This question may seem somewhat esoteric, but it is motivated by recent work in the field of representation stability, in particular the theory of twisted commutative algebras; see Section 1C. It is also connected to certain uniformity questions in commutative algebra in the spirit of (the now resolved) Stillman's conjecture; see Section 1B.

For $d \leq 2$ the question is easy since one can explicitly determine the GL_∞ orbits on $P_d(\boldsymbol{k}^\infty)$. For $d \geq 3$ this is not possible, and the problem is much harder. The purpose of this paper is to settle the $d = 3$ case.

Derksen was supported by NSF grant DMS-1601229. Snowden was supported by NSF grants DMS-1303082 and DMS-1453893 and a Sloan Fellowship.

MSC2010: primary 13A50; secondary 13E05.

Keywords: noetherian, cubic, twisted commutative algebra.

Theorem 1.2. *Question 1.1 has an affirmative answer for $d = 3$.*

In fact, we prove a quantitative result in finitely many variables that implies the theorem in the limit. This may be of independent interest; see Section 1A for details.

1A. ***Overview of the proof.*** The key concept in the proof, and the focus of most of this paper, is the following notion of rank for cubic forms.

Definition 1.3. Let $f \in P_3(\boldsymbol{k}^n)$ with $n \leq \infty$. We define the *q-rank*[1] of f, denoted $\mathrm{qrk}(f)$, to be the minimal nonnegative integer r for which there is an expression $f = \sum_{i=1}^{r} \ell_i q_i$ with $\ell_i \in P_1(\boldsymbol{k}^n)$ and $q_i \in P_2(\boldsymbol{k}^n)$, or ∞ if no such r exists (which can only happen if $n = \infty$).

Example 1.4. For $n \leq \infty$, the cubic

$$x_1 y_1 z_1 + x_2 y_2 z_2 + \cdots + x_n y_n z_n = \sum_{i=1}^{n} x_i y_i z_i$$

has q-rank n. This is proved in Section 4. In particular, infinite q-rank is possible when $n = \infty$.

Example 1.5. The cubic $x^3 + y^3$ has q-rank 1, as follows from the identity

$$x^3 + y^3 = (x + y)(x^2 - xy + y^2).$$

The cubic $\sum_{i=1}^{2n} x_i^3$ therefore has q-rank at most n, and we expect it is exactly n.

Remark 1.6. The notion of q-rank is similar to some other invariants in the literature:

(a) Ananyan and Hochster [2016] defined a homogeneous polynomial to have *strength* $\geq k$ if it does not belong to an ideal generated by k forms of strictly lower degree. For cubics, q-rank is equal to strength plus one.

(b) A definition similar to strength also appears in [Kazhdan and Ziegler 2017].

(c) Davenport and Lewis [1964] defined an invariant h of cubics that is exactly q-rank.

(d) Inspired by Tao's blog post [2016], [Blasiak et al. 2017] introduced the notion of *slice rank* for tensors. Q-rank is basically a symmetric version of this.

Let $P_3(\boldsymbol{k}^\infty)_{\leq r}$ be the locus of forms f with $\mathrm{qrk}(f) \leq r$. This is the image of the map

$$P_2(\boldsymbol{k}^\infty)^r \times P_1(\boldsymbol{k}^\infty)^r \to P_3(\boldsymbol{k}^\infty), \qquad (q_1, \ldots, q_r, \ell_1, \ldots, \ell_r) \mapsto \sum_{i=1}^{r} \ell_i q_i.$$

The main theorem of [Eggermont 2015] implies that the domain of the above map is GL_∞-noetherian, and so, by standard facts (see [Draisma 2010, §3]), its image

[1] The q here is meant to indicate the presence of quadrics in the expression for f.

$P_3(\boldsymbol{k}^\infty)_{\le r}$ is as well. It follows that any GL_∞-stable closed subset of $P_3(\boldsymbol{k}^\infty)$ of bounded q-rank is cut out by finitely many orbits of equations. Theorem 1.2 then follows from the following result:

Theorem 1.7. *Any GL_∞-stable subset of $P_3(\boldsymbol{k}^\infty)$ containing forms of arbitrarily high q-rank is Zariski dense.*

To prove this theorem, one must show that if $f_1, f_2, \dots$ is a sequence in $P_3(\boldsymbol{k}^\infty)$ of unbounded q-rank then for any d there is a k such that the orbit closure of f_k projects surjectively onto $P_3(\boldsymbol{k}^d)$. We prove a quantitative version of this statement:

Theorem 1.8. *Let $f \in P_3(\boldsymbol{k}^n)$ have q-rank $r \gg 0$ (in fact, $r > \exp(240)$ suffices), and suppose $d < \frac{1}{3}\log(r)$. Then the orbit closure of f surjects onto $P_3(\boldsymbol{k}^d)$.*

The proof of this theorem is really the heart of the paper. The idea is as follows. Suppose that $f = \sum_{i=1}^m \ell_i q_i$ has large q-rank. We establish two key facts. First, after possibly degenerating f (i.e., passing to a form in the orbit closure) one can assume that the ℓ_i and the q_i are in separate sets of variables, while maintaining the assumption that f has large q-rank. This is useful when studying the orbit closure, as it allows us to move the ℓ_i and the q_i independently. Second, we show that the q_i have large rank in a very strong sense: namely, that within the linear span of the q_i there is a large-dimensional subspace such that every nonzero element of it has large rank. The results of [Eggermont 2015] then imply that the orbit closure of $(q_1, \dots, q_m; \ell_1, \dots, \ell_m)$ in $P_2(\boldsymbol{k}^n)^m \times P_1(\boldsymbol{k}^n)^m$ surjects onto $P_2(\boldsymbol{k}^d)^m \times P_1(\boldsymbol{k}^d)^m$, and this yields the theorem.

1B. ***Uniformity in commutative algebra.*** We now explain one source of motivation for Question 1.1. An *ideal invariant* is a rule that assigns to each homogeneous ideal I in each standard-graded polynomial $\boldsymbol{k}$-algebra A (in finitely many variables) a quantity $\nu_A(I) \in \mathbb{Z} \cup \{\infty\}$, such that $\nu_A(I)$ only depends on the pair (A, I) up to isomorphism. We say that ν is *cone-stable* if $\nu_{A[x]}(I[x]) = \nu_A(I)$, i.e., adjoining a new variable does not affect ν. The main theorem of [Erman et al. $\ge$ 2017] is (in part):

Theorem 1.9 [Erman et al. $\ge$ 2017]. *The following are equivalent*:

(a) *Let ν be a cone-stable ideal invariant that is upper semicontinuous in flat families, and let $\boldsymbol{d} = (d_1, \dots, d_r)$ be a tuple of nonnegative integers. Then there exists an integer B such that $\nu_A(I)$ is either infinite or at most B whenever I is an ideal generated by r elements of degrees $d_1, \dots, d_r$. (Crucially, B does not depend on A.)*

(b) *For every $\boldsymbol{d}$ as above, the space*

$$P_{d_1}(\boldsymbol{k}^\infty) \times \cdots \times P_{d_r}(\boldsymbol{k}^\infty)$$

is GL-*noetherian.*

Remark 1.10. Define an ideal invariant ν by taking $\nu_A(I)$ to be the projective dimension of I as an A-module. This is cone-stable and upper semicontinuous in flat families. The boundedness in Theorem 1.9(a) for this ν is exactly Stillman's conjecture, proved in [Ananyan and Hochster 2016].

