THE ABC CONJECTURE IMPLIES ROTH'S THEOREM AND MORDELL'S CONJECTURE

MACHIEL VAN FRANKENHUYSEN

ABSTRACT. We present in a unified way proofs of Roth's theorem and an effective version of Mordell's conjecture, using the ABC conjecture. We also show how certain stronger forms of the ABC conjecture give information about the type of approximation to an algebraic number.

1. INTRODUCTION

In 1991, Noam D. Elkies showed that the ABC conjecture implies Mordell's conjecture [5]. And in 1994, Enrico Bombieri showed that the ABC conjecture implies Roth's theorem about Diophantine approximation of algebraic numbers [3]. The proofs of these two implications are very similar (see §§6.4, 6.7), and in §6.8, we formulate a theorem that implies both Roth's theorem and Mordell's conjecture.

We formulate the ABC conjecture in §2. In §2.4, we introduce the 'type function', which allows us to formulate certain stronger forms of the ABC conjecture. In §4, we formulate Roth's theorem and define the 'type' of an algebraic number, and in §5, we formulate Mordell's conjecture and 'effective Mordell'. §6.3 is devoted to Belyĭ's construction of an algebraic function which is ramified over 0, 1 and ∞ alone [1]. The application of this construction to $\mathbf{P}^1$ yields Roth's theorem, §6.4, and the application to a curve C of genus 2 or higher yields Mordell's conjecture, §6.7.

Both Roth's theorem and Mordell's conjecture are theorems, see [10, 16] and [2,4,6,7,21] respectively, and from this point of view it seems uninteresting to have conditional proofs of these theorems, depending on the ABC conjecture, whose validity is still unknown. However, the proofs of these theorems using ABC are much simpler and transparent, and point out very clearly the relationship between the theory of Diophantine approximation and the theory of points on curves of high genus. More importantly, using ABC, one can prove considerably stronger versions of the two theorems. Specifically, ABC implies *effective* Mordell (see §5.1), and a certain stronger form of the ABC conjecture implies a certain refinement of Roth's theorem (see §4.1).

1991 *Mathematics Subject Classification.* Primary 11D75, 11G30, 11J68; Secondary 14Hxx, 30D35.

Key words and phrases. ABC conjecture, ABC conjecture with type function, Diophantine approximation, Roth's theorem, type of an algebraic number, Mordell's conjecture, effective Mordell.

This is an elaborated version of the talk presented at the XV Escola de Álgebra, held in Canela, Brazil, July 26–August 1. This work was supported by the Marie Curie fellowship ERBFM-BICT960829 of the European Community. We would like to thank the Institut des Hautes Études Scientifiques (IHÉS), of which we were a member during part of this research work. We thank P. Vojta for valuable comments about §6.6.

Regarding this refinement of Roth's theorem, in the sixties S. Lang conjectured that Roth's theorem could be improved to $-\log|\alpha - p/q| - 2\log q \le (1+\varepsilon)\log\log q$, see [11, p. 214]. Indeed, this is supported by the analogous inequality in the case of meromorphic functions, see §3.2. However, the strongest possible form of the ABC conjecture only yields

$$-\log\left|\alpha - \frac{p}{q}\right| - 2\log q \le K\frac{\sqrt{\log q}}{\log\log q}, \tag{1.1}$$

for some constant K depending on α, see Theorem 4.3 and §2.4. Therefore, we want to raise the question whether this is the strongest possible form of Roth's theorem:

> For (some or every?) algebraic number α of degree ≥ 3 over $\mathbf{Q}$, (1.1) cannot be improved; that is, for some value of $K > 0$, the opposite inequality is satisfied for infinitely many p and q.

In terms of the coefficients a_n of the continued fraction expansion of α, this would imply that $\log a_n \ge \kappa\sqrt{n}$ infinitely often, for some $\kappa > 0$. There is some numerical evidence for this question, see [13, 18]. For example, the continued fraction of the real root of $x^3 - 8x - 10 = 0$, discussed in [18], has some very large coefficients. When these computations are pushed further though, this continued fraction seems to behave randomly after the 161st coefficient, with coefficients of order $\log a_n = O(\log n)$. On the other hand, one must be careful in interpreting these data: if the constant κ above is very small, one starts to find large coefficients only for very large values of n. Moreover, large coefficients will be very rare. We will address this question in subsequent work.

Finally, we point out that to simplify this exposition, we have restricted ourselves to the rational numbers. But the ABC conjecture, Mordell's conjecture and Roth's theorem can be formulated for any finite extension of $\mathbf{Q}$, and ABC implies Roth and Mordell in these more general situations as well. The proofs presented here generalize with little modification.

1.1. **Notations and conventions.** Throughout, we use the following notations and letters with a special meaning:

$x \vee y$	denotes the maximum of x and y,
$\#V$	denotes the number of elements of the set V,
p	is a prime number,
v	is a valuation, either v_p or v_∞, see §2.2,
S	is a finite set of valuations,
$w \in S$	valuations in S are denoted by w,
C	is an algebraic curve,
$C(\mathbf{Q})$	denotes the points on C with rational coordinates,
$\mathbf{P}^1$	is the projective line, see §2.3,
$\bar{\mathbf{Q}}$	is the field containing all numbers algebraic over $\mathbf{Q}$.

We think of C as the set of points $(x_0 : \cdots : x_n) \in \mathbf{P}^n$ that satisfy the homogeneous equations

$$p_1(x_0, \ldots, x_n) = 0, \ldots, p_k(x_0, \ldots, x_n) = 0, \tag{1.2}$$

with $k \ge n-1$ and p_i irreducible. The set of complex solutions of these equations is the *Riemann surface* $C(\mathbf{C})$. If the coefficients of $p_1, \ldots, p_k$ lie in $\mathbf{Q}$, we say that

C is *defined over* $\mathbf{Q}$. A map $f\colon C \longrightarrow \mathbf{P}^m$ will be given by $m+1$ homogeneous polynomials of the same degree,

$$f\colon\ (x_0 : \cdots : x_n) \longmapsto (f_0\,(x_0,\ldots,x_n) : \cdots : f_m\,(x_0,\ldots,x_n))\,. \tag{1.3}$$

If the coefficients of $f_0,\ldots,f_m$ lie in $\mathbf{Q}$, we say that f is *defined over* $\mathbf{Q}$.

2. The ABC conjecture

Given a sum $a+b=c$, with $a,b,c \in \mathbf{Z}$, coprime and $a,b,c \neq 0$, we define the **height** and the **radical** of this sum by

$$\begin{aligned} h(a,b,c) &= \max\{\log|a|, \log|b|, \log|c|\}\,, \\ r(a,b,c) &= \sum_{p\mid abc} \log p, \end{aligned} \tag{2.1}$$

where p runs over all prime divisors of a, b and c.

For example,

(2.2)

$a+b=c$	height	radical
$2+3=5$	$\log 5$	$\log 30$
$9+16=25$	$\log 25$	$\log 30$
$3+125=128$	$\log 128$	$\log 30$
$19\cdot 1307 + 7\cdot 29^2\cdot 31^8 = 2^8\cdot 3^{22}\cdot 5^4$	$36.15\ldots$	$22.26\ldots$

We see that in the last two examples, the height is larger than the radical. The ABC conjecture says that the height cannot be much larger than the radical.

Conjecture 2.1 (ABC conjecture). For every $\varepsilon > 0$ there exists a constant $K(\varepsilon)$ such that

$$h(a,b,c) \leq r(a,b,c) + \varepsilon h(a,b,c) + K(\varepsilon),$$

for every sum $a+b=c$ of coprime nonzero integers.

Equivalently, we may write this inequality as

$$h(a,b,c) \leq \frac{1}{1-\varepsilon}\,r(a,b,c) + \frac{K(\varepsilon)}{1-\varepsilon}. \tag{2.3}$$

For example, this inequality expresses the fact that if one fixes the radical, i.e., if one considers sums of integers composed of a fixed set of prime numbers, then there are only finitely many such sums, and the summands satisfy an a priori bound.

