BIRKHÄUSER

Progress in Mathematics
Volume 277

Timothy D. Browning

Quantitative Arithmetic of Projective Varieties

Birkhäuser
Basel · Boston · Berlin

Author:

Timothy D. Browning
School of Mathematics
University of Bristol
Bristol BS8 1TW
United Kingdom
e-mail: t.d.browning@bristol.ac.uk

2000 Mathematics Subject Classification: 11D45, 11G35, 11D72, 11E76, 11P55, 14G05, 14G10, 14G25

Library of Congress Control Number: 2009934090

Bibliographic information published by Die Deutsche Bibliothek.
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data is available in the Internet at http://dnb.ddb.de

ISBN 978-3-0346-0128-3 Birkhäuser Verlag AG, Basel · Boston · Berlin

Basel · Boston · Berlin
P.O. Box 133, CH-4010 Basel, Switzerland
Part of Springer Science+Business Media
Printed on acid-free paper produced from chlorine-free pulp. TCF ∞
Printed in Germany

ISBN 978-3-0346-0128-3 e-ISBN 978-3-0346-0129-0

9 8 7 6 5 4 3 2 1 www.birkhauser.ch

Ferran Sunyer i Balaguer (1912–1967) was a self-taught Catalan mathematician who, in spite of a serious physical disability, was very active in research in classical mathematical analysis, an area in which he acquired international recognition. His heirs created the Fundació Ferran Sunyer i Balaguer inside the Institut d'Estudis Catalans to honor the memory of Ferran Sunyer i Balaguer and to promote mathematical research.

Each year, the Fundació Ferran Sunyer i Balaguer and the Institut d'Estudis Catalans award an international research prize for a mathematical monograph of expository nature. The prize-winning monographs are published in this series. Details about the prize and the Fundació Ferran Sunyer i Balaguer can be found at

`http://www.crm.es/FSBPrize/ffsb.htm`

This book has been awarded the Ferran Sunyer i Balaguer 2009 prize.

The members of the scientific commitee of the 2009 prize were:

Hyman Bass
University of Michigan

Núria Fagella
Universitat de Barcelona

Joan Verdera
Universitat Autònoma de Barcelona

Alan Weinstein
University of California at Berkeley

Ferran Sunyer i Balaguer Prize winners since 1998:

1999 Patrick Dehornoy
Braids and Self-Distributivity, PM 192

2000 Juan-Pablo Ortega and Tudor Ratiu
Hamiltonian Singular Reduction, PM 222

2001 Martin Golubitsky and Ian Stewart
The Symmetry Perspective, PM 200

2002 André Unterberger
Automorphic Pseudodifferential Analysis and Higher Level Weyl Calculi, PM 209

Alexander Lubotzky and Dan Segal
Subgroup Growth, PM 212

2003 Fuensanta Andreu-Vaillo, Vincent Caselles and José M. Mazón
Parabolic Quasilinear Equations Minimizing Linear Growth Functionals, PM 223

2004 Guy David
Singular Sets of Minimizers for the Mumford-Shah Functional, PM 233

2005 Antonio Ambrosetti and Andrea Malchiodi
Perturbation Methods and Semilinear Elliptic Problems on R^n, PM 240

José Seade
On the Topology of Isolated Singularities in Analytic Spaces, PM 241

2006 Xiaonan Ma and George Marinescu
Holomorphic Morse Inequalities and Bergman Kernels, PM 254

2007 Rosa Miró-Roig
Determinantal Ideals, PM 264

2008 Luis Barreira
Dimension and Recurrence in Hyperbolic Dynamics, PM 272

To my wife Sinead

Contents

Preface

Over the millennia Diophantine equations have supplied an extremely fertile source of problems. Their study has illuminated ever increasing points of contact between very different subject areas, including algebraic geometry, mathematical logic, ergodic theory and analytic number theory. The focus of this book is on the interface of algebraic geometry with analytic number theory, with the basic aim being to highlight the rôle that analytic number theory has to play in the study of Diophantine equations.

Broadly speaking, analytic number theory can be characterised as a subject concerned with counting interesting objects. Thus, in the setting of Diophantine geometry, analytic number theory is especially suited to questions concerning the "distribution" of integral and rational points on algebraic varieties. Determining the arithmetic of affine varieties, both qualitatively and quantitatively, is much more complicated than for projective varieties. Given the breadth of the domain and the inherent difficulties involved, this book is therefore dedicated to an exploration of the projective setting.

