CHAPTER 1

Overview

In mathematics and its applications, we are often faced with a system of multivariate polynomial equations whose solutions we need to study or to find. Systems that arise naturally typically possess some geometric or combinatorial structure that may be exploited to better understand their solutions. Such structured systems are studied in enumerative algebraic geometry, which has given us the deep and powerful tools of intersection theory [**54**] to count and analyze their *complex* solutions. A companion to this theoretical work are algorithms, both symbolic (based on Gröbner bases [**154**] or resultants) and numerical (many based on numerical homotopy continuation [**137**]) for solving and analyzing systems of polynomial equations. An elegant and elementary introduction into algebraic geometry, algorithms, and its applications is given in the two-volume series [**31, 30**].

Despite these successes, this line of research largely sidesteps the often primary goal of formulating problems as solutions to systems of equations—namely to determine or study their real solutions. This deficiency is particularly acute in applications, from control [**25**], to kinematics [**22**], statistics [**114**], and computational biology [**110**], for it is typically the real solutions that are needed in applications. One reason that traditional algebraic geometry ignores the real solutions is that there are few elegant theorems or general results available to study real solutions. Nevertheless, the demonstrated importance of understanding the real solutions to systems of equations demands our attention.

In the 19th century and earlier, many elegant and powerful methods were developed to study the real roots of univariate polynomials (Sturm sequences, Budan-Fourier Theorem, Routh-Hurwitz criterion), which are now standard tools in some applications of mathematics. These and other results lead to a rich algorithmic theory of real algebraic geometry, which is developed in [**4**]. In contrast, it has only been in the past few decades that serious attention has been paid toward understanding the real solutions to systems of multivariate polynomial equations.

This recent work has concentrated on systems possessing some, often geometric, structure. The reason for this is two-fold: Not only do systems from nature typically possess some special structure that should be exploited in their study, but it is highly unlikely that any results of substance hold for general or unstructured systems. From this work, a story has emerged of bounds (both upper and lower) on the number of real solutions to certain classes of systems, as well as the discovery and study of systems that have only real solutions. This overview chapter will sketch this emerging landscape and the subsequent chapters will treat these ongoing developments in more detail.

We will use the notations $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$, to denote the natural numbers, integers, rational numbers, real numbers, and complex numbers. We write $\mathbb{R}_>$ for the positive real numbers, and $\mathbb{R}^*$ (or $\mathbb{T}_\mathbb{R}$) and $\mathbb{C}^*$ (or $\mathbb{T}$) for the nonzero real and complex numbers, respectively. For a positive integer n, write $[n]$ for the set $\{1,\ldots,n\}$, and let $\mathbb{Z}^n$ be the free abelian group of rank n (a lattice), and $\mathbb{Q}^n$, $\mathbb{R}^n$, and $\mathbb{C}^n$, vector spaces of dimension n over the indicated fields. Likewise $\mathbb{P}^n$ and $\mathbb{RP}^n$ are complex and projective spaces of dimension n, and $\mathbb{R}^n_>$, $(\mathbb{R}^*)^n$, and $(\mathbb{C}^*)^n = \mathbb{T}^n$ for the groups of n-tuples of positive, nonzero real, and nonzero complex numbers, respectively. These groups, vector spaces, and projective spaces, all have distinguished ordered bases. We will use $\mathbb{Z}^{\mathcal{A}}, \mathbb{R}^{\mathcal{A}}, \mathbb{T}^{\mathcal{A}}, \mathbb{P}^{\mathcal{A}}, \ldots$ to denote the groups and spaces with distinguished bases indexed by the elements of a set $\mathcal{A}$.

1.1. Introduction

Our goal is to say something meaningful about the real solutions to a system of multivariate polynomial equations. For example, consider a system

$$f_1(x_1,\ldots,x_n) \;=\; f_2(x_1,\ldots,x_n) \;=\; \cdots \;=\; f_N(x_1,\ldots,x_n) \;=\; 0\,, \tag{1.1}$$

of N real polynomials in n variables. Let r be its number of real solutions and let d be its number of complex solutions. We always assume that our systems are *generic* in the sense that all of their solutions are *nondegenerate*. Specifically, the differentials df_i of the polynomials at each solution span $\mathbb{C}^n$, so that each solution has algebraic multiplicity 1. Our systems will come in families whose generic member is nondegenerate and has d complex solutions. Since every real number is complex, and since nonreal solutions come in complex conjugate pairs, we have the following trivial inequality,

$$d \;\geq\; r \;\geq\; d \mod 2 \;\in\; \{0,1\}\,. \tag{1.2}$$

We can say nothing more unless the equations have some structure, and a particularly fruitful class of structures are those which come from geometry. The main point of this book is that we can identify structures in equations that will allow us to do better than this trivial inequality (1.2).

Our discussion will have three themes:

(I) Sometimes, there is a smaller bound on r than d.
(II) For many problems from enumerative geometry, the upper bound is sharp.
(III) The lower bound on r may be significantly larger than $d \mod 2$.

A major theme will be the Shapiro Conjecture (Mukhin, Tarasov, and Varchenko Theorem [**104**]) and its generalizations, which is a situation where the upper bound of d is also the lower bound—all solutions to our system are real. This also occurs in Example 9.7.

