An Algebraic View of Discrete Geometry

by

Frank de Zeeuw

Doctoraal (B.Sc. and M.Sc.) in Mathematics, Rijksuniversiteit Groningen, 2006
B.Sc. in Artifial Intelligence, Rijksuniversiteit Groningen, 2006
Master of Advanced Study (MASt) in Mathematics, University of Cambridge, 2007

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY

in

The Faculty of Graduate Studies

(Mathematics)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

October 2011

Abstract

This thesis includes three papers and one expository chapter as background for one of the papers. These papers have in common that they combine algebra with discrete geometry, mostly by using algebraic tools to prove statements from discrete geometry. Algebraic curves and number theory also recur throughout the proofs and results. In Chapter 1, we will detail these common threads.

In Chapter 2, we prove that an infinite set of points in $\mathbb{R}^2$ such that all pairwise distances are rational cannot be contained in an algebraic curve, except if that curve is a line or a circle, in which case at most 4 resp. 3 points of the set can be outside the line or circle. In the proof we use the classification of curves by their genus, and Faltings' Theorem.

In Chapter 3, we informally present an elementary method for computing the genus of a planar algebraic curve, illustrating some of the techniques in Chapter 2.

In Chapter 4, we prove a bound on the number of unit distances that can occur between points of a finite set in $\mathbb{R}^2$, under the restriction that the line segments corresponding to these distances make a rational angle with the horizontal axis. In the proof we use graph theory and an algebraic theorem of Mann.

In Chapter 5, we give an upper bound on the length of a simultaneous arithmetic progression (a two-dimensional generalization of an arithmetic progression) on an elliptic curve, as well as for more general curves. We give a simple proof using a theorem of Jarník, and another proof using the Crossing Inequality and some bounds from elementary algebraic geometry, which gives better explicit bounds.

Preface

This thesis is based on 3 papers, of which two have been published and one has been accepted for publication. The results are reproduced in this thesis with permission from the coauthors and journals. All work involved in these publications was shared in equal parts between the authors.
Chapter 2 was published as:

J. Solymosi and F. de Zeeuw,
On a Question of Erdős and Ulam,
Discrete and Computational Geometry 43 (2010), 393–401.

Chapter 3 is a writeup for a course taught at UBC by Jozsef Solymosi on Additive Combinatorics.
Chapter 4 was accepted for publication:

R. Schwartz, J. Solymosi, and F. de Zeeuw,
Rational distances with rational angles,
accepted for publication in Mathematika;
`arXiv:1008.3671v2`, [math.CO], 2011.

Chapter 5 was published as:

R. Schwartz, J. Solymosi, and F. de Zeeuw,
Simultaneous arithmetic progressions on algebraic curves,
International Journal of Number Theory 7 (2011), 921–931.

The mathematical content of these papers is unchanged in this thesis, though the exposition has undergone some changes.

Table of Contents

Abstract ii

Preface iii

Table of Contents iv

Acknowledgements vi

1 Introduction 1
- 1.1 Outline 1
- 1.2 Problems 1
- 1.3 Tools 4

2 Rational sets 8
- 2.1 Background 8
- 2.2 Main results 9
- 2.3 Proof of Theorem 2.2.1 11
- 2.4 Proof of Theorem 2.2.2 17

3 Computing the genus 18
- 3.1 Introduction 18
- 3.2 Basics from algebraic geometry 22
- 3.3 Computing the genus, Part I 23
- 3.4 Blowing up singularities 25
- 3.5 Computing the genus, Part II 28

4 Rational distances with rational angles 31
- 4.1 Background 31
- 4.2 Main results and proof sketch 32
- 4.3 Mann's Theorem 33
- 4.4 Proofs of main theorems 36
- 4.5 Lower bounds 40

5 Simultaneous arithmetic progressions 42
5.1 Introduction 42
5.2 First proof for algebraic curves 43
5.3 Second proof for algebraic curves 45
5.4 SAPs on elliptic curves 48

Bibliography 51

Acknowledgements

I would like to thank my advisor Jozsef Solymosi, fellow student and coauthor Ryan Schwartz, my supervisory committee, other UBC faculty, staff and students, my family and friends, and Julia, for all their guidance, cooperation, support and entertainment. I would also like to thank Kalle Karu, Jirka Matoušek, Trevor Wooley and three anonymous referees for help with the individual papers.

Chapter 1

Introduction

1.1 Outline

This thesis is based on 3 papers. What these papers have in common is that they combine discrete geometry with algebra (or algebraic geometry), with a touch of number theory. In Chapters 2 and 4, the problems are from discrete geometry (recognizable from the fact that Erdős coined them), and the tools are, more or less, from algebra. In both cases, we translate the geometric problem into algebraic equations, and apply some algebraic result, in the form of a bound on the number of solutions to the equations.

Chapter 5 is different: its problem is algebraic (a statement about algebraic curves), and our main tool is discrete (graph theory). In Chapter 3, we present an elementary method for computing the genus of an algebraic curve, as background for Chapter 2.

All three papers also involve some number theory: the problems revolve around the rationality or integrality of numbers, and the tools partly belong to number theory (in particular the theorems of Faltings, Mann, and Jarník).

In the rest of this introduction, we describe the three problems that form the basis of this thesis in Section 1.2, and in Section 1.3 we list the tools that we will apply to these problems.

1.2 Problems

1.2.1 Rational distance sets

We define a *rational distance set* to be a set $S \subset \mathbb{R}^2$ such that the distance between any two points in S is a rational number. We are interested in the existence of infinite rational distance sets on algebraic curves.

This notion was introduced by Erdős in 1945, when he proved with Anning [4] that any infinite set with all distances *integral* must be contained in a line. For a rational distance set, the same is not true, because one can find infinite rational distance sets contained in a circle. In fact, rational distance sets can be *dense* on lines or circles. In 1960 Ulam wrote [53] that he believed that an infinite rational distance set cannot be dense in the plane, but no progress has been made on this question to date. Erdős conjectured [23] that an infinite rational distance set must be very restricted, but wrote that it was probably a very deep problem. We prove the following.

Theorem 1.2.1. *A rational distance set S can have only finitely many points in common with an algebraic curve defined over $\mathbb{R}$, unless the curve has a component which is a line or a circle.*

If S has infinitely many points on a line or on a circle, then all but at most 4 resp. 3 points of S are on the line or on the circle.

Our proof will use the genus of algebraic curves and Faltings' Theorem (see Section 1.3.2). Roughly, the idea of the proof is that a curve with an infinite rational distance set on it lets us construct a new curve with higher genus, and with a set of rational points on it corresponding to the points in the rational distance set. Faltings' Theorem then tells us that this set of points must be finite, hence the rational distance set could not have been infinite.

1.2.2 Unit distances

A famous problem of Erdős concerns the maximal number of unit distances among n points in the plane; we will denote this number by $u(n)$. Erdős [18] showed that $u(n) > n^{1+\frac{c}{\log\log n}}$. The best known upper bound is $u(n) < cn^{4/3}$, due to Spencer, Szemerédi and Trotter [51].

We will approach this problem under the restriction that we only consider unit distances that have *rational angle*, by which we mean that the line through the pair of points makes a rational angle in degrees with the x-axis (or equivalently, its angle in radians divided by π is rational). This leads to the following result.

Theorem 1.2.2. *Suppose we have n points in $\mathbb{R}^2$. Then the number of pairs of points with unit distance and rational angle is at most $n^{1+6/\sqrt{\log n}}$.*

In the proof we view each unit distance line segment between two of the points as a complex number, which is then a root of unity thanks to the rational angle condition. Given a large set with many unit distances with rational angles, we deduce linear equations with many solutions in roots of unity. Comparing with a uniform upper bound on the number of such solutions, derived from Mann's Theorem (see Section 1.3.1), then leads to our bound.

In fact, this proof works for more general situations, leading to the following three theorems.

Theorem 1.2.3. *Suppose we have n points in $\mathbb{R}^2$, no three of which are on a line. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+6/\sqrt{\log n}}$.*

Theorem 1.2.4. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $0 < \alpha < 1/2$. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+\alpha+6/\sqrt{\log n}}$.*

Theorem 1.2.5. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $1/2 \leq \alpha \leq 1$. Then the number of pairs of points with rational distance and rational angle is at most $4n^{1+\alpha}$.*

1.2.3 Simultaneous arithmetic progressions on curves

There are interesting problems in number theory related to arithmetic progressions on elliptic curves (see Section 1.3.2 and Chapter 3 for more about elliptic curves). An example of such an open problem is to find the maximum number (if it exists) of rational points on an elliptic curve such that their x-coordinates are in arithmetic progression [8]; these are for instance related to 3×3 magic squares with square entries [7].

In [30], Garcia-Selfa and Tornero looked instead for "simultaneous" arithmetic progressions on elliptic curves, which are defined as follows.

Definition. A *simultaneous arithmetic progression* (SAP) of length k consists of points $(x_i, y_{\sigma(i)})$ in $\mathbb{R}^2$, where $x_i = a_1 + i \cdot d_1$ and $y_i = a_2 + i \cdot d_2$ for $i = 0, 1, \ldots, k-1$ are arithmetic progressions, and σ is a permutation of the numbers $0, 1, \ldots, k-1$.

Garcia-Selfa and Tornero gave examples of elliptic curves over $\mathbb{Q}$ that contain an SAP of length 6, and showed that no SAP of length 7 exists on an elliptic curve. However, it did not seem computationally feasible to extend their methods to longer SAPs. The final open problem they suggested in their paper is finding a universal bound for the length of SAP's on elliptic curves over $\mathbb{Q}$. We prove the following (unlikely to be optimal) bound to this effect.

Theorem 1.2.6. *Consider an elliptic curve over a subfield of $\mathbb{R}$ given by $y^2 + axy + by = x^3 + cx^2 + dx + e$. Then the length of an SAP on this curve is ≤ 4319.*

In our proof of this bound we construct a graph out of translates of the curve in a grid. Then we derive an algebraic upper bound on the number of intersections, from Bézout's Theorem (see Section 1.3.3), and compare it with a combinatorial lower bound on the number of crossings in the graph, obtained using the Crossing Inequality (see Section 1.3.3). We then extend this proof to algebraic curves of any degree. It turns out that the structure of an SAP is not crucial; our proof only uses the fact that an SAP has k points on a $k \times k$ grid. This led to the following more general theorem.

Theorem 1.2.7. *If f is a real plane algebraic curve of degree $d \geq 2$ with no linear factor, and f contains at least k points from a $k \times k$ grid, then there are absolute constants c and m such that $k \leq cd^m$.*

The proof outlined above gives $m = 7$ for an arbitrary curve, and $m = 4$ when the curve is irreducible.

We also give an alternative proof of this theorem, using Jarník's Theorem (see Section 1.3.3), which gives an upper bound on the number of integer grid points on a curve. This is a more direct proof, but the resulting bound is weaker: $m = 9$ in general and $m = 6$ for irreducible curves.

1.3 Tools

1.3.1 Linear equations

The simplest type of equation that we might derive from our geometric problems are linear equations

$$\sum_{i=1}^{k} a_i x_i = 0, \quad a_i \in \mathbb{C}.$$

Of course, without restrictions on the values that the variables can take, there is not much to say about the solutions. Under suitable restrictions, however, it is possible to obtain bounds on the number of solutions, that are even *uniform*, in the sense that they do not depend on the coefficients. The grandest example of this is the following theorem of Evertse, Schlickewei and Schmidt [26], where the solutions are restricted to a multiplicative subgroup of $\mathbb{C}$ of finite rank. It is a corollary of the Subspace Theorem proved by Schmidt [45].

Theorem 1.3.1. *Let Γ be a subgroup of $\mathbb{C}^*$ of finite rank r, and let $a_i \in \mathbb{C}^*$. Then the equation $\sum_{i=1}^{k} a_i x_i = 1$ has at most $S(k,r) = e^{(6k)^{5k}(r+1)}$ non-degenerate solutions.*

The proof of this theorem is very difficult, and we did not use the theorem in this thesis. But we mention it because it did lead us to consider a special case which has an elementary proof. We take the group in $\mathbb{C}^*$ to be *finite*, which means it must consist of roots of unity. In that case, the following result was proved by Mann in 1965 [41].

Theorem 1.3.2 (Mann). *The number of k-tuples $\{\zeta_i\}_{i=1}^k$ of roots of unity, with $\zeta_1 = 1$, that satisfy $\sum_{i=1}^k \zeta_i = 0$, but no shorter such equation, is bounded by a constant $C(k)$.*

Because the proof of this theorem uses only basic algebra, we were able to modify it to obtain an extension which exactly suits our needs in the proof of Theorem 1.2.2.

1.3.2 Algebraic curves

We will also encounter equations of higher degree, which will lead us to consider *algebraic curves*. For these we will not have the same kind of explicit bound as for linear equations, but much can still be said about their solution sets, in particular for points with rational coordinates. To explain this, we need to outline the classification of algebraic curves by *genus*.

By a *curve* C_f we mean the zero set in $\mathbb{R}^2$ of a polynomial $f \in \mathbb{R}[x, y]$:

$$C_f = C_f(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}.$$

The curve is *irreducible* if f is irreducible. In our problems we will also consider the set of *rational points* $C_f(\mathbb{Q}) = \{(x, y) \in \mathbb{Q}^2 : f(x, y) = 0\}$ of a curve C_f. We will frequently denote the curve by its polynomial f. Note that in this thesis we are restricting ourselves to *plane* curves, except in Section 2.3, where we use curves in $\mathbb{R}^3$.

The *genus* g is an important invariant of an irreducible curve, and we have devoted Chapter 3 to it. It has several equivalent definitions, but they are not elementary, and it would take us too far afield to introduce them here. Computationally, we can use the following formula in terms of the degree d of the curve:

$$g = \frac{(d-1)(d-2)}{2} - \sum_P \delta_P,$$

where the sum ranges over the *singularities* P, points of the curve where both derivatives of f vanish (possibly including projective points). The $\delta_P \geq 0$ are nonnegative invariants associated to singularities. In Chapter 3 we will specify what the δ_P are, and how one can compute them in an elementary way. But we can already see that the genus of a nonsingular curve is simply given by $g = \frac{(d-1)(d-2)}{2}$.

It turns out that the genus of a curve C is related to the structure of the set of rational points $C(\mathbb{Q})$ on the curve, as follows:

$$\begin{aligned} g = 0 &\iff \text{the curve is rational} \\ g = 1 &\iff \text{the curve is an elliptic curve} \\ g \geq 2 &\implies C(\mathbb{Q}) \text{ is finite} \end{aligned}$$

In Chapter 3 we will go into more detail about this classification. For now, the important thing is that curves with $g = 0$ or $g = 1$ have a well-understood structure (we will explain what rational and elliptic curves are in Chapter 3; we will meet elliptic curves again in Chapter 5). Note that the linear equations from Section 1.3.1 are curves with genus 0, but there we consider solutions taken from a group of finite rank, rather than $\mathbb{Q}$, which allows us to say more about their size.

