THE ARITHMETIC AND GEOMETRY OF BRANCHED COVERINGS: THEOREMS OF BELYI AND DARMON–GRANVILLE

ANAND DEOPURKAR

1. Introduction

Consider a polynomial equation, say

$$x^4 + y^4 = 31xy + 1,$$

or

$$xyz(x + y + z) = 1.$$

What can we say about its solutions—values of the variables that satisfy the equation? The answer, of course, depends on *where* we take the values. If we take the values in $\mathbf{C}$, the set of complex numbers, then there are infinitely many solutions. The same is true if we take the values in $\mathbf{R}$, the set of real numbers. Figure 1 shows the real solutions to the two equations above. If we restrict to number systems that are more algebraically constrained, like $\mathbf{Q}$, the set of rational numbers, then the question becomes much more difficult.

A fascinating observed phenomenon is that the geometry of the space of solutions over an algebraically unconstrained field like $\mathbf{C}$ influences the existence of solutions over an algebraically constrained field like $\mathbf{Q}$. For example, the complex solutions to the first equation $x^4 + y^4 = 31xy + 1$ define a surface of genus 3—an example of a *hyperbolic* surface. A famous theorem of Faltings, conjectured by Mordell, says that equations that define a hyperbolic surface over $\mathbf{C}$ can only have finitely many rational solutions.

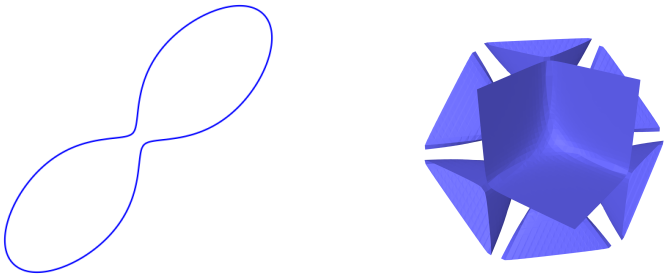

Figure 1. Real solutions of $x^4 + y^4 = 31xy + 1$ and $xyz(x + y + z) = 1$

Faltings theorem applies to equations whose complex solutions have one complex dimension. For equations that have solutions with dimension 2 or more, the precise connection between the geometry and arithmetic is still conjectural. If the space of solutions over $\mathbf{C}$ is hyperbolic in a certain sense (the precise condition is called "of general type"), then a conjecture of Bombieri and Lang predicts that the rational solutions are sparse [7, § F.5.2]. For other kinds of solution spaces over $\mathbf{C}$, our understanding of the rational points remains limited. For example, the complex solutions to the second equation $xyz(x+y+z) = 1$ define a 4-manifold called a K3 surface. We still do not completely understand in complete generality the rational solutions of equations that define a K3 surface.

The goal of this article is to demonstrate an example of the connection between geometry and arithmetic sketched above. More precisely, we prove the following.

Theorem 1.1 (Darmon and Granville)**.** *Let p, q, r be positive integers and let A, B, C be non-zero integers. If*

$$\frac{1}{p} + \frac{1}{q} + \frac{1}{r} < 1,$$

then the equation

$$Ax^p + By^q = Cz^r$$

has finitely many integral solutions (x, y, z) with $\gcd(x, y, z) = 1$.

Since we have one equation in 3 variables, the set of complex solutions to $Ax^p + By^q = Cz^r$ has dimension two—it is a complex algebraic surface. However, the equation is quasi-homogeneous. That is, it remains essentially unchanged if take any scalar t and multiply x, y, and z by t^{qr}, t^{pr}, and t^{pq}, respectively. The quotient of the solution space modulo this symmetry is an algebraic curve. Theorem 1.1 is essentially proved by applying Faltings' theorem to this quotient curve.

To handle the quotient curve above, we use a cover of the projective line $\mathbf{P}^1$ branched at three points. The crucial fact we need is that such covers can be defined by equations whose coefficients lie in a number field. There is a famous theorem due to Belyi that says that the converse is also true.

Theorem 1.2 (Belyi)**.** *Let X be a compact Riemann surface. The following two conditions are equivalent.*

1. *There exists a finite map $f\colon X \to \mathbf{P}^1$ that is unbranched on the complement of 3 points in $\mathbf{P}^1$.*
2. *X is defined over $\overline{\mathbf{Q}}$, that is, by a system of equations whose coefficients are algebraic over $\mathbf{Q}$.*

Although we do not strictly need the full strength of Theorem 1.2 for Theorem 1.1, we give include a complete treatment because Theorem 1.2 is interesting in its own right.

The deepest result we use, without any indication of its proof, is Faltings's theorem on the finiteness of rational points of a hyperbolic curve. Modulo this result, everything should be accessible to a reader with a foundational knowledge of algebraic geometry, on par with Chapter 1 and 2 of Hartshorne's book [6]. Some of the further remarks in Section 4 assume more background.

