Chapter 3
The theorems of Thue and Siegel

3.1 Thue's equation

One of the first finiteness results on Diophantine equations was proved by Axel Thue in 1909 [58]. It constitutes the starting point of the modern theories of Diophantine equations and Diophantine approximation.

Theorem 3.1.1 (Thue, 1909)**.** *Let $F(X,Y) \in \mathbb{Z}[X,Y]$ be a homogeneous irreducible polynomial of degree ≥ 3. Let $c \in \mathbb{Z}$ be a non-zero integer. The diophantine equation*

$$F(x,y) = c \tag{3.1.2}$$

has only finitely many solutions in integers $(x,y) \in \mathbb{Z}^2$.

We provide two proofs of this theorem, the second of which uses Siegel's theorem for open sets of $\mathbb{P}_1$, i.e. Corollary 3.2.4 from next section.

Proof. Our first proof follows Thue's original path. Let us suppose by contradiction that $n \mapsto (x_n, y_n)$ is an infinite sequence of integral solutions to (3.1.2), with $|y_n| \to \infty$ (it is clear that there are only finitely many solutions for each given y). We factor the form $F(X,Y)$ in $\bar{\mathbb{Q}}[X,Y]$ by writing

$$F(X,Y) = \prod_{i=1}^{d} (\beta_i X - \alpha_i Y),$$

where $d = \deg F$ and $(\alpha_i, \beta_i) \in \bar{\mathbb{Q}}^2$ are such that $F(\alpha_i, \beta_i) = 0$. Since $F(X,Y)$ is irreducible (over the rationals), the determinants $\alpha_i\beta_j - \beta_i\alpha_j$ do not vanish for any $i \neq j$. Also $\beta_1, \ldots, \beta_d$ are all non-zero. From the equation (3.1.2) we obtain, by taking absolute values,

$$\prod_{i=1}^{d} \left| \beta_i \frac{x_n}{y_n} - \alpha_i \right| = \frac{|c|}{|y_n|^d} \to 0.$$

P. Corvaja, *Integral Points on Algebraic Varieties*, HBA Lecture Notes in Mathematics, DOI 10.1007/978-981-10-2648-5_3

Then, up to extracting a subsequence from the sequence $n \mapsto y_n$ and reordering indices, we can suppose that the sequence of rational numbers $n \mapsto (x_n/y_n)$ tends to α_1/β_1. From the above relation we also obtain the inequality

$$\left| \frac{x_n}{y_n} - \frac{\alpha_1}{\beta_1} \right| \leq \frac{c_1}{|y_n|^d},$$

holding for all large n in an infinite subsequence, where c_1 is any number larger than $|c\beta_1^{d-1}| \max_i(|\alpha_1\beta_i - \alpha_i\beta_1|^{1-d})$. The above inequality contradicts Roth's Theorem, since $d > 2$, finishing the proof. □

As promised, we give a second proof of Thue's theorem.

Proof. Consider the algebraic curve $\mathcal{C} \subset \mathbb{A}^2$ defined by Thue's equation $F(X, Y) = c$. Let $U \subset \mathbb{P}_1$ be the open set $F(X, Y) \neq 0$. Then U is the complement of $d \geq 3$ points in $\mathbb{P}_1$. The map $\mathcal{C} \to U$ sending $\mathcal{C} \ni (x, y) \mapsto (x : y) \in U$ is a (unramified) cover of U, so if $\mathcal{C}$ had infinitely many integral points the same would be true of U, by Chevalley-Weil. An application of Theorem 3.2.4 gives the desired finiteness. □

Some remarks are in order:

(1) Thue did not use Roth's Theorem, which was not yet known at the time, but he used instead a weaker version that he proved in the same article; it is the lower bound

$$\left| \alpha - \frac{p}{q} \right| > \max(|p|, |q|)^{-\frac{d}{2}-1-\epsilon},$$

where $d = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ and $\epsilon > 0$, holding for all but finitely many rational numbers p/q.

(2) The same proof, using Roth's Theorem, applies without changes to prove the finiteness of integral solutions to equations of the form

$$F(x, y) = g(x, y)$$

where $F(X, Y)$ is an irreducible form and the total degree of the polynomial $g(X, Y)$ satisfies $\deg g < \deg F - 2$.

(3) All curves defined by a Thue's equation of the form (3.1.2) have d (distinct) points at infinity; if $d \geq 3$, and only in this case, they have non-zero genus, in other words they are not rational (see below). In contrast, when $d = 2$, the conic of equation $F(X, Y) = 0$ has two points at infinity, and has genus zero. The example of Pell's equation $x^2 - ay^2 = 1$, where $a > 0$ is a positive non-square integer, shows that the assumption that $d \geq 3$ cannot be omitted.

(4) Replacing Roth's Theorem by its generalized version, e.g. Theorem 2.1.8, one can deduce in the same way the more general

Theorem 3.1.3 (Thue-Mahler Theorem)**.** *Let κ be a number field, $\mathcal{O}_S \subset \kappa$ a ring of S-integers, $F(X, Y) \in \mathcal{O}_S[X, Y]$ be a binary homogeneous form with*

S-integral coefficients. Suppose that $F(X,Y)$ has at least three pairwise non-proportional linear factors in $\bar{\kappa}[X,Y]$. Then there are only finitely many pairs $(x,y) \in \mathcal{O}_S^2$, up to multiplicative constants, such that

$$F(x,y) \in \mathcal{O}_S^*. \tag{3.1.4}$$

Let us sketch an independent proof of the Thue-Mahler Theorem, which does not use directly Diophantine approximations methods, but rather the S-unit equation theorem in two variables.

