ABC implies the radicalized Vojta height inequality for curves

Machiel van Frankenhuijsen

Department of Mathematics, Utah Valley State College, 800 West University Parkway, Orem, UT 84058-5999, USA

Received 25 July 2005; revised 18 November 2005
Available online 29 April 2007
Communicated by David Goss

Abstract

The truncated or radicalized counting function of a meromorphic function \( f : \mathbb{C} \to \mathbb{P}^1(\mathbb{C}) \) counts the number of times that \( f \) takes a value \( a \), but without multiplicity. By analogy, one also defines this function for numbers. In this sequel to [M. van Frankenhuijsen, The ABC conjecture implies Vojta’s height inequality for curves, J. Number Theory 95 (2002) 289–302], we prove the radicalized version of Vojta’s height inequality, using the ABC conjecture. We explain the connection with a conjecture of Serge Lang about the different error terms associated with Vojta’s height inequality and with the radicalized Vojta height inequality.

© 2007 Elsevier Inc. All rights reserved.

MSC: 11J97; 11D45; 11D75; 14G40

Keywords: ABC conjecture; The error term in the ABC conjecture; Radicalized Vojta height inequality; Diophantine approximation; Roth’s theorem; Type of an algebraic number; Mordell’s conjecture; Effective Mordell

1. Introduction

It has long been known that there is an analogy between conjectures and theorems for numbers and for functions. For example, Roth’s theorem corresponds to Nevanlinna’s second main theorem, and Mordell’s conjecture corresponds to the fact that for an analytic map \( f : \Delta_r \to C \) of the disc of radius \( r \) into a Riemann surface of genus two or higher, there is an upper bound on \( f'(0)r \), depending only on \( C \). Paul Vojta was the first to make this analogy quantitative, by
formulating his height inequalities, here referred to as the VHI and the RVHI. (See Section 2 for the definitions.)

**Conjecture 1.1 (Vojta’s Height Inequality [V98, Conjecture 2.1]).** Let $C$ be a smooth complete curve and let $D$ be a finite set of points of $C$, defined over a number field $F$. Let $S$ be a finite set of valuations of $F$, including the archimedean valuations. Let $K$ be a canonical line sheaf and let $A$ be a big line sheaf on $C$. Let $\varepsilon > 0$ and $r \geq 1$. Then there exists a finite subset $Z$ of $C(\bar{F})$, depending on $C, D, S, A, \varepsilon, r$ and $F$, such that

$$h_K(x) + m_{S,D}(x) \leq d(F(x)) + \varepsilon h_A(x)$$

(1.1)

for all $x \in C(\bar{F}) - Z$ for which $[F(x): F] \leq r$.

With the $d(F(x))$ term, defined below in (2.5), one further conjectures that $Z$ is independent of $F$ and $r$.

This conjecture is inspired by the analogous theorem in Nevanlinna theory (see [LC90] for definitions),

Let $f : \mathbb{C} \to C$ be an analytic map into the Riemann surface $C$. Let $D$ be a finite set of points on $C$ and let $A$ be ample. Then there exists an open subset $Z$ of $[0, \infty)$ of finite total length such that

$$h_K(f, \rho) + m_D(f, \rho) + N_{\text{ram}}(f, \rho) \leq \log h_A(f, \rho) + O(\log \log h_A(f, \rho))$$

(1.2)

for all $\rho \in [0, \infty) - Z$.

The error term in this theorem, $\log h_A + O(\log \log h_A)$, is better than the error term in Vojta’s conjecture, $\varepsilon h_A$ for every $\varepsilon > 0$. By analogy, Lang conjectured that the error term in (1.1) is $O(\log h_A(x))$ instead of $o(h_A(x))$. Indeed, when applied to the projective line and $D$ consisting of a single algebraic point $\alpha$, we obtain Roth’s theorem [R55]. Contrary to [vF02, Remark 5.5], there is some numerical evidence that the error term in Roth’s theorem is as in Nevanlinna theory. In the case $F = \mathbb{Q}$, in terms of the partial quotients $a_n$ of the continued fraction expansion of the real algebraic number $\alpha$, this would mean that $a_n = O(n^\kappa)$, for some $\kappa > 0$.