2. Belyi's theorem

Belyi's theorem is remarkable because it faithfully translates a purely topological condition into a purely algebraic condition.

Let us call covers $f\colon X \to \mathbf{P}^1$ unbranched away from three points *Belyi covers*. Let us analyse such covers using topology. First, recall that by a fractional linear transformation, any three points on $\mathbf{P}^1$ may be taken to any other three points, say $\{0, 1, \infty\}$. Thus, we may assume without loss of generality, that f is unbranched outside $\{0, 1, \infty\}$. Set

$$X^\circ = X - f^{-1}(\{0, 1, \infty\}).$$

Then $X^\circ \to \mathbf{P}^1 - \{0, 1, \infty\}$ is a finite covering space. Conversely, by the Riemann existence theorem, any finite covering space of $\mathbf{P}^1 - \{0, 1, \infty\}$ may be completed to a branched cover of $f\colon X \to \mathbf{P}^1$. Covering spaces of $\mathbf{P}^1 - \{0, 1, \infty\}$ of degree d are defined by index d subgroups of $\pi_1(\mathbf{P}^1 - \{0, 1, \infty\})$, which is the free group on two generators.

More explicitly, let

$$U \to \mathbf{P}^1 - \{0, 1, \infty\}$$

be the universal cover. Then U is biholomorphic to the upper half plane. We can identify the fundamental group of $\mathbf{P}^1 - \{0, 1, \infty\}$ with the congruence subgroup $\Gamma(2) \subset \mathrm{PSL}_2(\mathbf{Z})$, consisting of invertible integral matrices up to ± 1 which reduce to the identity modulo 2. The action of $\Gamma(2)$ on U is by fractional linear transformations

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} : z \mapsto \frac{az + b}{cz + d},$$

and $\mathbf{P}^1 - \{0, 1, \infty\}$ is the quotient. We obtain connected covering spaces $X^\circ \to \mathbf{P}^1 - \{0, 1, \infty\}$ of degree d as quotients U/G, where $G \subset \Gamma(2)$ is a subgroup of index d. Particular examples of such finite index subgroups G are the congruence subgroups $G = \Gamma(2n)$, consisting of matrices that reduce to the identity modulo $2n$. The quotients $U/\Gamma(2n)$ are modular curves. These are particular examples of Riemann surfaces that can be expressed as Belyi covers.

2.1. Belyi's theorem: (1) $\implies$ (2).

Let us prove the "easy half" of Belyi's theorem. This follows from a general principle.

Proposition 2.1. *Let M be a variety defined over $\mathbf{Q}$, which has a finite number of $\mathbf{C}$ points. Then every $\mathbf{C}$-point of M is defined over a finite extension of $\mathbf{Q}$.*

Proof. Without loss of generality, assume that M is affine (otherwise, pass to an affine cover). Say $M = \operatorname{Spec} A$, where A is a finitely generated $\mathbf{Q}$-algebra. Since M has finitely many points over $\mathbf{C}$, it follows that $A \otimes_{\mathbf{Q}} \mathbf{C}$ is a finite dimensional $\mathbf{C}$ vector space. But then A is a finite dimensional $\mathbf{Q}$ vector space.

A $\mathbf{C}$ point of M is a homomorphism $\phi\colon A \to \mathbf{C}$. Let $K \subset \mathbf{C}$ be the image of ϕ. Since A is a finite dimensional $\mathbf{Q}$ vector space, so is K, and therefore $\mathbf{Q} \to K$ is a finite extension. The $\mathbf{C}$-point defined by ϕ is defined over K. □

Fix a positive integer d. For a field K of characteristic 0, let $\mathcal{M}(K)$ denote the set of isomorphism classes of Belyi covers defined over K. It turns out that there is a natural bijection between $\mathcal{M}(K)$ and the set of K-points of a space M, of finite type over $\mathbf{Q}$. If $K = \mathbf{C}$, then our topological reformulation of Belyi covers shows that $\mathcal{M}(K)$ is finite. By Proposition 2.1, it follows that every $\mathbf{C}$-point of M is defined over a finite extension of $\mathbf{Q}$. Thanks to the natural isomorphism between K-points of M and elements of $\mathcal{M}(K)$, we conclude that every Belyi cover is defined over a finite extension of $\mathbf{Q}$.

The argument above is correct in spirit, but not in the details. The "space" M invoked above is not a variety, as required in Proposition 2.1, but a Deligne–Mumford stack. Nevertheless, Proposition 2.1 remains true for Deligne–Mumford stacks.

We now give a direct proof of (1) $\implies$ (2) that avoids stacks, but retains the spirit of the proof above. In fact, let us prove the following more general statement.

