FERMAT DESCENT

SANTIAGO ARANGO-PIÑEROS

ABSTRACT. Descent theory (a modern formulation of Fermat's classical method of *infinite descent*) is a powerful tool in arithmetic geometry. In this article, we reinterpret descent theory through the lens of quotient stacks and apply it in the setting where it first arose: the Diophantine study of generalized Fermat equations

$$A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c = 0. \tag{1}$$

We focus on understanding the arithmetic of the stacks that arise from the study of primitive integral solutions to Equation (1), rather than on solving any particular instance of the equation.

CONTENTS

1. INTRODUCTION

1.1. **The framework of Poonen–Schaefer–Stoll.** Poonen, Schaefer, and Stoll [PSS07, Theorem 1.1] provably computed the finite set of primitive integral solutions of the generalized Fermat equation

$$F\colon \mathsf{x}^2 + \mathsf{y}^3 + \mathsf{z}^7 = 0. \tag{2}$$

(Recall that $(x, y, z) \in \mathbb{Z}^3$ is called primitive when $\gcd(x, y, z) = 1$.) If $\mathcal{U}$ is the punctured cone associated to $F \subset \mathbb{A}^3_{\mathbb{Z}}$ (i.e., the subscheme obtained by deleting $\{\mathsf{x} = \mathsf{y} = \mathsf{z} = 0\}$ to F), then $\mathcal{U}(\mathbb{Z})$ is identified with the set of primitive integral solutions to Equation (2). A preliminary step in their method is to consider the quotient stack $[\mathcal{U}/\mathbb{G}_{\mathrm{m}}]$, where the multiplicative group $\mathbb{G}_{\mathrm{m}}$ acts by

$$(x, y, z) \cdot \lambda := (\lambda^{21}x, \lambda^{14}y, \lambda^{6}z).$$

After inverting the bad primes $\mathcal{S} = \{2, 3, 7\}$, the stack $[\mathcal{U}/\mathbb{G}_{\mathrm{m}}]$ becomes isomorphic to the stack $\mathbb{P}^1(2, 3, 7)$; this is the projective line $\mathbb{P}^1_{\mathbb{Z}}$ rooted at the irreducible horizontal divisors $0, 1,$ and ∞ with multiplicities $2, 3,$ and 7, respectively.

$$[\mathcal{U}/\mathbb{G}_{\mathrm{m}}]_{\mathbb{Z}[1/42]} \cong \mathbb{P}^1(2,3,7)_{\mathbb{Z}[1/42]}. \tag{3}$$

After this preliminary result, the first step in the method is to find a geometrically Galois Belyi map $\phi\colon X \to \mathbb{P}^1$, with ramification indices $2, 3, 7$ above $0, 1, \infty$. To find $\mathcal{U}(\mathbb{Z})$, it is enough to calculate the sets of rational points on the curves X_τ for an explicit finite set of twists $\phi_\tau\colon X_\tau \to \mathbb{P}^1$ of the map $\phi\colon X \to \mathbb{P}^1$. Our interpretation of the method of [PSS07, Section 3] tersely summarized here is henceforth referred to as Fermat descent.

1.2. **Our main theorem.** To generalize the method of Fermat descent to arbitrary generalized Fermat equations

$$F\colon A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^d = 0, \tag{4}$$

one must start by finding the correct analog of the isomorphism (3). This is our main contribution.

Let $\mathcal{S}$ be set of primes dividing the integer $a \cdot b \cdot c \cdot A \cdot B \cdot C \neq 0$, and denote by R the ring of $\mathcal{S}$-integers. Let $\mathcal{U}$ be the punctured cone associated to Equation (4). Let H be the subgroup of $\mathbb{G}_{\mathrm{m}}^3$ given on points by those $(\lambda_0, \lambda_1, \lambda_\infty)$ such that $\lambda_0^a = \lambda_1^b = \lambda_\infty^c$. The group H visibly acts on $\mathcal{U}$ by coordinate-wise multiplication

$$(x, y, z) \cdot (\lambda_0, \lambda_1, \lambda_\infty) := (\lambda_0 x, \lambda_1 y, \lambda_\infty z).$$

Finally, let $\mathbb{P}^1(a,b,c)$ denote the iterated root stack of $\mathbb{P}^1_{\mathbb{Z}}$ at the divisors $0, 1, \infty$ with multiplicities a, b, c. This is the stacky version of Darmon's M-curve $\mathbf{P}^1_{a,b,c}$ [Dar97, p. 4] (see Figure 1).

Theorem 1.1. *The map*

$$j\colon \mathcal{U}_R \to \mathbb{P}^1_R, \quad (x,y,z) \mapsto (-Ax^a : Cz^c) \tag{5}$$

induces an isomorphism $[\mathcal{U}_R/\mathsf{H}_R] \cong \mathbb{P}^1(a,b,c)_R$.

Warning 1.2. For a general triple (a,b,c) of positive integers, the group H is not isomorphic to $\mathbb{G}_{\mathrm{m}}$. We show in Lemma 4.3 that this is only the case when $\gcd(bc, ab, ac) = 1$. For example, in the important case of $(a,b,c) = (p,p,p)$ for some prime p, the group scheme H is isomorphic to $\mathbb{G}_{\mathrm{m}} \times \mu_p \times \mu_p$.

To further motivate the method of Fermat descent and introduce notation, we revisit the one of the very first instances of the method of *infinite descent* from this point of view.

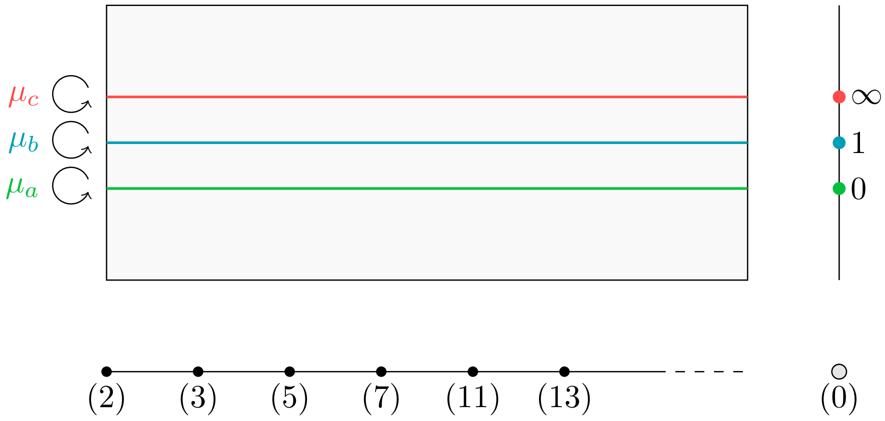

FIGURE 1. The Belyi stack of signature (a, b, c).

1.3. **Fermat's last theorem for $n = 4$.** The term *infinite descent* was coined in a letter from Fermat to Carcavi:

> And because the ordinary methods found in the books were insufficient to prove such difficult propositions, I finally discovered a completely novel path to reach them. I called this method of proof *infinite or indefinite descent*, etc.; at first, I used it only to prove negative propositions, such as, for example:
>
> ...
>
> That there is no right triangle with integer sides whose area is a square number.
>
> Fermat, 1659

Fermat actually proved the claim above (see [Dic66, Chapter XXII, p. 615]). Closely related to this problem is his famous "Last Theorem" in the case of exponent $n = 4$. In his controversial marginal notes to the *Arithmetica* of Diophantus, Fermat states that "the sum of two biquadrates is never a biquadrate or a square." (Excellent expositions of the proofs of these results via infinite descent are given in [Cona, Conb].)

At first glance, the method of infinite descent appears to be a simple reversed form of the principle of mathematical induction. In fact, it is much deeper. We present an overly complicated proof of Fermat's last theorem for exponent $n = 4$ using the method of Femat descent with the objective of illustrating the hidden geometry and introducing notation. The foreign definitions and constructions presented here will be introduced later on.

Fermat's last theorem for exponent $n = 4$ follows from the following stronger statement.

Theorem 1.3 (Fermat). *The primitive integral solutions* (x, y, z) *to the generalized Fermat equation* $F\colon \mathsf{x}^4 + \mathsf{y}^4 - \mathsf{z}^2 = 0$ *satisfy* $x \cdot y \cdot z = 0$. *They are the eight triples*

$$\pm(1, 0, 1), \pm(1, 0, -1), \pm(0, 1, 1), \pm(0, 1, -1).$$

Proof sketch. Let $\mathcal{U}$ be the punctured cone associated to F. Our goal is to show that $\mathcal{U}(\mathbb{Z})$ contains only the eight elements listed above. The group scheme H defined above the statement of Theorem 1.1 (and discussed in detail in Section 4.1) acts on $\mathcal{U}$ by coordinate-wise multiplication, and it is not hard to see there are non-trivial stabilizers. For instance, the $\bar{\mathbb{Q}}$-points can be described explicitly as

$$\mathsf{H}(\bar{\mathbb{Q}}) = \left\{(\xi_0\lambda, \xi_1\lambda, \xi_\infty\lambda^2) : \lambda \in \bar{\mathbb{Q}}^\times, \xi_0, \xi_1 \in \mu_2(\bar{\mathbb{Q}}), \xi_\infty \in \mu_4(\bar{\mathbb{Q}})\right\}.$$

From this description, we see that any point $(x, y, z) \in \mathcal{U}(\bar{\mathbb{Q}})$ with $x = 0$, $y = 0$, or $z = 0$ has $\mu_2(\bar{\mathbb{Q}}), \mu_2(\bar{\mathbb{Q}})$, or $\mu_4(\bar{\mathbb{Q}})$ stabilizers in $\mathsf{H}(\bar{\mathbb{Q}})$, respectively. On the other hand, any point $(x, y, z) \in \mathcal{U}(\bar{\mathbb{Q}})$ with $x \cdot y \cdot z \neq 0$ has no non-trivial stabilizers.

The idea is to understand the arithmetic of $\mathcal{U}$ by studying instead the arithmetic of the stack quotient $[\mathcal{U}/\mathsf{H}]$.

Remark 1.4 (Notation). If $\mathcal{X}$ is a stack and R is a ring, we denote by $\mathcal{X}(R)$ the *groupoid* of R-points, and by $\mathcal{X}\langle R\rangle$ the *set* of R-points, (see Section 2.1).

Descent theory (see Theorem 2.16) gives the partition

$$[\mathcal{U}/\mathsf{H}]\langle\mathbb{Z}\rangle = \bigsqcup_{\delta \in \mathrm{H}^1(\mathbb{Z}, \mathsf{H})} \mathcal{U}_\delta(\mathbb{Z})/\mathsf{H}(\mathbb{Z}) \supset \mathcal{U}(\mathbb{Z})/\mathsf{H}(\mathbb{Z}), \tag{6}$$

where the (fppf) cohomology group $\mathrm{H}^1(\mathbb{Z}, \mathsf{H})$ is finite, and the $\mathcal{U}_\delta$ are certain twists of $\mathcal{U}$ arising from generalized Fermat equations F_δ. (In fact, $\mathrm{H}^1(\mathbb{Z}, \mathsf{H})$ is trivial (Lemma 4.4).)

Remark 1.5. The group H is not the only (fppf) group scheme acting on $\mathcal{U}$. For instance, one can also consider the image $\mathbb{G}_{\mathrm{m}}(1, 1, 2)$ of the map $\mathbb{G}_{\mathrm{m}} \to \mathbb{G}_{\mathrm{m}}^3\colon \lambda \mapsto (\lambda, \lambda, \lambda^2)$. The quotient $[\mathcal{U}/\mathbb{G}_{\mathrm{m}}(1, 1, 2)]$ has the technical advantage of being a closed substack of the weighted projective stack $\mathcal{P}(1, 1, 2)$. The reason for choosing H over $\mathbb{G}_{\mathrm{m}}(1, 1, 2)$ will become clear shortly.

Step 0: In this case, $p = 2$ is the only bad prime, so $\mathcal{S} = \{2\}$ and we abbreviate $R = \mathbb{Z}[1/2]$. From Theorem 1.1, we have an isomorphism of stacks

$$[\mathcal{U}_R/\mathsf{H}_R] \cong \mathbb{P}^1(4,4,2)_R. \tag{7}$$

The notable consequence of this isomorphism is that, by the definition of the Belyi stack $\mathbb{P}^1(4,4,2)$, geometrically Galois Belyi maps $\phi\colon X \to \mathbb{P}^1_{\mathbb{Q}}$ with signature $(4,4,2)$ factor through étale covers $\phi\colon X \to \mathbb{P}^1(4,4,2)_{\mathbb{Q}}$, and thus we learn new information about $\mathcal{U}(\mathbb{Z})$ by studying the sets of rational points $X(\mathbb{Q})$ arising from such maps.