This book is based on a short graduate course given by the author at the I.C.T.P *School and Conference on Analytic Number Theory*, during the period 23rd April to 11th May, 2007. It is a pleasure to thank Professors Balasubramanian, Deshouillers and Kowalski for organising this meeting. Thanks are also due to Michael Harvey and Daniel Loughran for spotting several typographical errors in an earlier draft of this book. Over the years, the author has greatly benefited from discussing mathematics with Professors de la Bretèche, Colliot-Thélène, Fouvry, Hooley, Salberger, Swinnerton-Dyer and Wooley. A sincere debt of thanks is owed to them all. Finally, it is essential to single out Professor Heath-Brown for special gratitude, both as a mathematical inspiration and for the generosity of his explanations.

Chapter 1

Introduction

The study of integer solutions to Diophantine equations is a topic that is almost as old as mathematics itself. Since its inception at the hands of Diophantus of Alexandria in 250 A.D., it has been found to relate to virtually every mathematical field. Suppose that we are given a polynomial $f \in \mathbb{Z}[x_1, \ldots, x_n]$ and write

$$S_f := \{\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{Z}^n \setminus \{\mathbf{0}\} : f(\mathbf{x}) = 0\} \tag{1.1}$$

for the corresponding locus of non-zero integer solutions. There are a number of basic questions that can be asked about the set S_f.

- When is S_f non-empty?
- How large is S_f when it is non-empty?
- When S_f is infinite can we describe the set in some way?

Much of our progress has been driven by trying to understand the situation for equations in only $n = 2$ or 3 variables, with the arithmetic of curves being central in our understanding of Diophantine equations. The terrain for equations in 4 or more variables remains relatively obscure, however, with only a scattering of results and conjectures available.

The focus of this book will be on quantitative aspects of the arithmetic of higher-dimensional projective varieties. Thus our interest lies with the second and third questions posed above, for Diophantine equations $f = 0$ in which f is homogeneous and the corresponding zero locus S_f is infinite. The main goal is to understand how the counting function

$$N(f; B) := \#\{\mathbf{x} \in S_f : \|\mathbf{x}\| \leqslant B\} \tag{1.2}$$

behaves, as $B \to \infty$. Here $\|\cdot\| : \mathbb{R}^n \to \mathbb{R}_{>0}$ is an arbitrary choice of norm. We will always reserve $|\cdot|$ for the norm $|\mathbf{x}| := \max_{1 \leqslant i \leqslant n} |x_i|$, for any $\mathbf{x} \in \mathbb{R}^n$.

Aside from being intrinsically interesting in their own right, the study of functions like $N(f; B)$ often helps determine whether or not the equation $f = 0$

has any non-trivial integer solutions at all. In many applications of the Hardy–Littlewood circle method, for example, one is able to prove that S_f is infinite by showing that $N(f;B) \to \infty$ as $B \to \infty$. In addition to the solubility of Diophantine equations, there are a number of other situations where a proper understanding of $N(f;B)$ is extremely desirable. We will return to this topic in Section 1.3.

During the course of this work we will meet numerous estimates of one kind or another. It seems worthwhile recording some of the basic notation here. We will write $A(x) = O(B(x))$ to mean that there exists a constant $c > 0$ and $x_0 \in \mathbb{R}$ such that $|A(x)| \leqslant cB(x)$ for all $x \geqslant x_0$. We will often use the alternative notation $A(x) \ll B(x)$ or $B(x) \gg A(x)$. Furthermore, we will take $A(x) \asymp B(x)$ to mean $A(x) \ll B(x) \ll A(x)$ and $A(x) = o(B(x))$ to mean

$$\lim_{x\to\infty} \frac{A(x)}{B(x)} = 0.$$

Finally the relation $A(x) \sim B(x)$ will mean

$$\lim_{x\to\infty} \frac{A(x)}{B(x)} = 1.$$

The implied constants in our work will be uniform unless explicitly indicated otherwise by an appropriate subscript. We will occasionally find it convenient to depart from this convention, but such deviations will be clearly highlighted.

1.1 A naive heuristic

Given our discussion above, it is useful to have a general idea of which homogeneous polynomials f, hitherto called *forms*, might have an infinite zero locus S_f. Suppose that $f \in \mathbb{Z}[x_1, \dots, x_n]$ is a form of degree $d \geqslant 1$. Then for the vectors $\mathbf{x} \in \mathbb{Z}^n$ counted by $N(f;B)$, the values of $f(\mathbf{x})$ will all be of order B^d. In fact a positive proportion of them will have exact order B^d. Thus the probability that a randomly chosen value of $f(\mathbf{x})$ should vanish might be expected to be of order B^{-d}. Since the number of $\mathbf{x}$ to be considered has order B^n, this leads us to the following general expectation.