When $g \geq 2$, the structure of $C(\mathbb{Q})$ is not so easy to understand, but thanks to the following theorem of Faltings [27], we do know that $C(\mathbb{Q})$ is finite:

Theorem 1.3.3 (Faltings). *An irreducible curve of genus $g \geq 2$, defined over a number field, contains only finitely many rational points.*

The proof is very complicated, but the statement is relatively easy to apply, and we will use it as a black box.

A related conjecture that should be mentioned here (although we have not used it in this thesis), is the Uniformity Conjecture from [11]. It says that for every number field K and integer $g \geq 2$ there exists an integer $B(K, g)$ such that no smooth curve of genus g defined over K has more than $B(K, g)$ rational points. If this is true, it could perhaps be used to improve our finiteness result in Chapter 2.

1.3.3 Points on curves

The tools that we will use in Chapter 5 can be roughly characterized as dealing with points on curves: intersection points of algebraic curves, crossings of planar graphs, and integer grid points on convex curves.

The first of these is a standard theorem from algebraic geometry, Bézout's Theorem. It will provide the upper bound that we use in the proof of Theorem 1.2.6.

Theorem 1.3.4 (Bézout)**.** *Suppose F and G are curves of degree m and n. If F and G do not have a common factor, then they intersect in at most mn points.*

The second tool is a well-known theorem from graph theory, bounding the crossing number of a graph. We will apply it to obtain the lower bound used in the proof of Theorem 1.2.6.

Given a simple graph G, we define the *crossing number* $\mathrm{cr}(G)$ as the minimum number of pairs of crossing edges in a planar drawing of G.

Theorem 1.3.5 (Crossing Inequality)**.** *Suppose G is a simple graph with n vertices and e edges. If $e > 7.5n$ then*

$$\mathrm{cr}(G) \geq \frac{e^3}{33.75n^2}.$$

The third tool is a theorem of Jarník [35], which will give us an alternative proof of Theorem 1.2.7.

Theorem 1.3.6 (Jarník)**.** *If f is a strictly convex differentiable curve of length N, then the number of integer points on f is less than $\alpha N^{2/3}$ for some constant α.*

Chapter 2

Rational sets

2.1 Background

Recall that in the introduction we defined a *rational distance set* to be a set $S \subset \mathbb{R}^2$ such that the distance between any two elements is a rational number. For briefness, in this chapter we will refer to rational distance sets as *rational sets*. We are interested in the existence of infinite rational sets on plane algebraic curves.

On any line, one can easily find an infinite rational set that is in fact dense. It is also possible to find an everywhere dense rational subset of the unit circle (see Section 2.2). On the other hand, it is not known if there is a rational set with 8 points *in general position*, i.e. no 3 on a line, no 4 on a circle. In 1945, Anning and Erdős proved that any infinite *integral* set, i.e. where all distances are integers, must be contained in a line [4]. Problems related to rational and integral sets became one of Erdős' favorite subjects in combinatorial geometry [19, 20, 21, 22, 23, 25].

In 1945, when Ulam heard Erdős' simple proof [17] of his theorem with Anning, he said that he believed there is no everywhere dense rational set in the plane (see Problem III.5. in [53], and also [24]). Erdős conjectured that an infinite rational set must be very restricted, but that it was probably a very deep problem [23, 24]. Not much progress has been made on Ulam's question. There were attempts to find rational sets on parabolas [10, 12], and there were some results on integral sets, in particular bounds were found on the diameter of integral sets [33, 49]. Recently Kreisel and Kurz found an integral set with 7 points in general position [38].

In this thesis, we prove that lines and circles are the only irreducible algebraic curves that contain infinite rational sets. Our main tool is Faltings' Theorem [27]. We will also show that if a rational set S has infinitely many points on a line or on a circle, then all but at most 4 resp. 3 points of S are on the line or on the circle. This answers questions of Guy, Problem D20 in [31], and Pach, Section 5.11 in [6].

2.2 Main results

2.2.1 Main Theorems

Our main result is the following.

Theorem 2.2.1. *A rational set in the plane can have only finitely many points in common with an algebraic curve defined over $\mathbb{R}$, unless the curve has a component which is a line or a circle.*

The two special cases, line and circle, are treated in the next theorem.

Theorem 2.2.2. *If a rational set S has infinitely many points on a line or on a circle, then all but at most 4 resp. 3 points of S are on the line or on the circle.*

These numbers are best possible: there are infinite rational sets with all but 4 points on a line, and there are infinite rational sets with all but 3 points on a circle. A construction of Huff [34, 43] gives an infinite rational set with all but 4 points on a line.

The circle case follows from the line case by applying an inversion (see Section 2.2.2) with rational radius and center one of the 4 points not on the line. This will give an infinite rational set on the circle, with 3 points off it (the fourth point, the one that was used as center, is lost because the inversion sends it to infinity).

We can formulate our Theorem 2.2.1 in a different way by using the term *curve-general position*: we say that a point set S of $\mathbb{R}^2$ is in curve-general position if no algebraic curve of degree d contains more than $d(d+3)/2$ points of S. Note that $d(d+3)/2$ is the number of points in general position that determine a unique curve of degree d (see Section 5.2 in [29]).

Corollary 2.2.3. *If S is an infinite rational set in general position, then there is an infinite $S' \subset S$ such that S' is in curve-general position.*

Proof. We will construct S' with a greedy approach: we consecutively add as many points of S as possible while maintaining curve-general position.

Let S_5 consist of any five points in S, and let T_5 be the set of finitely many points on the unique conic through those 5 points. Continue recursively; at step n, add a point from $S \backslash T_{n-1}$ to S_{n-1} to get S_n. For each d such that $d(d+3)/2 \leq n$, let T_n be the union of T_{n-1} and the set of points of S that are on a curve of degree d through any $d(d+3)/2$ points in S_n. Since each T_n is finite by Theorem 2.2.1, we can always add another point. Then the infinite union of the sets S_n is an infinite subset of S with the required property. □

2.2.2 Two lemmas

Rationality of distances in $\mathbb{R}^2$ is clearly preserved by translations, rotations and uniform scaling ($(x, y) \mapsto (\lambda x, \lambda y)$ with $\lambda \in \mathbb{Q}$). More surprisingly, rational sets are preserved under certain central inversions. This will be an important tool in our proof below.

Lemma 2.2.4. *If we apply inversion to a rational set S, with center a point $x \in S$ and rational radius, then the image of $S \backslash \{x\}$ is a rational set.*

Proof. We may assume the center to be the origin and the radius to be 1. The properties of inversion are most easily seen in complex notation, where the map is $z \mapsto 1/z$ (ignoring a reflection). Suppose we have two points z_1, z_2 with rational distances $|z_1|$, $|z_2|$ from the origin, and with $|z_2 - z_1|$ rational. Then the distance between their images is also rational:

$$\left| \frac{1}{z_1} - \frac{1}{z_2} \right| = \left| \frac{z_2 - z_1}{z_1 z_2} \right| = \frac{|z_2 - z_1|}{|z_1||z_2|}.$$

□

A priori, points in a rational set could take any form. But after moving two of the points to fixed rational points by translating, rotating, and scaling, the points in the set are almost rational. This fact is well-known among researchers working with integer sets, and as far as we know, it was first proved by Kemnitz [36].

Note that the lemma implies that any curve of degree d that contains at least $d(d+3)/2$ points from a rational set containing $(0, 0)$ and $(1, 0)$ is defined over $\mathbb{Q}(\sqrt{k})$, for some $k \in \mathbb{N}$.

Lemma 2.2.5. *For any rational set S containing $(0, 0)$ and $(1, 0)$ there is a square-free integer k such that any point in S is of the form*

$$(r_1, r_2\sqrt{k}), \quad r_1, r_2 \in \mathbb{Q}.$$

Proof. Suppose we have a rational set containing $(0, 0)$ and $(1, 0)$, and a third point (x, y). Then $x^2 + y^2 \in \mathbb{Q}$ and $(x-1)^2 + y^2 = x^2 + y^2 - 2x + 1 \in \mathbb{Q}$, hence by subtracting we have $-2x + 1 \in \mathbb{Q}$, so $x \in \mathbb{Q}$. Combining that with $x^2 + y^2 \in \mathbb{Q}$ gives $y = s\sqrt{k}$ for $s \in \mathbb{Q}$ and a square-free integer k.

It remains to show that k is the same for different points. Let $(r_1, s_1\sqrt{k})$ and $(r_2, s_2\sqrt{l})$ be two points of our rational set. Then their distance,

$$(r_1 - r_2)^2 + (s_1\sqrt{k} - s_2\sqrt{l})^2 = (r_1 - r_2)^2 + s_1^2 k + s_2^2 l - 2s_1 s_2\sqrt{kl},$$

is rational, which implies that $\sqrt{kl} \in \mathbb{Q}$. If both k and l are square-free integers, that is only possible if $k = l$. □

2.3 Proof of Theorem 2.2.1

2.3.1 Outline

Our proof relies on the following theorem of Faltings [27].

Theorem 2.3.1 (Faltings)**.** *An irreducible curve of genus ≥ 2, defined over a number field K, contains only finitely many K-rational points.*

See Chapter 3 or [48] for definitions. As in the Introduction, in this chapter by *curve* (defined over a field $K \subset \mathbb{R}$) we usually mean the zero set in $\mathbb{R}^2$ of a polynomial in two variables with coefficients from K. But when we consider the genus of a curve, we are actually talking about the projective variety defined by the polynomial. In the proof we will also briefly encounter curves in $\mathbb{R}^3$; for definitions we again refer to [48], especially for their genus, which is not covered in Chapter 3.

Before we start to outline the proof, note that Theorem 2.2.1 allows reducible curves, but if the statement holds for irreducible curves, it follows for reducible curves. So in our proof we need only consider irreducible curves, and in the rest of this chapter we will assume that all curves are irreducible.

For a curve C of genus ≥ 2, defined over $\mathbb{R}$, we can prove our theorem very quickly. Suppose we have an infinite rational set S on such a C. We can move two points in S to $(0,0)$ and $(0,1)$, so that by Lemma 2.2.5 the elements of S are of the form $(r_1, r_2\sqrt{k})$. Then by the remark right before Lemma 2.2.5, the curve is defined over the number field $\mathbb{Q}(\sqrt{k})$. By Faltings' Theorem, S must be finite.

In the following sections we will treat the cases with genus 0 or 1. The main idea of the proof is that if we have an infinite rational set S on a curve C_1 of genus 0 or 1, which is not a line or a circle, then we can construct a curve C_2 of genus ≥ 2 with infinitely many rational points, contradicting Faltings' Theorem.

But the cases split into two groups, for which the details of the proofs are very different. For a curve C_1 with genus 1, or with genus 0 and degree ≥ 4, a point $(r_1, r_2\sqrt{k})$ from the rational set on C_1 will give a point $(r_1, r_2\sqrt{k}, r_3)$ on a curve C_2 in $\mathbb{R}^3$. Then using the Riemann-Hurwitz formula we can show that the genus of C_2 is ≥ 2.

For curves with genus 0 and degree < 4, this does not work, and we need a different approach. First we apply a central inversion: by Lemma 2.2.4 this preserves the rational set, but it transforms most curves into curves with degree ≥ 4, so that the previous proof applies. There are two exceptions:

first of all, lines and circles (as we would hope!), and second, a set of curves with $d = 3$. The curves in this second group have such specific polynomials that we can use them to explicitly construct hyperelliptic curves with infinitely many rational points (similar to the constructions in Section 3.1.2). That finishes the proof.

2.3.2 Curves of genus 1

Let $C_1 : f(x, y) = 0$ be an irreducible algebraic curve of genus $g_1 = 1$, and suppose that there is an infinite set S on C_1 with pairwise rational distances. Assume that the points $(0, 0)$ and $(1, 0)$ are on C_1 and in S, and that $(0, 0)$ is not a singularity of C_1. Below we will be allowed to make any other assumptions on C_1 that we can achieve by translating, rotating or scaling it, as long as we also satisfy the assumptions above. In particular, we can use any of the infinitely many rotations about the origin that put a different point of S on the x-axis.

We wish to show that the intersection curve C_2 of the surfaces

$$\begin{aligned} X &: f(x, y) = 0, \\ Y &: x^2 + y^2 = z^2, \end{aligned}$$

has genus $g_2 \geq 2$. Graphically, Y is a cone around the z-axis, and X is a "cylinder" above (and below) C_1, if C_1 is viewed as a curve in the $z = 0$ plane. The intersection curve C_2 then consists of the points on the cone that are directly above or below C_1.

Define the map $\pi : C_2 \to C_1$ by $(x, y, z) \mapsto (x, y)$, i.e. the restriction to C_2 of the projection from the cone Y to the $z = 0$ plane. The preimage of a point (x, y) consists of two points $(x, y, \pm\sqrt{x^2 + y^2})$, except when $x^2 + y^2 = 0$, which in $\mathbb{C}^2$ happens on the two lines $x + iy = 0$ and $x - iy = 0$.

We can determine (or at least bound from below) the genus of C_2 using the Riemann-Hurwitz formula applied to π (see [48], Ch. I, for the formula and for the definitions of the degree of a map and the ramification index e_P of π at a point),

$$2g_2 - 2 \geq \deg \pi \cdot (2g_1 - 2) + \sum_{P \in C_2} (e_P - 1).$$

This formula is usually stated with equality for smooth curves, but we are allowing C_1 and C_2 to have singularities. To justify this, observe that the map π corresponds to a map $\widetilde{\pi} : \widetilde{C_1} \to \widetilde{C_2}$ between the normalizations of the curves, for which Riemann-Hurwitz holds. The normalizations have the same genera as the original curves, and $\widetilde{\pi}$ has the same degree. Furthermore a ramification point of π away from any singularities gives a ramification point of $\widetilde{\pi}$. It is enough for us to have this inequality, but there could be more ramification points for $\widetilde{\pi}$, above where the singularities used to be.

Applying this formula with $g_1 = 1$, $\deg \pi = 2$, we have

$$g_2 \geq 1 + \frac{1}{2} \sum_{P \in C_2} (e_P - 1),$$

so to get $g_2 \geq 2$, we only need to show that π contains some point P with $e_P \geq 2$, i.e. a ramification point.

The potential ramification points are above the intersection points of C_1 with the lines $x \pm iy = 0$, of which there are $2d$ by Bézout's Theorem (Theorem 5.3.4), counting with multiplicities. Such an intersection point P can only fail to have a ramification point above it if the curve has a singularity at P, or if the curve is tangent to the line there. We will show that there are only finitely many lines through the origin on which one of those two things happens. Then certainly one of the infinitely many rotations of C_1 that we allowed above will give an intersection point of C_1 with $x \pm iy = 0$ that has a ramification point above it.

