

# JOURNAL

de Théorie des Nombres  
de BORDEAUX

*anciennement Séminaire de Théorie des Nombres de Bordeaux*

Enrico BOMBIERI

**Problems and results on the distribution of algebraic points on algebraic varieties**

Tome 21, n° 1 (2009), p. 41-57.

<[http://jtnb.cedram.org/item?id=JTNB\\_2009\\_\\_21\\_1\\_41\\_0](http://jtnb.cedram.org/item?id=JTNB_2009__21_1_41_0)>

© Université Bordeaux 1, 2009, tous droits réservés.

L'accès aux articles de la revue « Journal de Théorie des Nombres de Bordeaux » (<http://jtnb.cedram.org/>), implique l'accord avec les conditions générales d'utilisation (<http://jtnb.cedram.org/legal/>). Toute reproduction en tout ou partie cet article sous quelque forme que ce soit pour tout usage autre que l'utilisation à fin strictement personnelle du copiste est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

cedram

*Article mis en ligne dans le cadre du*  
*Centre de diffusion des revues académiques de mathématiques*  
<http://www.cedram.org/>

# Problems and results on the distribution of algebraic points on algebraic varieties

par ENRICO BOMBIERI

RÉSUMÉ. Cet article est un exposé de plusieurs résultats sur la distribution des points algébriques sur les variétés algébriques.

ABSTRACT. This is a survey paper on the distribution of algebraic points on algebraic varieties.

## 1. Introduction

In this expository talk I will restrict my attention to questions about classical heights, in the simplest arithmetic and geometric cases. Finiteness theorems, notwithstanding their appeal and importance, will not be discussed here. Also, the reader will find no analysis of the counterpart of the topics discussed here in the setting of abelian or semi-abelian varieties, nor there will be a report of very recent work by several authors (some of it still unpublished) about the structure of multiplicative relations in subvarieties of linear tori, or their counterpart for subvarieties of semi-abelian varieties or Shimura varieties. Each of these topics would require at least one separate lecture to give it justice. The bibliography will be limited to papers directly relevant to what will be treated here and has no pretense of being complete.

The theory of heights has been a very powerful tool in studying the distribution of rational and algebraic points on algebraic varieties. The main questions relate to finiteness and density of solutions, sometimes restricted by integrality conditions or by asking that they belong to finitely generated groups, typically a group of  $S$ -units or a subgroup of a group variety.

The simplest notion of height is the Weil absolute logarithmic height in the standard model of projective space, defined as follows.

Let  $K$  be a number field. The height of a point  $\mathbf{x} = (x_0 : x_1 : \cdots : x_n)$  in  $\mathbb{P}^n$  (homogeneous coordinates) with  $x_i \in K$ ,  $i = 0, \dots, n$ , is given by

$$(1.1) \quad h(\mathbf{x}) = \sum_{v \in M_K} \max_i \log |x_i|_v$$

where  $M_K$  is the set of all places of  $K$  and  $|\cdot|_v$  is an associated absolute value normalized so that for  $x \in \mathbb{Q} \setminus \{0\}$  one has

$$(1.2) \quad |x|_v = \|x\|_p^{[K_v:\mathbb{Q}_p]/[K:\mathbb{Q}]}.$$

Here  $p \in M_{\mathbb{Q}}$  is the place of  $\mathbb{Q}$  with  $v|p$  and  $\|\cdot\|_p$  denotes the usual  $p$ -adic or archimedean absolute value of  $\mathbb{Q}$ .

This height is well defined, *i.e.* it does not depend on the representative  $\mathbf{x}$ . The height of an algebraic number  $x$  is the height of the point  $(1 : x) \in \mathbb{P}^1$ . With this normalization the height  $h(\mathbf{x})$  is invariant by base change (hence the name absolute height) and can be used in an algebraic closure of  $K$ .

The main properties of this height are:

- a) Product formula: For  $x \in K \setminus \{0\}$  it holds

$$(1.3) \quad \sum_{v \in M_K} \log |x|_v = 0.$$

- b) Multiplicativity: For  $\mathbf{x} \in \mathbb{P}^m$ ,  $\mathbf{y} \in \mathbb{P}^n$ , it holds

$$(1.4) \quad h(\mathbf{x} \otimes \mathbf{y}) = h(\mathbf{x}) + h(\mathbf{y}).$$

In particular,  $h(x^m) = |m| h(x)$ . Also, for  $x, y \in K \setminus \{0\}$  it holds

$$(1.5) \quad h(xy) \leq h(x) + h(y).$$

- c) For  $x_1, \dots, x_n \in K - \{0\}$  it holds

$$(1.6) \quad h(x_1 + \dots + x_n) \leq h(x_1) + \dots + h(x_n) + \log n.$$

If  $M_{naive}(\alpha)$  is the maximum of the absolute value of the coefficients of a minimal equation over  $\mathbb{Z}$  of a non-zero algebraic number  $\alpha$ , then one has

$$(1.7) \quad \deg(\alpha)h(\alpha) = \lim_{m \rightarrow \infty} \frac{1}{m} \log M_{naive}(\alpha^m).$$

Also of importance is the Mahler measure

$$(1.8) \quad M(P) = \exp \left( \frac{1}{2\pi} \int_0^{2\pi} \log |P(e^{i\theta})| d\theta \right)$$

of a polynomial  $P(x) \in \mathbb{C}[x]$ . By its very definition,

$$(1.9) \quad M(PQ) = M(P)M(Q).$$

By Jensen's formula, if  $f$  is the minimal polynomial over  $\mathbb{Z}$  of an algebraic number  $\alpha$ , it holds

$$(1.10) \quad \deg(\alpha)h(\alpha) = \log M(f).$$

It is well known that the Mahler measure of a polynomial with integer coefficients is comparable to the naive measure  $M_{naive}$  defined above and

in fact if  $P$  has degree  $d$  then

$$(1.11) \quad \left( \binom{d}{\lfloor d/2 \rfloor} \right)^{-1} M_{naive}(P) \leq M(f) \leq \sqrt{d+1} M_{naive}(P).$$

We note here Mahler's discriminant inequality. If  $f$  is the minimal polynomial over  $\mathbb{Z}$  of an algebraic number  $\alpha$  of degree  $d$  and  $D_f$  is the discriminant of  $f$ , it holds

$$(1.12) \quad \frac{1}{d} \log |D_f| \leq \log d + (2d - 2) h(\alpha).$$

It is also convenient to consider the exponential of the logarithmic height

$$(1.13) \quad H(x) = e^{h(x)}$$

and we refer to this as "Height" with the capital  $H$ , to distinguish it from the logarithmic height  $h$  (a suggestion of Paula Cohen Tretkoff).