Theorem 1.9 shows that Question 1.1 is intimately connected to uniformity questions in commutative algebra in the style of Stillman's conjecture. The results of [Erman et al. $\geq$ 2017] are actually more precise: if (b) holds for a single $\boldsymbol{d}$ then (a) holds for the corresponding $\boldsymbol{d}$. Thus, combined with Theorem 1.2, we obtain:

Theorem 1.11. *Let ν be a cone-stable ideal invariant that is upper semicontinuous in flat families. Then there exists an integer B such that $\nu(I)$ is either infinite or at most B, whenever I is generated by a single cubic form.*

The following two consequences of Theorem 1.11 are taken from [Erman et al. $\geq$ 2017].

Corollary 1.12. *Given a positive integer c there is an integer B such that the following holds: if $Y \subset \mathbb{P}^{n-1}$ is a cubic hypersurface containing finitely many codimension c linear subspaces then it contains at most B such subspaces.*

Corollary 1.13. *Given a positive integer c there is an integer B such that the following holds: if $Y \subset \mathbb{P}^{n-1}$ is a cubic hypersurface whose singular locus has codimension c then its singular locus has degree at most B.*

It would be interesting if these results could be proved by means of classical algebraic geometry. It would also be interesting to determine the bound B for some small values of c.

1C. *Twisted commutative algebras.* In this section we put $\boldsymbol{k} = \mathbb{C}$. Our original motivation for considering Question 1.1 came from the theory of twisted commutative algebras. Recall that a *twisted commutative algebra* (tca) over the complex numbers is a commutative unital associative $\mathbb{C}$-algebra A equipped with a polynomial action of GL_∞; see [Sam and Snowden 2012] for background. The easiest examples of tca's come by taking the symmetric algebra on a polynomial representation of GL_∞, for example $\mathrm{Sym}(\mathbb{C}^\infty)$ or $\mathrm{Sym}(\mathrm{Sym}^2(\mathbb{C}^\infty))$.

In recent years, tca's have appeared in several applications, for instance:

- Modules over the tca $\mathrm{Sym}(\mathbb{C}^\infty)$ are equivalent to **FI**-modules, as studied in [Church et al. 2015]. The structure of the module category was worked out in great detail in [Sam and Snowden 2016].
- Finite length modules over the tca $\mathrm{Sym}(\mathrm{Sym}^2(\mathbb{C}^\infty))$ are equivalent to algebraic representations of the infinite orthogonal group [Sam and Snowden 2015].
- Modules over tca's generated in degree 1 were used to study Δ-modules in [Snowden 2013], with applications to syzygies of Segre embeddings.

A tca A is *noetherian* if its module category is locally noetherian; explicitly, this means that any submodule of a finitely generated A-module is finitely generated. A major open question in the theory, first raised in [Snowden 2013], is as follows:

Question 1.14. Is every finitely generated tca noetherian?

So far, our knowledge on this question is extremely limited. For tca's generated in degrees ≤ 1 (or more generally, "bounded" tca's), noetherianity was proved in [Snowden 2013]. (It was later reproved in the special case of **FI**-modules in [Church et al. 2015].) For the tca's $\mathrm{Sym}(\mathrm{Sym}^2(\mathbb{C}^\infty))$ and $\mathrm{Sym}\big(\bigwedge^2(\mathbb{C}^\infty)\big)$, noetherianity was proved in [Nagpal et al. 2016]. No other cases are known. We remark that these known cases of noetherianity, limited though they are, have been crucial in applications.

Since noetherianity is such a difficult property to study, it is useful to consider a weaker notion. A tca A is *topologically noetherian* if every radical ideal is the radical of a finitely generated ideal. The results of [Eggermont 2015] show that tca's generated in degrees ≤ 2 are topologically noetherian. Topological noetherianity of the tca $\mathrm{Sym}(\mathrm{Sym}^d(\mathbb{C}^\infty))$ is equivalent to the noetherianity of the space $P_d(\mathbb{C}^\infty)$ appearing in Question 1.1. Thus Theorem 1.2 can be restated as follows:

Theorem 1.15. *The tca* $\mathrm{Sym}(\mathrm{Sym}^3(\mathbb{C}^\infty))$ *is topologically noetherian.*

This is the first noetherianity result for an unbounded tca generated in degrees ≥ 3.

1D. ***A result for tensors.*** Using similar methods, we can prove the following result:

Theorem 1.16. *The space* $P_1(\boldsymbol{k}^\infty) \mathbin{\widehat{\otimes}} P_1(\boldsymbol{k}^\infty) \mathbin{\widehat{\otimes}} P_1(\boldsymbol{k}^\infty)$ *is noetherian with respect to the action of the group* $\mathrm{GL}_\infty \times \mathrm{GL}_\infty \times \mathrm{GL}_\infty$, *where* $\widehat{\otimes}$ *denotes the completed tensor product.*

We plan to write a short note containing the proof.

1E. ***Draisma's theorem.*** After this paper appeared, Draisma [2017] answered Question 1.1 affirmatively for all d; in fact, he proved topological noetherianity of all polynomial representations, not just symmetric powers. While this result subsumes our Theorem 1.2, his proof does not give the more precise results found in Theorems 1.7 and 1.8. We believe these more precise results should hold in greater generality, and that they could be quite useful. We plan to pursue this matter in future work.

1F. ***Outline of paper.*** In Section 2 we establish a number of basic facts about q-rank. In Section 3 we use these facts to prove the main theorem. Finally, in Section 4, we compute the q-rank of the cubic in Example 1.4. This example is not used in the proof of the main theorem, but we thought it worthwhile to include one nontrivial computation of our fundamental invariant.

1G. ***Notation and terminology.*** Throughout we let $\boldsymbol{k}$ be an algebraically closed field of characteristic $\neq 2, 3$. The symbols E, V, and W always denote $\boldsymbol{k}$-vector spaces, perhaps infinite dimensional. We write $P_d(V) = \mathrm{Sym}^d(V)^*$ for the space of degree d polynomials on V equipped with the Zariski topology. Precisely, we identify $P_d(V)$ with the $\boldsymbol{k}$-points of the spectrum of the ring $\mathrm{Sym}(\mathrm{Sym}^d(V))$. When V is infinite dimensional the elements of $P_d(V)$ are certain infinite series and the functions on $P_d(V)$ are polynomials in coefficients. Whenever we speak of the orbit of an element of $P_d(V)$, we mean its $\mathrm{GL}(V)$ orbit.