The intersection of a line $y = ax$ with $f(x, y) = 0$ is given by $p_a(x) = f(x, ax) = 0$, and if the point of intersection is a singularity or a point of tangency, then $p_a(x)$ has a multiple root. We can detect such multiple roots by taking the discriminant of $p_a(x)$, which will be a polynomial in a that vanishes if and only if $p_a(x)$ has a multiple root. Hence for all but finitely many values of a the line $y = ax$ has d simple intersection points with $f(x, y) = 0$. So indeed there is an allowed rotation after which π is certain to have a ramification point.

2.3.3 Curves of genus 0, $d \geq 4$

Let $C_1 : f(x, y) = 0$ be an algebraic curve of genus $g_1 = 0$, and again assume that it passes through the origin, but does not have a singularity there. The Riemann-Hurwitz formula with the same map π as above gives

$$g_2 \geq -1 + \frac{1}{2} \sum_{P \in C_2} (e_P - 1),$$

so to get $g_2 \geq 2$ it suffices to show that there are at least 5 ramification points. As above, we can ensure that the lines $x \pm iy$ each have d simple points of intersection. Discounting the intersection point of the two lines, this gives $2d-2$ ramification points. Hence if the degree of f is $d \geq 4$ we are done.

2.3.4 Curves of genus 0, $d = 2, 3$.

Let $f(x, y) = 0$ be an irreducible algebraic curve of genus $g_1 = 0$ and degree $d = 2$ or $d = 3$, and again assume that it passes through the origin, but does not have a singularity there. Consider applying inversion with the origin as center to the curve. This is a birational transformation, so does not change the genus. Therefore, when inversion increases the degree of f to above 4, we are done by the previous section.

Algebraically, inversion in the circle around the origin with radius 1 is given by

$$(x, y) \mapsto \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2} \right),$$

and since this map is its own inverse, the curve $f(x, y) = 0$ is sent to the curve

$$C_3 : (x^2+y^2)^k \cdot f\left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2} \right) = 0,$$

where $k \leq d$ is the lowest integer that makes this a polynomial. This curve is irreducible if and only if the original curve is irreducible. Since f does not have a singularity at the origin, it has a linear term $ax + by$ with a, b not both zero. After inversion this gives a highest degree term

$$(ax + by)(x^2 + y^2)^{k-1}.$$

If $d = 3$ and $k = 3$, then the curve C_3 has degree $2k - 1 = 5$, and we are done.

This does not work if $d = 3$ and $k = 2$, which happens if $x^2 + y^2$ divides the leading terms of f. We will treat this case in a completely different way in the next section.

If $d = 2$, then applying inversion will give a curve of degree 3, unless its leading terms are a multiple of $x^2 + y^2$. But that exactly means that the curve is a circle! So we treat this case by going to the $d = 3$ case, leaving only circles behind, as we should expect. Note that the curves with $d = 3$ resulting from this inversion have $x^2 + y^2$ dividing their leading terms, so are of the type that is handled in the next section.

2.3.5 Curves of genus 0, $d = 3$, $k = 2$.

Since f has genus 0 and degree 3, it must have a singularity (see Chapter 3). The singularity need not be in our rational set, but it is always a rational point, so we can move it to the origin, while maintaining the almost-rational form of the points in our rational set. Then f must have the form

$$(ax+by)(x^2+y^2)+cx^2+dy^2+exy.$$

Note that this is exactly what we get if we apply inversion to a quadratic that is not a circle and goes through the origin.

In fact, we can ensure that $(1,0)$ is on the curve again, so that $a+c=0$. Then if we divide by c, f is of the form

$$(-x+by)(x^2+y^2)+x^2+dy^2+exy.$$

We can parametrize this curve using lines $x = ty$, giving the parametrization

$$y(t)=\frac{t^2+et+d}{(t-b)(t^2+1)}=:\frac{p(t)}{q(t)}, \qquad x(t)=t\cdot y(t).$$

If we let t_j be a value of t that gives one of the points from our rational set, it follows that for infinitely many t,

$$(y(t)-y(t_j))^2+(x(t)-y(t_j))^2=\left(\frac{p(t)}{q(t)}-\frac{p(t_j)}{q(t_j)}\right)^2+\left(t\cdot\frac{p(t)}{q(t)}-t_j\cdot\frac{p(t_j)}{q(t_j)}\right)^2$$

is a square. Then we can multiply by $q(t)^2q(t_j)^2$ to get infinitely many squares of the form

$$\left(p(t)q(t_j)-p(t_j)q(t)\right)^2+\left(tp(t)q(t_j)-t_jp(t_j)q(t)\right)^2.$$

This polynomial has degree 6 in t. It has a factor $(t-t_j)^2$, and a factor t^2+1, since taking $t=\pm i$ gives (using $q(\pm i)=0$)

$$(p(\pm i)q(t_j))^2+(\pm i\cdot p(\pm i)q(t_j))^2=0.$$

Factoring these out, we get a quadratic polynomial $Q_j(t)$ in t. Its leading coefficient is

$$(t_j^2+1)((1+(e+b)^2)t_j^2+2(bd-b+de)t_j+d^2+b^2).$$

and its constant term is

$$(t_j^2+1)((d^2+b^2)t_j^2+2(b^2e+db-d^2b)t_j+b^2e^2+b^2d^2+d^2+2ebd),$$

These polynomials in t_j are not identically zero (if b and d were both 0, then f would be reducible), hence we can pick many t_j so that they are not zero. Then in turn $Q_j(t)$ is a proper quadratic polynomial, and since it is essentially a distance function in the real plane, it cannot have real roots other than the $t = t_j$ that we have already factored out, so it has two distinct imaginary roots.

Therefore an infinite rational set gives infinitely many solutions to equations of the form

$$z_j^2 = (t^2 + 1) \cdot Q_j(t).$$

Multiplying three of these together, and moving $(t^2 + 1)^2$ into the square on the left, we get infinitely many solutions to

$$z^2 = (t^2 + 1)Q_1(t)Q_2(t)Q_3(t).$$

If there are no multiple roots on the right, then this is a hyperelliptic curve of degree 8, so it has genus 3, hence cannot have infinitely many solutions, a contradiction.

The one thing we need to check is that we can choose the t_j so that the Q_j do not have roots in common. We need some notation: write

$$Q_j(t) = c_2(t_j)t^2 + c_1(t_j)t + c_0(t_j),$$

where

$$\begin{aligned}
c_2(t_j) &= (1 + (e + b)^2)t_j^2 + 2(bd + de - b)t_j + d^2 + b^2\\
c_1(t_j) &= 2(bd + de - b)t_j^2 + 2(b^2 + d^2 - bed - bd - be - d)t_j + 2(bd + b^2e - bd^2)\\
c_0(t_j) &= (d^2 + b^2)t_j^2 + 2(b^2e + db - d^2b)t_j + b^2e^2 + b^2d^2 + d^2 + 2ebd.
\end{aligned}$$

Suppose that for infinitely many t_j the polynomial $Q_j(t)$ has the same roots x_1 and x_2. Then for each of those infinitely many t_j we have

$$c_1(t_j) = -(x_1 + x_2) \cdot c_2(t_j), \quad c_0(t_j) = x_1 \cdot x_2 \cdot c_2(t_j),$$

which implies that the corresponding coefficients of these polynomials in t_j are equal. If we look at the coefficients of the linear t_j terms, we see that

$$\begin{aligned}
-x_1 - x_2 &= \frac{2(b^2 + d^2 - bed - bd - be - d)}{2(bd - b + de)} = -b - \frac{be + d - d^2}{bd + de - b},\\
x_1 \cdot x_2 &= \frac{2(b^2e + db - d^2b)}{2(bd + de - b)} = b \cdot \frac{be + d - d^2}{bd + de - b}.
\end{aligned}$$

Here we can read off that the roots are $x_1 = b$ and $x_2 = \frac{be+d-d^2}{bd+de-b}$, which is a contradiction, since the roots had to be imaginary (or since plugging them into $Q_j(t)$ does not always give zero).

2.4 Proof of Theorem 2.2.2

We will prove that if a rational set has infinitely many points on a line, then it can have at most 4 points off the line. The corresponding statement for 3 points off a circle then follows by applying an inversion.

More precisely, suppose we have a rational set S with infinitely many points on a circle C and at least 4 points off that circle. Assume that the origin is one of the points in $S \cap C$, and apply inversion with the origin as center, and with some rational radius. That turns C into a line L, and we get a rational set with infinitely many points on L, and 4 other points. Moreover, the new origin can be added to S, so that we get 5 points off the line, contradicting what we will prove below. To see that the new origin has rational distance to all points in S, observe that in complex notation the distances $|z|$ to the old origin were rational for all $z \in S$, and that the distances to the new origin are $1/|z|$.

To prove the statement for a line, our main tool will again be Faltings' Theorem, but now applied to a hyperelliptic curve

$$y^2 = \prod_{i=1}^{6}(x - \alpha_i),$$

which has genus 2 if and only if the α_i are distinct (see Chapter 3).

Suppose we have a rational set S with infinitely many points on a line, say the x-axis, and 5 or more points off that line. Then we can assume that 3 of those points are above the x-axis, say (a_1, b_1), (a_2, b_2), and (a_3, b_3). Note that we are taking 3 points on one side of the line, because we want to avoid having one point a reflection of another. If we had, say, $(a_2, b_2) = (a_1, -b_1)$, the argument below would break down.

Take a point $(x, 0)$ of S on the x-axis, with $x \neq 0, a_1, a_2, a_3$. Then

$$(x - a_1)^2 + b_1^2, \quad (x - a_2)^2 + b_2^2, \text{and} \quad (x - a_3)^2 + b_3^2$$

are rational squares, so that we get a rational point (x, y) on the curve

$$y^2 = ((x - a_1)^2 + b_1^2)((x - a_2)^2 + b_2^2)((x - a_2)^2 + b_2^2).$$

This is a curve of genus 2, since the roots on the right hand side distinct: they are $x = a_i \pm \sqrt{-b_i^2}$ for $i = 1, 2, 3$, which are distinct by the assumptions on the points (a_i, b_i).

Therefore the curve has genus 2, and cannot contain infinitely many rational points, contradicting the fact that S has infinitely many points on the line.

Chapter 3

Computing the genus

3.1 Introduction

In this chapter we will informally describe an elementary method for computing the genus from the equation for any planar algebraic curve without multiple components. This can often be done with the simple formula $g = (d-1)(d-2)/2$, but this only holds for nonsingular curves. For singular curves, the corresponding formula is

$$g = \frac{(d-1)(d-2)}{2} - \sum_P \delta_P.$$

The trouble here is in the *delta-invariant* δ_P: it is zero for nonsingular points, but it is not always easy to determine for singular points. This formula is mentioned in many books on algebraic geometry (for instance [32], p. 393), but usually no elementary method for computing δ_P in hard cases is given. One exception is the wonderful book [1] (the formula is on p. 148), which was the main source for this chapter.

A specific target we aim for is to compute the genus of hyperelliptic curves (curves of the form $y^2 = f(x)$, where f has no multiple roots) in this way, which we do in Section 3.5.2; these are important in Chapter 2 (in particular in Sections 2.3.5 and 2.4) and in the examples from additive combinatorics that we will give in Section 3.1.2.

It should be noted that the exposition here is intentionally informal, so definitions and claims may not be fully rigorous. In particular, we will not formally define the genus, although the formula above could be taken as a definition, together with the method that we will describe for computing the δ_P. We will also not give proofs for general claims, but we will work out the examples in detail. To be more precise would take up far more space, and be redundant; our goal here is to give a swift and practical introduction (the chapter is based on a writeup for a graduate course, entitled ”Genus for Dummies”).

3.1.1 Definitions

To repeat, by a *curve* we will mean the zero set in $\mathbb{R}^2$ of a polynomial $f \in \mathbb{R}[x, y]$:

$$C_f = C_f(\mathbb{R}) = \{(x, y) \in \mathbb{R}^2 : f(x, y) = 0\}.$$

For number-theoretic questions we mostly consider the set of *rational points* $C_f(\mathbb{Q}) = \{(x, y) \in \mathbb{Q}^2 : f(x, y) = 0\}$ of a curve C_f. In this chapter, we will assume all curves to be irreducible.

An important invariant of a curve is its *genus*, a nonnegative integer g. For nonsingular curves, one can think of it as the number of holes in the surface that we get when we view the curve as a subset of $\mathbb{R}^4$. There are many ways to define and interpret the genus for possibly singular curves, for instance using differential forms, the Hilbert polynomial, or cohomology.

We will try to give a view of the genus of a (possibly singular) curve that is more elementary and doesn't use any of these difficult words. What's more, this will give an algorithm for computing the genus given any irreducible polynomial that is straightforward and usually doable on paper. In particular, this approach works for singular curves without leaving the plane, whereas with the more advanced methods one will have to find an equivalent nonsingular curve, which may live in higher dimensions. On the other hand, the more advanced methods will often be more insightful, and generalize better to non-planar curves and higher-dimensional varieties.

For number theory, the interest in the genus lies in the following classification:

$$\begin{aligned} g = 0 \quad &\Longleftrightarrow \quad \text{rational curves} \\ g = 1 \quad &\Longleftrightarrow \quad \text{elliptic curves} \\ g \geq 2 \quad &\Longrightarrow \quad C(\mathbb{Q}) \ \text{ is finite} \end{aligned}$$

So when a curve has genus 0, it is *rational*, i.e. birationally equivalent to a projective line, hence we can parametrize it by rational functions, which makes the curve easy to understand, if not boring.

Curves with genus 1 are *elliptic curves*, which have a group structure on $C(\mathbb{Q})$, and with that we can understand its set of points very well. Elliptic curves occur as objects of study in Chapter 5, but in this thesis we will not use any elliptic curve techniques (except that in the next section we mention their outcome as an illustration).

When the genus is 2 or higher, it is not so easy to understand the rational points, but one does know that there are only finitely many, which is what Faltings' Theorem (Theorem 2.3.1) says.

An important class of curves is that of *hyperelliptic curves*, which have the form

$$y^2 = \varphi(x), \quad \varphi \in \mathbb{Q}[x] \text{ no multiple roots and monic.}$$

Their form makes them convenient to handle, and there is an easy formula for their genus:

$$\deg \varphi = 2k+1 \ \text{ or } \ 2k+2 \quad \Longrightarrow \quad g = k.$$

We used this formula in Sections 2.3.5 and 2.4. We will prove it in Section 3.5.2.

3.1.2 An application: progressions in powers

Here is an example of a problem from additive combinatorics where being able to compute the genus of singular curves comes in handy. These are not new results or techniques (see for instance [13] and [14] for a similar approach). We will not fully prove the conclusions, especially since we do not want to include elliptic curve methods in this thesis; we will merely indicate where elliptic curve methods could be applied. But these calculations should illustrate why this approach is useful for discrete problems; in fact they were the inspiration for our work in Chapter 2 (in particular see Sections 2.3.5 and 2.4).