2.1. **Original interest of the ABC conjecture.** The ABC conjecture was formulated in 1983 by Masser and Oesterlé, as a possible approach to Fermat's conjecture (Wiles' theorem):

> For $n \geq 3$, the equation $x^n + y^n = z^n$ has no solutions in positive integers x, y, z.

Indeed, this is a simple consequence of the ABC conjecture. Let $x^n + y^n = z^n$ be a solution. This is a sum of integers, and the height of this sum is $h = \log z^n$. The radical is composed of the prime factors of $x^n y^n z^n$, hence of the prime factors of xyz. Thus $r = \sum_{p\mid xyz} \log p \leq \log xyz < \log z^3$. We use formulation (2.3) of the ABC conjecture, with $\varepsilon = 1/2$. Dividing by $\log z$, we obtain $n < 6 + \frac{2K(1/2)}{\log z}$. Since it is known that there are no solutions for $n = 3$, 4, 5 or 6, this leaves only finitely many values of x, y, z and n to check.

For the later applications, we need to reformulate the definition of the height and the radical, so that we do not need to assume that a, b, c are integers and coprime. To do this, we introduce the projective plane and the valuations of $\mathbf{Q}$.

2.2. **The valuations of Q.** A **valuation** of $\mathbf{Q}$ is a function $v\colon \mathbf{Q} \to \mathbf{R} \cup \{-\infty\}$ satisfying, for some constant K,

$$v(x) = -\infty \text{ only for } x = 0,$$
$$v(xy) = v(x) + v(y) \text{ for all } x, y \in \mathbf{Q}^*,$$
$$v(x + y) \le K + (v(x) \vee v(y)) \text{ for all } x, y \in \mathbf{Q}.$$

Here, $v(x) \vee v(y)$ is the maximum of $v(x)$ and $v(y)$.

Given a prime number p, we denote the number of factors p of the rational number x by $\operatorname{ord}_p(x)$. Then we define the **p-adic valuation** of $\mathbf{Q}$ as

$$v_p(x) = -\operatorname{ord}_p(x) \log p,$$

and the **valuation at** ∞ as

$$v_\infty(x) = \log |x|.$$

For example, $v_2(4/3) = -2\log 2$, $v_3(4/3) = \log 3$, $v_p(4/3) = 0$ for all other p-adic valuations, and $v_\infty(4/3) = \log 4/3$.

One checks that these functions are indeed valuations. The p-adic valuations are *nonarchimedean*, i.e., $v_p(x + y) \le v_p(x) \vee v_p(y)$, for every $x, y \in \mathbf{Q}$. The valuation at infinity satisfies $v_\infty(x+y) \le \log 2 + (v_\infty(x) \vee v_\infty(y))$, for every $x, y \in \mathbf{Q}$, and we call it *archimedean*. It is known that these are all the valuations of $\mathbf{Q}$, except for the trivial valuation, $v(0) = -\infty$ and $v(x) = 0$ for $x \ne 0$.

Every nonzero rational number has a factorization into prime factors,

$$|x| = \prod_p p^{\operatorname{ord}_p(x)}.$$

Taking logarithms, we obtain the following important relation between the valuations of $\mathbf{Q}$. Here, and in the rest of this paper, $\sum_v$ denotes *summation over all valuations of* $\mathbf{Q}$, except the trivial one.

Proposition 2.2 (Sum formula). *For $x \in \mathbf{Q}^*$,*

$$\sum_v v(x) = 0.$$

In other words, the sum of all the valuations of $\mathbf{Q}$ *is the trivial valuation.*

2.2.1. *Extension of a valuation to an algebraic extension of* $\mathbf{Q}$. An algebraic number α is usually viewed as a complex root of its minimal polynomial. Then $|\alpha|$ is just the modulus of this complex number, and this extends v_∞ to a valuation of $\mathbf{Q}(\alpha)$.

To extend a finite valuation v_p is less easy. But if one is willing to accept the p-adic closure $\mathbf{Q}_p$ of $\mathbf{Q}$ and the algebraic closure $\mathbf{C}_p$ of $\mathbf{Q}_p$, with the corresponding extension of v_p to $\mathbf{C}_p$, this becomes just as easy as for v_∞. Namely, every embedding $\sigma\colon \mathbf{Q}(\alpha) \to \mathbf{C}_p$ gives an extension of v_p defined by $v_p(\beta) = v_p(\sigma(\beta))$, for $\beta \in \mathbf{Q}(\alpha)$.

2.3. **The projective plane over Q.** We denote by $\mathbf{P}^2(\mathbf{Q})$ the **projective plane** over $\mathbf{Q}$, i.e., the set of triples $(x:y:z)$, for $x,y,z \in \mathbf{Q}$ not all zero, where for $\lambda \in \mathbf{Q}^*$, the triples $(x:y:z)$ and $(\lambda x : \lambda y : \lambda z)$ denote the same point of $\mathbf{P}^2(\mathbf{Q})$. Subsets of $\mathbf{P}^2(\mathbf{Q})$ may be given by *homogeneous* equations. In particular, we will consider the subset given by the equation $x+y=z$, which is a *line* in $\mathbf{P}^2(\mathbf{Q})$.

For each point of $\mathbf{P}^2(\mathbf{Q})$, we have many different ways to denote this point. For example, given a point $(x:y:z)$, we may choose λ such that λx, λy and λz are *coprime integers*. Another useful choice is to divide by z if $z \neq 0$, to obtain the coordinates $(f:g:1)$ for the point $(x:y:z)$, where $f = x/z$ and $g = y/z$.

The triple $(0:0:0)$ is not a point of $\mathbf{P}^2(\mathbf{Q})$. In §6.6, we still need to consider it, and then we call it **indeterminate**.

The **height** of the point $P = (a:b:c) \in \mathbf{P}^2(\mathbf{Q})$ is defined by

$$h(P) = h(a:b:c) = \sum_v \max\{v(a), v(b), v(c)\},$$

where, as always, v runs over all valuations of $\mathbf{Q}$. If a, b and c are nonzero, the **radical** of P is defined by

$$r(P) = r(a:b:c) = \sum_{p:\ \#\{v_p(a), v_p(b), v_p(c)\} \geq 2} \log p.$$

One needs to check that these definitions do not depend on the choice of coordinates for P. For the radical, this is easy. For the height, one needs Proposition 2.2 to do this. Then one can choose relatively prime integer coordinates for P to see that these new definitions coincide with (2.1).

We define the **error term** of P as

$$e(P) = e(a:b:c) = \max\{h(P) - r(P), 0\}.$$

With these definitions, we reformulate the ABC conjecture.

Conjecture 2.3 (Reformulation of Conjecture 2.1). For all $\varepsilon > 0$ there exists a constant $K(\varepsilon)$ such that

$$e(P) \leq \varepsilon h(P) + K(\varepsilon), \tag{2.4}$$

for every point $P = (a:b:c) \in \mathbf{P}^2(\mathbf{Q})$ on the line $a+b=c$ with $abc \neq 0$.

2.4. **The type of the error term.** In view of (3.3) below, one might think that the ABC conjecture for $\mathbf{Q}$ could be improved to $e(P) \leq \log h(P) + K \log\log h(P)$, for every $P \in \mathbf{P}^2(\mathbf{Q})$ on the line $a+b=c$ with $abc \neq 0$. However, in [8], the author constructs an infinite sequence of such points P with

$$e(P) \geq 6.07 \frac{\sqrt{h(P)}}{\log h(P)}. \tag{2.5}$$

(See also [19, Theorem 2]). This result thus provides an upper bound for the strongest possible version of the ABC conjecture. Ignoring the factor $\log h(P)$, the conjecture that $e(P) \leq K\sqrt{h(P)}$ is indeed supported by numerical data. For example, the fourth sum in (2.2) has $e(P) = 13.88\ldots$, which is $2.309\ldots$ times $\sqrt{h(P)}$.

Assume that in (2.4), we know $K(\varepsilon)$ explicitly as a function of ε. Then we determine, for every value of h, the minimum $\psi(h)$ of $\varepsilon h + K(\varepsilon)$,

$$\psi(h) = \min_{\varepsilon > 0} \varepsilon h + K(\varepsilon).$$

This then allows us to formulate Conjecture 2.3 as

$$e(P) \le \psi(h(P)), \tag{2.6}$$

for some function $\psi(h) = o(h)$.