We will not describe how to actually find the solutions to a system (1.1) and there will be little discussion of algorithms and no complexity analysis. The book of Basu, Pollack, and Roy [**4**] is an excellent place to learn about algorithms for computing real algebraic varieties and finding real solutions. We remark that some techniques employed to study real solutions underlie numerical algorithms to compute the solutions [**137**]. Also, ideas from toric geometry [**52, 61**], Gröbner bases [**154**], combinatorial commutative algebra [**100**], and Schubert Calculus [**53**] permeate this book. Other background material may be found in [**31, 30**].

1.2. Polyhedral bounds

When $N = n$, the most fundamental bound on the number of complex solutions is due to Bézout: d is at most the product of the degrees of the polynomials f_i. When the polynomials have a sparse, or polyhedral structure, the smaller BKK bound applies.

Integer vectors $a = (a_1, \ldots, a_n) \in \mathbb{Z}^n$ are exponents for (Laurent) monomials

$$\mathbb{Z}^n \ni a \;\leftrightarrow\; x^a \;:=\; x_1^{a_1} x_2^{a_2} \cdots x_n^{a_n} \;\in\; \mathbb{C}[x_1, \ldots, x_n, x_1^{-1}, \ldots, x_n^{-1}] .$$

We will often identify a monomial with its exponent vector and thus will just call elements of $\mathbb{Z}^n$ *monomials*. Let $\mathcal{A} \subset \mathbb{Z}^n$ be a finite set of monomials. A linear combination

$$\sum_{a \in \mathcal{A}} c_a x^a \qquad c_a \in \mathbb{R}$$

of monomials from $\mathcal{A}$ is a *sparse polynomial* with *support $\mathcal{A}$*. Sparse polynomials naturally define functions on the complex torus $\mathbb{T}^n := (\mathbb{C}^*)^n$. A system (1.1) of $N = n$ polynomials in n variables, where each polynomial has support $\mathcal{A}$, will be called a *system* (of polynomials) with *support $\mathcal{A}$*. These are often called *unmixed systems* in contrast to *mixed systems* where each polynomial may have different support. While sparse systems occur naturally—multilinear or multihomogeneous polynomials are an example—they also occur in problem formulations for the simple reason that we humans seek simple formulations of problems, and this often means polynomials with few terms.

A fundamental result about unmixed systems is the Kushnirenko bound on their number of complex solutions. The *Newton polytope* of a polynomial f with support $\mathcal{A}$ is the convex hull $\Delta_{\mathcal{A}}$ of the set $\mathcal{A}$ of monomials of f. Write $\text{volume}(\Delta)$ for the Euclidean volume of a polytope Δ.

THEOREM 1.1 (Kushnirenko [**10**]). *A system of n polynomials in n variables with support $\mathcal{A}$ has at most $n!\,\text{volume}(\Delta_{\mathcal{A}})$ isolated solutions in $(\mathbb{C}^*)^n$, and exactly this number when the polynomials are generic polynomials with support $\mathcal{A}$.*

Bernstein generalized this to mixed systems. The Minkowski sum $P+Q$ of two polytopes in $\mathbb{R}^n$ is their pointwise sum as sets of vectors in $\mathbb{R}^n$. Let $P_1, \ldots, P_n \subset \mathbb{R}^n$ be polytopes. The volume

$$\text{volume}(t_1 P_1 \;+\; t_2 P_2 \;+\; \cdots \;+\; t_n P_n)$$

is a homogeneous polynomial of degree n in the variables $t_1, \ldots, t_n$ [**63**, Exercise 15.2.6]. The *mixed volume* $\text{MV}(P_1, \ldots, P_n)$ of $P_1, \ldots, P_n$ is the coefficient of the monomial $t_1 \cdots t_n$ in this polynomial.

THEOREM 1.2 (Bernstein [**11**]). *A system of n polynomials in n variables where the polynomials have supports $\mathcal{A}_1, \ldots, \mathcal{A}_n$ has at most $\text{MV}(\Delta_{\mathcal{A}_1}, \ldots, \Delta_{\mathcal{A}_n})$ isolated solutions in $(\mathbb{C}^*)^n$, and exactly this number when the polynomials are generic for their given support.*

Since $\text{MV}(P_1, \ldots, P_n) = n!\,\text{volume}(P)$ when $P_1 = \cdots = P_n = P$, this generalizes Kushnirenko's Theorem. We will prove these theorems in Chapters 3 and 4.

The bound of Theorem 1.1 and its generalization Theorem 1.2 is often called the *BKK bound* for Bernstein, Khovanskii, and Kushnirenko [**10**].

1.3. Upper bounds

While the number of complex roots of a univariate polynomial is typically equal to its degree, the number of real roots depends upon the length of the expression for the polynomial. Indeed, by Descartes's rule of signs [**34**] (see Section 2.1), a univariate polynomial with $m+1$ terms has at most m positive roots, and thus at most $2m$ nonzero real roots. For example, the polynomial $x^d - a$ with $a \neq 0$ has 0, 1, or 2 real roots, but always has d complex roots. Khovanskii generalized this type of bound to multivariate polynomials with his fundamental *fewnomial bound*.