Heuristic. *When $n \geqslant d$ we have*

$$B^{n-d} \ll N(f;B) \ll B^{n-d}. \tag{1.3}$$

As a crude first approximation, therefore, this heuristic tells us that we might expect polynomials whose degree is less than the number of variables to have infinitely many solutions. Unfortunately there are a number of things that can conspire to upset this heuristic expectation. First and foremost, local conditions will often provide a reason for $N(f;B)$ to be identically zero no matter the values of d and n. By local obstructions we mean that the obvious necessary conditions for S_f to be non-empty fail. These are the conditions that the equation $f(\mathbf{x}) = 0$ should have a non-zero real solution $\mathbf{x} \in \mathbb{R}^n$, and secondly, that the congruence

$$f(\mathbf{x}) \equiv 0 \pmod{p^k}$$

should be soluble for every prime power p^k, with $p \nmid \mathbf{x}$.

It is quite easy to construct examples that illustrate the failure of these local conditions. For example, the equation

$$x_1^{2d} + \cdots + x_n^{2d} = 0$$

does not have any integer solutions, since it patently does not have any real solutions. Let us now exhibit an example, due to Mordell [94], of a polynomial equation that fails to have integer solutions because it fails to have solutions as a congruence modulo prime powers. Let K be a number field of degree d over $\mathbb{Q}$, with ring of integers $\mathscr{O}_K$, such that the rational prime p is inert in $\mathscr{O}_K$. Write

$$\mathbf{N}(y_1, \ldots, y_d) := N_{K/\mathbb{Q}}(y_1\omega_1 + \cdots + y_d\omega_d)$$

for the corresponding norm form, where $\omega_1, \ldots, \omega_d$ is a basis for K over $\mathbb{Q}$. Then $\mathbf{N}$ is a homogeneous polynomial of degree d, with coefficients in $\mathbb{Z}$. Exercise 1.1 shows that $p \mid \mathbf{N}(\mathbf{y})$ if and only if $p \mid \mathbf{y}$, for any $\mathbf{y} \in \mathbb{Z}^d$. We define the form

$$f_1 := \mathbf{N}(x_1, \ldots, x_d) + p\mathbf{N}(x_{d+1}, \ldots, x_{2d}) + \cdots + p^{d-1}\mathbf{N}(x_{d^2-d+1}, \ldots, x_{d^2}), \quad (1.4)$$

which has degree d and d^2 variables. We claim that the only integer solution to the equation $f_1(\mathbf{x}) = 0$ is the trivial solution $\mathbf{x} = \mathbf{0}$. To see this we argue by contradiction. Thus we suppose there to be a vector $\mathbf{x} \in \mathbb{Z}^{d^2}$ such that $f_1(\mathbf{x}) = 0$, with $\gcd(x_1, \ldots, x_{d^2}) = 1$. Viewed modulo p we deduce that $p \mid \mathbf{N}(x_1, \ldots, x_d)$, whence $p \mid (x_1, \ldots, x_d)$. Writing $x_i = py_i$ for $1 \leqslant i \leqslant d$, and substituting into the equation $f_1 = 0$, we find that

$$p^{d-1}\mathbf{N}(y_1, \ldots, y_d) + \mathbf{N}(x_{d+1}, \ldots, x_{2d}) + \cdots + p^{d-2}\mathbf{N}(x_{d^2-d+1}, \ldots, x_{d^2}) = 0.$$

But then we deduce in a similar fashion that $p \mid (x_{d+1}, \ldots, x_{2d})$. We may clearly continue in this fashion, ultimately concluding that $p \mid (x_1, \ldots, x_{d^2})$, which is a contradiction.