Step 1: (Covering) The first task of the method of Fermat descent is to find a geometrically Galois Belyi map ϕ of signature $(4,4,2)$ with good reduction outside of $\mathcal{S} = \{2\}$ (see Definition 3.19). The LMFDB (beta) [LMF25b] gave us the elliptic curve $E_{\mathbb{Q}}\colon v^2w = u^3 - uw^2 \subset \mathbb{P}^2_{\mathbb{Q}}$ defined over $\mathbb{Q}$ with j-invariant 1728 and the map

$$\phi\colon E_{\mathbb{Q}} \to \mathbb{P}^1_{\mathbb{Q}}, \quad (u : v : w) \mapsto (u^2 : u^2 - w^2). \tag{8}$$

The same equations define an R-model $\Phi\colon E \to \mathbb{P}^1_R$. Let $\mathbf{Aut}(\Phi)$ be the automorphism group scheme over $\operatorname{Spec} R$ of Φ. In this situation, we have

$$\mathbb{P}^1(4,4,2)_R \cong [E/\mathbf{Aut}(\Phi)]. \tag{9}$$

Step 2: (Twisting) It turns out that $\mathbf{Aut}(\Phi) = \mathbf{Aut}(E)$. In addition, $\mathbf{Aut}(E) \cong \mu_4 = \operatorname{Spec} R[t]/\langle t^4 - 1\rangle$ by [Sil09, Corollary III.10.2]. By Kummer theory, we know that the cohomology group $\mathrm{H}^1(R, \mu_4)$ is isomorphic to the finite group $R^\times/(R^\times)^4 \cong \{\pm 1, \pm 2, \pm 4, \pm 8\}$. Once again, descent theory gives a partition

$$[E/\mathbf{Aut}(\Phi)]\langle R\rangle = \bigsqcup_{\tau \in \mathrm{H}^1(R,\mu_4)} \phi_\tau(E_\tau(R)) = \bigsqcup_{d \in R^\times/(R^\times)^4} \phi_d(E_d(\mathbb{Q})).$$

The last equality follows because the twists E_d are all proper and thus $E_d(R) = E_d(\mathbb{Q})$ by the valuative criterion.

The final remaining task in this step is to calculate the quartic twists $\Phi_d\colon E_d \to \mathbb{P}^1_R$ (see Section 2.3). By recalling that μ_4 acts on the elliptic surface $E\colon v^2 = u^3 - u$ via

$$(u, v) \cdot \zeta := (\zeta^2 u, \zeta^3 v),$$

and that for each $d \in \mathrm{H}^1(R, \mu_4)$ the corresponding (left fppf) μ_4-torsor is $T_d := \operatorname{Spec} R[t]/\langle t^4 - d\rangle \to \operatorname{Spec} R$ with action

$$\zeta \cdot \sqrt[4]{d} := \zeta \sqrt[4]{d},$$

an invariant calculation gives that $E_d\colon v^2w = u^3 - duw^2 \subset \mathbb{P}^2_R$ and

$$\Phi_d\colon E_d \to \mathbb{P}^1_R, \quad (u:v:w) \mapsto (u^2 : u^2 - dw^2). \tag{10}$$

This coincides with the Galois cohomology perspective in [Sil09, Proposition X.5.4]. Indeed, it turns out that $\mathrm{H}^1(R, \mathbf{Aut}(\Phi))$ is isomorphic to the Galois cohomology group $\mathrm{H}^1_{\{2\}}(\mathbb{Q}, \mu_4(\bar{\mathbb{Q}}))$ parametrizing isomorphism classes of Galois étale $\mathbb{Q}$-algebras unramified outside $\{2\}$, and with Galois group $C_4 = \mu_4(\bar{\mathbb{Q}})$.

Step 3: (Sieving) To summarize, we have shown that the set $j(\mathcal{U}(\mathbb{Z}))$ is contained in

$$j(\mathcal{U}(R)) \subset \bigsqcup_{d \in R^\times/(R^\times)^4} \phi_d(E_d(\mathbb{Q})).$$

Our objective now is to sieve out the points in $\mathcal{U}(R)$ that are not in $\mathcal{U}(\mathbb{Z})$. The best case scenario here would be that the sets $E_d(\mathbb{Q})$ are finite and that $\phi_d(E_d(\mathbb{Q}))$ are contained in $\{0, 1, \infty\} \subset \mathbb{P}^1(\mathbb{Q})$. Indeed, this would imply that $j(\mathcal{U}(\mathbb{Z})) \subset \{0, 1, \infty\}$, which can easily be seen to imply Theorem 1.3. Unfortunately, this is not the case: the elliptic curves E_2 and E_{-8} (with LMFDB labels `256.b1` and `256.b2`) have infinitely many $\mathbb{Q}$-rational points. We are forced to take a closer look at the rational points on the projective line arising from $\mathcal{U}(\mathbb{Z})$ and $E_d(\mathbb{Q})$ simultaneously.

For any choice of $d \in \{\pm 1, \pm 2, \pm 4, \pm 8\}$, let us consider a point Q in the intersection $j(\mathcal{U}(\mathbb{Z})) \cap \phi_d(E_d(\mathbb{Q})) \subset \mathbb{P}^1(\mathbb{Q})$, so

$$Q = (x^4 : z^2) = (u^2 : u^2 - dw^2)$$

for some primitive integral solution (x, y, z) to $\mathsf{x}^4 + \mathsf{y}^4 = \mathsf{z}^2$, and some rational point $P = (u : v : w) \in E_d(\mathbb{Q})$. Note that if $Q = 0$, $P \in E_d(\mathbb{Q})$ is the point at infinity and $(x, y, z) = \pm(0, 1, 1), \pm(0, 1, -1)$. We assume that $P \neq (0 : 1 : 0)$. It follows that there is some $\lambda \in \mathbb{Z}_{>1}$ such that $x^4 = \lambda^2 u^2$, and $z^2 = \lambda^2(u^2 - d)$. Moreover, $y^4 = z^2 - x^4 = -\lambda^2 d$. But this forces $d \in \{-1, -4\}$. Fortunately, both $E_{-1}(\mathbb{Q})$ and $E_{-4}(\mathbb{Q})$ are finite (see `64.a4`, `32.a4`), and a calculation gives:

$$\begin{aligned} \phi_{-1}(E_{-1}(\mathbb{Q})) &= \{1, \infty\}, \\ \phi_{-4}(E_{-4}(\mathbb{Q})) &= \{1, \infty, (1:2)\}. \end{aligned}$$

The points $Q = 1, \infty$ are the expected ones, and the point $Q = (1 : 2)$ is ruled out by the same arguments above: $x^4 = \lambda^2$ and $y^4 = 4\lambda^2$ imply that $\lambda^2 = 1$ by primitivity, and 4 is not a fourth power. □

1.4. **Summary of contributions.** This article contributes a thorough exposition of the stack-theoretic approach to the Diophantine study of generalized Fermat equations, complemented by a number of new results. The stack-theoretic perspective offers the distinct advantage of being more conceptual. The exposition is intended to serve as a reference for future work applying the method of Fermat descent to tackle numerous open problems in this area.

- In Section 2, we provide a concise review of quotient stacks and use it to present the fundamentals of *descent theory*, in the sense of [Sko01, Chapter 2], from this vantage point. The main result of this section is Theorem 2.16, the *descent theory partition*. It is stated in greater generality in [San22b, Lemma 2.4]. While this is certainly well known, we find our proof to be succinct and illuminating.
- In Section 3, we begin by reviewing the root stack construction, which is due to Cadman [Cad07]. Our exposition borrows from Olsson's treatment in [Ols16, Section 10.3]. The purpose of this review is to understand the arithmetic of the Belyi stack $\mathbb{P}^1(a,b,c)$, which provides the stack-theoretic interpretation of Darmon's M-curve $\mathbf{P}^1_{a,b,c}$ [Dar97, p. 4]. It is well known to experts that Darmon's M-curves can be interpreted as root stacks over their coarse moduli spaces. This perspective is hinted at by Poonen in [Poo06], and addressed more directly by Santens in this MathOverflow post [San22a] and [San22b, Lemma 2.1]. In turn, this is an instance of a general feature of Deligne–Mumford stacks [GS17]. The main result of this section is Lemma 3.14, which explicitly characterizes the PID points on a Belyi stack.
- In Section 4, we prove our main result: Theorem 4.6. In spirit, this theorem captures a striking connection between:
  - **Arithmetic:** the Diophantine equations $A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c = 0$, and
  - **Geometry:** the Riemann sphere $\mathbb{CP}^1$ with three orbifold points $(0, 1, \infty)$ of multiplicities (a, b, c).

  This connection was already identified by [DG95], and formulated in terms of stacks by [Poo06], [PSS07], [VZB22, Example 5.4.7], and [Poo23]. The originality of our contribution lies in Section 4.1, where we work out the combinatorics of the general case, when $\gcd(bc, ac, ab)$ is possibly greater than one.

Acknowledgements. This work is part of the author's PhD thesis. We thank David Zureick-Brown, John Voight, and Andrew Kobin for many enlightening conversations on this topic and for their valuable feedback. We are also grateful to Bjorn Poonen for agreeing to serve

on the thesis committee and for his detailed and insightful comments on an earlier draft.

2. Quotient stacks and descent

2.1. **Conventions on stacks.** Recall that a morphism of schemes is fppf if it is faithfully flat and locally of finite presentation (see [Poo17, Definition 3.4.1]). For a choice of base scheme S, we work on the big fppf site $S_{\mathtt{fppf}} = (\mathtt{Sch}_{/S})_{\mathtt{fppf}}$. This is the category $\mathtt{Sch}_{/S}$ of schemes over S where the open coverings are families $\{U_i \to U\}$ of S-morphisms such that $\bigsqcup_i U_i \to U$ is fppf.

Definition 2.1. A category over S is a pair $(\mathcal{X}, \mathrm{p})$ where $\mathcal{X}$ is a category and $\mathrm{p}\colon \mathcal{X} \rightsquigarrow \mathtt{Sch}_{/S}$ is a functor. A morphism $f\colon y \to z$ in $\mathcal{X}$ is called cartesian if given any morphism $g\colon x \to z$ and a factorization $\mathrm{p}(f) \circ \phi\colon \mathrm{p}(x) \to \mathrm{p}(y) \to \mathrm{p}(z)$ of $\mathrm{p}(g)$, there exists a unique morphism $h\colon x \to y$ such that $\mathrm{p}(h) = \phi$ and $g = h \circ f$.

$$\begin{tikzcd}
\mathcal{X} \arrow[d, squiggly, |->, "\mathrm{p}"'] & x \arrow[r, dashed, "h"] \arrow[rr, bend left, "g"] \arrow[d, squiggly, |->] & y \arrow[r, "f"] \arrow[d, squiggly, |->] & z \arrow[d, squiggly, |->] \\
\mathtt{Sch}_{/S} & \mathrm{p}(x) \arrow[r, "\phi"] & \mathrm{p}(y) \arrow[r, "\mathrm{p}(f)"] & \mathrm{p}(z)
\end{tikzcd} \tag{11}$$

Definition 2.2. Let $(\mathcal{X}, \mathrm{p})$ be a category over S. If $f\colon y \to z$ is a cartesian morphism, the object $y \in \mathcal{X}$ is called a pullback of z along $\mathrm{p}(f)$. Given an S-scheme U, the category of U-points in $\mathcal{X}$, denoted $\mathcal{X}(U)$, is the category of pullbacks over the identity. That is,

Objects: objects u in $\mathcal{X}$ such that $\mathrm{p}(u) = U$.
Morphisms: morphisms $\phi\colon v \to u$ in $\mathcal{X}$ such that $\mathrm{p}(\phi) = \mathrm{id}_U$.

Definition 2.3. A fibered category over S is a category $(\mathcal{X}, \mathrm{p})$ over S such that for every S-morphism of schemes $\Phi\colon V \to U$ and u in $\mathcal{X}(U)$, there exists a cartesian morphism $\phi\colon v \to u$ such that $\mathrm{p}(\phi) = \Phi$. In particular, this implies that v is in $\mathcal{X}(V)$.

Fibered categories over S assemble into a 2-category (see [Ols16, Definition 3.1.3]). Indeed, there are natural notions of (i) morphisms between fibered categories over S, and (ii) morphisms between morphisms of fibered categories over S. Moreover, there is a version of the Yoneda lemma (see [Ols16, Chapter 3.2]) in this context that justifies calling $U \mapsto \mathcal{X}(U)$ a "functor" of points.

Definition 2.4. Recall that a groupoid is a category in which every morphism is an isomorphism. A category fibered in groupoids over S if a fibered category $\mathcal{X}$ over S, such that for every S-scheme U, the category $\mathcal{X}(U)$ is a groupoid. Given a category fibered in groupoids $\mathcal{X}$ over S and an S-scheme U, we denote by $\mathcal{X}\langle U\rangle$ the set of isomorphism classes of the groupoid $\mathcal{X}(U)$.

Since our focus will be on the arithmetic of stacks, thinking about stacks in terms of their groupoids/sets of U-points will be enough for most of our applications. When we use the word *stack*, we mean an algebraic stack in the following sense.

Definition 2.5. Let $\mathcal{X}$ be a category fibered in groupoids over S.

(i) $\mathcal{X}$ is a stack if for every fppf cover $\{U_i \to U\}$, the induced descent functor $\mathcal{X}(U) \to \mathcal{X}(\{U_i \to U\})$ is an equivalence of categories. See [Ols16, Section 4.2.4].

(ii) A stack $\mathcal{X}$ is algebraic if the diagonal $\Delta \colon \mathcal{X} \to \mathcal{X} \times_S \mathcal{X}$ is representable by an algebraic space, and $\mathcal{X}$ admits a smooth surjection $X' \to \mathcal{X}$ from an S-scheme X'. The map $X' \to \mathcal{X}$ is called a smooth presentation of $\mathcal{X}$. See [Ols16, Section 8.1].