2. Basic properties of q-rank

In this section, we establish a number of basic facts about q-rank. Throughout, V denotes a vector space and f a cubic in $P_3(V)$. Initially we allow V to be infinite dimensional, but following Proposition 2.5 it will be finite dimensional (though this is often not necessary).

Our first result is immediate, but worthwhile to write out explicitly.

Proposition 2.1 (subadditivity). *Suppose $f, g \in P_3(V)$. Then*

$$\mathrm{qrk}(f+g) \leq \mathrm{qrk}(f) + \mathrm{qrk}(g).$$

We defined q-rank from an algebraic point of view (number of terms in a certain sum). We now give a geometric characterization of q-rank that can, at times, be more useful.

Proposition 2.2. *We have $\mathrm{qrk}(f) \leq r$ if and only if there exists a linear subspace W of V of codimension at most r such that $f|_W = 0$.*

Proof. First suppose $\mathrm{qrk}(f) \leq r$, and write $f = \sum_{i=1}^r \ell_i q_i$. Then we can take $W = \bigcap_{i=1}^r \ker(\ell_i)$. This clearly has the requisite properties.

Now suppose W of codimension r is given. Let $v_{r+1}, v_{r+2}, \ldots$ be a basis for W, and complete it to a basis of V be adding vectors $v_1, \ldots, v_r$. Let $x_i \in P_1(V)$ be dual to v_i. We can then write $f = g + h$, where every term in g uses one of the variables $x_1, \ldots, x_r$, and these variables do not appear in h. Since $f|_W = 0$ by assumption and $g|_W = 0$ by its definition, we find $h|_W = 0$. But h only uses the variables $x_{r+1}, x_{r+2}, \ldots$, and these are coordinates on W, so we must have $h = 0$. Thus every term of f has one of the variables $\{x_1, \ldots, x_r\}$ in it, and so we can write $f = \sum_{i=1}^r x_i q_i$ for appropriate $q_i \in P_2(V)$, which shows $\mathrm{qrk}(f) \leq r$. $\square$

Remark 2.3. In the above proposition, $f|_W = 0$ means that the image of f in $P_3(W)$ is 0. It is equivalent to ask that $f(w) = 0$ for all $w \in W$.

The next result shows that one does not lose too much q-rank when passing to subspaces.

Proposition 2.4. *Suppose $W \subset V$ has codimension d. Then for $f \in P_3(V)$ we have*

$$\mathrm{qrk}(f) - d \le \mathrm{qrk}(f|_W) \le \mathrm{qrk}(f).$$

Proof. If $f = \sum_{i=1}^r \ell_i q_i$ then we obtain a similar expression for $f|_W$, which shows that $\mathrm{qrk}(f|_W) \le \mathrm{qrk}(f)$. Suppose now that $\mathrm{qrk}(f|_W) = r$, and let $W' \subset W$ be a codimension r subspace such that $f|_{W'} = 0$ (Proposition 2.2). Then W' has codimension $r + d$ in V, and so $\mathrm{qrk}(f) \le r + d$ (Proposition 2.2 again). □

Our next result shows that if V is infinite dimensional, then the q-rank of $f \in P_3(V)$ can be approximated by the q-rank of $f|_W$ for a large finite dimensional subspace W of V. This will be used at a key juncture to move from an infinite dimensional space down to a finite dimensional one.

Proposition 2.5. *Suppose $V = \bigcup_{i \in I} V_i$ (directed union). Then*

$$\mathrm{qrk}(f) = \sup_{i \in I} \mathrm{qrk}(f|_{V_i}).$$

We first give two lemmas. In what follows, for a finite dimensional vector space W we write $\mathrm{Gr}_r(W)$ for the Grassmannian of codimension r subspaces of W. For a $\boldsymbol{k}$-point x of $\mathrm{Gr}_r(W)$, we write E_x for the corresponding subspace of W. By "variety" we mean a reduced scheme of finite type over $\boldsymbol{k}$.

Lemma 2.6. *Let $W \subset V$ be finite dimensional vector spaces, and let $Z \subset \mathrm{Gr}_r(V)$ be a closed subvariety. Suppose that for every $\boldsymbol{k}$-point z of Z, the space $E_z \cap W$ has codimension r in W. Then there is a unique map of varieties $Z \to \mathrm{Gr}_r(W)$ that on $\boldsymbol{k}$-points is given by the formula $E \mapsto E \cap W$.*

Proof. Let $\mathrm{Hom}(V, \boldsymbol{k}^r)$ be the scheme of all linear maps $V \to \boldsymbol{k}^r$, and let $\mathrm{Surj}(V, \boldsymbol{k}^r)$ be the open subscheme of surjective linear maps. We identify $\mathrm{Gr}_r(V)$ with the quotient of $\mathrm{Surj}(V, \boldsymbol{k}^r)$ by the group GL_r. The quotient map $\mathrm{Surj}(V, \boldsymbol{k}^r) \to \mathrm{Gr}_r(V)$ sends a surjection to its kernel. Let $\widetilde{Z} \subset \mathrm{Surj}(V, \boldsymbol{k}^r)$ be the inverse image of Z. There is a natural map $\mathrm{Hom}(V, \boldsymbol{k}^r) \to \mathrm{Hom}(W, \boldsymbol{k}^r)$ given by restricting. By assumption, every closed point of $\widetilde{Z}$ maps into $\mathrm{Surj}(W, \boldsymbol{k}^r)$ under this map. Since $\mathrm{Surj}(W, \boldsymbol{k}^r)$ is open, it follows that the map $\widetilde{Z} \to \mathrm{Hom}(W, \boldsymbol{k}^r)$ factors through a unique map of schemes $\widetilde{Z} \to \mathrm{Surj}(W, \boldsymbol{k}^r)$. Since this map is GL_r-equivariant, it descends to the desired map $Z \to \mathrm{Gr}_r(W)$. If z is a $\boldsymbol{k}$-point of Z then it lifts to a $\boldsymbol{k}$-point $\tilde{z}$ of $\widetilde{Z}$, and the corresponding map $\varphi : V \to \boldsymbol{k}^r$ has $\ker(\varphi) = E_z$. The image of z in $\mathrm{Gr}_r(W)$ is $\ker(\varphi|_W) = E_z \cap W$, which establishes the stated formula for our map. □

Lemma 2.7. *Let $\{Z_i\}_{i \in I}$ be an inverse system of nonempty proper varieties over $\boldsymbol{k}$. Then $\varprojlim Z_i(\boldsymbol{k})$ is nonempty.*

Proof. If $\boldsymbol{k} = \mathbb{C}$ then $Z_i(\mathbb{C})$ is a nonempty compact Hausdorff space, and the result follows from the well-known (and easy) fact that an inverse limit of nonempty compact Hausdorff spaces is nonempty.