We see that the geometry, $C, D, K$ and $A$, of Conjecture 1.1 and its counterpart in Nevanlinna theory, inequality (1.2), is the same, and the analysis (the proximity and height functions) is completely analogous in the formulation, with the exception of $d(F(x))$ and the counting function of the ramification of $f$,

$$N_{\text{ram}}(f, \rho) = \sum_{|x| < \rho: \ord(f,x) \geq 2} (\ord(f,x) - 1) \log \frac{\rho}{|x|}.$$

It is not entirely clear what the analogue for numbers is of this function (see however [V99, §8]), but one can combine $m_D$ and $N_{\text{ram}}$ by using

$$m_D(f, \rho) + N_{\text{ram}}(f, \rho) = h_D(f, \rho) - N_D(f, \rho) + N_{\text{ram}}(f, \rho) \geq h_D(f, \rho) - N_D^{(1)}(f, \rho)$$
(see [LC90] for the definitions). Here, the radicalized counting function

\[ N_D^{(1)}(f, \rho) = \sum_{|x| < \rho: f(x) \in D} \log \frac{\rho}{|x|} \]

counts the points \( x \in \mathbb{C} \) for which \( f(x) \) is a point in \( D \) with a weight \( \log \frac{\rho}{|x|} \), but without multiplicity. Using this, we obtain a corollary to inequality (1.2) which does not involve \( N_{\text{ram}} \) but instead the radicalized counting function. The analogue of this corollary in number theory does not follow directly from Conjecture 1.1, because Conjecture 1.1 does not have such a ramification term.

**Conjecture 1.2 (Radicalized Vojta Height Inequality [V98, Conjecture 2.3]).** Conjecture 1.1 holds with (1.1) replaced by

\[ h_{K+D}(x) \leq N^{(1)}_{S,D}(x) + d(F(x)) + \varepsilon h_A(x) \]  

(1.3)

for all \( x \in C(\tilde{F}) - Z \) for which \( [F(x) : F] \leq r \).

Again, we conjecture in addition that \( Z \) does not depend on \( F \) and \( r \).

In [N96, p. 500], Noguchi observes that the ABC conjecture follows from Conjecture 1.2, applied to the projective line and \( D = (0) + (1) + (\infty) \). It is known that in the ABC conjecture, \( \varepsilon h_A \) cannot be replaced by \( O(\log h_A) \). Indeed, the best possible error term is \( O(\sqrt{h_A}/\log h_A) \) (see Remark 3.3 below). Accordingly, Lang conjectured that the error term in Conjecture 1.1 follows Nevanlinna theory, on which it is inspired, whereas the error term in the radicalized version, Conjecture 1.2, follows the error term of the ABC conjecture. That is, Conjecture 1.1 may hold with \( \varepsilon h_A(x) \) replaced by \( O(\log h_A(x)) \), and Conjecture 1.2 may hold with \( \varepsilon h_A(x) \) replaced by \( O(\psi(h_A(x))) \), where \( \psi \) is as in Conjecture 3.1.

In [vF02], we proved that Conjecture 1.1, with an error term as in the ABC conjecture, follows from the ABC conjecture. We show here that the ABC conjecture implies Conjecture 1.2. This was already pointed out by Vojta in [V98]. However, Vojta’s proof uses a geometric construction that is quite involved, and therefore loses track of the error term. Our formulation of the ABC conjecture contains a function \( \psi(h) \) for the error term, which is conjectured to be \( o(h) \) and is known to be at least \( O(\sqrt{h}/\log h) \). The only geometric construction that we use is a Belyi function, and we therefore obtain an error term of the form \( \psi(h_A) \). Thus, the ABC conjecture and Conjecture 1.2 are equivalent with the same error term, and both imply Conjecture 1.1, with an error term as in the ABC conjecture. Conversely, Conjecture 1.1 implies the ABC conjecture, but with an \( \varepsilon h_A \) error term [V92], even if Conjecture 1.1 would be known with an \( O(\log h_A) \) error term.