Proposition 2.2. *Fix a positive integer d. Let K be a number field and $B \subset \mathbf{P}^1_K$ a divisor defined over K. Let $f\colon Y \to \mathbf{P}^1_{\mathbf{C}}$ be a finite cover of degree d whose branch divisor is B. Then Y and f are defined over a finite extension of K.*

Proof. It suffices to prove that Y and f are defined over $\overline{\mathbf{Q}}$. Let $E = f_*\mathcal{O}_Y$. Then E is a vector bundle on $\mathbf{P}^1$. All vector bundles on $\mathbf{P}^1$ are direct sums of line bundles, and hence are defined over $\overline{\mathbf{Q}}$.

An *algebra structure* on E consists of $\mathcal{O}_{\mathbf{P}^1}$-linear maps $m\colon E \otimes E \to E$ and $i\colon \mathcal{O}_{\mathbf{P}^1} \to E$ such that with m as the multiplication and i as the structure map, E becomes a commutative and associative $\mathcal{O}_{\mathbf{P}^1}$-algebra. An algebra structure on E yields a finite flat morphism $\phi\colon \operatorname{Spec} E \to \mathbf{P}^1$. Associated to ϕ, we have the trace map $\operatorname{tr}\colon E \to \mathcal{O}_{\mathbf{P}^1}$. We also have the branch divisor $\operatorname{br}(m,i)$, defined as the determinant of the map $E \to E^\vee$ adjoint to $\operatorname{tr}\circ m\colon E \otimes E \to \mathcal{O}_{\mathbf{P}^1}$.

The conditions that (m,i) define an algebra structure are polynomial equations on the affine space $\operatorname{Hom}(E \otimes E, E) \times \operatorname{Hom}(\mathcal{O}_{\mathbf{P}^1}, E)$ The branch divisor $\operatorname{br}(m,i)$ is also defined by a polynomial expression in (m,i). So there exists a scheme of finite type over $\overline{\mathbf{Q}}$ whose points represent algebra structures on E with branch divisor B. It has an open subscheme M whose points represent E such that $\operatorname{Spec} E$ is non-singular.

The algebraic group $G = \operatorname{Aut}(E)$ acts on M. Every point of M has a finite stabiliser, and therefore all orbits are closed. Over $\mathbf{C}$, using topology we know that up to isomorphism, there are finitely many covers of degree d with branch divisor B. So M has finitely many G-orbits, and hence each G-orbit is open and closed.

The given cover $f\colon Y \to \mathbf{P}^1_{\mathbf{C}}$ represents a $\mathbf{C}$-point of M. The connected component of M containing this point must also contain a $\overline{\mathbf{Q}}$-point. This gives a cover of $\mathbf{P}^1$ defined over $\overline{\mathbf{Q}}$ isomorphic to f over $\mathbf{C}$. □

2.2. **Belyi's theorem: (2) $\implies$ (1).** Let X be a projective algebraic curve defined over $\overline{\mathbf{Q}}$. We must show that there exists a Belyi map $f\colon X \to \mathbf{P}^1$. The existence of such a map does not follow from any general principles. In fact, when it was discovered, it came as a surprise to the experts.

We construct f in three steps, following [8].

Step 1: Pick any $f\colon X \to \mathbf{P}^1$. Since X is defined over $\overline{\mathbf{Q}}$, it can be expressed as a finite cover of $\mathbf{P}^1$ over $\overline{\mathbf{Q}}$. Indeed, take any line bundle L on X defined over $\overline{\mathbf{Q}}$ of sufficiently high degree so that it has at least two linearly independent global sections, say F and G. Then $[F : G]$ defines a finite map $X \to \mathbf{P}^1$.

Step 2: Modify f so that its branch points are rational. Take an $f\colon X \to \mathbf{P}^1$ defined over $\overline{\mathbf{Q}}$. Let $S \subset \mathbf{P}^1(\overline{\mathbf{Q}})$ be the set of branch points of f. Then f is unramified over $\mathbf{P}^1 - S$.

Proposition 2.3. *There exists $g\colon \mathbf{P}^1 \to \mathbf{P}^1$ defined over $\mathbf{Q}$ such that all the branch points of $g \circ f\colon X \to \mathbf{P}^1$ are rational (contained in $\mathbf{P}^1(\mathbf{Q})$).*

Proof. We prove a slightly stronger statement. Let a finite set $S \subset \mathbf{P}^1(\overline{\mathbf{Q}})$ be given. Then there exists a finite map $g\colon \mathbf{P}^1 \to \mathbf{P}^1$, defined over $\mathbf{Q}$, such that the branch points of g and the images under g of all points of S are all points of $\mathbf{P}^1(\mathbf{Q})$. To deduce the proposition, we apply this to the S which is the set of branch points of f.