Another absolute height is the Arakelov height induced by the Fubini-Study metric on  $\mathbb{P}^n$ . This is done by changing in the definition of the height the terms with  $v$  an archimedean valuation, replacing  $\max$  by an  $L^2$ -norm, namely

$$(1.14) \quad \frac{[K_v : \mathbb{R}]}{[K : \mathbb{Q}]} \log \|\mathbf{x}\|$$

where  $\|\cdot\|$  is the euclidean length of the vector  $\mathbf{x}$ .

Other heights can be defined in a similar way, replacing the archimedean norms by more general ones. The fundamental property of all these heights is that bounded sets in number fields are finite.

## 2. Varieties with many points

It is an interesting and difficult problem to determine asymptotically the number  $N(V(K), X; H)$  of points  $\mathbf{x} \in V(K)$  rational over the field  $K$ , with Height  $H(\cdot)$  bounded by  $X$ .

The simplest case  $V = \mathbb{P}^n$ ,  $K = \mathbb{Q}$ , counting the number of rational points with Weil Height at most  $X$  on  $\mathbb{P}^n/\mathbb{Q}$  ( $n \geq 1$ ) is classical and goes back to Dedekind and Weber; the asymptotic is  $N(\mathbb{P}^n(\mathbb{Q}), X; H) \sim X^{n+1}/\zeta(n+1)$ . The case of a general number field  $K$  was treated by Schanuel [23] in his thesis, who proved

$$(2.1) \quad N(\mathbb{P}^n(K), X; H) = c_n(K) X^{m(n+1)} + O\left(X^{m(n+1)-1} \log X\right)$$

where

$$(2.2) \quad c_n(K) = \frac{h_K R_K}{w_K \zeta_K(n+1)} \left( \frac{2^r (2\pi)^s}{\sqrt{D}} \right)^{n+1} (n+1)^{r+s-1}$$

with  $h_K$ ,  $R_K$ ,  $w_K$ , and  $D_K$  the class number, regulator, number of roots of unity, and absolute discriminant of  $K$ , and with  $r$  and  $s$  the number of

real and complex places of  $K$ . The term  $\log X$  can be omitted if either  $m$  or  $n$  are greater than 1.

W.M. Schmidt was able to treat the case when  $V$  is a Grassmannian [25].

Franke, Manin, and Tschinkel [15] treated the case of  $V$  a flag manifold, by considering the zeta function determined by the height of rational points and applying a tauberian theorem, thereby getting an asymptotic formula. J.L. Thunder [31] obtained asymptotics for flag varieties and quite general heights, with explicit error terms (valid for large height).

The Hardy and Littlewood circle method can be used with success when  $V$  is defined by equations of additive type, such as Fermat type hypersurfaces

$$(2.3) \quad \sum_{i=0}^n a_i x_i^d = 0.$$

If the coefficients  $a_i$  are non-zero rational integers, it turns out that the number of solutions in integers  $|x_i| \leq X$  is  $cX^{n+1-d} + o(X^{n+1-d})$  with  $c$  a constant depending on the vector  $\mathbf{a}$  of coefficients of (2.3), provided  $n$  is sufficiently large as a function of  $d$ . The constant  $c$  itself is the infinite product, over all primes  $p$  (including  $p = \infty$ ), of the  $p$ -adic volumes  $V_p$  of  $V/\mathbb{Q}_p$  with respect to the measure induced by the standard measure on the ambient affine space  $\mathbb{A}^{n+1}$ . Note that  $V_p = 0$  occurs if there is failure of the local  $p$ -adic solubility of the equation (2.3). The best known result today for large  $d$  is due to K.B. Ford [14], obtaining the expected asymptotic if  $n > d^2(\log d + \log \log d) + O(d^2)$ .

Notwithstanding the many results obtained by the circle method, some of them quite impressive, its weakness stems from the fact that its success usually depends on the additivity of the defining equations, because only in this case the underlying harmonic analysis is done in the much easier dimension 1. For general non-singular hypersurfaces of degree  $d$  in  $\mathbb{A}^{n+1}$ , the results obtained by the circle method require  $n$  to be very large, even if the final result has the same form as described before. At any rate, one should not expect a Hardy and Littlewood asymptotic formula to hold if  $n \leq d$ . Even if  $n \geq d + 1$ , such a formula need not hold in general.

The following example is due to Hooley [17]. Consider the equation

$$(2.4) \quad x_0^3 + x_1^3 + x_2^3 = x_3^3 + x_4^3 + x_5^3$$

to be solved in integers  $|x_i| \leq X$ , thus  $d = 3$  and  $n = 5$ . In this case, we have the obvious solutions where  $x_0, x_1, x_2$  are given arbitrary values and  $x_3, x_4, x_5$  are a permutation of  $x_0, x_1, x_2$ . The number of such solutions bounded by  $X$  is easily seen to be asymptotic to  $aX^3 + O(X^2)$  for a certain constant  $a > 0$ . Geometrically, the associated projective variety, of dimension 4, contains several projective planes defined over  $\mathbb{Q}$ , necessarily with many rational points. On the other hand, the Fourier analysis of the circle

method is insensitive to what happens on subsets of positive codimension of a variety, so the expected counting of solutions  $cX^3 + o(X^3)$  does not take into account the anomalous density of rational points on these planes. Hooley [17] has given convincing heuristic evidence that, in the example above, the correct counting of solutions is  $(a + c)X^3 + o(X^3)$ .