For example, if $K(\varepsilon) = \exp(1/\varepsilon)$, we find $\psi(h) \approx h/\log h$. Likewise, $K(\varepsilon) = K/\varepsilon$ corresponds to $\psi(h) = 2\sqrt{Kh}$. Finally, $K(\varepsilon) = -\log \varepsilon$ gives $\psi(h) = \log h + 1$, which is impossible by inequality (2.5) above.

3. Algebraic and meromorphic functions

This section is not needed to understand the rest of this paper. We have assembled the facts about algebraic functions that we do need in §§6.2, 6.5. The reader unfamiliar with the basic theory of algebraic functions may read §6.2 before reading this section.

Proposition 2.2 expresses the fact that $\mathbf{Q}$ is a *global field*. Other global fields are the function field $\mathbf{C}(C)$ of a curve C and the field $\mathcal{M}$ of meromorphic functions. For every global field one can formulate the ABC conjecture, Roth's theorem and Mordell's conjecture.

3.1. Algebraic functions.

Let C be an algebraic curve, of genus g. The field of maps $f\colon C(\mathbf{C}) \longrightarrow \mathbf{P}^1(\mathbf{C})$ has the valuations $v_x(f) = -\operatorname{ord}_x(f)$, for each point $x \in C(\mathbf{C})$. The analogue of Proposition 2.2 is $\sum_x v_x(f) = 0$.

For a non-constant map $f\colon C(\mathbf{C}) \longrightarrow \mathbf{P}^1(\mathbf{C})$, we define the height and the radical of $P = (f : 1-f : 1) \in \mathbf{P}^2(\mathbf{C}(C))$ by $h(P) = \deg f$, and $r(P) = \# f^{-1}\{0, 1, \infty\}$. By (6.6) and (6.7), counting only the ramification above 0, 1 and ∞, we obtain $2g-2 \ge -2\deg f + \sum_{f(x)=0,1,\infty} (e_x(f) - 1) = \deg f - \sum_{f(x)=0,1,\infty} 1$. Thus

$$h(P) \le r(P) + 2g - 2,$$

which is the analogue of the ABC conjecture for algebraic functions. The only question that remains is whether this inequality is sharp. In other words, does there exist a map $f\colon C \longrightarrow \mathbf{P}^1$ that is only ramified over 0, 1 and ∞? This question is answered in Theorem 6.1.

By (6.7), there does not exist an algebraic function $p\colon \mathbf{P}^1 \longrightarrow C$ if $g \ge 1$. This is the analogue of Mordell's conjecture[1]. The analogue of Roth's theorem is the following theorem (see [10, Theorem 1.1, Ch. 7]):

> Let A be a finite set of points of $C(\mathbf{C})$. Choose, for each $a \in A$, a curve C_a, a covering $p_a\colon C_a \longrightarrow C$, a point $x_a \in C_a(\mathbf{C})$ above a and a map $\alpha_a\colon C_a(\mathbf{C}) \longrightarrow \mathbf{P}^1(\mathbf{C})$. Let w_a be the valuation associated with x_a, and $\lambda_a(\alpha, \beta) = \max\{0, -w_a(\alpha - \beta)\}$. Let $\varepsilon > 0$. Then there exists a constant K such that
>
> $$\sum_{a \in A} \lambda_a(f \circ p_a, \alpha_a) \le 2\deg f + \varepsilon \deg f + K,$$
>
> for every map $f\colon C(\mathbf{C}) \longrightarrow \mathbf{P}^1(\mathbf{C})$.

[1] The analogue of Mordell's conjecture for function fields F over $\mathbf{C}$ is that all but finitely many points of $C(F)$ are constant, i.e., they lie in $C(\mathbf{C})$. See [11, Theorem 2.3, Ch. I] and [15, Lecture II, p. 39].

3.2. **Meromorphic functions.** This situation has the flavor of both algebraic functions and rational numbers. This is due to the fact that the field $\mathcal{M}$ of meromorphic functions has *archimedean* valuations. Also the nonarchimedean valuations are scaled by a certain factor, the analogue of $\log p$ for the valuation v_p. For each complex number x with $|x| < \rho$, we have a valuation $v_x(f,\rho) = -\operatorname{ord}_x(f)\log\frac{\rho}{|x|}$, and for each $|z| = \rho$, we have the valuation $v_z(f,\rho) = \log|f(z)|$. See [12, 20] for an introduction to Nevanlinna theory and these ideas.

For $a \in \mathbf{P}^1(\mathbf{C})$ and $f \in \mathcal{M}$ with $f(0) \neq a$ or ∞, we have,

$$\begin{aligned} N_a(f,\rho) + \lambda_a(f,\rho) + \log\frac{|f(0)-a|}{\sqrt{1+|a|^2}\sqrt{1+|f(0)|^2}} &= \\ = N_\infty(f,\rho) + \lambda_\infty(f,\rho) - \log\sqrt{1+|f(0)|^2}&, \end{aligned} \tag{3.1}$$

where N_a and N_∞ respectively count the a-points and the poles of f,

$$N_a(f,\rho) = \sum_{|x|<\rho,\, f(x)=a} \operatorname{ord}_x(f-a)\log\frac{\rho}{|x|},$$
$$N_\infty(f,\rho) = \sum_{|x|<\rho,\, f(x)=\infty} -\operatorname{ord}_x(f)\log\frac{\rho}{|x|},$$

and λ_a and λ_∞ respectively measure the closeness of f to a and to ∞ on large circles,

$$\lambda_a(f,\rho) = \int_{|z|=\rho} -\log\frac{|f(z)-a|}{\sqrt{1+|a|^2}\sqrt{1+|f(z)|^2}}\,\frac{dz}{2\pi i z},$$
$$\lambda_\infty(f,\rho) = \int_{|z|<\rho} \log\sqrt{1+|f(0)|^2}\,\frac{dz}{2\pi i z}.$$

The height, $h(f,\rho)$, of f is defined by the right hand side of (3.1). Further, the ramification of f is measured by $R(f,\rho) = \sum_{|x|<\rho}(e_x(f)-1)\log\frac{\rho}{|x|}$.

For $a = 0$, equality (3.1) expresses the analogue of Proposition 2.2. The analogue of Roth's theorem is usually called the 'second main theorem of Nevanlinna theory':

> For every finite subset A of $\mathbf{P}^1(\mathbf{C})$ there exist a constant K and an open set $V \subset (0,\infty)$ of finite total length such that
>
> $$\sum_{a\in A}\lambda_a(f,\rho) + R(f,\rho) \leq 2h(f,\rho) + \log h(f,\rho) + K\log\log h(f,\rho), \tag{3.2}$$
>
> for all $\rho \notin V$.

The height and the radical of a point $P = (f : 1-f : 1) \in \mathbf{P}^2(\mathcal{M})$ with $f(0) \neq 0$, 1 or ∞ are defined by $h(P,\rho) = h(f,\rho)$ and $r(P,\rho) = \sum_{f(x)=0,1,\infty}\log\frac{\rho}{|x|}$. Thus $r(P,\rho) \geq N_0(f,\rho) + N_1(f,\rho) + N_\infty(f,\rho) - R(f,\rho)$. Taking $A = \{0,1,\infty\}$ in (3.2) yields the ABC conjecture for meromorphic functions, with a bound $2\log h(f,\rho) + K\log\log h(f,\rho)$ for the error term,

$$h(P,\rho) \leq r(P,\rho) + 2\log h(P,\rho) + K\log\log h(P,\rho), \tag{3.3}$$

for all $\rho > 0$ outside a set of finite total length.

The analogue of Mordell's conjecture is that there does not exist a non-constant holomorphic map $f\colon \mathbf{C} \longrightarrow C(\mathbf{C})$ when the genus of C is at least two.

4. Roth's theorem

In 1955, K.J. Roth proved the following theorem, see [16],

Theorem 4.1. *Let α be algebraic over $\mathbf{Q}$ and $\varepsilon > 0$. Then*

$$\left|\alpha - \frac{s}{t}\right| < \frac{1}{t^{2+\varepsilon}}$$

for only finitely many rational numbers s/t.

