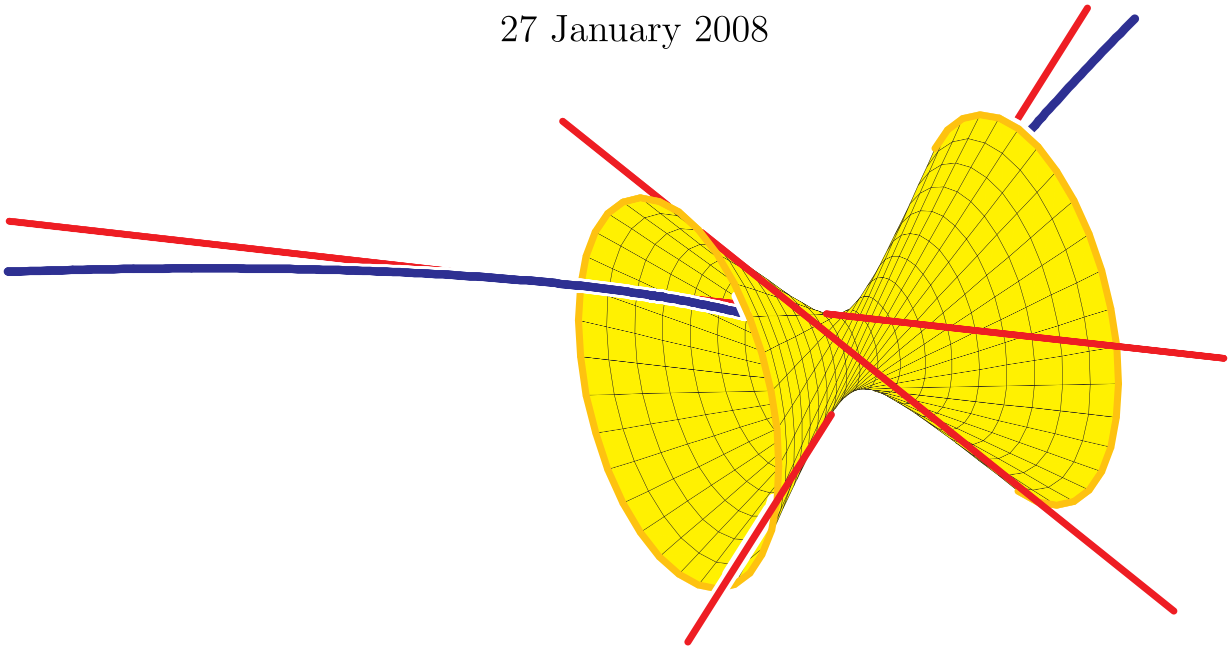
Real Solutions to Equations From Geometry

Frank Sottile
Department of Mathematics
Texas A&M University
College Station
TX 77843
USA
sottile@math.tamu.edu
http://www.math.tamu.edu/~sottile

27 January 2008



Contents

Abstract

Understanding, finding, or even deciding on the existence of real solutions to a system of equations is a very difficult problem with many applications. While it is hopeless to expect much in general, we know a surprising amount about these questions for systems which possess additional structure. Particularly fruitful—both for information on real solutions and for applicability—are systems whose additional structure comes from geometry. Such equations from geometry for which we have information about their real solutions are the subject of these lecture notes.

We will focus on equations from toric varieties and homogeneous spaces, particularly Grassmannians. Not only is much known in these cases, but they encompass some of the most common applications. The results we discuss may be grouped into three themes:

(I) Upper bounds on the number of real solutions.

(II) Geometric problems that can have all solutions be real.

(III) Lower bounds on the number of real solutions.

Upper bounds as in (I) bound the complexity of the set of real solutions—they are one of the sources for the theory of o-minimal structures which are an important topic in real algebraic geometry. The existence (II) of geometric problems that can have all solutions be real was initially surprising, but this phenomena now appears ubiquitous. Lower bounds as in (III) give an existence proof for real solutions. Their most spectacular manifestation is the nontriviality of the Welschinger invariant, which was computed via tropical geometry.

I thank the Institut Henri Poincaré, where a preliminary version of these notes were produced during a course I taught there in November 2005. These notes were revised and expanded during course at Texas A&M University in 2007 and a lecture series at the Centre Interfacultaire Bernoulli at EPFL in 2008. I also thank Prof. Dr. Peter Gritzmann of the Technische Universität München, whose hospitality enabled the completion of the first version of these notes. During this period, my research was supported by NSF grants DMS-0701059 and CAREER grant DMS-0538734. Lastly, I thank the students, both young and old, who attended the lectures. do not forget: Askold, Marie-Françoise Roy, and my collaborators!

Frank Sottile
01.28.08, College Station

Chapter 1: Overview

Introduction

Our goal will be to say something meaningful about the real solutions to a system of 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 d its number of complex solutions†. Since every real number is complex, and since nonreal solutions come in conjugate pairs, we have the following trivial inequality

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

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

Our discussion will have three themes:

(I) Sometimes, the upper bound d is not sharp and there is a smaller bound for r.

(II) For many problems from enumerative geometry, the upper bound is sharp.

(III) The lower bound for r may be significantly larger than $d \mod 2$.

A lot of time will be spent on the Shapiro Conjecture (Theorem of Mukhin, Tarasov, and Varchenko [71])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.

We will not describe how to actually find the solutions to a system (1.1) and there will be no discussion of algorithms nor any complexity analysis. The book of Basu, Pollack, and Roy [2] is an excellent place to learn about algorithms for computing real algebraic varieties and finding real solutions. We remark however that some of the techniques we employ to study real solutions underlie numerical algorithms to find all complex solutions.

One class of systems that we will study are systems of sparse polynomials. 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 shall always assume that our systems are *generic* in the sense that all of their solutions occur with multiplicity 1, and the number d of complex solutions is stable under certain allowed perturbations of the coefficients.

Sometimes, we 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

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

of monomials from $\mathcal{A}$ is a *sparse polynomial* with *support $\mathcal{A}$*. Sparse polynomials naturally define functions on the complex torus $(\mathbb{C}^\times)^n$. 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 $\mathcal{A}$*. Often, these are called *unmixed systems* in contrast to *mixed systems* where each polynomial may have different support.

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

Theorem 1.1 (Kushnirenko [6]) *A system of n polynomials in n variables with common support $\mathcal{A}$ has at most $n!\mathrm{vol}(\Delta_{\mathcal{A}})$ isolated solutions in $(\mathbb{C}^\times)^n$, 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

$$\mathrm{vol}(t_1P_1 \;+\; t_2P_2 \;+\; \cdots \;+\; t_nP_n)$$

is a homogeneous polynomial of degree n in the variables $t_1, \ldots, t_n$ [45, Exercise 15.2.6]. The *mixed volume* $\mathrm{MV}(P_1, \ldots, P_n)$ of $P_1, \ldots, P_n$ is the coefficient of the monomial $t_1t_2\cdots t_n$ in this polynomial.

Theorem 1.2 (Bernstein [7]) *A system of n polynomials in n variables where the polynomials have supports $\mathcal{A}_1, \ldots, \mathcal{A}_n$ has at most $\mathrm{MV}(\Delta_{\mathcal{A}_1}, \ldots, \Delta_{\mathcal{A}_n})$ isolated solutions in $(\mathbb{C}^\times)^n$, and exactly this number when the polynomials are generic for their given support.*

Since $\mathrm{MV}(P_1, \ldots, P_n) = n!\mathrm{vol}(P)$ when $P_1 = \cdots = P_n = P$, this generalizes Kushnirenko's Theorem. We will prove Kushnirenko's Theorem in Chapter 3, but will not present a proof of Bernstein's Theorem. For a proof, we suggest two excellent sources by Sturmfels. Both are similar, but the first is self-contained and superbly written.

- [103] *Polynomial equations and convex polytopes*, Amer. Math. Monthly **105** (1998), no. 10, 907–922.
- [104] Chapter 3 in *Solving systems of polynomial equations*, CBMS, vol. 97, American Mathematical Society, Providence, RI, 2002.

The bound of Theorem 1.1 and its generalization Theorem 1.2 is often called the *BKK bound* for Bernstein, Khovanskii, and Kushnirenko [6].

1.1 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 [23] (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 a bound to multivariate polynomials with his fundamental *fewnomial bound*.

Theorem 1.3 (Khovanskii [59]) *A system of n polynomials in n variables having a total of $l+n+1$ distinct monomials has at most*

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

nondegenerate positive real 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 extra trivial zeroes which differ from the positive solutions only by some sign patterns.

This is the first of many results which verified the principle of Berenstein 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 their degree. Khovanskii's work was also a motivation for the notion of o-minimal structures. The main point of Khovanskii's theorem is the existence of such a bound and not the actual bound itself. 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 $l+n+1$ monomials. Khovanskii's Theorem gives a bound for this number, but the bound is enormous. For example, when $l = n = 2$, it is 5184. Because of this, it 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 [66]) *Two trinomials in two variables have at most 5 nondegenerate positive real solutions.*

This bound sharp. Haas [46] had shown that the system of two trinomials in x and y

$$10x^{106} + 11y^{53} - 11y \;=\; 10y^{106} + 11x^{53} - 11x \;=\; 0\,, \tag{1.2}$$

has 5 positive solutions.

Since we may multiply one of the trinomials by an arbitrary monomial without changing the solutions, we can assume that the two trinomials (1.2) 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 5 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 gives strong evidence that Khovanskii's bound may be improved. Such an improved bound was recently found [14].

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 1 positive real solution, so $X(0,n) = 1$. Theorem 1.5 was inspired by the sharp bound of Theorem 1.6 [9] when $l = 1$. A set $\mathcal{A}$ of exponents is *primitive* if $\mathcal{A}$ affinely spans the full integer lattice $\mathbb{Z}^n$.

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 system 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 real solutions, not just positive solutions. We will discuss this in Section 4.2. Further analysis gives the sharp bound for $X(1,n)$.

Theorem 1.7 (Bihan [11]) $X(1,n) = n+1$.

In contrast to these results establishing absolute upper bounds for 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 [100] proved the following.

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 7. The state of affairs in 2001 was presented in [93].

1.2 The Wronski map and the Shapiro Conjecture

The *Wronskian* of univariate polynomials $f_1, f_2, \ldots, f_k$ in the variable t is the determinant

$$W(f_1,\ldots,f_k) \;:=\; \det\bigl((\tfrac{\partial}{\partial t})^{j-1}f_i(t)\bigr)_{i,j=1,\ldots,k}\,.$$

When the polynomials f_i have degree d and are linearly independent, the Wronskian has degree at most $k(d{+}1{-}k)$. For example, if $k = 2$, then $W(f,g) = f'g - fg'$, which has degree $2d{-}2$ as the coefficients of t^{2d-1} in this expression cancel. Up to a scalar, the Wronskian depends only upon the linear span of the polynomials $f_1, f_2, \ldots, f_k$. Removing these ambiguities gives the *Wronski map*,

$$\mathrm{Wr} \;:=\; \mathrm{Gr} \;\longrightarrow\; \mathbb{P}^{k(d+1-k)}\,, \tag{1.3}$$

where Gr is the *Grassmannian* of k-dimensional subspaces of the linear space of degree d polynomials in the variable t, and $\mathbb{P}^{k(d+1-k)}$ is the projective space of polynomials of degree at most $k(d{+}1{-}k)$, which has dimension equal to the the dimension of the Grassmannian.

Work of Schubert in 1886 [81], combined with a result of Eisenbud and Harris in 1983 [29] shows that the Wronski map is surjective and the general polynomial $\Phi \in \mathbb{P}^{k(d+1-k)}$ has

$$\#_{d,k} \;:=\; \frac{1!2!\cdots(k-1)!\cdot[k(d+1-k)]!}{(d+1-k)!(d+2-k)!\cdots d!} \tag{1.4}$$

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. This was proven when $\min(k, d+1-k) = 2$ by Eremenko and Gabrielov [33], and finally settled by Mukhin, Tarasov, and Varchenko [71]. There are now two more anounced proofs [72, 43].

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

We will discuss a proof of this in Chapter 10. 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 8, 9, and 11. We now mention two of its interpretations.

1.2.1 The problem of four lines

A geometric interpretation of the Wronski map and the Shapiro Conjecture when $d = 3$ and $k = 2$ is a variant of the classical problem of the lines in space which meet four given lines. The *moment curve* γ in 3-dimensional space is the curve with parametrization

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

Given a polynomial $\Phi(t)$ of degree 4 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 the points corresponding to 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 parametrization of the moment curve is realized by a projective linear transformation of 3-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 3-dimensional space, the moment curve becomes the curve with parametrization

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

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 Q of one sheet, which is defined by

$$1 - x_1^2 + x_2^2 - x_3^2 \;=\; 0\,. \tag{1.5}$$

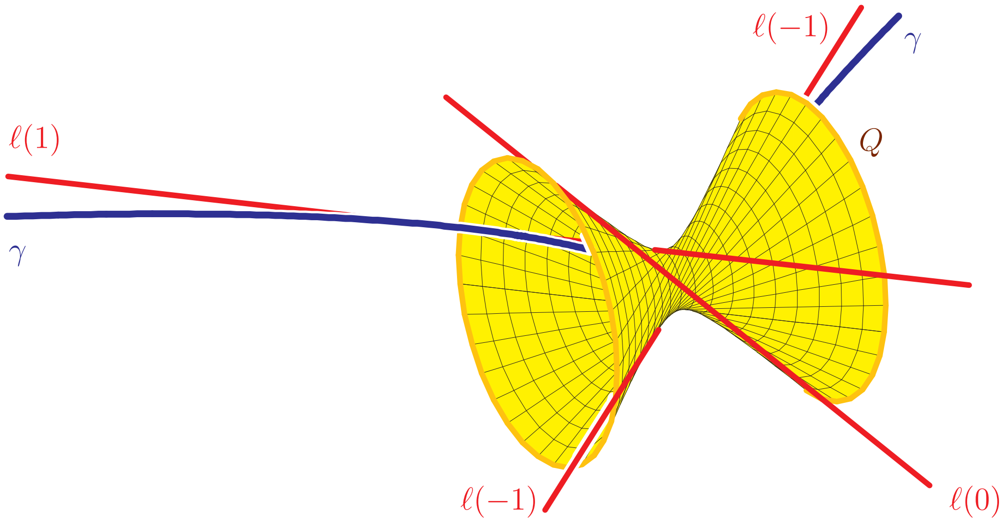


Figure 1.1: Quadric containing three lines tangent to γ.

We display this geometric configuration in Figure 1.1. There, $\ell(i)$ is the line tangent to γ at the point $\gamma(i)$. The quadric Q has two rulings. One ruling contains our three tangent lines and the other ruling (which is drawn on Q) 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 parametrization

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

Substituting this parametrization into (1.5) 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 quadric Q. 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.

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 Q 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 Shapiro's conjecture for $d = 3$ and $k = 2$.

1.2.2 Rational functions with real critical points

When $k = 2$, the Shapiro conjecture may be interpreted in terms of rational functions. A rational function $r(t) = f(t)/g(t)$ is a quotient of two univariate polynomials, f and g. This defines a map $r\colon \mathbb{P}^1 \to \mathbb{P}^1$ whose critical points are those t for which $r'(t) = 0$. Since

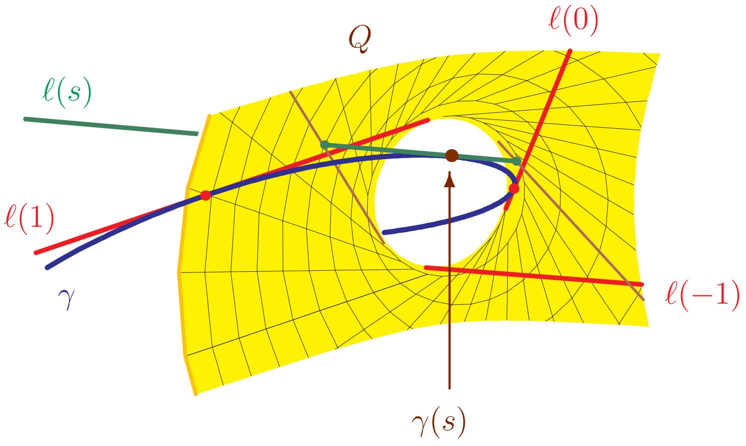


Figure 1.2: The fourth tangent line meets Q in two real points.

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

Eremenko and Gabrielov [33] proved exactly this statement in 2002, and thereby established the Shapiro Conjecture in the case $k = 2$.

Theorem 1.10 *A rational function with only real critical points is equivalent to a real rational function.*

In Chapter 9 we will present an elementary proof of this result that Eremenko and Gabrielov found in 2005 [31].

1.3 Lower bounds

We begin with some of perhaps the most exciting recent development in real algebraic geometry. It begins with the fundamental observation of Euclid that two points determine a line. Many people who have studied geometry know that five points on the plane determine a conic. In general, if you have m random 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 m is too big), or an infinite number of solutions (if m is too small), or a finite number of solutions (if m is just right). It turns out that "m just right" means $m = 3d-1$ ($m = 2$ for lines and $m = 5$ for conics).

A harder question is, if $m = 3d-1$, *how many* rational curves of degree d interpolate the points? Let's 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 $N_3 = 12$, and in 1873 Zeuthen [110] showed that $N_4 = 620$. That was where matters stood until 1989, when Ran [74] gave a recursion for these numbers. About ten years ago, Kontsevich and

Manin [63] used associativity in quantum cohomology of $\mathbb{P}^2$ to give the elegant recursion

$$N_d \;=\; \sum_{a+b=d} N_a N_b \left(a^2 b^2 \binom{3d-4}{3a-2} - a^3 b \binom{3d-4}{3a-1} \right) \;, \tag{1.6}$$

which begins with the Eucidean declaration that two points determine a line ($N_1 = 1$).

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 [109] did. He found an invariant which does not depend upon the choice of points.

A rational curve in the plane is necessarily singular—typically it has $\binom{d-1}{2}$ nodes. Real curves have three types of nodes. Two types are visible in $\mathbb{RP}^2$, and we are familiar with them from rational cubics. 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 node is a pair of complex conjugate nodes, which are not visible in $\mathbb{RP}^2$.

Theorem 1.11 (Welschinger [109]) *The sum,*

$$\sum (-1)^{\#\{\textit{solitary points, } \bullet, \textit{in } C\}} \,, \tag{1.7}$$

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 the points.

Set W_d to be the sum (1.7). The absolute value of this *Welschinger invariant* is then a lower bound for 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 [22, Proposition 4.7.3] (see also Section 7.2.2) 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 [54, 55], who used Mikhalkin's Tropical Correspondence Theorem [70] to show

1. If $d > 0$, then $W_d \geq \frac{d!}{3}$. (Hence W_d is positive.)

2. $\displaystyle\lim_{d\to\infty} \frac{\log N_d}{\log W_d} = 1$. (Note that $\log N_d \sim 3d \log d$.)

In particular, there are always quite a few real rational curves of degree d interpolating $3d-1$ points in $\mathbb{RP}^2$. Recently, Itenberg, Kharlamov, and Shustin announced a recursive formula for the Welschinger invariant which is based upon Gathmann and Markwig's [42] tropicalization of the Caporaso-Harris [17] formula. Solomon [85] has also found a moduli space interpretation for these invariants.

These ideas recently found an application. Gahleitner, Jüttler, and Schicho [41] proposed a method to compute an approximate parametrization of a plane curve using rational cubics. Later, Fiedler-Le Touzé [35] used the result of Degtyarev and 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 $\sigma_{d,k}$ 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 $\Phi \in \mathbb{RP}^{k(d+1-k)}$. If d is even, then we let p be the minimum of k and $d+1-k$, and set m to be the maximum of these numbers. With these definitions, set

$$\sigma_{d,k} := \frac{1!2!\cdots(p-1)!(m-1)!(m-2)!\cdots(m-p+1)!(\frac{mp}{2})!}{(m-p+2)!(m-p+4)!\cdots(m+p-2)!\left(\frac{m-p+1}{2}\right)!\left(\frac{m-p+3}{2}\right)!\cdots\left(\frac{m+p-1}{2}\right)!}\ .$$

If d is odd, then set $\sigma_{d,k} = 0$.

Theorem 1.12 (Eremenko-Gabrielov [32]) *If $\Phi(t) \in \mathbb{RP}^{k(d+1-k)}$ is a generic real polynomial of degree $k(d+1-k)$ (a regular value of the Wronski map), then there are at least $\sigma_{d,k}$ real k-dimensional subspaces of polynomials of degree d with Wronskian Φ.*

Remark 1.13 Recall that the number of complex points in $\mathrm{Wr}^{-1}(\Phi)$ is $\#_{d,k}$ (1.4). It is instructive to compare these numbers. We show them for $d = 10$ and $k = 2, \ldots, 5$.

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

We also have $\sigma_{12,6} \approx 3.4 \times 10^4$ and $\#_{12,6} \approx 9.5 \times 10^{18}$. Despite this disparity in their magnitudes, the asymptotic ratio of $\log(\sigma_{d,k})/\log(\#_{d,k})$ appears to be close to 1/2. We display this ratio in the table below, for different values of d and k.

$\frac{\log(\sigma_{d,k})}{\log(\#_{d,k})}$		k					
		2	$d/10$	$2d/10$	$3d/10$	$4d/10$	$5d/10$
d	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 seems to be 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 [82] 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.