(iii) An algebraic stack $\mathcal{X}$ is Deligne–Mumford if the smooth presentation above is in fact étale. See [Ols16, Section 8.3].

2.2. **Review of quotient stacks.** We focus on an concrete kind of stacks that arise from groups acting on schemes.

Situation 2.6. We place ourselves in the following situation for the rest of Section 2.2.

- Let S be a fixed base a scheme.
- Let Z be a scheme over S.
- Let G be an fppf S-group scheme.
- Suppose that we have $Z \times_S G \to Z$ a right action of G over S.
- We abbreviate $\mathrm{H}^1(S, G) = \check{\mathrm{H}}^1_{\mathtt{fppf}}(S, G)$ for the Čech cohomology set on the big fppf site of S, as in [Poo17, Section 6.4.4].

Definition 2.7 (Torsor scheme)**.** Let $G \to S$ be an fppf group scheme. A right fppf G-torsor over S is an S-scheme $T \to S$ together with a right action $T \times_S G \to T$ such that the following conditions hold:

(1) $T \to S$ is fppf.

(2) The map $T \times_S G \to T \times_S T$ defined by $(t, g) \mapsto (t, t \cdot g)$ is an isomorphism.

A morphism of G-torsors is a G-equivariant morphism of S-schemes.

An obvious yet important example is the trivial G-torsor. This is the fppf scheme $G \to S$ itself, with the right G-action given by the multiplication law. In fact, all G-torsors are locally trivial.

Lemma 2.8. *Let $T \to S$ be an S-scheme, equipped with a G-action $T \times_S G \to T$ satisfying Item 2 in Definition 2.7. The following conditions are equivalent.*

(a) $T \to S$ is fppf.
(b) $T \to S$ is fppf locally isomorphic to the trivial G-torsor.
(c) $T \to S$ admits a section fppf locally.

If we care about understanding group actions (i.e., quotients), we must leave the world of schemes. For many interesting examples, the sheafification of $U \mapsto Z(U)/G(U)$ is not representable by a scheme. Quotient stacks elegantly and succinctly solve this problem in terms of torsors.

Definition 2.9 (Quotient stack)**.** Define the quotient stack of Z by G, denoted $[Z/G]$, to be the algebraic stack over $S_{\texttt{fppf}}$ with:
Objects: triples (U, T, ϕ)

$$\begin{tikzcd}
T \arrow[r, "G\text{-equivariant}", "\phi"'] \arrow[d, "G_U\text{-torsor}"'] & Z \arrow[ddr, bend left] & \\
U \arrow[drr, bend right] & & \\
& & S
\end{tikzcd}$$

where

(i) U is an S-scheme,
(ii) $T \to U$ is a right fppf G_U-torsor, and
(iii) $\phi \colon T \to Z$ is a G-equivariant S-morphism.

Morphisms: $(U', T', \phi') \to (U, T, \phi)$ are pairs (f, h), where

(iv) $f \colon U' \to U$ is an S-morphism of schemes, and
(v) $h \colon T' \to T$ is a G-equivariant morphism over f inducing an isomorphism of $G_{U'}$-torsors $T' \cong T \times_{f,U} U'$, such that $\phi' = \phi \circ h$.

$$\begin{tikzcd}
T' \arrow[r, "h"] \arrow[d] \arrow[drr, "\phi'"] & T \arrow[d, dashed] \arrow[dr, "\phi"] & \\
U' \arrow[r, "f"] \arrow[drr] & U \arrow[dr] & X \arrow[d] \\
& & S
\end{tikzcd} \tag{12}$$

In particular, for any given S-scheme U, the groupoid $[Z/G](U)$ consists of pairs (T, ϕ) with $T \to U$ a G_U-torsor, and $\phi\colon T \to Z$ a G-equivariant S-morphism; and isomorphisms $h\colon (T_1, \phi_1) \to (T_2, \phi_2)$ are simply isomorphisms $h\colon T_1 \to T_2$ of G_U-torsors, compatible with the maps to Z.

$$\begin{tikzcd}
T_1 \arrow[rr, "h"] \arrow[dr] \arrow[drrr, "\phi_1"] & & T_2 \arrow[dl, dashed] \arrow[dr, "\phi_2"] & \\
& U \arrow[drr] & & Z \arrow[d] \\
& & & S
\end{tikzcd} \tag{13}$$

2.3. **Descent theory revisited.** In this section, we summarize the basics of descent theory from the point of view of quotient stacks. We follow Skorobogatov's book [Sko01, Section 2.2], but with inverted handedness. We place ourselves in the following situation until the end of Section 2.3.

Situation 2.10. Suppose we are in Situation 2.6.

- Assume that Z is quasi-projective.
- Let T denote an S-scheme with a **left** G action over S.
- Let $q\colon Z \to [Z/G]$ be the natural projection map.
- The structure map $G \to S$ is affine, and S is locally noetherian.

Remark 2.11. The last assumption ensures that there is a bijective correspondence between $\mathrm{H}^1(S, G)$ and isomorphism classes of G-torsor schemes, as a consequence of [Poo17, Theorem 6.5.10].

Remark 2.12. It is possible to work in greater generality (see [San22b, Section 2.5]) if we are not concerned with representability. In general, the contracted product of an S scheme Z with a left G-torsor will be an algebraic space.

Recall that when G is not commutative, $\mathrm{H}^1(S, G)$ is only a pointed set (the distinguished element corresponds to the class of the trivial

G-torsor) and not an abelian group. Nevertheless, we can still perform certain algebraic operations in this pointed set in terms of corresponding geometric operations on torsors.

Definition 2.13 (Contracted product)**.** The contracted product $Z\overset{G}{\times}T$ is defined as the quotient stack $[Z \times_S T/G]$, where G acts on the **right** on $Z \times_S T$ via

$$(z, t) \cdot g := (z \cdot g, g^{-1} \cdot t).$$

A crucial application of this definition is the pushforward operation on torsors. Given $\varphi\colon G \to H$ a homomorphism of fppf group schemes over S, we consider the left action of G on H. If $P \to S$ is a G-torsor, the contracted product

$$\varphi_* P := P\overset{G}{\times}H, \qquad \text{where } (p, h) \cdot g := (p \cdot g, \varphi(g)^{-1} \cdot h),$$

turns out to be an H-torsor, called the pushforward of P by φ. As an application of this construction, we have the following lemma (see [Ols16, Exercise 10.F]).

Lemma 2.14 (Induced maps on quotient stacks)**.** *Let S be a scheme, and $\varphi\colon G \to H$ a homomorphism of fppf group schemes over S. Let X be an S-scheme with a right G-action, and Y and S-scheme with a right H-action. Suppose that there is an S-morphism $f\colon X \to Y$ that is compatible with the group actions. Then, f induces a morphism of algebraic stacks $\bar{f}\colon [X/G] \to [Y/H]$.*

Proof sketch. Let $U \to S$ be an S-scheme. At the level of U-points, the functor $\bar{f}(U)\colon [X/G](U) \to [Y/H](U)$ is defined in the following way. Recall that a U-point on $[X/G]$ is a triple (U, T, ϕ) as in Definition 2.9. First, consider the pullback $\varphi_U\colon G_U \to H_U$. Pushing forward the G_U-torsor T via φ_U gives a triple $(U, (\varphi_U)_* T, (\varphi_U)_*(f \circ \phi))$.

$$\begin{tikzcd}
T \arrow[r, "\phi"] \arrow[dd] & X \arrow[dd] & & & (\varphi_U)_*T \arrow[r] \arrow[dd] & Y \arrow[dd] \\
& & {} \arrow[r, mapsto] & {} & & \\
U \arrow[r] & {[X/G]} & & & U \arrow[r] & {[Y/H]}
\end{tikzcd}$$

□

The following lemma is a restatement of [Poo17, Section 6.5.6].

Lemma 2.15 (Twisting by fppf descent)**.** *Given $\tau \in \mathrm{H}^1(S, G)$, let $T \to S$ be a **left** fppf G-torsor corresponding to τ. Then:*

(i) The contracted product $Z\overset{G}{\times}T$ is represented by a quasi-projective S-scheme Z_τ. We call this the twist of Z by τ.

(ii) If $T = G$ is the trivial left G-torsor, then $Z_\tau \cong Z$ as S-schemes with a right G-action.

(iii) Taking $Z = G$ acting on itself by conjugation, the twist $Z_\tau = G_\tau$ is an affine fppf group scheme over S. It is called the inner twist of G by τ.

(iv) The twist Z_τ is a right fppf G_τ-torsor over S. Moreover, there is an isomorphism $[Z/G] \cong [Z_\tau/G_\tau]$. In particular, there is an induced map $q_\tau \colon Z_\tau \to [Z/G]$, called the twist of q by τ.

(v) The S-scheme T is a (G, G_τ)-bitorsor. The same S-scheme with the inverse G action $t \cdot g := g^{-1} \cdot t$ is a (G_τ, G)-bitorsor. We denote inverse (right) G-torsor by T^{-1}.

(vi) Finally, the contracted product $T^{-1} \overset{G}{\times} T$ is isomorphic to the trivial G-torsor.

Proof. (i) The representability of $Z_\tau = Z \overset{G}{\times} T$ is an application of fppf descent. See [Sko01, Lemma 2.2.3] for a proof when Z is affine.

(ii) We have the morphism $Z \times_S G \to Z \times_S Z$ given by $(z, g) \mapsto (z, z \cdot g)$. Observe that it is G-equivariant for the twisted action on $Z \times_S G$, and the action $(z_1, z_2) \cdot g := (z_1 \cdot g, z_2)$ on $Z \times_S Z$. This gives a morphism of quotient stacks $Z_\tau = [Z \times_S G/G] \to [Z \times_S Z/G]$. Since the first projection $Z \times_S Z \to Z$ is G-equivariant for the trivial G-action on Z, we get a map $\psi \colon Z_\tau \to Z$. On the other hand, we have a morphism $\phi \colon Z \to Z_\tau$ induced by $Z \to Z \times_S G$. To see that these are mutual inverses, it is enough to realize that the following diagram is commutative

$$\begin{tikzcd}
Z \arrow[d] \arrow[dr, bend left, "\phi"] & \\
Z \times_S G \arrow[r] \arrow[d] & Z_\tau \arrow[d, "\psi"] \\
Z \times_S Z \arrow[r, "\mathrm{pr}_1"'] & Z.
\end{tikzcd}$$

(iii) Since $G \to S$ is affine, the same will be true for $G_\tau \to S$ by fppf descent. Checking that G_τ is an S-group is a matter of pulling back the group operations to T and verifying that they are G-equivariant under the twisted action. For example, consider the inverse morphism $\iota \colon G \to G$ pulled back to $\iota \times_S T \colon G \times_S T \to G \times_S T$. Then, we have that $(g, t) \cdot h = (h^{-1} g h, h^{-1} t)$ maps to $(h^{-1} g^{-1} h, h^{-1} t) = (g^{-1}, t) \cdot h$. We obtain the twisted inverse morphism $G_\tau \to G_\tau$ by passing to the quotient.

(iv) Consider the morphism $\phi \colon (Z \times_S T) \times_S (G \times_S T) \to Z \times_S T$ given on points by $(z, x, g, t) \mapsto (z \cdot g, t)$. Note that $(z, x, g, t) \cdot h =$

$(z \cdot h, h^{-1}x, h^{-1}gh, h^{-1}t)$ maps to $(z \cdot gh, h^{-1}t) = (z \cdot g, t) \cdot h$, so ϕ induces a morphism $Z_\tau \times_S G_\tau \to Z_\tau$. One similarly verifies the G-equivariance of the diagrams that descend to the group action axioms on $Z_\tau \times_S G_\tau \to Z_\tau$. For the third statement, note that we have a morphism $Z \times_S T \to Z$ compatible with $\varphi\colon G_\tau \to G$, namely the first projection $(z, t) \mapsto z$. From Lemma 2.14, we get a map $[Z_\tau/G_\tau] \to [Z/G]$ is the induced map of quotient stacks. This map is in fact as an isomorphism. We obtain q_τ as the composition $Z_\tau \to [Z_\tau/G_\tau] \cong [Z/G]$.

(v) Follows directly from (iv).

(vi) This is a particular instance of the general fact that $Z \times_{[Z/G]} Z \cong Z \times_S G$. Indeed, taking $Z = T^{-1} \times_S T$ shows that Z_τ has a section fppf locally, implying that it is the trivial G-torsor by Lemma 2.8. □

In general, the sets $[Z/G]\langle S\rangle$ and $Z(S)/G(S)$ are not the same. Nevertheless, the former is contained in the latter, and the difference is accounted for by the quotients $Z_\tau(S)/G(S)$ as τ ranges over $\mathrm{H}^1(S, G)$.