THEOREM 1.3 (Khovanskii [**83**]). *A system of n polynomials in n variables having a total of $1+l+n$ distinct monomials has at most*

$$2^{\binom{l+n}{2}}(n+1)^{l+n}$$

nondegenerate positive solutions.

There are two reasons for this restriction to positive solutions. Most fundamentally is that Khovanskii's proof requires this restriction. This restriction also excludes the following type of trivial zeroes: Under the substitution $x_i \mapsto x_i^2$, each positive solution becomes 2^n real solutions, one in each of the 2^n orthants. More subtle substitutions lead to similar trivial zeroes which differ from the positive solutions only by some sign patterns.

This is the first of many results verifying the principle of Bernstein and Kushnirenko that the topological complexity of a set defined by real polynomials should depend on the number of terms in the polynomials and not on the degrees of the polynomials. Khovanskii's work was also a motivation for the notion of o-minimal structures [**160, 113**]. The main point of Khovanskii's theorem is the existence of such a bound and not the actual bound itself.

Nevertheless, it raises interesting questions about such bounds. For each $l, n \geq 1$, we define the *Khovanskii number $X(l,n)$* to be the maximum number of nondegenerate positive solutions to a system of n polynomials in n variables with $1+l+n$ monomials. Khovanskii's Theorem gives a bound on $X(l,n)$, but that bound is enormous. For example, when $l = n = 2$, the bound is 5184. Because of this, Khovanskii's bound was expected to be far from sharp. Despite this expectation, the first nontrivial improvement was only given in 2003.

THEOREM 1.4 (Li, Rojas, and Wang [**94**]). *Two trinomials in two variables have at most five nondegenerate positive solutions.*

This bound is sharp. Haas [**64**] had shown that the system of two trinomials in x and y

$$10x^{106} + 11y^{53} - 11y \;=\; 10y^{106} + 11x^{53} - 11x \;=\; 0\,, \tag{1.3}$$

has five positive solutions.

Since we may multiply one of the trinomials in (1.3) by an arbitrary monomial without changing the solutions, we can assume that the two trinomials (1.3) share a common monomial, and so there are at most $3+3-1 = 5 = 2+2+1$ monomials between the two trinomials, and so two trinomials give a fewnomial system with $l = n = 2$. While five is less than 5184, Theorem 1.4 does not quite show that $X(2,2) = 5$ as two trinomials do not constitute a general fewnomial system with $l = n = 2$. Nevertheless, Theorem 1.4 gave strong evidence that Khovanskii's fewnomial bound may be improved. Such an improved bound was given in [**17**].

THEOREM 1.5. $X(l,n) < \frac{e^2+3}{4} 2^{\binom{l}{2}} n^l$.

For small values of l, it is not hard to improve this. For example, when $l = 0$, the support $\mathcal{A}$ of the system is a simplex, and there will be at most one positive real solution, so $X(0,n) = 1$. Theorem 1.5 was inspired by the sharp bound of Theorem 1.7 when $l = 1$ [**15**]. A set $\mathcal{A}$ of exponents is *primitive* if $\mathcal{A}$ affinely spans the full integer lattice $\mathbb{Z}^n$. That is, the differences of vectors in $\mathcal{A}$ generate $\mathbb{Z}^n$.

THEOREM 1.6. *If $l = 1$ and the set $\mathcal{A}$ of exponents is primitive, then there can be at most $2n{+}1$ nondegenerate nonzero real solutions, and this is sharp in that for any n there exist systems with $n{+}2$ monomials and $2n{+}1$ nondegenerate real solutions whose exponent vectors affinely span $\mathbb{Z}^n$.*

Observe that this bound is for all nonzero real solutions, not just positive solutions. We will discuss this in Section 5.3. Further analysis by Bihan gives the sharp bound on $X(1,n)$.

THEOREM 1.7 (Bihan [**15**]). $X(1,n) = n + 1$.

These fewnomial bounds are discussed and proven in Chapters 5 and 6.

In contrast to these results establishing absolute upper bounds on the number of real solutions which improve the trivial bound of the number d of complex roots, there are a surprising number of problems that come from geometry for which all solutions can be real. For example, Sturmfels [**153**] proved the following. (Regular triangulations are defined in Section 4.2, and we give his proof in Section 4.4.)

THEOREM 1.8. *Suppose that a lattice polytope $\Delta \subset \mathbb{Z}^n$ admits a regular triangulation with each simplex having minimal volume $\frac{1}{n!}$. Then there is a system of sparse polynomials with support $\Delta \cap \mathbb{Z}^n$ having all solutions real.*

For many problems from enumerative geometry, it is similarly possible that all solutions can be real. This will be discussed in Chapter 9.