We define the **height** of $x = s/t$, where $s, t \in \mathbf{Z}$ are coprime, as

$$h(x) = \max\{\log|s|, \log|t|\}.$$

Given a valuation w of $\mathbf{Q}$ and an algebraic number α, we extend w to a valuation of $\mathbf{Q}(\alpha)$. The function

$$\lambda_w(x,\alpha) = \max\{0, -w(x-\alpha)\},$$
$$\lambda_w(x,\infty) = \max\{0, w(x)\}.$$

measures the *w-adic closeness* of x to α and to ∞, respectively. We now formulate a generalization of Roth's theorem, see [10, Theorem 1.1, Ch. 7],

Theorem 4.2. *Let S be a finite set of valuations of $\mathbf{Q}$. Let α_w, for each $w \in S$, be an algebraic number or ∞, and extend w to a valuation of $\mathbf{Q}(\alpha_w)$. Let $\varepsilon > 0$. Then there exists a constant K such that*

$$\sum_{w\in S} \lambda_w\left(x, \alpha_w\right) \leq 2h(x) + \varepsilon h(x) + K, \tag{4.1}$$

for every $x \in \mathbf{Q}$.

4.1. **The type of an algebraic number.** In view of (3.2), one could hope to improve (4.1) to $\sum_{w\in S} \lambda_w\left(x,\alpha_w\right) \leq 2h(x) + \psi\left(h(x)\right)$, with $\psi(h) = \log h + K \log\log h$, where K may depend on $\{\alpha_w : w \in S\}$. In general, a function ψ for which this inequality is satisfied is called a **type** of $\{\alpha_w\}$. In §6.4, we will show,

Theorem 4.3. *Assume the ABC conjecture with bound $\psi(h)$, see (2.6). Then there exist constants d and K such that (4.1) is satisfied for every $x \in \mathbf{Q}$, with $\varepsilon(h(x))+K$ replaced by $\psi(dh(x)) + K$.*

Thus, in view of (2.5), the best possible version of Roth's theorem that can be obtained using the ABC conjecture (and the method of §6.4) is (1.1).

5. Mordell's conjecture

In 1922, L.J. Mordell made the following conjecture, see [14],

Conjecture 5.1. Let C be an algebraic curve defined over $\mathbf{Q}$ of genus $g \geq 2$. Then $C(\mathbf{Q})$ is finite.

For example, the curve given by the equation $y^2 = x^5 + x + 1$ in the plane has genus 2. Thus Mordell's conjecture says that $x^5 + x + 1$ is the square of a rational number for only finitely many rational values of x.

Mordell's conjecture was proved by Faltings in 1983 [6]. In 1991, Vojta gave a proof along the lines of Diophantine approximation, see [2, 7, 21]. As Vojta points out in [21], the known proofs of this conjecture are *ineffective*, in the sense that given an algebraic curve, one can obtain an explicit upper bound for the number of points in $C(\mathbf{Q})$, but not for their height. In the above example, this means that one does not know an upper bound for the numerator and denominator of x.

5.1. **ABC implies effective Mordell.** Using the ABC conjecture, we obtain an algorithm to find all points of $C(\mathbf{Q})$ as follows:

- Construct a special map $f\colon C \longrightarrow \mathbf{P}^1$ (see §6.3);
- Then every point $x \in C(\mathbf{Q})$ either has $f(x) = 0$, 1 or ∞, or the height of $f(x)$ is bounded by an explicit constant (see §6.7).

It thus remains to look for points of C in the fibers of f above 0, 1 and ∞, and above finitely many other points of $\mathbf{P}^1(\mathbf{Q})$.

6. ABC implies Roth's theorem and Mordell's conjecture

We collect facts from the theory of heights in §6.1, see [10, 17]. In §§6.2, 6.5–6.6, we collect facts from algebraic geometry, see [9, 10, 17].

6.1. **The theory of heights.** For $x = (x_0 : \cdots : x_n) \in \mathbf{P}^n(\mathbf{Q})$, the **height** of x is defined by

$$h(x) = \sum_v \max\{v(x_0) : \cdots : v(x_n)\}.$$

The first, and most important fact about the height is that for every $B > 0$, *the number of $x \in \mathbf{P}^1(\mathbf{Q})$ with $h(x) \leq B$ is finite.*

Let m be a polynomial in two variables of total degree d. Then there exists a constant K such that,

$$\log|m(s,t)| \leq d\,h(s:t) + K, \tag{6.1}$$

for $s, t \in \mathbf{Z}$ coprime.

A map $f\colon \mathbf{P}^1(\mathbf{Q}) \longrightarrow \mathbf{P}^m(\mathbf{Q})$ of degree d is given by (1.3), with $n = 1$ and $f_i \in \mathbf{Q}[x_0, x_1]$ of degree d. Then $h(f(x)) \leq dh(x) + K$. The converse inequality is more difficult to prove, but it holds too. Thus,

$$|h(f(x)) - dh(x)| \leq K, \tag{6.2}$$

for some constant K. For example, for the map $P\colon (a : b) \longmapsto (a : b : a + b)$ we have

$$h(x) \leq h(P(x)) \leq h(x) + \log 2. \tag{6.3}$$

In §6.7, we need the theory of heights on a curve C. To define a height function on $C(\mathbf{Q})$, we choose a map $f\colon C \longrightarrow \mathbf{P}^1$. If f has degree d, we define the height $h(x) = h_f(x)$ of $x \in C(\mathbf{Q})$ by

$$h(x) = h_f(x) = \frac{1}{d}h(f(x)). \tag{6.4}$$

If $g\colon C \longrightarrow \mathbf{P}^1$ is another map, there exists a constant K such that,

$$|h_f(x) - h_g(x)| \leq K\sqrt{h_f(x)}, \tag{6.5}$$

for all $x \in C\left(\bar{\mathbf{Q}}\right)$, the points with algebraic coordinates on C.

6.2. **Ramification and Hurwitz's formula.** Consider a map $f\colon C \longrightarrow C'$ between nonsingular algebraic curves. Allowing complex values for the coordinates, we obtain a map $f\colon C(\mathbf{C}) \longrightarrow C'(\mathbf{C})$ of Riemann surfaces. For a point $y \in C'(\mathbf{C})$, the preimage $f^{-1}\{y\}$ contains in general a certain number, say d, points. Only for finitely many points y, the preimage contains a different number of points, and then it contains less than d points. This number d is called the **degree** of f, denoted by $\deg f$. When $\#f^{-1}\{y\} < \deg f$, we say that f is **ramified over** y, or **above** y.

For a point $x \in C(\mathbf{C})$, in general, f maps a small enough neighborhood of x in $C(\mathbf{C})$ one-to-one to a small neighborhood of $f(x)$ in $C'(\mathbf{C})$. Only for finitely many points x, the map f is not one-to-one on any neighborhood of x. For such points x, we say that f is **ramified at** x. In that case, there exists a number $e \geq 2$ and a small neighborhood U of x in $C(\mathbf{C})$ such that the restriction of f to $U - \{x\}$ is e-to-one. This number e is called the **multiplicity** of f at x, denoted by $e_x(f)$. Also, f is not ramified at x if and only if $e_x(f) = 1$.

If $f\colon C \longrightarrow \mathbf{P}^1$ maps to the Riemann sphere $\mathbf{P}^1(\mathbf{C})$, we have that $e_x(f) = \operatorname{ord}_x(f - f(x))$ if $f(x) \neq \infty$, and $e_x(f) = -\operatorname{ord}_x(f)$ if $f(x) = \infty$.

We can check whether f is ramified at x with the derivative. Let $\Delta \subset \mathbf{C}$ be the unit disc and let $\varphi\colon \Delta \longrightarrow U$ be an analytic bijection, with U as above and $\varphi(0) = x$. Likewise, let $\psi\colon f(U) \longrightarrow \mathbf{C}$ be analytic one-to-one. Then $g = \psi \circ f \circ \varphi\colon \Delta \longrightarrow \mathbf{C}$ is analytic and e-to-one outside 0. Thus $g(z) = g(0) + g_e \cdot z^e + \dots$, and f is ramified if and only if $g'(0) = 0$.