For a general field $\boldsymbol{k}$, we argue as follows. (We thank Bhargav Bhatt for this argument.) Let $|Z_i|$ be the Zariski topological space underlying the scheme Z_i, and let Z be the inverse limit of the $|Z_i|$. Since each $|Z_i|$ is a nonempty spectral space and the transition maps $|Z_i| \to |Z_j|$ are spectral (being induced from a map of varieties), Z is also a nonempty spectral space [Stacks 2005–, Lemmas 5.24.2 and 5.24.5]. It therefore has some closed point z. Let z_i be the image of z in $|Z_i|$.

We claim that z_i is closed for all i. Suppose not, and let $0 \in I$ be such that z_0 is not closed. Passing to a cofinal set in I, we may as well assume 0 is the unique minimal element. Let $\boldsymbol{k}(z_i)$ be the residue field of z_i, and let K be the direct limit of the $\boldsymbol{k}(z_i)$. The point z_i is then the image of a canonical map of schemes $a_i : \operatorname{Spec}(K) \to Z_i$. Since z_0 is not closed, it admits some specialization, so we may choose a valuation ring R in K and a nonconstant map of schemes $b_0 : \operatorname{Spec}(R) \to Z_0$ extending a_0. Since Z_i is proper, the map a_i extends uniquely to a map $b_i : \operatorname{Spec}(R) \to Z_i$. By uniqueness, the b_i are compatible with the transition maps, and so we get an induced map $b : |\operatorname{Spec}(R)| \to Z$ extending the map $a : |\operatorname{Spec}(K)| \to Z$. Since $|b_0|$ is induced from b, it follows that b is nonconstant. The image of the closed point in $\operatorname{Spec}(R)$ under b is then a specialization of z, contradicting the fact that z is closed. This completes the claim that z_i is closed.

Since z_i is closed, it is the image of a unique map $\operatorname{Spec}(\boldsymbol{k}) \to Z_i$ of $\boldsymbol{k}$-schemes. By uniqueness, these maps are compatible, and so give an element of $\varprojlim Z_i(\boldsymbol{k})$. $\square$

Proof of Proposition 2.5. First suppose that V_i is finite dimensional for all i. For $i \le j$ we have $\operatorname{qrk}(f|_{V_i}) \le \operatorname{qrk}(f|_{V_j})$ by Proposition 2.4, and so either $\operatorname{qrk}(f|_{V_i}) \to \infty$ or $\operatorname{qrk}(f|_{V_i})$ stabilizes. If $\operatorname{qrk}(f|_{V_i}) \to \infty$ then $\operatorname{qrk}(f) = \infty$ by Proposition 2.4 and we are done. Thus suppose $\operatorname{qrk}(f|_{V_i})$ stabilizes. Replacing I with a cofinal subset, we may as well assume $\operatorname{qrk}(f|_{V_i})$ is constant, say equal to r, for all i. We must show $\operatorname{qrk}(f) = r$. Proposition 2.4 shows that $r \le \operatorname{qrk}(f)$, so it suffices to show $\operatorname{qrk}(f) \le r$.

Let $Z_i \subset \operatorname{Gr}_r(V_i)$ be the closed subvariety consisting of all codimension r subspaces $E \subset V_i$ such that $f|_E = 0$. This is nonempty by Proposition 2.2 since $f|_{V_i}$ has q-rank r. Suppose $i \le j$ and z is a $\boldsymbol{k}$-point of Z_j, that is, E_z is a codimension r subspace of V_j on which f vanishes. Of course, f then vanishes on $V_i \cap E_z$, which has codimension at most r in V_i. Since $f|_{V_i}$ has q-rank exactly r, it cannot vanish on a subspace of codimension less than r (Proposition 2.2), and so $V_i \cap E_z$ must have codimension exactly r. Thus by Lemma 2.6, intersecting with V_i defines a map of varieties $Z_j \to \operatorname{Gr}_r(V_i)$. This maps into Z_i, and so for $i \le j$ we have a map $Z_j \to Z_i$. These maps clearly define an inverse system.

Appealing to Lemma 2.7 we see that $\varprojlim Z_i(\boldsymbol{k})$ is nonempty. Let $\{z_i\}_{i \in I}$ be a point in this inverse limit, and put $E_i = E_{z_i}$. Thus E_i is a codimension r subspace of V_i on which f vanishes, and for $i \le j$ we have $E_j \cap V_i = E_i$. It follows that $E = \bigcup_{i \in I} E_i$ is a codimension r subspace of V on which f vanishes, which shows $\operatorname{qrk}(f) \le r$ (Proposition 2.2).

We now treat the general case, where the V_i may not be finite dimensional. Write $V_i = \bigcup_{j \in J_i} W_j$ with W_j finite dimensional. Then $V = \bigcup_{i \in I} \bigcup_{j \in J_i} W_j$, so

$$\operatorname{qrk}(f) = \sup_{i \in I} \sup_{j \in J_i} \operatorname{qrk}(f|_{W_j}) = \sup_{i \in I} \operatorname{qrk}(f|_{V_i}).$$

This completes the proof. □

For the remainder of this section we assume that V is finite dimensional. If V is d-dimensional then the q-rank of any cubic in $P_3(V)$ is obviously bounded above by d. The next result gives an improved bound, and will be crucial in what follows.

Proposition 2.8. *Suppose* $\dim(V) = d$. *Then* $\operatorname{qrk}(f) \le d - \xi(d)$, *where*

$$\xi(d) = \left\lfloor \frac{\sqrt{8d+17}-3}{2} \right\rfloor.$$

Note that $\xi(d) \approx \sqrt{2d}$.

Proof. Let k be the largest integer such that $\binom{k+1}{2} + k - 1 \le d$. Then the hypersurface $f = 0$ contains a linear subspace of dimension at least k by [Harris et al. 1998, Lemma 3.9]. It follows from Proposition 2.2 that $\operatorname{qrk}(f) \le d - k$. Some simple algebra shows that $k = \xi(d)$. □

Suppose that $f = \sum_{i=1}^{n} \ell_i q_i$ is a cubic. Eventually, we want to show that if f has large q-rank then its orbit under $\mathrm{GL}(V)$ is large. For studying the orbit, it would be convenient if the ℓ_i and the q_i were in separate sets of variables, as then they could be moved independently under the group. This motivates the following definition.