It runs as follows: after factoring

$$F(X,Y) = \prod_{i=1}^{k} (\beta_i X - \alpha_i Y)^{e_i} \tag{3.1.5}$$

where $\beta_i X - \alpha_i Y$ for $i = 1, \ldots, k$ are the distinct prime divisors of $F(X,Y)$ in $\bar{\mathbb{Q}}[X,Y]$, we can suppose, after enlarging κ and S if necessary, that the β_i, α_i belong to κ and the determinants $\beta_i \alpha_j - \beta_j \alpha_i$ are S-units. Then for every coprime S-integers x, y, the values $\beta_i x - \alpha_i y$, for $i = 1, \ldots, k$, are pairwise coprime; if (x,y) is a solution to (3.1.4), the product of the $\beta_i x - \alpha_i y$ is a unit, so each term is a unit. Let us write

$$u_i = \beta_i x - \alpha_i y,$$

for $i = 1, \ldots, k$; since by our hypothesis $k \geq 3$, we can consider the first three terms u_1, u_2, u_3. Eliminating x and y from the relations above we obtain a linear relation of the form $a_1 u_1 + a_2 u_2 + a_3 u_3 = 0$, for some non-zero constant coefficients a_1, a_2, a_3, holding for all the solutions (x,y). An application of the S-unit equation theorem (Theorem 3.2.1) gives the desired result. □

It is worthwhile to look for a geometric interpretation of the last proof; it will turn out that this is precisely the second proof of Thue's theorem given above.

We can view the solutions (x,y) to (3.1.4) as integral points on $\mathbb{A}^2$, which moreover are integral with respect to the curve of equation $F(x,y) = 0$ in $\mathbb{A}^2$. The latter is a union of k lines intersecting at the origin. Viewing the point $(x,y) \in \mathbb{A}^2$ as a point $(x : y : 1) \in \mathbb{P}_2$, it becomes integral also with respect to the line at infinity. Hence we are considering integral points in $\mathbb{P}_2$ with respect to a configuration of $k+1$ lines, the first k passing through a single point and the last one, the line at infinity, being in general position with respect to the previous k. This variety is isomorphic to the product $\mathbb{A}^1 \times (\mathbb{P}_1 \setminus \{k \text{ points}\})$, the projection on the last factor being given by $(x : y : 1) \mapsto (x : y)$. Hence its points are degenerate and moreover they lie on finitely many lines $x = \lambda_i y$; this gives the required finiteness statement.

3.2 Hyperelliptic curves and sums of two units

The aim of this section is proving the following two theorems and showing their interdependence.

Theorem 3.2.1 (S-unit Equation Theorem in two variables)**.** *Let $\Gamma \subset \bar{\mathbb{Q}}^*$ be a finitely generated multiplicative group. Then the equation*

$$u + v = 1$$

has only finitely many solutions $(u, v) \in \Gamma \times \Gamma$.

This theorem is indeed equivalent to Siegel's Theorem 3.3.1 for the particular curve $\mathcal{C} = \mathbb{P}_1 \setminus \{0, 1, \infty\}$, as explained in §1.2. The following result is Siegel's theorem in the case of the so-called hyperelliptic curves, where the points (or the point) at infinity are fixed by the hyperelliptic involution.

Theorem 3.2.2. *Let $\mathcal{O}_S$ be a ring of S-integers in a number field κ; let $f(X) \in \mathcal{O}_S[X]$ be a polynomial with at least three simple roots in $\bar{\kappa}$. Then the equation*

$$y^2 = f(x) \tag{3.2.3}$$

has only finitely many solutions $(x, y) \in \mathcal{O}_S \times \kappa$.

In other words, there exist only finitely many $x \in \mathcal{O}_S$ such that the value $f(x)$ is a square in κ. Note that if $f(X)$ has two roots, the conclusion does not hold in general, as the example of the polynomial $f(X) = 2X^2 + 1$ shows already for $\kappa = \mathbb{Q}$ and $\mathcal{O}_S = \mathbb{Z}$. Also, if $f(X)$ has two simple roots, and no other root in $\bar{\kappa}$, then the curve of equation $y^2 = f(x)$ is rational, isomorphic to $\mathbb{G}_m$ over a suitable extension of the number field κ. Hence it contains infinitely many integral points, over a suitable extension of the ring of integers $\mathcal{O}_S$.