We have that f is ramified above y if and only if f is ramified at some point x with $f(x) = y$. Let $g\colon C' \longrightarrow C''$ be another map. Then $\deg(g \circ f) = \deg f \cdot \deg g$ and $g \circ f$ is ramified exactly at each point where f is ramified and at each point $x \in C(\mathbf{C})$ such that g is ramified at $f(x)$. Also, $g \circ f$ is ramified exactly over each point over which g is ramified and over each point $z \in C''$ such that f is ramified over some point in $g^{-1}\{z\}$.

If we count the points in $f^{-1}\{y\}$ with multiplicity, their number is always $\deg f$,

$$\text{for every } y \in C'(\mathbf{C})\colon \quad \sum_{x\colon\, f(x)=y} e_x(f) = \deg f. \tag{6.6}$$

We also need Hurwitz's formula, which relates the ramification of f with the genus of C and C':

$$2g(C) - 2 = (2g(C') - 2) \cdot \deg f + \sum_{x \in C(\mathbf{C})} (e_x(f) - 1)\,. \tag{6.7}$$

Note that the sum on the right is finite, since $e_x > 1$ for only finitely many x. For example, the genus of $\mathbf{P}^1$ is 0. This follows from Hurwitz's formula, applied to the map $z \longmapsto z^2$ from $\mathbf{P}^1$ to $\mathbf{P}^1$.

Finally, for a map that is only ramified over 0, 1 and ∞, we have,

$$\begin{aligned} 2g(C) - 2 &= -2\deg f + \sum_{f(x)=0,\,1,\,\infty} (e_x(f) - 1) \\ &= \deg f - \#f^{-1}\{0, 1, \infty\}, \end{aligned} \tag{6.8}$$

by (6.6) and (6.7).

6.3. **Belyĭ's construction.** The following theorem answers the question whether equality is possible in the ABC conjecture for algebraic functions, see §3.1.

Theorem 6.1 (Belyĭ [1]). *Given an algebraic curve C defined over* $\mathbf{Q}$ *and a finite subset* Σ *of algebraic points of* C*, there exists a map* $f\colon C \longrightarrow \mathbf{P}^1$*, defined over* $\mathbf{Q}$*, such that* f *is only ramified over* 0*,* 1 *and* ∞*, and* $f(\Sigma) \subseteq \{0,1,\infty\}$*.*

Proof. The proof is given in three steps.

Step 1. Reduction to $C = \mathbf{P}^1$*.* Let g be any map $C \longrightarrow \mathbf{P}^1$ defined over $\mathbf{Q}$ and consider the finite subset of $\mathbf{P}^1$,

$$\Sigma' = g(\Sigma) \cup \left\{ x \in \mathbf{P}^1 : g \text{ is ramified over } x \right\}.$$

If $f'\colon \mathbf{P}^1 \longrightarrow \mathbf{P}^1$ is the map of the theorem applied to $\mathbf{P}^1$ and Σ', then we take $f = f' \circ g$. We assume now that $C = \mathbf{P}^1$ and that $\Sigma \subset \mathbf{P}^1$ is a finite set of algebraic points.

Step 2. Reduction of the degree of $\alpha \in \Sigma$*.* Let d be the maximal degree over $\mathbf{Q}$ of the elements of Σ and choose $\alpha \in \Sigma$ of degree d. The algebraic number α is root of a polynomial $m(X)$ of degree d with rational coefficients. We obtain a map $m\colon \mathbf{P}^1 \longrightarrow \mathbf{P}^1$,

$$m\colon\ (x_0 : x_1) \longmapsto \left(x_1^d m\left(x_0/x_1 \right) : x_1^d \right),$$

which is ramified at ∞ and at every point x where the derivative $m'(x)$ vanishes. Consider the set

$$\Sigma' = m(\Sigma) \cup \{m(x)\colon m'(x) = 0\} \cup \{\infty\}.$$

Now $m(\alpha) = 0$, and for every $\beta \in \Sigma$, the degree of $m(\beta)$ is at most the degree of β. Moreover, since m' has degree $d-1$, $m(x)$ has degree at most $d-1$ over $\mathbf{Q}$ for a root x of m'. Thus Σ' contains less elements of degree d than Σ.

Repeating this step, eventually Σ will contain only rational points. We may then assume that $\{0,1,\infty\} \subseteq \Sigma$. Namely, if $a \in \Sigma$, then $z \mapsto \frac{1}{z-a}$ is nowhere ramified and it maps a to ∞. Then, if $\{a,\infty\} \subseteq \Sigma$, the map $z \mapsto z - a$ is nowhere ramified and it maps a, ∞ to $0, \infty$, and finally, if $\{a, 0, \infty\} \subseteq \Sigma$, then $z \mapsto z/a$ is nowhere ramified and it maps $a, 0, \infty$ to $1, 0, \infty$.

Step 3. Reduction of the number of elements of Σ*.* Suppose that Σ contains 0, 1, ∞ and a fourth point a/c, with $a, c \neq 0$ and $a \neq c$. Consider the function

$$\varphi(x) = \lambda x^a (1-x)^{c-a}.$$

This map is (possibly) ramified at 0, 1 and ∞, and at points x with $\varphi'(x) = 0$. Moreover, $\varphi(x) = 0$ or ∞ only for $x = 0$, 1 or ∞. Thus for $x \neq 0, 1, \infty$, $\varphi'(x) = 0$ if and only if $\varphi'(x)/\varphi(x) = 0$. Now,

$$\frac{\varphi'(x)}{\varphi(x)} = \frac{a}{x} - \frac{c-a}{1-x},$$

and we find that $\varphi'(x) = 0$ for $x = a/c$. Choose λ such that $\varphi(a/c) = 1$. Then φ is only ramified over 0, 1 and ∞, and since $\varphi\{0,1,\infty\} = \{0,\infty\}$, $\varphi(\Sigma)$ contains less elements than Σ. After repeating this step, Σ will eventually only contain 0, 1 and ∞. $\square$

6.4. ABC implies Roth's theorem.

Proof of Theorem 4.3 using ABC. Let $f\colon \mathbf{P}^1 \longrightarrow \mathbf{P}^1$ be Belyĭ's map associated to $C = \mathbf{P}^1$ and $\Sigma = \{(\alpha_w : 1) : w \in S\}$. Since f is a rational function defined over $\mathbf{Q}$, we can write it as a quotient of relatively prime homogeneous polynomials with integer coefficients,

$$f\left(x_0 : x_1\right) = \left(a\left(x_0, x_1\right) : c\left(x_0, x_1\right)\right),$$

where $a, c \in \mathbf{Z}[x_0, x_1]$ are homogeneous of degree $d = \deg f$. Let $b\left(x_0, x_1\right) = c\left(x_0, x_1\right) - a\left(x_0, x_1\right)$. The polynomials a, b and c factorize into irreducible homogeneous factors over $\mathbf{Z}[x_0, x_1]$,

$$\begin{aligned} a\left(x_0, x_1\right) &= m_1^{e_1}\left(x_0, x_1\right) \cdot \ldots \cdot m_i^{e_i}\left(x_0, x_1\right), \\ b\left(x_0, x_1\right) &= m_{i+1}^{e_{i+1}}\left(x_0, x_1\right) \cdot \ldots \cdot m_j^{e_j}\left(x_0, x_1\right), \\ c\left(x_0, x_1\right) &= m_{j+1}^{e_{j+1}}\left(x_0, x_1\right) \cdot \ldots \cdot m_k^{e_k}\left(x_0, x_1\right). \end{aligned}$$

We write $d_\nu = \deg m_\nu$. Thus $\#f^{-1}\{0\} = \sum_{\nu=1}^{i} d_\nu$, $\#f^{-1}\{1\} = \sum_{\nu=i+1}^{j} d_\nu$, $\#f^{-1}\{\infty\} = \sum_{\nu=j+1}^{k} d_\nu$, and by (6.8),

$$\sum_{\nu=1}^{k} d_\nu = d + 2. \tag{6.9}$$

Since for each $w \in S$, $f\left(\alpha_w\right) = 0$, 1 or ∞, the point α_w is a root of one of the m_ν, i.e., for some μ, $m_\mu(\alpha_w, 1) = 0$, or $m_\mu(1, 0) = 0$ if $\alpha_w = \infty$. Since the polynomials m_ν are coprime, there is only one such m_μ.