Theorem 2.16 (Descent theory partition)**.** *Then, the* ***set*** *of S-points on the quotient stack $[Z/G]$ is partitioned by the images of the S-points of the twists of $q\colon Z \to [Z/G]$.*

$$[Z/G]\langle S\rangle = \bigsqcup_{\tau \in \mathrm{H}^1(S,G)} q_\tau(Z_\tau(S)).$$

Proof. Recall that a map $S \to [Z/G]$ is the data of a triple (S, T^{-1}, ϕ) where T^{-1} is a right fppf G-torsor over S, and $\phi\colon T^{-1} \to Z$ is a G-equivariant map of S-schemes. We want to show that every map $(T^{-1}, \phi)\colon S \to [Z/G]$ factors through a twist $q_\tau\colon Z_\tau \to [Z/G]$ of the canonical quotient $q\colon Z \to [Z/G]$, where τ is completely determined by the isomorphism class of the point (T, ϕ). Indeed, in this setting, we have the *evaluation map* $\zeta\colon (T^{-1}, \phi) \mapsto \tau := [T \to S]$ from $[Z/G]\langle S\rangle$ to $\mathrm{H}^1(S, G)$, where τ is the cohomology class corresponding to the left G-torsor $T \to S$. Since $T^{-1} \overset{G}{\times} T$ is isomorphic to the trivial G-torsor, we have a section $e\colon S \to G \cong T^{-1} \overset{G}{\times} T$ that realizes the factorization of our map (T^{-1}, ϕ) by the commutativity of the diagram in Figure 2. The map $T^{-1} \overset{G}{\times} T \to Z_\tau$ is the one induced by the G-equivariant S-morphism $\phi \times_S \mathrm{id}_T\colon T^{-1} \times_S T \to Z \times_S T$. □

As a reality check, let us calculate the set of R-points on the projective line over a principal ideal domain R using Theorem 2.16.

Example 2.17 (PID points on the projective line)**.** Recall the greatest common divisor of two elements s, t in R is a generator of the ideal $sR + tR$. Let $\mathcal{V} := \mathbb{A}^2 - \mathbf{0}$, so that $\mathcal{V}(R) \cong \{(s, t) \in R^2 : sR + tR = R\}$.

$$\begin{tikzcd}
T^{-1}\overset{G}{\times}T \arrow[rr, "\phi\overset{G}{\times}\mathrm{id}_{\mathrm{T}}"] & & Z_\tau \arrow[dd, "q_\tau"] \\
T^{-1} \arrow[r, "\phi"] \arrow[d] & Z \arrow[d, "q"] & \\
S \arrow[r] \arrow[uu, dashed, bend left, "e"] & {[Z/G]} \arrow[r, equal] & {[Z_\tau/G_\tau]}
\end{tikzcd}$$

Figure 2. Proof of the method of descent.

We have that $\mathbb{P}^1(R) \cong \{(s,t) \in R^2 : sR + tR = R\} / R^\times$. One can see this using the fact that $\mathbb{P}^1$ is the quotient stack $[\mathcal{V}/\mathbb{G}_{\mathrm{m}}]$. Indeed, since $\operatorname{Pic} R$ is trivial,

$$\mathbb{P}^1(R) = [\mathcal{V}/\mathbb{G}_{\mathrm{m}}]\langle R\rangle = \bigsqcup_{\tau\in \mathrm{H}^1(R,\mathbb{G}_{\mathrm{m}})} \mathcal{V}_\tau(R)/\mathbb{G}_{\mathrm{m}}(R) = \mathcal{V}(R)/R^\times.$$

Indeed, a point $Q \in \mathbb{P}^1(R)$ is (isomorphic to a) cartesian square

$$\begin{tikzcd}
\mathbb{G}_{\mathrm{m}} \arrow[r, "\phi"] \arrow[d] & \mathcal{V} \arrow[d] \\
\operatorname{Spec} R \arrow[r] & \mathbb{P}^1,
\end{tikzcd}$$

where ϕ is a $\mathbb{G}_{\mathrm{m}}$-equivariant map. Composing the identity section $e\colon \operatorname{Spec} R \to \mathbb{G}_{\mathrm{m}}$ with ϕ we obtain a point in $\mathcal{V}(R)$, i.e., a pair $(s,t) \in R^2$ such that $sR + tR = R$. Any other isomorphic square comes from a $\mathbb{G}_{\mathrm{m}}$-equivariant map $\phi'\colon \mathbb{G}_{\mathrm{m}} \to \mathcal{V}$ giving rise to a point (s',t') such that $(s',t') = (us, ut)$ for some $u \in R^\times$.

3. Root stacks and the Belyi stack

3.1. **Review of the root stack construction.** An effective Cartier divisor on a scheme X is a closed subscheme $D \subset X$ such that the corresponding ideal sheaf $\mathcal{O}_X(-D)$ is a line bundle [Sta24, Tag 01WR]. Equivalently, a closed subscheme is an effective Cartier divisor if and only if it is locally cut out by a single element which is a nonzero divisor [Sta24, Tag 01WS]. Denote by $j_D\colon \mathcal{O}_X(-D) \hookrightarrow \mathcal{O}_X$ the natural inclusion morphism of $\mathcal{O}_X$-modules.

Definition 3.1 ([Ols16, Definition 10.3.2])**.** A generalized effective Cartier divisor on a scheme X is a pair $(\mathcal{L}, \rho)$, where $\mathcal{L}$ is a line bundle on X, and $\rho\colon \mathcal{L} \to \mathcal{O}_X$ is a morphism of $\mathcal{O}_X$-modules. An isomorphism of generalized Cartier divisors $(\mathcal{L}', \rho') \cong (\mathcal{L}, \rho)$ is an isomorphism of line bundles $\sigma\colon \mathcal{L}' \to \mathcal{L}$ such that the following triangle commutes

$$\begin{tikzcd}
\mathcal{L}' \arrow[rr, "\sigma"] \arrow[dr, "\rho'"'] & & \mathcal{L} \arrow[dl, "\rho"] \\
& \mathcal{O}_X & .
\end{tikzcd}$$

We can multiply generalized effective Cartier divisors $(\mathcal{L}, \rho)$ and $(\mathcal{L}', \rho')$ by declaring $(\mathcal{L}, \rho) \cdot (\mathcal{L}', \rho') := (\mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{L}', \rho \otimes \rho')$, where $\rho \otimes \rho'$ is the morphism of $\mathcal{O}_X$-modules given by the composition

$$\mathcal{L} \otimes_{\mathcal{O}_X} \mathcal{L}' \to \mathcal{O}_X \otimes_{\mathcal{O}_X} \mathcal{O}_X \cong \mathcal{O}_X.$$

Example 3.2 (Effective Cartier divisors)**.** Given an effective Cartier divisor $D \subset X$, the pair $(\mathcal{O}_X(-D), j_D)$ is a generalized effective Cartier divisor. By definition, two effective Cartier divisors $D', D \subset X$ are isomorphic as generalized effective Cartier divisors if and only if they are equal and the isomorphism is therefore unique.

Example 3.3 (Generalized effective Cartier divisors on affine schemes)**.** In light of the equivalence between R-modules and quasicoherent $\mathcal{O}_X$-modules on $X = \operatorname{Spec} R$, a generalized effective Cartier divisor on an affine scheme is of the from $(\widetilde{M}, \widetilde{\lambda})$ for a projective R-module M of rank one, and a morphism $\lambda\colon M \to R$ of R-modules. In particular, $\lambda(M)$ is an ideal in R. Two generalized effective Cartier divisors (M', λ') and (M, λ) on $\operatorname{Spec} R$ are isomorphic if and only if there exists an R-module isomorphism $\sigma\colon M' \to M$ such that $\lambda' = \lambda \circ \sigma$. In particular, note that such a pair gives rise to the same ideal $\lambda'(M') = \lambda(\sigma(M')) = \lambda(M)$.

Definition 3.4 (Root stack)**.** Fix an effective Cartier divisor D on a scheme X, and a positive integer r. Let $\sqrt[r]{X; D}$ be the fibered category over $X_{\texttt{fppf}}$ with:

Objects: triples $(f\colon T \to X, (\mathcal{M}, \lambda), \sigma)$ where $f\colon T \to X$ is an X-scheme, $(\mathcal{M}, \lambda)$ is a generalized effective Cartier divisor on T, and $\sigma\colon (\mathcal{M}^{\otimes r}, \lambda^{\otimes r}) \to (f^*\mathcal{O}_X(-D), f^*j_D)$ is an isomorphism of generalized effective Cartier divisors on T.

Morphisms: a morphism

$$(f'\colon T' \to X, (\mathcal{M}', \lambda'), \sigma') \to (f\colon T \to X, (\mathcal{M}, \lambda), \sigma)$$

is the data of a pair $(h, h^\flat)$ where $h\colon T' \to T$ is an X-morphism, and $h^\flat\colon (\mathcal{M}', \lambda') \to (h^*\mathcal{M}, h^*\lambda)$ is an isomorphism of generalized effective

Cartier divisors on T' such that the following diagram commutes

$$\begin{tikzcd}
\mathcal{M}'^{\otimes r} \arrow[r, "h^{\flat \otimes r}"] \arrow[d, "\sigma'"'] & h^*\mathcal{M}^{\otimes r} \arrow[d, "h^*\sigma"] \\
(f')^*\mathcal{O}_X(-D) \arrow[r, "\sim"] & h^*f^*\mathcal{O}_X(-D).
\end{tikzcd}$$

Remark 3.5 (Points on a root stack)**.** Usually, the base scheme X is itself defined over a different base scheme S. If $\mathcal{X} = \sqrt[r]{X; D}$, it is common to abuse notation and write $\mathcal{X}(S)$. What we mean is that we are considering $\mathcal{X}$ as a stack over S via the forgetful map $X_{\mathtt{fppf}} \to S_{\mathtt{fppf}}$. In particular, it follows that the groupoid $\mathcal{X}(S)$ is the disjoint union over $x \in \mathrm{Hom}_S(S, X) = X(S)$ of the groupoids $\mathcal{X}(x)$.

Remark 3.6 (Rooting a scheme at an effective Cartier divisor)**.** We are interested in the special case in which we root a scheme at a good old effective Cartier divisor D. We abbreviate $\sqrt[r]{X; (\mathcal{O}_X(-D), j_D)}$ by $\sqrt[r]{X; D}$. In particular, given an X-scheme $f\colon T \to X$, the groupoid $\sqrt[r]{X; D}(f)$ consists of:

Objects: triples $(f\colon T \to X, (\mathcal{M}, \lambda), \sigma)$ where $(\mathcal{M}, \lambda)$ is a generalized effective Cartier divisor on T, and $\sigma\colon (\mathcal{M}^{\otimes r}, \lambda^{\otimes r}) \to (f^*\mathcal{O}_X(-D), f^*j_D)$ is an isomorphism of generalized effective Cartier divisors on T.

Isomorphisms: $(f\colon T \to X, (\mathcal{M}', \lambda'), \sigma') \to (f\colon T \to X, (\mathcal{M}, \lambda), \sigma)$ consist of pairs $(h, h^\flat)$ where $h \in \mathrm{Aut}(T)$ satisfies $f = f \circ h$, and $h^\flat\colon (\mathcal{M}', \lambda') \to (h^*\mathcal{M}, h^*\lambda)$ is an isomorphism of generalized effective Cartier divisors on T such that the following diagram commutes

$$\begin{tikzcd}
\mathcal{M}'^{\otimes r} \arrow[r, "h^{\flat \otimes r}"] \arrow[d, "\sigma'"'] & h^*\mathcal{M}^{\otimes r} \arrow[d, "h^*\sigma"] \\
(f)^*\mathcal{O}_X(-D) \arrow[r, "\sim"] & h^*f^*\mathcal{O}_X(-D).
\end{tikzcd}$$

Finally, we arrive at the main definition of this section.

Definition 3.7 (Iterated root stack)**.** Let X be a scheme. Take a finite list $P_1, \ldots, D_r$ of effective Cartier divisors on X, and let $n_1, \ldots, n_r$ be positive integers. The iterated root stack of X at the divisors $P_1, \ldots, D_r$ with multiplicities $n_1, \ldots, n_r$ is the fiber product

$$\sqrt[n_1]{X; P_1} \times_X \cdots \times_X \sqrt[n_r]{X; D_r} \to X. \tag{14}$$

3.2. **The projective line rooted at a point.** Our first concrete non trivial example of a root stack is the projective line rooted at a single point $\mathcal{X} := \sqrt[n]{\mathbb{P}^1; P}$.

Definition 3.8. Let R be a principal ideal domain, and choose $P = (c : d)$ and $Q = (a : b)$ in $\mathbb{P}^1(R)$. Define the intersection ideal of P with Q as $I(P, Q) := (ad - bc)R \subset R$.