Let us prove Theorem 3.2.1 by using Corollary 2.1.10. First we can find a number field κ and a ring of S-integers $\mathcal{O}_S \subset \kappa$ such that $\Gamma \subset \mathcal{O}_S^*$. Suppose by contradiction that there exist infinitely many solutions (u, v) to the equation $u + v = 1$ of the Theorem. By symmetry, we can suppose that for all our solutions $H(v) \geq H(u)$. Now, for each solution (u, v), let $T = T(u, v)$ be the set of places $\nu \in S$ such that $|v|_\nu < 1$. Since S is finite, there are only finitely many possibilities for the subset T. So, after extracting a suitable infinite subsequence, we can and shall suppose that T is fixed. Putting $\gamma := -u/v$ we obtain $\gamma - 1 = -v^{-1}$ so

$$\prod_{\nu \in T} |\gamma - 1|_\nu = \prod_{\nu \in T} \frac{1}{|v|_\nu} = \prod_{\nu \in S} \max(1, |v^{-1}|_\nu) = H(v^{-1})^{-1} = H(v)^{-1},$$

where the last equality follows from the product formula. Since $H(\gamma) \leq H(u) \cdot H(v) \leq H(v)^2$, we obtain

$$\prod_{\nu \in T} |\gamma - 1|_\nu \leq H(\gamma)^{-1/2}.$$

Hence Corollary 2.1.10, applied with any number $\epsilon < 1/2$, gives the desired contradiction, finishing the proof. Note that by inserting on T also the places for which u is small, we could have ended with the inequality $\prod_{\nu\in T} |\gamma - 1|_\nu \ll H(\gamma)^{-1}$, so a much weaker result than Corollary 2.1.10 would suffice.

Let us now prove Theorem 3.2.2 by using Theorem 3.2.1. Of course, if we prove finiteness of solutions of (3.2.3) for x in a ring larger than $\mathcal{O}_S$, our theorem will be proved. Hence we can enlarge the number field κ so that the roots of $f(X)$ become rational and we also enlarge the ring of S-integers $\mathcal{O}_S$ so that it becomes a Principal Ideal Domain. Now equation (3.2.3) can be written as

$$y^2 = a \cdot (x - \alpha_1)(x - \alpha_2)(x - \alpha_3)g(x)$$

where $a, \alpha_1, \alpha_2, \alpha_3 \in \kappa$, $a \neq 0$, the α_i are pairwise distinct and the polynomial $g(X) \in \kappa[X]$ does not vanish at α_i for any $i = 1, 2, 3$. Since the two polynomials $h(X) := a(X-\alpha_1)(X-\alpha_2)(X-\alpha_3)$ and $g(X)$ are coprime, in the ring $\kappa[X]$, they generate the unit ideal; in other words, there exist polynomials $\varphi(X), \psi(X) \in \kappa[x]$ such that

$$\varphi(X)h(X) + \psi(X)g(X) = 1.$$

Up to enlarging if necessary the ring of S-integers $\mathcal{O}_S$, we can suppose that $a \in \mathcal{O}_S^*$ and that all the coefficients of all the four polynomials $h(X), g(X), \varphi(X), \psi(X)$ are S-integers. Now, for every $x \in \mathcal{O}_S$, from the above identity it follows that the two S-integers $h(x), g(x)$ generate the unit ideal $(1) = \mathcal{O}_S$, so they must be coprime. Hence, due to unique factorization in $\mathcal{O}_S$, whenever the product $f(x) = h(x)g(x)$ is a square in κ, both factors should be squares, up to units. Since the quotient of the group of units modulo squares is finite, we obtain from the infinitude of the set of solutions to (3.2.3) that for at least one unit $\gamma \in \mathcal{O}_S^*$ the equation

$$y^2 = \gamma(x - \alpha_1)(x - \alpha_2)(x - \alpha_3)$$

has infinitely many solutions $(x, y) \in \mathcal{O}_S \times \kappa$. We observe that $x - \alpha_i, x - \alpha_j$ are essentially coprime for $i \neq j$, actually they are coprime whenever the discriminant $((\alpha_1 - \alpha_2)(\alpha_2 - \alpha_3)(\alpha_3 - \alpha_1))^2$ is a unit, which we can suppose to hold after enlarging S; so, by repeating the previous argument, we deduce that there exist units $\gamma_1, \gamma_2, \gamma_3$ such that for infinitely many $x \in \mathcal{O}_S$ and each $i = 1, 2, 3$, the elements $\gamma_i(x - \alpha_i)$ are squares in κ. After enlarging κ we can suppose that the γ_i are also squares, so that for infinitely many $x \in \mathcal{O}_S$ there exist $y_1, y_2, y_3 \in \kappa$ such that we can write

$$y_i^2 = x - \alpha_i \qquad \text{for } i = 1, 2, 3.$$

Eliminating x from the first two relations (i.e. those corresponding to $i = 1, 2$) we obtain

$$y_1^2 - y_2^2 = (y_1 - y_2)(y_1 + y_2) = \alpha_2 - \alpha_1$$

Recall that $\alpha_2 - \alpha_1$ is a unit, since the discriminant of $h(X)$ was supposed to be a unit; then $y_1 - y_2$ (and also $y_1 + y_2$) must be a unit. Then, using the same

relations for the other pairs of distinct indices i, j in $\{1, 2, 3\}$, we obtain that $y_1 - y_3$ and $y_2 - y_3$ are also unit. Writing

$$\begin{cases} y_1 - y_2 = u_3 \\ y_2 - y_3 = u_1 \\ y_3 - y_1 = u_2 \end{cases}$$

for suitable units u_1, u_2, u_3, we obtain the homogeneous S-unit equation

$$u_1 + u_2 + u_3 = 0.$$

Putting $u := -u_1/u_3$, $v := -u_2/u_3$, we get the relation $u + v = 1$ and Theorem 3.2.1 gives the finiteness of the ratios $u_1/u_3, u_2/u_3$. This in turn gives the finiteness of the triples (y_1, y_2, y_3) up to multiplicative constants and from this and the relation $y_1^2 - y_2^2 = \alpha_2 - \alpha_1$ it is easy to deduce the finiteness of the solutions x.