The idea behind the proof of Theorem 1.12 is to compute the topological degree of the real Wronski map

$$\mathrm{Wr}_{\mathbb{R}} \;:=\; \mathrm{Wr}|_{\mathrm{Gr}_{\mathbb{R}}} \;:\; \mathrm{Gr}_{\mathbb{R}} \;\longrightarrow\; \mathbb{RP}^{k(d+1-k)}\,,$$

which maps the Grassmannian of real subspaces to the space of real Wronski polynomials. Recall[†] that the topological 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 homology cycles of X and Y, respectively, and f_* is the functorial map in homology. When f is differentiable, this may be computed as follows. Let $y \in Y$ be a regular value of f so that at any point x in the fiber $f^{-1}(y)$ above y the derivative $d_xf\colon T_xX \to T_yY$ is an isomorphism. Since X and Y are oriented, the isomorphism d_xf either preserves the orientation or it reverses it. Let p be the number of points $x \in f^{-1}(y)$ at which d_xf preserves the orientation and r be the number of points where the orientation is reversed. Then the degree of f is the difference $p - r$.

There is a slight problem in computing the degree of $\mathrm{Wr}_{\mathbb{R}}$, as neither the real Grassmannian nor the real projective space is orientable when d is even, and thus the topological degree of $\mathrm{Wr}_{\mathbb{R}}$ is not defined when d is even. 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 $\mathbb{RP}^{k(d+1-k)}$, which 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 [85] showed how to orient a moduli space of rational curves with marked points so that the Welschinger incariant 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 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 [86] set out to develop a theory of lower bounds for sparse polynomial systems, using the approach of Eremenko and Gabrielov via topological degree. This is a first step toward practical applications of these ideas. Chapter 6 will elaborate this theory. Here is an outline of this theory.

(*i*) Realize the solutions to a system of polynomials as the fibers of a map from a toric variety.

(*ii*) Give a condition which implies that 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 and the systems of (*iv*) to recover the result of Eremenko and Gabrielov.

[†]cite a basic topology text

Example 1.14 We close this Chapter with one example from this theory. Let w, x, y, z be indeterminates, and consider a sparse polynomial of the form

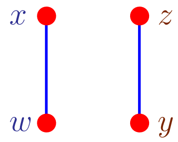
$$\begin{gathered} c_4\, wxyz \\ + c_3(wxz + xyz) \\ + c_2(wx + xz + yz) \\ + c_1(x + z) \\ + c_0\,, \end{gathered} \tag{1.8}$$

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

Theorem 1.15 *A system of four equations involving polynomials of the form* (1.8) *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 (1.8) 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 $wxyz$ where w precedes x and y precedes z.

One result (*ii*) gives conditions on the Newton polytope which imply that any polynomial system with that Newton polytope has a lower bound, and 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 (1.8) 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	$wxyz$	$wyxz$	$ywxz$	$wyzx$	$ywzx$	$yzwx$	sum
sign	$+$	$-$	$+$	$+$	$-$	$+$	2

This shows that the lower bound in Theorem 1.15 is 2.

We record the frequency of the different root counts in each of 10,000,000 instances of this polynomial system, where the coefficients were chosen uniformly from $[-200, 200]$.

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

This computation took 198548 CPU-seconds, or 5:30:54 on a AMD Athlon 64-bit dual core processor 4600+ (2.4 GHz, 512KB cache), using a single core.

The apparent gap in the numbers of real solutions (4 does not seem a possible number of real solutions) is proven for the system of Example 1.14. This is the first instance we have seen of this phenomena of gaps in the numbers of real solutions. More are found in [86] and [78].

Chapter 2: Real solutions to univariate polynomials

Introduction

Before we study the real solutions to systems of multivariate polynomials, it will be instructive to review some of what is known for univariate polynomials. The strength and precision of results concerning real solutions to univariate polynomials forms the gold standard in this subject of real roots to systems of polynomials. We will only discuss two results about univariate polynomials: Descartes' rule of signs and Sturm's Theorem. Descartes' rule of signs, or rather its generalization the Budan-Fourier Theorem, gives a bound for the number of roots in an interval, counted with multiplicity. Sturm's theorem is topological—it gives the number of roots of a univariate polynomial in an interval and does not count their multiplicity. From Sturm's Theorem we obtain a symbolic algorithm to count the number of real solutions to a system of multivariate polynomials in many cases. We underscore the topological nature of Sturm's Theorem by presenting a new and very elementary proof due to Burda and Khovanskii [16]. These and other fundamental results about real roots of univariate polynomials were established in the 19th century. In contrast, the main results about real solutions to multivariate polynomials were only established in recent decades.

2.1 Descartes' rule of signs

Descartes' rule of signs [23] is fundamental for real algebraic geometry. Suppose that f is a univariate polynomial and write its terms in increasing order of their exponents,

$$f \;=\; c_0 x^{a_0} + c_1 x^{a_1} + \cdots + c_n x^{a_m}\,, \tag{2.1}$$

where $c_i \neq 0$ and $a_0 < a_1 < \cdots < a_m$.

Theorem 2.1 (Descartes' rule of signs) *The number, r, of positive roots of f, counted with multiplicity, is at most the variation in sign of the coefficients of f,*

$$\#\{i \mid 1 \leq i \leq m \text{ and } c_{i-1}c_i < 0\} \;\leq\; r\,,$$

and the difference between the variation and r is even.

We will prove a generalization, the Budan-Fourier Theorem, which provides a similar estimate for any interval in $\mathbb{R}$. We first formalize this notion of variation in sign that appears in Descartes' rule.

The *variation* $\mathrm{var}(c)$ in a finite sequence c of real numbers is the number of times that consecutive elements of the sequence have opposite signs, after we remove any 0s in the sequence. For example, the first sequence below has four variations, while the second has three.

$$8, -4, -2, -1, 2, 3, -5, 7, 11, 12 \qquad\qquad -1, 0, 1, 0, 1, -1, 1, 1, 0, 1\,.$$

Suppose that we have a sequence $F = (f_0, f_1, \ldots, f_k)$ of polynomials and a real number $a \in \mathbb{R}$. Then $\mathrm{var}(F, a)$ is the variation in the sequence $f_0(a), f_1(a), \ldots, f_k(a)$. This notion also makes sense when $a = \pm\infty$: We set $\mathrm{var}(F, \infty)$ to be the variation in the sequence of leading coefficients of the $f_i(t)$, which are the signs of $f_i(a)$ for $a \gg 0$, and $\mathrm{var}(F, -\infty)$ is the variation in the leading coefficients of $f_i(-t)$.

Given a univariate polynomial $f(t)$ of degree k, let δf be the sequence of its derivatives,

$$\delta f \;:=\; (f(t), f'(t), f''(t), \ldots, f^{(k)}(t))\,.$$

For $a, b \in \mathbb{R} \cup \{\pm\infty\}$, let $r(f, a, b)$ be the number of roots of f in the interval $(a, b]$, counted with multiplicity. We prove a version of Descartes' rule due to Budan [15] and Fourier [36].

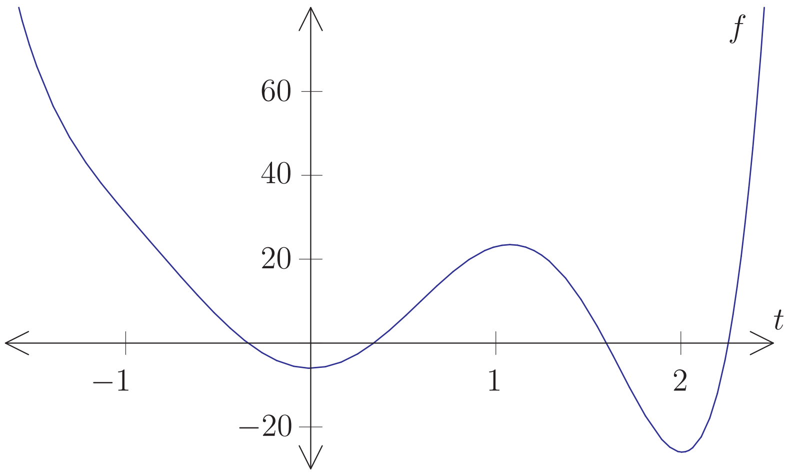
Theorem 2.2 (Budan-Fourier) *Let $f \in \mathbb{R}[t]$ be a univariate polynomial and $a < b$ two numbers in $\mathbb{R} \cup \{\pm\infty\}$. Then*

$$\mathrm{var}(\delta f, a) \;-\; \mathrm{var}(\delta f, b) \;\geq\; r(f, a, b)\,,$$

and the difference is even.

We may deduce Descartes' rule of signs from the Budan-Fourier Theorem once we observe that for the polynomial $f(t)$ (2.1), $\mathrm{var}(\delta f, 0) = \mathrm{var}(c_0, c_1, \ldots, c_m)$, while $\mathrm{var}(\delta f, \infty) = 0$, as the leading coefficients of δf all have the same sign.

Example 2.3 The the sextic $f = 5t^6 - 4t^5 - 27t^4 + 55t^2 - 6$ whose graph is displayed below



has four real zeroes at approximately -0.339311, 0.340401, 1.59753, 2.25615. If we evaluate the derivatives of f at 0 we obtain

$$-6,\ 0,\ 110,\ 0,\ -648,\ -480,\ 3600\,,$$

which has 3 variations in sign. If we evaluate the derivatives of f at 2, we obtain

$$-26,\ -4,\ 574,\ 2544,\ 5592,\ 6720,\ 3600\,,$$

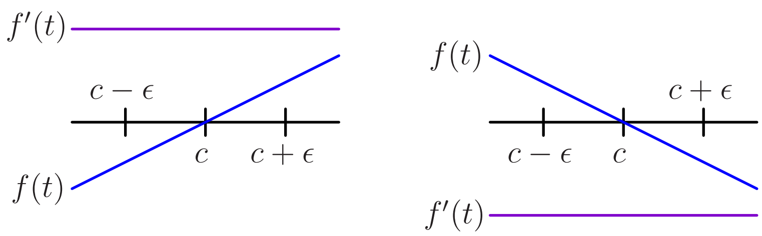
which has one sign variation. Thus, by the Budan-Fourier Theorem, f has eiter 2 or 0 roots in the interval $(0,2)$, counted with multiplicity. This agrees with our observation that f has 2 roots in the interval $[0,2]$.

Proof of Budan-Fourier Theorem. Observe that $\mathrm{var}(\delta f, t)$ can only change when t passes a root c of some polynomial in the sequence δf of derivatives of f. Suppose that c is a root of some derivative of f and let $\epsilon > 0$ be a positive number such that no derivative $f^{(i)}$ has a root in the interval $[c-\epsilon, c+\epsilon]$, except possibly at c. Let m be the order of vanishing of f at c. We will prove that

$$\begin{array}{ll} (1) & \mathrm{var}(\delta f, c) = \mathrm{var}(\delta f, c+\epsilon), \ \text{ and} \\ (2) & \mathrm{var}(\delta f, c-\epsilon) \geq \mathrm{var}(\delta f, c) + m, \quad \text{and the difference is even.} \end{array} \tag{2.2}$$

We deduce the Budan-Fourier theorem from these conditions. As t ranges from a to b, both $r(f,a,t)$ and $\mathrm{var}(\delta f, t)$ only change when t passes a root c of f. At such a point, $r(f,a,t)$ jumps by the multiplicity m of that root, while $\mathrm{var}(\delta f, t)$ drops by m, plus a nonnegative even integer. Thus the sum $r(f,a,t) + \mathrm{var}(\delta f, t)$ can only change at roots c of f, where it drops by an even integer. Since this sum equals $\mathrm{var}(\delta f, a)$ when $t = a$, the Budan-Fourier Theorem follows.

Let us now prove our claim about the behavior of $\mathrm{var}(\delta f, t)$ in a neighborhood of a root c of some $f^{(i)}$. We argue by induction on the degree of f. When f has degree 1, then we are in one of the following two cases, depending upon the sign of f'



In both cases, $\mathrm{var}(\delta f, c-\epsilon) = 1$, but $\mathrm{var}(\delta f, c) = \mathrm{var}(\delta f, c+\epsilon) = 0$, which proves the claim when f is linear.

Now suppose that the degree of f is greater than 1 and let m be the order of vanishing of f at c. We treat first the case when $f(c) = 0$, and hence $m > 0$ so that f' vanishes at c to order $m-1$. We apply our induction hypothesis to f' and obtain that

$$\mathrm{var}(\delta f', c) \ = \ \mathrm{var}(\delta f', c+\epsilon), \qquad \text{and} \qquad \mathrm{var}(\delta f', c-\epsilon) \ \geq \ \mathrm{var}(\delta f', c) \ + \ (m-1)\,,$$

and the difference is even. By Lagrange's Mean Value Theorem applied to the intervals $[c-\epsilon, c]$ and $[c, c+\epsilon]$, f and f' must have opposite signs at $c-\epsilon$, but the same signs at $c+\epsilon$, and so

$$\begin{aligned} \mathrm{var}(\delta f, c) \ &= \ \mathrm{var}(\delta f', c) \ = \ \mathrm{var}(\delta f', c+\epsilon) \ = \ \mathrm{var}(\delta f, c+\epsilon)\,, \\ \mathrm{var}(\delta f, c-\epsilon) \ &= \ \mathrm{var}(\delta f', c-\epsilon) + 1 \ \geq \ \mathrm{var}(\delta f', c) + (m-1) + 1 \ = \ \mathrm{var}(\delta f, c) + m\,, \end{aligned}$$

and the difference is even. This proves the claim when $f(c) = 0$.

Now suppose that $f(c) \neq 0$ so that $m = 0$. Let n be the order of vanishing of f' at c. We apply our induction hypothesis to f' to obtain that

$$\operatorname{var}(\delta f', c) \;=\; \operatorname{var}(\delta f', c+\epsilon), \qquad \text{and} \qquad \operatorname{var}(\delta f', c-\epsilon) \;\geq\; \operatorname{var}(\delta f', c) \;+\; n\,,$$

and the difference is even. We have $f(c) \neq 0$, but $f'(c) = \cdots = f^{(n)}(c) = 0$, and $f^{(n+1)}(c) \neq 0$. Multiplying f by -1 if necessary, we may assume that $f^{(n+1)}(c) > 0$. We consider each of four cases n even or odd, and $f(c)$ positive or negative.

Suppose that n is even. Then both $f'(c-\epsilon)$ and $f'(c+\epsilon)$ are positive and so for each $t \in \{c-\epsilon, c, c+\epsilon\}$ the first nonzero term in the sequence

$$f'(t),\;\; f''(t),\;\; \ldots,\;\; f^{(k)}(t) \tag{2.3}$$

is positive. When $f(c)$ is positive, this implies that $\operatorname{var}(\delta f, t) = \operatorname{var}(\delta f', t)$ and when $f(c)$ is negative, that $\operatorname{var}(\delta f, t) = \operatorname{var}(\delta f', t)+1$. This proves the claim as it implies that $\operatorname{var}(\delta f, c) = \operatorname{var}(\delta f, c+\epsilon)$ and also that

$$\operatorname{var}(\delta f, c-\epsilon) \;-\; \operatorname{var}(\delta f, c) \;=\; \operatorname{var}(\delta f', c-\epsilon) \;-\; \operatorname{var}(\delta f', c)\,,$$

but this last difference exceeds n by an even number, and so is even as n is even.

Now suppose that n is odd. Then $f'(c-\epsilon) < 0 < f'(c+\epsilon)$ and so the first nonzero term in the sequence (2.3) has sign $-,+,+$ as $t = c-\epsilon, c, c+\epsilon$. If $f(c)$ is positive, then $\operatorname{var}(\delta f, c-\epsilon) = \operatorname{var}(\delta f', c-\epsilon) + 1$ and the other two variations are unchanged, but if $f(c)$ is negative, then the variation at $t = c-\epsilon$ is unchanged, but it increases by 1 at $t = c$ and $t = c+\epsilon$. This again implies the claim, as $\operatorname{var}(\delta f, c) = \operatorname{var}(\delta f, c+\epsilon)$, but

$$\operatorname{var}(\delta f, c-\epsilon) - \operatorname{var}(\delta f, c) \;=\; \operatorname{var}(\delta f', c-\epsilon) - \operatorname{var}(\delta f', c) \;\pm\; 1\,.$$

Since the difference $\operatorname{var}(\delta f', c-\epsilon) - \operatorname{var}(\delta f', c)$ is equal to the order n of the vanishing of f' at c plus a nonnegative even number, if we add or subtract 1, the difference is a nonnegative even number. This completes the proof of the Budan-Fourier Theorem.

2.2 Sturm's Theorem

Let f, g be univariate polynomials. Their *Sylvester sequence* is the sequence of polynomials

$$f_0 \;:=\; f,\;\;\; f_1 \;:=\; g,\;\;\; f_2,\;\; \ldots,\;\; f_k\,,$$

where f_k is a greatest common divisor of f and g, and

$$-f_{i+1} \;:=\; \operatorname{remainder}(f_{i-1}, f_i)\,,$$

the usual remainder from the Euclidean algorithm. Note the sign. We remark that we have polynomials $q_1, q_2, \ldots, q_{k-1}$ such that

$$f_{i-1}(t) \;=\; q_i(t) f_i(t) \;-\; f_{i+1}(t)\,, \tag{2.4}$$

and the degree of f_{i+1} is less than the degree of f_i. The *Sturm sequence* of a univariate polynomial f is the Sylvester sequence of f, f'.

Theorem 2.4 (Sturm's Theorem) *Let f be a univariate polynomial and $a, b \in \mathbb{R} \cup \{\pm\infty\}$ with $a < b$ and $f(a), f(b) \neq 0$. Then the number of zeroes of f in the interval (a, b) is the difference*

$$\mathrm{var}(F, a) \quad - \quad \mathrm{var}(F, b)\,,$$

where F is the Sturm sequence of f.

Example 2.5 The sextic f of Example 2.3 has Sturm sequence

$$\begin{aligned}
f &= 5t^6 - 4t^5 - 27t^4 + 55t^2 - 6 \\
f_1 := f'(t) &= 30t^5 - 20t^4 - 108t^3 + 110t \\
f_2 &= \tfrac{84}{9}t^4 + \tfrac{12}{5}t^3 - \tfrac{110}{3}t^2 - \tfrac{22}{9}t + 6 \\
f_3 &= \tfrac{559584}{36125}t^3 + \tfrac{143748}{1445}t^2 - \tfrac{605394}{7225}t - \tfrac{126792}{7225} \\
f_4 &= \tfrac{229905821875}{724847808}t^2 + \tfrac{1540527685625}{4349086848}t + \tfrac{7904908625}{120807968} \\
f_5 &= -\tfrac{280364022223059296}{58526435357253125}t + \tfrac{174201756039315072}{292632176786265625} \\
f_6 &= -\tfrac{17007035533771824564661037625}{162663080627869030112013128}\,.
\end{aligned}$$

Evaluating the Sturm sequence at $t = -0$ gives

$$-6,\ 0,\ 6,\ -\tfrac{126792}{7225},\ \tfrac{174201756039315072}{292632176786265625},\ -\tfrac{17007035533771824564661037625}{162663080627869030112013128}\,,$$

which has 4 variations in sign, while evaluating the Sturm sequence at $t = 2$ gives

$$-26,\ -4,\ \tfrac{1114}{45},\ \tfrac{3210228}{36125},\ -\tfrac{1076053821625}{2174543424},\ -\tfrac{2629438466191277888}{292632176786265625},\ -\tfrac{17007035533771824564661037625}{162663080627869030112013128}\,,$$

which has 2 variations in sign. Thus by Sturm's Theorem, we see that f has 2 roots in the interval $[0, 2]$, which we have already seen by other methods.

An application of Sturm's Theorem is to isolate real solutions to a univariate polynomial f by finding intervals of a desired width that contain a unique root of f. When $(a, b) = (-\infty, \infty)$, Sturm's Theorem gives the total number of real roots of a univariate polynomial. In this way, it leads to an algorithm to investigate the number of real roots of generic systems of polynomials. We briefly describe this algorithm here. This algorithm was used in an essential way to get information on real solutions which helped to formulate many results discussed in later chapters.

Suppose that we have a system of real multivariate polynomials

$$f_1(x_1, \ldots, x_n) \;=\; f_2(x_1, \ldots, x_n) \;=\; \cdots \;=\; f_N(x_1, \ldots, x_n) \;=\; 0\,, \tag{2.5}$$

whose number of real roots we wish to determine. Let $I \subset \mathbb{R}[x_1, \ldots, x_n]$ be the ideal generated by the polynomials $f_1, f_2, \ldots, f_N$. If (2.5) has finitely many complex zeroes, then the dimension of the quotient ring (the *degree* of I) is finite, and for each variable x_i, there is a univariate polynomial $g(x_i) \in I$ of minimal degree, called an *eliminant* for I.

Proposition 2.6 *The roots of $g(x_i) = 0$ form the set of ith coordinates of solutions to* (2.5).

The main tool here is a consequence of the Shape Lemma [4].