Thus the best one could hope for is that the number  $N(V(K), X; H)$  of rational points on a projective variety  $V$  should be a formula

$$(2.5) \quad N(V(K), X; H) = N(V^o(K), X; H) + N(W(K), X; H)$$

where  $V^o$  is a Zariski open subset of  $V$  and  $W$  is the closed complement of  $V^o$ , and where the number of points  $N(V^o(K), X; H)$  is determined asymptotically by the Hardy and Littlewood prediction using the circle method. Even in this form, other conditions will be needed, since the product  $\prod V_p$  of local volumes need not be absolutely convergent. Also, rational points on  $V$  may concentrate on a countable union subvarieties of lower dimension, with the union of these varieties being Zariski dense on  $V$ ; for example, a cone over an elliptic curve  $C/K$  and vertex defined over  $K$  has all its rational points over  $K$  lying on the countably many lines determined by the vertex and a rational point on  $C$ . Thus it is important to consider not just projective varieties  $V$ , but more generally constructible subsets  $U$  of  $V$ . Then  $N(U(K), X; H)$  can be defined as before.

Changing the projective embedding of a projective variety  $V$  will change the height function and therefore the counting function. In some cases, it is possible to define canonical polarizations of a variety. An interesting class of varieties for which this occurs is the class of Fano varieties, defined by the property that the anticanonical bundle  $O(-K_V)$  is ample. Batyrev and Manin [3] noticed that Fano varieties tend to have a large number of rational points over a number field  $K$  and formulated three main conjectures about their density. Since Fano varieties have a canonical polarization, it is natural to consider the problem of counting points relative to a height  $H_{O(-K_V)}(\cdot)$  determined by a choice of a metric on  $O(-K_V)$ . A naive elementary way of defining such a height is to consider a very ample multiple  $O(-mK)$ , a choice of a basis  $\mathbf{s}$  of  $\Gamma(V, O(-mK_V))$  yielding an embedding of  $V$  as a projective variety in a projective space  $\mathbb{P}^N$ , and a height  $h_{\mathbf{s}}(\cdot)$  on  $V$  induced by a height equivalent to the standard height in  $\mathbb{P}^N$ . Then the heights  $m^{-1}h_{\mathbf{s}}(\cdot)$  are all equivalent and can be taken as a height relative to the anticanonical bundle  $O(-K)$ . (Note however that a definition obtained directly from an adelic metric on  $O(-K_V)$  is both more general and more intrinsic.)

The basic invariant is

$$(2.6) \quad \tilde{\beta}(U(K); H) = \limsup_{X \rightarrow \infty} \frac{\log N(U(K), X; H)}{[K : \mathbb{Q}] \log X}$$

A non-empty subvariety  $W$  of  $V$  is called a *strong accumulator* if for every non-empty open subset  $U'$  of  $W$  defined over  $K$  there is a non-empty open subset  $U$  of  $V$ , also defined over  $K$ , such that

$$(2.7) \quad \tilde{\beta}(U'(K); H) > \tilde{\beta}(U(K); H).$$

Thus if  $V$  admits a strong accumulator over  $K$  the majority of points of  $V(K)$  must lie in a closed proper subset of  $V$ .

If we have only the weaker statement that

$$(2.8) \quad \limsup_{X \rightarrow \infty} \frac{N(U'(K), X; H)}{N(U(K), X; H)} > 0 \quad \text{and} \quad N(U(K), X; H) \rightarrow \infty$$

then  $W$  is called a *weak accumulator*.

The following conjectural example is characteristic of some of the problems which occur in presence of accumulators, already in the case of surfaces. The construction is classical (I learned it from Aldo Andreotti in 1964 but certainly it is much older). Consider a  $K3$  surface  $V$  defined over  $\mathbb{Q}$  such that the even quadratic form given by the self-intersection on the Néron–Severi group represents 2. Then every divisor  $D$  with  $D^2 = 2$  gives rise to a linear system  $|D|$  of curves of genus 2 without base points, of dimension 2. Let  $x$  be a general point on  $V$ . The linear system  $|D - x|$  of such curves passing through  $x$  has dimension 1 and any two irreducible curves  $D_1, D_2$ , in this linear system intersect in two points  $x, y$ , where  $y$  is independent of the choice of the curves (otherwise, fixing  $D_1$  and varying  $D_2$  we would obtain a dominant map  $\mathbb{P}^1 \rightarrow D_1$ , which is impossible because  $D_1$  has genus 2). Therefore, every divisor  $D$  (or  $-D$ ) of self-intersection 2 on a  $K3$  surface  $V$  determines an involution. The automorphism group of  $V$  generated by such involutions is infinite.

Now any  $K3$  surface contains a rational curve  $C$  (see the reference in [3]). Then the transforms of  $C$  by the automorphism group give rise to countably many rational curves on  $V$ , of unbounded degree. By going to a suitable finite extension  $L$  of the base field  $K$  and to an infinite subgroup of the automorphism group, we may also ensure that such a countable set  $\{C_i\}_{i \in \mathbb{N}}$  can be achieved over  $L$ . Each rational curve obtained in this way will have  $\tilde{\beta}(C_i(L); H) > 0$ , but  $\tilde{\beta}(C_i(L); H) \rightarrow 0$  as the degree of  $C_i$  increases (this follows from the easy fact that the number of rational curves on  $V$  of fixed degree is bounded). It is expected that this construction will give a sequence of open subsets  $U_i$  of  $V$  such that  $\tilde{\beta}(U_i(L); H) \rightarrow 0$  and therefore  $V$  will admit a countable infinite sequence of strong accumulators.

There is an important geometric invariant defined by Batyrev and Manin which plays a big role in the study of the growth of  $N(U(K), X; H)$ . Let  $N_{eff}^1(V)$  be the closed cone generated by all line bundles  $O(D)$  determined

by effective divisors  $D$  on  $V$ . Then define for a line bundle  $L$  on  $V$

$$(2.9) \quad \alpha(L) = \inf\{\gamma \in \mathbb{R} : \gamma L + O(K_V) \in N_{eff}^1(V)\}$$

if it exists.

The first conjecture of Batyrev and Manin (stated here in a slightly weaker form for the invariant  $\tilde{\beta}$  rather than their invariant  $\beta$ ) states that for  $L$  ample

$$(2.10) \quad \inf_{U \subset V} \tilde{\beta}(U(K); H_L) \leq \max(\alpha(L), 0)$$

where  $U$  ranges over the non-empty open subsets of  $V$ . In all cases where one has been able to compute the function  $\tilde{\beta}$ , we have that if  $V \subset \mathbb{P}^N$  is a non-singular complete intersection of type  $(d_1, \dots, d_k)$  then

$$(2.11) \quad \inf_U \tilde{\beta}(U(K); H_{O(1)}) \leq \max(N + 1 - \sum_{i=1}^k d_i, 0) = \alpha(O(1))$$

where the inf is taken over all non-empty open subsets of  $V$ , with equality if  $K \supset K_0$  for a suitable number field  $K_0$ . This is compatible with the circle method philosophy and with the well-known Lang conjecture on rational points on varieties of general type.