The ideal $I(P, Q)$ cuts out the locus in $\operatorname{Spec} R$ over which P and Q intersect. Indeed, the pullback of the diagonal $\mathbb{P}^1 \to \mathbb{P}^1 \times \mathbb{P}^1$ by $(P, Q)\colon \operatorname{Spec} R \to \mathbb{P}^1 \times \mathbb{P}^1$ gives the closed subscheme $\operatorname{Spec} R/I(P, Q)$. From the magic square, $I(P, Q)$ can equivalently be defined by the cartesian square

$$\begin{tikzcd}
\operatorname{Spec} R/I(P,Q) \arrow[r] \arrow[d] & \operatorname{Spec} R \arrow[d, hook, "Q"] \\
\operatorname{Spec} R \arrow[r, hook, "P"'] & \mathbb{P}^1_R\,.
\end{tikzcd} \tag{15}$$

Warning 3.9. The pullback $P^*\mathcal{O}_{\mathbb{P}^1}(-Q)$ does not coincide with the sheaf corresponding to $I(P, Q)$. More generally, the pulback of a quasicoherent ideal sheaf need not coincide with the ideal sheaf of the pulled back closed subscheme. Nevertheless, we have the following commutative diagram of sheaves on $\operatorname{Spec} R$ with exact rows

$$\begin{tikzcd}
 & P^*\mathcal{O}_{\mathbb{P}^1}(-Q) \arrow[r] \arrow[d] \arrow[dr, "\widetilde{\lambda}"] & P^*\mathcal{O}_{\mathbb{P}^1} \arrow[r] \arrow[d, equal] & P^*Q_*\widetilde{R} \arrow[r] \arrow[d, equal] & 0 \\
0 \arrow[r] & \widetilde{I(P,Q)} \arrow[r] & \widetilde{R} \arrow[r] & \widetilde{R/I(P,Q)} \arrow[r] & 0\,.
\end{tikzcd} \tag{16}$$

Proposition 3.10. *Let R be a principal ideal domain with fraction field K. Let $\mathbb{P}^1 = \operatorname{Proj} R[\mathsf{s}, \mathsf{t}]$. Fix a point $P \in \mathbb{P}^1(R)$, and a positive integer n. Let $\mathcal{X} := \sqrt[n]{\mathbb{P}^1; P}$ be the n^{th} root stack of $\mathbb{P}^1$ at P, defined over $\operatorname{Spec} R$. Then,*

$$\mathcal{X}(R) = \bigsqcup_{Q \in \mathbb{P}^1(R)} \mathcal{X}(Q),$$

where

- *(i) The fiber $\mathcal{X}(P)$ contains one object up to isomorphism, with automorphism group isomorphic to $\mu_n(R) = \{u \in R^\times : u^n = 1\}$.*
- *(ii) For $Q \neq P$ the ideal $I(P, Q)$ is nonzero, and the fiber $\mathcal{X}(Q)$ contains one object with trivial automorphism group if and only if $I(P, Q) = J^n$ for some ideal $0 \neq J \subsetneq R$, and is empty otherwise.*

In particular, when $R = K$, we have that $\mathcal{X}\langle K \rangle \cong \mathbb{P}^1(K)$.

Proof. Let $\mathcal{X} := \sqrt[n]{\mathbb{P}^1; P}$. As explained in Remark 3.5, the groupoid $\mathcal{X}(R)$ is the disjoint union of the groupoids $\mathcal{X}(Q)$, ranging over $Q \in \mathbb{P}^1(R)$. We proceed to describe each groupoid $\mathcal{X}(Q)$.

To start, consider the pullback of the ideal sheaf $\mathcal{O}_{\mathbb{P}^1}(-Q) = \widetilde{I_Q}$ via the map $P\colon \operatorname{Spec} R \to \mathbb{P}^1$, where $I_Q = (a\mathsf{t} - b\mathsf{s})R[\mathsf{s},\mathsf{t}] \subset R[\mathsf{s},\mathsf{t}]$. This is a line bundle on $\operatorname{Spec} R$ corresponding to a certain free R-module of rank one $M(P,Q)$. Moreover, the pullback of the generalized effective Cartier divisor $j_Q\colon \mathcal{O}_{\mathbb{P}^1}(-Q) \hookrightarrow \mathcal{O}^1_{\mathbb{P}}$ corresponds to an R-module homomorphism $\lambda(P,Q)\colon M(P,Q) \to R$ with image $I(P,Q)$, as illustrated in Diagram 16.

The **objects** in $\mathcal{X}(Q)$ are triples $(Q,(M,\lambda),\sigma)$, where

- (M,λ) is a generalized effective Cartier divisor on $\operatorname{Spec} R$ (see Example 3.3). Since R is a principal ideal domain, M is a free R-module of rank one and $\lambda\colon M \to R$ is an R-module homomorphism.
- $\sigma\colon (M^{\otimes n}, \lambda^{\otimes n}) \to (M(P,Q), \lambda(P,Q))$ is an isomorphism of generalized effective Cartier divisors on $\operatorname{Spec} R$, that is, a commutative triangle of R-modules

(17)
$$\begin{tikzcd}
M^{\otimes n} \arrow[rr, "\cong", "\sigma"'] \arrow[dr, "\lambda^{\otimes n}"'] & & M(P,Q) \arrow[dl, "\lambda(P,Q)"] \\
& R. &
\end{tikzcd}$$

By definition, an **isomorphism** $(Q,(M',\lambda'),\sigma') \to (Q,(M,\lambda),\sigma)$ in $\mathcal{X}(Q)$ is a pair $(h, h^\flat)$, where

- $h\colon \operatorname{Spec} R \to \operatorname{Spec} R$ is a morphism over $\operatorname{Spec} R$, so it must be the identity.
- $h^\flat\colon M' \to M$ is an isomorphism of R-modules such that $\lambda' = \lambda \circ h^\flat$ and the following diagram commutes

(18)
$$\begin{tikzcd}
M'^{\otimes n} \arrow[rr, "h^{\flat\otimes n}"] \arrow[dr] \arrow[drrr, "\sigma'"'] & & M^{\otimes n} \arrow[dl] \arrow[dr, "\sigma"] & \\
& R^{\otimes r} \arrow[drr, "\cong"] & & M(P,Q) \arrow[d, "\lambda(P,Q)"] \\
& & & R.
\end{tikzcd}$$

(i) When $P = Q$, then $I(P,Q) = 0$ and this forces every map $\lambda\colon M \to R$ to be the zero map. In particular, the bottom part of diagram (18) imposes no restriction and the isomorphisms of $\mathcal{X}(P)$ are precisely the isomorphisms of R-modules $h^\flat\colon M' \to M$

such that

$$\begin{tikzcd}
(M')^{\otimes n} \arrow[rr, "\cong", "h^{\flat\otimes n}"'] \arrow[dr, "\sigma'"'] & & M^{\otimes n} \arrow[dl, "\sigma"] \\
& M(P,P). &
\end{tikzcd}$$

In particular, any triple $(P,(M,\lambda),\sigma)$ in $\mathcal{X}(P)$ has $\mu_n(R)$ automorphisms.

(ii) When $P \neq Q$, the commutativity of (17) requires that the nonzero ideal $I(P,Q)$ is the n^{th} power of the ideal $\lambda(M)$ in R. This condition is also sufficient. Indeed, if $I(P,Q) = J^n$ for some nonzero ideal $\lambda\colon J \subset R$, then take an isomormphism of R-modules $\sigma\colon I(P,Q) \to M(P,Q)$ and note that

$$(Q,(J,\lambda),\sigma\colon J^n \to M(P,Q)) \tag{19}$$

is an object of $\mathcal{X}(Q)$, and every object in $\mathcal{X}(Q)$ is isomorphic to it. To calculate the automorphism group of this object, note that the only possible isomorphism $h^\flat\colon J \to J$ of R-modules such that $\lambda = h^\flat \circ \lambda\colon J \hookrightarrow R$, is the identity. Thus, the automorphism groups in $\mathcal{X}(Q)$ are trivial.

□

3.3. **The Belyi stack.** In this section, we summarize a few geometric and arithmetic properties of the Belyi stack $\mathbb{P}^1(a,b,c)$. This is the stack corresponding to Darmon's M-curve $\mathbf{P}^1_{a,b,c}$ in [Dar97, p. 4].

Situation 3.11. Let

- $(a,b,c) \in \mathbb{Z}^3$ be a triple of positive integers,
- $\mathbb{P}^1 = \operatorname{Proj} \mathbb{Z}[\mathsf{s},\mathsf{t}]$, and
- $P_0 = V(\mathsf{s}), P_1 = V(\mathsf{s}-\mathsf{t}), P_\infty = V(\mathsf{t}) \in \operatorname{Div}(\mathbb{P}^1_{\mathbb{Z}})$.

Definition 3.12 (Belyi stack)**.** We define the Belyi stack $\mathbb{P}^1(a,b,c)$ as the iterated root stack of $\mathbb{P}^1_{\mathbb{Z}}$ at the divisors P_0, P_1, P_∞ with multiplicities a, b, c.

$$\mathbb{P}^1(a,b,c) := \left(\sqrt[a]{\mathbb{P}^1; P_0}\right) \times_{\mathbb{P}^1} \left(\sqrt[b]{\mathbb{P}^1; P_1}\right) \times_{\mathbb{P}^1} \left(\sqrt[c]{\mathbb{P}^1; P_\infty}\right).$$

We start by summarizing some straightforward geometric properties of the Belyi stack. See [VZB22, Definition 11.2.1] and [VZB22, Definition 5.2.1] for the definition of a (relative) stacky curve.

Lemma 3.13. *The following statements hold.*

(i) The Belyi stack $\mathbb{P}^1(a,b,c)$ is a relative stacky curve over $\mathbb{Z}$ with coarse space $\mathbb{P}^1$. The coarse space morphism $\mathbb{P}^1(a,b,c) \to \mathbb{P}^1$ is an isomorphism over the open set $U = \mathbb{P}^1 - P_0 \cup P_1 \cup P_\infty$.

(ii) *Let $R = \mathbb{Z}[1/abc]$. Then the base change $\mathbb{P}^1(a,b,c)_R$ is tame.*

(iii) *For every geometric point $s\colon \operatorname{Spec} k \to \operatorname{Spec} R$, the fiber $\mathbb{P}^1(a,b,c)_s$ is a stacky curve over k. Moreover, the Euler characteristic of $\mathbb{P}^1(a,b,c)_s$ is*

$$\chi(\mathbb{P}^1(a,b,c)_s) = \tfrac{1}{a} + \tfrac{1}{b} + \tfrac{1}{c} - 1.$$

We define this common value to be the Euler characteristic of $\mathbb{P}^1(a,b,c)$.

We now turn to the arithmetic of the Belyi stack. We want to understand the set of $\mathbb{Z}$-points on $\mathbb{P}^1(a,b,c)$. The first step is to understand the set of $\mathbb{Z}$-points of the projective line rooted at a single point.

Lemma 3.14 (R-points on the Belyi stack)**.** *Let R be a principal ideal domain. Let $\mathbb{P}^1(a,b,c)$ be the base extension of the Belyi stack to R. The set $\mathbb{P}^1(a,b,c)\langle R\rangle$ is in bijection with the subset of $Q = (s : t) \in \mathbb{P}^1(R) = \mathbb{P}^1(k)$ such that $Q \in \{P_0, P_1, P_\infty\}$, or:*

- *$I(P_0, Q) = sR$ is a a^{th} power.*
- *$I(P_1, Q) = (s-t)R$ is a b^{th} power.*
- *$I(P_\infty, Q) = tR$ is a c^{th} power.*

Proof. Let $\mathcal{X}$ denote the Belyi stack. As with any fiber product of groupoids (see [Ols16, Section 3.4.9]), $\mathcal{X}\langle R\rangle$ is the fiber product of sets

$$\left(\sqrt[a]{\mathbb{P}^1; P_0}\right)\langle R\rangle \times_{\mathbb{P}^1(R)} \left(\sqrt[b]{\mathbb{P}^1; P_1}\right)\langle R\rangle \times_{\mathbb{P}^1(R)} \left(\sqrt[c]{\mathbb{P}^1; P_\infty}\right)\langle R\rangle,$$

so the result follows from the description of the R-points of the n^{th} root stack of the projective line at a given point P given in Proposition 3.10. □

As Darmon observed in [Dar97, p. 5], the integral points on the Belyi stack $\mathbb{P}^1(a,b,c)$ correspond to primitive integral solutions to generalized Fermat equations of signature (a,b,c), up to some *sloppiness* in the signs (i.e., $A, B, C \in \{\pm 1\} = \mathbb{Z}^\times$). If we consider $\mathbb{Z}[\mathcal{S}^{-1}]$-points instead, the same is true but allowing the coefficients $A, B, C \in \mathbb{Z}[\mathcal{S}^{-1}]^\times$.

Lemma 3.15. *Let $\mathcal{S}$ be a finite (possibly empty) set of primes, and let $R = \mathbb{Z}[\mathcal{S}^{-1}]$. Then, every point in $Q = \mathbb{P}^1(a,b,c)\langle R\rangle$ arises from a primitive integral solution to a generalized Fermat equation $F\colon A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c = 0$, where $A \cdot B \cdot C \in R^\times$, as $j(x,y,z) = Q$.*

Proof. Explicitly, an object in the groupoid $\mathcal{X}(Q)$ is a triple

$$(Q, [(M_0, \lambda_0), (M_1, \lambda_1), (M_\infty, \lambda_\infty)], (\sigma_0, \sigma_1, \sigma_\infty))$$

where $(M_0, \lambda_0), (M_1, \lambda_1), (M_\infty, \lambda_\infty)$ are generalized effective Cartier divisors on $\operatorname{Spec} R$, and $\sigma_0, \sigma_1, \sigma_\infty$ are isomorphisms of generalized effective Cartier divisors on $\operatorname{Spec} R$

$$\begin{aligned}
\sigma_0 &\colon (M_0^{\otimes a}, \lambda_0^{\otimes a}) \to Q^*(\mathcal{O}_{\mathbb{P}^1}(-P_0), j_0) = (M(P_0, Q), \lambda(P_0, Q)),\\
\sigma_1 &\colon (M_1^{\otimes b}, \lambda_1^{\otimes b}) \to Q^*(\mathcal{O}_{\mathbb{P}^1}(-P_1), j_1) = (M(P_1, Q), \lambda(P_1, Q)),\\
\sigma_\infty &\colon (M_\infty^{\otimes c}, \lambda_\infty^{\otimes c}) \to Q^*(\mathcal{O}_{\mathbb{P}^1}(-P_\infty), j_\infty) = (M(P_\infty, Q), \lambda(P_\infty, Q)).
\end{aligned}$$

If $Q = (s : t)$, it follows from the assumption that R is a principal ideal domain and Lemma 3.14 that $s = -A \cdot x^a$, $t = C \cdot z^c$, and $s - t = B \cdot y^b$ for some $A, B, C \in R^\times$ and $x, y, z \in R$. Since $-s + (s - t) + t = 0$, this implies that $Ax^a + By^c + Cz^c = 0$. Moreover, since $sR + tR = R$, we also have that $x^a R + z^c R = R$. □

3.4. **Triangle groups and Belyi maps.**

Situation 3.16.