The above proof might seem complicated and unnatural, but it can be enlightened using a geometric view-point. Let $\mathcal{C}$ be the affine curve defined by our equation (3.2.3). The main point of the proof is the observation that the three rational functions $x - \alpha_i$, for $i = 1, 2, 3$, take perfect square values at integral points (x, y) (after a fixed enlargement of the ring $\mathcal{O}_S$). This is of course connected with the Chevalley-Weil theorem: the three functions in question are locally a square everywhere, so the function field extension $\kappa(\mathcal{C})(\sqrt{x - \alpha_i})/\kappa(\mathcal{C})$ is unramified over $\mathcal{C}$ (it might ramify at infinity, depending on the parity of $\deg f$). Hence, by Chevalley-Weil, each integral point on $\mathcal{C}$ lifts to an integral point on $\mathcal{C}'$, where $\mathcal{C}'$ is the affine curve corresponding to the integral closure of the ring $\kappa[\mathcal{C}][\sqrt{x - \alpha_1}, \sqrt{x - \alpha_2}, \sqrt{x - \alpha_3}]$. Now, letting y_i be square roots of $x - \alpha_i$ in $\kappa[\mathcal{C}']$, we have that the regular functions $u_3 := y_1 - y_2, u_1 := y_2 - y_3, u_2 := y_3 - y_2$ have all their zeros and poles at infinity, so they send $\mathcal{C}' \to \mathbb{G}_m$. So we obtain the morphism $\mathcal{C}' \to \mathbb{G}_m^2$ by sending $\mathcal{C}' \ni p \mapsto (\frac{-u_1(p)}{u_3(p)}, \frac{-u_2(p)}{u_3(p)})$ whose image is the line $u + v = 1$ inside the torus $\mathbb{G}_m^2$; this closed set of the torus is isomorphic to $\mathbb{P}_1 \setminus \{0, 1, \infty\}$, so it contains only finitely many integral points and the proof is finished.

It could be proved that for a general affine hyperelliptic curve $\mathcal{C}$ defined by an equation (3.2.3) there exists no non-constant morphism $\mathcal{C} \to \mathbb{G}_m^2$, and, when such a morphism does exist, its image is a translate of a subtorus; so Theorem 3.2.1 or its generalization Theorem 1.2.2 cannot be used directly. However, the two-variables S-unit equation theorem can be used, and has been used in the above proof, after taking an unramified cover $\mathcal{C}' \to \mathcal{C}$ of the original curve, since the curve $\mathcal{C}'$ does admit such a non-trivial morphism to $\mathbb{G}_m^2$.

We end by showing that Theorem 3.2.1 immediately implies Siegel's theorem in the rational case:

Corollary 3.2.4. *Let $X \subset \mathbb{P}_1$ be an algebraic open set with $\mathbb{P}_1 \setminus X$ consisting of at least three points. Then $X(\mathcal{O}_S)$ is finite, for every ring of S-integers $\mathcal{O}_S$.*

Proof. We just repeat the argument given in the introduction. We can suppose, up to enlarging the field of definition κ, that three of the points of the complement of X in $\mathbb{P}_1$ are $0, 1, \infty$. Then the algebra $\kappa[X]$ contains the functions $x, 1/x, 1/(x-1)$. Hence for every integral point $p \in X(\mathcal{O}_S)$ we obtain a solution $(u, v) = (x(P), 1 - x(P))$ to the S-equation $u + v = 1$ of Theorem 3.2.1. □

3.3 Siegel's Theorem on curves

As mentioned, a general theorem of Siegel-Mahler, which we give here in the most general formulation provided by Lang [39], asserts the finiteness of S-integral points on a large class of curves, namely all those satisfying the assumption of log-general type appearing in Vojta's conjecture:

Theorem 3.3.1 (Siegel's Theorem on curves)**.** *Let $\mathcal{C}$ be an affine curve defined over a number field. Suppose either that it has genus > 0 or that it has at least three points at infinity. Then for each ring of S-integers $\mathcal{O}_S$, the set $\mathcal{C}(\mathcal{O}_S)$ is finite.*

Note that we do not assume the curve is smooth; however, the theorem in the possibly singular case would follow easily from the particular case of smooth curves.

In the sequel, we shall suppose that $\mathcal{C}$ is smooth affine and define $\tilde{\mathcal{C}}$ to be its smooth compactification; let g be the genus of $\tilde{\mathcal{C}}$. We denote by D the complement: $D = \tilde{\mathcal{C}} \setminus \mathcal{C}$.

So Siegel's theorem asserts that whenever $\deg D \geq 3$ or $g \geq 1$, then the set $\mathcal{C}(\mathcal{O}_S)$ is finite. On the other hand, we have already observed in Chapter 1 that if $g = 0$ and $\deg D = 1$ or 2 the corresponding curve, which is either $\mathbb{G}_a = \mathbb{A}^1$ or $\mathbb{G}_m = \mathbb{A}^1 \setminus \{0\}$ has infinitely many integral points (over a suitable ring of S-integers). Hence Siegel's theorem provides a complete classification of the algebraic curves admitting infinitely many integral points.