The situation for asymptotics is more complicated. Franke, Manin, and Tschinkel formulated the interesting conjecture that for a Fano variety (*i.e.* with ample anticanonical bundle) a formula

$$(2.12) \quad N(V(K), X; H_{O(-K_V)}) \sim CN (\log N)^{b-1}$$

holds for some constant  $C$  and  $C > 0$  if  $K \supset K_0$  with  $K_0$  a sufficiently large number field. In many cases where one could prove an asymptotic formula, it turned out that  $b = \text{rank} NS(V/K)$ , the Néron–Severi group of  $V/K$ .

We refer to (2.12) with  $b = \text{rank} NS(V/K)$  as the Manin conjecture.

Batyrev and Tschinkel later exhibited several examples for which the above conjecture does not hold. Here is the simplest, see also Peyre [20]. Consider the equation

$$(2.13) \quad \sum_{i=0}^3 x_i y_i^3 = 0$$

This defines a non-singular hypersurface  $V$  in  $\mathbb{P}^3 \times \mathbb{P}^3$ . The canonical bundle of  $V$  is  $O(K_V) = O(-3, -1)$ . The Néron–Severi group has rank 2 and no torsion and the conjecture predicts

$$(2.14) \quad N(U(K), X; H_{O(-K_V)}) \sim CN \log N.$$

Since  $O(K_V) = O(-3, -1)$ , the height associated to the anticanonical bundle is simply

$$(2.15) \quad H_{O(-K_V)}(\mathbf{x}, \mathbf{y}) = H(\mathbf{x})^3 H(\mathbf{y})$$



where in the right-hand side we have the Weil height in projective space. For fixed  $\mathbf{x}$  with  $x_0x_1x_2x_3 \neq 0$ , the subvariety  $V_{\mathbf{x}}$  of  $V$  defined by  $\sum x_i y_i^3 = 0$  is a non-singular cubic surface. If all  $x_i$  are cubes and  $K$  contains the third roots of unity, then the Néron–Severi group of  $V_{\mathbf{x}}/K$  has rank 7. Again by the conjecture, we would have

$$(2.16) \quad N(U_{\mathbf{x}}(K), X; H_{O(-K_{V_{\mathbf{x}}})}) \sim CN (\log N)^6$$

for any open subset of  $V_{\mathbf{x}}$ .

On the other hand, since  $O(-K_{V_{\mathbf{x}}}) = O(1)$  the anticanonical height on the Fano cubic surface is proportional to the induced Weil height. Since the points  $\mathbf{x}^3$  with  $\mathbf{x} \in K^4$  are dense in  $\mathbb{P}^3$ , this implies that for any open set  $U \subset V$  it holds

$$(2.17) \quad N(U(K), X; H_{O(-K_V)}) \gg N (\log N)^6$$

in contradiction with (2.14).

The failure in this example is due to the presence of infinitely many weak accumulators, defined over  $K$ , in the variety  $V$ .

The constant in the Manin conjecture should be compatible with what is expected from the circle method (*i.e.* the Tamagawa number coincides with the Hardy–Littlewood singular series) and E. Peyre [21] has made a deep analysis of the situation. The conjecture has been proved in many cases, such as classes of toric varieties and many Del Pezzo surfaces. We refer to [21, 11, 22, 20] for an overview of the ample literature on the subject.

### 3. A special case: The cubic threefold

A very interesting variety is a non-singular cubic threefold  $V$  in  $\mathbb{P}^4$ . Its canonical bundle is the restriction of  $O(-2)$ , hence  $V$  is a Fano threefold with very ample anticanonical bundle. Manin’s conjecture predicts an asymptotic

$$(3.1) \quad N(V(K), X; H) \sim c(V, K)X^{2d}$$

where  $H$  is induced by the standard height on  $\mathbb{P}^4$  and  $d = [K : \mathbb{Q}]$  is the degree of the extension  $K/\mathbb{Q}$ .

The solution of the problem of estimating  $N(V(K), X; H)$  with  $V$  a cubic threefold or fourfold is of particular importance in the study of cubic Weyl sums, namely

$$(3.2) \quad S(\alpha) = \sum_{n=1}^N e^{2\pi i \alpha f(n)}$$

where  $f(x)$  is a monic cubic polynomial with real coefficients.

The simplest case  $f(n) = n^3$  occurs in the study of Waring’s problem for cubes and estimates for the moments of the associated Weyl sum are of importance for Waring’s problem of decomposing an integer into a sum of

cubes of positive integers. For the sixth moment, there is the conjectural asymptotic estimate

$$(3.3) \quad \int_0^1 |S(\alpha)|^6 d\alpha \sim cN^3$$

for some positive constant  $c$ , which would follow from the Manin conjecture for the cubic fourfold (2.4). (The integral is the number of integer solutions in the positive octant  $0 < x_i \leq N$ .)

So far, no improvement in the exponent over L.K.Hua's estimate of 1938

$$(3.4) \quad \int_0^1 |S(\alpha)|^6 d\alpha \ll N^{3+\frac{1}{2}+\epsilon}$$

has been obtained. Hua's exponent is the critical one, any lowering of it would have interesting consequences.

According to the Manin conjecture, a non-singular cubic threefold  $V/K$  should satisfy an asymptotic relation

$$(3.5) \quad N(V(K), X; H) \sim C(V, K)X^{2d}$$

where  $d = [K : \mathbb{Q}]$ . The cubic threefold should not have accumulators but it may very well contain lines contributing an amount of order  $X^2$  to the counting of points. A more precise analysis is as follows.