1.4. The Wronski map and the Shapiro Conjecture

The *Wronskian* of univariate polynomials $f_1(t), \ldots, f_m(t)$ is the determinant

$$\mathrm{Wr}(f_1, f_2, \ldots, f_m) \;:=\; \det\left(\left(\tfrac{d}{dt}\right)^{i-1} f_j(t)\right)_{i,j=1,\ldots,m} .$$

When the polynomials f_i have degree $m{+}p{-}1$ and are linearly independent, the Wronskian has degree at most mp. For example, if $m = 2$, then $\mathrm{Wr}(f,g) = f'g - fg'$, which has degree $2p$ as the coefficients of t^{2p+1} in this expression cancel. Up to a scalar, the Wronskian depends only upon the linear span of the polynomials $f_1, \ldots, f_m$. Removing these ambiguities gives the *Wronski map*,

$$\mathrm{Wr} \,:\, \mathrm{Gr}(m, \mathbb{C}_{m+p-1}[t]) \longrightarrow \mathbb{P}(\mathbb{C}_{mp}[t]) \simeq \mathbb{P}^{mp}, \tag{1.4}$$

where $\mathrm{Gr}(m, \mathbb{C}_{m+p-1}[t])$ is the *Grassmannian* of m-dimensional subspaces of the linear space $\mathbb{C}_{m+p-1}[t]$ of complex polynomials of degree $m{+}p{-}1$ in the variable t, and $\mathbb{P}(\mathbb{C}_{mp}[t])$ is the projective space of complex polynomials of degree at most mp, which has dimension mp, equal to the dimension of the Grassmannian.

Work of Schubert in 1886 [**130**], combined with a result of Eisenbud and Harris in 1983 [**40**] shows that the Wronski map is surjective and the general polynomial $\Phi \in \mathbb{P}^{mp}$ has

$$\#_{m,p} \;:=\; \frac{1!2!\cdots(m{-}1)!\cdot(mp)!}{m!(m{+}1)!\cdots(m{+}p{-}1)!} \tag{1.5}$$

preimages under the Wronski map. These results concern the complex Grassmannian and complex projective space.

Boris Shapiro and Michael Shapiro made a conjecture in 1993/4 about the Wronski map from the real Grassmannian to real projective space.

THEOREM 1.9. *If the polynomial $\Phi \in \mathbb{P}^{mp}$ has only real zeroes, then every point in $\mathrm{Wr}^{-1}(\Phi)$ is real. Moreover, if Φ has mp simple real zeroes then there are $\#_{m,p}$ real points in $\mathrm{Wr}^{-1}(\Phi)$.*

This was proven when $\min(m,p) = 2$ by Eremenko and Gabrielov [**46**], who subsequently found a second, elementary proof [**42**], which we present in Chapter 11. It was finally settled by Mukhin, Tarasov, and Varchenko [**104**], who showed that every point in the fiber is real. We sketch their proof in Chapter 12. The second statement, about there being the expected number of real roots, follows from this by an argument of Eremenko and Gabrielov that we reproduce in Chapter 13 (Theorem 13.2). It also follows from a second proof of Mukhin, Tarasov, and Varchenko, in which they directly show transversality [**106**], which is equivalent to the second statement. This *Shapiro Conjecture* has appealing geometric interpretations, enjoys links to several areas of mathematics, and has many theoretically satisfying generalizations which we will discuss in Chapters 10, 11, 13, and 14. We now mention two of its interpretations.

EXAMPLE 1.10 (The problem of four lines). A geometric interpretation of the Wronski map and the Shapiro Conjecture when $m = p = 2$ is a variant of the classical problem of the lines in space which meet four given lines. Points in $\mathrm{Gr}(2, \mathbb{C}_3[t])$ correspond to lines in $\mathbb{C}^3$ as follows. The *moment curve* γ in $\mathbb{C}^3$ is the curve with parameterization

$$\gamma(t) \ := \ (t, t^2, t^3)\,.$$

A cubic polynomial f is the composition of γ and an affine-linear map $\mathbb{C}^3 \to \mathbb{C}$, and so a two-dimensional space of cubic polynomials is a two-dimensional space of affine-linear maps whose common kernel is the corresponding line in $\mathbb{C}^3$. (This description is not exact, as some points in $\mathrm{Gr}(2, \mathbb{C}_3[t])$ correspond to lines at infinity.)

Given a polynomial $\Phi(t)$ of degree four with distinct real roots, points in the fiber $\mathrm{Wr}^{-1}(\Phi)$ correspond to the lines in space which meet the four lines tangent to the moment curve γ at its points coming from the roots of Φ. There will be two such lines, and the Shapiro Conjecture asserts that both will be real.

It is not hard to see this directly. Any fractional linear change of parameterization of the moment curve is realized by a projective linear transformation of three-dimensional space which stabilizes the image of the moment curve. Thus we may assume that the polynomial $\Phi(t)$ is equal to $(t^3 - t)(t - s)$, which has roots $-1, 0, 1$, and s, where $s \in (0, 1)$. Applying an affine transformation to three-dimensional space, the moment curve becomes the curve with parameterization

$$\gamma \ : \ t \ \longmapsto \ (6t^2 - 1,\ \tfrac{7}{2}t^3 + \tfrac{3}{2}t,\ \tfrac{3}{2}t - \tfrac{1}{2}t^3)\,. \tag{1.6}$$

Then the lines tangent to γ at the roots $-1, 0, 1$ of Φ have parameterizations

$$(5 - s,\ -5 + s,\ -1)\,,\ (-1,\ s,\ s)\,,\ (5 + s\,, 5 + s\,, 1) \qquad s \in \mathbb{R}\,.$$