Theorem 2.7 (Shape Lemma) *Suppose that I has an eliminant $g(x_i)$ whose degree is equal to the degree of I. Then the number of real solutions to* (2.5) *is equal to the number of real roots of g.*

Suppose that the coefficients coefficients of the polynomials f_i in the system (2.5) lie in a computable subfield of $\mathbb{R}$, for example, $\mathbb{Q}$ (e.g. if the coefficients are integers). Then the degree of I may be computed using Gröbner bases, and we may also use Gröbner bases to compute an eliminant $g(x_i)$. Since Buchberger's algorithm does not enlarge the field of the coefficients, so that $g(x_i) \in \mathbb{Q}[x_i]$ has real coefficients, and so we may use Sturm sequences to compute the number of its real roots.

Algorithm
GIVEN: $I = \langle f_1, \ldots, f_N \rangle \subset \mathbb{Q}[x_1, \ldots, x_n]$

- Use Gröbner bases to compute the degree d of I.
- Use Gröbner bases to compute an eliminant $g(x_i) \in I \cap \mathbb{Q}[x_i]$ for I.
- If $\deg(g) = d$, then use Sturm sequences to compute the number r of real roots of $g(x_i)$, and output "The ideal I has r real solutions."
- Otherwise output "The ideal I does not satisfy the hypotheses of the Shape Lemma."

2.2.1 Traditional Proof of Sturm's Theorem

We prove Sturm's Theorem by looking at the variation $\mathrm{var}(F, t)$ as t increases from a to b. This variation can only change when t passes a number c where some member f_i of the Sturm sequence has a root, for then the sign of f_i could change. We will show that if $i > 0$, then this has no effect on the variation of the sequence, but when c is a root of $f = f_0$, then the variation decreases by exactly 1 as t passes c. Since multiplying a sequence by a nonzero number does not change its variation, we will at times make an assumption on the sign of some value $f_j(c)$ to reduce the number of cases we examine.

Observe first that by (2.4), if $f_i(c) = f_{i+1}(c) = 0$, then f_{i-1} also vanishes at c, as do the other polynomials f_j. In particular $f(c) = f'(c) = 0$, so f has a multiple root at c. Suppose first that this does not happen, either that $f(c) \neq 0$ or that c is a simple root of f.

Suppose that $f_i(c) = 0$ for some $i > 0$. The vanishing of f_i at c, together with (2.4) implies that $f_{i-1}(c)$ and $f_i(c)$ have opposite signs. Then, whatever the sign of $f_i(t)$ for t near c, there is exactly one variation in sign coming from the subsequence $f_{i-1}(t), f_i(t), f_{i+1}(t)$, and so the vanishing of f_i at c has no effect on the variation as t passes c. Note that this argument works equally well for any Sylvester sequence.

Now we consider the effect on the variation when c is a simple root of f. In this case $f'(c) \neq 0$, so we may assume that $f'(c) > 0$. But then $f(t)$ is negative for t to the left of c and positive for t to the right of c. In particular, the variation $\mathrm{var}(F, t)$ decreases by exactly 1 when t passes a simple root of f and does not change when f does not vanish.

We are left with the case when c is a multiple root of f. Suppose that its multiplicity is $m + 1$. Then $(t - c)^m$ divides every polynomial in the Sturm sequence of f. Consider the sequence of quotients,

$$G \;=\; (g_0, \ldots, g_k) \;:=\; (f/(t-c)^m,\; f'/(t-c)^m,\; f_2/(t-c)^m,\; \cdots,\; f_k/(t-c)^m)\;.$$

Note that $\operatorname{var}(G,t) = \operatorname{var}(F,t)$ when $t \neq c$, as multiplying a sequence by a nonzero number does not change its variation. Observe also that G is a Sylvester sequence. Since $g_1(c) \neq 0$, not all polynomials g_i vanish at c. But we showed in this case that there is no contribution to a change in the variation by any polynomial g_i with $i > 0$.

It remains to examine the contribution of g_0 to the variation as t passes c. If we write $f(t) = (t-c)^{m+1}h(t)$ with $h(c) \neq 0$, then

$$f'(t) \;=\; (m+1)(t-c)^m h(t) \;+\; (t-c)^{m+1}h'(t)\,.$$

In particular,

$$g_0(t) \;=\; (t-c)h(t) \qquad \text{and} \qquad g_1(t) \;=\; (m+1)h(t) \;+\; (t-c)h'(t)\,.$$

If we assume that $h(c) > 0$, then $g_1(c) > 0$ and $g_0(t)$ changes from negative to positive as t passes c. Once again we see that the variation $\operatorname{var}(F,t)$ decreases by 1 when t passes a root of f. This completes the proof of Sturm's Theorem.

2.3 A topological proof of Sturm's Theorem

We present a second, very elementary, proof of Sturm's Theorem due to Burda and Khovanskii [16] whose virtue is in its tight connection to topology. We first recall the definition of topological degree of a continuous function $\varphi\colon \mathbb{RP}^1 \to \mathbb{RP}^1$. Since $\mathbb{RP}^1$ is isomorphic to the quotient $\mathbb{R}/\mathbb{Z}$, we may pull φ back to the interval $[0,1]$ to obtain a map $[0,1] \to \mathbb{RP}^1$. This map lifts to the universal cover of $\mathbb{RP}^1$ to obtain a map $\psi\colon [0,1] \to \mathbb{R}$. Then the *mapping degree*, $\operatorname{mdeg}(\varphi)$, of φ is simply $\psi(1) - \psi(0)$, which is an integer. We call this mapping degree to distinguish it from the usual algebraic degree of a polynomial or rational function.

The key ingredient in this proof is a formula to compute the mapping degree of a rational function $\varphi\colon \mathbb{RP}^1 \to \mathbb{RP}^1$. Any rational function $\varphi = f/g$ where $f, g \in \mathbb{R}[t]$ are polynomials has a continued fraction expansion of the form

$$\varphi \;=\; q_0 \;+\; \cfrac{1}{q_1 + \cfrac{1}{q_2 + \cfrac{1}{\ddots + \cfrac{1}{q_k}}}} \tag{2.6}$$

where $q_0, \ldots, q_k$ are polynomials. Indeed, this continued fraction is constructed recursively. If we divide f by g with remainder h, so that $f = q_0 g + h$ with the degree of h less than the degree of g, then

$$\varphi \;=\; q_0 + \frac{h}{g} \;=\; q_0 + \cfrac{1}{\cfrac{g}{h}}\,.$$

We may again divide g by h with remainder, $g = q_1 h + k$ and obtain

$$\varphi \;=\; q_0 + \cfrac{1}{q_1 + \cfrac{1}{\cfrac{h}{k}}}\,.$$

As the degrees of the numerator and denominator drop with each step, this process terminates with an expansion (2.6) of φ.

For example, if $f = 4t^4 - 18t^2 - 6t$ and $g = 4t^3 + 8t^2 - 1$, then

$$\frac{f}{g} \;=\; t - 2 \;+\; \cfrac{1}{-2t + 1 \;+\; \cfrac{1}{-2t - 3 \;+\; \cfrac{1}{t+1}}}$$

This continued fraction expansion is just the Euclidean algorithm in disguise.

Suppose that $q = c_0 + c_1 t + \cdots + c_d t^d$ is a real polynomial of degree d. Define

$$[q] \;:=\; \operatorname{sign}(c_d) \cdot (d \mod 2) \;\in\; \{\pm 1, 0\}\,.$$

Theorem 2.8 *Supose that φ is a rational function with continued fraction expansion* (2.6)*. Then the mapping degree of φ is*

$$[q_1] - [q_2] + \cdots + (-1)^{k-1}[q_k]\,.$$

We may use this to count the roots of a real polynomial f by the following lemma.

Lemma 2.9 *The number of roots of a polynomial f, counted without multiplicities is the mapping degree of the rational function f/f'.*

We deduce Sturm's Theorem from Lemma 2.9. Let $f_0, f_1, f_2 \ldots, f_k$ be the Sturm sequence for f. Then $f_0 = f$, $f_1 = f'$, and for $i > 1$, $-f_{i+1} :=$ remainder(f_{i-1}, f_i). That is, $\deg(f_i) < \deg(f_{i-1})$ and there are univariate polynomials $g_1, g_2, \ldots, g_k$ with

$$f_{i-1} \;=\; g_i f_i - f_{i+1} \qquad \text{for} \quad i = 1, 2, \ldots, k{-}1\,.$$

We relate these polynomials to those obtained from the Euclidean algorithm applied to f, f' and thus to the continued fraction expansion of f/f'. It is clear that the f_i differ only by a sign from the remainders in the Euclidean algorithm. Set $r_0 := f$ and $r_1 = f'$, and for $i > 1$, $r_i :=$ remainder(r_{i-2}, r_{i-1}). Then $\deg(r_i) < \deg(r_{i-1})$, and there are univariate polynomials $q_1, q_2, \ldots, q_k$ with

$$r_{i-i} \;=\; q_i r_i + r_{i+1} \qquad \text{for} \quad i = 1, \ldots, k{-}1\,.$$

We leave the proof of the following lemma as an exercise for the reader.

Lemma 2.10 *We have $g_i = (-1)^{i-1} q_i$ and $f_i = (-1)^{\lfloor \frac{i}{2} \rfloor} r_i$, for $i = 1, 2, \ldots, k$.*

Write F for the Sturm sequence $(f_0, f_1, f_2 \ldots, f_k)$ for f. Then $\mathrm{var}(F, \infty)$ is the variation in the leading coefficients $(f_0^{\mathrm{top}}, f_1^{\mathrm{top}}, \ldots, f_k^{\mathrm{top}})$ of the polynomials in F. Similarly, $\mathrm{var}(F, -\infty)$ is the variation in the sequence

$$((-1)^{\deg(f_0)} f_0^{\mathrm{top}}, (-1)^{\deg(f_1)} f_1^{\mathrm{top}}, \ldots, (-1)^{\deg(f_k)} f_k^{\mathrm{top}})\,.$$

Note that the variation in a sequence $(c_0, c_1, \ldots, c_k)$ is just the sum of the variations in each subsequence (c_{i-1}, c_i) for $i = 1, \ldots, k$. Thus

$$\begin{aligned} &\mathrm{var}(F, -\infty) - \mathrm{var}(F, \infty) \\ &\qquad = \sum_{i=1}^{k} \left(\mathrm{var}((-1)^{\deg(f_{i-1})} f_{i-1}^{\mathrm{top}}, (-1)^{\deg(f_i)} f_i^{\mathrm{top}}) \;-\; \mathrm{var}(f_{i-1}^{\mathrm{top}}, f_i^{\mathrm{top}})\right)\,. \end{aligned} \tag{2.7}$$

Since $f_{i-1} = g_i f_i - f_{i+1}$ and $\deg(f_{i+1}) < \deg(f_i) < \deg(f_{i-1})$, we have

$$f_{i-1}^{\mathrm{top}} \;=\; g_i^{\mathrm{top}} f_i^{\mathrm{top}} \qquad \text{and} \qquad \deg(f_{i-1}) \;=\; \deg(g_i) + \deg(f_i)\,.$$

Thus we have

$$\begin{aligned} \mathrm{var}(f_{i-1}^{\mathrm{top}}, f_i^{\mathrm{top}}) &\;=\; \mathrm{var}(g_i^{\mathrm{top}}, 1)\,, \qquad \text{and} \\ \mathrm{var}((-1)^{\deg(f_{i-1})} f_{i-1}^{\mathrm{top}}, (-1)^{\deg(f_i)} f_i^{\mathrm{top}}) &\;=\; \mathrm{var}((-1)^{\deg(g_i)} g_i^{\mathrm{top}}, 1)\,. \end{aligned}$$

Thus the summands in (2.7) are

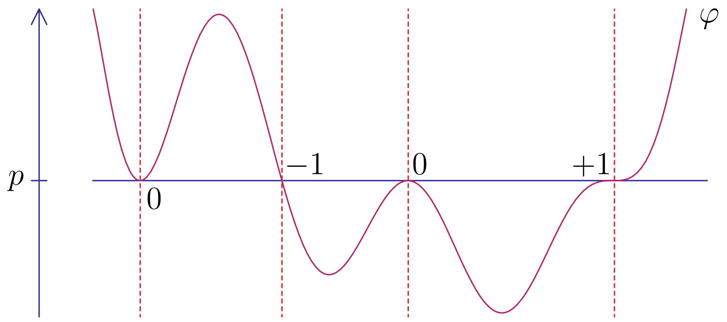
$$\begin{aligned} \mathrm{var}((-1)^{\deg(g_i)} g_i^{\mathrm{top}}, 1) \;-\; \mathrm{var}(g_i^{\mathrm{top}}, 1) &\;=\; \mathrm{sign}\,(g_i^{\mathrm{top}})(\deg(g_i) \mod 2) \\ &\;=\; [g_i] \;=\; (-1)^{i-1}[q_i]\,, \end{aligned}$$

This proves that

$$\begin{aligned} \mathrm{var}(F, -\infty) - \mathrm{var}(F, \infty) &\;=\; [g_1] + [g_2] + \cdots + [g_k] \\ &\;=\; [q_1] - [q_2] + \cdots + (-1)^{k-1}[q_k]\,. \end{aligned}$$

But this proves Sturm's Theorem, as this is the number of roots of f, by Lemma 2.9.

The key to the proof of Lemma 2.9 is an alternative formula for the mapping degree of a continuous function $\varphi\colon \mathbb{RP}^1 \to \mathbb{RP}^1$. Suppose that $p \in \mathbb{RP}^1$ is a point with finitely many inverse images $\varphi^{-1}(p)$. To each inverse image we associate an index that records the behavior of $\varphi(t)$ as t increases past the inverse image. The index is $+1$ if $\varphi(t)$ increases past p, it is -1 if $\varphi(t)$ decreases past p, and it is 0 if φ stays on the same side of p. (Here, increase/decrease are taken with respect to the orientation of $\mathbb{RP}^1$.) For example, here is a graph of a function φ in relation to the value p with the indices of inverse images indicated.



With this definition, the mapping degree of φ is the sum of the indices of the points in a fiber $\varphi^{-1}(p)$, whenever the fiber is finite. That is,

$$\mathrm{mdeg}(\varphi) \;=\; \sum_{a\in\varphi^{-1}(p)} \text{index of } a\,.$$

Proof of Lemma 2.9. The zeroes of the rational function $\varphi := f/f'$ coincide with the zeroes of f. Suppose $f(a)=0$ so that a lies in $\varphi^{-1}(0)$. The lemma will follow once we show that a has index $+1$. Then we may write $f(t) = (t-a)^d h(t)$, where h is a polynomial with $h(a)\neq 0$. We see that $f'(t) = d(t-a)^{d-1}h(t) + (t-a)^d h'(t)$, and so

$$\varphi(t) \;=\; \frac{f(t)}{f'(t)} \;=\; \frac{(t-a)h(t)}{dh(t)+(t-a)h'(t)} \;\approx\; \frac{t-a}{d}\,,$$

the last approximation being valid for t near a as $h(t)\neq 0$. Since d is positive, we see that the index of the point a in the fiber $\varphi^{-1}(0)$ is $+1$.

Proof of Theorem 2.8. Suppose first that φ and ψ are rational functions with no common poles. Then

$$\mathrm{mdeg}(\varphi+\psi) \;=\; \mathrm{mdeg}(\varphi)+\mathrm{mdeg}(\psi)\,.$$

To see this, note that $(\varphi+\psi)^{-1}(\infty)$ is just the union of the sets $\varphi^{-1}(\infty)$ and $\psi^{-1}(\infty)$, and the index of a pole of φ equals the index of the same pole of $(\varphi+\psi)$.

Next, observe that $\mathrm{mdeg}(\varphi) = -\mathrm{mdeg}(1/\varphi)$. For this, consider the behavior of φ and $1/\varphi$ near the level set 1. If $\varphi > 1$ than $1/\varphi < 1$ and vice-versa. The two functions have index 0 at the same points, and opposite index at the remaining points in the fiber $\varphi^{-1}(1) = (1/\varphi)^{-1}(1)$.

Now consider the mapping degree of $\varphi = f/g$ as we construct its continued fraction expansion. At the first step $f = f_0 g + h$, so that $\varphi = f_0 + h/g$. Since f_0 is a polynomial, its only pole is at ∞, but as the degree is h is less than the degree of g, h/g does not have a pole at ∞. Thus the mapping degree of φ is

$$\mathrm{mdeg}(f_0+h/g) \;=\; \mathrm{mdeg}(f_0)+\mathrm{mdeg}\big(\tfrac{h}{g}\big) \;=\; \mathrm{mdeg}(f_0)-\mathrm{mdeg}\big(\tfrac{g}{h}\big)\,.$$

The theorem follows by induction, as $\mathrm{mdeg}(f_0) = [f_0]$.

We close this chapter with an application of this method. Suppose that we are given two polynomials f and g, and we wish to count the zeroes p of f where $g(p) > 0$. If $g = (x-a)(x-b)$ with $a<b$, then this will count the zeroes of f in the interval $[a,b]$, but this question is much more general, and it is not clear how to use the methods in the first two sections of this chapter to solve this problem.

A first step toward solving this problem is to compute the mapping degree of the rational function

$$\varphi \;:=\; \frac{f}{gf'}\,.$$

We consider the indices of its zeroes. First, the zeroes of φ are those zereos of f that are not zeroes of g, together with a zero at infinity if $\deg(g) > 1$. If $f(a)=0$ but $g(a)\neq 0$, then $f = (t-a)^d h(t)$ with $h(a)\neq 0$. For t near a,

$$\varphi(t) \;\approx\; \frac{t-a}{d\cdot g(a)}\,,$$

and so the preimage $a \in \varphi^{-1}(0)$ has index $\operatorname{sign}(g(a))$. If $\deg(g) = e > 1$ and $\deg(f) = d$ then the asymptotic expansion of φ for t near infinity is

$$\varphi(t) \;\approx\; \frac{1}{dg_e t^{e-1}}\,,$$

where g_e is the leading coefficient of g. Thus the index of $\infty \in \varphi^{-1}(0)$ is $\operatorname{sign}(g_e)(e-1 \mod 2) \;=\; [g']$. We summarize this discussion.

Lemma 2.11 *If* $\deg(g) > 1$, *then*

$$\sum_{\{a|f(a)=0\}} \operatorname{sign}(g(a)) \;=\; \operatorname{mdeg}(\varphi) \;-\; [g']\,,$$

and if $\deg(g) = 1$, *the correction term* $-[g']$ *is omitted.*

Since $\operatorname{mdeg}(\varphi) = -\operatorname{mdeg}(1/\varphi)$, we have the alternative expression for this sum.

Lemma 2.12 *Let* $q_1, q_2, \ldots, q_k$ *be the successive quotients in the Euclidean algorithm applied to the division of* $f'g$ *by* f. *Then*

$$\sum_{\{a|f(a)=0\}} \operatorname{sign}(g(a)) \;=\; [q_2] - [q_3] + \cdots + (-1)^k [q_k]\,.$$

Proof. We have

$$\operatorname{mdeg}\frac{f}{f'g} \;=\; -\operatorname{mdeg}\frac{f'g}{f} \;=\; -[q_1] + [q_2] - \cdots + (-1)^k[q_k]\,,$$

by Theorem 2.8. Note that we have $f'g = q_1 f + r_1$. If we suppose that $\deg(f) = d$ and $\deg(g) = e$, then $\deg(q_1) = e - 1$. Also, the leading term of q is dg_e, where g_e is the leading term of g, which shows that $[q_1] = [g']$. Thus the lemma follows from Lemma 2.11, when $\deg(g) \geq 2$.

But it also follows when $\deg(g) < 2$ as $[q_1] = 0$ in that case.

Chapter 3: Sparse Polynomial Systems

find a way to cite the book GKZ

Introduction

Consider a system of n polynomials in n variables

$$f_1(x_1,\dots,x_n) \;=\; f_2(x_1,\dots,x_n) \;=\; \cdots \;=\; f_n(x_1,\dots,x_n) \;=\; 0\,, \tag{3.1}$$

where the polynomial f_i has total degree d_i. By Bézout's Theorem [10], this system has at most $d_1 d_2 \cdots d_n$ isolated complex solutions, and exactly that number if the polynomials are generic among all polynomials with the given degrees.

Polynomials in nature (e.g. from applications) are not necessarily generic—they often have some additional structure which we would like our count of solutions to reflect.

Example 3.1 Consider the system of two polynomials in the variables x and y

$$f \;:=\; x^2y + 2xy^2 + xy - 1 \;=\; 0 \quad\text{and}\quad g \;:=\; x^2y - xy^2 - xy + 2 \;=\; 0\,.$$

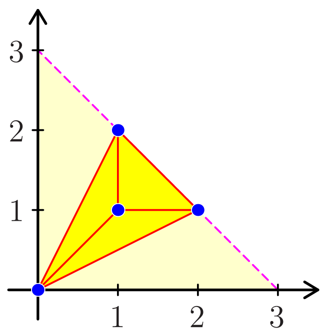
Since these equations have the algebraic consequences,

$$\begin{aligned} f\cdot(y-x+1) + g\cdot(x+2y+1) \;&=\; 3y+3x+1\,, \qquad \text{and}\\ f\cdot(3y^3-3xy^2+5y^2-2xy+2y-3) &\\ +\; g\cdot(6y^3+3xy^2+7y^2+2xy+2y+3) \;&=\; 9y^3+9y^2+2y+9\,, \end{aligned}$$

we see that the system has three solutions. (The linear consequence allows us to recover f and g from the cubic in y.)