Suppose we have an arithmetic progression of length 3 (a 3AP) that consists of squares (in the integers). Then we can write them in the form x, $x+d$, $x+2d$, with $d \neq 0$, so that since they are squares, multiplying them all together gives a square, and we have a solution to

$$y^2 = x(x+d)(x+2d).$$

As we saw in Section 3.3.1, this curve has genus $g = 1$, so it is an elliptic curve, and we could find out a lot about its solutions.

This is merely a first idea, and not too fruitful, but there are many variants. For instance, given a 4-term AP of squares x^2, x^2+d, x^2+2d, x^2+3d, we have a solution to

$$y^2 = (x^2+d)(x^2+2d)(x^2+3d),$$

which is a hyperelliptic curve with genus 2 (as we will prove in 3.5.2), hence we immediately know that there are only finitely many 4APs with given common difference in the squares.

That's still not too exciting, because with a bit more ingenuity we could show that there are no 4APs whatsoever in the squares, as was first proven by Euler (see [15]) using very different methods. Let x_1^2, x_2^2, x_3^2, x_4^2 form a 4AP. That means that

$$x_2^2 - x_1^2 = x_3^2 - x_2^2 = x_4^2 - x_3^2$$

$$\Rightarrow x_1^2 = 2x_2^2 - x_3^2, \quad x_4^2 = 2x_3^2 - x_2^2$$

$$\Rightarrow \left(\frac{x_1}{x_3}\right)^2 = 2\left(\frac{x_2}{x_3}\right)^2 - 1, \quad \left(\frac{x_4}{x_3}\right)^2 = 2 - \left(\frac{x_2}{x_3}\right)^2,$$

hence we have a rational point on the curve

$$y^2 = (2x^2 - 1)(2 - x^2).$$

This is an elliptic curve (as we will see in Section 3.5.2; note that the formula $(d-1)(d-2)/2$ would give the wrong genus). Basic elliptic curve methods would tell us that this curve has only a few rational points, and these only give trivial 4APs, proving that there are no 4APs in the squares (see [14], where this is done via a different curve).

We can also apply this approach to progressions of higher powers; we only need to look at 3APs. If x_1^k, $x_2^k = x_1^k + d$, $x_3^k = x_1^k + 2d$ is a 3AP of kth powers, then we have

$$2x_2^k - 2x_1^k = 2d = x_3^k - x_1^k \quad \Rightarrow \quad 2x_2^k = x_1^k + x_3^k \quad \Rightarrow \quad 2 = \left(\frac{x_1}{x_2}\right)^k + \left(\frac{x_3}{x_2}\right)^k,$$

so we get a rational point on the curve

$$x^k + y^k = 2.$$

We'll see in Section 3.3.1 that for $k = 3$, this curve has genus 1, so we could determine the solutions. For $k \geq 4$, we have genus ≥ 2, hence there are only finitely many 3APs of such powers. That also implies there are no arbitrarily long progressions for higher powers.

Actually, using Wiles-type methods it has been shown [16] that $x^k + y^k = 2z^k$ has no nontrivial integer solutions for $k \geq 3$, which implies that there are no 3APs of higher powers at all.

3.2 Basics from algebraic geometry

We will informally introduce a few notions from algebraic geometry that are necessary here. We follow the notation of [48].

3.2.1 Projective curves

The most natural setting for algebraic curves is actually not the affine plane $\mathbb{C}^2$, but the projective plane $\mathbb{P}^2(\mathbb{C})$, which is defined by taking $\mathbb{C}^3$, identifying two points if they are scalar multiples of each other, and removing $(0, 0, 0)$. As notation for an equivalence class of points we will write $[x : y : z]$, where (x, y, z) is some representative from the class. The points $[x : y : 1]$ will be considered as the points from the "affine" plane, and the points $[x : 1 : 0]$ and $[1 : 0 : 0]$ make up the projective line at infinity.

Given a curve $f \in \mathbb{C}[x, y]$ with $\deg f = d$, we can view it projectively by homogenizing the polynomial:

$$F(X, Y, Z) = Z^d \cdot f(X/Z, Y/Z);$$

then the curve is given projectively as the points $[X : Y : Z] \in \mathbb{P}^2(\mathbb{C})$ where $F(X, Y, Z) = 0$.

If we want to take a closer look at how f behaves around a point at infinity like $P = [0 : 1 : 0]$, we put $g(x, z) = F(x, 1, z)$, and consider $g = 0$ around $P = (0, 0)$ in the xz-plane.

3.2.2 Morphisms and rational maps

We'll need to know what a rational map is to be able to make sense of the important statement "genus is a birational invariant of a curve". We say that a map

$$f : C_1 \to C_2, \quad (x, y) \mapsto (f_1(x, y), f_2(x, y))$$

is a *morphism* if f_1 and f_2 are polynomials from $\mathbb{Q}[x, y]$, and an *isomorphism* if it has an inverse which is also a morphism.

Basically, such a map is a *rational map* if f_1 and f_2 are not polynomials but rational functions from $\mathbb{Q}(x, y)$, though that doesn't quite make sense because a rational function might have poles. So a rational map is only a partial map, defined on all but finitely many points. A *birational equivalence* is then a rational map with an inverse that's also a rational map.

Note that we are leaving out a number of subtleties, e.g. that because these maps are only defined at points satisfying some equation, what looks like a rational function might actually be a polynomial. The thing that matters here is that birational equivalences do not change the genus. This can be seen pretty easily from the differential form concept of genus, but is harder to prove in an elementary way.

3.2.3 Singularities

The affine curve given by f has a *singularity* at (a, b) if $f(a, b) = 0$ and the partial derivatives $f_x(a, b) = 0$ and $f_y(a, b) = 0$. Projectively, F has a singularity at $[A : B : C]$ if $F(A, B, C) = 0$, $F_X(A, B, C) = 0$, $F_Y(A, B, C) = 0$, and $F_Z(A, B, C) = 0$ (this might seem like one more condition, but a homogeneous polynomial satisfies $XF_X + YF_Y + ZF_Z = \deg(F) \cdot F$). For example, the curves $y^2 = x^3$ and $y^2 = x^2(x + 1)$ both have a singularity at $(0, 0)$.

A hyperelliptic curve $y^2 = \varphi(x)$ has no affine singularity, since it would have to have $y = 0$ and x a root of φ', but by definition φ and φ' have no root in common. However, if $d = \deg \varphi > 3$, it does have a singularity when $Z = 0$: since its homogenization is $F = Y^2Z^{d-2} - Z^d\varphi(X/Z)$, we get $X = 0$ from $F_X = 0$, and then $F_Y = F_Z = 0$ follow, hence it has the singularity $[0 : 1 : 0]$ at infinity.

3.2.4 Multiplicity of a singularity

With a singularity we can associate a number m that says how "bad" it is. Assume the singularity is $(0, 0)$ and write $f = \sum a_{ij}x^iy^j$. Then having a singularity at $(0, 0)$ is the same as having $a_{00} = a_{01} = a_{10} = 0$. The multiplicity is defined by

$$m = \min\{k : i + j = k, a_{ij} \neq 0\},$$

i.e. the lowest total order of a nonzero term. For instance, $x^4 - y^4 + 3y^7x^2$ has a singularity with multiplicity $m = 4$ at $(0, 0)$.

3.3 Computing the genus, Part I

3.3.1 Easy Cases

Now we can explain how to compute the genus in the easier cases. First of all, if the curve is nonsingular, there is a simple formula for the genus in terms of the degree d of the curve:

$$g = \frac{(d - 1)(d - 2)}{2}.$$

So for instance a curve of the form $y^2 = x^3 + ax + b$, with no multiple roots on the right, has no singularities, hence has genus $\frac{(3-1)(2-1)}{2} = 1$, so is an elliptic curve. As a quick consequence, we see that there are no nonsingular plane curves of genus 2, since $(d-1)(d-2)/2$ never equals 2; the same holds for other values not assumed by $(d-1)(d-2)/2$.

For the curves $x^k + y^k = 2$ that we obtained for 3APs of kth powers in 3.1.2, there are no singularities, so we can apply the formula above to see that the genus is $(k-1)(k-2)/2$, which gives the values that we used in 3.1.2.

We can extend this formula to curves with only singularities that have all tangents distinct. More precisely, if we move the singularity to the origin and it has multiplicity m, and we write

$$f = \sum_{i+j=m} a_{ij}x^iy^j + \sum_{i+j>m} a_{ij}x^iy^j = \prod_{k=1}^{m}(\alpha_k x + \beta_k y) + \sum_{i+j>m} a_{ij}x^iy^j,$$

then the tangents at the origin are the lines $\alpha_k x + \beta_k y = 0$. If these tangents are all distinct lines, then the genus of f is given by

$$g = \frac{(d-1)(d-2)}{2} - \sum_P \frac{m_P(m_P-1)}{2}.$$

Here m_P is the multiplicity of the curve at the point P, which is 1 when P is on the curve but not a singularity, and 0 when P is not on the curve, so this sum really only runs over the singularities.

For instance, if $f = y^2 - x^3 - x^2$, then its only singularity is the origin, where it has distinct tangents $x \pm y = 0$ and $m = 2$, hence its genus is $\frac{(3-1)(2-1)}{2} - \frac{2\cdot 1}{2} = 0$.

3.3.2 Not all cases are easy

Here is an example of a curve for which the above formula fails:

$$y^2 = x^4 - x^5.$$

It has a singularity with $m = 2$ at the origin (but not with distinct tangents), and it has a singularity at $[0:1:0]$, where its equation is $z^3 = x^4z - x^5$, so $m = 3$. Then the formula above would give $g = \frac{4\cdot 3}{2} - \frac{2\cdot 1}{2} - \frac{3\cdot 2}{2} = 6 - 1 - 3 = 2$. However, we can see that the genus must be 0: we can rewrite the equation to $y = x^2\sqrt{1-x}$, hence $x = \cos^2\theta$, $y = \sin\theta\cos^4\theta$ is a parametrization. Since we can rationally parametrize cos and sin by $\cos\theta = \frac{1-t^2}{1+t^2}$, $\sin\theta = \frac{2t}{1+t^2}$, we can rationally parametrize this curve, hence its genus must be zero. We will see in the next section how to get this genus right.

3.4 Blowing up singularities

3.4.1 Local blowup

To extend the formula above, we need to "blow up" the singularities. The idea is to untangle the singularity by separating the branches of the curve according to their tangents.

For simplicity, we will illustrate this process for the example

$$C : y^2 = x^2(x+1),$$

by blowing up its singularity at the origin. We will end up with a birational equivalence $C'' \to C$ from a nonsingular curve C'', but we will need several steps to construct it. Fortunately, afterwards we will see that the equation for C'' is quite easy to calculate.

Define the "blowup surface"

$$B = \{(x, y, t) : y = xt\} \subset \mathbb{C}^3,$$

with the corresponding projection $\pi : B \to \mathbb{C}^2$, $(x, y, t) \mapsto (x, y)$. Then the inverse images of points under this map are pretty simple:

$$\pi^{-1}(x, y) = \begin{cases} \{(x, y, \frac{y}{x})\} & \text{if } x \neq 0 \\ \{(0, 0, t) : t \in \mathbb{C}\} & \text{if } x = 0 \end{cases}.$$

So π is bijective away from the origin, and collapses the line

$$L = \{(0, 0, t) : t \in \mathbb{C}\}$$

to the origin. The inverse image of C under π is

$$\begin{aligned} \pi^{-1}(C) &= \{(x, y, t) : y = xt,\ y^2 = x^2(x+1)\} = \{(x, xt, t) : x^2t^2 = x^2(x+1)\} \\ &= \{(x, xt, t) : x^2 = 0\} \cup \{(x, xt, t) : t^2 = x+1\} \\ &= L \cup C', \end{aligned}$$

where

$$C' = \{(x, y, t) : y = xt,\ t^2 = x+1\}$$

is a new curve in $\mathbb{C}^3$. Now the restricted map $\pi : C' \to C$ is a birational equivalence, which is bijective away from the line L, and sends the two points $(0, 0, \pm 1)$ on L to the origin. We say that these points "lie above" the singularity $(0, 0)$ of C.

We can simplify this map a bit. We project C' onto a curve in xt-space with $\rho : C' \to C''$, $(x, y, t) \mapsto (x, t)$, where

$$C'' = \{(x, t) : t^2 = x + 1\}.$$

This projection has a well defined inverse $\rho^{-1} : (x, t) \mapsto (x, xt, t)$. Then the blowup of C, with center the origin, is the map

$$\pi \circ \rho^{-1} : C'' \to C, \quad (x, t) \mapsto (x, xt).$$

This blowup is a birational map from a nonsingular curve to our original curve, which is bijective away from the origin, and the inverse image of the singularity at the origin consists of the two nonsingular points $(0, 1)$ and $(0, -1)$. If we took a closer look, we would see that these points correspond to the two tangent directions the original curve had there; the blowup has "untangled" the singularity.

In practice, we do not have to consider all these maps explicitly: given the curve $y^2 = x^2(x + 1)$, we plug in $y = xt$ to get $t^2x^2 = x^2(x + 1)$; then cancel out x^2 (which corresponds to removing the line $x^2 = 0$) to get $t^2 = x + 1$, which is the blown-up curve. Since this curve is nonsingular, its genus is $\frac{(2-1)(2-2)}{2} = 0$, hence our original curve also has genus 0.

To show this for another example, let's blow up the singularity at the origin of the cusp $y^2 = x^3$. We plug in $y = xt$ to get $x^2t^2 = x^3$, then remove the new line $x^2 = 0$ to obtain the blowup $t^2 = x$, which is nonsingular and birational to $y^2 = x^3$. Since $t^2 = x$ has genus 0, the genus of $y^2 = x^3$ must also be 0.

In general, the points lying above the blown-up singularity are the points of C'' that are on the t-axis. Note that it is possible for these to be singularities as well. For example, if we start with $y^2 = x^5$, then blowing up gives $t^2 = x^3$, which still has a singularity at the origin. But another blowup (with for instance $t = xu$) then gives the birational nonsingular curve $u^2 = x$ with genus 0, hence $y^2 = x^5$ has genus 0 as well. More generally, for $y^2 = x^{2k+1}$, blowing up k times will show that the genus is 0. In fact, one can show that any singularity can be resolved in finitely many steps this way ([1], p.137).

Finally, note that the curve C should not have a vertical tangent, since this would correspond to a point on C'' with $t = \infty$. But this can be easily avoided by applying a rational rotation.

3.4.2 Resolutions

The above kind of blowup is not quite satisfactory, because it does not behave well on the line at infinity; in fact, it might *create* new singularities. For instance, the curve $x^4 + y^3 + y^2 - x^2$ has a singularity at the origin, and no others, including at infinity. Blowing up like above gives $x^2 + xt^3 + t^2 - 1$, which is nonsingular in the affine plane, but has a new singularity at infinity, at $[1 : 0 : 0]$.