Our proof is very similar to the one in [vF02], the only difference being in the first displayed inequality in [vF02, p. 299], which we replace by the stronger inequality (5.1).

2. The theory of heights

Let \( F \) be a number field with algebraic closure \( \tilde{F} = \bar{Q} \). The completion of \( F \) at the valuation \( v \) is denoted \( F_v \), and, if \( v \) is nonarchimedean, \( F(v) \) denotes the field of residue classes. The *height*
of a valuation is defined by
\[ h_F(v) = \begin{cases} 0, & \text{if } v \text{ is archimedean}, \\ \frac{1}{[F : \mathbb{Q}]} \log \#F(v), & \text{if } v \text{ is nonarchimedean}. \end{cases} \tag{2.1} \]

We normalize the valuations by
\[ v(2) = \frac{[F_v : \mathbb{R}]}{[F : \mathbb{Q}]} \log 2 \text{ if } v \text{ is archimedean, and by } v(p) = -\frac{[F_v : \mathbb{Q}_p]}{[F : \mathbb{Q}]} \log p \text{ if } v \text{ is nonarchimedean and } p \text{ is the rational prime with } v(p) < 0. \]

We refer to [V87, V98, Hu93, vF02] for the basic definitions of the height. We recall here the properties that we need. Let \( R \) be the ring of integers of \( F \) and let \( R_v \) denote the completion of \( R \) at a nonarchimedean valuation \( v \). Let \( C \) be a curve defined over \( F \) and choose a model \( C \) of \( C \) over \( R \). Given a divisor \( D \), we denote the corresponding line bundle also by \( D \). The line bundle is positive if the constant function \( 1 \) is a global section, via this identification of functions with sections. For a positive line bundle \( D \) over \( C \), we let \( D \) be the (noncanonically) corresponding line bundle over \( C \) together with metrics \( \| \cdot \|_v \) on \( D \otimes F_v \) for every archimedean valuation \( v \). For a point \( x \in C(F) \), not in the support of \( D \), we define
\[ \deg_v x^*D = \begin{cases} -\frac{[F_v : \mathbb{R}]}{[F : \mathbb{Q}]} \log \|1\|_v, & \text{if } v \text{ is archimedean}, \\ \frac{1}{[F : \mathbb{Q}]} \log \#((R_v \otimes x^*D)/R_v), & \text{if } v \text{ is nonarchimedean}. \end{cases} \]

Note that \( \deg_v x^*D \) is a multiple of \( h_F(v) \) for every nonarchimedean valuation. The (logarithmic) height of \( x \) with respect to \( D \) is then defined by
\[ h_D(x) = \sum_v \deg_v x^*D, \]
where \( v \) runs over all valuations of \( F \). In general, \( h_D \) is defined by linearity in \( D \).

Let \( S \) be a finite set of places of \( F \). The counting function of \( x \) in \( D \) is defined by
\[ N_{S,D}(x) = \sum_{v \not\in S} \deg_v x^*D, \]
and the proximity of \( x \) to \( D \) is \( m_{S,D}(x) = h_D(x) - N_{S,D}(x) \). The radicalized counting function counts if \( x \in D \), but without multiplicity,
\[ N^{(1)}_{F,S,D}(x) = \sum_{v \not\in S, \left| \deg_v x^*D \right| > 0} h_F(v). \tag{2.2} \]