To prove the stronger statement, we first assume that S is invariant under the Galois group $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$ (if not, enlarge S to make it so). Let $S^{\mathrm{irr}} \subset S$ be the subset consisting of the points of S that are not rational. We induct on the size of S^{irr}. The base case, when the size is 0, is vacuous. Otherwise, consider the map $p\colon \mathbf{P}^1 \to \mathbf{P}^1$ on $\mathbf{A}^1$ by the polynomial

$$p(x) = \prod_{s \in S^{\mathrm{irr}}} (x - s).$$

Since S^{irr} is preserved by $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$, the polynomial p has rational coefficients. By construction, it maps S^{irr} to 0. Let $R \subset \mathbf{A}^1(\overline{\mathbf{Q}})$ be the set of roots of $p'(x)$. Then the branch points of $p\colon \mathbf{P}^1 \to \mathbf{P}^1$ that are not rational are contained in $p(R)$. By construction, the size of R, and hence the size of $p(R)$ is less that the size of S^{irr}. Also, R, and hence $p(R)$, is preserved by $\mathrm{Gal}(\overline{\mathbf{Q}}/\mathbf{Q})$. We now proceed by induction. □

Arrange the branch set to be $\{0, 1, \infty\}$. By the preceding two steps, we have $f\colon X \to \mathbf{P}^1$ whose branch set is contained in $\mathbf{P}^1(\mathbf{Q})$.

Proposition 2.4. *There exists a finite map $g\colon \mathbf{P}^1 \to \mathbf{P}^1$, defined over $\mathbf{Q}$, such that the branch points of $g \circ f\colon X \to \mathbf{P}^1$ are contained in $\{0, 1, \infty\}$.*

Proof. Again, we prove something slightly stronger. Given any finite set $S \subset \mathbf{P}^1(\mathbf{Q})$, we construct a finite map $g\colon \mathbf{P}^1 \to \mathbf{P}^1$, defined over $\mathbf{Q}$, such that the union of $g(S)$ and the branch set of g is contained in $\{0, 1, \infty\}$.

The magic ingredient is the fuction $q\colon \mathbf{P}^1 \to \mathbf{P}^1$ defined by the polynomial

$$q(x) = c \cdot x^m (1 - x)^n,$$

where m, n are positive integers and $c \in \mathbf{Q}$ is a constant. The branch points of q are $\{0, 1, \infty, \frac{m}{m+n}\}$. The map q sends 0 and 1 to 0 and ∞ to ∞. By choosing an appropriate c, we can ensure that it sends $\frac{m}{m+n}$ to 1. Then the branch points of q lie in $0, 1, \infty$.

Now, we may assume that S contains at least 3 points. After applying a fractional linear transformation, we may assume that S contains $\{0, 1, \infty\}$. Write $S = \{0, 1, \infty\} \cup T$, where T is disjoint from $\{0, 1, \infty\}$. We induct on the size of T. If T is empty, there is nothing

to prove. Otherwise, pick a $t \in T$. Using a combination of $z \mapsto 1-z$ and $z \mapsto 1/z$, both of which preserve the triplet $\{0,1,\infty\}$, we may assume that $t \in \mathbf{Q}$ satisfies $0 < t < 1$. Then $t = \frac{m}{m+n}$ for some positive integers m, n. Let $q\colon \mathbf{P}^1 \to \mathbf{P}^1$ be defined as above. Set $S' = q(T) \cup \{0,1,\infty\}$. Then $S' = \{0,1,\infty\} \cup T'$, where $T' \subset q(T - \{t\})$ has fewer elements than T. By the inductive hypothesis, there is a $g\colon \mathbf{P}^1 \to \mathbf{P}^1$, defined over $\mathbf{Q}$, such that the union of $g(q(S'))$ and the branch set of g is contained in $\{0,1,\infty\}$. Then $f = g \circ q$ achieves the desired property. □

3. Fermat like equations

Recall that we have positive integers p,q, r and integers A, B, and C. We want to prove that the equation

$$Ax^p + By^q = Cz^r \tag{1}$$

has finitely many integer solutions (x,y,z) with $\gcd(x,y,z) = 1$. We call tuples of integers with no common divisor *primitive*. So we want to prove that (1) has finitely many primitive integer solutions.

We first describe the main idea and then dive into the detials.

3.1. **Main idea.** Let X be a projective algebraic curve defined over $\mathbf{Z}$. Suppose $X_{\mathbf{C}}$ is a smooth algebraic curve, that is, a compact a Riemann surface. Whether X has finitely many or infinitely many rational solutions depends on the genus of $X_{\mathbf{C}}$. If the genus of $X_{\mathbf{C}}$ is zero, then X has infinitely many points, perhaps not over $\mathbf{Q}$, but certainly over some finite extension $K/\mathbf{Q}$. (To see why a finite extension is necessary, consider the curve defined by $x^2+y^2+z^2 = 0$ in $\mathbf{P}^2$. To see why a finite extension is sufficient, see Section 4.1.) If the genus of $X_{\mathbf{C}}$ is one, then the same holds—X has infinitely many points over some finite extension $K/\mathbf{Q}$ (but this is much less obvious than the genus zero case; see Section 4.2). But if the genus of $X_{\mathbf{C}}$ is 2 or more, then we have the following.