These lie on a hyperboloid of one sheet, which is defined by

$$1 - x_1^2 + x_2^2 - x_3^2 \ = \ 0\,. \tag{1.7}$$

We display this geometric configuration in Figure 1.1. There, $\ell(i)$ is the line tangent to γ at the point $\gamma(i)$. The hyperboloid has two rulings. One ruling contains our

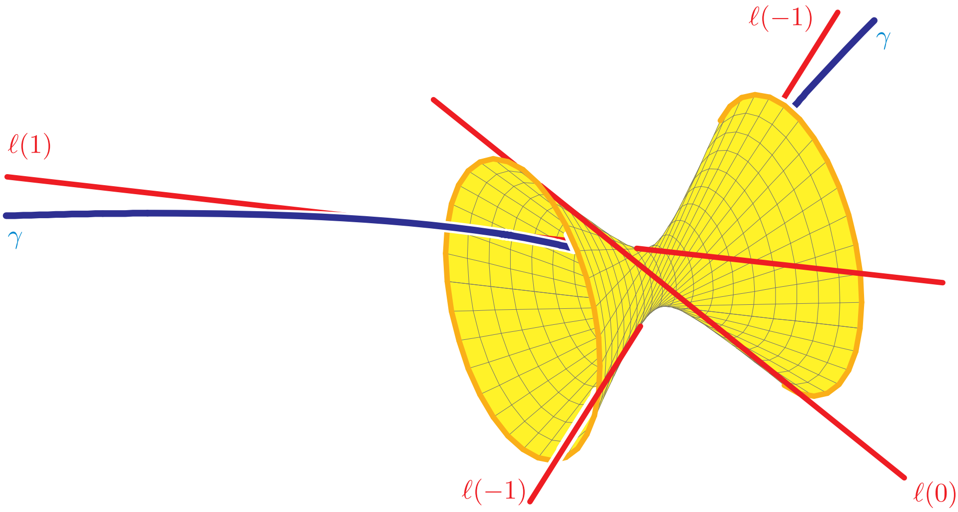


FIGURE 1.1. hyperboloid containing three lines tangent to γ.

three tangent lines and the other ruling (which is drawn on hyperboloid) consists of the lines which meet our three tangent lines.

Now consider the fourth line $\ell(s)$ which is tangent to γ at the point $\gamma(s)$. This has the parameterization

$$\ell(s) \;=\; \left(6s^2-1\,,\ \tfrac{7}{2}s^3+\tfrac{3}{2}s\,,\ \tfrac{3}{2}s-\tfrac{1}{2}s^3\right) \;+\; t\left(12s\,,\ \tfrac{21}{2}s^2+\tfrac{3}{2}\,,\ \tfrac{3}{2}-\tfrac{3}{2}s^2\right).$$

We compute the intersection of the fourth line with the hyperboloid. Substituting its parameterization into (1.7) and dividing by -12 gives the equation

$$(s^3-s)(s^3-s+t(6s^2-2)+9st^2) \;=\; 0\,.$$

The first (nonconstant) factor s^3-s vanishes when $\ell(s)$ is equal to one of $\ell(-1)$, $\ell(-0)$, or $\ell(1)$–for these values of s every point of $\ell(s)$ lies on the hyperboloid. The second factor has solutions

$$t \;=\; -\frac{3s^2-1\pm\sqrt{3s^2+1}}{9s}\,.$$

Since $3s^2+1>0$ for all s, both solutions will be real. In fact, for $s\neq\sqrt{-1/3}$, this will have exactly two solutions.

We may also see this geometrically. Consider the fourth line $\ell(s)$ for $0<s<1$. In Figure 1.2, we look down the throat of the hyperboloid at the interesting part of this configuration. This picture demonstrates that $\ell(s)$ must meet the hyperboloid in two real points. Through each point, there is a real line in the second ruling which meets all four tangent lines, and this proves the Shapiro Conjecture for $m=p=2$.

EXAMPLE 1.11 (Rational functions with real critical points). When $m=2$, the Shapiro Conjecture may be interpreted in terms of rational functions. A rational function $\rho(t)=f(t)/g(t)$ is a quotient of two univariate polynomials, f and g. This

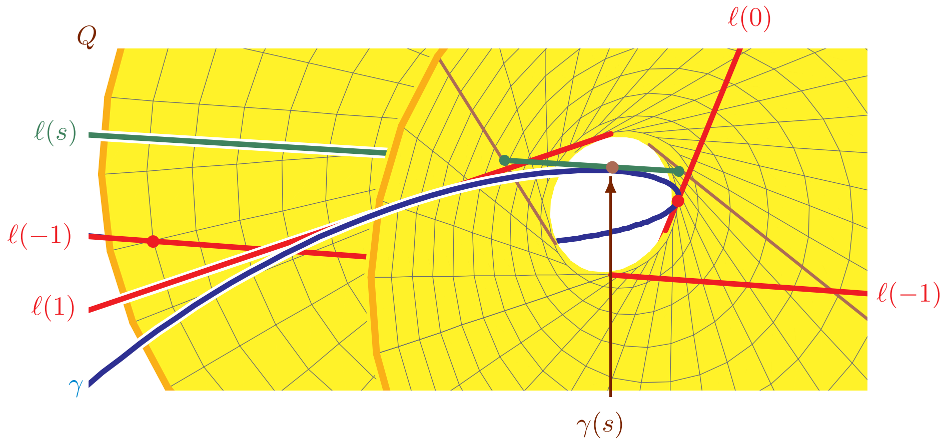


FIGURE 1.2. The fourth tangent line meets hyperboloid in two real points.