In the following, if $w = v_p$ or $w = v_\infty$ is a specific valuation, we shall write α_p or α_∞ instead of α_w.

Let $x \in \mathbf{P}^1(\mathbf{Q})$ be a point such that $f(x) \neq 0, 1, \infty$. Write $x = (s : t)$ with $s, t \in \mathbf{Z}$ coprime. We apply the ABC conjecture to the point

$$P = (f(x) : 1 - f(x) : 1) = (a(s,t) : b(s,t) : c(s,t)).$$

By (6.3) and (6.2), for the height of P, there exists a constant K_0 such that

$$h(P) \geq h(f(x)) \geq d \cdot h(x) - K_0. \tag{6.10}$$

For the radical, we obtain the estimate,

$$r(P) \leq \sum_{p \mid a(s,t) b(s,t) c(s,t)} \log p = \sum_{p \mid m_1(s,t) \ldots m_k(s,t)} \log p. \tag{6.11}$$

If S contains only the valuation v_∞, we may proceed as follows. Let α_∞ be a root of m_μ. Then, by (6.11), (6.1) and case (i) of Lemma 6.2 below,

$$\begin{aligned} r(P) &\leq \sum_{\nu=1}^{k} \log \left|m_\nu(s,t)\right| \\ &\leq \sum_{\nu=1}^{k} d_\nu h(x) - \lambda_\infty\left(x, \alpha_\infty\right) + K, \end{aligned} \tag{6.12}$$

for some constant K. By the ABC conjecture, (6.10) and (6.9), we conclude,

$$\lambda_\infty\left(x, \alpha_\infty\right) \leq 2h(x) + K + \psi(dh(x)).$$

This implies Theorem 4.1. If we know the ABC conjecture with a better bound for the error term than $\psi(h) = \varepsilon h + K(\varepsilon)$, we also obtain a type for α_∞.

In general, S contains more valuations. By (6.11), if a prime p contributes $\log p$ to the radical, then $p|m_\nu(s,t)$ for some ν. This contribution is then bounded by $-v_p\left(m_\nu(s,t)\right)$. If $v_p \in S$ and α_p is a root of m_μ, we apply case (ii) of Lemma 6.2 below to get a stronger bound for the contribution of p to the radical,

$$\log p \leq -v_p\left(m_\mu(s,t)\right) - \lambda_p\left(x, \alpha_p\right) + K_p,$$

for some constant K_p.

Thus the contribution of v_p to the radical is bounded by

$$\sum_{\nu=1}^{k} -v_p\left(m_\nu(s,t)\right),$$

if $v_p \notin S$, and by

$$\left(\sum_{\nu=1}^{k} -v_p\left(m_\nu(s,t)\right)\right) - \lambda_p\left(x, \alpha_p\right) + K_p,$$

if $v_p \in S$. Adding all these contributions, we obtain, by Proposition 2.2,

$$r(P) \leq \sum_{\nu=1}^{k} \log\left|m_\nu(s,t)\right| - \sum_{w \in S,\ \text{finite}} \lambda_w\left(x, \alpha_w\right) + \sum_{w \in S,\ \text{finite}} K_w.$$

We conclude, as in (6.12) above, that

$$r(P) \leq \sum_{\nu=1}^{k} d_\nu h(x) - \sum_{w \in S} \lambda_w\left(x, \alpha_w\right) + K, \tag{6.13}$$

for some constant K. Combining this with (6.10) and (6.9), the ABC conjecture gives us,

$$\sum_{w \in S} \lambda_w\left(x, \alpha_w\right) \leq 2h(x) + K + \psi(dh(x)).$$

This proves Theorem 4.3. □

Lemma 6.2. *Let α be algebraic over $\mathbf{Q}$ of degree d, or $\alpha = \infty$, in which case $d = 1$. Let $m(x_0, x_1) \in \mathbf{Z}[x_0, x_1]$ be the minimal homogeneous polynomial such that $m(\alpha, 1) = 0$ (or $m(x_0, x_1) = x_1$ if $\alpha = \infty$). Let w be a valuation, and extend w to a valuation of $\mathbf{Q}(\alpha)$. Then there exists a constant K such that for all $x = s/t \in \mathbf{Q}$, with $s, t \in \mathbf{Z}$ coprime,*

(i) *if $w = v_\infty$,*

$$v_\infty(m(s,t)) \leq dh(x) - \lambda_\infty\left(x, \alpha\right) + K, \tag{6.14}$$

(ii) *if $w = v_p$,*

$$v_p(m(s,t)) \leq -\lambda_p\left(x, \alpha\right) + K. \tag{6.15}$$

Proof. For $\alpha = \infty$, this follows directly from the definitions.

Assume $\alpha \neq \infty$. If $w(x - \alpha) \geq 0$, we have $\lambda_w(x, \alpha) = 0$. Then (i) follows from (6.1), and (ii) follows since the valuation of an integer is ≤ 0.

If $w(x-\alpha)<0$, we factorize $m(x,1)$ as $(x-\alpha)P(x)$, for some polynomial P over $\mathbf{Q}(\alpha)$. Since $w(x)\le w(\alpha)+\log 2$, we obtain $w(P(x))\le K$ for some constant K. Then for $x=s/t$,

$$\begin{aligned} w(m(s,t)) &= dw(t)+w(m(x,1)) \\ &\le dw(t)+w(x-\alpha)+K \\ &= dw(t)-\lambda_w(x,\alpha)+K. \end{aligned}$$

Case (i) and (ii) follow since $dv_\infty(t)\le dh(x)$ and $dv_p(t)\le 0$, respectively. □

6.5. **The theory of divisors.** Let $C(\mathbf{C})$ be a Riemann surface. A finite sum

$$D=e_1(x_1)+\cdots+e_k(x_k),$$

with $x_1,\dots,x_k\in C(\mathbf{C})$ and $e_1,\dots,e_k\in\mathbf{Z}$, is called a **divisor**. We write $e_i=\operatorname{ord}_{x_i}(D)$, the **order** of D at x_i. Thus we may write $D=\sum_{x\in C(\mathbf{C})}\operatorname{ord}_x(D)(x)$. The **degree** of D is $\deg D=e_1+\cdots+e_k$. A divisor is **positive**, $D\ge 0$, if $\operatorname{ord}_x(D)\ge 0$ for every $x\in C(\mathbf{C})$. We also write $D\le D'$ for $D'-D\ge 0$. The **support** of D is $\sup D=\{x\in C(\mathbf{C})\colon \operatorname{ord}_x(D)\ne 0\}$.

For a map $f\colon C\longrightarrow \mathbf{P}^1$ and $a\in\mathbf{P}^1(\mathbf{C})$, we get a positive divisor

$$f^*(a)=\sum_{x\in f^{-1}\{a\}} e_x(f)(x),$$

the a-**divisor** of f. Note that $\deg f^*(a)=\deg f$.

Let C be defined over $\mathbf{Q}$. Then every embedding $\sigma\colon\bar{\mathbf{Q}}\to\mathbf{C}$ gives an embedding of $C\left(\bar{\mathbf{Q}}\right)$ in $C(\mathbf{C})$. We say that a positive divisor D is **defined over Q** if the image σD does not depend on σ. The positive divisor D defined over $\mathbf{Q}$ is **irreducible** if it cannot be written as sum of positive divisors defined over $\mathbf{Q}$.