This can be avoided, by doing projective blowups and gluing them together. We won't do this, but the result is a *global blowup*: a projective birational map which creates no new singularities, and "improves" at least one of the singularities, by which we mean that on an affine neighborhood around that singularity it is just a local blowup like above.

The main result is then that any curve C has a *resolution*

$$C \leftarrow C_1 \leftarrow \cdots \leftarrow C_n,$$

where each map is a global blowup, and C_n is nonsingular ([1], p.137, or [32], p. 391).

So one way to find the genus of any curve would be to find this resolution, and then just compute $\frac{(d-1)(d-2)}{2}$ for C_n. However, finding that resolution is not so easy. Luckily there is a shortcut: locally, on an affine neighborhood of a singularity, these global blowups act just like local blowups. So if we can find the "contribution" to the genus of every such local blowup, putting all of those together we can compute the genus of C, without having to do all the global blowups.

In other words, for each singularity P, we repeatedly apply local blowups, until we have resolved it, i.e. the singularity is replaced by nonsingular points. We refer to the singularities Q that appear during this process as *lying above* P. Note that this needn't be a linear process like the resolution above: blowing up a singularity might split it into two new singularities (or more), and then one has to continue by blowing up each of those. But again one can prove that this process will terminate in finitely many steps.

For each such singularity Q lying above P, take its multiplicity m_Q, and compute

$$\delta_P = \sum_Q \frac{m_Q(m_Q - 1)}{2}.$$

As it turns out, this quantity is what we have to subtract from the degree formula to get the correct genus:

$$g = \frac{(d-1)(d-2)}{2} - \sum_P \delta_P.$$

A different way of looking at it is to think of these Q as points "infinitely near" P, and then the genus formula is

$$g = \frac{(d-1)(d-2)}{2} - \sum_Q \frac{m_Q(m_Q-1)}{2},$$

where the sum runs over all infinitely near points Q.

3.5 Computing the genus, Part II

3.5.1 Examples

Let's consider $y^2 = x^4 - x^5$ from Section 3.3.2 again. We showed that it has a singularity with $m = 2$ at the origin, and one with $m = 3$ at $[0 : 1 : 0]$.

First let's blow up the origin. Putting $y = xt$ gives $x^2t^2 = x^4 - x^5$, hence the new curve is $t^2 = x^2 - x^3$. This has a singularity at the origin with $m = 2$ (and we should check that there are no others on the t-axis), and there the tangent lines are $t \pm x = 0$, so we don't have to blow up again, because there wouldn't be any new singularities. So the origin has $\delta = \frac{2\cdot 1}{2} + \frac{2\cdot 1}{2} = 2$.

For the singularity at infinity, the local equation is $z^3 = x^4z - x^5$, and blowing up with $z = xt$ gives $t^3 = x^2t - x^2$, which has a singularity at $(0, 0)$ with $m = 2$. Blowing up again with $t = xs$, we get $xs^3 = xs - 1$, which has no singularities on the s-axis, so we are done. Hence the singularity at infinity has $\delta = \frac{3\cdot 2}{2} + \frac{2\cdot 1}{2} = 4$. Now the genus is

$$\frac{(5-1)(5-2)}{2} - 2 - 4 = 0,$$

as we expected. Apparently, when we miscalculated this genus to be 2 in Section 3.3.2, we missed out on the two infinitely near points, for which we should have subtracted 1 each.

For another example, consider the four-leaved clover $(x^2+y^2)^3 = x^2y^2$. It has a singularity at the origin, and at $[1:i:0]$ and $[1:-i:0]$. Those two at infinity can easily be seen to have $\delta = 1$.

The singularity with $m = 4$ at the origin is more complicated. Since it has vertical and horizontal tangents, we will have to rotate the curve before we can blow up. Put for instance $x = u+v$, $y = u-v$, so that $x^2+y^2 = 2(u^2+v^2)$, and $xy = u^2-v^2$, so that the curve becomes

$$8(u^2+v^2)^3 = (u^2-v^2)^2.$$

Now its tangents at the origin are $u \pm v = 0$, each double. We blow up with $v = tu$, which gives the new curve

$$2u^2(1+t^2)^3 = (1-t^2)^2.$$

We only care about its points on the t-axis, which are at $t = \pm 1$, both singular with $m = 2$. To analyze them we need to move them to the origin, so put for instance $t = w+1$, then we get

$$2u^2(w^2+2w+2)^3 = w^2(w+2)^2.$$

This has lowest terms $16u^2 - 4w^2$, so here the curve has distinct tangents, and we do not need to blow up any further. Same for $t = -1$. Hence the origin of the clover has $\delta = \frac{4\cdot 3}{2} + \frac{2\cdot 1}{2} + \frac{2\cdot 1}{2} = 8$.

Hence the curve has genus $(6-1)(6-2)/2 - 1 - 1 - 8 = 0$. That makes sense, as this is a polar curve ($2r = \sin 2\theta$), hence has a parametrization.

Note that both these examples happen to have genus 0, which is of course not representative. But they show how complicated singularities can get, and we can check our genus calculation by finding a parametrization.

3.5.2 Hyperelliptic curves

Finally we will prove the following formula for the genus g of a hyperelliptic curve $y^2 = \varphi(x)$, where φ has no multiple roots and is monic:

$$\deg(\varphi) = 2k+1 \text{ or } 2k+2 \quad \Longrightarrow \quad g = k.$$

We only need to consider $d \geq 3$. First we assume that $\deg \varphi$ is even, so that we can write the equation as $y^2 - x^{2k+2} - h(x) = 0$, $\deg h = l \leq 2k+1$. Then homogenizing and setting $y = 1$ gives

$$z^{2k} - x^{2k+2} - z \cdot H(x,z) = 0,$$

where $H(x,z) = z^{2k+1-l}h(x/z)$ is homogeneous and $\deg H = 2k+1$. Then the singularity is at $(0,0)$ and has multiplicity $m = 2k$.

Applying the blowup $z = xt$ and factoring out x^{2k} gives

$$t^{2k} - x^2 - x^2 t \cdot H_0 = 0,$$

where $H_0 = H(x,xt)/x$ is a polynomial since H is homogeneous. The only singularity on the t-axis is at $(0,0)$, and has $m = 2$. Next we apply $x = st$ and factor out t^2 to get

$$t^{2k-2} - s^2 - s^2 t \cdot H_0 = 0.$$

Again we have a singularity with $m = 2$. Repeating this, we see that we get k infinitely near singularities with $m = 2$. Hence the genus is

$$\begin{aligned} g &= \frac{(2k+1)2k}{2} - \frac{2k(2k-1)}{2} - k \cdot \frac{2 \cdot 1}{2} \\ &= k\left((2k+1) - (2k-1) - 1\right) = k, \end{aligned}$$

as desired.

For $\deg \varphi = 2k+1$ the calculation is similar. From

$$z^{2k-1} - x^{2k+1} - z \cdot H(x,z) = 0$$

with $m = 2k-1$ we get

$$t^{2k-1} - x^2 - x^2 t \cdot H_0 = 0$$

with a singularity with $m = 2$. Repeating, we get $k-1$ such singularities with $m = 2$, so that the genus is

$$\begin{aligned} g &= \frac{2k(2k-1)}{2} - \frac{(2k-1)(2k-2)}{2} - (k-1) \cdot \frac{2 \cdot 1}{2} \\ &= (2k-1)\left(k - (k-1)\right) - (k-1) = k. \end{aligned}$$

Chapter 4

Rational distances with rational angles

4.1 Background

A famous problem of Erdős from 1946 [18] concerns the maximum number of unit distances among n points in the plane; we will denote this number by $u(n)$. He showed that $u(n) > n^{1+c/\log\log n}$, using a $\sqrt{n} \times \sqrt{n}$ piece of a scaled integer lattice, and conjectured that this was the true magnitude. The best known upper bound is $u(n) < cn^{4/3}$, first proved by Spencer, Szemerédi and Trotter in 1984 [51]. This bound has several other proofs, the simplest of which was the proof by Székely [52], using a lower bound for the crossing number of graphs (the very same that we use in Section 5.3). A recent result of Matoušek [42] shows that the number of unit distances is bounded above by $cn \log n \log\log n$ for most norms. As a general reference for work done on the unit distances problem, see [6].

We will show that the upper bound $n^{1+6/\sqrt{\log n}}$ holds if we only consider unit distances that have *rational angle*, by which we mean that the line through the pair of points makes a rational angle in degrees with the x-axis (or equivalently, its angle in radians, divided by π, is rational). Under this restriction, we can use an algebraic theorem of Henry Mann [41], Theorem 4.3.1, to get a uniform bound on the number of paths between two fixed vertices in the unit distance graph, which will lead to a contradiction if there are too many unit distances with rational angle between the points.

In fact, our proof also shows that the bound $n^{1+6/\sqrt{\log n}}$ holds for the number of rational distances with rational angles, if we have no three points on a line. The lower bound, $n^{1+c/\log\log n}$, of Erdős does not apply in this case as we are restricted to rational angles. But a construction of Erdős and Purdy [25] gives a superlinear lower bound for unit (and hence rational) distances with rational angles (see Section 4.5).

If instead we allow up to n^α points on a line where $1/2 \leq \alpha \leq 1$, the number of rational distances with rational angles is bounded by $4n^{1+\alpha}$. This bound is tight up to a constant factor with the lower bound now coming from an $n^{1-\alpha} \times n^{\alpha}$ square grid. If we allow up to n^{α} points on a line where $0 < \alpha < 1/2$, the number of rational distances with rational angles is bounded above by $n^{1+\alpha+6/\sqrt{\log n}}$. We get a lower bound of $cn^{1+\alpha}$ from $n^{1-\alpha}$ horizontal lines each containing n^{α} rational points so that no three points on different lines are collinear (see Section 4.5).

In Section 4.2 we will state our main results and give an outline of the proof. Section 4.3 contains the algebraic tools that we will use, including, for completeness, a proof of Mann's Theorem. In Section 4.4 we use the bounds obtained from Mann's Theorem and some graph theory to prove our main results. In Section 4.5 we give lower bounds for the main results.

4.2 Main results and proof sketch

We will say that a pair of points in $\mathbb{R}^2$ *has rational angle* if the line segment between them, viewed as a complex number $z = re^{\pi i \gamma}$, has $\gamma \in \mathbb{Q}$. Our first result is the following.

Theorem 4.2.1. *Given n points in $\mathbb{R}^2$, the number of pairs of points with unit distance and rational angle is at most $n^{1+6/\sqrt{\log n}}$.*

Roughly speaking, our proof goes as follows. Given n points in the plane, we construct a graph with the points as vertices, and as edges the unit line segments that have rational angle. We can represent these unit line segments as complex numbers, which must be roots of unity because of the rational angle condition. Then if this graph has many edges, it should have many cycles of a given length k, and each such cycle would give a solution to the equation

$$\sum_{i=1}^{k} \zeta_i = 0,$$

with ζ_i a root of unity. Using Mann's Theorem, we could give a uniform bound on the number of such solutions, depending only on k (under the non-degeneracy condition that no subsum vanishes). If the number of non-degenerate cycles goes to infinity with n, this would give a contradiction.

However, dealing with cycles of arbitrary length is not so easy, so instead in our proof we count non-degenerate paths of length k between two fixed vertices, which correspond to solutions of the equation

$$\sum_{i=1}^{k} \zeta_i = a,$$

where $a \in \mathbb{C}, a \neq 0$, corresponds to the line segment between the two points. We have extended Mann's Theorem to this type of equation, giving a similar upper bound and proving our result.

In fact, in our proof it turns out that it is not necessary for the lengths to be 1, but that they only need to be rational. This is because our extension of Mann's Theorem also works for equations of the type

$$\sum_{i=1}^{k} a_i \zeta_i = a,$$

where $a_i \in \mathbb{Q}$ and $a \in \mathbb{C}, a \neq 0$. This leads to the following results (the first supersedes 4.2.1; we have stated both because 4.2.1 answers our initial question).

Theorem 4.2.2. *Suppose we have n points in $\mathbb{R}^2$, no three of which are on a line. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+6/\sqrt{\log n}}$.*

Theorem 4.2.3. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $0 < \alpha < 1/2$. Then the number of pairs of points with rational distance and rational angle is at most $n^{1+\alpha+6/\sqrt{\log n}}$.*

Theorem 4.2.4. *Suppose we have n points in $\mathbb{R}^2$, with no more than n^α on a line, where $1/2 \leq \alpha \leq 1$. Then the number of pairs of points with rational distance and rational angle is at most $4n^{1+\alpha}$.*

The constants 6 and 4 in these theorems are not optimal, but they are the smallest integers that followed directly from our proof.

4.3 Mann's Theorem

For completeness we provide a proof of Mann's Theorem. We then prove the extension that we will need to prove the main result in the next section.

Theorem 4.3.1 (Mann)**.** *Suppose we have*

$$\sum_{i=1}^{k} a_i \zeta_i = 0,$$

with $a_i \in \mathbb{Q}$, the ζ_i roots of unity, and no subrelations $\sum_{i \in I} a_i \zeta_i = 0$ where $\emptyset \neq I \subsetneq [k]$. Then

$$(\zeta_i/\zeta_j)^m = 1$$

for all i, j, with $m = \prod_{\substack{p \leq k \\ p \text{ prime}}} p$.

Proof. We can assume that $\zeta_1 = 1$ and $a_1 = 1$, so that we have $1 + \sum_{i=2}^{k} a_i \zeta_i = 0$. We take a minimal m such that $\zeta_i^m = 1$ for each i. We will show that m must be squarefree, and that a prime p that divides m must satisfy $p \leq k$. Together these facts prove the theorem.

Let p be a prime dividing m. Write $m = p^j \cdot m^*$ with $(p, m^*) = 1$, and use that to factor each ζ_i as follows:

$$\zeta_i = \rho^{\sigma_i} \cdot \zeta_i^*,$$

with ρ a primitive p^jth root of unity so

$$\rho^{p^j} = 1, \quad (\zeta_i^*)^{p^{j-1} m^*} = 1, \quad 0 \leq \sigma_i \leq p-1.$$

Now reorganize the equation as follows:

$$0 = 1 + \sum_{i=2}^{k} a_i \zeta_i = 1 + \sum_{l=0}^{p-1} \alpha_\ell \rho^\ell = f(\rho),$$

where the coefficients are of the form

$$\alpha_\ell = \sum_{i \in I_\ell} a_i \zeta_i^* \in \mathbb{Q}(\zeta_2^*, \ldots, \zeta_k^*) = K,$$

with $I_\ell = \{i \in [k] : \sigma_i = \ell\}$. So f is a polynomial over the field K of degree $\leq p-1$ and $f(\rho) = 0$. The polynomial f isn't identically zero, since that would give a subrelation containing strictly fewer than k terms. To see this, observe that we must have $\sigma_i \geq 1$ for at least one i, otherwise $\zeta_i^{m/p} = 1$ for each i, contradicting the minimality of m.