- k denotes a perfect field.
- K denotes a number field, with ring of integers $\mathcal{O}_K$.
- For any prime $\mathfrak{p} \in \operatorname{Spec} \mathcal{O}_K$, let $K_\mathfrak{p}$ be the $\mathfrak{p}$-adic completion of K, and let $\mathbf{k}(\mathfrak{p})$ be the corresponding residue field.
- Z_k denotes a nice (smooth, projective, geometrically integral) curve (separated scheme of finite type over a field), defined over k.
- $\phi\colon Z_k \to \mathbb{P}^1_k$ will denote a k-morphism.

The fundamental group of the thrice-punctured Riemann sphere $\mathbb{CP}^1 - \{0, 1, \infty\}$ is the free group on three generators; these generators are represented by loops $\gamma_0, \gamma_1, \gamma_\infty$ going around the punctures. Introducing the stackyness imposes the relations

$$\gamma_0^a = \gamma_1^b = \gamma_\infty^c = \gamma_0\gamma_1\gamma_\infty = 1$$

on the generators. This is the fundamental group of the Belyi orbifold $\mathbb{P}^1(a, b, c)(\mathbb{C})$. The abstract group defined by these generators and relations is the triangle group $\overline{\Delta}(a, b, c)$. For more on this topic see [CV19, Section 2], [Mag74, Chapter II]. (More generally, the fundamental groups of any orbifold curve can be calculated via van Kampen's theorem [BN06, Proposition 5.6].)

Definition 3.17. Let Z_k be a nice curve defined over a perfect field k. A k-Belyi map is a finite k-morphism $\phi\colon Z_k \to \mathbb{P}^1_k$ that is unramified outside $\{0, 1, \infty\} \subset \mathbb{P}^1(k)$.

Remark 3.18. These remarkable covers of the projective line are named after the Ukrainian mathematician G. V. Belyi , who famously proved that a complex algebraic curve can be defined over a number field if and only if it admits a $\mathbb{C}$-Belyi map [Bel79, Bel02]. For this reason, it is customary to require that $k \subset \mathbb{C}$ to use the term *Belyi* map. We ignore this convention, and allow k to have positive characteristic.

Since $\pi_1(\mathbb{P}^1(a,b,c)(\mathbb{C}))$ is the triangle group $\bar{\Delta}(a,b,c)$, the Riemann Existence Theorem guarantees that monodromy groups of Galois Belyi maps are always finite quotients of triangle groups.

Definition 3.19. Let $\phi\colon Z_k \to \mathbb{P}^1_k$ be a k-Belyi map with automorphism k-group scheme $\mathrm{Aut}(\phi)$. We say that ϕ is geometrically Galois with Galois group G if the extension of function fields $\mathbf{k}(Z_{\bar{k}}) \supset \mathbf{k}(\mathbb{P}^1_{\bar{k}})$ is Galois, with Galois group G. Equivalently, ϕ is geometrically Galois if the monodromy group $\mathrm{Aut}(\phi)(\bar{k})$ is isomorphic to G and acts transitively on the set of critical points $\phi^{-1}\{0,1,\infty\} \subset Z(\bar{k})$. This is the case if and only if $\#\,\mathrm{Aut}(\phi)(\bar{k}) = \#G = \deg\phi$.

Definition 3.20. The signature of a geometrically Galois k-Belyi map $\phi\colon Z_k \to \mathbb{P}^1_k$ is the triple (e_0, e_1, e_∞) where e_P is the ramification index $e_\phi(z)$ of any critical point $z \in Z_k$ with critical value $P \in \{0,1,\infty\}$. The Euler characteristic of ϕ is the quantity

$$\chi(\phi) := \tfrac{1}{e_0} + \tfrac{1}{e_1} + \tfrac{1}{e_\infty} - 1. \tag{20}$$

As a consequence of the Riemann Existence Theorem, there exist Galois Belyi maps of any signature. See [DG95, Proposition 3.1] and [Poo05, Lemma 2.5].

Proposition 3.21. *For any positive integers $a, b, c > 1$, there exists a number field K and a geometrically Galois K-Belyi map $\phi\colon Z_K \to \mathbb{P}^1_K$ of signature $(e_0, e_1, e_\infty) = (a,b,c)$. Let g be the genus of Z_K, and G be the monodromy group of ϕ. Then $2 - 2g = \deg\phi \cdot \chi(\phi)$. In particular,*

(i) If $\chi(\phi) > 0$, then $g = 0$ and $\deg\phi = \#G(\bar{K}) = 2/\chi(\phi)$.
(ii) If $\chi(\phi) = 0$, then $g = 1$.
(iii) If $\chi(\phi) < 0$, then $g > 1$.

The definition of the Belyi stack implies the following.

Lemma 3.22. *Let $\phi\colon Z_K \to \mathbb{P}^1_K$ be a geometrically Galois K-Belyi map of signature (a,b,c). Then, there exists an étale $\mathrm{Aut}(\phi)$-torsor $\psi\colon Z_K \to \mathbb{P}^1(a,b,c)_K$ such that Diagram (21) commutes.*

$$\begin{tikzcd}
Z_K \arrow[dd, "\phi", "\text{geometrically Galois Belyi}"'] \arrow[dr, dashed, "\text{étale } \mathrm{Aut}(\phi)\text{-torsor}", "\psi"'] & \\
& \mathbb{P}^1(a,b,c)_K \arrow[dl, "\text{coarse}"] \\
\mathbb{P}^1_K\,. &
\end{tikzcd} \tag{21}$$

We wish to find integral models for our geometrically Galois Belyi maps defined over number fields with certain good reduction properties. Informally, we want to spread out Lemma 3.22 to the ring of $\mathcal{T}$-integers in $\mathcal{O}_K$ for a certain finite set of primes (containing the archimedean primes). To accomplish this, we rely on the work of Beckmann [Bec89, Bec91], which has been expanded and refined by [Con00], [DG11], [DG12], and more recently by [BCLV25].

Lemma 3.23 (Good reduction)**.** *Let $\phi\colon Z_K \to \mathbb{P}^1_K$ be a geometrically Galois K-Belyi map, with Galois group G and signature (a,b,c). Then, there exists a finite set of primes $\mathcal{T}$ in K and a model $\Phi\colon Z \to \mathbb{P}^1_R$, defined over $R = \mathcal{O}_K[\mathcal{T}^{-1}]$, such that for every $\mathfrak{p} \notin \mathcal{T}$:*

(1) ϕ has good reduction at $\mathfrak{p}$ (meaning that $\phi \times_K K_{\mathfrak{p}}$ has good reduction in the sense of [BCLV25, Definition 4.1]), and

(2) the special fiber $\Phi_{\mathfrak{p}}\colon Z_{\mathbf{k}(\mathfrak{p})} \to \mathbb{P}^1_{\mathbf{k}(\mathfrak{p})}$ is a geometrically Galois $\mathbf{k}(\mathfrak{p})$-Belyi map with Galois group G and signature (a,b,c).

More over, under these conditions, there exists an étale $\mathrm{Aut}(\Phi)$-torsor $\Psi\colon Z \to \mathbb{P}^1(a,b,c)_R$ such that Diagram (22) commutes.

$$\begin{tikzcd}
Z \arrow[dd, "\Phi"'] \arrow[dr, dashed, "\text{étale } \mathrm{Aut}(\Phi)\text{-torsor}", "\Psi"'] & \\
& \mathbb{P}^1(a,b,c)_R \arrow[dl, "\text{coarse}"] \\
\mathbb{P}^1_R\,. &
\end{tikzcd} \tag{22}$$

Proof. The existence of the finite set $\mathcal{T}$ for which (1) holds is proved in [BCLV25, Lemma 5.1]: note that if $L \supseteq K$ is the smallest field extension over which ϕ_L is Galois, then $Z_L \to Z_K \to \mathbb{P}^1$ is the Galois closure of ϕ. For (2) and the final statement, we can apply [BCLV25, Theorem 5.3]. □

4. The stack $[\mathcal{U}/\mathsf{H}]$

4.1. **The group H.** For this section we will need some basic notions from the theory of diagonalizable group schemes of multiplicative type. See the notes of Oésterle [Oes14] and Conrad [Con14, Appendix B].

Given a base scheme S, and a finitely generated $\mathbb{Z}$-module M, we define $\mathbf{D}_S(M)$ to be the S-group scheme $\operatorname{Spec} \mathcal{O}_S[M]$ representing the functor $\underline{\mathrm{Hom}}_{S-\mathtt{GrpSch}}(M_S, \mathbb{G}_{\mathrm{m}})$ of characters of the constant S-group scheme M_S. An S-group scheme is called diagonalizable if it is isomorphic to $\mathbf{D}_S(M)$ for some finitely generated $\mathbb{Z}$-module M. Moreover, $\mathbf{D}_S$ gives a contravariant functor between finitely generated $\mathbb{Z}$-modules and the category of diagonalizable S-group schemes satisfying certain exactness properties that are summarized in [Oes14, 5.3].

Situation 4.1. Let

- $\mathbf{D}$ denote the functor described above, over the base scheme $S = \operatorname{Spec} \mathbb{Z}$.
- (a, b, c) be a triple of positive integers.
- $m := \gcd(bc, ac, ab)$, and define the weight vector of (a, b, c) by $\mathbf{w} = (w_0, w_1, w_\infty)$, where $w_0 = bc/m$, $w_1 = ac/m$ and $w_\infty = ab/m$.
- $\mathbb{G}_{\mathrm{m}}(\mathbf{w})$ be the image of the (injective) homomorphism $\mathbb{G}_{\mathrm{m}} \to \mathbb{G}_{\mathrm{m}}^3$ given by $\lambda \mapsto (\lambda^{w_0}, \lambda^{w_1}, \lambda^{w_\infty})$.
- $\bar{\triangle}(a, b, c)$ denote the triangle group
$$\bar{\triangle}(a, b, c) = \langle \gamma_0, \gamma_1, \gamma_\infty : \gamma_0^a = \gamma_1^b = \gamma_\infty^c = \gamma_0\gamma_1\gamma_\infty = 1 \rangle.$$

Definition 4.2. Consider the finitely generated $\mathbb{Z}$-module

$$M := \langle (a, -b, 0), (0, b, -c), (-a, 0, c) \rangle \subset \mathbb{Z}^3. \tag{23}$$

Define H to be the subgroup $\mathbf{D}(\mathbb{Z}^3/M)$ of $\mathbb{G}_{\mathrm{m}}^3 = \mathbf{D}(\mathbb{Z}^3)$.

The diagonalizable group H admits a maximal torus corresponding to the $\mathbb{Z}$-free part of $\mathbb{Z}^3/M$. Moreover, we have the following characterization. An important formula to keep in mind is

$$\operatorname{lcm}(a, b, c) = \frac{abc}{\gcd(bc, ac, ab)}. \tag{24}$$

Lemma 4.3 (The structure of H)**.** *Let $\mathsf{K} = \mathbf{D}(\bar{\triangle}(a, b, c)^{\mathrm{ab}})$, and recall that $m = \gcd(bc, ac, ab)$.*

(1) The $\mathbb{Z}$-module $\mathbb{Z}^3/M$ is isomorphic to $\mathbb{Z} \oplus \bar{\triangle}(a, b, c)^{\mathrm{ab}}$.

(2) Let K be the kernel of the map
$$\mu_a \times \mu_b \times \mu_c \to \mu_{\operatorname{lcm}(a,b,c)}, \quad (\xi_0, \xi_1, \xi_\infty) \mapsto \xi_0 \cdot \xi_1 \cdot \xi_\infty.$$
Then $\mathsf{K} \cong \mathbf{D}(\bar{\triangle}(a, b, c)^{\mathrm{ab}})$.

(3) The group scheme H is equal to $\mathbb{G}_m(\mathbf{w}) \cdot \mathsf{K}$ and isomorphic to $\mathbb{G}_m \times \mathsf{K}$.

(4) In particular, when $m = 1$, $\mathsf{H} = \mathbb{G}_m(\mathbf{w}) \cong \mathbb{G}_m$.