Definition 2.9. We say that a cubic $f \in P_3(V)$ is *separable*[2] if there is a direct sum decomposition $V = V_1 \oplus V_2$ and an expression $f = \sum_{i=1}^{n} \ell_i q_i$ with $\ell_i \in P_1(V_1)$ and $q_i \in P_2(V_2)$.

Now, if we have a cubic f of high q-rank we cannot conclude, simply based on its high q-rank, that it is separable. Fortunately, the following result shows that if we are willing to degenerate f a bit (which is fine for our ultimate applications), then we can make it separable while retaining high q-rank.

Proposition 2.10. *Suppose that* $f \in P_3(V)$ *has q-rank* r. *Then the orbit closure of* f *contains a separable cubic* g *satisfying* $\frac{1}{2}\xi(r) \le \operatorname{qrk}(g)$.

Proof. Let $\{x_i\}$ be a basis for $P_1(V)$. After possibly making a linear change of variables, we can write $f = \sum_{i=1}^{r} x_i q_i$. Write $f = f_1 + f_2 + f_3$, where f_i is homogeneous of degree i in the variables $\{x_1, \ldots, x_r\}$. Since f_3 has degree 3 in the variables $\{x_1, \ldots, x_r\}$, it can contain no other variables, and can thus be regarded as an element of $P_3(\boldsymbol{k}^r)$. Therefore, by Proposition 2.8, we have $\operatorname{qrk}(f_3) \le r - \xi(r)$.

[2] This notion of separable is unrelated to the notion of separability of univariate polynomials. We do not expect this to cause confusion.

After possibly making a linear change of variables in $\{x_1, \dots, x_r\}$, we can write $f_3 = \sum_{i=\xi(r)+1}^{r} x_i q_i'$ for some q_i'. Let f' and f_j' be the result of setting $x_i = 0$ in f and f_j, respectively, for $\xi(r) < i \le r$. We have $\mathrm{qrk}(f') \ge \xi(r)$ by Proposition 2.4. Of course, $f_3' = 0$, so $f' = f_1' + f_2'$. By subadditivity (Proposition 2.1), at least one of f_1' or f_2' has q-rank $\ge \frac{1}{2}\xi(r)$.

We have $f_1 = \sum_{i=1}^{r} x_i q_i''$, where q_i'' is a quadratic form in the variables x_i with $i > r$. Thus f_1 and f_1' are separable. We have $f_2 = \sum_{1 \le i \le j \le r} x_i x_j \ell_{i,j}$, where $\ell_{i,j}$ is a linear form in the variables x_i with $i > r$. Thus f_2 and f_2' are separable.

To complete the proof, it suffices to show that f_1' and f_2' belong to the orbit closure of f, as we can then take $g = f_1'$ or $g = f_2'$. It is clear that f' is in the orbit closure of f, so it suffices to show that f_1' and f_2' are in the orbit closure of f'. Consider the element γ_t of GL_n defined by

$$\gamma_t(x_i) = \begin{cases} t^2 x_i, & 1 \le i \le r, \\ t^{-1} x_i, & r < i \le n. \end{cases}$$

Then $\gamma_t(f_1') = f_1'$ and $\gamma_t(f_2') = t^3 f_2'$. Thus $\lim_{t \to 0} \gamma_t(f') = f_1'$. A similar construction shows that f_2' is in the orbit closure of f'. $\square$

Suppose that $f = \sum_{i=1}^{n} \ell_i q_i$ is a cubic of high q-rank. One would like to be able to conclude that the q_i then have high ranks as well. We now prove two results along this line. For a linear subspace $Q \subset P_2(V)$, we let $\mathrm{maxrank}(Q)$ be the maximum of the ranks of elements of Q, and we let $\mathrm{minrank}(Q)$ be the minimum of the ranks of the nonzero elements of Q (or 0 if $Q = 0$).

Proposition 2.11. *Suppose $f = \sum_{i=1}^{n} \ell_i q_i$ has q-rank r, and let $Q \subset P_2(V)$ be the span of the q_i. Then for every subspace Q' of Q we have*

$$\mathrm{codim}(Q : Q') + \mathrm{maxrank}(Q') \ge r.$$

Proof. We may as well assume that ℓ_i and q_i are linearly independent. Thus $\dim(Q) = n$. Let Q' be a subspace of dimension $n - d$. After making a linear change of variables in the q_i and ℓ_i, we may as well assume that Q' is the span of $q_1, \dots, q_{n-d}$. Let $t = \mathrm{maxrank}(Q')$. We must show that $d + t \ge r$. Let $q' \in Q'$ have rank t. Choose a basis $\{x_i\}$ of $P_1(V)$ so that $q' = x_1^2 + \cdots + x_t^2$. If some q_i for $1 \le i \le n - d$ had a term of the form $x_j x_k$ with $j, k > t$ then some linear combination of q_i and q' would have rank $> t$, a contradiction. Thus every term of q_i, for $1 \le i \le n-d$, has a variable of index $\le t$, and so we can write $q_i = \sum_{j=1}^{t} x_j m_{i,j}$, where $m_{i,j} \in P_1(V)$. But now

$$f = \sum_{i=1}^{n-d} \ell_i q_i + \sum_{i=n-d+1}^{n} \ell_i q_i = \sum_{j=1}^{t} x_j q_j' + \sum_{i=n-d+1}^{n} \ell_i q_i,$$

where $q_j' = \sum_{i=1}^{n-d} \ell_i m_{i,j}$. This shows $r = \mathrm{qrk}(f) \le t + d$, completing the proof. $\square$

In our eventual application, it is actually minrank that is more important than maxrank. Fortunately, the above result on maxrank automatically gives a result for minrank, thanks to the following general proposition.

Proposition 2.12. *Let $Q \subset P_2(V)$ be a linear subspace and let r be a positive integer. Suppose that*

$$\operatorname{codim}(Q : Q') + \operatorname{maxrank}(Q') \geq r$$

holds for all linear subspaces $Q' \subset Q$. Let k and s be positive integers satisfying

$$(2^k - 1)(s - 1) + k \leq r. \tag{2.13}$$

Then there exists a k-dimensional linear subspace $Q' \subset Q$ with $\operatorname{minrank}(Q') \geq s$.