Let us analyze this classification in view of the Chevalley-Weil theorem. Recall that given two (smooth, affine) curves $\mathcal{C}_1, \mathcal{C}_2$ admitting a dominant morphism $\pi : \mathcal{C}_1 \to \mathcal{C}_2$, if $\mathcal{C}_1(\mathcal{O}_S)$ is infinite, also $\mathcal{C}_2(\mathcal{O}_S)$ will be infinite. On the other hand, if $\pi : \mathcal{C}_1 \to \mathcal{C}_2$ is an unramified cover, then the two finiteness properties are equivalent; more precisely, if $\mathcal{C}_2(\mathcal{O}_S)$ is finite for every ring of S-integers $\mathcal{O}_S$, the same is true of $\mathcal{C}_1(\mathcal{O}_S)$. Let us write, as usual, $\mathcal{C}_i = \tilde{\mathcal{C}}_i \setminus D_i$ and denote by g_i the genus of $\tilde{\mathcal{C}}_i$ and by $s_i = \deg D_i$. The inequalities $g_1 \geq g_2$ and $s_1 \geq s_2$ hold for every dominant morphism $\pi : \mathcal{C}_1 \to \mathcal{C}_2$; moreover, for unramified morphism the equality holds if either $\deg \pi = 1$ (which is certainly the case if $\mathcal{C}_1 = \mathbb{A}^1$) or $\mathcal{C}_2 = \mathbb{G}_m$ (in which case necessarily $\mathcal{C}_1 = \mathbb{G}_m$). Hence, Siegel's finiteness theorem can be stated as follows:

Theorem 3.3.2 (Siegel's Theorem - alternate version)**.** *Let $\mathcal{C}$ be a smooth affine curve defined over a number field κ. The following are equivalent:*

(i) *the set $\mathcal{C}(\mathcal{O}_S)$ is finite for every ring of S-integers;*
(ii) *there exists an unramified cover $\mathcal{C}' \to \mathcal{C}$ of $\mathcal{C}$ such that the genus of $\mathcal{C}'$ is strictly larger than the genus of $\mathcal{C}$;*

(iii) *for every integer g, there exists an unramified cover $\mathcal{C}' \to \mathcal{C}$ of $\mathcal{C}$ such that the genus of $\mathcal{C}'$ is larger than g;*
(iv) *there exists an unramified cover $\mathcal{C}' \to \mathcal{C}$ of $\mathcal{C}$ such that $\mathcal{C}'$ has strictly more points at infinity than $\mathcal{C}$;*
(v) *for every integer N there exists an unramified cover $\mathcal{C}' \to \mathcal{C}$ of $\mathcal{C}$ such that $\mathcal{C}'$ has at least N points at infinity;*
(vi) *the fundamental group of the topological space $\mathcal{C}(\mathbb{C})$ is not abelian.*

By Chevalley-Weil Theorem and topological classification of algebraic curves, this means that an apparently weaker statement than Siegel's, namely the finiteness of integral points on curves of sufficiently large genus (say genus > 100) would imply via Chevalley-Weil theorem the full statement in Theorem 3.3.1. The same could be said about the requirement on the number of points at infinity: the finiteness of integral points on all the curves with at least, say, one hundred points at infinity, over every ring of S-integers, would imply the finiteness of integral points on curves with at least three points at infinity, as well as on those which have positive genus and at least one point at infinity (i.e. are affine).

Siegel's proofs. We give a sketch of a proof of Siegels' Theorem similar to the original one (but we should mention that Siegel did not treat arbitrary S-integers in his 1929 paper [52]; the generalization to arbitrary S-integers is due to Mahler and Lang, see [39]). Actually, Siegel provided two different proofs; we recommend the paper [63], which we are following now, for a careful discussion of the different tools needed in the various approaches.

Let us suppose that an affine curve $\mathcal{C}$ (say embedded in $\mathbb{A}^N$) of genus > 0 admits infinitely many S-integral points. Then we can extract an infinite sequence $P_1, P_2, \dots$ in $\mathcal{C}(\mathcal{O}_S)$ converging in the projective completion $\tilde{\mathcal{C}}$ for every place $\nu \in S$ (recall that $\tilde{\mathcal{C}}(\kappa_\nu)$ is compact).

For each point $P \in \mathbb{A}^N(\kappa)$ and each place $\nu \in S$, denote by $|P|_\nu$ the sup-norm of P. Then the height of an S-integer point $P \in \mathbb{A}^N(\mathcal{O}_S)$ is

$$H(P) = \prod_{\nu \in S} \max\{1, |P|_\nu\},$$

so $H(P) \leq \max\{1, |P|_{\nu_0}\}^{\sharp(S)}$ where ν_0 is such that $|P|_{\nu_0} \geq |P|_\nu$ for any other place ν. We can suppose that for our points in the sequence $P_1, P_2, \dots$ the place ν_0 is one and the same. Let $Q = \lim_{n\to\infty} P_n$, the limit being taken in the ν_0-adic sense. Then for a suitable local parameter $t \in \kappa(\mathcal{C})$ at Q and a positive real number δ, we shall have

$$|t(P_n)| =: \operatorname{dist}(P_n, Q) \ll |P_n|^{-\delta} \ll H(P_n)^{-\delta/\sharp(S)}. \tag{3.3.3}$$