Theorem 3.1 (Faltings' theorem, Mordell's conjecture)**.** *Let X be a smooth projective curve over a number field K of genus ≥ 2. Then X has finitely many K-points.*

Let $Z \subset \mathbf{A}^3$ be the the algebraic variety defined by

$$Ax^p + By^q = Cz^r,$$

excluding the point $(0,0,0)$. Note that Z is a surface, not a curve, and hence Faltings' theorem does not apply directly. For simplicity, assume that p, q, r are pairwise relatively prime. The surface Z admits an action of the multiplicative group $\mathbf{G}_m$ given by

$$t \cdot (x,y,z) \mapsto (t^{qr}, t^{pr}, t^{pq}).$$

The quotient of Z by $\mathbf{G}_m$ turns out to be the projective line $\mathbf{P}^1$. But if the quotient is $\mathbf{P}^1$, then Faltings' theorem does not apply!

On closer inspection, we see that the "correct" quotient is not $\mathbf{P^1}$, but something slightly different. The action of $\mathbf{G}_m$ on Z is not free. It has non-trivial stabilisers on precisely three orbits: the orbit corresponding to $x = 0$, where the stabiliser group is μ_p, the orbit corresponding to $y = 0$, where the stabiliser group is μ_q, and the orbit corresponding to

$z = 0$, where the stabiliser group is μ_r. A more refined quotient of Z by $\mathbf{G}_m$ that takes into account the non-freeness of the action exists in the category of orbifolds, or Deligne–Mumford stacks. The Deligne–Mumford stack quotient P agrees with the usual quotient $\mathbf{P}^1$ at all points except the three points, but not at the three points. But at the three points, it "remembers" the non-trivial stabilisers.

A proper introduction to Deligne–Mumford stacks will lead us too far astray, so we stop at the informal picture. Let K_P be the canonical bundle of P. It turns out that

$$\deg K_P = 1 - \left(\frac{1}{p} + \frac{1}{q} + \frac{1}{r}\right).$$

If the quantity in parentheses is less than 1, then the degree of the canonical bundle is positive, just like an algebraic curve of genus ≥ 2. As a result, we would like say that by Faltings' theorem (more precisely, its analogue for orbifold curves), P has finitely many rational points, and hence finitely many integral points.

There is an additional wrinkle that we must address. For smooth proper curves, there is no distinction between rational points and integral points. For smooth proper *orbifold* curves, however, there is a distinction. It is *not* true that a smooth, proper, orbifold curve with negative canonical bundle has finitely many rational points; see Section 4.3 for an example. Nevertheless, it is true that such a curve has finitely many integral points. See Section 4.4 for the sketch of proof.

3.2. **Proof of Theorem 1.1.** We now give a self-contained proof of Theorem 1.1, without appealing to orbifolds. The proof follows Darmon and Granville [3].

We use a general principle encoded in the following lemma. First, let us set up some notation. For a number field K, denote by $\mathcal{O}_K \subset K$ the ring of integers. Given a K-scheme X, separated and of finite type, a model over $\mathcal{O}_K$ is an $\mathcal{O}_K$-scheme $\mathcal{X}$, separated and of finite type, such that $\mathcal{X}_K = X$.

Lemma 3.2. *Let K be a number field and $X \to W$ a finite étale morphism of separated K-schemes of finite type. Fix an $\mathcal{O}_K$-model $\mathcal{W}$ of W. There exists a field extension $K \subset L$ such that every $\mathcal{O}_K$-point of $\mathcal{W}$ is the image of an L-point of X.*

Proof. Choose an $\mathcal{O}_K$-model $\mathcal{X}$ of X with a finite map $\mathcal{X} \to \mathcal{W}$. Since $\mathcal{X}_K \to \mathcal{W}_K$ is étale, there exists a finite set of primes S of $\mathcal{O}_K$ such that

$$\mathcal{X} \otimes \mathcal{O}_K[S^{-1}] \to \mathcal{W} \otimes \mathcal{O}_K[S^{-1}]$$

is étale. Given $p\colon \operatorname{Spec} \mathcal{O}_K \to \mathcal{W}$, the fiber product $\operatorname{Spec} \mathcal{O}_K \times_{\mathcal{W}} \mathcal{X}$ is the spectrum of a finite $\mathcal{O}_K$-algebra A. Observe that $A \otimes K$ is a product of extensions of K of degree at most $d = \deg(X \to W)$, and each extension in this product is unramified over S. There are finitely many such extensions. We take L to be their compositum. □

For now, we work over $\mathbf{C}$. Consider the vanishing locus L of the homogeneous polynomial $X + Y - Z$ in $\mathbf{P}^2$. The map

$$[X : Y : Z] \mapsto [X : Z]$$

gives an isomorphism $L \to \mathbf{P}^1$. Let us represent the point $[x : z] \in \mathbf{P}^1$ by the value x/z, including ∞. Then our isomorphism sends the points $[0 : 1 : 1]$, $[1 : 0 : 1]$, and $[1 : -1 : 0]$ to 0, 1, and ∞, respectively. From now on, we identify L with $\mathbf{P}^1$ by this isomorphism.