If  $V/\mathbb{C}$  is a non-singular cubic threefold, there are exactly 6 lines through a generic point  $P$  of  $V$  (*i.e.* for  $P$  in a dense Zariski open subset). If  $V$  is defined over a number field  $K$ , and  $P \in V(K)$ , some of these lines may contain another point  $Q$  rational over  $K$ , giving rise to a subspace  $\mathbb{P}^1/K$  contained in  $V$ . By Schanuel's theorem, such a projective space contains  $\sim cX^{2d}$  points rational over  $K$  and Height at most  $X$ , for some positive constant  $c$ . Therefore, the contribution to  $N(V(K), X; H)$  due to rational lines is at least of the order predicted by the Manin conjecture. If  $V$  contains infinitely many lines defined over  $K$ , which certainly occurs in some cases, it is a non-trivial problem to show that the overall contribution due to such lines is not too large.

The lines of  $V$  are naturally parametrized by a projective surface  $\Sigma/K$  in the Grassmannian  $G(2, 5)$  of lines in projective space  $\mathbb{P}^4$ . The surface  $\Sigma$  is absolutely irreducible, with very ample canonical bundle yielding back the embedding of  $\Sigma$  in the Grassmannian, *via* a corresponding canonical embedding in  $\mathbb{P}^9 \supset G(2, 5)$ . The standard embedding of  $G(2, 5)$  into  $\mathbb{P}^9$  given by the Grassmann coordinates defines, by restriction of the standard Arakelov Height in  $\mathbb{P}^9$ , an Arakelov Height on  $G(2, 5)$  and, again by restriction, an Arakelov Height on  $\Sigma$ .

Let  $L(V/K) = \overline{\text{lines}(V)/K}$  denote the Zariski closure in  $V$  of the subset of rational lines in  $V$  defined over  $K$ . The interest of this set is twofold. On the one hand, it gives a sizable contribution to the total counting of rational points of  $V/K$ . On the other hand, on the complement  $U$  of  $L(V/K)$

in  $V$  there is a well-defined composition law for rational points: Given two rational points  $P$  and  $Q$  in  $U(K)$ , the line through these points is not contained in  $V$  by the very definition of  $L(V/K)$ , hence the residual intersection  $R$  with the cubic is a well-defined rational point  $R$ , again in  $U(K)$ . The hope is that a clever use of this composition law may help in studying the distribution, and the density, of points in  $U(K)$ .

The following theorem was proved in [5], see Th.11.10.11 there and its proof.

**Theorem 3.1.** *The following statements hold:*

- a) *For every number field  $K$ , the closed set  $L(V/K)$  is the union of a finite set of lines and of not more than 30 cubic cones over elliptic curves, defined over  $K$ .*
- b) *Let  $N_{Ar}(L(V/K), X)$  be the number of rational points of  $L(V/K)$  defined over  $K$ , of Arakelov Height at most  $X$ . Let  $d = [K : \mathbb{Q}]$  be the degree of the number field  $K$ . Then for every fixed  $\varepsilon > 0$  and  $X \geq 1$  it holds*

$$(3.6) \quad N_{Ar}(L(V/K), X) = c_2 \gamma X^{2d} + O(X^{2d-1+\varepsilon})$$

where  $c_2$  is Schanuel's constant for a projective line (relative to the standard Arakelov Height)

$$(3.7) \quad c_2 = \frac{h_K R_K}{w_K \zeta(2)} \frac{2^{d-1} \pi^d}{D_K}$$

and  $\gamma$  is given by the convergent series

$$(3.8) \quad \sum_{Q \in \Sigma(K)} H_{Ar}(Q)^{-d}.$$

The idea of the proof is as follows. As shown in [9, 6], the surface  $\Sigma$  (already studied by Fano) parametrizing the lines of  $V$  is immersed in its Albanese variety, which is an abelian variety of dimension 5. Rational lines correspond to rational points on  $\Sigma$ . By Faltings's Big Theorem [13],  $\Sigma(K)$  is a finite union

$$(3.9) \quad \Sigma(K) = \bigcup \{x_i + B_i(K)\}$$

of the rational points of translates  $x_i + B_i \subset \Sigma$  of finitely many abelian subvarieties of the ambient abelian variety. Since  $\Sigma$  is of general type it contains only finitely many elliptic curves  $x_i + E_i$ . Thus the rational lines on  $V/K$  are, apart a finite set, parametrized by the rational points  $x_i + E_i(K)$ . The contribution of each line to the counting of rational points is clearly of order  $X^2$  and if we have infinitely many rational lines, which is the case if  $E_i(K)$  has positive rank, it is no longer sufficient to count naively an estimate based only on the main term of the asymptotic formula, the error terms also must be controlled.

It turns out that the totality of lines parametrized by  $x_i + E_i$  is a cone with rational vertex and with directrix an elliptic curve isomorphic to  $E_i$ . The maximal number of such cones is 30, see H. Clemens and P.A. Griffiths [9]. The height of the line itself is the height of the corresponding point in  $\Sigma$ , hence the height of any other rational point on this line is, up to a constant factor, at least the height of this point. Since the height of rational points on an elliptic curve grows very fast, only  $O((\log X)^{r/2})$  points may contribute to the counting (here  $r$  is the rank of the Mordell–Weil group of the elliptic curve). In order to conclude the proof, one needs an asymptotic bound for the number of points of bounded height on a line in  $\mathbb{P}^4$ , with an explicit error term valid in all ranges of  $X$ . An explicit estimate for the number of points in linear subspaces of projective space is in Thunder’s paper [31], but it cannot be applied directly because it is valid only for sufficiently large  $X$ . However, a simple additional argument using geometry of number suffices for filling the gap in a satisfactory manner (see Christensen and Gubler [8]), yielding at the end the result we wanted.

The situation is better for the cubic complete intersection given by the equations

$$(3.10) \quad \begin{aligned} x_0^3 + x_1^3 + x_2^3 &= x_3^2 + x_4^3 + x_5^3 \\ x_0 + x_1 + x_2 &= x_3 + x_4 + x_5 \end{aligned}$$

which is a linear section of the cubic fourfold (2.4). This is associated to the sixth moment of the Weyl sum  $\sum e^{2\pi i(\alpha n^3 + \beta n)}$  taken with respect to  $\alpha$  and  $\beta$ . The variety  $V$  so defined is a cubic threefold with 10 double nodes and with 15 planes.