defines a map $\rho\colon \mathbb{P}^1 \to \mathbb{P}^1$ whose critical points are those t for which $\rho'(t) = 0$. Since $\rho'(t) = (f'g - g'f)/g^2$, we see that the critical points are the roots of the Wronskian of f and g. Composing the rational function $\rho\colon \mathbb{P}^1 \to \mathbb{P}^1$ with an automorphism of the target $\mathbb{P}^1$ gives an equivalent rational function, and the equivalence class of ρ is determined by the linear span of its numerator and denominator. Thus the Shapiro Conjecture asserts that a rational function having only real critical points is equivalent to a real rational function.

Eremenko and Gabrielov [**46**] proved exactly this statement in 2002, thereby establishing the Shapiro Conjecture in the case $m = 2$.

THEOREM 1.12. *A rational function with only real critical points is equivalent to a real rational function.*

In Chapter 11 we will present an elementary proof of this result that Eremenko and Gabrielov gave in 2005 [**42**].

1.5. Lower bounds

We begin with perhaps the most exciting recent development in real algebraic geometry. This starts with the fundamental observation of Euclid that two points determine a line. Slightly less elementary is that five points in the plane with no three collinear determine a conic. In general, if you have n general points in the plane and you want to pass a rational curve of degree d through all of them, there may be no solution to this interpolation problem (if n is too big), or an infinite number of solutions (if n is too small), or a finite number of solutions (if n is just right). It turns out that "n just right" means $n = 3d-1$ ($n = 2$ for lines where $d = 1$, and $n = 5$ for conics where $d = 2$).

A harder question is, if $n = 3d-1$, *how many* rational curves of degree d interpolate the points? Call this number N_d, so that $N_1 = 1$ and $N_2 = 1$ because the line and conic of the previous paragraph are unique. It has long been known that $N_3 = 12$ (see Example 9.3 for a proof), and in 1873 Zeuthen [**164**] showed that $N_4 = 620$. That was where matters stood until 1989, when Ran [**118**] gave a recursion for these numbers. In the 1990's, Kontsevich and Manin [**88**] used associativity in quantum cohomology of $\mathbb{P}^2$ to give the elegant recursion

$$N_d \;=\; \sum_{a+b=d} N_a N_b \left(a^2 b^2 \binom{3d-4}{3a-2} - a^3 b \binom{3d-4}{3a-1} \right) \,, \tag{1.8}$$

which begins with the Euclidean declaration that two points determine a line ($N_1 = 1$). These numbers grow quite fast, for example $N_5 = 87304$.

The number of real rational curves which interpolate a given $3d-1$ points in the real plane $\mathbb{RP}^2$ will depend rather subtly on the configuration of the points. To say anything about the real rational curves would seem impossible. However this is exactly what Welschinger [**162**] did, by finding an invariant which does not depend upon the choice of points.

A rational curve in the plane is necessarily singular—typically it has $\binom{d-1}{2}$ ordinary double points. Real curves have three types of ordinary double points. Only two types are visible in $\mathbb{RP}^2$, and we are familiar with them from rational cubics, which typically have an ordinary double point. The curve on the left below has a *node* with two real branches, and the curve on the right has a *solitary point* '$\bullet$', where two complex conjugate branches meet.

The third type of ordinary double point is a pair of complex conjugate ordinary double points, which are not visible in $\mathbb{RP}^2$.

Theorem 1.13 (Welschinger [**162**]). *The sum,*

$$\sum (-1)^{\#\{\textit{solitary points in } C\}} \,, \tag{1.9}$$

over all real rational curves C of degree d interpolating $3d-1$ general points in $\mathbb{RP}^2$ does not depend upon the choice of points.

Set W_d to be the sum (1.9). The absolute value of this *Welschinger invariant* is then a lower bound on the number of real rational curves of degree d interpolating $3d-1$ points in $\mathbb{RP}^2$. Since $N_1 = N_2 = 1$, we have $W_1 = W_2 = 1$. Prior to Welschinger's discovery, Kharlamov [**33**, Proposition 4.7.3] (see also Example 9.3) showed that $W_3 = 8$. The question remained whether any other Welschinger invariants were nontrivial. This was settled in the affirmative by Itenberg, Kharlamov, and Shustin [**77, 78**], who used Mikhalkin's Tropical Correspondence Theorem [**99**] to show

1. If $d > 0$, then $W_d \geq \frac{d!}{3}$. (Hence W_d is positive.)
2. $\lim_{d\to\infty} \dfrac{\log N_d}{\log W_d} = 1$. (In fact for d large, $\log N_d \sim 3d \log d \sim \log W_d$.)