Recall that $Z \subset \mathbf{A}^3 - \{0\}$ is the algebraic variety defined by $Ax^p + By^q = Cz^r$. We have a map $\pi\colon Z \to \mathbf{P}^1$ given by $(x, y, z) \mapsto (Ax^p : By^q : Cz^r)$. We have an action of $\mathbf{G}_m$ on Z defined so that $t \in \mathbf{G}_m$ acts by

$$t\colon (x, y, z) \mapsto (xt^{qr}, yt^{pr}, zt^{pq}).$$

Observe that π is invariant with respect to this action. That is, we have

$$\pi(t \cdot (x, y, z)) = \pi(x, y, z).$$

Also observe that π is a smooth map over $\mathbf{P}^1 - \{0, 1, \infty\}$.

Let $f\colon Y \to \mathbf{P}^1$ be a branched cover of connected Riemann surfaces satisfying the following properties:

(1) f is unramified over the complement of $\{0, 1, \infty\}$;
(2) the monodromy of f around 0, 1, and ∞ is a product of disjoint p-cycles, disjoint q-cycles, and disjoint r-cycles, respectively.

We say that such a cover has *signature* (p, q, r). It is not obvious whether such a cover exists. Indeed, it exists if and only if the equation $\alpha\beta\gamma = 1$ has a solution for α, β, γ in the symmetric group S_d where α, β, γ are products of disjoint p, q, r-cycles, respectively. It can be proved that this is possible for all (p, q, r) with $p, q, r > 1$; see Section 4.5. By the Riemann–Hurwitz formula, it follows that Y has positive genus.

Let W be the normalisation of the fiber product $Z \times_{\mathbf{P}^1} Y$. Also denote by π the induced map $\pi\colon W \to Y$. The $\mathbf{G}_m$ action on Z induces a $\mathbf{G}_m$ action on W such that the map $\pi\colon W \to Y$ is $\mathbf{G}_m$-invariant.

Lemma 3.3. *The map $W \to Z$ is finite and étale.*

Proof. Let $U = \mathbf{P}^1 - \{0, 1, \infty\}$ and $V = f^{-1}(U) \subset Y$. Then, $W|_V \to V$ is the pull-back of the smooth map $Z|_U \to U$. In particular, $W|_V$ is non-singular. Also, the map $W|_V \to Z|_U$ is the pull-back of the étale map $V \to U$. So $W|_V \to Z|_U$ is étale.

We need to check the statement in a neighborhood of a point in the complement of V. Since we are working over $\mathbf{C}$, we may rescale our variables so that $A = B = C = 1$; this will make the equations simpler. Let $y \in Y - V$ lie over $0 \in \mathbf{P}^1$; the case of points lying over 1 and ∞ is similar. We work in the analytic topology. Using u as a local coordinate around 0 for $\mathbf{P}^1$, the map $Z \to \mathbf{P}^1$ is given by $(x, y, z) \mapsto u = x^p/z^r$. So we can write Z as

$$uz^r + y^q = z^r \text{ and } uz^r = x^p.$$

In a suitable local coordinate v around y for Y, the map $Y \to \mathbf{P}^1$ is given by $u = v^p$. Then the fiber product $Z \times_{\mathbf{P}^1} Y$ is defined by

$$v^p z^r + y^q = z^r \text{ and } v^p z^r = x^p.$$

Observe that $w = x/v$ lies in the integral closure, and adjoining it leads to the equations

$$v^p z^r + y^q = z^r \text{ and } z^r = w^p.$$

It is easy to check that these equations define a non-singular space, and hence must be the equations of the normalisation.