If $\delta > 2\sharp(S)$, a direct application of Roth's Theorem would be sufficient to conclude. If, however, that $\delta < 2\sharp(S)$, inequality 3.3.3 would not suffice. Siegel's trick to overcome this difficulty consists in taking an unramified covering $\tilde{\mathcal{C}}' \to \tilde{\mathcal{C}}$ of $\mathcal{C}$. By Chevalley-Weil theorem, the integral points P_n lift to integral points $P'_n \in \mathcal{C}'_i(\mathcal{O}_S)$, in one of the finitely many twists of $\mathcal{C}'_i$ of $\mathcal{C}'$. We can suppose,

since we may dispose of infinitely many integral points, that all of them lift to integral points on a same curve $\mathcal{C}'$. The rate of convergence at infinity of the P'_n is the same, since the given covering is unramified (even at infinity). On the other hand, the height of the new points P'_n is smaller then that of the P_n by a factor equal to the degree of the cover. Working with a cover of degree $> 2\sharp(S)/\delta$ one is in the situation of applying Roth's theorem and we may conclude.

In connection with Siegels' Theorem, we end this section by showing that the results about Thue-Mahler and hyperelliptic equations fit into this frame.

Let us start by proving that the algebraic curve defined by Thue's equation (3.1.2) is non-rational, as soon as the hypotheses appearing in the statement of Thue's Theorem are satisfied: this is the content of the following

Theorem 3.3.4. *Given a homogeneous form $F(X,Y) \in \mathbb{C}[X,Y]$ of degree $\deg F = d \geq 3$, with no repeated linear factors, for every non-zero complex number $c \in \mathbb{C}^*$ the equation*

$$F(x(t), y(t)) = c$$

has no solutions $(x(t), y(t)) \in \mathbb{C}(t)^2$ in non-constant rational functions.

Proof. Homogenizing, we are reduced to showing that the homogeneous equation

$$F(X,Y) = cZ^d$$

has no non-constant solutions in *coprime polynomials* $x(t), y(t), z(t) \in \mathbb{C}[t]$. Factoring the homogeneous form as in (3.1.5) and dividing all factors by $y(t)$ we obtain

$$\prod_{i=1}^{d} \left(\frac{x(t)}{y(t)} - \frac{\alpha_i}{\beta_i} \right) = C \left(\frac{z(t)}{y(t)} \right)^d,$$

where $C = \frac{c}{\beta_1 \cdots \beta_d}$. Here we are assuming that all β_i are non-zero, but the proof would not be really different if (at most) one β_i vanishes. As mentioned, the points $\gamma_i := \alpha_i/\beta_i$, $i = 1, \ldots, d$, are pairwise distinct. Each time the rational function $f(t) := x(t)/y(t)$ takes one of these values, the function $z(t)/y(t)$ takes the value zero. Since $\deg(z(t)/y(t)) \leq \deg f$, the cardinality of the set $f^{-1}(\{\gamma_1, \ldots, \gamma_d\})$ cannot exceed $\deg f$; on the other hand, the pre-image of a set of cardinality d has at least $d \deg(f) - R$ points, where R is the degree of the ramification divisor of f; the latter is equal to $2\deg(f) - 2$ by Riemann-Hurwitz formula or direct computation. Hence $(d-3)\deg f + 2 \leq 0$ from which it follows that $d \leq 2$, finishing the proof. □

It is also easy to see that the number of points at infinity is precisely d; so the curves defined by Thue's equations have two good reasons for the set of their integral points to be finite.

We now consider the geometry of the algebraic curve defined by the hyperelliptic equation

$$y^2 = f(x), \tag{3.3.5}$$

where $f(X) \in \mathbb{C}[X]$ is a polynomial with no repeated factors. The above equation defines a smooth affine curve in the plane $\mathbb{A}^2$; however, whenever $\deg f \geq 4$ its natural completion in $\mathbb{P}_2$ turns out to be singular at its only point at infinity; its desingularization has two points at infinity. Let us denote by $\tilde{\mathcal{C}}$ this smooth projective model.

Theorem 3.3.6. *Let $f(X) \in \mathbb{C}[X]$ be, as before, a non-constant polynomial without repeated roots and let $\tilde{\mathcal{C}}$ be a smooth complete model of the affine curve defined by the above equation* (3.3.5)*. If $\deg f \geq 3$, then $\tilde{\mathcal{C}}$ is non-rational.*

Proof. One could apply the well-known genus formula to prove that the genus of $\tilde{\mathcal{C}}$ is $\frac{d}{2} - 1$ if $d = \deg f$ is even, $\frac{d-1}{2}$ if d is odd: hence it is > 0 whenever $d \geq 3$. Nevertheless, we prefer a proof which is closer in spirit to our proof of the finiteness of integral solutions. We exibit a non-zero class in $\mathrm{H}^1(\tilde{\mathcal{C}}, \{\pm 1\})$, recalling that this group is isomorphic to the quotient

$$\{f \in \mathbb{C}(\tilde{\mathcal{C}})^* \,:\, \mathrm{ord}_p(f) \equiv 0 \pmod 2 \quad \forall p \in \tilde{\mathcal{C}}\}/\{f^2 \,:\, f \in \mathbb{C}(\tilde{\mathcal{C}})^*\}.$$