In particular, there are always quite a few real rational curves of degree d interpolating $3d-1$ points in $\mathbb{RP}^2$. Since then, Itenberg, Kharlamov, and Shustin [**79**] gave a recursive formula for the Welschinger invariant which is based upon Gathmann and Markwig's [**60**] tropicalization of the Caporaso-Harris [**26**] formula. This shows that $W_4 = 240$ and $W_5 = 18264$. Solomon [**136**] has also found an intersection-theoretic interpretation for these invariants.

These ideas have also found an application. Gahleitner, Jüttler, and Schicho [**58**] proposed a method to compute an approximate parameterization of a plane curve using rational cubics. Later, Fiedler-Le Touzé [**49**] used the result of Kharlamov (that $W_3 = 8$), and an analysis of pencils of plane cubics to prove that this method works.

While the story of this interpolation problem is fairly well-known, it was not the first instance of lower bounds in enumerative real algebraic geometry. In their investigation of the Shapiro Conjecture, Eremenko and Gabrielov found a similar invariant $\sigma_{m,p}$ which gives a lower bound on the number of real points in the inverse image $\mathrm{Wr}^{-1}(\Phi)$ under the Wronski map of a real polynomial $\Phi \in \mathbb{RP}^{mp)}$. Assume that $p \leq m$. If $m+p$ is odd, set $\sigma_{m,p}$ to be

$$\frac{1!2!\cdots(m-1)!(p-1)!(p-2)!\cdots(p-m+1)!(\frac{mp}{2})!}{(p-m+2)!(p-m+4)!\cdots(p+m-2)!\left(\frac{p-m+1}{2}\right)!\left(\frac{p-m+3}{2}\right)!\cdots\left(\frac{p+m-1}{2}\right)!}\,. \tag{1.10}$$

If $m+p$ is even, then set $\sigma_{m,p} = 0$. If $p > m$, then set $\sigma_{m,p} := \sigma_{p,m}$.

THEOREM 1.14 (Eremenko-Gabrielov [**45**]). *If a polynomial $\Phi(t) \in \mathbb{RP}^{mp}$ of degree mp is a regular value of the Wronski map, then there are at least $\sigma_{m,p}$ real m-dimensional subspaces of polynomials of degree $m+p-1$ with Wronskian Φ.*

REMARK 1.15. The number of complex points in $\mathrm{Wr}^{-1}(\Phi)$ is $\#_{m,p}$ (1.5). It is instructive to compare these numbers. We show them for $m+p = 11$ and $m = 2,\ldots,5$.

m	2	3	4	5
$\sigma_{m,p}$	14	110	286	286
$\#_{m,p}$	4862	23371634	13672405890	396499770810

We also have $\sigma_{7,6} \approx 3.4 \cdot 10^4$ and $\#_{7,6} \approx 9.5 \cdot 10^{18}$. Despite this disparity in their magnitudes, the asymptotic ratio of $\log(\sigma_{m,p})/\log(\#_{m,p})$ appears to be close to 1/2. We display this ratio in the table below, for different values of m and p.

$\frac{\log(\sigma_{m,p})}{\log(\#_{m,p})}$		m					
		2	$\frac{m+p-1}{10}$	$2\frac{m+p-1}{10}$	$3\frac{m+p-1}{10}$	$4\frac{m+p-1}{10}$	$5\frac{m+p-1}{10}$
$m+p-1$	100	0.47388	0.45419	0.43414	0.41585	0.39920	0.38840
	1000	0.49627	0.47677	0.46358	0.45185	0.44144	0.43510
	10000	0.49951	0.48468	0.47510	0.46660	0.45909	0.45459
	100000	0.49994	0.48860	0.48111	0.47445	0.46860	0.46511
	1000000	0.49999	0.49092	0.48479	0.47932	0.47453	0.47168
	10000000	0.50000	0.49246	0.48726	0.48263	0.47857	0.47616

Thus, the lower bound on the number of real points in a fiber of the Wronski map appears asymptotic to the square root of the number of complex solutions.

It is interesting to compare this to the the result of Shub and Smale [**132**] that the expected number of real solutions to a system of n Gaussian random polynomials in n variables of degrees $d_1,\ldots,d_n$ is $\sqrt{d_1\cdots d_n}$, which is the square root of the number of complex solutions to such a system of polynomials. Thus $\frac{1}{2}$ is the ratio of the logarithm of the expected number of complex solutions to the logarithm of the expected number of real solutions.

The idea behind the proof of Theorem 1.14 is to compute the topological degree of the real Wronski map, which is the restriction of the Wronski map to real subspaces of polynomials,

$$\mathrm{Wr}_{\mathbb{R}} \ := \ \mathrm{Wr}|_{\mathrm{Gr}(m,\mathbb{R}_{m+p-1}[t])} \ : \ \mathrm{Gr}(m,\mathbb{R}_{m+p-1}[t]) \ \longrightarrow \ \mathbb{RP}^{mp}\,.$$

This maps the Grassmannian of real subspaces of polynomials of degree $m+P-1$ to the space of real polynomials of degree mp. Recall that the topological degree (or mapping degree) of a map $f\colon X \to Y$ between two oriented manifolds X and Y of the same dimension is the number d such that $f_*[X] = d[Y]$, where $[X]$ and $[Y]$ are the fundamental cycles of X and Y in homology, respectively, and f_* is the functorial map in homology. When f is differentiable, this mapping degree may be computed as follows. Let $y \in Y$ be a regular value of f so that the derivative map on tangent spaces $df_x\colon T_xX \to T_yY$ is an isomorphism at any point x in the fiber $f^{-1}(y)$ above y. Since X and Y are oriented, the isomorphism df_x either preserves the orientation of the tangent spaces or it reverses the orientation. Let P be the number of points $x \in f^{-1}(y)$ at which df_x preserves the orientation and R be the number of points where the orientation is reversed. Then the mapping degree of f is the difference $P - R$.