Lemma 2.14. *Let $q_1, \ldots, q_n \in P_2(V)$ be quadratic forms of rank $< s$. Suppose there is a linear combination of the q_i that has rank at least t. Then there is a linear combination q' of the q_i satisfying $t \leq \operatorname{rank}(q') \leq t + s - 2$.*

Proof. Let $q' = \sum_{i=1}^{k} a_i q_i$ be a linear combination of the q_i with rank $\geq t$ and k minimal. Since $\operatorname{rank}(q_k) \leq s - 1$, it follows that $\operatorname{rank}(q' - a_k q_k) \geq \operatorname{rank}(q') - (s - 1)$. Thus if $\operatorname{rank}(q') \geq t + s - 1$ then $\sum_{i=1}^{k-1} a_i q_i$ would have rank $\geq t$, contradicting the minimality of k. Therefore $\operatorname{rank}(q') \leq t + s - 2$. □

Proof of Proposition 2.12. Suppose that $q_1, \ldots, q_n$ forms a basis for Q such that $(\operatorname{rank}(q_1), \ldots, \operatorname{rank}(q_n))$ is lexicographically minimal. In particular, this implies that $\operatorname{rank}(q_1) \leq \cdots \leq \operatorname{rank}(q_n)$. If $\operatorname{rank}(q_{n-k+1}) \geq s$, then lexicographic minimality ensures that any nontrivial linear combination of $q_{n-k+1}, \ldots, q_n$ has rank at least s, and so we can take Q' to be the span of these forms. Thus suppose that $\operatorname{rank}(q_{n-k+1}) < s$. In what follows, we put $m_i = (2^i - 1)(s - 1) + 1$. Note that $m_k \leq r$. In fact, $n - r + m_k \leq n - k + 1$, and so $\operatorname{rank}(q_{n-r+m_k}) < s$.

For $1 \leq \ell \leq k$, consider the following statement:

(S_ℓ) There exist linearly independent $p_1, \ldots, p_\ell$ such that: (i) p_i is a linear combination of $q_1, \ldots, q_{n-r+m_i}$; (ii) $m_i \leq \operatorname{rank}(p_i) \leq m_i + s - 2$; and (iii) the span of $p_1, \ldots, p_\ell$ has minrank at least s.

We prove (S_ℓ) by induction on ℓ. Of course, (S_k) implies the proposition.

First consider the case $\ell = 1$. The statement (S_1) asserts that there exists a nonzero linear combination p of $q_1, \ldots, q_{n-r+s}$ such that $s \leq \operatorname{rank}(p) \leq 2s - 2$. Since the span of $q_1, \ldots, q_{n-r+s}$ has codimension $r - s$ in Q, our assumption guarantees that some linear combination p of these forms has rank at least s. Since each form has rank $< s$, Lemma 2.14 ensures we can find p with $\operatorname{rank}(p) \leq s + (s - 2)$.

We now prove (S_ℓ) assuming ($S_{\ell-1}$). Let $(p_1, \ldots, p_{\ell-1})$ be the tuple given by ($S_{\ell-1}$). The span of $q_1, \ldots, q_{n-r+m_\ell}$ has codimension $r - m_\ell$ in Q, and so our assumption guarantees that some linear combination p_ℓ has rank at least m_ℓ. By

Lemma 2.14, we can ensure that this p_ℓ has rank at most $m_\ell + s - 2$. Thus (i) and (ii) in (S_ℓ) are established.

We now show that any nontrivial linear combination $\sum_{i=1}^{\ell} \lambda_i p_i$ has rank at least s, which will show that the p_i are linearly independent and establish (iii) in (S_ℓ). If $\lambda_\ell = 0$ then the rank is at least s by the assumption on $(p_1, \ldots, p_{\ell-1})$. Thus assume $\lambda_\ell \neq 0$. We have

$$\operatorname{rank}\left(\sum_{i=1}^{\ell-1} \lambda_i p_i\right) \leq \sum_{i=1}^{\ell-1} \operatorname{rank}(p_i) \leq \sum_{i=1}^{\ell-1} (m_i + s - 2) = m_\ell - s.$$

Since $\operatorname{rank}(p_\ell) \geq m_\ell$, we thus see that $\sum_{i=1}^{\ell} \lambda_i p_i$ has rank at least s, which completes the proof. □

Remark 2.15. Proposition 2.12 is not specific to ranks of quadratic forms; it applies to any subadditive invariant on a vector space.

Combining the Propositions 2.11 and 2.12, we obtain:

Corollary 2.16. *Suppose $f = \sum_{i=1}^{n} \ell_i q_i$ has q-rank r, let Q be the span of the q_i, and let k and s be positive integers such that* (2.13) *holds. Then there exists a k-dimensional linear subspace $Q' \subset Q$ with* $\operatorname{minrank}(Q') \geq s$.

3. Proof of Theorem 1.2

We now prove the main theorems of the paper. We require the following result; see [Eggermont 2015, Proposition 3.3] and its proof.

Theorem 3.1. *Let x be a point in $P_2(V)^n \times P_1(V)^m$, with V finite dimensional. Write x as $(q_1, \ldots, q_n; \ell_1, \ldots, \ell_m)$, and let $Q \subset P_2(V)$ be the span of the q_i. Let W be a d-dimensional subspace of V. Suppose that $\ell_1, \ldots, \ell_m$ are linearly independent and that* $\operatorname{minrank}(Q) \geq dn2^n + 2(n+1)m$. *Then the orbit closure of x surjects onto $P_2(W)^n \times P_1(W)^m$.*

We begin by proving an analog of the above theorem for $P_3(V)$.

Theorem 3.2. *Suppose V is finite dimensional. Let $f \in P_3(V)$ have q-rank r and let W be a d-dimensional subspace of V with*

$$(2^d - 1)(d^2 2^d + 2(d+1)d - 1) + d \leq \tfrac{1}{2}\xi(r).$$

Then the orbit closure of f surjects onto $P_3(W)$.

Proof. Applying Proposition 2.10, let g be a separable cubic in the orbit closure of f satisfying $\frac{1}{2}\xi(r) \leq \operatorname{qrk}(g)$. Write $g = \sum_{i=1}^{n} \ell_i q_i$, where $\ell_i \in P_1(V_1)$ and $q_i \in P_2(V_2)$, with $V = V_1 \oplus V_2$, and the ℓ_i and q_i are linearly independent. Let Q be the span of the q_i. Put $s = d^2 2^d + 2(d+1)d$ and $k = d$. Note that

$$(2^k - 1)(s - 1) + k \leq \tfrac{1}{2}\xi(r).$$

By Corollary 2.16 we can therefore find a $k = d$ dimensional subspace Q' of Q with $\operatorname{minrank}(Q') \geq s$. Making a linear change of variables, we can assume Q' is the span of $q_1, \dots, q_d$. Let $g' = \sum_{i=1}^{d} \ell_i q_i$. This is in the orbit closure of g (and thus f) since it is obtained by setting $\ell_i = 0$ for $i > d$. It is crucial here that the q_i and the ℓ_i are in different sets of variables, so that setting some of the ℓ_i to 0 does not change the q_i. By Theorem 3.1, the orbit closure of $(q_1, \dots, q_d; \ell_1, \dots, \ell_d)$ in $P_2(V)^d \times P_1(V)^d$ surjects onto $P_2(W)^d \times P_1(W)^d$. Now let $h \in P_3(W)$. Since $\dim(W) = d$ we can write $h = \sum_{i=1}^{d} \ell_i' q_i'$ with $\ell_i' \in P_1(W)$ and $q_i' \in P_2(W)$. Pick $\gamma_t \in \mathrm{GL}(V)$ such that $(q_1', \dots, q_d'; \ell_1', \dots, \ell_d')$ is in the image of

$$\lim_{t \to 0} \gamma_t \cdot (q_1, \dots, q_d; \ell_1 \dots, \ell_d).$$

Then h is the image of $\lim_{t \to 0} \gamma_t \cdot g'$, which completes the proof. $\square$