The height and the counting function do not depend on \( F \), but \( N^{(1)}_{F,S,D} \) does depend on the number field. Indeed, for an extension \( E \) of \( F \), one has
\[ h_E(w) = \frac{[E(w) : F(v)]}{[E : F]} h_F(v). \]
Thus the contribution \( \sum_{w|v} h_E(w) \) of the valuations above \( v \) to the radical satisfies
\[
\frac{1}{[E : F]} h_F(v) \leq \sum_{w|v} h_E(w) \leq h_F(v),
\]
(2.3)

with equality on the right if \( v \) is unramified in \( E \). Hence

\[
0 \leq N_{F,S,D}^{(1)}(x) - N_{E,S,D}^{(1)}(x) \leq d(E) - d(F),
\]
(2.4)

where

\[
d(F) = \frac{1}{[F : \mathbb{Q}]} \log |\text{disc}(F)|.
\]
(2.5)

**Remark 2.1.** \( \deg_v x^s D \geq 0 \) for every nonarchimedean valuation \( v \). For an appropriate choice of metrics, this also holds for the archimedean valuations. We call this a choice of positive metrics, and we say that \( D \) is positive.

Moreover, if the positive divisor \( D = D' + D'' \) is a sum of divisors for which the underlying divisors \( D' \) and \( D'' \) are positive, then one can choose positive metrics on \( D' \) and \( D'' \) such that their sum equals the metric on \( D \).

### 3. The ABC conjecture with error term

Denote a point of \( \mathbb{P}^2 \) by \((a : b : c)\). Consider the divisor \([b = 0]\) of \( \mathbb{P}^2 \) with metrics at the archimedean valuations\(^1\)

\[
\|s\|_v(a : b : c) = \frac{|bs(a : b : c)|_v}{\sqrt{|a|^2_v + |b|^2_v + |c|^2_v}},
\]
(3.1)

for a section \( s \). We choose similar metrics for the divisors \([a = 0]\) and \([c = 0]\).

We let \( h(P) \) be the height of the point \( P = (a : b : c) \in \mathbb{P}^2(F) \) with respect to the divisor \([b = 0]\) with these metrics. Thus

\[
h(P) = \sum_v h_v(a, b, c),
\]

where

\[
h_v(a, b, c) = \frac{[F_v : \mathbb{R}]}{[F : \mathbb{Q}]} \log \sqrt{|a|^2_v + |b|^2_v + |c|^2_v}, \quad \text{if } v \text{ is archimedean,}
\]
\[
h_v(a, b, c) = \frac{1}{[F : \mathbb{Q}]} \max\{v(a), v(b), v(c)\}, \quad \text{if } v \text{ is nonarchimedean.}
\]

Note that the height does not depend on the choice of coordinates for \( P \), even though the local contributions do. Moreover, the height does not depend on the number field \( F \).

\(^1\) The norm is defined by \( |x|_v = \exp(v(x)/[F : \mathbb{R}]) \), which is the ordinary distance of \( x \) to 0 on the real line or complex plane.
The radical of \( P = (a : b : c) \) with \( abc \neq 0 \) is defined by

\[
r_F(P) = \sum_v r_{F,v}(P),
\]

where the contribution of the valuation \( v \) to the radical is given by

\[
r_{F,v}(P) = 0 \quad \text{if } v(a) = v(b) = v(c),
\]

\[
r_{F,v}(P) = h_F(v) \quad \text{otherwise (see (2.1) for } h_F(v)).
\]

The radical depends on the number field, and as for the truncated counting function, we have the bounds

\[
0 \leq r_F(P) - r_E(P) \leq d(E) - d(F),
\]

(3.2)

for an extension \( E \) of \( F \).

Originally, the ABC conjecture was formulated (for \( F = \mathbb{Q} \)) in 1983 by Masser and Oesterlé with \( \varepsilon h(P) + K \) instead of an error term \( \psi(h(P)) \) as below (see [O88]).

**Conjecture 3.1** (ABC Conjecture with Type \( \psi \) [vF99,vF00]). There exists a positive increasing function \( \psi \) with \( \psi(h) = o(h) \) such that

\[
h(P) \leq r_F(P) + d(F) + \psi(h(P)),
\]

(3.3)

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

**Remark 3.2.** The function \( \psi \) may depend on the number field. However, we make the further conjecture that \( \psi \) does not depend on \( F \), so that the term \( d(F) \) takes care of the dependence on the number field.