But we can compute the degree of ρ over K to be

$$\deg_K(\rho) = \frac{\phi(m)}{\phi(p^{j-1} m^*)} = \frac{\phi(p^j)}{\phi(p^{j-1})} = \begin{cases} p-1 & \text{if } j = 1 \\ p & \text{if } j > 1. \end{cases}$$

This is a contradiction unless $j = 1$, which proves that m is squarefree.

Knowing that m is squarefree, we have $m = p \cdot m^*$ with $(p, m^*) = 1$, and

$$\zeta_i = \rho^{\sigma_i} \cdot \zeta_i^*, \quad \rho^p = 1, \quad (\zeta_i^*)^{m^*} = 1, \quad 0 \leq \sigma_i \leq p-1.$$

Still $f(\rho) = 0$ for $f(x)$ a polynomial over K, not identically zero. But we know ([39], Ch. VI.3) that the minimal irreducible polynomial of ρ over K is $F(x) = x^{p-1} + x^{p-2} + \cdots + x + 1$, hence we must have $f(x) = cF(x)$ for some $c \in K$. In particular, f has p terms, which implies that our original relation had at least p terms, so $k \geq p$. □

In our proof in the next section we will not use Mann's Theorem itself, but the following modified version and its corollary.

Theorem 4.3.2. *Suppose we have*

$$\sum_{i=1}^{k} a_i\zeta_i = a, \qquad \sum_{j=1}^{k} a_j^*\zeta_j^* = a,$$

with $a \in \mathbb{C}, a \neq 0$, $a_i \in \mathbb{Q}$, *roots of unity* ζ_i, *and no subrelations* $\sum_{i\in I} a_i\zeta_i = 0$ *or* $\sum_{j\in J} a_j^*\zeta_j^* = 0$ *where* $\emptyset \neq I \subsetneq [k]$ *and* $\emptyset \neq J \subsetneq [k]$. *Then for any* ζ_j^* *there is a* ζ_i *such that*

$$\left(\zeta_j^*/\zeta_i\right)^m = 1$$

with $m = \prod_{\substack{p\leq 2k \\ p \text{ prime}}} p$.

Proof. We have $\sum a_i\zeta_i = a = \sum a_j^*\zeta_j^*$, which gives the single equation

$$\sum_{i=1}^{k} a_i\zeta_i - \sum_{j=1}^{k} a_j^*\zeta_j^* = 0. \tag{4.1}$$

Mann's Theorem does not apply immediately, because there might be subrelations. But we can break the equation up into minimal subrelations

$$\sum_{i\in I_\ell} a_i\zeta_i - \sum_{j\in I_\ell^*} a_j^*\zeta_j^* = 0, \tag{4.2}$$

where each $I_\ell \neq \emptyset$, $I_\ell^* \neq \emptyset$, and there are no further subrelations.

Given ζ_j^*, there is such a minimal subrelation of length $\leq 2k$ in which it occurs, and which must also contain some ζ_i. Applying Mann's Theorem to this equation gives $\left(\zeta_j^*/\zeta_i\right)^m = 1$ with $m = \prod_{\substack{p\leq 2k \\ p \text{ prime}}} p$. □

Note that in the above proof we require $a \neq 0$. If $a = 0$ and there is no proper subrelation as in (4.2) then (4.1) still has the subrelations

$$\sum_{i=1}^{k} a_i\zeta_i = 0, \qquad \sum_{j=1}^{k} a_j^*\zeta_j^* = 0,$$

so we cannot use Mann's Theorem to get a relation between a ζ_i and ζ_j^*.

For $a \in \mathbb{C}, a \neq 0, k \in \mathbb{Z}, k > 0$ we define Z_a^k to be the set of k-tuples of roots of unity $(\zeta_1, \ldots, \zeta_k)$ for which there are $a_i \in \mathbb{Q}$ such that $\sum_{i=1}^k a_i\zeta_i = a$ with no subrelations, i.e.:

$$Z_a^k = \{(\zeta_1, \ldots, \zeta_k) \mid \exists a_i \in \mathbb{Q} : \sum_{i=1}^k a_i\zeta_i = a, \sum_{i\in I} a_i\zeta_i \neq 0 \text{ for } \emptyset \neq I \subset [k]\}.$$

Corollary 4.3.3. *Given $a \in \mathbb{C}$ with $a \neq 0$, we have $|Z_a^k| \leq (k \cdot C(k))^k$, where $C(k) = \prod_{\substack{p \leq 2k \\ p \text{ prime}}} p$.*

Proof. Fix an element $(\zeta_1, \ldots, \zeta_k) \in Z_a^k$ and let $m = C(k)$ and $M_i = \zeta_i^{-m}$ for $1 \leq i \leq k$. Then for ζ_j^* in any element of Z_a^k, we have an i such that $M_i \left(\zeta_j^*\right)^m = 1$. In other words ζ_j^* is a solution of $M_i x^m = 1$. Each of these k equations has $m = C(k)$ solutions, hence there are at most $k \cdot m = k \cdot C(k)$ choices for each ζ_j^*. □

4.4 Proofs of main theorems

4.4.1 Preparation

We are now in a position to prove the main results.

Suppose we have a graph $G = G(V, E)$ on $v(G) = n$ vertices and $e(G) = cn^{1+\alpha}$ edges. We will denote the minimum degree in G by $\delta(G)$. The following lemma assures us that we can remove low-degree vertices from our graph without greatly affecting the number of edges.

Lemma 4.4.1. *Let G be as above. Then G contains a subgraph H, with $e(H) \geq (c/2)n^{1+\alpha}$ edges and $v(H) \geq (c/2)n^{\alpha}$ vertices, such that the minimum degree $\delta(H) \geq (c/2)n^{\alpha}$.*

Proof. We iteratively remove vertices from G of degree less than $(c/2)n^{\alpha}$. Then the resulting subgraph H has $\delta(H) \geq (c/2)n^{\alpha}$, and we removed fewer than $(c/2)n^{1+\alpha}$ edges, so H contains more than $(c/2)n^{1+\alpha}$ edges. □

Suppose we are given a path $P_k = p_0 p_1 \dots p_k$ on k edges in this graph. We will denote by $\overrightarrow{p_i p_j}$ the complex number representing the vector between the points in $\mathbb{R}^2$ corresponding to the vertices p_i, p_j. We call the path *irredundant* if

$$\sum_{i \in I} \overrightarrow{p_i p_{i+1}} \neq 0$$

for any $\emptyset \neq I \subset \{0, 1, \dots, k-1\}$.

4.4.2 Proof of Theorem 4.2.2

Let G be the graph with the n points in the plane as vertices and the rational distances with rational angles between pairs of points as edges. Suppose there are $n^{1+f(n)}$ such distances for some positive function f. Then $e(G) \geq n^{1+f(n)}$. We will count the number of irredundant paths P_k in G, for a fixed k that we will choose later. By Lemma 4.4.1 we can assume that $e(G) \geq (1/2)n^{1+f(n)}, v(G) \geq (1/2)n^{f(n)}$ and $\delta(G) \geq (1/2)n^{f(n)}$.

The number of irredundant paths P_k starting at any vertex v is at least

$$N = \prod_{\ell=0}^{k-1} (\delta(G) - 2^\ell + 1),$$

since, if we have constructed a subpath P_ℓ of P_k, then at most $2^\ell - 1$ of the at least $\delta(G)$ continuations are forbidden. Thus the total number of irredundant paths P_k is at least

$$\frac{nN}{2} \geq (n/2) \prod_{\ell=0}^{k-1} ((1/2)n^{f(n)} - 2^\ell + 1) \geq \frac{n^{kf(n)+1}}{2^{2k+1}}$$

if $2^k \leq (1/2)n^{f(n)}$, which is true as long as $k < f(n) \log n / \log 2$. It follows that there are two vertices v and w with at least

$$\frac{N}{n} \geq (1/n) \prod_{\ell=0}^{k-1} ((1/2)n^{f(n)} - 2^\ell + 1) \geq \frac{n^{kf(n)-1}}{4^k}$$

irredundant paths P_k between them. We will call the set of these paths $\mathcal{P}_{vw}$, so that we have $|\mathcal{P}_{vw}| \geq n^{kf(n)-1}/4^k$.

Given $P_k \in \mathcal{P}_{vw}$, $P_k = p_0 p_1 \dots p_k$, consider the k-tuple $(\zeta_1, \dots, \zeta_k)$ where ζ_i is the root of unity in the direction from p_{i-1} to p_i, i.e. $\zeta_i = \overrightarrow{p_{i-1}p_i}/|\overrightarrow{p_{i-1}p_i}|$. Note that $(\zeta_1, \dots, \zeta_k) \in Z_a^k$, because P_k is irredundant. Since there are no three points on a line, this process gives an injective map from $\mathcal{P}_{vw}$ to Z_a^k. Hence $|\mathcal{P}_{vw}| \leq (k \cdot C(k))^k$ by Corollary 4.3.3, and we get

$$\frac{n^{kf(n)-1}}{4^k} \leq (k \cdot C(k))^k \Longrightarrow n^{kf(n)-1} \leq (4k \cdot C(k))^k.$$

This gives

$$(kf(n)-1)\log n \leq k \log(4k \cdot C(k)) \Longrightarrow f(n) \leq \frac{\log(4k) + \log(C(k))}{\log n} + \frac{1}{k}.$$

The term $\log(C(k))$ is the log of the product of the primes less than or equal to $2k$. This is a well known number-theoretic function called the Chebyshev function and denoted by ϑ, specifically $\vartheta(2k) = \log(C(k))$. We use the bound $\vartheta(x) < 4x \log 2 < 3x$ (for a proof see [3]). This gives

$$f(n) < \frac{\log(4k) + 6k}{\log n} + \frac{1}{k} < \frac{7}{\log n}k + \frac{1}{k}.$$

Let k be an integer such that $f(n)\log n/18 < k < f(n)\log n/14$ (possible since otherwise $f(n) = O(1/\log n)$ giving $n^{f(n)} = O(1)$). Then the condition that $k < f(n)\log n/\log 2$ is clearly satisfied, and we get

$$f(n) < \frac{7}{\log n} \cdot \frac{f(n)\log n}{14} + \frac{18}{f(n)\log n} \Longrightarrow f(n) < \frac{6}{\sqrt{\log n}}.$$

This completes the proof.

4.4.3 Proof of Theorem 4.2.1

Theorem 4.2.1 follows from the same proof as 4.2.2: the condition that no three points are on a line was used to show that the map from $\mathcal{P}_{vw}$ to Z_a^k is injective, which now follows from the fact that the edges in the graph all have unit length.

4.4.4 Proof of Theorem 4.2.3

Consider a path $P_k = p_0 p_1 \ldots p_k$. If the distance from p_{i-1} to p_i is less than the distance from p_{i-1} to any vertex on the line connecting p_{i-1} and p_i and not in $P_{i-1} = p_0 p_1 \ldots p_{i-1}$ then P_k is called a *shortest path.*

This proof is almost the same as the proof of Theorem 4.2.2 except that instead of considering all irredundant paths P_k, we only consider shortest irredundant paths. Suppose there are $n^{1+\alpha+f(n)}$ edges in the rational distance graph. Since there are at most n^α points on a line, we get that from any vertex v there are at least

$$N = \prod_{\ell=0}^{k-1}\left(\frac{\delta(G)}{n^\alpha} - 2^\ell + 1\right) \geq \frac{n^{kf(n)}}{4^k}$$

shortest irredundant paths P_k, if $k < f(n)\log n/\log 2$. For any two vertices v, w let $\mathcal{P}_{v,w}$ be the set of shortest irredundant paths P_k between v and w. Then there are two vertices v, w such that the number of shortest irredundant paths between v and w is at least

$$|\mathcal{P}_{vw}| \geq \frac{n^{kf(n)-1}}{4^k}.$$

By Mann's Theorem, since we are looking at shortest irredundant paths, $|\mathcal{P}_{vw}| \leq (k \cdot C(k))^k$. Let k be an integer such that $f(n)\log n/18 < k < f(n)\log n/14$. Then

$$\frac{n^{kf(n)-1}}{4^k} \leq (k \cdot C(k))^k \Longrightarrow f(n) < \frac{6}{\sqrt{\log n}}.$$

That completes the proof.

4.4.5 Proof of Theorem 4.2.4

Assume we have a configuration of n points with at most n^α on a line, $1/2 \leq \alpha \leq 1$, and $n^{1+\alpha+f(n)}$ rational distances with rational angles, for some positive function $f(n)$.

The graph G on these points has $e(G) = n^{1+\alpha+f(n)}$. By Lemma 4.4.1 we can assume that $e(G) \geq n^{1+\alpha+f(n)}/2$, $v(G) \geq n^{\alpha+f(n)}/2$ and $\delta(G) \geq n^{\alpha+f(n)}/2$. We now count irredundant paths P_2 of length 2. Note that an irredundant path on two edges is just a noncollinear path.

For any vertex v, since we have at most n^α points on a line, v is the midpoint of at least

$$N = \delta(G)(\delta(G) - n^\alpha) \geq \frac{n^{2(\alpha+f(n))}}{8}$$

paths P_2 if $f(n) \geq \log 4/\log n$ (if $f(n) < \log 4/\log n$ then $n^{f(n)} < 4$, completing the proof.) Thus there are two vertices v and w with at least $(1/8)n^{2(\alpha+f(n))-1}$ noncollinear paths P_2 between them.

But by Corollary 4.3.3 there is a constant number of directions from each of v and w. Since we are looking at noncollinear paths P_2, the direction from v and the direction from w uniquely determine the midpoint for a path P_2. Thus there are at most $(k \cdot C(k))^k = 144$ noncollinear paths P_2 between v and w, since $k = 2$.

Putting the upper and lower bounds together we get $n^{2(\alpha+f(n))-1} \leq 2^7 3^2$. This gives

$$f(n) \leq \frac{7\log 2 + 2\log 3}{2\log n} + \frac{1}{2} - \alpha \leq \frac{7\log 2 + 2\log 3}{2\log n} < \frac{4}{\log n},$$

since $\alpha \geq 1/2$. But this gives $n^{f(n)} < 4$, completing the proof.

4.5 Lower bounds

In this section we give lower bounds for the theorems given in Section 4.2. We will be a little more informal, since these are known constructions.

The bounds in Theorems 4.2.1 and 4.2.2 are not far from optimal as the following construction of Erdős and Purdy [25] shows.