The normalisation map $W \to Z \times_{\mathbf{P}^1} Y$ and the map $Z \times_{\mathbf{P}^1} Y \to Z$ are both finite, so $W \to Z$ is finite. Since both W and Z are non-singular, $W \to Z$ is étale if the pre-image of every point of Z is a finite union of reduced points of W. In coordinates, $W \to Z$ is given by

$$(v, w, y, z) \mapsto (x, y, z) = (wv, y, z)$$

We already know that $W \to Z$ is étale in the complement of $u = 0$. The points of Z over $u = 0$ are $(0, y_0, z_0)$ with $y_0^q = z_0^r$. The pre-image of such a point in W is defined by

$$wv = 0, \quad y = y_0, \quad z = z_0, \quad w^p = z_0^r,$$

which is a union of p reduced points. □

Having discussed the geometry over $\mathbf{C}$, we now pay closer attention to the fields of definition. By Belyi's theorem (Theorem 1.2), $f\colon Y \to \mathbf{P}^1$ is defined over a number field K. (This is the only place where we use Theorem 1.2, and observe that this is the "easy" direction of the theorem). Then W is also defined over K. We have the following diagram of varieties over K

$$\begin{array}{ccc} W & \xrightarrow[\text{étale}]{\phi} & Z \\ {\scriptstyle\pi}\big\downarrow & \mathbf{G}_m\text{-invariant} & \\ Y. & & \end{array}$$

Denote by 0 the 0-section of $\mathbf{A}^3_{\mathbf{Z}} \to \operatorname{Spec} \mathbf{Z}$. Let $\mathcal{Z} \subset \mathbf{A}^3_{\mathbf{Z}} - \{0\}$ be the scheme defined by

$$Ax^p + By^q = Cz^r.$$

Our goal is to prove that $\mathcal{Z}$ has finitely many $\mathbf{Z}$-points. By Lemma 3.2, there exists a number field L such that every $\mathbf{Z}$-point of $\mathcal{Z}$ is the image of an L-point of W. Although W may have infinitely many L-points, this infinitude is essentially due to the $\mathbf{G}_m$-action. The curve Y is of genus greater than 1, so by Faltings' theorem, has finitely many L points. The finiteness of $Y(L)$ implies the finiteness of $Z(\mathbf{Z})$, as we now show.

First of all, consider two $\mathbf{Z}$-points of Z, say (x_1, y_1, z_1) and (x_2, y_2, z_2). Considered as $\overline{\mathbf{Q}}$-points of Z, they lie in the same $\mathbf{G}_m(\overline{\mathbf{Q}})$-orbit if and only if $x_1 = \pm x_2$ and $y_1 = \pm y_2$ and $z_1 = \pm z_2$. So it suffices to show that, up to the action of $\mathbf{G}_m(\overline{\mathbf{Q}})$, we have finitely many integer points on Z. Let n be the maximum number of $\mathbf{G}_m$ orbits in the geometric fibers of $W \to Y$. Then the maximum number of L-points of W such that no two lie in the same $\mathbf{G}_m(\overline{\mathbf{Q}})$-orbit is $n|Y(L)|$. But every $\mathbf{Z}$-point of Z is the image of an L-point of W. So, up to the action of $\mathbf{G}_m(\overline{\mathbf{Q}})$, we have at most $n|Y(L)|$ integer points on Z.

4. Further remarks

4.1. **Rational points on curves of genus 0.** Let X be a smooth geometrically connected curve of genus 0 over a number field K. Then $X_{\mathbf{C}}$ is isomorphic to $\mathbf{P}^1$. However, it is possible that X is not isomorphic to $\mathbf{P}^1$ over K. For example, the curve defined by the

homogeneous equation $X^2 + Y^2 + Z^2 = 0$ is isomorphic to $\mathbf{P}^1$ over $\mathbf{C}$ but it does not even have a point defined over $\mathbf{Q}$, and hence cannot be isomorphic to $\mathbf{P}^1$ over $\mathbf{Q}$.

Nevertheless, the following statements are equivalent:

(1) X has a point over K,
(2) X is isomorphic to $\mathbf{P}^1$ over K,
(3) X has infinitely many points over K.

Indeed, the only non-trivial assertion is that if X has a point over K, then $X \cong \mathbf{P}^1$ over K. Let $p \in X(K)$. Consider the line bundle $L = \mathcal{O}_X(p)$, defined over K. Observe that $L_{\mathbf{C}} \cong \mathcal{O}(1)$. So $H^0(X, L)$ is a two dimensional vector space. Its two sections yield an isomorphism $X \to \mathbf{P}^1$ over K.

Even if X has no K points, there exist degree 2 extensions $K \subset L$ such that X has L points (and hence $X_L \cong \mathbf{P}^1_L$). To see this, consider $M = K_X^\vee$, the dual of the sheaf of Kähler differentials on X. Then $M_{\mathbf{C}} \cong \mathcal{O}(2)$, and hence $H^0(X, M)$ is 3-dimensional vector space. The zero locus of a global section of M is a subscheme of X of length 2. Being of length 2, this subscheme necessarily has a point over a quadratic extension of K.

4.2. **Rational points on curves of genus 1.** Let X be a smooth geometrically connected curve of genus 1 over a number field K. Then $X(K)$ is an abelian group. If it contains a point of infinite order, then it must be infinite. For any given n, there are finitely many points on X of order n. So, there are countably many points finite order on X. Over an uncountable field, we readily conclude that X must have a point of infinite order. Over a countable field, however, this is much less clear. It is not obvious, for example, that X has a point of infinite order over $\overline{\mathbf{Q}}$, and therefore a point of infinite order over any number field. If it does not, then for any number field K, the group $X(K)$ will be finitely generated and torsion, and hence finite.