**Remark 3.3.** In [vF00], following [ST86, Theorem 2], the author shows that there exist infinitely many rational ABC examples such that

\[
h(P) \geq r_Q(P) + 6.07 \frac{\sqrt{h(P)}}{\log h(P)}.
\]

This result provides an upper bound for the strongest possible version of the ABC conjecture. Thus if a function as in Conjecture 3.1 exists, then \( \psi(h) \geq 6.07 \sqrt{h}/\log h \). Indeed, ignoring the factor \( \log h \), numerical data seems to indicate that in general we may take

\[
\psi(h) = K \sqrt{h},
\]

for some constant \( K \), independent of the number field.
4. Belyi’s construction

In [Be80, Theorem 4], Belyi constructs a function with the following property:

**Theorem 4.1.** Let $C$ be an algebraic curve defined over a number field $F$ and let $D \subseteq C(\overline{\mathbb{Q}})$ be a finite set of algebraic points on $C$. Then there exists a morphism $f : C \rightarrow \mathbb{P}^1$ defined over $F$ such that $f(D) \subseteq \{0, 1, \infty\}$ and $f$ is only ramified over $\{0, 1, \infty\}$.

Belyi’s construction provides us with examples of equality in the ABC theorem for function fields [Sto81, vF99]. Let $f : C \rightarrow \mathbb{P}^1$ be a morphism of a complete nonsingular curve $C$ to the projective line. Let $e(x)$ denote the multiplicity of $f$ at the point $x$ on $C$. Thus, for every point $a$ in $\mathbb{P}^1$, the total multiplicity of the fiber over $a$ is constant, $\sum_{x \in f^{-1}(a)} e(x) = \deg f$. By Hurwitz’ formula, writing $g$ for the genus of $C$,

$$2 - 2g = 2\deg f - \sum_{x \in C} (e(x) - 1) = 2\deg f - \sum_{a \in \mathbb{P}^1} \sum_{x \in f^{-1}(a)} (e(x) - 1).$$

Counting only the ramification above 0, 1 and $\infty$, we obtain

$$\deg f \leq 2g - 2 + \#f^{-1}\{0, 1, \infty\}. \quad (4.1)$$

This is the ABC theorem for function fields. The function $f$ corresponds to the ABC sum $(f : 1 - f : 1)$ of height $\deg f$ and radical $\#f^{-1}\{0, 1, \infty\}$. Equality holds if and only if $f$ is a Belyi function. In that case, a canonical divisor of $C$ is given by

$$K = (df/f) = f^*(1) - f^{-1}\{0, 1, \infty\}. \quad (4.2)$$

We use this canonical divisor in the next section.

5. The ABC conjecture implies Vojta’s height inequality

Let $C$ be a curve over $F$, and let $D$ be a finite set of algebraic points of $C$, defined over $F$. Let $f : C \rightarrow \mathbb{P}^1$ be a Belyi map for $D$. The divisors $A = f^*(0)$, $B = f^*(1)$ and $C = f^*(\infty)$ have a decomposition over $F$ into irreducible divisors,

$$A = e_1\mathcal{P}_1 + \cdots + e_i\mathcal{P}_i, \quad B = e_{i+1}\mathcal{P}_{i+1} + \cdots + e_j\mathcal{P}_j, \quad C = e_{j+1}\mathcal{P}_{j+1} + \cdots + e_k\mathcal{P}_k.$$

Choose metrics at the archimedean places on $A$, $B$ and $C$ as in (3.1) and choose positive metrics on each $\mathcal{P}_v$, compatible with this decomposition, as in Remark 2.1. Let $K = B - \sum_{v=1}^k \mathcal{P}_v$ be the canonical line bundle (4.2). Note that as a line bundle without metrics, $D = \sum_{\mu} \mathcal{P}_{\mu}$, where the summation is restricted to the components of $D$. We write this symbolically as $D = \sum_{\mu \in D} \mathcal{P}_{\mu}$.

Let $D = \sum_{\mu \in D} \mathcal{P}_{\mu}$ be the corresponding metrized line bundle.