Corollary 3.3 (Theorem 1.8). *Suppose that $f \in P_3(V)$ has q-rank $r > \exp(240)$ and let W be a subspace of V of dimension d with $d < \frac{1}{3} \log r$. Then the orbit closure of f surjects onto $P_3(W)$.*

Proof. By definition of ξ, we have $a \leq \xi(r)$ (for an integer a) if and only if $\binom{a+1}{2} + a - 1 \leq r$. So the condition in Theorem 3.2 is equivalent to $\binom{D+1}{2} + D - 1 \leq r$, where

$$D = 2(2^d - 1)(d^2 2^d + 2(d+1)d - 1) + 2d$$

is twice the left side of the inequality in Theorem 3.2. Now, $\binom{D+1}{2} + D - 1$ is equal to $4 \cdot d^4 \cdot 16^d$ plus lower order terms, and is therefore less than 20^d for $d \gg 0$; in fact, $d > 80$ is sufficient. Thus for $d > 80$ it is enough that $d < \log r / \log 20$; since $\log(20) < 3$, it is enough that $d < \frac{1}{3} \log(r)$. Thus for $80 < d < \frac{1}{3} \log(r)$, the orbit closure of f surjects onto $P_3(W)$. But it obviously then surjects onto smaller subspaces as well, so we only need to assume $80 < \frac{1}{3} \log(r)$. $\square$

Theorem 3.4 (Theorem 1.7). *Let V be infinite dimensional. Suppose $Z \subset P_3(V)$ is Zariski closed*, $\mathrm{GL}(V)$*-stable, and contains elements of arbitrarily high q-rank. Then $Z = P_3(V)$.*

Proof. It suffices to show that Z surjects onto $P_3(W)$ for all finite dimensional $W \subset V$. Thus let W of dimension d be given. Let r be sufficiently large so that the inequality in Theorem 3.2 is satisfied and let $f \in Z$ have q-rank at least r. By Proposition 2.5, there exists a finite dimensional subspace V' of V containing W such that $f|_{V'}$ has q-rank at least r. Theorem 3.2 implies that the orbit closure of $f|_{V'}$ surjects onto $P_3(W)$. Since Z surjects onto the orbit closure of $f|_{V'}$, the result follows. $\square$

It was explained in the introduction how this implies Theorem 1.2, so the proof is now complete.

4. A computation of q-rank

Fix a positive integer n, and consider the cubic

$$f = x_1 y_1 z_1 + \cdots + x_n y_n z_n$$

in the polynomial ring $\boldsymbol{k}[x_i, y_i, z_i]_{1 \le i \le n}$ introduced in Example 1.4. We now show:

Proposition 4.1. *The above cubic f has q-rank n.*

It is clear that $\mathrm{qrk}(f) \le n$. To prove equality, it suffices by Proposition 2.2 to show that $f|_V \ne 0$ if V is a codimension $n-1$ subspace of $\boldsymbol{k}^{3n}$. This is exactly the content of the following proposition.

Proposition 4.2. *Let V be a vector space of dimension $2n+1$ and $(x_i, y_i, z_i)_{1 \le i \le n}$ a collection of elements that span $P_1(V)$. Then $f = x_1 y_1 z_1 + \cdots + x_n y_n z_n \in P_3(V)$ is nonzero.*

Proof. Arrange the given elements in a matrix as follows:

$$\begin{pmatrix} x_1 & y_1 & z_1 \\ \vdots & \vdots & \vdots \\ x_n & y_n & z_n \end{pmatrix}.$$

Note that we are free to permute the rows and apply permutations within a row without changing the value of f, e.g., we can switch the values of x_1 and y_1, or switch (x_1, y_1, z_1) with (x_2, y_2, z_2), without changing f. We now proceed to find a basis for V among the elements in the matrix according to the following three-phase procedure.

Phase 1. Find a nonzero element of the matrix, and move it (using the permutations mentioned above) to the x_1 position. Now in rows $2, \ldots, n$ find an element that is not in the span of x_1 (if one exists) and move it to the x_2 position. Now in rows $3, \ldots, n$ find an element that is not in the span of x_1 and x_2 (if one exists) and move it to the x_3 position. Continue in this manner until it is no longer possible; suppose we go r steps. At this point, $x_1, \ldots, x_r$ are linearly independent, and x_i, y_i, and z_i, for $r < i$ all belong to their span.

Phase 2. From rows $1, \ldots, r$ find an element in the second or third column not in the span of $x_1, \ldots, x_r$ and move it (using permutations that fix the set $\{x_1, \ldots, x_r\}$) to the y_1 position. Next, from rows $2, \ldots, r$ find an element in the second or third column not in the span of $x_1, \ldots, x_r, y_1$ and move it to the y_2 position. Continue in this manner until it is no longer possible; suppose we go s steps. At this point, $x_1, \ldots, x_r, y_1, \ldots, y_s$ form a linearly independent set, and the elements y_i, z_i for $s < i \le r$ belong to their span. The conclusion from Phase 1 still holds as well.

Phase 3. Now carry out the same procedure in the third column. That is, from rows $1, \dots, s$ find an element in the third column not in the span of $x_1, \dots, x_r, y_1, \dots, y_s$ and move it (by permuting rows) to the z_1 position. Then from rows $2, \dots, s$ find an element in the third column not in the span of $x_1, \dots, x_r, y_1, \dots, y_s, z_1$ and move it to the z_2 position. Continue in this manner until it is no longer possible; suppose we go t steps. At this point, $x_1, \dots, x_r, y_1, \dots, y_s, z_1, \dots, z_t$ forms a basis of V. The conclusions from Phases 1 and 2 still hold.

For clarity, we write $X_1, \dots, X_r, Y_1, \dots, Y_s, Z_1, \dots, Z_t$ for our basis. We note that because $\dim(V) > 2n$ we must have $t \geq 1$. The ring $\mathrm{Sym}(V^*)$ is identified with the polynomial ring in the X, Y, Z variables. We now determine the coefficient of $X_1 Y_1 Z_1$ in $m_i = x_i y_i z_i$. If $i > r$ then m_i has degree 3 in the X variables, and so the coefficient is 0. If $s < i \leq r$ then m_i has degree 0 in the Z variables, and so again the coefficient is 0. Finally, suppose that $i \leq s$. Then $m_i = X_i Y_i z_i$. The only way this can contain $X_1 Y_1 Z_1$ is if $i = 1$. We thus see that the coefficient of $X_1 Y_1 Z_1$ in m_i is 0 except for $i = 1$, in which case it is 1, and so $f = \sum_{i=1}^n m_i$ is nonzero. □

Remark 4.3. It follows from the above results and Proposition 2.5 that the cubic $\sum_{i=1}^\infty x_i y_i z_i$ has infinite q-rank.