Suppose we have n points, no three on a line, with the maximum possible number of unit distances with rational angles; we call this number $f(n)$. Consider these points as a set $\{z_1, \ldots, z_n\}$ of complex numbers. For $a \in \mathbb{C}$ with $|a| = 1$ and $a \neq z_i - z_j$ for any $i \neq j$, the set $\{z_1, \ldots, z_n, z_1+a, \ldots, z_n+a\}$ contains at least $2f(n) + n$ unit distances, since there are $f(n)$ among each of the sets $\{z_1, \ldots, z_n\}$ and $\{z_1 + a, \ldots, z_n + a\}$ and $|z_i - (z_i + a)| = 1$ for each i. This new set may have three points on a line, but we show that we can choose a appropriately so that this is not the case.

Consider a pair of points z_i and z_j. For each z_k, the set $\{z_k + a : |a| = 1\}$ intersects the line through z_i and z_j in at most two points. So there are at most two values of a that will give three points on a line. There are $\binom{n}{2}$ pairs of points and n choices for z_k so there are at most $2n\binom{n}{2} = n^2(n-1)$ values of a that make a point $z_k + a$ collinear with two points z_i and z_j. Similarly we have $n^2(n-1)$ values of a that make a point z_k collinear with $z_i + a$ and $z_j + a$. Thus there are only finitely many values of a that give three points on a line, but there are infinitely many choices for a, so we are done.

This shows that $f(2n) \geq 2f(n) + n$ for $n > 2$ and clearly $f(2) = 1$. From this we get that $f(2^k) \geq 2^{k-1}(k-1) = 2^{k-1}\log_2(2^{k-1})$. Taking $2^k \leq n < 2^{k+1}$ we get that $f(n) \geq cn\log n$ for all n. This construction gives a lower bound for Theorems 4.2.1 and 4.2.2.

The bound in Theorem 4.2.3 is not far from optimal. In fact we can get a lower bound of $cn^{1+\alpha}$. Consider $n^{1-\alpha}$ lines parallel to the x-axis, and choose n^α rational points on each line such that no three points on different lines are collinear (this can always be done since there are infinitely many rational points to choose from). There are $cn^{2\alpha}$ rational distances on each horizontal line and $n^{1-\alpha}$ such lines giving at least $cn^{1+\alpha}$ rational distances with rational angles (all the angles are zero).

The bound in Theorem 4.2.4 is tight up to a constant factor as can be seen by considering an $n^{1-\alpha} \times n^\alpha$ square grid. Then there are at least $cn^{2\alpha}$ rational distances on each of the $n^{1-\alpha}$ horizontal lines in the grid containing n^α points. This gives at least $cn^{1+\alpha}$ rational distances with rational angles.

Chapter 5

Simultaneous arithmetic progressions

5.1 Introduction

There are interesting problems in number theory related to arithmetic progressions on elliptic curves (see Section 5.4 for the definition). An example of such an open problem is: what is the maximum number (if it exists) of rational points on an elliptic curve such that their x-coordinates are in arithmetic progression? In [8], Bremner found elliptic curves in Weierstrass form with arithmetic progressions of length 8 on them, and Campbell [10] found elliptic curves of the form $y^2 = f(x)$, with f a quartic, that contain arithmetic progressions of length 12. In [7], Bremner described how these arithmetic progressions are related to 3×3 magic squares with square entries. Bremner, Silverman, and Tzanakis noted in [9] that points in arithmetic progression on elliptic curves are often independent with respect to the group structure, which suggests a relation with the much-researched rank of the curve.

In [30], Garcia-Selfa and Tornero looked instead for "simultaneous" arithmetic progressions on elliptic curves, which are defined as follows.

Definition 5.1.1. A *simultaneous arithmetic progression* (SAP) of length k consists of points $(x_i, y_{\sigma(i)})$ in $\mathbb{R}^2$, where $x_i = a_1 + i \cdot d_1$ and $y_i = a_2 + i \cdot d_2$ for $i = 0, 1, \ldots, k-1$ are arithmetic progressions, and σ is a permutation on the numbers $0, 1, \ldots, k-1$.

Note that the appearance of this permutation is quite natural, since points with both coordinates in arithmetic progression would all lie on a line. Garcia-Selfa and Tornero gave examples of elliptic curves over $\mathbb{Q}$ that contain an SAP of length 6. They also showed that there are only finitely many such curves, and there are none with an SAP of length 7. Extending their methods to SAP's of length 8 did not seem computationally feasible, and they were not able to find an elliptic curve with an SAP of length 8, or prove that none exists. The final open problem they suggested is finding a universal bound for the length of SAP's on elliptic curves over $\mathbb{Q}$.

In Section 4 we prove that 4319 is an upper bound for the length of an SAP on an elliptic curve over $\mathbb{R}$ in Weierstrass form, using a combinatorial approach. This solves the open problem above.

We first approach the more general problem of bounding the k for which an algebraic curve (as defined in section 1.3.2) can contain k points from a $k \times k$ grid. We do this in two different ways: in Section 2 we give a short proof based on a theorem of Jarník, which will give the bound $k \leq cd^9$, where d is the degree of the curve. Then in Section 3 we give a different proof using the well-known crossing inequality from graph theory, in combination with Bézout's Theorem. This will result in the improved bound $k \leq cd^7$. We also extend this result to complex plane algebraic curves. Finally in Section 4 we specialize the second proof to the case of an elliptic curve in Weierstrass form, with some adjustments, resulting in our upper bound of 4319.

5.2 First proof for algebraic curves

Our first theorem shows that the length of an SAP on a curve is bounded by a function of the degree d of the curve. The proof is an application of a theorem of Jarník. In the next section the dependence of the bound on d will be improved with a different approach.

By a $k \times k$ *grid* we will mean the cartesian product of two arithmetic progressions of length k; so an SAP consists of k elements from a $k \times k$ grid with exactly one element on each row and on each column. In the theorems below, we will not specifically deal with SAP's, but more generally with any collection of k points from a $k \times k$ grid.

Theorem 5.2.1. *If f is a curve of degree $d \geq 2$ with no linear factor, and f has at least k points from a $k \times k$ grid, then there is an absolute constant c such that $k \leq cd^9$. If f is also irreducible, the bound improves to $k \leq c'd^6$.*

The dependence on d of the bound in the theorem cannot be removed. Consider any k points from a $k \times k$ grid, and take d such that $k \leq d(d+3)/2$. Then there is a curve of degree d passing through all k of the points (see Section 5.2 in [29]).

Our proof uses the following result of Jarník [35].

Theorem 5.2.2 (Jarník)**.** *If f is a strictly convex differentiable curve of length N, then the number of integer points on f is less than $\alpha N^{2/3}$ for some constant α.*

For algebraic curves, this bound is by no means optimal. In fact, Bombieri and Pila proved in [5] that we can get the bound $c(d,\varepsilon)N^{1/d+\varepsilon}$ for any $\varepsilon > 0$ if the curve is irreducible. This clearly gives a better bound for large degree.

To be able to apply Jarník's bound, we need to break up our algebraic curve into convex or concave pieces (of course Jarník's Theorem works equally well for concave curves). This can be achieved by cutting the curve at all inflection points and singularities (points on f where the first derivatives f_x and f_y both vanish). We will show that the resulting number of pieces is bounded by a function of the degree of f, by reducing to a bounded number of irreducible curves, and then using the following bounds.

Lemma 5.2.3. *Suppose f is an irreducible curve of degree d. Then f has at most $3d(d-2)$ inflection points, and at most $(d-1)(d-2)/2$ singularities.*

For a proof of the bound on inflection points see Proposition 3.33 in [37], and for the bound on singularities see Section 5.4 in [29]. Note that the bound on singularities also follows from the genus formula in Chapter 3, since singularities have $\delta_P \geq 1$ and genus cannot be negative.

It follows that we need at most $3d(d-2) + (d-1)(d-2)/2 < 4d^2$ cuts to break an irreducible f into convex parts.

Proof of Theorem 5.2.1. First assume that f is irreducible and contains k points from a $k \times k$ grid. We scale and translate f so that the gap in the $k \times k$ grid is 1 in both the x- and y-directions, and the points of the grid are integral. Now we can separate f into convex parts using at most $4d^2$ cuts. One of these parts has at least the average number $k/4d^2$ of points from the grid on f. Since the grid has gap 1 and length k we can bound the length of this part of the curve by $2k$. Thus we get, by Theorem 5.2.2, that

$$\frac{k}{4d^2} < \alpha(2k)^{2/3}.$$

This gives $k < Cd^6$ for an irreducible curve f.

For a reducible curve f with k points from a $k \times k$ grid, we have a factorization $f = f_1^{\alpha_1} f_2^{\alpha_2} \dots f_r^{\alpha_r}$ where each f_i is irreducible of degree d_i. Since f has no linear factor, we have $d_i \geq 2$ for all i, as well as $r \leq d/2$. We take the factor f_j with the most points from the grid on it, which is at least $\frac{k}{r} \geq \frac{2k}{d}$. Repeating the argument above with $2k/d$ points from a $k \times k$ grid, we get the inequality $\frac{2k/d}{4d^2} < \alpha(2k)^{2/3}$, which leads to $k < cd^9$. □

5.3 Second proof for algebraic curves

In this section we obtain the following improvement of Theorem 5.2.1, using graph theory.

Theorem 5.3.1. *If f is a curve of degree $d \geq 2$ with no linear factor, such that f has at least k points from a $k \times k$ grid, then there is an absolute constant c such that $k \leq cd^7$. If f is also irreducible, the bound improves to $k \leq c'd^4$.*

The idea behind the proof is to construct a graph G (actually, a multigraph) out of translates of the curve in a grid, with the grid points as vertices, and with edges between points which occur consecutively on a translate. To get the stated upper bound for k, we obtain a lower bound on the number of intersections of G from the Crossing Inequality (Theorem 5.3.2), and compare this with an upper bound that we get from Bézout's Theorem (5.3.4).

Construction of the graph G. For convenience suppose that the grid is $[1, k] \times [1, k]$, and extend it to $[1, 2k] \times [1, 2k]$; this will be the vertex set of our graph. Consider the k^2 translates of f obtained by shifting i in the x-direction and j in the y direction, for all pairs $i, j \in [1, k]$. We will basically draw an edge along the curve between two grid points if they occur as consecutive points on some translate of f, but there are several things we have to watch out for.

First note that some parts of the curve do not occur among the edges, namely the parts that go off to infinity, and components of f that have no grid point or a single grid point on it. Especially this last case will have to be accounted for in the proof.

Second, we may have more than one edge connecting two vertices. We show that the maximum edge multiplicity for a pair of vertices is d^2. If we have more than one edge connecting two vertices then these two vertices appear as consecutive points on different translates. These points are given as (l, m) and (l', m') for some $l, l', m, m' \in \mathbb{Z}$. If these points appear on r translates then the vector $(l - l', m - m')$ occurs as a difference vector between r pairs of points on the original elliptic curve. But this is equivalent to having r points on the original curve intersecting r points on its translate by $(l - l', m - m')$. Hence $r \leq d^2$ by Bézout's Theorem.

Third, since we are considering translates of a curve in a grid, a vertex may occur on a number of translates. Suppose v_1 and v_2 are consecutive points on a translate. We may have a point v_3 on another translate which is actually between v_1 and v_2 on the first translate. In this case the edge from v_1 to v_2 passes through the vertex v_3. This is not allowed in a graph so we have to alter our graph slightly. In this case we remove the edge in consideration from v_1 to v_2 and add an edge from v_1 to v_3. Performing this change where necessary we end up with a graph with the same number of vertices and edges but without the problem of an edge passing through a vertex to which it is not adjacent. We call this graph G.

We now introduce the results from graph theory that we will use to get a lower bound on the crossing number of our graph.

Given a simple graph G, the *crossing number* $\mathrm{cr}(G)$ is the minimum number of pairs of crossing edges in a planar drawing of G.

Theorem 5.3.2 (Crossing Inequality)**.** *Suppose G is a simple graph with n vertices and e edges. If $e > 7.5n$ then*

$$\mathrm{cr}(G) \geq \frac{e^3}{33.75n^2}.$$

The crossing inequality was first proved independently by Ajtai, Chvátal, Newborn and Szemerédi [2] and by Leighton [40]. The version with the best bound to date, presented above, was given by Pach and Tóth [44].

Pach and Tóth also gave the following crossing inequality for multigraphs, which is the result we will use here.

Theorem 5.3.3 (Crossing inequality for multigraphs)**.** *Suppose G is a multigraph with n vertices and e edges (counting multiplicity). Suppose there are at most m edges between any pair of vertices in G. If $e > 7.5mn$ then*

$$\mathrm{cr}(G) \geq \frac{e^3}{33.75mn^2}.$$

To get an upper bound on the number of intersections in our graph, we use Bézout's Theorem. For details see [29].

Theorem 5.3.4 (Bézout)**.** *Suppose F and G are curves of degree m and n. If F and G do not have a common factor, then they intersect in at most mn points.*

The main consequence that we will use is that if $f(x, y) = 0$ is an irreducible curve of degree d, then f and a translate of f intersect in at most d^2 points. Finally, we will need the following result. For details, see [28], p. 218.

Lemma 5.3.5 (Harnack)**.** *The number of connected components in $\mathbb{R}^2$ of an irreducible curve is at most $\frac{(d-1)(d-2)}{2}+1$.*

Proof of Theorem 5.3.1. We first assume that f is irreducible. Suppose that f contains k points from a $k \times k$ grid.

Typically, a component of f with m vertices on it will give m edges in our graph, except that we get $m-1$ edges if the component does not form a closed loop (i.e. contains a point at infinity), and 0 edges if the component contains only one vertex. Fortunately, by Harnack's result there are at most $(d-1)(d-2)/2+1$ components, and by Bézout the line at infinity and our curve can have at most d intersections, so the number of such "bad" components is $\leq d^2/2$. If we assume $k \geq d^2$, then at least $k/2$ of the vertices are on "good" components, hence give us $k/2$ edges. All translates together then give $k^3/2$ edges in our graph.

The crossing number of the graph is the number of intersections between translates plus the number of self-intersections of translates. By Bézout's Theorem, for any pair of translates there are at most d^2 intersections. A self-intersection is a singularity, hence the number of these is bounded above by $(d-1)(d-2)/2$ by Lemma 5.2.3. There are k^2 translates and $\binom{k^2}{2}$ pairs of translates so we get

$$cr(G) \leq d^2\binom{k^2}{2} + k^2\frac{(d-1)(d-2)}{2} = \frac{1}{2}k^2(d^2k^2-3d+2) \leq \frac{1}{2}d^2k^4.$$

By the crossing inequality for multigraphs we get the lower bound:

$$\frac{(k^3/2)^3}{33.75d^2(4k^2)^2} \leq \frac{e^3}{33.75d^2v^2} \leq cr(G).$$

Combining these we get $k \leq C_1d^4$, with $C_1 = 2^6 \cdot 33.75$, for an irreducible curve.