The polynomial (1.4) illustrates that for any d it is possible to construct examples of homogeneous polynomials in d^2 variables that have no non-zero integer solutions. The construction is purely local, relying upon showing that the polynomial fails to have a non-zero solution in $\mathbb{Q}_p^{d^2}$. It was conjectured by Artin that $\mathbb{Q}_p$ is a C_2 field, so that f should have a non-trivial p-adic zero as soon as $n > d^2$. The latter property is certainly true of forms of degree at most 3. However, Artin's conjecture is now known to be false, with Terjanian [118] having provided a counterexample with $p = 2, d = 4$ and $n = 18$. In a positive direction, Ax and Kochen [1] have used methods from mathematical logic to show that for every d

there is a number $p(d)$ such that f has a non-trivial p-adic zero provided $n > d^2$ and $p > p(d)$. When no restriction is placed on the size of p we know that there is a number v_d such that the form f has a non-trivial p-adic zero as soon as $n > v_d$. Brauer [8] achieved the first result in this direction using an elementary argument based on multiply nested inductions. The resulting value of v_d was too large to write down, but the central ideas have since been revisited and improved upon by Wooley [124], with the outcome that we may take $v_d \leqslant d^{2^d}$.

So far we have only seen examples of polynomials f for which the zero locus S_f is empty. In this case the corresponding counting function $N(f;B)$ is particularly easy to estimate! There are also examples which show that $N(f;B)$ may grow in quite unexpected ways, even when $n \geqslant d$. An equation that illustrates excessive growth is provided by the polynomial

$$f_2 := x_1^d - x_2(x_3^{d-1} + \cdots + x_n^{d-1}). \tag{1.5}$$

Here there are *trivial* solutions of the type $(0, 0, a_3, \ldots, a_n)$ which already contribute $\gg B^{n-2}$ to the counting function $N(f;B)$, whereas (1.3) predicts that we should have exponent $n - d$.

It is also possible to construct examples of varieties which demonstrate inferior growth, as observed by Wooley [125]. Let $n > d^2$ and choose any d^2 linear forms $L_1, \ldots, L_{d^2} \in \mathbb{Z}[x_1, \ldots, x_n]$ that are linearly independent over $\mathbb{Q}$. Consider the form

$$f_3 := f_1(L_1(x_1, \ldots, x_n), \ldots, L_{d^2}(x_1, \ldots, x_n)),$$

where f_1 is given by (1.4). Then it is clear that $N(f_3;B)$ has the same order of magnitude as the counting function associated to the system of linear forms $L_1 = \cdots = L_{d^2} = 0$. Since these forms are linearly independent we deduce that $N(f_3;B)$ has order of magnitude B^{n-d^2}, whereas (1.3) led us to expect an exponent $n - d$.

We have seen several reasons why (1.3) might fail — how about some evidence supporting it? One of the most outstanding achievements in this direction is the following very general result due to Birch [6].

Theorem 1.1. *Suppose $f \in \mathbb{Z}[x_1, \ldots, x_n]$ is a non-singular homogeneous polynomial of degree d in $n > (d-1)2^d$ variables. Assume that $f(\mathbf{x}) = 0$ has non-trivial solutions in $\mathbb{R}$ and each p-adic field $\mathbb{Q}_p$. Then there is a constant $c_f > 0$ such that*

$$N(f;B) \sim c_f B^{n-d},$$

as $B \to \infty$.

We will discuss the proof of this result for the case $d = 4$ in Section 8.2. Birch's result does not apply to either of the polynomials f_2, f_3 that we considered above, since both of these contain a rather large singular locus. Since generic homogeneous polynomials are non-singular, Birch's result answers our initial questions completely for typical forms with $n > (d-1)2^d$. It would be of considerable interest to reduce the lower bound for n, but except for $d \leqslant 4$ this has not been done. Theorem 1.1 is established using the Hardy–Littlewood circle method, and exhibits a common feature of all Diophantine problems successfully tackled via this machinery: the number of variables involved needs to be large compared to the degree. In particular, there is an obvious disparity between the range for n in Birch's result and the range for n in (1.3).

Before coming to the Hardy–Littlewood circle method, we will also discuss some of the other technology that has been brought to bear on the quantitative analysis of homogeneous Diophantine equations. Whereas the circle method is geared towards forms for which the number of variables n is large compared to the degree d, we will also meet machinery to deal with equations in which n is comparable in size with d, in addition to those equations for which n is much smaller than d.

1.2 The basic counting function

It turns out that phrasing things in terms of single homogeneous polynomial equations is far too restrictive. It is much more satisfactory to work with arbitrary projective algebraic varieties $V \subseteq \mathbb{P}^{n-1}$. All of the varieties that we will work with are assumed to be cut out by a finite system of homogeneous equations defined over $\mathbb{Q}$. Moreover, whenever we speak of a variety as being *irreducible* we will henceforth take this to mean that the variety is geometrically reduced and irreducible. In the case of varieties cut out by a single equation this is equivalent to the underlying polynomial being irreducible over the complex numbers.