In 1995 Vaughan and Wooley [32] considered the positive integer solutions of the above system. In their notation, solutions in the union of the 15 planes form the *diagonal set* of solutions. They proved that the diagonal of  $V/\mathbb{Q}$  is a strong accumulator and that outside of the diagonal the number of solutions has the precise order of magnitude

$$(3.11) \quad N(U(\mathbb{Q}), X; H_{O(1)}) \asymp X^2(\log X)^5.$$

This agrees with the Manin conjecture and the components of the diagonal are the only strong accumulators. Very recently, in a remarkable paper, R. de la Bréteche [11] was able to prove the expected asymptotics for  $N(V(\mathbb{Q}), X; H)$ . The proof, which starts from the approach propounded by Vaughan and Wooley, requires a combination of deep techniques from algebraic geometry and analytic number theory and is rather delicate.

Results of this type have important applications to additive number theory and to the theory of trigonometric sums.

#### 4. The Northcott property

The first basic result about heights was proved by Northcott, who showed that there are only finitely many algebraic numbers of bounded degree and bounded height. This is immediate by looking at the minimal polynomial over  $\mathbb{Z}$  of a given algebraic number. In spite of its simplicity, Northcott's theorem is very useful.

We say that a set  $\mathcal{A}$  of algebraic numbers has the *Northcott property* if bounded subsets for the height function above are finite. Thus the set  $\mathcal{A}(m)$  of all algebraic numbers of degree at most  $m$  has the Northcott property.

This raises the question of studying the distribution of points of degree  $m$  over a number field  $K$ , lying on a variety  $V$ . We denote by  $N(m, V/K, X)$  the number of such points of degree  $m$  over  $K$ . By Northcott's theorem,  $N(m, V/K, X)$  is a finite quantity. The determination of asymptotics for  $N(m, V/K, X)$  is a difficult problem.

The case  $m = 2$ ,  $V = \mathbb{P}^n$ ,  $K = \mathbb{Q}$ , was solved by W.M. Schmidt [26], obtaining

$$(4.1) \quad N(2, \mathbb{P}^n/\mathbb{Q}, X) = \begin{cases} c(2, n)X^{2(n+1)} + O(X^{2n+1}) & \text{if } n \geq 3, \\ c(2, 2)X^6 \log X + O(X^6 \sqrt{\log X}) & \text{if } n = 2, \\ c(2, 1)X^6 + O(X^5 \log X) & \text{if } n = 1; \end{cases}$$

the constants  $c(2, n)$  are given explicitly, with  $c(2, 1) = 8/\zeta(3)$  and  $c(2, 2) = (48+4\pi^2)/\zeta(3)^2$ . This was followed by X. Gao [16] who obtained in his thesis an asymptotic

$$(4.2) \quad N(m, \mathbb{P}^n/\mathbb{Q}, X) \sim c(m, n)X^{m(n+1)}$$

for  $n \geq m \geq 3$  and also showed that if  $n < m$  then the precise order of magnitude of  $N(m, \mathbb{P}^n, X)$  is  $X^{m(m+1)}$ .

D. Masser and J.D. Vaaler [19] have established for  $m \geq 2$  an asymptotic

$$(4.3) \quad N(m, \mathbb{P}^1/K, X) \sim c(m, 1; K)X^{dm(m+1)}$$

with an explicit error term, corresponding to Schanuel's if  $m = 1$ . The intermediate cases  $1 < n \leq m$  remain open.

Let  $K$  be a number field and denote by  $K^{(d)}$  the compositum of all extension fields  $L/K$  of relative degree at most  $d$ . It was proved in [7] (see also [5], Ch.4, for an exposition) that  $K^{(2)}$  again has the Northcott property, and, more generally, the maximal abelian subfield  $K_{ab}^{(d)}$  of  $K^{(d)}$  also has the Northcott property. This provides examples of extensions of  $K$  of infinite degree with the Northcott property.

It remains an open problem to determine whether the Northcott property holds for  $K^{(d)}$  if  $d \geq 3$  and, more generally, to determine workable conditions for its validity.

## 5. Small points

In 1933 D.H. Lehmer raised the question whether there was a positive lower bound  $c > 1$  for the Mahler measure of a non-zero algebraic number not a root of unity. This statement is nowadays known as the Lehmer conjecture and it has proved to be a very stubborn problem to solve. Interestingly enough, the Lehmer conjecture admits many variants both for classical and more sophisticated heights as well, such as the Néron–Tate height on abelian varieties. It has become clear that the Lehmer conjecture is the prototype of a group of basic questions on algebraic numbers and algebraic varieties over number fields and, as such, it deserves a lot of attention.

For the normalized Weil height, the Lehmer conjecture states that there is an absolute constant  $c > 0$  such that

$$(5.1) \quad \deg(\alpha) h(\alpha) \geq c$$

unless  $\alpha$  is zero or a root of unity. The simple example  $\alpha = 2^{1/d}$  shows that if the conjecture is true we must have  $c \leq \log 2$ . In the other direction, it is easy to show that  $h(\alpha) \geq (\log 2)/\deg(\alpha)$  unless  $\alpha \neq 0$  is a unit (*i.e.* both  $\alpha$  and  $1/\alpha$  are algebraic integers).

C.J. Smyth [27] has shown that the minimum of  $M(\alpha)$  when the minimal polynomial of  $\alpha \neq 0$  is not reciprocal occurs for the cubic number with minimal polynomial  $x^3 - x - 1$ , hence the Mahler conjecture holds if  $\alpha$  is not reciprocal.

The best general lower bound is due to Dobrowolski [12]. It states that there is an absolute constant  $c > 0$  such that for  $\alpha \neq 0$  an algebraic number of degree  $d$  and not a root of unity it holds

$$(5.2) \quad h(\alpha) \geq \frac{c}{d} \left( \frac{\log \log d}{\log d} \right)^3.$$

The constant  $c$  can be taken as  $c = 1 + \varepsilon$  if we assume  $d$  sufficiently large as function of  $\varepsilon > 0$ . R. Louboutin [18] improved the last statement to  $c = \frac{9}{4} + \varepsilon$  for large  $d$ , by a variant of Dobrowolski method. It was remarked by Vaaler and the author (unpublished) that Dobrowolski's argument can be refined to Louboutin's result by using the full force of Siegel's lemma (*i.e.* the successive minima) in the proof.