Both polynomials f and g have degree three, but they only have three common solutions, which is fewer than the nine predicted by Bézout's Theorem. The key idea behind this deficit of $6 = 9-3$ is illustrated by plotting the exponent vectors of monomials which occur in the polynomials f and g.

$$\begin{array}{rcl} x^2y & \leftrightarrow & (2,1)\\ xy^2 & \leftrightarrow & (1,2)\\ xy & \leftrightarrow & (1,1)\\ 1 & \leftrightarrow & (0,0) \end{array}$$



(3.2)

The *Newton polytope* of f (and of g) is the convex hull of these exponent vectors. This triangle has area $\frac{3}{2}$ and is covered by three lattice triangles. We will see why this number of lattice triangles equals the number of solutions to the original system.

3.1 Kushnirenko's Theorem

The polynomial system in Example 3.1 is a sparse system whose support is the set of integer points in the triangle of (3.2). Let $\mathcal{A} \subset \mathbb{Z}^n$ be a finite set of exponent vectors which affinely spans $\mathbb{R}^n$. A *sparse polynomial* f with *support* $\mathcal{A}$ is a linear combination

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

of monomials from $\mathcal{A}$. Let $\Delta_{\mathcal{A}} \subset \mathbb{R}^n$ be the convex hull of the vectors in $\mathcal{A}$. Let $\mathbb{T} := \mathbb{C}^\times := \mathbb{C} - \{0\}$ be the algebraic torus. We recall Kushnirenko's Theorem from Chapter 1.

Kushnirenko's Theorem. *A system (3.1) of n polynomials in n variables with common support $\mathcal{A}$ has at most $n!\mathrm{vol}(\Delta_{\mathcal{A}})$ isolated solutions in $\mathbb{T}^n$, and exactly this number if the polynomials are generic given their support $\mathcal{A}$.*

We will give two proofs of this result, one using algebraic geometry that is due to Khovanskii and another which is algorithmic. Each proof introduces some important geometry related to sparse systems of polynomials.

It is worth remarking that while sparse polynomials occur naturally—multilinear or multihomogeneous polynomials are an example—they also occur in problem formulations due to Human psychology. Most of us are unable to write down or reason with polynomials having thousands of terms, and we instead seek problem formulations with fewer terms.[†]

3.1.1 The geometry of sparse polynomial systems

Consider the map

$$\varphi_{\mathcal{A}} \;:\; \mathbb{T}^n \ni x \;\longmapsto\; [x^a \mid a \in \mathcal{A}] \;\in\; \mathbb{P}^{\mathcal{A}}\,,$$

where $\mathbb{P}^{\mathcal{A}}$ is the projective space with coordinates $(z_a \mid a \in \mathcal{A})$ indexed by $\mathcal{A}$. This map factors

$$\mathbb{T}^n \;\longrightarrow\; \mathbb{T}^{\mathcal{A}} \;\longrightarrow\!\!\!\!\!\rightarrow\; \mathbb{T}^{\mathcal{A}}/\delta(\mathbb{T}) \;\subset\; \mathbb{P}^{\mathcal{A}}\,,$$

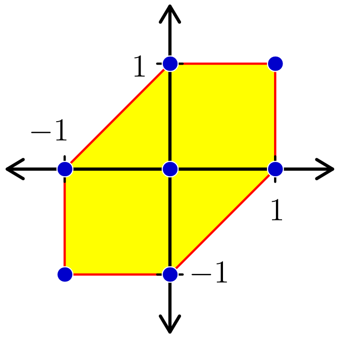
where $\mathbb{T}^{\mathcal{A}} \simeq (\mathbb{C}^\times +^{|\mathcal{A}|}$ is the torus with coordinates indexed by $\mathcal{A}$ and $\delta(\mathbb{T}) \subset \mathbb{T}^{\mathcal{A}}$ is the diagonal torus. The quotient $\mathbb{T}^{\mathcal{A}}/\delta(\mathbb{T})$ is the dense torus in the projective space $\mathbb{P}^{\mathcal{A}}$. It consists of those points $[z_a \mid a \in \mathcal{A}] \in \mathbb{P}^{\mathcal{A}}$ with no coordinate zero. Notice that $\varphi_{\mathcal{A}}$ is a homomorphism into this dense torus. It is often convenient to identify $\mathcal{A}$ with the $n \times |\mathcal{A}|$ matrix whose columns are the exponent vectors in $\mathcal{A}$.

Example 3.2 Suppose that $\mathcal{A}$ consists of the seven exponent vectors $(0,0)$, $(1,0)$, $(0,1)$, $(1,1)$, $(-1,0)$, $(0,-1)$, and $(-1,-1)$. Here is the corresponding matrix,

$$\begin{pmatrix} 0 & 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & -1 \end{pmatrix}.$$

[†] Maybe a modified version of this comment should appear in the introduction?

The convex hull $\Delta_{\mathcal{A}}$ of these points is the hexagon,



and the map $\varphi_{\mathcal{A}}$ is

$$\varphi_{\mathcal{A}} \,:\, (x,y) \in \mathbb{T}^2 \;\longmapsto\; [1, x, y, xy, x^{-1}, y^{-1}, x^{-1}y^{-1}] \in \mathbb{P}^{\mathcal{A}} \;=\; \mathbb{P}^6\,.$$

Consider a linear form Λ on $\mathbb{P}^{\mathcal{A}}$,

$$\Lambda \;=\; \sum_{a\in\mathcal{A}} c_a z_a\,.$$

Its pullback $\varphi^*_{\mathcal{A}}(\Lambda)$ along $\varphi_{\mathcal{A}}$ is a polynomial with support $\mathcal{A}$,

$$\varphi^*_{\mathcal{A}}(\Lambda) \;=\; \sum_{a\in\mathcal{A}} c_a x^a\,.$$

This provides a bijective correspondence between linear forms on $\mathbb{P}^{\mathcal{A}}$ and sparse polynomials with support $\mathcal{A}$. Under $\varphi_{\mathcal{A}}$, the zero set of a sparse polynomial is mapped to a hyperplane section of $\varphi_{\mathcal{A}}(\mathbb{T}^n)$ (the hyperplane is where the corresponding linear form vanishes). Since n general linear forms on $\mathbb{P}^{\mathcal{A}}$ cut out a linear subspace L of codimension n, a general system (3.1) of sparse polynomials with support $\mathcal{A}$ is the pullback along $\varphi_{\mathcal{A}}$ of a codimension n linear section of $\varphi_{\mathcal{A}}(\mathbb{T}^n)$. That is, it equals $\varphi^{-1}_{\mathcal{A}}(L)$ or $\varphi^{-1}_{\mathcal{A}}(\varphi(\mathbb{T}^n)\cap L)$.

The closure of the image of $\varphi_{\mathcal{A}}$ is the *toric variety*[†] $X_{\mathcal{A}}$ parametrized by the monomials $\mathcal{A}$. Since $\varphi_{\mathcal{A}}$ is a homomorphism, the number $d_{\mathcal{A}}$ of solutions to a general sparse system with support $\mathcal{A}$ is the product

$$d_{\mathcal{A}} \;=\; |\ker(\varphi_{\mathcal{A}})|\cdot \deg(X_{\mathcal{A}})\;. \tag{3.3}$$

Indeed, if L is a general linear subspace of $\mathbb{P}^{\mathcal{A}}$ of codimension n, then Bertini's theorem implies that $\varphi_{\mathcal{A}}(\mathbb{T}^n)\cap L = X_{\mathcal{A}}\cap L$ and this intersection is transverse[§]. When the intersection is not transverse, $d_{\mathcal{A}}$ will be the sum of the multiplicities of the solutions. The number of points in such a linear section is the degree $\deg(X_{\mathcal{A}})$ of $X_{\mathcal{A}}$, and each point pulls back under $\varphi_{\mathcal{A}}$ to $|\ker(\varphi_{\mathcal{A}})|$ solutions of the sparse system corresponding to the linear section.

[†]There are competing notions of toric variety. We follow the convention from symbolic computation [102], rather than from algebraic geometry [38]. In particular, we do not assume that $X_{\mathcal{A}}$ is normal.

[§]For those not familiar with Bertini's Theorem, this transversality will follow from the arguments given in our second proof of Kushnirenko's Theorem.

3.1.2 Algebraic-geometric proof of Kushnirenko's Theorem

We prove Kushnirenko's Theorem by showing that

$$n!\cdot \mathrm{vol}(\Delta_{\mathcal{A}}) \;=\; |\ker(\varphi_{\mathcal{A}})|\cdot \deg(X_{\mathcal{A}}) \;=\; d_{\mathcal{A}}\ .$$

This proof is due to Khivanskii ??cite??. The same idea of proof is used in ??Khovanskii—Kuimars???, where they extend the theory of Newton polyhedra for functions on $\mathbb{T}^n$ to convex bodies associated to functions on an arbitrary affine variety.

We first determine the kernel of the map $\varphi_{\mathcal{A}}$, which is the composition

$$\begin{array}{ccccccc} \mathbb{T}^n & \longrightarrow & \mathbb{T}^{\mathcal{A}} & \longrightarrow & \mathbb{T}^{\mathcal{A}}/\delta(\mathbb{T}) & \subset & \mathbb{P}^{\mathcal{A}} \\ x & \longmapsto & (x^a \mid a\in\mathcal{A})\ . & & & & \end{array}$$

To facilitate this computation, we assume that $0\in\mathcal{A}$. This is no loss of generality, for if $0\not\in\mathcal{A}$, then we simply translate $\mathcal{A}$ so that one of its exponent vectors is the origin. This has the effect of multiplying each point in $\varphi_{\mathcal{A}}(\mathbb{T}^n)$ by a scalar, and so it does not change $X_{\mathcal{A}}$. It also multiplies each polynomial in (3.1) by a common monomial, which affects neither the solutions in $\mathbb{T}^n$ nor their number $d_{\mathcal{A}}$. By the relation (3.3), this translation does not change the cardinality of the kernel of $\varphi_{\mathcal{A}}$.

Suppose that $0\in\mathcal{A}$. Then the z_0-coordinate of $\varphi_{\mathcal{A}}$ is constant ($x^0=1$) and so the map which sends $x\in\mathbb{T}^n$ to $(x^a\mid a\in\mathcal{A})$ maps $\mathbb{T}^n$ into $1\times\mathbb{T}^N\subset\mathbb{T}^{\mathcal{A}}$, where $|\mathcal{A}|=N+1$. The composition of the two maps

$$1\times\mathbb{T}^N \;\longrightarrow\; \mathbb{T}^{\mathcal{A}} \;\longrightarrow\; \mathbb{T}^{\mathcal{A}}/\delta(\mathbb{T}) \;\bigl(\subset\; \mathbb{P}^{\mathcal{A}}\bigr)$$

is an isomorphism. Thus it is sufficient to compute the kernel of the map

$$\begin{array}{rccc} \psi_{\mathcal{A}}\,:\; & \mathbb{T}^n & \longrightarrow & \mathbb{T}^{\mathcal{A}}\,, \\ & x & \longmapsto & (x^a\mid a\in\mathcal{A})\,, \end{array}$$

when $0\in\mathcal{A}$.

Let $\mathbb{Z}\mathcal{A}\subset\mathbb{Z}^n$ be the sublattice spanned by the exponent vectors in $\mathcal{A}$. It has full rank, by our assumption that $\mathcal{A}$ affinely spans $\mathbb{R}^n$. Then the quotient $\mathbb{Z}^n/\mathbb{Z}\mathcal{A}$ is a finite abelian group whose order is the lattice index $[\mathbb{Z}^n:\mathbb{Z}\mathcal{A}]$. Its group $\mathrm{Hom}(\mathbb{Z}^n/\mathbb{Z}\mathcal{A},\mathbb{T})$ of characters (homomorphisms to $\mathbb{T}$) is also called its *Pontryagin dual*. We have the sequence of abelian groups,

$$\{0\}\;\longrightarrow\;\mathbb{Z}\mathcal{A}\;\longrightarrow\;\mathbb{Z}^n\;\twoheadrightarrow\;\mathbb{Z}^n/\mathbb{Z}\mathcal{A}\;\longrightarrow\;\{0\}\,.$$

This gives rise to the sequence of Pontryagin duals,

$$0\;\longrightarrow\;\mathrm{Hom}(\mathbb{Z}^n/\mathbb{Z}\mathcal{A},\mathbb{T})\;\longrightarrow\;\mathrm{Hom}(\mathbb{Z}^n,\mathbb{T})\longrightarrow\;\mathrm{Hom}(\mathbb{Z}\mathcal{A},\mathbb{T})\,,$$

which is exact in that the image of $\mathrm{Hom}(\mathbb{Z}^n/\mathbb{Z}\mathcal{A},\mathbb{T})$ in $\mathrm{Hom}(\mathbb{Z}^n,\mathbb{T})$ is the kernel of the last map. If we identify $\mathrm{Hom}(\mathbb{Z}^n,\mathbb{T})$ with $\mathbb{T}^n$, then this sequence of groups, or direct calculation, shows that

$$\mathrm{Hom}(\mathbb{Z}^n/\mathbb{Z}\mathcal{A},\mathbb{T}) \;=\; \{x\in\mathbb{T}^n\mid x^a=1\ \ \forall a\in\mathcal{A}\} \;=\; \ker(\varphi_{\mathcal{A}})\,.$$

Thus $\ker(\varphi_{\mathcal{A}})$ is the Pontryagin dual to the quotient $\mathbb{Z}^n/\mathbb{Z}\mathcal{A}$. Since $\mathbb{Z}^n/\mathbb{Z}\mathcal{A}$ is a finite abelian group, we see that

$$|\ker(\varphi_{\mathcal{A}})| \;=\; [\mathbb{Z}^n : \mathbb{Z}\mathcal{A}] \;=\; |\mathbb{Z}^n/\mathbb{Z}\mathcal{A}|\,. \tag{3.4}$$

The *Hilbert polynomial* $h_X(d)$ a projective variety X is the polynomial which is eventually equal to the dimension of the dth graded piece of the homogeneous coordinate ring $\mathbb{C}[X]$ of X,

$$h_X(d) \;=\; \dim_{\mathbb{C}} \mathbb{C}_d[X]\,, \ \text{ for all } d \text{ sufficiently large.}$$

The Hilbert polynomial contains many numerical invariants of X. For example, the degree of the Hilbert polynomial is the dimension n of X and its leading coefficient is $\deg(X)/n!$. For a discussion of Hilbert polynomials, see Section 9.3 of [21].

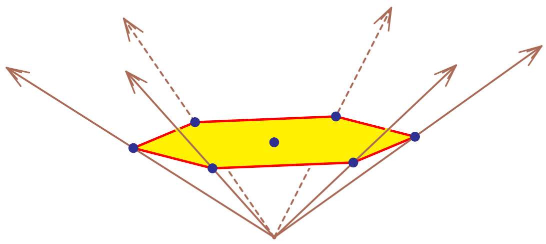
We determine the Hilbert polynomial of the toric variety $X_{\mathcal{A}}$. For this, it is helpful to consider the homogeneous version of the parametrization map $\varphi_{\mathcal{A}}$. We lift $\mathcal{A} \subset \mathbb{Z}^n$ to a homogenized set of exponent vectors $\mathcal{A}^+ \subset 1 \times \mathbb{Z}^n$ by prepending a component of 1 to each vector in $\mathcal{A}$. That is,

$$\mathcal{A}^+ \;:=\; \{(1,a) \mid a \in \mathcal{A}\}\,.$$

The matrix A^+ is obtained from the matrix A by adding a new first row of 1s. For the hexagon of Example 3.2, this is

$$\mathcal{A}^+ \;=\; \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 0 & -1 & -1 & 0 \\ 0 & 0 & 1 & 1 & 0 & -1 & -1 \end{pmatrix}\,.$$

Here is the lifted hexagon, where the first coordinate is vertical.



The map $\varphi_{\mathcal{A}^+}$ on $\mathbb{T}^n$ has the same image in $\mathbb{P}^{\mathcal{A}}$ as does $\varphi_{\mathcal{A}}$. The advantage is that the image of the map $\varphi_{\mathcal{A}^+}$ in $\mathbb{C}^{\mathcal{A}}$ is stable under multiplication by scalars—this is built into it via the new first coordinate of $\mathcal{A}^+$. If we let t be the first coordinate of $\mathbb{T}^{1+n}$, then the pullback of the coordinate ring of $\mathbb{P}^{\mathcal{A}}$ to the ring of Laurent polynomials (the coordinate ring of $\mathbb{T}^{1+n}$) is

$$S_{\mathcal{A}} \;:=\; \mathbb{C}[tx^a \mid a \in \mathcal{A}] \;\simeq\; \mathbb{C}[\mathbb{N}\mathcal{A}^+]\,.$$

This is also the homogeneous coordinate ring of the toric variety $X_{\mathcal{A}} = \overline{\varphi_{\mathcal{A}}(\mathbb{T}^{1+n})}$ (the closure taken in $\mathbb{P}^{\mathcal{A}}$).

The grading on $S_{\mathcal{A}}$ is given by the exponent of the variable t. It follows that the dth graded piece of $S_{\mathcal{A}}$ has a basis of monomials

$$\{t^d \cdot x^a \mid a \in d\mathcal{A}\}\,,$$

where $d\mathcal{A}$ is the set of d-fold sums of vectors in $\mathcal{A}$. This is just the set of integer points in $d\Delta_{\mathcal{A}}$ which lie in the nonnegative integer span $\mathbb{N}\mathcal{A}$ of $\mathcal{A}$. If we let $H_{\mathcal{A}}(d)$ be the dimension of the dth graded piece of the homogeneous coordinate ring of $X_{\mathcal{A}}$ (also called the *Hilbert function* of $X_{\mathcal{A}}$), then these arguments show that

$$H_{\mathcal{A}}(d) \;=\; |d\Delta_{\mathcal{A}} \cap \mathbb{N}\mathcal{A}|.$$

We will estimate this Hilbert function, which will enable us to determine the leading coefficient of the Hilbert polynomial, as the Hilbert function and Hilbert polynomial agreee for d sufficiently large.

Let $\Lambda \simeq \mathbb{Z}^n$ be a lattice in $\mathbb{R}^n$ and Δ be a polytope with vertices in Λ. Ehrhart [27] showed that the counting function

$$P_{\Delta} \,:\, \mathbb{N} \ni d \;\longmapsto\; |d\Delta \cap \Lambda|$$

for the points of Λ contained in positive integer multiples of the polytope Δ is a polynomial in d. This polynomial is called the *Ehrhart polynomial* of the polytope Δ, and its degree is the dimension of the affine span of Δ. When Δ has dimension n, its leading coefficient is the volume of Δ, normalized so that a fundamental domain of the lattice Λ has volume 1. That is, it is the Euclidean volume divided by the lattice index $[\mathbb{Z}^n : \mathbb{Z}\mathcal{A}]$. When $\Lambda = \mathbb{Z}^n$, this is the ordinary Euclidean volume of Δ.

Now suppose that $\Delta = \Delta_{\mathcal{A}}$, the convex hull of $\mathcal{A}$. Then $\Lambda = \mathbb{Z}\mathcal{A}$, and we see that

$$P_{\Delta_{\mathcal{A}}}(d) \;\geq\; H_{\mathcal{A}}(d)\,. \tag{3.5}$$

We give a lower bound for $H_{\mathcal{A}}(d)$. Let $\mathcal{B}$ be the points b in $\mathbb{Z}\mathcal{A}$ which may be written as

$$b \;=\; \sum_{a\in\mathcal{A}} c_a a\,,$$

where c_a is a rational number in $[0, 1)$. Fix an expression for each $b \in \mathcal{B}$ as an integer linear combination of elements of $\mathcal{A}$, and let $-N$ with $N \geq 0$ be an integer lower bound for the coefficients in these expressions for the finitely many elements of $\mathcal{B}$.

For $d \geq N|\mathcal{A}|$, we claim that translation by the vector $N\sum_{a\in\mathcal{A}} a$ defines a map

$$\mathbb{Z}\mathcal{A} \cap (d - N|\mathcal{A}|)\Delta_{\mathcal{A}} \;\longrightarrow\; \mathbb{N}\mathcal{A} \cap d\Delta_{\mathcal{A}}\,.$$

Indeed, a point $v \in \mathbb{Z}\mathcal{A} \cap (d - N|\mathcal{A}|)\Delta_{\mathcal{A}}$ is a nonnegative rational combination of the vectors in $\mathcal{A}$. Taking fractional parts gives $v = b + c$, where $b \in \mathcal{B}$ and $c \in \mathbb{N}\mathcal{A}$. Adding $N\sum_{a\in\mathcal{A}} a$ to the fixed integral expression of b gives a positive integral expression, which proves the claim. This shows that

$$H_{\mathcal{A}}(d) \;\geq\; P_{\Delta_{\mathcal{A}}}(d - N|\mathcal{A}|)\,.$$

If we combine this estimate with (3.5), and use the result that the Hilbert function equals the Hilbert polynomial for d large enough, then we have shown that the Hilbert polynomial $h_{\mathcal{A}}$ of $X_{\mathcal{A}}$ has the same degree and leading coefficient as the Ehrhart polynomial $P_{\Delta_{\mathcal{A}}}$.