Our main interest lies with those varieties V for which we expect the set $V(\mathbb{Q}) = V \cap \mathbb{P}^{n-1}(\mathbb{Q})$ to be infinite. Let $x = [\mathbf{x}] \in \mathbb{P}^{n-1}(\mathbb{Q})$ be a projective rational point, with $\mathbf{x} \in \mathbb{Z}^n$ chosen so that $\gcd(x_1, \dots, x_n) = 1$. Then we define the *height* of x to be

$$H(x) := \|\mathbf{x}\|.$$

This therefore defines a function $H : \mathbb{P}^{n-1}(\mathbb{Q}) \to \mathbb{R}_{>0}$, and is none other than the exponential height function metrized by the choice of norm $\|\cdot\|$. Given any locally closed subset $U \subseteq V$, we may then define the counting function

$$N_U(B) := \#\{x \in U(\mathbb{Q}) : H(x) \leqslant B\}, \tag{1.6}$$

for each $B \geqslant 1$. All known examples of asymptotic formulae for the counting function $N_U(B)$ take the shape

$$N_U(B) \sim cB^a (\log B)^b,$$

as $B \to \infty$, for $a, b, c \geqslant 0$ such that $a \in \mathbb{Q}$ and $b \in \frac{1}{2}\mathbb{Z}$. In Chapter 2 we will encounter an attempt to interpret these quantities in terms of the underlying geometry of V.

The main difference between the counting function $N_U(B)$ and the quantity introduced in (1.2) is that we are now only interested in *primitive* integer solutions, by which we mean that the components of the vector $\mathbf{x} \in \mathbb{Z}^n$ should share no common prime factors. This formulation has the advantage of treating all scalar multiples of a given non-zero integer solution as a single point. We will henceforth write $\mathbb{Z}^n_{\text{prim}}$ for the set of primitive vectors in $\mathbb{Z}^n$.

Recall the definition of the *Möbius function* $\mu : \mathbb{N} \to \{0, \pm 1\}$, which is given by

$$\mu(n) := \begin{cases} 0, & \text{if } p^2 \mid n \text{ for some prime } p, \\ 1, & \text{if } n = 1, \\ (-1)^r, & \text{if } n = p_1 \cdots p_r \text{ for distinct primes } p_1, \ldots, p_r. \end{cases}$$

The Möbius function is a multiplicative arithmetic function that is of fundamental importance in analytic number theory. It is frequently engaged via the simple identity

$$\sum_{d \mid n} \mu(d) = \begin{cases} 1, & \text{if } n = 1, \\ 0, & \text{if } n \in \mathbb{Z}_{>1}. \end{cases}$$

It is through this rôle as a characteristic function that it figures in the quantitative study of Diophantine equations. We illustrate the procedure by showing how it allows us to relate the counting function (1.6) to our earlier counting function $N(f; B)$ in (1.2), when $U = V$ and $V \subset \mathbb{P}^{n-1}$ is a hypersurface with underlying form $f \in \mathbb{Z}[x_1, \ldots, x_n]$. On noting that $\mathbf{x}$ and $-\mathbf{x}$ represent the same point in $\mathbb{P}^{n-1}$, it follows that

$$\begin{aligned} N_V(B) &= \frac{1}{2} \#\{\mathbf{x} \in \mathbb{Z}^n_{\text{prim}} : f(\mathbf{x}) = 0, \ \|\mathbf{x}\| \leqslant B\} \\ &= \frac{1}{2} \sum_{k=1}^{\infty} \mu(k) \#\{\mathbf{x} \in \mathbb{Z}^n : f(\mathbf{x}) = 0, \ k \mid \mathbf{x}, \ \|\mathbf{x}\| \leqslant B\}. \end{aligned}$$

But then a simple change of variables furnishes

$$N_V(B) = \frac{1}{2} \sum_{k=1}^{\infty} \mu(k) N(f; k^{-1}B). \tag{1.7}$$

This process of using the Möbius function will henceforth be termed *Möbius inversion*.

The simplest sort of subvariety in $\mathbb{P}^{n-1}$ is obtained by taking f to be identically zero. This corresponds to taking $V = \mathbb{P}^{n-1}$. Schanuel [107] has obtained an asymptotic formula for $N_{\mathbb{P}^{n-1}}(B)$. There is a natural way to define a height function on $\mathbb{P}^{n-1}(K)$ for any algebraic number field K, and it is to this more general context that Schanuel's result applies. It will be instructive to present a proof of this result in the case $K = \mathbb{Q}$.