By Mahler inequality, if  $\alpha \neq 0$  has height  $h(\alpha) < \kappa/d$  then the minimal equation  $f$  of  $\alpha$  has discriminant

$$(5.3) \quad |D_f| < (e^{2\kappa}d)^d$$

which is relatively small. Although it is known that there is an absolute constant  $C > 1$  such that there are infinitely many number fields  $K$  with

$|D_K| \leq C^{[K:\mathbb{Q}]}$  (the proof of this is not easy), the discriminant of an irreducible equation  $f$  can be much larger than the discriminant of the associated number field. To the best of my knowledge, the problem of determining

$$(5.4) \quad C^* := \liminf_{\deg(f) \rightarrow \infty} \frac{\log |D_f|}{\deg(f)}$$

when  $f$  ranges over all irreducible polynomials with integer coefficients, and in particular the question whether  $C^*$  is a finite constant or not, remains open.

There is a very interesting multidimensional version of the Dobrowolski theorem, due to Amoroso and David [1], which has found several applications.

**Theorem 5.1.** *There is a positive constant  $c(n) > 0$  with the following property. Let  $\alpha_1, \dots, \alpha_n$  be non-zero multiplicatively independent algebraic numbers and let  $D = [\mathbb{Q}(\alpha_1, \dots, \alpha_n) : \mathbb{Q}]$ . Then*

$$(5.5) \quad h(\alpha_1) \cdots h(\alpha_n) \geq \frac{c(n)}{D} (\log(3D))^{-n\kappa(n)}$$

where  $\kappa(n) = (n+1)(n+1)!^n - 1$ .

Moreover, the degree  $D$  can be replaced by the smallest degree of a hypersurface in  $\mathbb{A}^n$ , defined over  $\mathbb{Q}$ , containing the point  $(\alpha_1, \dots, \alpha_n)$ .

The proof of this theorem is much harder than the proof of Dobrowolski's theorem and requires deep tools from arithmetic geometry. A consequence of this theorem is the validity of Lehmer's conjecture for  $\alpha$  if  $\mathbb{Q}(\alpha)$  is a Galois extension of  $\mathbb{Q}$ .

A nice theorem of Amoroso and Dvornicich [2] states that if  $K$  is an abelian extension of  $\mathbb{Q}$  and  $\alpha \in K$  is not 0 or a root of unity then

$$(5.6) \quad h(\alpha) \geq \frac{\log 5}{12}.$$

Thus the same lower bound holds for the infinite cyclotomic extension of  $\mathbb{Q}$  generated by all roots of unity. This gap phenomenon with the height is quite remarkable and we refer to it as the *Bogomolov property*.

Another infinite extension of  $\mathbb{Q}$  with the Bogomolov property is the field of all totally real numbers. Here the sharp bound was found by A. Schinzel [24], who proved that if  $\alpha \neq 0$  is totally real and not a root of unity then

$$(5.7) \quad h(\alpha) \geq \frac{1}{2} \log \left( \frac{1 + \sqrt{5}}{2} \right).$$

The minimum is attained if  $\alpha = \pm(1 \pm \sqrt{5})/2$ .

C.J. Smyth [29] observed that if  $\alpha$  is totally real then  $\beta - \beta^{-1} = \alpha$  yields another totally real number of degree at most  $2 \deg(\alpha)$ . Starting with  $\alpha = 1$ ,

he obtained a sequence  $\alpha_1, \alpha_2, \dots$  of totally real numbers with small height, accumulating at  $\lambda = \lim h(\alpha_n) = 0.2732831\dots$

The minimum  $h(\alpha_1)$  is isolated and so are the next three values at  $h(\alpha_2)$ ,  $h(\alpha_3)$ , and  $h(2 \cos(2\pi/7))$ . It is conjectured that  $\lambda$  is the first limit point of the set of values of  $h(\alpha)$  for totally real  $\alpha$ .

Let  $p$  be a prime number. We say that  $\alpha$  is totally  $p$ -adic if the rational prime  $p$  splits completely in the field  $\mathbb{Q}(\alpha)$ . Then again the field  $L$  of totally  $p$ -adic numbers has the Bogomolov property, see [7].

The Bogomolov property is intimately related to the general question of the distribution of points of small height on varieties. The basic result here is the uniform distribution with respect to Galois orbits. The statement for the standard height is due to Yu.F. Bilu [4]. We need a definition.

For an algebraic number  $\xi$ , let  $\delta_\xi$  be the uniform probability measure supported by the Galois orbit of  $\xi$ , namely

$$(5.8) \quad \delta_\xi = \frac{1}{[\mathbb{Q}(\xi) : \mathbb{Q}]} \sum_{\sigma: \mathbb{Q}(\xi) \rightarrow \mathbb{C}} \delta_{\sigma\xi}$$

where  $\delta_a$  is the Dirac measure at  $a$ . Bilu's theorem states

**Theorem 5.2.** *Let  $\{\xi_i\}_{i \in \mathbb{N}}$  be an infinite sequence of distinct non-zero algebraic numbers such that  $h(\xi_i) \rightarrow 0$  as  $i \rightarrow \infty$ . Then the sequence of probability measures  $\{\delta_i\}_{i \in \mathbb{N}}$  converges in the weak\*-topology to the uniform probability measure  $\mu_{\mathbb{T}} = d\theta/(2\pi)$  on the unit circle  $\mathbb{T} = \{e^{i\theta} : 0 \leq \theta < 2\pi\}$  in  $\mathbb{C}$ .*

Bilu's elegant proof exploits in a clever way Mahler's discriminant inequality (1.12).

This theorem can be easily extended to the higher dimensional setting. The corresponding result on an abelian variety is much harder to prove and was obtained at the same time by L. Szpiro, E. Ullmo, S. Zhang [30] in a well-known paper, proving a conjecture of Bogomolov that algebraic points on a curve of genus at least 2 had the gap property with respect to the Néron–Tate height. The corresponding result for the Weil height, also in the case of algebraic varieties, had been proved earlier by S. Zhang [33] in a landmark paper. Bilu's theorem yields, as an easy consequence, Zhang's theorem that the Bogomolov property holds for the set of algebraic points on the open set  $X^*$  of a variety  $X \subset \mathbb{G}_m^N$  obtained by removing all torus cosets of positive dimension contained in  $X$ .