Thus the Hilbert polynomial has degree n and its leading coefficient is $\mathrm{vol}(\Delta_{\mathcal{A}})/[\mathbb{Z}^n\!: \mathbb{Z}\mathcal{A}]$, which is the normalized volume of the polytope $\Delta_{\mathcal{A}}$ with respect to the lattice $\mathbb{Z}\mathcal{A}$. We conclude that the degree of $X_{\mathcal{A}}$ is

$$n!\,\frac{\mathrm{vol}(\Delta_{\mathcal{A}})}{[\mathbb{Z}^n\!: \mathbb{Z}\mathcal{A}]}\,.$$

Recall (3.4) that the kernel of $\varphi_{\mathcal{A}}$ has order $[\mathbb{Z}^n\colon \mathbb{Z}\mathcal{A}]$. Then the formula (3.3) for the number $d_{\mathcal{A}}$ of solutions to a sparse system (3.1) with support $\mathcal{A}$ becomes

$$d_{\mathcal{A}} \;=\; |\ker(\varphi_{\mathcal{A}})|\cdot \deg(X_{\mathcal{A}}) \;=\; [\mathbb{Z}^n\colon \mathbb{Z}\mathcal{A}]\cdot n!\,\frac{\mathrm{vol}(\Delta_{\mathcal{A}})}{[\mathbb{Z}^n\colon \mathbb{Z}\mathcal{A}]} \;=\; n!\,\mathrm{vol}(\Delta_{\mathcal{A}})\,,$$

which proves Kushnirenko's Theorem.

3.2 Algorithmic proof of Kushnirenko's Theorem

We present a second proof of Kushnirenko's Theorem whose advantage is that it introduces more geometry connected to toric varieties (which will be useful in subsequent chapters), in particular, the proof uses *toric degenerations* (sometimes called Gröbner degenerations). These toric degenerations underlie the method of Viro [108, 101], which is important for many constructions in real algebraic geometry. We first analyze the simplest case of Kushnirenko's Theorem—when $|\mathcal{A}| = n+1$.

3.2.1 Kushnirenko's Theorem for a simplex

Suppose that $|\mathcal{A}| = n+1$ so that $\Delta_{\mathcal{A}}$ is a simplex with vertices $\mathcal{A}$. Then $X_{\mathcal{A}} = \mathbb{P}^n = \mathbb{P}^{\mathcal{A}}$. As we saw, the solutions to any sparse system (3.1) with support $\mathcal{A}$ have the form $\varphi_{\mathcal{A}}^{-1}(L)$, where $L \subset \mathbb{P}^n$ is the codimension-n plane cut out by the liniear forms which define the polynomials of the system. In this case L is simply a point $\beta \in \mathbb{P}^n$ (which lies in the dense torus as the equations are general), and the solutions have the form $\varphi_{\mathcal{A}}^{-1}(\beta)$. Since $\varphi_{\mathcal{A}}$ is a homomorphism to the dense torus of $\mathbb{P}^{\mathcal{A}}$, these solutions form a single coset of $\ker(\varphi_{\mathcal{A}})$.

We may determine these solutions explicitly. Assume that $0 \in \mathcal{A}$. Since $|\mathcal{A}| = n+1$, we may write the sparse system (3.1) as

$$C\cdot (x^{a_1}, x^{a_2}, \ldots, x^{a_n})^T \;=\; b\,, \tag{3.6}$$

where $\mathcal{A} - \{0\} = \{a_1, a_2, \ldots, a_n\}$, C is the n by n matrix of coefficients, and $b \in \mathbb{C}^n$. If our system is generic, then C is invertible and we may perform row operations on C and hence on the system (3.6) to obtain an equivalent binomial system

$$x^{a_i} \;=\; \beta_i \qquad \text{for } i = 1, \ldots, n\,. \tag{3.6$'$}$$

where $\beta_1, \ldots, \beta_n \in \mathbb{T}$, and so the system has the form $\varphi_{\mathcal{A}}^{-1}(\beta)$. (In fact, the requirements that C be invertible and that the resulting constants $\beta_i \in \mathbb{T}$ are the conditions for genericity of this system.)

Let A be the n by n matrix whose columns are the exponent vectors in $\mathcal{A} - \{0\}$. We will see how the integer linear algebra of the matrix A is used to solve the system (3.6)$'$.

Example 3.3 Consider the system of equations with support $\mathcal{A} = \{(16, 14), (22, 18), (0, 0)\}$

$$\begin{aligned} 23x^{16}y^{14} \;-\; x^{22}y^{18} \;&=\; -27\,, \qquad \text{and} \\ 35x^{16}y^{14} \;-\; x^{22}y^{18} \;&=\; -9\,. \end{aligned} \tag{3.7}$$

Subtracting the two equations gives the binomal,

$$12x^{16}y^{14} \;=\; 36 \qquad \text{or} \qquad x^{16}y^{14} \;=\; 3\,,$$

and solving for $x^{22}y^{18}$ gives $x^{22}y^{18} = 96$. Thus we have the equivalent binomial system

$$x^{16}y^{14} \;=\; 3 \qquad \text{and} \qquad x^{22}y^{18} \;=\; 96\,. \tag{3.8}$$

Under the invertible substitution (the inverse is given by $u = x^8y^7$ and $v = (xy)^{-1}$)

$$x \;=\; uv^7 \qquad \text{and} \qquad y \;=\; u^{-1}v^{-8}\,, \tag{3.9}$$

our equations become triangular

$$\begin{array}{lllllll} (uv^7)^{16}(u^{-1}v^{-8})^{14} & = & u^{16}v^{112}u^{-14}v^{-112} & = & u^2 & = & 3\,, \quad \text{and} \\ (uv^7)^{22}(u^{-1}v^{-8})^{18} & = & u^{22}v^{154}u^{-18}v^{-144} & = & u^4v^{10} & = & 96\,. \end{array} \tag{3.10}$$

While the solution is now immediate via back substitution, we make one further simplifying substitution. Write f for the first equation and g for the second. Replacing g by gf^{-2} yields a diagonal system which is now completely trivial to solve

$$\begin{array}{rcrcl} f & : & u^2 & = & 3\,, \\ gf^{-2} & : & v^{10} & = & 32/3\,. \end{array}$$

The solutions are $u = \pm\sqrt{3}$ and $v = \zeta\sqrt{2}$, where ζ runs over all 10th roots of $\frac{1}{3}$. Substituting these into (3.9), gives the 20 solutions to our original system of equations (3.7).

Underlying these simplifications is the relation between the integer linear algebra of n by n matrices and (multiplicative) coordinate changes in $\mathbb{T}^n$. Since $\mathbb{T}^n = \mathrm{Hom}(\mathbb{Z}^n, \mathbb{T})$, its automorphism group is $GL(n, \mathbb{Z})$, the group of invertible n by n integer matrices, and this is the source of that relation.

Example 3.4 The monomials in (3.8) correspond to the columns of the matrix

$$A \;=\; \begin{pmatrix} 16 & 22 \\ 14 & 18 \end{pmatrix}\,,$$

and the coordinate change (3.9) corresponds to left multiplication (hence row operations) by the matrix

$$\begin{pmatrix} 1 & -1 \\ 7 & -8 \end{pmatrix}\,.$$

Indeed,

$$\begin{pmatrix} 1 & -1 \\ 7 & -8 \end{pmatrix}\begin{pmatrix} 16 & 22 \\ 14 & 18 \end{pmatrix} \;=\; \begin{pmatrix} 2 & 4 \\ 0 & 10 \end{pmatrix}\,,$$

which corresponds to the exponent vectors in the triangular system (3.10). This upper triangular matrix is the *Hermite normal form* of the matrix A—the row reduced echelon form over $\mathbb{Z}$. This notion makes sense for matrices whose entries lie in any principal ideal domain.

Multiplicative reductions using the equations correspond to multiplicative coordinate changes in the target torus $\mathbb{T}^2$ and are represented by column operations, or right multiplication by integer matrices. Indeed

$$\begin{pmatrix} 2 & 4 \\ 0 & 10 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 10 \end{pmatrix},$$

which is the *Smith normal form* of the integer matrix A.

The Smith normal form of an n by n matrix A is the diagonal matrix with entries $d_1, d_2, \ldots, d_n$ where d_i is the greatest common divisor of all i by i subdeterminants (minors) of A. These are called the *invariant factors* of A, and we have $d_1|d_2|\cdots|d_n$. Then

$$\mathbb{Z}^n/\mathbb{Z}\mathcal{A} \;\simeq\; \mathbb{Z}/d_1\mathbb{Z} \times \mathbb{Z}/d_2\mathbb{Z} \times \cdots \times \mathbb{Z}/d_n\mathbb{Z},$$

the canonical way to write a finite abelian group. As with the Hermite normal form, the Smith normal form makes sense for presentations of modules over any principal ideal domain.

3.2.2 Regular triangulations and toric degenerations

Let $\omega\colon \mathcal{A} \to \mathbb{N}$. This is a weight for $\mathbb{T}$ acting diagonally on the space $\mathbb{C}^{\mathcal{A}}$, if $z = (z_a \mid a \in \mathcal{A}) \in \mathbb{C}^{\mathcal{A}}$, then $t.z = (t^{\omega(a)} z_a \mid a \in \mathcal{A})$. This induces a similar action on $\mathbb{P}^{\mathcal{A}}$, and a dual action on the homogeneous coordinate ring $\mathbb{C}[z_a \mid a \in \mathcal{A}]$ of $\mathbb{P}^{\mathcal{A}}$, namely

$$t.z_a \;=\; t^{-\omega(a)} z_a \qquad \text{for a} \in \mathcal{A},$$

where z_a is a variable (coordinate function on $\mathbb{C}^{\mathcal{A}}$). Let $\mathcal{X}_{\mathcal{A}} \subset \mathbb{C} \times \mathbb{P}^{\mathcal{A}}$ be the closure of the family over $\mathbb{T}$ of deformations of the toric variety $X_{\mathcal{A}}$ under this action,

$$\mathcal{X}_{\mathcal{A}} \;:=\; \overline{\{(t,\, t.z) \subset \mathbb{C} \times \mathbb{P}^{\mathcal{A}} \mid t \in \mathbb{T},\; z \in X_{\mathcal{A}}\}}\,.$$

This is a flat family over $\mathbb{C}$ [28, Ch. 15]. (This technical fact implies that all fibers have the same Hilbert polynomial.) The fiber of $\mathcal{X}_{\mathcal{A}}$ over a point $t \in \mathbb{T}$ is the translated toric variety $t.X_{\mathcal{A}}$, and the fiber over $0 \in \mathbb{C}$ is called the *scheme-theoretic limit* of the family $t.X_{\mathcal{A}}$, which is written

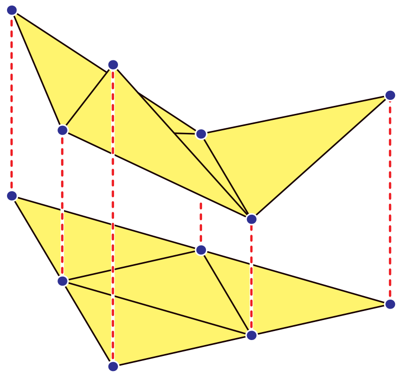
$$\lim_{t \to 0} t.X_{\mathcal{A}}\,.$$

We use geometric combinatorics to study this limit. The passage from $X_{\mathcal{A}}$ to such a scheme-theoretic limit of an action of $\mathbb{T}$ is called a *toric degeneration* or a *Gröbner degeneration*.

We may also use the weight $\omega\colon \mathcal{A} \to \mathbb{N}$ to lift the vector configuration $\mathcal{A}$ into $\mathbb{R}^{1+n}$. Consider the convex hull P_ω in $\mathbb{R}^{1+n}$ of the lifted vectors

$$P_\omega \;:=\; \operatorname{conv}\{(w(a), a) \mid a \in \mathcal{A}\}\,. \tag{3.11}$$

The *lower facets* of this polytope are those whose outward-pointing normal vector has a negative first coordinate. Projecting these lower facets back to $\mathbb{R}^n$ gives the facets in the *regular polyhedral subdivision* Δ_ω of the convex hull of $\mathcal{A}$ given by the lifting function ω. The vertices in this subdivision are some of the vectors in $\mathcal{A}$. While much of the following makes sense for general polyhedral subdivisions, we shall henceforth assume that Δ_ω is a *triangulation* in that each lower facet is a simplex. (This is a mild genericity assumption on ω.) We display the lower facets and the resulting triangulation Δ_ω for a weight function ω on where $\mathcal{A} = \{(0,0),(1,0),(0,1),(2,0),(1,1),(0,2)\}$.



Exponent vectors α of monomials in the coordinate ring $\mathbb{C}[z_a \mid a \in \mathcal{A}]$ of $\mathbb{P}^{\mathcal{A}}$ are elements of $\mathbb{N}^{\mathcal{A}}$, so ω induces a linear form on $\mathbb{N}^{\mathcal{A}}$,

$$\omega \cdot \alpha \ := \ \omega \cdot (\alpha_a \mid a \in \mathcal{A}) \ = \ \sum_{a \in \mathcal{A}} \omega(a)\alpha_a \ .$$

The *initial form* $\mathrm{in}_\omega(g)$ of a homogeneous form $g(z) \in \mathbb{C}[z_a \mid a \in \mathcal{A}]$ is the sum of the terms $c_\alpha z^\alpha$ of g for which $\omega \cdot \alpha$ is maximal among all terms of g. Let $\omega(g)$ be this maximal value. Multiplying $t.g$ by $t^{\omega(g)}$ we see that

$$t^{\omega(g)}(t.g(z)) \ = \ \mathrm{in}_\omega g(z) \ + \ h\,,$$

where h is divisible by t. Thus

$$\lim_{t \to 0} t^{\omega(g)}(t.g(z)) \ = \ \mathrm{in}_\omega g(z)\,.$$

Write $I_{\mathcal{A}}$ for the ideal defining $X_{\mathcal{A}}$. (This *toric* or *lattice ideal* has a linear basis of binomials $z^\alpha - z^\beta$ such that $A^+\alpha = A^+\beta$, where A is the matrix whose columns are the exponent vectors in $\mathcal{A}^+$ [102, Lemma 4.1].) These binomials have the following geometric interpretation. The product $A^+\alpha$ is a positive linear combination of the vectors in $\mathcal{A}^+$, so binomials $z^\alpha - z^\beta$ in $I_{\mathcal{A}}$ record vectors in $\mathbb{N}\mathcal{A}$ that have two distinct representations as positive linear combinations of the vectors in $\mathcal{A}^+$. If we divide by the the initial coordinate of the vector $A^+\alpha$, we obtain a vector of the form $(1,x)$, where x is a rational point which lies in the convex hull of $\mathcal{A}$. When $A^+\alpha = A^+\beta$, this pont x has two distinct rational representations as a convex combination of elements of $\mathcal{A}$. Thus x lies in the convex hulls of two different subsets of $\mathcal{A}$. Conversely, any such point gives rise to a binomial $z^\alpha - z^\beta \in I_{\mathcal{A}}$.

When $\omega \cdot \alpha > \omega \cdot \beta$, then z^α is the initial term $\mathrm{in}_\omega(z^\alpha - z^\beta)$. The *initial ideal* of $I_{\mathcal{A}}$ is

$$\mathrm{in}_\omega(I_{\mathcal{A}}) \ = \ \{\mathrm{in}_\omega(g) \mid g \in I_{\mathcal{A}}\}\,.$$

Since the ideal $I(t.X_\mathcal{A})$ of $t.X_\mathcal{A}$ is

$$I(t.X_\mathcal{A}) \;=\; \{t.g(y) \mid g \in I(X_\mathcal{A})\} \;=\; t.I_\mathcal{A}\,,$$

we see that

$$\lim_{t\to 0} t.I_\mathcal{A} \;=\; \mathrm{in}_\omega(I_\mathcal{A})\,.$$

Thus this initial ideal is the ideal of the scheme-theoretic limit of the family $t.X_\mathcal{A}$.

3.2.3 Kushnirenko's Theorem via toric degenerations

Since the family $\mathcal{X}_\mathcal{A} \to \mathbb{C}$ is flat, every fiber has the same degree, and so Kushnirenko's Theorem follows if we can prove that the degree of the limit scheme $\mathrm{in}_\omega(X_\mathcal{A})$ is $n!\mathrm{vol}(\Delta_\mathcal{A})$ divided by the degree $[\mathbb{Z}^n\colon \mathbb{Z}\mathcal{A}]$ of $\varphi_\mathcal{A}$. It is proved in Chapter 8 of [102] that

$$\sqrt{\mathrm{in}_\omega(I_\mathcal{A})} \;=\; \bigcap_\tau \langle z_a \;\mid\; a \notin \tau\rangle\;, \tag{3.12}$$

the intersection over all facet n-simplices τ of the regular triangulation of $\Delta_\mathcal{A}$. This result is not so hard. If the segment $\overline{ab}$ for $a, b \in \mathcal{A}$ is not a face of the triangulation Δ_ω, then it crosses a minimal face $\mathrm{conv}(\tau)$ of the triangulation. This implies that there is a binomial $z_a^N z_b^M - z^\gamma$, where M, N are positive integers and the monomial z^γ involves the variables in τ. By the construction of the triangulation Δ_ω, the corresponding lifted segment $\overline{(\omega(a), a), (\omega(b), b)}$ of P_ω lies above the lift of the face $\mathrm{conv}(\tau)$, and thus the initial term of this binomial is $z_a^N z_b^M$, and so $z_a z_b$ lies in the radcal $\sqrt{\mathrm{in}_\omega(I_\mathcal{A})}$ of the initial ideal.

It follows that the limit scheme $\mathrm{in}_\omega(X_\mathcal{A})$ is supported on the union of coordinate n-planes $\mathbb{P}^\tau$, one for each facet n-simplex τ of the regular triangulation. (Here $\mathbb{P}^\tau$ is the coordinate plane which is spanned by the coordinates indexed by τ.) The degree of this initial scheme is then

$$\deg(\mathrm{in}_\omega(X_\mathcal{A})) \;=\; \sum_\tau \mathrm{mult}_{\mathbb{P}^\tau}(\mathrm{in}_\omega(X_\mathcal{A}))\;,$$

the sum over these facets τ of the algebraic multiplicity of the limit scheme $\mathrm{in}_\omega(X_\mathcal{A})$ along the coordinate n-plane $\mathbb{P}^\tau$.

In [102, Chapter 8], and under the (mild) assumption that $\mathcal{A}$ is primitive ($\mathbb{Z}\mathcal{A} = \mathbb{Z}^n$), Sturmfels shows that this multiplicity is $n!\mathrm{vol}(\Delta_\tau)$. Since these facets cover $\Delta_\mathcal{A}$, the degree of the limit scheme $\mathrm{in}_\omega(X_\mathcal{A})$ is $n!\mathrm{vol}(\Delta_\mathcal{A})$. As the family $\mathcal{X}_\mathcal{A}$ is flat, this degree is the degree of $X_\mathcal{A}$, and we may deduce Kushnirenko's Theorem from this.

We give a different proof, one which is related to the polyhedral homotopy method [50, 107], is adapted from [86], and does not assume primitivity. Its limitation is that it only works over $\mathbb{C}$ and not an arbitrary algebraically closed field.

The main idea is to fix a general linear subspace L of codimension n in $\mathbb{P}^\mathcal{A}$ and consider the family of linear sections $L \cap t.X_\mathcal{A}$ of fibers of the flat family $\mathcal{X}_\mathcal{A}$, for t near 0. This is illustrated in Figure 3.1. The subspace L will meet each facet n-plane $\mathbb{P}^\tau$ in a single point p_τ, and the points of $L \cap t.X_\mathcal{A}$ for t small will be clustered near the different points p_τ.

We could determine the number of points clustered near one of the p_τ, which is the algebraic multiplicity, $\mathrm{mult}_{\mathbb{P}^\tau}(\mathrm{in}_\omega(X_\mathcal{A}))$. It is in fact easier to determine the number of

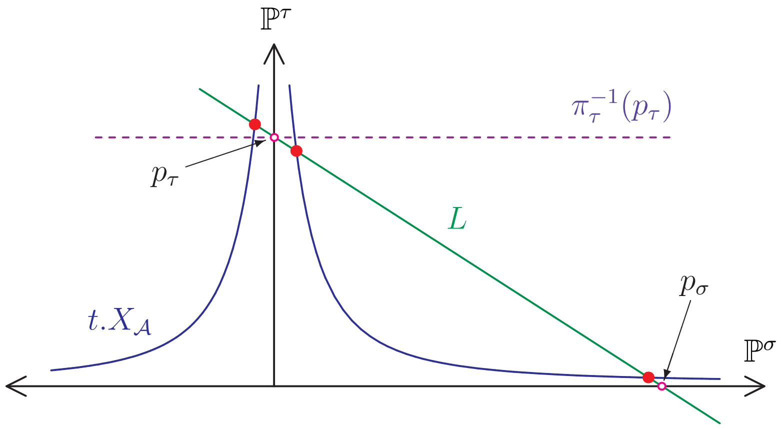


Figure 3.1: Points in $L \cap t.X_{\mathcal{A}}$ near $\mathbb{P}^\tau$ for small t.

points in $\mathbb{T}^n$ of the form $\varphi_{\mathcal{A}}^{-1}(t^{-1}.x)$, for x a point clustered near one of the p_τ and when t is small. This is more direct, and it bypasses computing this algebraic multiplicity. This is also where we avoid the assumption of primitivity, but must work over $\mathbb{C}$.