For a reducible curve f with k points from a $k \times k$ grid we proceed as in Section 5.2. We can find an irreducible factor of f with at least $2k/d$ grid points on it. That gives us k^3/d edges in our graph, so that the inequality becomes

$$\frac{(k^3/d)^3}{33.75d^2(4k^2)^2} \leq \frac{1}{2}d^2k^4,$$

which gives us $k \leq C_2d^7$, with $C_2 = 33.75 \cdot 2^3$. □

Theorem 5.3.1 can be extended to any complex plane algebraic curve, i.e. where $f \in \mathbb{C}[x, y]$ and we consider the zero set

$$C_f(\mathbb{C}) = \{(x, y) \in \mathbb{C}^2 : f(x, y) = 0\}.$$

We use the result for the reals to prove the result for this complex case. By a $k \times k$ grid in $\mathbb{C}^2$ we mean a cartesian product of two arithmetic progressions in $\mathbb{C}$; by an arithmetic progression in $\mathbb{C}$ we mean points $\alpha + i\beta$ with $\alpha, \beta \in \mathbb{C}$ and $i = 0, 1, \ldots, k-1$.

Theorem 5.3.6. *If f is a* complex *plane algebraic curve of degree $d \geq 2$ with no linear factor such that $C_f(\mathbb{C})$ contains at least k points from a $k \times k$ grid, then there is an absolute constant c such that $k \leq cd^7$.*

Proof. Suppose $f(w, z)$ is our complex curve with k points on a $k \times k$ grid, given by $\alpha + j\beta$ in one direction and $\gamma + j\delta$ in the other direction, where $\alpha, \beta, \gamma, \delta \in \mathbb{C}$ and $j = 0, 1, \ldots, k-1$. Now consider the polynomial $g(x, y) = f(\alpha + x\beta, \gamma + y\delta)$. This is a polynomial with complex coefficients in two real variables, and the curve $g(x, y) = 0$ has k points on the $k \times k$ grid consisting of the points (i, j) with $i, j = 0, 1, \ldots, k-1$. The real and imaginary parts of g are real algebraic curves of degree $\leq d$, each having k points on the grid. Thus Theorem 5.3.1 gives the stated bound. $\square$

5.4 SAPs on elliptic curves

In this section we use the method from Section 5.3, with some modifications, to give a universal bound on the size of an SAP on a real elliptic curve, answering the question of Garcia-Selfa and Tornero. By *elliptic curve*, we will mean a nonsingular irreducible cubic curve, with Weierstrass equation $y^2 + axy + by = x^3 + cx^2 + dx + e$; however, we will not make any use of the arithmetic theory of elliptic curves.

The upper bound that we would get using Jarník's Theorem as in Section 5.2 is roughly $4 \cdot 10^5$. With the approach from Section 5.3 we would get an upper bound of roughly $2 \cdot 10^5$, using $k \leq C_1 \cdot d^4$ from the proof. For the specific case of elliptic curves, we can make several improvements. First of all, we know that the curve is irreducible, and that there are no singularities. Second, we know exactly what the components can look like. Finally, the following lemma shows that we can do better than the intersection bound $d^2 = 9$ Bézout's Theorem.

Lemma 5.4.1. *An elliptic curve and a translate of that curve can intersect in at most* 4 *points (excluding points at infinity.)*

Proof. Suppose the curve is given by $y^2 + axy + by = x^3 + cx^2 + dx + e$. A translate is given by $(y+v)^2 + a(x+u)(y+v) + b(y+v) = (x+u)^3 + c(x+u)^2 + d(x+u) + e$ where at least one of u, v does not equal 0. If (x, y) is an intersection point of these curves then subtracting one equation from the other we get

$$2vy + v^2 + avx + auy + auv + bv = 3ux^2 + 3u^2x + u^3 + 2cux + cu^2 + du.$$

If $2v + au = 0$ then all terms involving y disappear. In this case we have a quadratic in x which can have at most 2 real roots. Putting these values into the original equation we get at most 4 intersection points. If $2v + au \neq 0$ then we can solve for y to get

$$y = \frac{3ux^2 + (3u^2 + 2cu - av)x + (u^3 + cu^2 + du - auv - bv - v^2)}{2v + au}.$$

Substituting this into the original equation we get $f(x) = 0$ where f is a quartic polynomial in x. This polynomial has at most 4 roots. Thus we cannot have more than 4 intersection points of our elliptic curve and its translate. □

The main result is:

Theorem 5.4.2. *Suppose we have an elliptic curve given by* $y^2+axy+by = x^3 + cx^2 + dx + e$*, containing an SAP of length* k*. Then* $k \leq 4319$.

Modified construction of G. The graph we will use is essentially the same, but we will make a small adjustment based on a closer analysis of the possible components. A cubic curve of the type above will either consist of a single component containing a point at infinity, or it will consist of two components, one a closed loop and the other containing a point at infinity. We will treat these two cases simultaneously. We will add the point at infinity to our graph, along with the two edges that approach it; this will improve our bounds below.

Suppose our curve has t points from the SAP on the loop component (if it exists; otherwise $t = 0$) and $k - t$ points on the infinite component. If there is more than one point on the loop, so $t > 1$, then it gives a t-cycle with t edges in the graph. If $t = 1$ or $t = 0$, then we get no edges from the loop component.

Consider the $k-t$ points on the component containing a point at infinity. They give us $k-t$ vertices and $k-t-1$ edges. We add the curve's point at infinity to the graph, then we connect the rightmost point on the top part of the curve to the point at infinity, and we do the same for the rightmost point on the bottom part of the curve. This component now contains $k-t+1$ vertices and $k-t+1$ edges. Together with the other component we end up with $k+1$ vertices in one translate, and we have $k+1$ edges if $t \neq 1$ and k edges if $t = 1$.

By the same argument as in the construction of G in the previous section and by Lemma 5.4.1, edges have multiplicity at most 4. The new edges (to the point at infinity) will not increase the multiplicity. The only way we can have more than one edge going from a point in the grid to the point at infinity is if that point is the rightmost point on the top half of one translate and the rightmost point on the bottom half of another translate. Thus these edges have multiplicity at most 2.

Proof of Theorem 5.4.2. Our graph now has $4k^2+1$ vertices. Let t be defined as above. If $t \neq 1$, then the number of edges, counting multiplicity, in G is $k^2(k+1)$, while if $t = 1$ then the number of edges is k^3. We need only consider the case with fewer edges, so we assume we have k^3 edges.

The crossing inequality now gives

$$\frac{(k^3)^3}{4(33.75)(4k^2+1)^2} \leq \operatorname{cr}(G).$$

By Lemma 5.4.1, any pair of translates intersects in at most 4 points in the grid, and there are $\binom{k^2}{2}$ such pairs. Thus the crossing number is bounded above by

$$\operatorname{cr}(G) \leq 4\binom{k^2}{2}.$$

Putting these two inequalities together we get

$$\frac{(k^3)^3}{4(33.75)(4k^2+1)^2} \leq 4\binom{k^2}{2}.$$

Solving for k in this inequality and noting that k is a positive integer, we get $k \leq 4319$.

Note that in Theorem 5.3.3 we require $e > 4(7.5)n = 30n$. This certainly holds since otherwise $e \leq 30n$, i.e. $k^2(k+1) \leq 30(4k^2+1)$. A quick calculation shows that this gives $k \leq 120$. □

Bibliography

[1] S.S. Abhyankar, *Algebraic geometry for scientists and engineers*, Mathematical Surveys and Monographs 38, AMS, 2000.

[2] M. Ajtai, V. Chvátal, M.M. Newborn, and E. Szemerédi, *Crossing-free subgraphs*, Theory and practice of combinatorics 60 (1982), 9–12.

[3] T.M. Apostol, *Introduction to analytic number theory*, Springer, 1976.

[4] N.H. Anning and P. Erdős, *Integral distances*, Bull. Amer. Math. Soc. 51 (1945), 598–600.

[5] E. Bombieri and J. Pila, *The number of integral points on arcs and ovals.*, Duke Mathematical Journal 59 (1989), 337–357.

[6] P. Brass, W. Moser, and J. Pach, *Research problems in discrete geometry*, Springer, 2006.

[7] A. Bremner, *On squares of squares.*, Acta Arithmetica 88 (1999), 289–297.

[8] A. Bremner, *On arithmetic progressions on elliptic curves.*, Experimental Mathematics 8 (1999), 409–413.

[9] A. Bremner, J. Silverman, and N. Tzanakis, *Integral points in arithmetic progression on $y^2 = x(x^2 - n^2)$*, J. Number Theory 80 (2000), 187–208.

[10] G. Campbell, *Points on $y = x^2$ at rational distance*, Math. Comp. 73 (2004), 2093–2108.

[11] L. Caporaso, J. Harris, and B. Mazur, *Uniformity of rational points*, J. Amer. Math. Soc. 10 (1997), 1–35.

[12] A. Choudhry, *Points at rational distances on a parabola*, Rocky Mountain J. Math. 36 (2006), 413–424.

[13] K. Conrad, *Arithmetic progressions of three squares*, Online expository paper, available at
`www.math.uconn.edu/~kconrad/blurbs/ugradnumthy/3squarearithprog.pdf`

[14] K. Conrad, *Arithmetic progressions of four squares*, Online expository paper, available at
`www.math.uconn.edu/~kconrad/blurbs/ugradnumthy/4squarearithprog.pdf`

[15] L.E. Dickson, *History of the theory of numbers. Vol. II: Diophantine analysis*, Chelsea Publishing Co., New York, 1966.

[16] H. Darmon and L. Merel, *Winding quotients and some variants of Fermat's last theorem,* J. reine angew. Math. 490 (1997), 81–100.

[17] P. Erdős, *Integral distances.*, Bull. Amer. Math. Soc. 51 (1945), 996.

[18] P. Erdős, *On sets of distances of n points*, Am. Math. Mon. 53 (1946), 248–250.

[19] P. Erdős, *Verchu niakoy geometritchesky zadatchy,*(On some geometric problems, in Bulgarian), Fiz.-Mat. Spis. Bŭlgar. Akad. Nauk. 5 (1962), 205–212.

[20] P. Erdős, *On some problems of elementary and combinatorial geometry*, Annali di Matematica pura ed applicata 4 (1975), 99–108.

[21] P. Erdős, *Néhány elemi geometriai problémáról* (On some problems in elementary geometry, in Hungarian), Köz. Mat. Lapok 61 (1980), 49–54.

[22] P. Erdős, *Combinatorial problems in geometry*, Math. Chronicle 12 (1983), 35–54.

[23] P. Erdős, *Some combinatorial and metric problems in geometry*, Colloquia Mathematica Societatis János Bolyai 48 (1985), 167–177.

[24] P. Erdős, *Ulam, the man and the mathematician*, J. Graph Theory 9 (1985) no. 4, 445–449.

[25] P. Erdős and G.B. Purdy, *Extremal problems in combinatorial geometry*, Handbook of Combinatorics, Elsevier Science (1995), 809–874.

[26] J.-H. Evertse, H.P. Schlickewei, and W.M. Schmidt, *Linear equations in variables which lie in a multiplicative group*, Ann. Math. 155 (2002), 807–836.

[27] G. Faltings, *Endlichkeitssätze für abelsche Varietäten über Zahlkörpern*, Invent. Math. 73 (3) (1983), 349–366.

[28] G. Fischer, *Plane Algebraic Curves*, Student Mathematical Library 15, AMS, 2001.

[29] W. Fulton, *Algebraic curves: an introduction to algebraic geometry*, 1969.

[30] I. Garcia-Selfa and J.M. Tornero, *On simultaneous arithmetic progressions on elliptic curves*, Experimental Mathematics 15 (2006), 471–478.

[31] R. Guy, *Unsolved problems in number theory*, Problem Books in Mathematics Subseries: Unsolved Problems in Intuitive Mathematics 1, Springer, 3rd ed., 2004.

[32] R. Hartshorne, *Algebraic geometry*, Graduate Texts in Mathematics 52, Springer, 1977.

[33] H. Harborth, A. Kemnitz, and M. Möller, *An upper bound for the minimum diameter of integral point sets*, Discrete & Comput. Geom. 9 (1993), 427–432.

[34] G.B. Huff, *Diophantine problems in geometry and elliptic ternary forms*, Duke Math. J. 15 (1948), 443–453.

[35] V. Jarník, *Über die Gitterpunkte auf konvexen Kurven*, Mathematische Zeitschrift 24 (1926), 500–518.

[36] A. Kemnitz, *Punktmengen mit ganzzahligen Abständen*, Habilitationsschrift, TU Braunschweig, 1988.

[37] F.C. Kirwan, *Complex algebraic curves*, Cambridge University Press, 1992.

[38] T. Kreisel and S. Kurz, *There are integral heptagons, no three points on a line, no four on a circle*, Discrete & Computational Geometry 39 (2008), 786–790.

[39] S. Lang, *Algebra*, Addison-Wesley, 3rd ed., 1994.

[40] F.T. Leighton, *New lower bound techniques for VLSI*, Math. Systems Theory 17 (1984), 47–70.

[41] H.B. Mann, *On linear relations between roots of unity*, Mathematika 12 (1965), 107–117.

[42] J. Matoušek, *The number of unit distances is almost linear for most norms*, Adv. Math. 226 (2011), 2618–2628.

[43] W. D. Peeples Jr., *Elliptic curves and rational distance sets*, Proc. Am. Math. Soc. 5 (1954), 29–33.

[44] J. Pach and G. Tóth, *Graphs drawn with few crossings per edge*, Combinatorica 17 (1997), 427–439.

[45] W.M. Schmidt, *Diophantine approximation*, LNM 785, Springer Verlag, 1980.

[46] R. Schwartz, J. Solymosi, and F. de Zeeuw, *Simultaneous arithmetic progressions on algebraic curves*, International Journal of Number Theory 7 (2011), 921–931.

[47] R. Schwartz, J. Solymosi, and F. de Zeeuw, *Rational distances with rational angles*, accepted for publication in Mathematika; `arXiv:1008.3671v2`, [math.CO], 2011.

[48] J. Silverman, *The arithmetic of elliptic curves*, Graduate Texts in Mathematics 106, Springer, 1994.

[49] J. Solymosi, *Note on integral distances*, Discrete & Comput. Geom. 30 (2003), 337–342.

[50] J. Solymosi and F. de Zeeuw, *On a Question of Erdős and Ulam*, Discrete and Computational Geometry 43 (2010), 393–401.

[51] J. Spencer, E. Szemerédi, and W. Trotter, *Unit distances in the Euclidean plane*, Graph Theory and Combinatorics: Proceedings of the Cambridge Combinatorial Conference, in Honour of Paul Erdős, Academic Press, 1984, 293–303.

[52] L. Székely, *Crossing numbers and hard Erdős problems in discrete geometry*, Combinatorics, Probability and Computing 6 (1997), 353–358.

[53] S.M. Ulam, *A collection of mathematical problems*, Interscience Tracts in Pure and Applied Mathematics 8, Interscience Publishers, 1960.