Proof. (1) We calculate the invariant factor decomposition of $\mathbb{Z}^3/M$ from the Smith normal form of the matrix having the generators of M as its rows [Sta16, Theorem 2.3]. Let

$$\mathsf{m} = \begin{bmatrix} a & -b & 0 \\ 0 & b & -c \\ -a & 0 & c \end{bmatrix}.$$

From Stanley's formula [Sta16, Theorem 2.4], we see that

$$\mathrm{SNF}(\mathsf{m}) = \begin{bmatrix} d & 0 & 0 \\ 0 & m/d & 0 \\ 0 & 0 & 0 \end{bmatrix},$$

where $d = \gcd(a, b, c)$ is the greatest common divisor of the 1×1 minors, and $m = \gcd(bc, ac, ab)$ is the greatest common divisor of the 2×2 minors. It follows that $\mathbb{Z}^3/M \cong \mathbb{Z} \oplus \mathbb{Z}/d\mathbb{Z} \oplus \mathbb{Z}/(m/d)\mathbb{Z}$.

It remains to show that $\mathbb{Z}/d\mathbb{Z} \oplus \mathbb{Z}/(m/d)\mathbb{Z}$ is isomorphic to $\bar{\triangle}(a, b, c)^{\mathsf{ab}}$. To this end, note that the group $\bar{\triangle}(a, b, c)^{\mathsf{ab}}$ is isomorphic to the quotient of $\mathbb{Z}^3$ by the subgroup

$$J = \langle (a, 0, 0), (0, b, 0), (0, 0, c), (1, 1, 1) \rangle.$$

As before, we calculate the invariant factor decomposition of $\mathbb{Z}^3/J$ via a Smith normal form computation.

$$\mathrm{SNF} \begin{bmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & d & 0 \\ 0 & 0 & m/d \\ 0 & 0 & 0 \end{bmatrix}.$$

We conclude that $\bar{\triangle}(a, b, c)^{\mathsf{ab}} \cong \mathbb{Z}^3/J \cong \mathbb{Z}/d\mathbb{Z} \oplus \mathbb{Z}/(m/d)\mathbb{Z}$.

(2) From the presentation given in Situation 4.1, we see that $\bar{\triangle}(a, b, c)^{\mathsf{ab}}$ is the cokernel of the map $\mathbb{Z}/l\mathbb{Z} \to \mathbb{Z}/a\mathbb{Z} \oplus \mathbb{Z}/b\mathbb{Z} \oplus \mathbb{Z}/c\mathbb{Z}$ taking $1 \bmod l \mapsto (1 \bmod a, 1 \bmod b, 1 \bmod c)$, where $l = \mathrm{lcm}(a, b, c)$. The result follows by applying the functor $\mathbf{D}$.

(3) The computation above shows that $\mathbb{Z}^3/M$ has $\mathbb{Z}$-rank one. The free part of $\mathbb{Z}^3/M$ corresponds to the (dual of the) kernel of the matrix m. That is, we want a generator for the subgroup of $\mathbf{v} \in \mathbb{Z}^3$ such that

$$\begin{bmatrix} a & -b & 0 \\ 0 & b & -c \\ -a & 0 & c \end{bmatrix} \mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

In other words, we are looking for minimal $v_1, v_2, v_3 \in \mathbb{Z}$ satisfying $av_1 = bv_2 = cv_3$. But this is precisely the property defining the weight vector $\mathbf{w}$ (see Situation 4.1). The equality $\mathsf{H} = \mathbb{G}_{\mathrm{m}}(\mathbf{w}) \cdot \mathsf{K}$ follows from the exact sequence $0 \to \mathbb{Z}^3/\langle \mathbf{w} \rangle \to \mathbb{Z}^3/M \to \mathbb{Z}^3/J \to 0$ and the exactness of the functor $\mathbf{D}$.

The statement that $\mathsf{H} \cong \mathbb{G}_{\mathrm{m}} \times \mathsf{K}$ follows from the fact that $\mathbb{Z}^3/\langle \mathbf{w} \rangle$ has $\mathbb{Z}$-rank one and the general fact that $\mathbf{D}(M_1 \oplus M_2) \cong \mathbf{D}(M_1) \times \mathbf{D}(M_2)$ for arbitrary finitely generated $\mathbb{Z}$-modules M_1, M_2. □

Lemma 4.4. *Let $\mathcal{S}$ be a finite set of rational primes, and let $R = \mathbb{Z}[\mathcal{S}^{-1}]$. Then $\mathrm{H}^1(R, \mathsf{H}_R)$ is finite. Moreover, $\mathrm{H}^1(\mathbb{Z}, \mathsf{H})$ is trivial.*

Proof. From $\mathsf{H} \cong \mathbb{G}_{\mathrm{m}} \times \mathsf{K}$, we obtain the exact sequence $\mathrm{H}^1(R, \mathbb{G}_{\mathrm{m}}) \to \mathrm{H}^1(R, \mathsf{H}) \to \mathrm{H}^1(R, \mathsf{K}) \to \mathrm{H}^2(R, \mathbb{G}_{\mathrm{m}})$. Since $\mathrm{H}^1(R, \mathbb{G}_{\mathrm{m}}) = \operatorname{Pic} \mathbb{Z}$ is trivial, we have that $\mathrm{H}^1(R, \mathsf{H})$ injects into the finite group $\mathrm{H}^1(R, \mathsf{K})$. In the special case of $R = \mathbb{Z}$, $\mathrm{H}^2(\mathbb{Z}, \mathbb{G}_{\mathrm{m}}) = \operatorname{Br} \mathbb{Z}$ is also trivial, and we obtain that $\mathrm{H}^1(\mathbb{Z}, \mathsf{H}) \cong \mathrm{H}^1(\mathbb{Z}, \mathsf{K})$. But Minkowski's theorem implies that $\mathrm{H}^1(\mathbb{Z}, \mathsf{K})$ is trivial. □

4.2. Proof of the main theorem.

We are ready to prove the main result.

Situation 4.5. We place ourselves in the following situation for the remainder of this section.

- Let $F := \operatorname{Spec} \mathbb{Z}[\mathsf{x}, \mathsf{y}, \mathsf{z}]/\langle A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c \rangle \subset \mathbb{A}^3$.
- Let $\mathcal{T}$ be set of primes dividing the integer $a \cdot b \cdot c \cdot A \cdot B \cdot C$.
- Let $R = \mathbb{Z}[\mathcal{T}^{-1}]$ be the ring of $\mathcal{T}$-integers.
- Let H be the affine group scheme introduced in Definition 4.2.
- Let $\mathcal{U}$ be the punctured cone associated to F, defined over R.
- Let $s\colon \operatorname{Spec} k \to \operatorname{Spec} R$ denote a geometric point.
- For a geometric object $\mathcal{X}$ defined over R, we let $\mathcal{X}_s := \mathcal{X} \times_s \operatorname{Spec} k$ denote the geometric fiber above s.

We recall the statement of our main theorem and then proceed to prove some preliminary lemmas.

Theorem 4.6. *The map*

$$j\colon \mathcal{U} \to \mathbb{P}^1_R, \quad (x, y, z) \mapsto (-Ax^a : Cz^c) \tag{25}$$

induces an isomorphism $\mathbf{j}\colon [\mathcal{U}/\mathsf{H}_R] \cong \mathbb{P}^1(a, b, c)_R$.

The reason we are interested in the group scheme H is that it arises as the stabilizer in $\mathbb{G}_{\mathrm{m}}^3$ of the punctured cone $\mathcal{U}$ associated to a generalized Fermat equation.

Lemma 4.7. *Let S be the stabilizer subgroup of $\mathcal{U}$ under the action of $\mathbb{G}_m^3$ on $\mathbb{A}^3_{\mathbb{Z}}$. Then, $\mathsf{H} \subset \mathsf{S}$ and $\mathsf{H}_R = \mathsf{S}_R$.*

Proof. By definition, $\mathsf{S} := \mathrm{Stab}_{\mathbb{G}_{\mathrm{m}}^3}(\mathcal{U})$ is the group scheme that takes any $\mathbb{Z}$-algebra B to the group

$$\mathsf{S}(B) = \left\{(\lambda_0, \lambda_1, \lambda_\infty) \in (B^\times)^3 : F(\lambda_0 \mathsf{x}, \lambda_1 \mathsf{y}, \lambda_\infty \mathsf{z})/F(\mathsf{x}, \mathsf{y}, \mathsf{z}) \in B^\times\right\},$$

and this group visibly contains

$$\mathsf{H}(B) = \left\{(\lambda_0, \lambda_1, \lambda_\infty) \in (B^\times)^3 : \lambda_0^a = \lambda_1^b = \lambda_\infty^c\right\}.$$

So we have an inclusion $\mathsf{H} \hookrightarrow \mathsf{S}$. For every geometric point $s\colon \operatorname{Spec} k \to \operatorname{Spec} R$, this inclusion pulls back to an equality $\mathsf{S}_s = \mathsf{H}_s$, so we conclude that $\mathsf{S}_R = \mathsf{H}_R$ by fpqc descent [Sta24, Tag 02L4] and spreading out. □

We start by considering the situation on the geometric fibers.

Lemma 4.8. *For every geometric point $s\colon \operatorname{Spec} k \to \operatorname{Spec} R$, the map*

$$j\colon \mathcal{U}_s \to \mathbb{P}^1_s, \quad (x, y, z) \mapsto (-Ax^a : Cz^c) \tag{26}$$

induces an isomorphism $\mathbf{j}_s\colon [\mathcal{U}_s/\mathsf{H}_s] \cong \mathbb{P}^1(a, b, c)_s$.

Proof. We omit the subscript "s" and work over k throughout. We start by showing that j induces a coarse map $j\colon [\mathcal{U}/\mathsf{H}] \to \mathbb{P}^1$. Recall that $\mathcal{R} = k[\mathsf{x}, \mathsf{y}, \mathsf{z}]/\langle A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c\rangle$ is the coordinate ring of F. Consider the affine open $D(\mathsf{z}) \subset F$, with corresponding coordinate ring $\mathcal{R}[1/\mathsf{z}]$. Note that $\mathcal{U} \cap D(\mathsf{z}) = D(\mathsf{z})$. Since $D(\mathsf{z}) = \operatorname{Spec} \mathcal{R}[1/\mathsf{z}]$ is affine, H is linearly reductive, and $[D(\mathsf{z})/\mathsf{H}]$ is tame, the natural map $[\operatorname{Spec} \mathcal{R}[1/\mathsf{z}]/\mathsf{H}] \to \operatorname{Spec} \mathcal{R}[1/\mathsf{z}]^H$ is a good moduli space and thus a coarse moduli space (see [Alp13, Theorem 13.2 and Remark 7.3]). Now, we calculate that $\mathcal{R}[1/\mathsf{z}]^H = k\left[\frac{-A\mathsf{x}^a}{C\mathsf{z}^c}\right]$. Applying the same argument to $D(\mathsf{x})$, the result follows by glueing the maps

$$[\mathcal{U} \cap D(\mathsf{x})/\mathsf{H}] \to \operatorname{Spec} k\left[\tfrac{-C\mathsf{z}^c}{A\mathsf{x}^a}\right], \quad [\mathcal{U} \cap D(\mathsf{z})/\mathsf{H}] \quad \to \operatorname{Spec} k\left[\tfrac{-A\mathsf{x}^a}{C\mathsf{z}^c}\right]$$

to obtain the coarse map $j\colon [\mathcal{U}/\mathsf{H}] \to \mathbb{P}^1$.

We proceed to show that $[\mathcal{U}/\mathsf{H}] \cong \mathbb{P}^1(a, b, c)$. By definition of $\mathbb{P}^1(a, b, c)$ as an iterated root stack, the map $j\colon \mathcal{U} \to \mathbb{P}^1$ induces a map $\mathbf{j}\colon [\mathcal{U}/\mathsf{H}] \to \mathbb{P}^1(a, b, c)$. Indeed, the map $j\colon \mathcal{U} \to \mathbb{P}^1$ satisfies

$$j^*\mathcal{O}_{\mathbb{P}^1}(-P_0) = \mathcal{L}_0^a, \quad j^*\mathcal{O}_{\mathbb{P}^1}(-P_1) = \mathcal{L}_1^b, j^*\mathcal{O}_{\mathbb{P}^1}(-P_\infty) = \mathcal{L}_\infty^c,$$

with $\mathcal{L}_0 = \mathsf{x} \cdot \mathcal{O}_{\mathcal{U}}, \mathcal{L}_1 = \mathsf{y} \cdot \mathcal{O}_{\mathcal{U}}$ and $\mathcal{L}_\infty = \mathsf{z} \cdot \mathcal{O}_{\mathcal{U}}$, and this gives rise an object in $\mathbb{P}^1(a, b, c)(\mathcal{U})$.