Note that in a neighborhood of $\mathbb{P}^\tau$, the linear space L is isotopic to $\pi_\tau^{-1}(p_\tau)$, which is a fiber of the coordinate projection $\pi_\tau \colon \mathbb{P}^{\mathcal{A}} - \to \mathbb{P}^\tau$. This is a rational map not defined on the linear span of coordinates $\{z_a \mid a \notin \tau\}$. It follows that the number of points in $\mathbb{T}^n$ coming from points in the linear section $L \cap t.X_{\mathcal{A}}$ near p_τ is equal to the number of points in $\mathbb{T}^n$ coming from points in the linear section $\pi_\tau^{-1}(p_\tau) \cap t.X_{\mathcal{A}}$ near p_τ. This is simply the degree of the map which is the composition of the parametrization $\varphi_{\mathcal{A}}$ of $X_{\mathcal{A}}$, the map $z \mapsto t.z$ on $\mathbb{P}^{\mathcal{A}}$, and this projection π_τ,

$$\mathbb{T}^n \xrightarrow{\varphi_{\mathcal{A}}} X_{\mathcal{A}} \xrightarrow{t} t.X_{\mathcal{A}} \xrightarrow{\pi_\tau} \mathbb{P}^\tau .$$

Since multiplication by t is homotopic to the identity and it commutes with the projection π_τ, we may assume now that $t = 1$, and so this composition is nothing other than the parametrization φ_τ of $\mathbb{P}^\tau$ by the monomials corresponding to integer points of τ. The degree of this map is the order of the kernel of φ_τ, which is $n! \cdot \mathrm{vol}(\Delta_\tau)$. Summing this quantity over all facets τ of the triangulation shows that there are

$$\sum_\tau n! \cdot \mathrm{vol}(\Delta_\tau) \;=\; n!\mathrm{vol}(\Delta_{\mathcal{A}})$$

points in $\mathbb{T}^n$ which are pullbacks under $\varphi_{\mathcal{A}}$ of the linear section $L \cap X_{\mathcal{A}}$. This completes our algorithmic proof of Kushnirenko's Theorem.

The reason that it is algorithmic is that it (more-or-less) counts the solutions to the system $L \cap t.X_{\mathcal{A}}$, for t small, while also giving enough information on their location and structure to determine them numerically so that they may become the input to the polyhedral homotopy method [50, 107] for computing the solutions at $t = 1$.

Note that this proof also shows that the intersection $X_{\mathcal{A}} \cap L$ is transverse when L is general, and thus give a proof of Bertini's Theorem in this context. Since, for t small, the intersection near p_τ may be deformed to the intersection of $X_{\mathcal{A}}$ with the horizontal subspaces

$\pi_\tau^{-1}(p_\tau)$, and this is deformed to the system $\varphi_\tau^{-1}(p_\tau)$, which consists of $n\ \mathrm{vol}(\Delta_\tau)$ distinct points, the general such intersection is transverse.

3.2.4 A brief aside about real solutions

Observe that if the triangulation Δ_ω is *unimodular*, in that each facet has minimal volume $1/n!$, then near each point p_τ there will be exactly one point of $L \cap t.X_{\mathcal{A}}$ and one corresponding solution in $\mathbb{T}^n$. If both L and t are real, then each p_τ and each nearby point in $t.X_{\mathcal{A}}$ will be real. Since

$$t^{-1}.(L \cap t.X_{\mathcal{A}}) \;=\; \big(t^{-1}.L\big) \cap X_{\mathcal{A}}\,,$$

and the points in the left hand side are all real, so are the points in the right hand side. This right hand side corresponds to a system of real polynomials with support $\mathcal{A}$. This proves a theorem of Sturmfels [100], and gives his argument in a nutshell.

Theorem 1.8 *If a lattice polytope $\Delta \subset \mathbb{Z}^n$ admits a regular unimodular triangulation, then there exist real polynomial systems with support $\Delta \cap \mathbb{Z}^n$ having all solutions real.*

A more careful analysis, which begins by examining real solutions when $|\mathcal{A}| = n + 1$, leads to the more refined result for not necessarily unimodular triangulations that appears in Sturmfels's paper.

Chapter 4: Upper Bounds

Reprise: Descartes's Rule

Recall Descartes rule of signs ([23] or Section 2.1), which gives a bound for the number of positive solutions to a univariate polynomial.

Theorem 2.1 (Descartes's rule of signs) *The number, r, of positive roots of a univariate polynomial*

$$f(x) \;=\; c_0x^{a_0} + c_1x^{a_1} + \cdots + c_mx^{a_m}\,, \tag{4.1}$$

counted with multilplicity, is bounded above by the number of variations in sign of the coefficients of f,

$$\#\{i \mid 1 \leq i \leq m \text{ and } c_{i-1}c_i < 0\} \;\leq\; r\,,$$

and the difference between the variation and r is even.

Nota Bene: In (4.1) we assume that $a_0 < a_1 < \cdots < a_m$ and no coefficient c_i vanishes.

Thus a univariate polynomial with $m+1$ monomials has at most m positive roots. This bound is sharp, as the polynomial

$$(x-1)(x-2)\cdots(x-m) \tag{4.2}$$

has $m+1$ distinct terms and m positive roots. Replacing x by x^2 gives a polynomial with m+1 terms and $2m$ nonzero real roots.

This chapter and the next will discuss extensions of this Descartes bound to systems of multivariate polynomials.

4.1 Khovanskii's fewnomial bound

Descartes's rule of signs suggests that the number of real roots to a system of polynomials depends not on its degree, but rather on the complexity of its description. D. Bernstein and A. Kouchnirenko formulated the principle that the topological complexity of a set in $\mathbb{R}^n$ defined by real polynomials is controlled by the complexity of the description of the polynomials, rather than by their degree or Newton polytopes. This is exactly what Khovanskii found in 1980 with his celebrated *fewnomial bound*.

Theorem 4.1 (Khovanskii [59]) *A system of n real polynomials in n variables involving $l+n+1$ distinct monomials will have at most*

$$2^{\binom{l+n}{2}} \cdot (n+1)^{l+n} \tag{4.3}$$

nondegenerate positive solutions.

We remark that nondegenerate solutions are isolated, and there are finitely many of them. This bound, like other bounds in this part of the subject, considers solutions in the *positive orthant* $\mathbb{R}^n_>$. A consequence of Khovanskii's bound is that for each l and n, there is a number $X(l,n)$ which is equal to the maximum number of positive solutions to a system of n polynomials in n variables having $l+n+1$ distinct monomials. A central question in this area is to determine the *Khovanskii number* $X(l,n)$ exactly, or give good bounds. Khovanskii's Theorem shows that $X(l,n)$ is bounded above by the quantity (4.3).

A complete proof of Theorem 4.1 may be found in Khovanskii's book [60], where much else is also developed. Chapter 1 of that book contains an accessible sketch. Benedetti and Risler [5, §4.1] have a careful and self-contained exposition of Khovanskii's fewnomial bound. We give a sketch of the main ideas in the exposition of Benedetti and Risler, to which we refer for further details (this is also faithful to Khovanskii's sketch). We remark that Sturmfels has also sketched ([104, pp. 39–40] and in [103]) a version of the proof. While that is in the spirit of Khovanskii's proof, it is not the same—it omits some contributions to the root count and is therefore regrettably incorrect.

Khovanskii looks for solutions in the positive orthant $\mathbb{R}^n_>$, proving a far more general result involving solutions in $\mathbb{R}^n$ of polynomial functions in logarithms of the coordinates and monomials. For this, he took logarithms of the coordinates. Set

$$z_i \ := \ \log(x_i) \qquad \text{and} \qquad y_j \ := \ e^{z\cdot a_j} \ = \ x^{a_j}\,, \tag{4.4}$$

where $i=1,\ldots,n$, $j=1,\ldots,k$, and $a_j\in\mathbb{R}^n$ can be real exponents. Consider a system of functions of the form

$$F_i(z_1,\ldots,z_n,\, y_1,\ldots,y_k) \ = \ 0 \qquad i=1,\ldots,n\,, \tag{4.5}$$

where each $y_j=y_j(z)$ is an exponential function $e^{z\cdot a_j}$ and the F_i are polynomials in $n{+}k$ indeterminates.

Theorem 4.2 (Khovanskii's Theorem) *The number of nondegenerate real solutions to the system* (4.5) *is at most*

$$\Bigl(\prod_{i=1}^n \deg F_i\Bigr)\cdot\bigl(1+\sum_{i=1}^n \deg F_i\bigr)^k\cdot 2^{\binom{k}{2}}\,. \tag{4.6}$$

Proof of Theorem 4.1. Given a system of n real polynomials in n variables involving $l{+}n{+}1$ distinct monomials, we may assume that one of the monomials is 1. Under the substitution (4.4), this becomes a system of the form (4.5), where each F_i is a degree 1 polynomial in $k=l+n$ variables. Then $\deg F_i=1$ and the bound (4.6) reduces to (4.3).

Sketch of proof of Theorem 4.2. We proceed by induction on k, skipping some technicalities involving Sard's Theorem. When $k=0$, there are no exponentials, and the system is just a system of n polynomials in n variables, whose number of nondegenerate isolated solutions is bounded above by the Bézout number,

$$\prod_{i=1}^n \deg F_i\,,$$

which is the bound (4.6) when $k = 0$.

Suppose that we have the bound (4.6) for systems of the form (4.5) with k exponentials, and consider a system with $k+1$ exponentials and one added variable, t.

$$\begin{array}{rcccl} G_i(z,t) & := & F_i(z_1,\ldots,z_n,\; y_1,\ldots,y_k,\; t\cdot y_{k+1}) & = & 0 \qquad i=1,\ldots,n \\ & & t & = & 1 \end{array} \tag{4.7}$$

The subsystem (4.7) defines an analytic curve C in $\mathbb{R}^{n+1}$, which we assume is smooth and transverse to the hyperplane at $t = 1$.

Write z_{n+1} for t and consider the vector field ξ in $\mathbb{R}^{n+1}$ whose rth component is

$$\xi_r \;:=\; (-1)^{n+1-r}\det\left(\frac{\partial G_i}{\partial z_j}\right)_{j=1,\ldots,\hat{r},\ldots,n+1}^{i=1,\ldots,n} . \tag{4.8}$$

This vector field is tangent to the curve C, and we write $\xi_t = \xi_{n+1}$ for its component in the t-direction. An important ingredient in our proof of Theorem 4.2 is a special case of the Khovanskii-Rolle Theorem [60, pp. 42–51].

Theorem 4.3 (Khovanskii-Rolle Theorem) *The number of points of C where $t = 1$ is bounded above by*

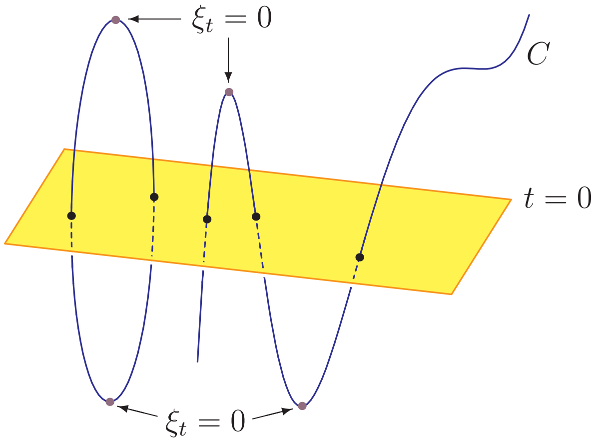
$$N \;+\; q\,,$$

where N is the number of points of C where $\xi_t = 0$ and q is the number of unbounded components of C.

Proof. Note that ξ_t varies continuously along C. Suppose that a and b are consecutive points along an arc of C where $t = 1$. Since C is transverse to the hyperplane $t = 1$, we have $\xi_t(a)\cdot\xi_t(b) < 0$, and so there is a point c of C on the arc between a and b with $\xi_t(c) = 0$.

The hyperplane $t = 1$ cuts a compact connected component of C into the same number of arcs as points where $t = 1$. Since the endpoints of each arc lie on the hyperplane $t = 1$, there is at least one point c on each arc with $\xi_t(c) = 0$. Similarly, the hyperplane $t = 1$ cuts a noncompact component into arcs, and each arc with two endpoints in the hyperplane $t = 1$ gives a point c with $\xi_t(c) = 0$. However, there will be one more point with $t = 1$ on this component than such arcs.

We illustrate the argument in the proof below.



The key to the induction in the proof of Khovanskii's formula is to replace the last exponential by a new variable. This substitution is omitted in Sturmfels's argument, which also does not use the Khovanskii-Rolle Theorem 4.3. Since we have

$$\begin{aligned}
\frac{\partial G_i}{\partial z_r} &= \frac{\partial F_i}{\partial z_r}(z_1, \ldots, z_n,\ y_1, \ldots, y_k, t y_{k+1}) \\
&\quad + \sum_{j=1}^{k} \frac{\partial F_i}{\partial y_j}(z_1, \ldots, z_n,\ y_1, \ldots, y_k, t\, y_{k+1}) \cdot a_{j,r}\, y_j \\
&\quad + \frac{\partial F_i}{\partial y_{k+1}}(z_1, \ldots, z_n,\ y_1, \ldots, y_k, t\, y_{k+1}) \cdot a_{k+1,r}\ t\, y_{k+1}\,,
\end{aligned}$$

if we set $u := t y_{k+1}$ and define $\phi(z, u)$ to be the expression for $\xi_t = \xi_{n+1}$ (4.8) written in terms of z and u, then the total degree (in $z_1, \ldots, z_n, y_1, \ldots, y_k, u$) of $\phi_t(z, u)$ is at most $\sum_{i=1}^{n} \deg F_i$.

Thus number N of Theorem 4.3 is the number of solutions to the system

$$\begin{aligned}
F_i(z_1, \ldots, z_n,\ y_1, \ldots, y_k,\ u) &= 0 \qquad i = 1, \ldots, n \\
\phi_t(z, u) &= 0\,.
\end{aligned} \tag{4.9}$$

This has the form (4.5) with k exponentials. Given any solution to the system (4.9), we use the substitution $u = t y_{k+1}$ to solve for t and get a corresponding point c on the curve C with $\xi_t(c) = 0$. We apply our induction hypothesis to the system (4.9) (which has k exponentials and $n{+}1$ equations in $n{+}1$ variables) to obtain

$$N \ \leq\ \prod_{i=1}^{n} \deg F_i \cdot \Big(\sum_{i=1}^{n} \deg F_i\Big) \cdot \big(1 + 2\sum_{i=1}^{n} \deg F_i\big)^k \cdot 2^{\binom{k}{2}}\,.$$

We similarly estimate the number q of noncompact components of C. We claim that this is bounded above by the maximum number of points of intersection of C with a hyperplane. Indeed, since each noncompact component has 2 infinite branches, there are $2q$ points (counted with multiplicity) on the sphere S^n corresponding to directions of accumulation points of the branches of C at infinity. Any hyperplane through the origin not meeting these points will have at least q of these points in one of the hemispheres into which it divides the sphere. If we translate this hyperplane sufficiently far toward infinity, it will meet the branches giving these accumulation points, and thus will meet C in at least q points.

Thus q is bounded by the number of solutions to a system of the form

$$\begin{aligned}
F_i(z_1, \ldots, z_n,\ y_1, \ldots, y_k,\ u) &= 0 \qquad i = 1, \ldots, n \\
l_0 + l_1 z_1 + l_2 z_2 + \cdots + l_n z_n + l_u u &= 0\,,
\end{aligned} \tag{4.10}$$

where $l_0, \ldots, l_n$, l_u are some real numbers. This again involves only k exponentials, and the last equation has degree 1, so we have

$$q \ \leq\ \prod_{i=1}^{n} \deg F_i \cdot 1 \cdot \big(1 + \sum_{i=1}^{n} \deg F_i\ + 1\big)^k \cdot 2^{\binom{k}{2}}\,.$$

Combining these estimates gives

$$N+q \;\le\; \prod_{i=1}^{n}\deg F_i\cdot 2^{\binom{k}{2}}\cdot\Bigl[\Bigl(\sum_{i=1}^{n}\deg F_i\Bigr)\cdot\bigl(1+2\sum_{i=1}^{n}\deg F_i\bigr)^k \;+\; \bigl(2+\sum_{i=1}^{n}\deg F_i\bigr)^k\Bigr].$$

We can obtain a simpler (but larger) estimate by bounding $N+q$ by the number of solutions to the single system of equations,

$$\begin{array}{rcl} F_i(z_1,\ldots,z_n,\, y_1,\ldots,y_k,\, u) &=& 0 \qquad i=1,\ldots,n\\ F_{n+1} \;:=\; (l_0+l_1z_1+l_2z_2+\cdots+l_nz_n+l_uu)\cdot\phi_t(z,u) &=& 0\,. \end{array}$$

By our induction hypothesis, we have

$$N+q \;\le\; \prod_{i=1}^{n+1}\deg F_i\cdot\bigl(1+\sum_{i=1}^{n+1}\deg F_i\bigr)^k\cdot 2^{\binom{k}{2}}\,.$$

But we saw that F_{n+1} has degree at most $\sum_{i=1}^{n}\deg F_i\; +1$, and so the number, M, of solutions to the system with $k+1$ exponentials is bounded by

$$\begin{aligned} M \;\le\; N+q \;\le\;& \prod_{i=1}^{n}\deg F_i\cdot\bigl(1+\sum_{i=1}^{n}\deg F_i\bigr)\cdot\bigl(1+\sum_{i=1}^{n}\deg F_i+1+\sum_{i=1}^{n}\deg F_i\bigr)^k\cdot 2^{\binom{k}{2}}\\ =\;& \prod_{i=1}^{n}\deg F_i\cdot\bigl(1+\sum_{i=1}^{n}\deg F_i\bigr)^{k+1}\cdot 2^{\binom{k+1}{2}}\,, \end{aligned}$$

whuch completes the proof of Theorem 4.2.

We see that the result of Theorem 4.2 is much more general than the statement of Theorem 4.1. Also, the bound is not sharp. While no one believed that Khovanskii's bound (4.3) was anywhere near the actual upper bound $X(l,n)$, it has been extremely hard to improve upon it. We close this chapter with the first steps in this direction.

4.1.1 Kushnirenko's conjecture

One of the first proposals for a more reasonable bound than Khovanskii's for the number of positive solutions to a system of polynomials due to Kushnirenko, and for many years experts believed that this may indeed be the truth.

Conjecture 4.4 (Kushnirenko) *A system $f_1=f_2=\cdots=f_n=0$ of real polynomials where each f_i has m_i+1 terms has at most $m_1m_2\cdots m_n$ nondegenerate positive solutions.*

This generalizes the bound given by Descartes's rule of signs. It easy to use the example (4.2) for the sharpness of Descartes's rule to construct systems of the form

$$f_1(x_1) \;=\; f_2(x_2) \;=\; \cdots \;=\; f_n(x_n) \;=\; 0\,,$$

which achieve the bound of Conjecture 4.4.

Soon after Kushnirenko made this conjecture, K. Sevostyanov found a counterexample which was unfortunately lost. Nevertheless, it passed into folklore until Haas [46] found a simple example of two trinomials ($3 = 2+1$) in variables x and y with $5\,(> 4 = 2\cdot 2)$ isolated nondegenerate positive solutions.

$$10x^{106} \;+\; 11y^{53} \;-\; 11y \;=\; 10y^{106} \;+\; 11x^{53} \;-\; 11x \;=\; 0\,. \tag{4.11}$$

There have been other attempts to find better bounds than the Khovanskii bound. Sturmfels [100] used the toric degenerations introduced in Chapter 3 to show how to construct systems with many real roots (the root count depends upon a mixture the geometry of the Newton polytopes and some combinatorics of signs associated to lattice points)†. This inspired Itenberg and Roy [53] to propose a multivariate version of Descartes's rule of signs, which was later found to be too optimistic [67]. An interesting part of this story is told in the cheeky paper of Lagarias and Richardson [64].

More recently, Li, Rojas, and Wang looked closely at Haas's counterexample to Kushnirenko's conjecture, seeking to obtain realistic bounds which depended only on the number of monomials in the different polynomials. For example, they showed that Haas's counterexample was the best possible.

Theorem 4.5 (Li, Rojas, and Wang [66]) *A system consisting of two trinomials in two variables has at most* 5 *nondegenerate positive solutions.*

Dickenstein, Rojas, Rusek, and Shih [25] have recently used exact formulas for $\mathcal{A}$-discriminants [24] to study systems of two trinomials in two variables which achieve this bound of 5 positive solutions. They gave the following example, which indicates how difficult it is to find solutios with many real points.

Do Maurice's example in careful detail....

To compare the trinomial bound in [66] to the fewnomial bound (4.3), note that we may multiply one of the polynomials by a monomial so that the two trinomials share a monomial. Then there are at most $5 = 2+2+1$ distinct monomials occurring in the two trinomials. The fewnomial bound for $l = n = 2$ is

$$X(2,2) \;\leq\; 2^{\binom{4}{2}}\cdot(2+1)^4 \;=\; 5184\,.$$

We remark that this system of Li, Rojas, and Wang is not quite a general fewnomial system with $l = n = 2$. Still, their bound of 5 real solutions lended credence to the belief $X(2,2)$ is closer to 5 than to 5184 and that Khovanskii's fewnomial bound (4.3) could be improved.