**Theorem 5.1.** Let $C$ be a curve with a finite set of points $D$, defined over a number field $F$, and let $f$ be a Belyi map for $D$ as above. Choose metrics to obtain the associated metrized line bundle $D$. Let $S$ be a finite set of places of $F$, including the archimedean places. Assume the ABC conjecture with type function $\psi$. Then
\[ h_{\mathcal{K}+\mathcal{D}}(x) \leq N_{F(x),S,D}(x) + d(F(x)) + \sum_{v \in S} h_F(v) + \psi(h(f(x) : 1 - f(x) : 1)), \]

for every \( x \in C(\overline{F}) - f^{-1}[0, 1, \infty]. \)

The error term \( \psi(h) \) comes from the ABC conjecture. As explained in the text after Conjecture 1.2, this error term cannot in general be improved.

**Proof of Theorem 5.1.** Let \( x \) be a point of \( C(\overline{F}) \), defined over \( E = F(x) \), 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. Note that \( h_B(x) = h(f(x) : 1 - f(x) : 1) \), so that \( h(P) = h_B(x) \).

We estimate the radical of \( P \). Let \( w \) be a valuation of \( E \) so that its restriction to \( F \) is not equivalent to an element of \( S \). Thus \( w \) is nonarchimedean, and it contributes \( h_E(w) \) to the radical if and only if \( w(f(x)) < 0, w(1 - f(x)) < 0 \) or \( w(f(x)) > 0 \). In other words, \( w \) contributes to the radical only if \( \deg_w x^A, \deg_w x^B \) or \( \deg_w x^C \) is positive. Since \( \deg_w x^A = \sum_{\mu=1}^{i} e_{\mu} \deg_w x^P_{\mu}, \) and similarly for \( B \) and \( C \), it follows that \( \deg_w x^P_{\mu} > 0 \) for some \( \mu \) (\( \mu = 1, \ldots, k \)). Since \( \deg_w x^P_{\mu} \) is a multiple of \( h_E(w) \) for every \( \mu \) and in view of Remark 2.1, the contribution of \( w \) to the radical of \( P \) is bounded by

\[
 r_{E,w}(P) \leq \min \left\{ h_E(w), \sum_{\mu \in D} \deg_w x^P_{\mu}, \right\} + \sum_{\mu \notin D} \deg_w x^P_{\mu}. \tag{5.1}
\]

For \( w \in S \), we obtain the bound

\[
 r_{E,w}(P) \leq \sum_{\mu \notin D} \deg_w x^P_{\mu} + h_E(w).
\]

Since the metrics are positive, this also holds for the archimedean valuations. Adding these contributions, we find

\[
 r_E(P) \leq N_{E,S,D}^{(1)}(x) + \sum_{\mu \notin D} h_\mu(x) + \sum_{w \mid S} h_E(w),
\]

where \( h_\mu \) denotes the height with respect to \( P_{\mu} \), and \( w \mid S \) means that \( w \) is a valuation of \( E \) whose restriction to \( F \) is equivalent to an element of \( S \).

By the ABC conjecture with type \( \psi \) and (2.3), we obtain

\[
 h_B(x) \leq N_{E,S,D}^{(1)}(x) + \sum_{\mu \notin b} h_\mu(x) + \sum_{v \in S} h_F(v) + d(E) + \psi(h(P)).
\]

By (4.2), we have \( h_{\mathcal{K}+\mathcal{D}}(x) = h_B(x) - \sum_{\mu \notin b} h_\mu(x) \). Thus we obtain the radicalized Vojta height inequality, with \( \epsilon h_A(x) \) replaced by \( \psi(h(P)) + \sum_{v \in S} h_F(v) \). \( \square \)

**Acknowledgments**

This work was supported by a Scholarly Activities Summer Stipend from Utah Valley State College. I thank Serge Lang for explaining his conjecture and making the preprint [L05] available. He has been a great teacher to me and an example of how to do mathematical research.
References