Acknowledgements

We thank Bhargav Bhatt, Jan Draisma, Daniel Erman, Mircea Mustata, and Steven Sam for helpful discussions.

References

[Ananyan and Hochster 2016] T. Ananyan and M. Hochster, "Small subalgebras of polynomial rings and Stillman's conjecture", preprint, 2016. arXiv

[Blasiak et al. 2017] J. Blasiak, T. Church, H. Cohn, J. A. Grochow, E. Naslund, W. F. Sawin, and C. Umans, "On cap sets and the group-theoretic approach to matrix multiplication", *Discrete Anal.* (2017), paper no. 3, 27 pp. MR

[Church et al. 2015] T. Church, J. S. Ellenberg, and B. Farb, "FI-modules and stability for representations of symmetric groups", *Duke Math. J.* **164**:9 (2015), 1833–1910. MR Zbl

[Davenport and Lewis 1964] H. Davenport and D. J. Lewis, "Non-homogeneous cubic equations", *J. London Math. Soc.* **39** (1964), 657–671. MR Zbl

[Draisma 2010] J. Draisma, "Finiteness for the k-factor model and chirality varieties", *Adv. Math.* **223**:1 (2010), 243–256. MR Zbl

[Draisma 2017] J. Draisma, "Topological Noetherianity of polynomial functors", preprint, 2017. arXiv

[Eggermont 2015] R. H. Eggermont, "Finiteness properties of congruence classes of infinite-by-infinite matrices", *Linear Algebra Appl.* **484** (2015), 290–303. MR Zbl

[Erman et al. $\geq$ 2017] D. Erman, S. Sam, and A. Snowden, "Connections between commutative algebra and twisted commutative algebras", in preparation.

[Harris et al. 1998] J. Harris, B. Mazur, and R. Pandharipande, "Hypersurfaces of low degree", *Duke Math. J.* **95**:1 (1998), 125–160. MR Zbl

[Kazhdan and Ziegler 2017] D. Kazhdan and T. Ziegler, "On the bias of cubic polynomials", preprint, 2017. arXiv

[Nagpal et al. 2016] R. Nagpal, S. V. Sam, and A. Snowden, "Noetherianity of some degree two twisted commutative algebras", *Selecta Math.* (*N.S.*) **22**:2 (2016), 913–937. MR Zbl

[Sam and Snowden 2012] S. V. Sam and A. Snowden, "Introduction to twisted commutative algebras", preprint, 2012. arXiv

[Sam and Snowden 2015] S. V. Sam and A. Snowden, "Stability patterns in representation theory", *Forum Math. Sigma* **3** (2015), e11, 108 pp. MR Zbl

[Sam and Snowden 2016] S. V. Sam and A. Snowden, "GL-equivariant modules over polynomial rings in infinitely many variables", *Trans. Amer. Math. Soc.* **368**:2 (2016), 1097–1158. MR Zbl

[Snowden 2013] A. Snowden, "Syzygies of Segre embeddings and Δ-modules", *Duke Math. J.* **162**:2 (2013), 225–277. MR Zbl

[Stacks 2005–] "The Stacks project", electronic reference, 2005–, http://stacks.math.columbia.edu.

[Tao 2016] T. Tao, "A symmetric formulation of the Croot–Lev–Pach–Ellenberg–Gijswijt capset bound", blog post, 2016, https://terrytao.wordpress.com/2016/05/18/a.

Communicated by Victor Reiner
Received 2017-02-08 Revised 2017-06-16 Accepted 2017-06-20

hderksen@umich.edu *Department of Mathematics, University of Michigan, Ann Arbor, MI, United States*

r.h.eggermont@tue.nl *Faculteit Wiskunde en Informatica, Eindhoven University of Technology, Eindhoven, The Netherlands*

asnowden@umich.edu *Department of Mathematics, University of Michigan, Ann Arbor, MI, United States*

Guidelines for Authors

Authors may submit manuscripts in PDF format on-line at the Submission page at the ANT website.

Originality. Submission of a manuscript acknowledges that the manuscript is original and and is not, in whole or in part, published or under consideration for publication elsewhere. It is understood also that the manuscript will not be submitted elsewhere while under consideration for publication in this journal.

Language. Articles in *ANT* are usually in English, but articles written in other languages are welcome.

Length There is no a priori limit on the length of an *ANT* article, but *ANT* considers long articles only if the significance-to-length ratio is appropriate. Very long manuscripts might be more suitable elsewhere as a memoir instead of a journal article.

Required items. A brief abstract of about 150 words or less must be included. It should be self-contained and not make any reference to the bibliography. If the article is not in English, two versions of the abstract must be included, one in the language of the article and one in English. Also required are keywords and subject classifications for the article, and, for each author, postal address, affiliation (if appropriate), and email address.

Format. Authors are encouraged to use LaTeX but submissions in other varieties of TeX, and exceptionally in other formats, are acceptable. Initial uploads should be in PDF format; after the refereeing process we will ask you to submit all source material.

References. Bibliographical references should be complete, including article titles and page ranges. All references in the bibliography should be cited in the text. The use of BibTeX is preferred but not required. Tags will be converted to the house format, however, for submission you may use the format of your choice. Links will be provided to all literature with known web locations and authors are encouraged to provide their own links in addition to those supplied in the editorial process.

Figures. Figures must be of publication quality. After acceptance, you will need to submit the original source files in vector graphics format for all diagrams in your manuscript: vector EPS or vector PDF files are the most useful.

Most drawing and graphing packages (Mathematica, Adobe Illustrator, Corel Draw, MATLAB, etc.) allow the user to save files in one of these formats. Make sure that what you are saving is vector graphics and not a bitmap. If you need help, please write to graphics@msp.org with details about how your graphics were generated.

White space. Forced line breaks or page breaks should not be inserted in the document. There is no point in your trying to optimize line and page breaks in the original manuscript. The manuscript will be reformatted to use the journal's preferred fonts and layout.

Proofs. Page proofs will be made available to authors (or to the designated corresponding author) at a Web site in PDF format. Failure to acknowledge the receipt of proofs or to return corrections within the requested deadline may cause publication to be postponed.

Algebra & Number Theory

Volume 11 No. 9 2017

1937-0652(2017)11:9;1-C