In addition to providing the counter example to Koushnirenko's conjecture, Sevostyanov also established the first result of fewnomial-type. He showed the existence of an absolute bound $\sigma(d,m)$ for the number of real solutions to a system

$$f(x,y) \;=\; g(x,y) \;=\; 0\,,$$

†Actually, he used the toric proof of Bernstein's Theorem.

where f is a polynomial of degree d and g has m terms. The proof of this result, like his couterexample, has unfortunately been lost. This result however, was the inspiration for Khovanskii to develop his theory of fewnomials.

Recently, Avendaño [1], established a precise version of a special case of Sevostyanov's theorem.

Theorem 4.6 *Suppose that $f(x,y)$ is linear and $g(x,y)$ has m terms. Then the system*

$$f(x,y) \;=\; g(x,y) = \;0\,,$$

has at most $6m-4$ real solutions.

4.2 Systems supported on a circuit

Restricting the analysis of Section 3.2.1 to real solutions shows that $X(0,n)=1$. Recently, it was shown that $X(1,n)=1+n$. We discuss this here.

A collection $\mathcal{A}$ of $n+2$ vectors in $\mathbb{Z}^n$ which affinely spans $\mathbb{R}^n$ is called a *circuit*. The circuit is *primitive* if its $\mathbb{Z}$-affine span is all of $\mathbb{Z}^n$. When $0\in\mathcal{A}$, this means that $\mathbb{Z}^n=\mathbb{Z}\mathcal{A}$.

Theorem 4.7 (Bertrand, Bihan, and Sottile [9]) *A polynomial system supported on a primitive circuit has at most $2n+1$ nondegenerate nonzero real solutions.*

Theorem 4.8 (Bihan [11]) *A polynomial system supported on a circuit has at most $n+1$ nondegenerate positive solutions, and there exist systems supported on a circuit having $n+1$ positive solutions.*

This can be used to construct fewnomial systems with relatively many positive solutions.

Corollary 4.9 (Bihan, Rojas, Sottile [12]) *There exist systems of n polynomials in n variables having $l+n+1$ monomials and $\lfloor\frac{n}{l}\rfloor^l$ positive solutions.*

Describe how this gives asymptotic sharpness, somehwere!

This gives a lower bound for $X(l,n)$ of $\lfloor\frac{n}{l}\rfloor^l$. It remains an open problem to give constructions with more solutions, or constructions with many† solutions when l is not fixed.

The construction establishing Corollary 4.9 is quite simple. Suppose that $n=kl$ is a multiple of l, and let

$$f_1(x_1,\ldots,x_k) \;=\; f_2(x_1,\ldots,x_k) \;=\; \cdots \;=\; f_k(x_1,\ldots,x_k) \;=\; 0\,,$$

be a system with $k+2$ monomials and $k+1$ positive solutions. Such systems exist, by Theorem 4.8. Write $F(x)=0$ for this system and assume that one of its monomials is a constant. For each $i=1,\ldots,l$, let $y^{(i)}=(y_1^{(i)},\ldots,y_k^{(i)})$ be a set of k variables. Then the system

$$F(y^{(1)}) \;=\; F(y^{(2)}) \;=\; \cdots \;=\; F(y^{(l)}) \;=\; 0\,,$$

†Give the easy construction and challenge the reader to improve this

has $(k+1)^l$ solutions, kl variables, and $1+l+kl$ monomials.

When $l=1$ the fewnomial bound (4.3) becomes

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

which is considerably larger than Bihan's bound of $n+1$. Replacing $l+n$ by n in the fewnomial bound, it becomes equal to Bihan's bound when $l=1$. When $l=0$, this same substitution in (4.3) yields 1, which is the sharp bound when $l=0$. In Chapter 5, we will outline generalizations of Theorems 4.7 and 4.8 to arbitrary l, giving the bound

$$X(l,n) \;\leq\; \frac{e^2+3}{4} 2^{\binom{l}{2}} n^l \tag{4.12}$$

for positive solutions and, when $\mathcal{A}$ is primitive, a bound for all real solutions,

$$\frac{e^4+3}{4} 2^{\binom{l}{2}} n^l\,.$$

This is only slightly larger—the difference is in the exponents 2 and 4 of e. These are proven in [3, 14].

Theorems 4.7 and 4.8 are related, and we will outline their proofs, following the papers in which they appear, where more details may be found. To begin, let

$$f_1(x_1,x_2,\ldots,x_n) \;=\; f_2(x_1,x_2,\ldots,x_n) \;=\; \cdots \;=\; f_n(x_1,x_2,\ldots,x_n) \;=\; 0 \tag{4.13}$$

be a system with support $\mathcal{A}$. Suppose that $0\in\mathcal{A}$ and list the elements of the circuit $\mathcal{A}=\{0,a_0,a_1,\ldots,a_n\}$. After a multiplicative change of coordinates (if necessary), we may assume that $a_0=\ell\mathbf{e}_n$, where $\mathbf{e}_n$ is the nth standard basis vector. Since the system (4.13) is generic, row operations on the equations put it into diagonal form

$$x^{a_i} \;=\; w_i \;+\; v_i x_n^{\ell} \qquad \text{for } i=1,\ldots,n\,. \tag{4.14}$$

When $\mathcal{A}$ was a simplex we used integer linear algebra to reduce the equations to a very simple system in Section 3.2.1. We use (different) integer linear algebra to simplify this system supported on a circuit.

4.2.1 Some arithmetic for circuits

Suppose that $\{0,a_0,a_1,\ldots,a_n\}\subset\mathbb{Z}^n$ is a primitive circuit. We assume here that it is nondegenerate—there is no affine dependency involving a subset. (The bounds in the degenerate case are lower, replacing n by the size of this smaller circuit.) After possibly making a coordinate change, we may assume that $a_0=\ell\cdot\mathbf{e}_n$, where $\mathbf{e}_n$ is the nth standard basis vector.

For each i, we may write $a_i=b_i+k_i\cdot\mathbf{e}_n$, where $b_i\in\mathbb{Z}^{n-1}$. Removing common factors from a nontrivial integer linear relation among the n vectors $\{b_1,\ldots,b_n\}\subset\mathbb{Z}^{n-1}$ gives us the *primitive relation* among them (which is well-defined up to multiplication by -1),

$$\sum_{i=1}^{p}\lambda_i b_i \;=\; \sum_{i=p+1}^{n}\lambda_i b_i\,.$$

Here, each $\lambda_i > 0$, and we assume that the vectors are ordered so that the relation has this form. We further assume that

$$N \ := \ \sum_{i=p+1}^{n} \lambda_i k_i \ - \ \sum_{i=1}^{p} \lambda_i k_i \ > \ 0\,.$$

Then we have

$$N\mathbf{e}_n \ + \ \sum_{i=1}^{p} \lambda_i a_i \ - \ \sum_{i=p+1}^{n} \lambda_i a_i \ = \ 0\,,$$

and so

$$x_n^N \cdot \prod_{i=1}^{p} (x^{a_i})^{\lambda_i} \ - \ \prod_{i=p+1}^{n} (x^{a_i})^{\lambda_i} \ = \ 0\,. \tag{4.15}$$

4.2.2 Elimination for circuits

Using (4.14) to substitute for x^{a_i} in (4.15) gives the univariate consequence of (4.14)

$$f(x_n) \ := \ x_n^N \prod_{i=1}^{p} (\alpha_i + \beta_i x_n^{\ell})^{\lambda_i} \ - \ \prod_{i=p+1}^{n} (\alpha_i + \beta_i x_n^{\ell})^{\lambda_i}\ . \tag{4.16}$$

Some further arithmetic of circuits (which may be found in [9]) shows that $f(x_n)$ has degree equal to $n!\mathrm{vol}(\Delta_{\mathcal{A}})$. This is in fact the eliminant of the system.

Lemma 4.10 *The association of a solution x of* (4.14) *to its nth coordinate x_n gives a bijection between the solutions of* (4.14) *and the roots of f* (4.16) *which restricts to a bijection between their real solutions/roots.*

While f is the eliminant of the system, we do not have a Gröbner basis or even a triangular system to witness this fact, and the proof proceeds by explicitly constructing a solution (4.14) from a root x_n of f.

The upper bound is found by writing $f = F - G$ as in (4.16) and then perturbing f,

$$f_t(y) \ = \ t \cdot F(y) \ - \ G(y)\,.$$

We simply estimate the number of changes in the the real roots of f_t as t passes from $-\infty$ to 1, which can occur only at the singular roots of f_t. While similar to the proof of Khovanskii's theorem, this is not inductive, but relies on the form of the Wronskian $F'G - G'F$ whose roots are the singular roots of f_t. This may also be seen as an application of Rolle's Theorem. We note that this estimation also uses Viro's construction for t near 0 and ∞.

Lastly, (for special circuits) polynomials f are constructed with $2n+1$ real solutions in [9] using Viro's methods for univariate polynomials. Bihan [11] constructs a system with $n+1$ using Grothendieck's *dessins d'enfants*.

4.2.3 A family of systems with a sharp bound

We give a family of systems that illustrate the result of Theorem 4.7 (actually of an extension of it) and which may be treated by hand. These systems come from a family of polytopes $\Delta \subset \mathbb{Z}^n$ for which we prove a nontrivial upper bound on the number of real solutions to polynomial systems with primitive support $\mathcal{A} := \Delta \cap \mathbb{Z}^n$. That is, the integer points $\mathcal{A}$ in Δ affinely span $\mathbb{Z}^n$, so that general systems supported on Δ have $n!\mathrm{vol}(\Delta)$ complex solutions, but there are fewer than $n!\mathrm{vol}(\Delta)$ real solutions to polynomial systems with support $\mathcal{A}$. This is intended to not only give a glimpse of the more general results in [9], but also possible extensions which are not treated in [14].

Let $l > k > 0$ and $n \geq 3$ be integers and $\epsilon = (\epsilon_1, \ldots, \epsilon_{n-1}) \in \{0,1\}^{n-1}$ have at least one nonzero coordinate. The polytope $\Delta^{\epsilon}_{k,l} \subset \mathbb{R}^n$ is the convex hull of the points

$$(0,\ldots,0),\ (1,0,\ldots,0),\ldots,(0,\ldots,0,1,0),\ (0,\ldots,0,k),\ (\epsilon_1,\ldots,\epsilon_{n-1},l)\,.$$

The configuration $\mathcal{A}^{\epsilon}_{k,l} = \Delta^{\epsilon}_{k,l} \cap \mathbb{Z}^n$ also includes the points along the last axis

$$(0,\ldots,0,1),\ (0,\ldots,0,2),\ \ldots,\ (0,\ldots,0,k{-}1)\,.$$

Since these points include the standard basis and the origin, $\mathcal{A}^{\epsilon}_{k,l}$ is primitive in that $\mathbb{Z}\mathcal{A} = \mathbb{Z}^n$.

Set $|\epsilon| := \sum_i \epsilon_i$. Then the volume of $\Delta^{\epsilon}_{k,l}$ is $l + k|\epsilon|$. Indeed, the configuration $\mathcal{A}^{\epsilon}_{k,l}$ can be triangulated into two simplices $\Delta^{\epsilon}_{k,l} \setminus \{(\epsilon_1,\ldots,\epsilon_{n-1},l)\}$ and $\Delta^{\epsilon}_{k,l} \setminus \{0\}$ with volumes k and $l - k + k|\epsilon|$, respectively. One way to see this is to apply the affine transformation

$$(x_1,\ldots,x_n) \ \longmapsto\ (x_1,\ldots,x_{n-1},x_n - k + k\sum_{i=1}^{n-1} x_i)\,.$$

Theorem 4.11 *The number, r, of real solutions to a generic system of n real polynomials with support $\mathcal{A}^{\epsilon}_{k,l}$ lies in the interval*

$$0 \ \leq\ r\ \leq\ k + k|\epsilon| + 2\,,$$

and every number in this interval with the same parity as $l + k|\epsilon|$ occurs.

This upper bound does not depend on l and, since $k < l$, it is smaller than or equal to the number $l + k|\epsilon|$ of complex solutions. We use elimination to prove this result.

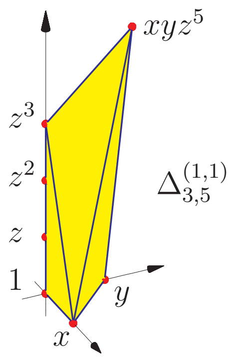
Example 4.12 Suppose that $n = k = 3$, $l = 5$, and $\epsilon = (1,1)$.

Then the system

$$\begin{aligned} x + y + xyz^5 + 1 + z + z^2 + z^3 &= 0\\ x + 2y + 3xyz^5 + 5 + 7z + 11z^2 + 13z^3 &= 0\\ 2x + 2y + xyz^5 + 4 + 8z + 16z^2 + 32z^3 &= 0 \end{aligned}$$

is equivalent to

$$\begin{aligned} x - (5 + 11z + 23z^2 + 41z^3) &= 0\\ y + (8 + 18z + 38z^2 + 72z^3) &= 0\\ xyz^5 - (2 + 6z + 14z^2 + 30z^3) &= 0 \end{aligned}$$



And thus its number of real roots equals the number of real roots of

$$z^5(5+11z+23z^2+41z^3)(8+18z+38z^2+72z^3)-(2+6z+14z^2+30z^3)\,,$$

which, as we invite the reader to check, is 3.

Proof of Theorem 4.11. A generic real polynomial system with support $\mathcal{A}^{\epsilon}_{k,l}$ has the form

$$\sum_{j=1}^{n-1} c_{ij}x_j \; + c_{in}x^{\epsilon}x_n^l \;\; + \;\; f_i(x_n) \;\; = \;\; 0 \qquad \text{for } i=1,\ldots,n\,,$$

where each polynomial f_i has degree k and x^{ϵ} is the monomial $x_1^{\epsilon_1}\cdots x_{n-1}^{\epsilon_{n-1}}$.

Since all solutions to our system are simple, we may perturb the coefficient matrix $(c_{ij})_{i,j=1}^n$ if necessary and then use Gaussian elimination to obtain an equivalent system

$$x_1 - g_1(x_n) \;\; = \;\; \cdots \;\; = \;\; x_{n-1} - g_{n-1}(x_n) \;\; = \;\; x^{\epsilon}x_n^l - g_n(x_n) \;\; = \;\; 0\,, \tag{4.17}$$

where each polynomial g_i has degree k. Using the first $n-1$ polynomials to eliminate the variables $x_1,\ldots,x_{n-1}$ gives the univariate polynomial

$$x_n^l \cdot g_1(x_n)^{\epsilon_1}\cdots g_{n-1}(x_n)^{\epsilon_{n-1}} \;\; - \;\; g_n(x_n)\,, \tag{4.18}$$

which has degree $l+k|\epsilon| = v(\Delta^{\epsilon}_{k,l})$. Any zero of this polynomial leads to a solution of the original system (4.17) by back substitution. This implies that the number of real roots of the polynomial (4.18) is equal to the number of real solutions to our original system (4.17).

The eliminant (4.18) has no terms of degree m for $k < m < l$, and so it has at most $k+k|\epsilon|+2$ nonzero real roots, by Descartes's rule of signs. This proves the upper bound.

We omit the construction which shows that this bound is sharp.

Chapter 5: Fewnomial upper bounds from Gale dual polynomial systems

Suppose that we have the system of polynomials,

$$\begin{aligned}
v^2w^3 - 11uvw^3 - 33uv^2w + 4v^2w + 15u^2v + 7 &= 0\,,\\
v^2w^3 \qquad\qquad + \ 5uv^2w - 4v^2w - \ 3u^2v + 1 &= 0\,,\\
v^2w^3 - 11uvw^3 - 31uv^2w + 2v^2w + 13u^2v + 8 &= 0\,.
\end{aligned} \tag{5.1}$$

If we solve them for the monomials v^2w^3, v^2w, and uvw^3, we obtain

$$\begin{aligned}
v^2w^3 &= 1 - u^2v - uv^2w\,,\\
v^2w &= \tfrac{1}{2} - u^2v + uv^2w\,,\\
uvw^3 &= \tfrac{10}{11}(1 + u^2v - 3uv^2w)\,.
\end{aligned} \tag{5.2}$$

Since

$$\begin{aligned}
\big(uv^2w\big)^3 \cdot \big(v^2w^3\big) &= u^3v^8w^6 = \big(u^2v\big)\cdot\big(v^2w\big)^3\cdot\big(uvw^3\big) \qquad \text{and}\\
\big(u^2v\big)^2 \cdot \big(v^2w^3\big)^3 &= u^4v^8w^9 = \big(uv^2w\big)^2\cdot\big(v^2w\big)\cdot\big(uvw^3\big)^2\,,
\end{aligned}$$

we may substitute the expressions on the right hand sides of (5.2) for the monomials v^2w^3, v^2w, and uvw^3 in these expressions to obtain the system

$$\begin{aligned}
\big(uv^2w\big)^3\cdot\big(1-u^2v-uv^2w\big) &= \big(u^2v\big)\cdot\big(\tfrac{1}{2}-u^2v+uv^2w\big)^3\cdot\big(\tfrac{10}{11}(1+u^2v-3uv^2w)\big) \qquad \text{and}\\
\big(u^2v\big)^2\cdot\big(1-u^2v-uv^2w\big)^3 &= \big(uv^2w\big)^2\cdot\big(\tfrac{1}{2}-u^2v+uv^2w\big)\cdot\big(\tfrac{10}{11}(1+u^2v-3uv^2w)\big)^2\,.
\end{aligned}$$

Writing x for u^2v and y for uv^2w and solving for 0, these become

$$\begin{aligned}
f &:= \ y^3(1-x-y) - x(\tfrac{1}{2}-x+y)^3\big(\tfrac{10}{11}(1+x-3y)\big) = 0\,, \quad \text{and}\\
g &:= x^2(1-x-y)^3 - y^2(\tfrac{1}{2}-x+y)\big(\tfrac{10}{11}(1+x-3y)\big)^2 = 0\,.
\end{aligned} \tag{5.3}$$

Figure 5.1 shows the curves these define and the lines given by the linear factors in (5.3).

It is clear that the solutions to (5.3) in the complement of the lines are consequences of solutions to (5.1). More, however, it true. The two systems define isomorphic schemes as complex or as real varieties, with the positive solutions to (5.1) corresponding to the solutions of (5.3) lying in the central pentagon. Gale duality, which generalizes this isomorphism, is a first step towards the new fewnomial bounds of [3, 14].

We remark that these new bounds are derived using the general method that Khovanskii developed in [59]. However, they take advantage of special geometry (encoded in Gale duality) available to systems of polynomials in a way that the proof of Khovanski's bound (Theorem 4.1) did not. Their main value is that they are sharp, in one asymptotic sense described after Corollary 4.9.

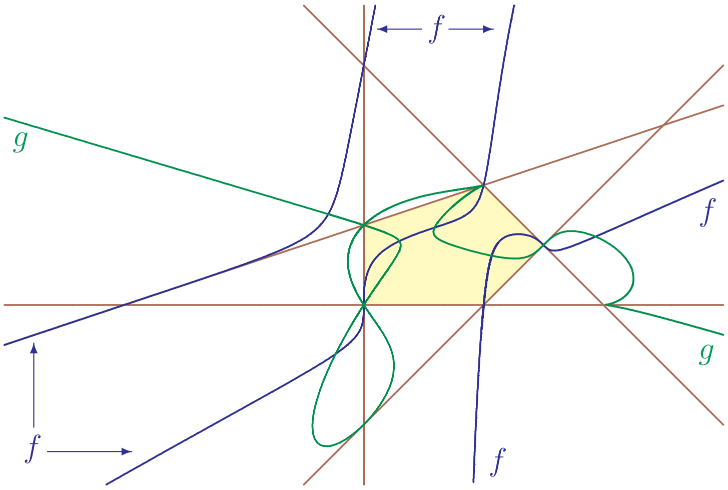


Figure 5.1: Curves and lines

5.1 Gale duality for polynomial systems

Gale duality is another way to view a sparse system of polynomials. It was developed in [13] in more generality than we will treat here. Let us work over the complex numbers. Let $\mathcal{A} = \{0, a_1, \ldots, a_{l+n}\} \subset \mathbb{Z}^n$ be integer vectors which span $\mathbb{R}^n$ and suppose that we have a system

$$f_1(x_n, \ldots, x_n) \;=\; f_2(x_1, \ldots, x_n) \;=\; \cdots \;=\; f_n(x_1, \ldots, x_n) \;=\; 0 \tag{5.4}$$

of polynomials with support $\mathcal{A}$. As in Section 3.1.1, the solutions may be interpreted geometrically as $\varphi_{\mathcal{A}}^{-1}(L)$, where $\varphi_{\mathcal{A}}$ is the map

$$\begin{array}{rcccl} \varphi_{\mathcal{A}} : \mathbb{T}^n & \longrightarrow & \mathbb{T}^{l+n} & \subset & \mathbb{C}^{l+n} \\ x & \longmapsto & (x^{a_1}, x^{a_2}, \ldots, x^{a_{l+n}})\,, & & \end{array}$$

and $L \subset \mathbb{C}^{l+n}$ is a codimension n plane defined by degree 1 polynomials $\Lambda_1, \ldots, \Lambda_n$ corresponding to the polynomials f_i. While we work here in $\mathbb{C}^{n+l}$, but used $\mathbb{P}^{\mathcal{A}} = \mathbb{P}^{n+l}$ in Section 3.1.1, there is no essential difference.