In fact, supposing for simplicity $d \equiv 0 \pmod 2$, $d \geq 4$, and writing $f(X) = a(X - \alpha_1) \cdots (X - \alpha_d)$, for complex numbers $\alpha_1, \ldots, \alpha_d, a \in \mathbb{C}$, $a \neq 0$, we see at once that each rational function $x - \alpha_i$ has a double zero at $(\alpha_i, 0)$. It has a simple pole at each of the two points at infinity; so the product $f = (x-\alpha_1)(x-\alpha_2)$ is a square locally everywhere. Let us show that it is not globally a square in $\mathbb{C}(\tilde{\mathcal{C}})$; if it were so, we would have $\mathbb{C}(\tilde{\mathcal{C}}) = \mathbb{C}(x)(\sqrt{f})$; however, this extension is unramified over $x = \alpha_3$, while the extension $\mathbb{C}(\tilde{\mathcal{C}})/\mathbb{C}(x)$ does ramify over $x = \alpha_3$. □

3.4 A Subspace Theorem approach to Siegel's Theorem

The aim of this section is to provide a complete proof of Siegel's Theorem on curves assuming the Subspace Theorem (in the version given in Theorem 2.2.1).

Let us go back to the proof of Thue's theorem. Recall that the equation under examination was

$$F(x, y) = c,$$

where $F(X, Y) = \prod_{i=1}^{d} (\beta_i X - \alpha_i Y)$, $d \geq 3$, the linear factors are pairwise coprime and $c \neq 0$. Letting $\mathcal{C}$ be the algebraic curve defined by the above equation, the main point of the proof consisted in considering one of the rational functions $\beta_i x - \alpha_i y$ on $\mathcal{C}$, viewed as a morphism $\mathcal{C} \to \mathbb{A}^1$. We can extend it to the complete curve $\tilde{\mathcal{C}}$ (defined by the homogeneous equation $F(X, Y) = cZ^d$) by sending $\tilde{\mathcal{C}} \ni (X : Y : Z) \mapsto (\beta_i X - \alpha_i Y : Z) \in \mathbb{P}_1$. Then we applied Roth's theorem, i.e. a result on Diophantine approximation on the line. The choice of such a rational functions was dictated by the fact that it is regular on $\mathcal{C}$ (i.e. its poles lie at infinity) and vanishes at sufficiently high degree on an accumulation point for an infinite sequence of integral points on $\mathcal{C}$ (supposed to exist).

This strategy does not work in general: for instance, if a curve $\mathcal{C}$ has only one point at infinity, such a point will be an accumulation point for every infinite sequence of integral points on $\mathcal{C}$, and there exist no non-constant regular function on $\mathcal{C}$ vanishing at infinity. Even if there are more points at infinity, it may be that no function with the desired property exists. Let us see a concrete example:

`Example`. Consider the algebraic curve of equation

$$\mathcal{C} : x^3 - 2y^3 = x + y + 1. \tag{3.4.1}$$

Its genus is one, and moreover it has three points at infinity, so by Siegel's theorem it should have only finitely many integral points. Each sequence (x_n, y_n), $n \in \mathbb{N}$ in $\mathcal{C}(\mathbb{Z})$ should converge to the point $A := (\sqrt[3]{2} : 1 : 0) \in \mathbb{P}_2$ (considering the natural compactification $\tilde{\mathcal{C}}$ of $\mathcal{C}$ given by the equation $X^3 - 2Y^3 = Z^2(X + Y) + Z^3$). The other two points at infinity are $B := (\zeta\sqrt[3]{2} : 1 : 0)$ and $\bar{B} = (\bar{\zeta}\sqrt[3]{2} : 1 : 0)$, where ζ is a primitive third root of unity. Every regular function $f \in \kappa[\mathcal{C}]$ is a polynomial function of $x = X/Z$, $y = Y/Z$. If $\kappa = \mathbb{Q}$, then, since $A, B, \bar{B}$ are Galois-conjugated over $\mathbb{Q}$, such a function must have poles at each of the three points or be constant. However, working over the cubic field $\kappa = \mathbb{Q}(\sqrt[3]{2})$ we can find a function having a zero at A, for instance the function $x + \sqrt[3]{2}y$. Now from the equation (3.4.1) we deduce that

$$(x - \sqrt[3]{2}y) = \frac{x + y + 1}{x^2 + \sqrt[3]{4}xy + y^2}.$$

When the pair (x, y) tends to infinity (i.e. to A) on the curve $\mathcal{C}$ the asymptotic estimations $|x + y + 1| \gg \max(|x|, |y|) = |x|$ and $|x^2 + \sqrt[3]{4}xy + y^2| \ll \max(|x|, |y|) = x^2$ hold. Hence the left hand side tends to zero asymptotically as x^{-1}, not faster; dividing by y one obtains $|x/y - \sqrt[3]{2}| \ll H(x/y)^{-2}$ which is not sufficient to deduce a contradiction via Roth's theorem.

We can, however, try to consider more functions $f_1, \ldots, f_r \in \mathbb{Q}(\sqrt[3]{2})[\mathcal{C}]$, giving rise to a morphism $\mathcal{C} \to \mathbb{A}^r$, and then try to apply Diophantine approximation results in the larger space $\mathbb{A}^r$, like the Subspace Theorem.

Let us now give the details, following [15]. Precisely, we want to prove the following

Theorem 3.4.2. *Let $\mathcal{C}$ be a smooth affine curve with $r \geq 3$ points at infinity, defined over a number field κ. Then for every ring of S-integers $\mathcal{O}_S \subset \kappa$, the set $\mathcal{C}(\mathcal{O}_S)$ is finite.*

The full Siegel's theorem then follows by applying Chevalley-Weil theorem.