Since $[\mathcal{U}/\mathsf{H}]\langle k\rangle = \mathbb{P}^1(k) = \mathbb{P}^1(a, b, c)\langle k\rangle$, and the map $[\mathcal{U}/\mathsf{H}](k) \to \mathbb{P}^1(a, b, c)(k)$ induces isomorphisms between the stabilizer groups of the stacky points

$$\begin{aligned} \mathrm{Stab}_{\mathsf{H}}(V(\mathsf{x})) &\cong \mu_a(k), \\ \mathrm{Stab}_{\mathsf{H}}(V(\mathsf{y})) &\cong \mu_b(k), \\ \mathrm{Stab}_{\mathsf{H}}(V(\mathsf{z})) &\cong \mu_c(k). \end{aligned}$$

The result follows from [VZB22, Lemma 5.3.10(a)]. □

Proof of Theorem 4.6. The R-morphism j is surjective (this can be checked on geometric fibers by fpqc descent [Sta24, Tag 02KV] and spreading out) and H_R-invariant. From Lemma 2.14, this induces a morphism $[\mathcal{U}/\mathsf{H}_R] \to \mathbb{P}^1_R$, which factors through the coarse map $\mathbb{P}^1(a,b,c)_R \to \mathbb{P}^1_R$ by the definition of the Belyi stack. Both $\mathbb{P}^1(a,b,c)_R$ and $[\mathcal{U}/\mathsf{H}_R]$ are tame relative stacky curves. To calculate the coarse space of $\mathcal{U}/\mathsf{H}$ of $[\mathcal{U}/\mathsf{H}]$, we use the same argument as in the proof of Lemma 4.8.

$$\begin{tikzcd}
 & \mathcal{U} \arrow[dl] \arrow[dr] \arrow[dd, "j"] & \\
{[\mathcal{U}/\mathsf{H}_R]} \arrow[rr, dashed, "\mathrm{j}"] \arrow[dr, "\text{coarse}"'] & & \mathbb{P}^1(a,b,c)_R \arrow[dl, "\text{coarse}"] \\
 & \mathbb{P}^1_R &
\end{tikzcd}$$

In summary, we have a morphism $\mathbf{j}\colon [\mathcal{U}/\mathsf{H}]_R \to \mathbb{P}^1(a,b,c)_R$ with the property that on each geometric fiber, the induced map on the coarse spaces $(\mathcal{U}/\mathsf{H})_s \to \mathbb{P}^1_s$ is an isomorphism inducing a stabilizer-preserving bijection between $[\mathcal{U}/\mathsf{H}]\langle\bar{k}\rangle$ and $\mathbb{P}^1(a,b,c)\langle\bar{k}\rangle$. [VZB22, Lemma 5.3.10 (a)] implies that $\widetilde{(j)}_s$ is an isomorphism for every geometric point of $\operatorname{Spec} R$, and this implies that the same is true globally. Alternatively, we can apply Santens' characterization of tame relative stacky curves [San22b, Lemma 2.1]. □

5. The method of Fermat descent

What follows is a brief discussion of the method of Fermat descent. This is not intended to be a comprehensive algorithm for finding the primitive integral solutions of an arbitrary generalized Fermat equation with integer coefficients. Rather, it is meant as an artisanal guide to the method, from the point of view developed in this article.

Situation 5.1. We fix the following setup for the remainder of this section.

- Let (a,b,c) be a triple of positive integers, with $a,b,c > 1$.
- Let $F\colon A\mathsf{x}^a + B\mathsf{y}^b + C\mathsf{z}^c = 0 \subset \mathbb{A}^3_{\mathbb{Z}}$ be a generalized Fermat equation.
- Let $\mathcal{S}$ be the set of primes p dividing $a \cdot b \cdot c \cdot A \cdot B \cdot C$, and $R := \mathbb{Z}[\mathcal{S}^{-1}]$.
- Let $\mathcal{U}$ denote the punctured cone associated to F.
- Let H be as in Definition 4.2.
- Let $j\colon \mathcal{U} \to \mathbb{P}^1$ be the morphism $(x,y,z) \mapsto (-Ax^a : Cz^c)$.

As a consequence of Theorem 4.6, we have that $j(\mathcal{U}(\mathbb{Z}))$ is contained in the set $\mathbb{P}^1(a,b,c)\langle R\rangle \subset \mathbb{P}^1(R) = \mathbb{P}^1(\mathbb{Q})$. This is our starting point. The method of Fermat descent consists of three steps: covering, twisting, and sieving.

5.1. **Covering.** The goal is to find a geometrically Galois Belyi map (Definition 3.19) $\phi\colon Z_K \to \mathbb{P}^1_K$. We know that one exists from Proposition 3.21, but we might have to base extend to a number field K to find it. Furthermore, the map ϕ admits an integral model Φ over $\operatorname{Spec} \mathcal{O}_K[\mathcal{T}^{-1}]$ for some finite set of primes $\mathcal{T}$ which we can arrange to contain the primes above $\mathcal{S}$. In practice, it is desirable to minimize the complexity of this data as much as possible.

5.2. **Twisting.** Now that we found a covering $\Phi\colon Z \to \mathbb{P}^1_{\mathcal{O}_K[\mathcal{T}^{-1}]}$ with automorphism group scheme $\mathbf{Aut}(\Phi)$, we are tasked with finding:

(1) The (Čech fppf) cohomology set $\mathrm{H}^1(\mathcal{O}_K[\mathcal{T}^{-1}], \mathbf{Aut}(\Phi))$.
(2) For each $\tau \in \mathrm{H}^1(\mathcal{O}_K[\mathcal{T}^{-1}], \mathbf{Aut}(\Phi))$, the twist $\Phi_\tau\colon Z_\tau \to \mathbb{P}^1$

For (1), the first observation is that there is an isomorphism of pointed sets between $\mathrm{H}^1(\mathcal{O}_K[\mathcal{T}^{-1}], \mathbf{Aut}(\Phi))$ is in bijective correspondence with the Galois cohomology set $\mathrm{H}^1_{\mathcal{T}}(K, \mathrm{Gal}(\phi))$, parametrizing isomorphism classes of Galois étale K-algebras with Galois group $\mathrm{Gal}(\phi) = \mathbf{Aut}(\phi)(\bar{K}) = \mathrm{Aut}(\phi_{\bar{K}})$, and unramified outside $\mathcal{T}$. The second observation is that we might not need the full cohomology set. More precisely, for our purposes of finding $\mathcal{U}(\mathbb{Z})$, we want to identify the subset of those τ for which $\Phi_\tau(Z_\tau(K)) \cap j(\mathcal{U}(\mathbb{Z})) \neq \emptyset$.

5.3. **Sieving.** Let $R' = \mathcal{O}_K[\mathcal{T}^{-1}]$. If we reach this step, we have found a subset $T \subset \mathrm{H}^1(R', \mathbf{Aut}(\Phi))$ such that

$$j(\mathcal{U}(\mathbb{Z})) \subset \bigsqcup_{\tau \in T} \Phi_\tau(Z_\tau(R')) = \bigsqcup_{\tau \in T} \phi_\tau(Z_\tau(K)).$$

We use the word "sieve" to mean that our goal is separate the points on the right hand side that do not come from primitive integral solutions to F. For this purpose, it might be helpful to recall that $j(\mathcal{U}(\mathbb{Z}))$ is also contained in $\mathbb{P}^1(a,b,c)\langle R\rangle$, where $R = \mathbb{Z}[\mathcal{S}]$ is a principal ideal domain, to apply Lemma 3.14.

References

[Alp13] Jarod Alper, *Good moduli spaces for Artin stacks*, Ann. Inst. Fourier (Grenoble) **63** (2013), no. 6, 2349–2402. MR 3237451

[BCLV25] Irmak Balçik, Stephanie Chan, Yuan Liu, and Bianca Viray, *Number fields generated by points in linear systems on curves*, arXiv e-prints (2025), arXiv:2503.07846v1.

[Bec89] Sybilla Beckmann, *Ramified primes in the field of moduli of branched coverings of curves*, J. Algebra **125** (1989), no. 1, 236–255. MR 1012673

[Bec91] ______, *On extensions of number fields obtained by specializing branched coverings*, J. Reine Angew. Math. **419** (1991), 27–53. MR 1116916

[Bel79] G. V. Belyi, *Galois extensions of a maximal cyclotomic field*, Izv. Akad. Nauk SSSR Ser. Mat. **43** (1979), no. 2, 267–276, 479. MR 534593

[Bel02] ______, *A new proof of the three-point theorem*, Mat. Sb. **193** (2002), no. 3, 21–24. MR 1913596

[BN06] Kai Behrend and Behrang Noohi, *Uniformization of Deligne-Mumford curves*, J. Reine Angew. Math. **599** (2006), 111–153. MR 2279100

[Cad07] Charles Cadman, *Using stacks to impose tangency conditions on curves*, Amer. J. Math. **129** (2007), no. 2, 405–427. MR 2306040

[Cona] Keith Conrad, *The congruent number problem*, https://kconrad.math.uconn.edu/blurbs/ugradnumthy/congnumber.pdf, Accessed: July 16, 2025.

[Conb] ______, *Proofs by descent*, https://kconrad.math.uconn.edu/blurbs/ugradnumthy/descent.pdf, Accessed: July 16, 2025.

[Con00] Brian Conrad, *Inertia groups and fibers*, J. Reine Angew. Math. **522** (2000), 1–26. MR 1759533

[Con14] ______, *Reductive group schemes*, Autour des schémas en groupes. Vol. I, Panor. Synthèses, vol. 42/43, Soc. Math. France, Paris, 2014, pp. 93–444. MR 3362641

[CV19] Pete L. Clark and John Voight, *Algebraic curves uniformized by congruence subgroups of triangle groups*, Trans. Amer. Math. Soc. **371** (2019), no. 1, 33–82. MR 3885137

[Dar97] H. Darmon, *Faltings plus epsilon, Wiles plus epsilon, and the generalized Fermat equation*, C. R. Math. Rep. Acad. Sci. Canada **19** (1997), no. 1, 3–14. MR 1479291

[DG95] Henri Darmon and Andrew Granville, *On the equations $z^m = F(x,y)$ and $Ax^p + By^q = Cz^r$*, Bull. London Math. Soc. **27** (1995), no. 6, 513–543. MR 1348707

[DG11] Pierre Dèbes and Nour Ghazi, *Specializations of Galois covers of the line*, “Alexandru Myller” Mathematical Seminar, AIP Conf. Proc., vol. 1329, Amer. Inst. Phys., Melville, NY, 2011, pp. 98–108. MR 2752504

[DG12] ______, *Galois covers and the Hilbert-Grunwald property*, Ann. Inst. Fourier (Grenoble) **62** (2012), no. 3, 989–1013. MR 3013814

[Dic66] Leonard Eugene Dickson, *History of the theory of numbers. Vol. I: Divisibility and primality*, Chelsea Publishing Co., New York, 1966. MR 245499

[GS17] Anton Geraschenko and Matthew Satriano, *A "bottom up" characterization of smooth Deligne-Mumford stacks*, Int. Math. Res. Not. IMRN **2017** (2017), no. 21, 6469–6483. MR 3719470

[LMF25a] The LMFDB Collaboration, *The L-functions and modular forms database*, https://www.lmfdb.org, 2025, [Online; accessed 21 July 2025].

[LMF25b] ______, *The L-functions and modular forms database (beta)*, https://beta.lmfdb.org/Belyi/4T1/4/4/2.2/a/, 2025, [Online; accessed 10 July 2025].

[Mag74] Wilhelm Magnus, *Noneuclidean tesselations and their groups*, Pure and Applied Mathematics, vol. Vol. 61, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1974. MR 352287

[Oes14] Joseph Oesterlé, *Schémas en groupes de type multiplicatif*, Autour des schémas en groupes. Vol. I, Panor. Synthèses, vol. 42/43, Soc. Math. France, Paris, 2014, pp. 63–91. MR 3362640

[Ols16] Martin Olsson, *Algebraic spaces and stacks*, American Mathematical Society Colloquium Publications, vol. 62, American Mathematical Society, Providence, RI, 2016. MR 3495343

[Poo05] Bjorn Poonen, *Unramified covers of Galois covers of low genus curves*, Math. Res. Lett. **12** (2005), no. 4, 475–481. MR 2155225

[Poo06] ______, *The projective line minus three fractional points*, Slides for the MSRI program: Rational and integral points on higher-dimensional varieties, July 2006.

[Poo17] ______, *Rational points on varieties*, Graduate Studies in Mathematics, vol. 186, American Mathematical Society, Providence, RI, 2017. MR 3729254

[Poo23] ______, *Some examples of stacks*, Slides for the AMS MRC: Explicit computations with stacks, June 2023.

[PSS07] Bjorn Poonen, Edward F. Schaefer, and Michael Stoll, *Twists of $X(7)$ and primitive solutions to $x^2 + y^3 = z^7$*, Duke Math. J. **137** (2007), no. 1, 103–158. MR 2309145

[San22a] Tim Santens, *Relation between stacky curves and M-curves*, MathOverflow: `https://mathoverflow.net/q/423432`, 2022, Accessed: 08-20-2025.

[San22b] Tim Santens, *The Brauer-Manin obstruction for stacky curves*, arXiv e-prints (2022), `arXiv:2210.17184v3`.

[Sil09] Joseph H. Silverman, *The arithmetic of elliptic curves*, second ed., Graduate Texts in Mathematics, vol. 106, Springer, Dordrecht, 2009. MR 2514094

[Sko01] Alexei Skorobogatov, *Torsors and rational points*, Cambridge Tracts in Mathematics, vol. 144, Cambridge University Press, Cambridge, 2001. MR 1845760

[Sta16] Richard P. Stanley, *Smith normal form in combinatorics*, J. Combin. Theory Ser. A **144** (2016), 476–495. MR 3534076

[Sta24] The Stacks project authors, *The stacks project*, `https://stacks.math.columbia.edu`, 2024.

[VZB22] John Voight and David Zureick-Brown, *The canonical ring of a stacky curve*, Mem. Amer. Math. Soc. **277** (2022), no. 1362, v+144. MR 4403928