Good quantitative estimates for lower bounds in these problems are hard to obtain and require sophisticated methods of arithmetic geometry. For problems about Mahler's measure, we refer to the excellent survey by C. Smyth [28]. We also refer to P. Philippon and S. David [10] for the best results obtained on the generalized Lehmer conjecture and related topics.



## References

- [1] F. AMOROSO, S. DAVID, *Le problème de Lehmer en dimension supérieure*. J. reine angew. Math. **513** (1999), 145–179.
- [2] F. AMOROSO, R. DVORNICICH, *A lower bound for the height in an abelian extension*. J. Number Th. **80** (2000), 260–272.
- [3] V.V. BATYREV, YU.I. MANIN, *Sur le nombre des points rationnels de hauteur bornée des variétés algébriques*. Math. Annalen **286** (1980), 27–43.
- [4] YU.F. BILU, *Limit distribution of small points on algebraic tori*. Duke Math. J. **89** (1997), 465–476.
- [5] E. BOMBIERI, W. GUBLER, *Heights in Diophantine Geometry*, Cambridge Univ. Press 2006, xvi+652pp.
- [6] E. BOMBIERI, H.P.F. SWINNERTON-DYER, *On the local zeta function of a cubic threefold*. Ann. Scuola Norm. Super. Pisa Sci. Fis. Mat. (3) **21** (1967), 1–29.
- [7] E. BOMBIERI, U. ZANNIER, *A note on heights in infinite extensions of  $\mathbb{Q}$* . Atti Accad. Naz. Lincei Cl. Sci. Fis. Mat. Natur. Rend. Lincei (9) Mat. Appl. **12** (2001), 5–14.
- [8] C. CHRISTENSEN, W. GUBLER, *Der relative Satz von Schanuel*. Manuscripta Math. **126** (2008), 505–525.
- [9] H. CLEMENS, P.A. GRIFFITHS, *The intermediate Jacobian of the cubic threefold*. Ann. of Math. (2) **95** (1972), 281–356.
- [10] S. DAVID, P. PHILIPPON, *Minorations des hauteurs normalisées des sous-variétés des tores*. Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **28** (1999), 489–543. Errata: *ibid.* (4) **29** (2000), 729–731.
- [11] R. DE LA BRETÈCHE, *Points rationnels sur la cubique de Segre*. Proc. London Math. Soc. (3) **95** (2007), 69–155.
- [12] E. DOBROWOLSKI, *On a question of Lehmer and the number of irreducible factors of a polynomial*. Acta Arith. **XXXIX** (1979), 391–401.
- [13] G. FALTINGS, *Diophantine approximation on abelian varieties*. Ann. of Math. (2) **133** (1991), 549–576.
- [14] K.B. FORD, *New estimates for mean values of Weyl sums*. Internat. Math. Res. Notices (1995), 155–171.
- [15] J. FRANKE, YU.I. MANIN, Y. TSCHINKEL, *Rational points of bounded height on Fano varieties*. Inventiones Math. **95** (1989), 421–435. Erratum: “Rational points of bounded height on Fano varieties”. *ibid.* **105** (1990), 463.
- [16] X. GAO, *On Northcott’s theorem*. Ph. D. Thesis, University of Colorado (1995).
- [17] C. HOOLEY, *On some topics connected with Waring’s problem*. J. Reine Angew. Math. **369** (1986), 110–153.
- [18] R. LOUBOUTIN, *Sur la mesure de Mahler d’un nombre algébrique*. C. R. Acad. Sci. Paris Sér. I Math. **296** (1983), 707–708.
- [19] D. MASSER, J.D. VAALER, *Counting algebraic numbers with large height*. Trans. Amer. Math. Soc. **359** (2007), 427–445.
- [20] E. PEYRE, *Points de hauteur bornée et géométrie des variétés (d’après Y. Manin et al.)*. Séminaire Bourbaki, Vol. 2000/2001, 323–344.
- [21] E. PEYRE, *Hauteurs et mesure de Tamagawa sur les variétés de Fano*. Duke Math. J. **79** (1995), 101–218.
- [22] P. SALBERGER, *Tamagawa measures on universal torsors and points of bounded height on Fano varieties*. Astérisque **251** (1998), 91–258.
- [23] S.H. SCHANUEL, *Heights in number fields*. Bull. Soc. Math. France **107** (1979), 433–449.
- [24] A. SCHINZEL, *On the product of conjugates outside the unit circle of an algebraic integer*. Acta Arith. **XXIV** (1973), 385–399. Addendum *ibid.* **XXVI** (1974/75), 329–331.
- [25] W.M. SCHMIDT, *Asymptotic formulae for point lattices of bounded determinant and subspaces of bounded height*. Duke Math. J. **35** (1968), 327–339.
- [26] W.M. SCHMIDT, *Northcott’s theorem on heights II. The quadratic case*. Acta Arith. **LXX** (1995), 343–375.

- [27] C.J. SMYTH, *On the product of conjugates outside the unit circle of an algebraic integer*. Bull. London Math. Soc. **3** (1971), 169–175.
- [28] C. SMYTH, *The Mahler measure of algebraic numbers: A survey*. In Number Theory & Polynomials conference proceedings, London Math. Soc., Lecture Note Ser. **352**, Cambridge Univ. Press, Cambridge 2008, 322–349.
- [29] C.J. SMYTH, *On the measure of totally real algebraic numbers, I*. J. Austral. Math. Soc. Ser. A **30** (1980/81), 137–149; *II*, Math. of Comp. **37** (1981), 205–208.
- [30] L. SZPIRO, E. ULLMO, S. ZHANG, *Equirépartition des petits points*. Inventiones Math. **127** (1997), 337–347.
- [31] J.L. THUNDER, *Asymptotic estimates for the number of rational points of bounded height on flag varieties*. Compos. Math. **88** (1993), 155–186.
- [32] R.C. VAUGHAN, T.D. WOOLEY, *On a certain nonary cubic form and related equations*. Duke Math. J. **80** (1995), 669–735.
- [33] S. ZHANG, *Positive line bundles on arithmetic surfaces*. Ann. of Math. (2) **136** (1995), 569–587.

Enrico BOMBIERI  
Institute for Advanced Study  
1 Einstein Drive  
Princeton, NJ 08540, USA  
*E-mail*: [eb@ias.edu](mailto:eb@ias.edu)  
*URL*: <http://www.math.ias.edu>