Proof. Let $Q_1, \ldots, Q_r$ be the points (valuations) at infinity of the curve $\mathcal{C}$. They are defined over a finite extension of κ. For a large integer N put

$$V_N = \mathrm{H}^0(\tilde{\mathcal{C}}, N(Q_1 + \ldots + Q_r)) = \{f \in \bar{\kappa}[\mathcal{C}] \,:\, (f) \geq -N(Q_1 + \ldots + Q_r)\}.$$

Let $f_1, \dots, f_d$, where $d = h^0(N(Q_1 + \dots + Q_r)) = rN + O(1)$, be a basis of V_N. Since the divisor $Q_1 + \dots + Q_r$ is defined over κ, we can choose $f_1, \dots, f_d$ defined over κ, i.e. with $f_i \in V_N \cap \kappa[\mathcal{C}]$ for $i = 1, \dots, d$.

As in the previous sketch of the proof, if $\mathcal{C}(\mathcal{O}_S)$ is infinite, we can find a sequence $P_1, P_2, \dots$ of integral points in $\mathcal{C}(\mathcal{O}_S)$ such that for each place $\nu \in S$ the sequence converges to a point $R_\nu \in \tilde{\mathcal{C}}(\kappa_\nu)$. We let S' to be the set of places for which the limit R_ν lies at infinity.

After multiplying the f_j by a suitable constant, we can suppose that $f_j(P_n) \in \mathcal{O}_S$ for all j, n.

For every $\nu \in S$, consider the filtration $V = W_{\nu,1} \supset W_{\nu,2} \supset \dots$ defined as

$$W_j = W_{\nu,j} = \{f \in V_N \,:\, \mathrm{ord}_{R_\nu} f \geq j - 1 - N\}.$$

We have $\dim(W_j/W_{j+1}) \leq 1$ for each j; in particular $\dim W_j \geq d - j + 1$.

Now, for each $\nu \in S'$, choose a basis of V_N containing a basis of each subspace $W_{\nu,j}$ (for each j such that $W_{\nu,j} \neq \{0\}$). These functions can be expressed as linear combinations of the basis $(f_1, \dots, f_d)$, i.e. as values of linear forms $L_{\nu,j}(f_1, \dots, f_d)$, where $L_{\mu,j}(X_1, \dots, X_d)$ has its coefficients in $\bar{\kappa}$. Clearly

$$\mathrm{ord}_{R_\nu} L_{\nu,j}(f_1, \dots, f_d) \geq j - N + 1.$$

For $\nu \in S \setminus S'$ we just put $L_{\nu,j}(f_1, \dots, f_d) = f_j$.

For each $\nu \in S'$ choose a local parameter $t_\nu \in \kappa(\mathcal{C})$ at R_ν. The above displayed inequality implies that

$$|L_{\nu,j}(f_1(P_n), \dots, f_d(P_n))|_\nu \ll |t_\nu(P_n)|_\nu^{j-1+N}.$$

Now, observe that we dispose of $d = rN + O(1)$ rational functions $L_{\nu,j}(f_1, \dots, f_d)$, of which at most N have poles and approximately $(r-1)N$ have zeros at R_ν. Estimating the order of the product $\prod_j L_{\nu,j}(f_1, \dots, f_d)$ we have that this order is positive, and actually $> (r-2)N + O(1)$ for large N (a stronger asymptotic estimates in fact holds, but we do not need it).

Put $\mathbf{x} = (f_1(P_n), \dots, f_d(P_n)) \in \mathcal{O}_S^d$ and let as before $|\mathbf{x}|_\nu$ be its sup-norm in the ν-adic absolute value. Observing that for $\nu \notin S'$ the absolute values of $f_j(P_n)$ are uniformly bounded, we can deduce that

$$\prod_{\nu \in S} \prod_{j=1}^{d} \frac{|L_{\nu,j}(\mathbf{x})|_\nu}{|\mathbf{x}|_\nu} \ll \prod_{\nu \in S'} \left(|t_\nu(P_n)|\right)^{(r-2)N}.$$

On the other hand, the height is easily estimated by $H(\mathbf{x}) \ll \prod_{\nu \in S'} \left(|t_\nu(P_n)|\right)^N$. Finally we obtain

$$\prod_{\nu \in S} \prod_{j=1}^{d} \frac{|L_{\nu,j}(\mathbf{x})|_\nu}{|\mathbf{x}|_\nu} \ll H(\mathbf{x})^{2-r}.$$

The Subspace Theorem then implies that infinitely many vectors $\mathbf{x}$ lie on a hyperplane; this is impossible, since the functions $f_1, \dots, f_d$ are linearly inde-

pendent, so every non-trivial linear combination of $f_1, \ldots, f_d$ can have only finitely many zeros.

Another approach to Siegel's theorem on integral points involving non-standard analysis has been proposed by Robinson and Roquette [46]. Their proof implicitly uses Mordell-Weil theorem on the Jacobian of the curve, although it does not mention explicitly Jacobians.

Finally, Gasbarri [33] gave a different proof of Siegel's theorem, which uses ideas coming from the proof of Thue-Siegel-Dyson-Gelfond theorem on Diophantine approximation. Basically, he reproves this approximation theorem for integral points lying on a curve, and deduces a finiteness statement whenever there are three points at infinity. □