It turns out that X always has a point over $\overline{\mathbf{Q}}$ of infinite order. As a result, there exists an extension L of K such that $X(L)$ is infinite. See, for example, [5, Theorem 10.1].

4.3. **A hyperbolic orbifold curve with infinitely many rational points.** Let $Z \subset \mathbf{A}^3 - \{0\}$ be defined by $x^2 + y^3 = z^7$. Let $\mathbf{G}_m$ act on Z by $t \cdot (x, y, z) \mapsto (t^{15}x, t^{10}y, t^6 z)$, and let $X = [Z/\mathbf{G}_m]$ be the stacky quotient. Then the canonical bundle of X has degree $1 - 1/2 - 1/3 - 1/7 > 0$, so X is hyperbolic. Nevertheless, we claim that X has infinitely many $\mathbf{Q}$-points.

For any $x \in \mathbf{Q}$ and $y \in \mathbf{Q}$, consider $w = x^2 + y^3$. We want to make w a perfect seventh power. For a prime p, let $\mathrm{val}_p(w)$ be the p-adic valuation of w. Then w is a perfect fifth power if and only if for all p, the integer $\mathrm{val}_p(w)$ is divisible by 7. For all but finitely many p, we have $\mathrm{val}_p(w) = 0$. If p is such that $\mathrm{val}_p(w) \neq 0$, then let $n \in \mathbf{Z}$ be divisible by 6 and congruent to $-\mathrm{val}_p(w)$ modulo 7. Replace x and y by $p^{n/2}x$ and $p^{n/3}y$. Repeat for all primes such that $\mathrm{val}_p(w)$ is non-zero. The result is a $\mathbf{Q}$-point of Z and hence a $\mathbf{Q}$-point of X. In this way, we can produce infinitely many $\mathbf{Q}$-points of X.

4.4. **Hyperbolic orbifold curves have finitely many integral points.** Let K be a number field with ring of integers $\mathcal{O}_K$. Let X be a smooth proper 1-dimensional Deligne–Mumford stack over K with trivial generic stabiliser such that $\deg K_X > 0$. Let $\mathcal{X}$ be a model of X over $\mathcal{O}_K$. Then $\mathcal{X}$ has finitely many $\mathcal{O}_K$-points.

The proof is similar to the proof sketched in Section 3.2. There exists a Riemann surface Y with a finite étale map $Y \to X_{\mathbf{C}}$ (see [1]). This Y must be defined over a finite extension of K. By an argument similar to Lemma 3.2, there exists a finite extension $K \subset L$ such that every $\mathcal{O}_K$-point of $\mathcal{X}$ is the image of an L point of Y. Since $\deg K_X > 0$ and $Y \to X_{\mathcal{C}}$ is étale, we have $\deg K_Y > 0$. By Faltings' theorem, Y has finitely many L points.

4.5. **Triangle groups and Fenchel's conjecture.** Given any integers p, q, r all greater than 1, there exists an n and permutations $\alpha, \beta, \gamma \in S_n$ such that A has order p, and B has order q, and C has order r, with $ABC = 1$. This is proved in [4] (also see [2] for a correction in the proof). Re-embed S_n in $S_{n!}$ as in the proof of Cayley's theorem—g maps to the permutation given by right-multiplication by g. Let α, β, γ be images of A, B, C, respectively. Then α, β, γ are products of disjoint p, q, r cycles, respectively, and $\alpha\beta\gamma = 1$.

References

[1] K. Behrend and B. Noohi. Uniformization of Deligne-Mumford curves. *J. Reine Angew. Math.*, 599:111–153, 2006.

[2] T. C. Chau. A note concerning Fox's paper on Fenchel's conjecture. *Proc. Am. Math. Soc.*, 88:584–586, 1983.

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

[4] R. H. Fox. On Fenchel's conjecture about F-groups. *Mat. Tidsskr. B*, 1952:61–65, 1952.

[5] G. Frey and M. Jarden. Approximation theory and the rank of Abelian varieties over large algebraic fields. *Proc. Lond. Math. Soc. (3)*, 28:112–128, 1974.

[6] R. Hartshorne. *Algebraic geometry*. Springer-Verlag, New York, 1977. Graduate Texts in Mathematics, No. 52.

[7] M. Hindry and J. H. Silverman. *Diophantine geometry. An introduction*, volume 201 of *Grad. Texts Math.* New York, NY: Springer, 2000.

[8] B. Köck. Belyi's theorem revisited. *Beitr. Algebra Geom.*, 45(1):253–265, 2004.