There is a slight problem in computing the mapping degree of $\mathrm{Wr}_{\mathbb{R}}$, as neither the real Grassmannian $\mathrm{Gr}_{\mathbb{R}}$ nor the real projective space $\mathbb{RP}^{mp}$ are orientable when $m+p$ is odd, and thus the mapping degree of $\mathrm{Wr}_{\mathbb{R}}$ is not defined when $m+p$ is odd. Eremenko and Gabrielov get around this by computing the degree of the restriction of the Wronski map to open cells of $\mathrm{Gr}_{\mathbb{R}}$ and $\mathbb{RP}^{mp}$, where $\mathrm{Wr}_{\mathbb{R}}$ is a proper map. They also show that it is the degree of a lift of the Wronski map to oriented double covers of both spaces. The degree bears a resemblance to the Welschinger invariant as it has the form $|\sum \pm 1|$, the sum over all real points in $\mathrm{Wr}_{\mathbb{R}}^{-1}(\Phi)$, for Φ a regular value of the Wronski map. This resemblance is no accident. Solomon [**136**] showed how to orient a moduli space of rational curves with marked points so that the Welschinger invariant is indeed the degree of a map.

While both of these examples of geometric problems possessing a lower bound on their numbers of real solutions are quite interesting, they are rather special. The existence of lower bounds for more general geometric problems or for more general systems of polynomials would be quite important in applications, as these lower bounds guarantee the existence of real solutions.

With Soprunova, we [**138**] set out to develop a theory of lower bounds for sparse polynomial systems, using the approach of Eremenko and Gabrielov via mapping degree. This is a first step toward practical applications of these ideas. Chapters 7 and 8 will elaborate this theory. Here is an outline:

- (*i*) Realize the solutions to a system of polynomials as the fibers of a map from a toric variety.
- (*ii*) Characterize when a toric variety (or its double cover) is orientable, thus determining when the degree of this map (or a lift to double covers) exists.
- (*iii*) Develop a method to compute the degree in some (admittedly special) cases.
- (*iv*) Give a nice family of examples to which this theory applies.
- (*v*) Use the sagbi degeneration of a Grassmannian to a toric variety [**154**, Ch. 11] and the systems of (*iv*) to reprove the result of Eremenko and Gabrielov.

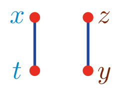
Example 1.16. We close this overview with one example from this theory. Let t, x, y, z be indeterminates, and consider a sparse polynomial of the form

$$c_4\,txyz \;+\; c_3(txz + xyz) \;+\; c_2(tx + xz + yz) \;+\; c_1(x + z) \;+\; c_0\,, \tag{1.11}$$

where the coefficients $c_0, \ldots, c_4$ are real numbers.

Theorem 1.17. *A system involving four polynomials of the form* (1.11) *has six solutions, at least two of which are real.*

We make some remarks to illustrate the ingredients of this theory. First, the monomials in the sparse system (1.11) are the integer points in the order polytope of the poset P,

$P \;:=$ 

.

That is, each monomial corresponds to an order ideal of P (a subset which is closed upwards). The number of complex roots is the number of linear extensions of the poset P. There are six, as each is a permutation of the word $txyz$ where t precedes x and y precedes z.

One result (*ii*) characterizes polytopes whose associated polynomial systems will have a lower bound, and many order polytopes satisfy these conditions. Another result (*iv*) computes that lower bound for certain families of polynomials with support an order polytope. Polynomials in these families have the form (1.11) in that monomials with the same total degree have the same coefficient. For such polynomials, the lower bound is the absolute value of the sum of the signs of the permutations underlying the linear extensions. We list these for P.

permutation	$txyz$	$tyxz$	$ytxz$	$tyzx$	$ytzx$	$yztx$	sum
sign	$+$	$-$	$+$	$+$	$-$	$+$	2

This shows that the lower bound in Theorem 1.17 is two.

Table 1.1 records the frequency of the different numbers of real solutions in each of 10,000,000 instances of this polynomial system, where the coefficients were chosen uniformly from $[-200, 200]$. This computation took 13 gigahertz-hours.

Table 1.1. Observed frequencies of numbers of real solutions.

number of real solutions	0	2	4	6
frequency	0	9519429	0	480571

The apparent gap in the numbers of real solutions in Table (1.1) (four does not seem a possible number of real solutions) is proven for the system of Example 1.16 in Section 8.3. This is the first instance we have seen of this phenomena of gaps in the numbers of real solutions. More are found in [**138**], [**123**], and some are presented in Chapters 8, 13, and 14. Many examples of lower bounds continue to be found, e.g. [**3**].