For a positive divisor D the maps $f\colon C\longrightarrow\mathbf{P}^1$ with $f^*(\infty)\le D$ form a vector space (respectively, a $\mathbf{Q}$-vector space if D is defined over $\mathbf{Q}$ and we consider only f that are defined over $\mathbf{Q}$). The dimension of this space over $\mathbf{C}$ (respectively, over $\mathbf{Q}$) is denoted by $l(D)$. We need the theorem of Riemann-Roch, in the form

$$l(D)=\deg D+1-g, \tag{6.16}$$

for every positive divisor D with $\deg D\ge 2g-1$ (respectively, defined over $\mathbf{Q}$). Here, g is the genus of C. We use (6.16) to deduce,

Lemma 6.3. *Let D be a positive divisor of the Riemann surface C, of genus g. If $\deg D\ge 2g$, then there exists a map $d\colon C\longrightarrow\mathbf{P}^1$ such that $D=d^*(0)$.*

Proof. The trivial divisor $D=0$ is zero divisor of a constant, nonzero map. This takes care of the case $g=0$ and $D=0$.

Let $D>0$ have degree $\ge 2g$. By (6.16), $l(D)=\deg D+1-g$. We consider a point x in the support of D. Since $\deg(D-(x))=\deg D-1\ge 2g-1$, we obtain by (6.16), $l(D-(x))=l(D)-1$. This means that there exists a function f with $f^*(\infty)\le D$ but not $f^*(\infty)\le D-(x)$. Thus f has a pole at x of order exactly the multiplicity of x in D.

For each point x in the support of D, we find a function f_x with a pole of order $\operatorname{ord}_x(D)$ at x, possible poles at the other points of D, of order at most that of D, and no other poles. Some linear combination of these functions, $f=\sum_{x\in\sup D}c_xf_x$, with $c_x\in\mathbf{Q}$, will have pole divisor D. Then for $d=1/f$ we have $D=d^*(0)$. □

6.6. **Primes of good reduction.** Consider a curve C and a map $f\colon C \longrightarrow \mathbf{P}^1$, defined over $\mathbf{Q}$. We multiply the defining equations (1.2) and (1.3) by a common denominator so that all polynomials involved have integral coefficients. Then, given a prime number p, we may take each coefficient modulo p, and consider only values modulo p for the variables. More generally, we set an algebraic number α to 0 if $v_p(\alpha) < 0$, and to ∞ if $v_p(\alpha) > 0$ for some extension of v_p. We use a bar to denote reduction modulo p.

Modulo some primes however, a defining equation of C may become $0 = 0$, or the map f may get a lower degree, or become indeterminate at some points, meaning that it maps a point to $(0 : 0)$. Also, given a divisor D of C, modulo p some points in D may coalesce or may become indeterminate. We exclude such primes.

We describe a procedure to find a finite set of primes that contains all primes of bad reduction for f. Choose a point $a \in \mathbf{P}^1(\mathbf{Q})$ such that f is not ramified over a and a point $b \in \mathbf{P}^1(\mathbf{Q})$ different from a. Exclude all primes for which a or b become indeterminate or for which they coalesce. Also exclude all primes of bad reduction for the divisors $f^*(a)$ and $f^*(b)$. Modulo a prime that has not been excluded, we have $\bar{a} \neq \bar{b}$, so $\bar{f}$ is not constant, and $\deg \bar{f} = \deg \bar{f}^*(\bar{a}) = \deg f$.

It may still happen that $\bar{f}$ maps a point of C to the indeterminate point. This happens for only finitely many primes, as we see as follows. By Hilbert's Nullstellensatz, there exist homogeneous polynomials $c_0, c_1, \ldots, c_{k+1}$, with integral coefficients, such that,

$$c_0 f_0 + c_1 f_1 + c_2 p_1 + \cdots + c_{k+1} p_k = c x_n^l,$$

for some exponent l and a nonzero integer c. Now C has finitely many points with $x_n = 0$, and we exclude the primes for which the image under f of some of these points becomes indeterminate[2]. We also exclude the prime divisors of c.

6.7. **ABC implies Mordell's conjecture.**

Proof of Conjecture 5.1 using ABC. Let $f\colon C \longrightarrow \mathbf{P}^1$ be Belyĭ's function associated to C and $\Sigma = \emptyset$. Then f is defined over $\mathbf{Q}$, and in particular, $f(x) \in \mathbf{P}^1(\mathbf{Q})$ for $x \in C(\mathbf{Q})$. The divisors $A = f^*(0)$, $B = f^*(1)$ and $C = f^*(\infty)$ have a decomposition into irreducible divisors,

$$\begin{aligned} A &= e_1 M_1 + \cdots + e_i M_i, \\ B &= e_{i+1} M_{i+1} + \cdots + e_j M_j, \\ C &= e_{j+1} M_{j+1} + \cdots + e_k M_k. \end{aligned}$$

We denote the degree of M_ν by d_ν and the degree of f by d. Thus $\#f^{-1}\{0\} = \sum_{\nu=1}^{i} d_\nu$, $\#f^{-1}\{1\} = \sum_{\nu=i+1}^{j} d_\nu$ and $\#f^{-1}\{\infty\} = \sum_{\nu=j+1}^{k} d_\nu$, and by (6.8),

$$\sum_{\nu=1}^{k} d_\nu = d + 2 - 2g < d. \tag{6.17}$$

To complete the argument, we only need that $\sum_{\nu=1}^{k} d_\nu < d$. Thus, we do not need the full power of Belyĭ's construction.

Let N be so large that for each ν, $N M_\nu$ is given as the zero divisor of a function,

$$N M_\nu = m_\nu^*(0).$$

(By Lemma 6.3, we may take $N = 2g$.)

[2] If C is contained in $[x_n = 0]$, we view C as embedded in $\mathbf{P}^{n-1}$, and we use x_{n-1} instead.

Let $x \in C(\mathbf{Q})$ be a point with rational coordinates such that $f(x) \neq 0, 1, \infty$. We apply the ABC conjecture to the point

$$P = (f(x) : 1 - f(x) : 1)$$

to deduce that the height of x is bounded. By (6.3), $h(P) \geq h(f(x))$. We choose the function f to define a height function $h(x)$ on C, by formula (6.4). Then

$$h(P) \geq dh(x). \tag{6.18}$$

We now estimate the radical. Let p be a prime of good reduction for C, f, each m_ν and each M_ν. We use a bar to denote reduction modulo p. The prime p contributes $\log p$ to the radical only if $v_p(f(x)) > 0$, $v_p(f(x)) < 0$ or $v_p(1-f(x)) < 0$, i.e., if $\bar{f}(\bar{x}) = \bar{\infty}$, $\bar{0}$ or $\bar{1}$. Then $\bar{x}$ is in the support of $\bar{A}$, $\bar{B}$ or $\bar{C}$. Since the decompositions of A, B and C remain the same when reducing modulo p, $\bar{x}$ is in the support of some $\bar{M}_\nu$. Thus $\bar{m}_\nu(\bar{x}) = \bar{0}$. This means that $v_p(m_\nu(x)) < 0$. Moreover, since $\bar{m}_\nu^*(\bar{0}) = N\bar{M}_\nu$, $v_p(m_\nu(x))$ is a multiple of $N \log p$. Thus the contribution of p to the radical is bounded by

$$-\frac{1}{N} v_p\left(m_\nu(x)\right) \leq \sum_{\nu=1}^{k} 0 \vee -\frac{1}{N} v_p\left(m_\nu(x)\right). \tag{6.19}$$

Also for v_∞ we obtain

$$0 \leq \sum_{\nu=1}^{k} 0 \vee -\frac{1}{N} v_\infty\left(m_\nu(x)\right). \tag{6.20}$$

Adding all these contributions, we find,

$$\begin{aligned} r(P) &\leq \sum_{p:\ \text{good}} \sum_{\nu=1}^{k} 0 \vee -\frac{1}{N} v_p\left(m_\nu(x)\right) + \sum_{p:\ \text{bad}} \log p \\ &\leq \sum_{\nu=1}^{k} \sum_{v} 0 \vee -\frac{1}{N} v\left(m_\nu(x)\right) + K \\ &= \sum_{\nu=1}^{k} \frac{1}{N} h\left(m_\nu(x)\right) + K, \end{aligned}$$

for some constant K. Now m_ν has degree Nd_ν, hence by (6.5), $h\left(m_\nu(x)\right) \leq Nd_\nu h(x) + O\left(\sqrt{h(x)}\right)$, and we obtain

$$r(P) \leq \sum_{\nu=1}^{k} d_\nu h(x) + K_0\sqrt{h(x)} + K_1, \tag{6.21}$$

for some constants K_0 and K_1. By the ABC conjecture, (6.17) and (6.18),

$$(2g-2)h(x) \leq K_0\sqrt{h(x)} + K_1 + \psi(dh(x)). \tag{6.22}$$

Since $2g - 2 > 0$ (or, more generally, $d - \sum_{\nu=1}^{k} d_\nu > 0$) and $\psi(dh(x)) = o\left(h(x)\right)$, we obtain that $h(x)$ is bounded. This proves Conjecture 5.1 and the algorithm in §5.1. $\square$

Note that compared with (6.13), we get an extra $\sqrt{h(x)}$ term in the estimate for the radical in (6.21). It is remarkable that for the strongest possible form of the ABC conjecture,

$$\psi(h) = 6.07\,\frac{\sqrt{h}}{\log h},$$

by (2.5), the third term on the right of inequality (6.22) is only slightly smaller than the first term on the right, and for weaker types of the error term, the third term is the larger one.

6.8. **Unification of Roth's theorem and Mordell's conjecture.** Let C be defined over $\mathbf{Q}$ of genus g. For an algebraic point $\alpha \in C(\bar{\mathbf{Q}})$ and $e = 2g$, there exists a function $\phi\colon C \longrightarrow \mathbf{P}^1$, defined over $\mathbf{Q}(\alpha)$, with $\phi^*(\infty) = e(\alpha)$. For a valuation w, extended to $\mathbf{Q}(\alpha)$, we define

$$\lambda_w(x,\alpha) = 0 \vee \frac{1}{e} w\left(\phi(x)\right), \tag{6.23}$$

for $x \in C(\mathbf{Q})$.

The following theorem implies both Theorem 4.3 and effective Mordell.

Theorem 6.4. *Let C be a curve of genus g, defined over $\mathbf{Q}$. Let S be a finite set of valuations of $\mathbf{Q}$ and let $\alpha_w \in C(\bar{\mathbf{Q}})$, for each $w \in S$. Extend w to a valuation of $\mathbf{Q}(\alpha_w)$. Assume the ABC conjecture with bound $\psi(h)$. Then there exist constants d, K_0 and K_1 such that,*

$$\sum_{w\in S} \lambda_w\left(x,\alpha_w\right) \le (2-2g)\,h(x) + K_0\sqrt{h(x)} + \psi(dh(x)) + K_1,$$

for every $x \in C(\mathbf{Q})$.

Proof. Let $f\colon C \longrightarrow \mathbf{P}^1$ be Belyĭ's function associated to C and $\Sigma = \{\alpha_w\colon w \in S\}$. We then follow the same argument as in §6.7. To estimate the radical, we also exclude the primes of bad reduction for the functions ϕ that we have used in (6.23) to define λ_w. If $v_p \in S$ contributes to the radical, and $v_p\left(m_\mu\left(\bar{x}\right)\right) < 0$, with $m_\mu\left(\alpha_p\right) = 0$, we obtain by Lemma 6.5 below,

$$\begin{aligned}\log p &\le -\frac{1}{N} v_p\left(m_\mu(x)\right) - \lambda_p\left(x,\alpha_p\right) + K_p\\ &\le \left(\sum_{\nu=1}^{k} 0 \vee -\frac{1}{N} v_p\left(m_\nu(x)\right)\right) - \lambda_p\left(x,\alpha_p\right) + K_p.\end{aligned}$$

Also for v_∞ we obtain by Lemma 6.5,

$$0 \le \left(\sum_{\nu=1}^{k} 0 \vee -\frac{1}{N} v_\infty\left(m_\nu(x)\right)\right) - \lambda_\infty\left(x,\alpha_p\right) + K_\infty.$$

We then finish the proof as in §6.7. □

Lemma 6.5. *Let $\alpha \in C(\bar{\mathbf{Q}})$ and $m\colon C \longrightarrow \mathbf{P}^1$ be defined over $\mathbf{Q}$ such that $m(\alpha) = 0$ with multiplicity at least N. Let w be a valuation, extended to $\mathbf{Q}(\alpha)$. Then there exists a constant K such that,*

$$0 \le \left(0 \vee -\frac{1}{N} w(m(x))\right) - \lambda_w(x,\alpha) + K,$$

for all $x \in C(\mathbf{Q})$.

Proof. If $\lambda_w(x,\alpha)$ is large, then $w(\phi(x))$ is large by (6.23). Thus $\phi(x)$ is close to ∞ in the w-adic topology. By an approximation procedure, we find a point y close to x such that $\phi(y)=\infty$, i.e., $y \in \sup \phi^*(\infty)$. But then $y=\alpha$ and hence x is close to α. Since m vanishes at α to order N at least, and ϕ has a pole of order e, the lemma follows. □

References

1. Belyĭ, G. V., *On Galois extensions of a maximal cyclotomic field*, Math. USSR Izvestija Vol. **14** (1980), No. 2, 247–256.
2. Bombieri, E., *The Mordell Conjecture Revisited*, Ann. Scu. Norm. Sup. Pisa **17**, 4 (1990), 615–640, and *Errata-Corrige*, *ibid* **18**, 3 (1991), 473.
3. Bombieri, E., *Roth's Theorem and the abc-Conjecture*, preprint ETH Zürich, 1994.
4. Edixhoven, B. and Evertse, J.-H., (Eds.), *Diophantine Approximation and Abelian Varieties*, Lect. Notes in Math. 1566, Springer-Verlag, New York, 1993.
5. Elkies, N. D., *ABC implies Mordell*, Intern. Math. Research Notices No. **7** (1991), 99–109.
6. Faltings, G., *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Inv. math. **73** (1983), 349–366.
7. Faltings, G., *Diophantine approximation on abelian varieties*, Annals of Mathematics **133** (1991), 549–576.
8. Frankenhuysen, M. van, *Good ABC Examples over Number Fields*, preprint, submitted for publication.
9. Hartshorne, R., *Algebraic geometry*, Graduate texts in mathematics **52**, Springer-Verlag, New York, 1977.
10. Lang, S., *Fundamentals of Diophantine Geometry*, Springer-Verlag, New York, 1983.
11. Lang, S., *Number Theory III*, Encyclopedia of Mathematical Sciences, vol. 60, Springer-Verlag, New York, 1991.
12. Lang, S. and Cherry, W., *Topics in Nevanlinna theory*, Lect. Notes in Math. 1433, Springer-Verlag, New York, 1990.
13. Lang, S. and Trotter, H., *Continued fractions for some algebraic numbers*, J. reine angew. Math. **255** (1972), 112–134, and *Addendum to "Continued fractions for some algebraic numbers"*, *ibid* **267** (1974), 219–220.
14. Mordell, L. J., *On the rational solutions of the indeterminate equation of the third and the fourth degree*, Math. Proc. Cambridge Philos. Soc. **21** (1922), 179–192.
15. Mumford, D., *Curves and their Jacobians*, The University of Michigan Press, Ann Arbor, 1976.
16. Roth, K. F., *Rational approximation to algebraic numbers*, Mathematika **2** (1955), 1–20.
17. Serre, J.-P., *Lectures on the Mordell-Weil theorem*, 2nd ed., Aspects of Mathematics, Vieweg, Wiesbaden, 1990.
18. Stark, H., *An explanation of some exotic continued fractions found by Brillhart*, Computers in number theory, Proc. Sci. Res. Council Atlas Symp. No. 2, Academic Press, New York, 1971, pp. 21–35.
19. Stewart, C. L. and Tijdeman, R., *On the Oesterlé-Masser conjecture*, Mh. Math. **102** (1986), 251–257.
20. Vojta, P., *Diophantine Approximations and Value Distribution Theory*, Lect. Notes in Math. 1239, Springer-Verlag, New York, 1987.
21. Vojta, P., *Siegel's theorem in the compact case*, Annals of Mathematics **133** (1991), 509–548.

Department of Mathematics, Sproul Hall, University of California, Riverside, CA 92521-0135, USA

E-mail address: machiel@math.ucr.edu