Suppose that $\mathcal{A}$ is primitive in that $\mathbb{Z}\mathcal{A} = \mathbb{Z}^n$, so that the homomorphism $\varphi_{\mathcal{A}}$ is injective. Then the subscheme of $\mathbb{T}^n$ defined by (5.4) is isomorphic to the subscheme $X := \varphi(\mathbb{T}^n) \cap L$ of $\mathbb{T}^{l+n}$ or $\mathbb{C}^{l+n}$. If we change our perspective and view X as the basic object, then the parametrization $\varphi_{\mathcal{A}}$ of $\varphi_{\mathcal{A}}(\mathbb{T}^n)$ realizes X (or rather $\varphi_{\mathcal{A}}^{-1}(X)$) as the subscheme of $\mathbb{T}^n$ defined by (5.4).

The main idea behind Gale duality for polynomial systems is to instead parameterize L with a map $\psi_p \colon \mathbb{C}^l \to L$ and then consider the subscheme $\psi_p^{-1}(X)$ of $\mathbb{C}^l$, which is isomorphic to X. This is in fact what we did in transforming (5.1) into (5.3). We will show that $\psi_p^{-1}(X)$ is defined in $\mathbb{C}^l$ by a system of master functions, which we define in the next section.

5.1.1 Master functions

Let $p_1(y), \ldots, p_{l+n}(y)$ be pairwise non-proportional degree 1 polynomials on $\mathbb{C}^l$. Their product $\prod_i p_i(y) = 0$ defines a hyperplane arrangement $\mathcal{H}$ in $\mathbb{C}^l$. Let $\beta \in \mathbb{Z}^{l+n}$ be an integer vector, called a *weight* for the arrangement $\mathcal{H}$. We use this to define a rational function p^β,

$$p^\beta \;=\; p(y)^\beta \;:=\; p_1(y)^{b_1}p_2(y)^{b_2}\cdots p_{l+n}(y)^{b_{l+n}}\,,$$

where $\beta=(b_1,\ldots,b_{l+n})$. This rational function $p(y)^\beta$ is a *master function* for the arrangement $\mathcal{H}$. As the components of β may be negative, its natural domain of definition is the complement $M_{\mathcal{H}}:=\mathbb{C}^l\setminus\mathcal{H}$ of the arrangement.

A *system of master functions* in $M_{\mathcal{H}}$ with weights $\mathcal{B}=\{\beta_1,\ldots,\beta_l\}$ is the system of equations in $M_{\mathcal{H}}$,

$$p(y)^{\beta_1} \;=\; p(y)^{\beta_2} \;=\; \cdots \;=\; p(y)^{\beta_l} \;=\; 1\,. \tag{5.5}$$

More generally, we could instead consider equations of the form $p(y)^\beta=\alpha$, where $\alpha\in\mathbb{T}$ is an arbitrary nonzero complex number. We may however absorb such constants into the polynomials $p_i(y)$, as there are $l{+}n$ such polynomials but only l constants in a system of master functions. This is the source of the factor $\frac{10}{11}$ in (5.3). We further assume that the system (5.5) defines a zero-dimensional scheme in $M_{\mathcal{H}}$. This implies that the weights $\mathcal{B}$ are linearly independent, and that the suppressed constants multiplying the $p_i(y)$ are sufficiently general.

As with sparse systems, a system of master functions may be realized geometrically through an appropriate map. The polynomials $p_1(y),\ldots,p_{l+n}(y)$ define an affine map

$$\begin{array}{rccl}\psi_p\,: & \mathbb{C}^l & \longrightarrow & \mathbb{C}^{l+n}\\ & y & \longmapsto & (p_1(y),p_2(y),\ldots,p_{l+n}(y))\,.\end{array}$$

This map is injective if and only if the polynomials $\{1,p_1(y),\ldots,p_{l+n}(y)\}$ span the space of degree 1 polynomials on $\mathbb{C}^l$, in which case the hyperplane arrangement $\mathcal{H}$ is called essential. The hyperplane arrangement $\mathcal{H}$ is the pullback along ψ_p of the coordinate hyperplanes $z_i=0$ in $\mathbb{C}^{l+n}$, and its complement $M_{\mathcal{H}}$ is the pullback of the torus $\mathbb{T}^{l+n}$ which is the complement of the coordinate hyperplanes in $\mathbb{C}^{l+n}$.

The weights $\mathcal{B}$ are *saturated* if they span a saturated subgroup of $\mathbb{Z}^{l+m}$, that is, if $\mathbb{Z}\mathcal{B}=\mathbb{Q}\mathcal{B}\cap\mathbb{Z}^{l+n}$. Linear independence of $\mathcal{B}$ is equivalent to the subgroup $\mathbb{G}$ of the torus $\mathbb{T}^{l+n}$ defined by the equations

$$z^{\beta_1} \;=\; z^{\beta_2} \;=\; \cdots \;=\; z^{\beta_l} \;=\; 1 \tag{5.6}$$

having dimension n and saturation is equivalent to $\mathbb{G}$ being connected. (Here, $z_1,\ldots,z_{l+n}$ are the coordinates for $\mathbb{C}^{l+n}$.) In this way, we see that the equations (5.5) describe the pullback $\psi_p^{-1}(\mathbb{G})$ of this subgroup $\mathbb{G}$. We summarize this discussion.

Proposition 5.1 *A system of master functions* (5.5) *in $M_{\mathcal{H}}$ is the pullback along ψ_p of the intersection of the linear space $\psi_p(\mathbb{C}^l)$ with a subgroup $\mathbb{G}$ of $\mathbb{T}^{l+n}$ of dimension n, and any such pullback defines a system of master functions in $M_{\mathcal{H}}$. When ψ_p is injective, it gives a scheme-theoretic isomorphism between the solutions of the system of master functions and the intersection $\mathbb{G}\cap\psi_p(\mathbb{C}^l)$.*

5.1.2 Gale duality

Proposition 5.1 is the new ingredient needed for the notion of Gale duality. Suppose that $\mathbb{G}\subset\mathbb{T}^{l+n}$ is a connected subgroup of dimension n and that $L\subset\mathbb{C}^{l+n}$ is a linear subspace of dimension l not parallel to any coordinate hyperplane. Then the intersection $\mathbb{G}\cap L$ has dimension 0.

Definition 5.2 Suppose that we are given

1. An isomorphism $\varphi_{\mathcal{A}}\colon \mathbb{T}^n \to \mathbb{G}$ for $\mathcal{A}=\{0,a_1,\dots,a_{l+n}\}$ and equations (5.6) defining $\mathbb{G}$ as a subgroup of $\mathbb{T}^{l+n}$. Necessarily $\mathcal{A}$ is primitive and $\mathcal{B}=\{\beta_1,\dots,\beta_l\}$ is saturated.

2. A linear isomorphism $\psi_p\colon \mathbb{C}^l \to L$ and degree 1 polynomials $\Lambda_1,\dots,\Lambda_n$ on $\mathbb{C}^{l+n}$ defining L.

Let $\mathcal{H}\subset\mathbb{C}^l$ be the pullback of the coordinate hyperplanes of $\mathbb{C}^{l+n}$ along ψ_p. We say that the system of sparse polynomials on $\mathbb{T}^n$

$$\varphi^*(\Lambda_1) \;=\; \varphi^*(\Lambda_2) \;=\; \cdots \;=\; \varphi^*(\Lambda_n) \;=\; 0 \tag{5.7}$$

with support $\mathcal{A}$ is *Gale dual* to the system of master functions

$$p(y)^{\beta_1} \;=\; p(y)^{\beta_2} \;=\; \cdots \;=\; p(y)^{\beta_l} \;=\; 1 \tag{5.8}$$

with weights $\mathcal{B}$ on the hyperplane complement $M_{\mathcal{H}}$ and vice-versa.

The following is immediate.

Theorem 5.3 *A pair of Gale dual systems* (5.7) *and* (5.8) *define isomorphic schemes.*

This notion of Gale duality involves two different linear algebraic dualities in the sense of linear functions annihilating vector spaces. In the first, the degree 1 polynomials $p_i(y)$ defining the map ψ_p are annihilated by the degree 1 polynomials Λ_i which define the system of sparse polynomials (5.7). The second is integral, as the weights $\mathcal{B}$ form a basis for the free abelian group of integer linear relations among the nonzero elements of $\mathcal{A}$. Writing the elements of $\mathcal{B}$ as the rows of a matrix, the $l+n$ columns form the *Gale dual* or Gale transform [45, §5.4] of the vector configuration $\mathcal{A}$—this is the source of our terminology.

If we restrict the domain of $\varphi_{\mathcal{A}}$ to the real numbers or to the positive reals, then we obtain the two forms of Gale duality which are relevant to us. Set $\mathbb{T}_{\mathbb{R}} := \mathbb{R}\setminus\{0\}$, the real torus and $\mathbb{R}_>$ to be the positive real numbers. Suppose that $\mathcal{A}$ is not necessarily primitive, but that the lattice index $[\mathbb{Z}^n\colon \mathbb{Z}\mathcal{A}]$ is odd. Then $\varphi_{\mathcal{A}}\colon \mathbb{T}^n_{\mathbb{R}} \to \mathbb{T}^{l+n}_{\mathbb{R}}$ is injective. Similarly, if $\mathbb{Z}\mathcal{B}$ has odd index in its saturation $\mathbb{Q}\mathcal{B}\cap\mathbb{Z}^{l+n}$, which is the group of integer linear relations holding on $\mathcal{A}$, then the equations (5.6) define a not necessarily connected subgroup $\mathbb{G}\subset\mathbb{T}^{l+n}$ whose real points $\mathbb{G}_{\mathbb{R}}$ lie in its connected component containing the identity. When the linear polynomials Λ_i of (5.7) and p_i of (5.8) are real and annihilate each other, then these two systems in—which do not necessarily define isomorphic schemes in $\mathbb{T}^n$ and $M_{\mathcal{H}}$—define isomorphic real analytic sets in $\mathbb{T}^n_{\mathbb{R}}$ for (5.7) and in the complement $M^{\mathbb{R}}_{\mathcal{H}} := \mathbb{R}^l\setminus\mathcal{H}_{\mathbb{R}}$ of the real hyperplanes defined by the p_i for (5.8).

In the version valid for the positive real numbers, we may suppose that the exponents $\mathcal{A}$ are real vectors, for if $r\in\mathbb{R}_>$ and $a\in\mathbb{R}$, then $r^a := \exp(a\log(x))$ is well-defined. In this case, the weights $\mathcal{B}$ should be a basis for the vector space of linear relations holding on $\mathcal{A}$, and the degree 1 polynomials Λ_i and p_i are again real and dual to each other. The equations 5.6 for $z \in \mathbb{R}^{l+n}_{>}$ define a connected subgroup of $\mathbb{R}^{l+n}_{>}$ which equals $\varphi_{\mathcal{A}}(\mathbb{R}^n_{>})$. In this generality, the polynomial system (5.7) makes sense only for $x \in \mathbb{R}^n_{>}$ and the system of master functions (5.8) only makes sense for y in the *positive chamber* Δ_p of the hyperplane complement $M^{\mathbb{R}}_{\mathcal{H}}$,

$$\Delta_p \ := \ \{y \in \mathbb{R}^l \mid p_i(y) > 0 \quad i = 1, \ldots, l{+}n\}\,,$$

and the two systems define isomorphic real analytic sets in $\mathbb{R}^n_{>}$ for (5.7) and in Δ_p for (5.8).

5.1.3 Algebra of Gale duality

The description of Gale duality in Definition 5.2 lends itself immediately to an algorithm for converting a system of sparse polynomials into an equivalent system of master functions. We describe this over $\mathbb{C}$, but it works equally well over $\mathbb{R}$ or over $\mathbb{R}_{>}$. Suppose that $\mathcal{A} \subset \mathbb{Z}^n$ is a primitive collection of integer vectors and suppose that

$$f_1(x_1, \ldots, x_n) \ = \ f_2(x_1, \ldots, x_n) \ = \ \cdots \ = \ f_n(x_1, \ldots, x_n) \ = \ 0 \qquad (5.9)$$

defines a zero dimensional subscheme of $\mathbb{T}^n$. In particular, the polynomials f_i are linearly independent. We may solve these equations for some of the monomials to obtain

$$\begin{array}{rcccl} x^{a_1} & = & g_1(x) & =: & p_1(x^{a_{n+1}}, \ldots, x^{a_{l+m+n}}) \\ \vdots & & & & \vdots \\ x^{a_n} & = & g_n(x) & =: & p_n(x^{a_{n+1}}, \ldots, x^{a_{l+m+n}}) \end{array} \qquad (5.10)$$

Here, for each $i = 1, \ldots, n$, $g_i(x)$ is a polynomial with support $\{0, a_{n+1}, \ldots, a_{l+n}\}$ which is a degree 1 polynomial function $p_i(x^{a_{n+1}}, \ldots, x^{a_{l+n}})$ in the given l arguments. For $i = n{+}1, \ldots, l{+}n$, set $p_i(x^{a_{n+1}}, \ldots, x^{a_{l+n}}) := x^{a_i}$.

An integer linear relation among the exponent vectors in $\mathcal{A}$,

$$b_1 a_1 \ + \ b_2 a_2 \ + \ \cdots \ + \ b_{l+n} a_{l+n} \ = \ 0\,,$$

is equivalent to the monomial identity

$$(x^{a_1})^{b_1} \cdot (x^{a_2})^{b_2} \cdots (x^{a_{l+n}})^{b_{l+n}} \ = \ 1\,,$$

which gives the consequence of the system (5.10)

$$\big(p_1(x^{a_{n+1}}, \ldots, x^{a_{l+n}})\big)^{b_1} \ \cdots \ \big(p_{l+n}(x^{a_{n+1}}, \ldots, x^{a_{l+n}})\big)^{b_{l+n}} \ = \ 1\,.$$

Define $y_1, \ldots, y_l$ to be new variables which are coordinates for $\mathbb{C}^l$. The degree 1 polynomials $p_i(y_1, \ldots, y_l)$ define a hyperplane arrangement $\mathcal{H}$ in $\mathbb{C}^l$. Let $\mathcal{B} := \{\beta_1, \ldots, \beta_l\} \subset \mathbb{Z}^{l+n}$ be a basis for the $\mathbb{Z}$-module of integer linear relations among the nonzero vectors in $\mathcal{A}$. These weights $\mathcal{B}$ define a system of master functions

$$p(y)^{\beta_1} \ = \ p(y)^{\beta_2} \ = \ \cdots \ = \ p(y)^{\beta_l} \ = \ 1 \qquad (5.11)$$

in the complement $M_{\mathcal{H}}$ of the arrangement.

Theorem 5.4 *The system of polynomials* (5.9) *in* $\mathbb{T}^n$ *and the system of master functions* (5.11) *in* $M_{\mathcal{H}}$ *define isomorphic schemes.*

Proof. Condition (1) in Definition 5.2 holds as $\mathcal{A}$ and $\mathcal{B}$ are both primitive and annihilate each other. The linear forms Λ_i that pull back along $\varphi_{\mathcal{A}}$ to define the system (5.10) are

$$\Lambda_i(z) \;=\; z_i \;-\; p_i(z_{n+1},\ldots,z_{l+n})\,,$$

which shows that condition (2) holds, and so the statement follows from Theorem 5.3.

The example at the beginning of this chapter illustrated Gale duality. However, the equations (5.3) are not of the form $p^\beta = 1$. They are, however, easily transformed into such equations, and we obtain

$$\frac{y^3(1-x-y)}{x(\frac{1}{2}-x+y)^3\left(\frac{10}{11}(1+x-3y)\right)} \;=\; \frac{x^2(1-x-y)^3}{y^2(\frac{1}{2}-x+y)\left(\frac{10}{11}(1+x-3y)\right)^2} \;=\; 1\,. \tag{5.12}$$

Systems of the form (5.3) may be obtained from systems of master functions by multiplying $p^\beta = 1$ by the terms of p^β with negative exponents to clear the denominators, to obtain $p^{\beta^+} = p^{\beta^-}$ and thus $p^{\beta^+} - p^{\beta^-} = 0$, where $\beta_\pm$ is the componentwise maximum of the pair of vectors $(0, \pm\beta)$.

5.2 New fewnomial bounds

Theorem 5.5 *A system* (5.4) *of* n *polynomials in* n *variables having a total of* $l+n+1$ *monomials with exponents* $\mathcal{A} \subset \mathbb{R}^n$ *has at most*

$$\frac{e^2+3}{4}\,2^{\binom{l}{2}}n^l$$

positive nondegenerate solutions.

If $\mathcal{A} \subset \mathbb{Z}^n$ *and* $\mathbb{Z}\mathcal{A}$ *has odd index in* $\mathbb{Z}^n$*, then the system has at most*

$$\frac{e^4+3}{4}\,2^{\binom{l}{2}}n^l$$

nondegenerate real solutions.

The first bound is proven in [14] and the second in [3]. By Gale duality, Theorem 5.5 is equivalent to the following.

Theorem 5.6 *Let* $p_1(y),\ldots,p_{l+n}(y)$ *be degree* 1 *polynomials on* $\mathbb{R}^l$ *that, with* 1*, span the space of degree* 1 *polynomials. For any linearly independent vectors* $\mathcal{B} = \{\beta_1,\ldots,\beta_l\} \subset \mathbb{R}^{l+n}$*, the number of solutions to*

$$1 \;=\; p(y)^{\beta_j} \qquad \text{for} \quad j = 1,\ldots,l$$

in the positive chamber Δ_p *is less than*

$$\frac{e^2+3}{4}\,2^{\binom{l}{2}}n^l\,.$$

If $\mathcal{B} \subset \mathbb{Z}^{l+n}$ has odd index in its satiration, than the number of solutions in $M_{\mathcal{H}}^{\mathbb{R}}$ is less than

$$\frac{e^{4}+3}{4}\, 2^{\binom{l}{2}} n^{l}\,.$$

We outline the proof of Theorem 5.6 in the next three sections.

5.2.1 One idea

The basic idea behind the proof of Theorem 5.6 is to use the Khovanskii-Rolle Theorem, but in a slightly different form than given in Theorem 4.3. Using it in this way to establish bounds for real solutions was first done in [40]. Given functions $g_1, \ldots, g_m$ defined on a domain D, let $V(g_1, \ldots, g_m)$ be their set of common zeroes. If C is a curve in D, let $\mathrm{ubc}(C)$ be its number of unbounded components in D.

Theorem 5.7 (Khovanskii-Rolle) *Let $g_1, \ldots, g_l$ be smooth functions defined on a domain $D \subset \mathbb{R}^l$ with finitely many common zeroes and suppose that $C := V(g_1, \ldots, g_{l-1})$ is a smooth curve. Set J to be the Jacobian determinant, $\det(\partial g_i/\partial y_j)$, of $g_1, \ldots, g_l$. Then we have*

$$|V(g_1, \ldots, g_l)| \;\leq\; \mathrm{ubc}(C) + |V(g_1, \ldots, g_{l-1},\, J)|\,. \tag{5.13}$$

Explain this better!

This Khovanskii-Rolle Theorem follows from the observation that along any arc of C connecting two zeroes of g_l, there must be a point in that arc where C is tangent to a level set of g_l. At that point, the normal dg_l to the level set is orthogonal to C. Since the normal directions to C are spanned by $dg_1, \ldots, dg_{l-1}$, this means that the Jacobian J must vanish at that point. Figure 5.2 illustrates the Khovanskii-Rolle Theorem when $k = 2$.

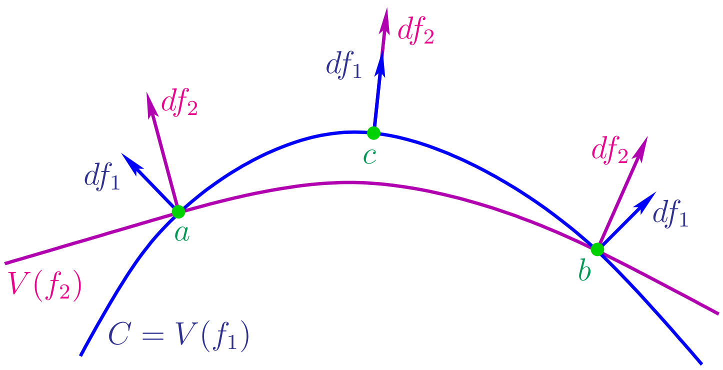


Figure 5.2: $df_1 \wedge df_2(a) < 0$, $\quad df_1 \wedge df_2(c) = 0$, $\quad$ and $df_1 \wedge df_2(b) > 0$.

As in the proof of Theorem 4.3, the estimate of (5.13) follows from the observation concerning consecutive zeroes of g_l along C.

5.2.2 A generalization, and two reductions

We first make an adjustment to the system of master functions in Theorem 5.6, replacing each master function $p(y)^\beta = p_1(y)^{b_1} \cdots p_{l+n}(y)^{b_{l+n}}$ by

$$|p(y)|^\beta \;:=\; |p_1(y)|^{b_1} \cdots |p_{l+n}(y)|^{b_{l+n}}\,.$$

For example, if we take absolute values in the system of master functions (5.12), we obtain

$$\frac{|y|^3|1-x-y|}{|x||\frac{1}{2}-x+y|^3|\frac{10}{11}(1+x-3y)|} \;=\; \frac{|x|^2|1-x-y|^3}{|y|^2|\frac{1}{2}-x+y||\frac{10}{11}(1+x-3y)|^2} \;=\; 1\,. \tag{5.14}$$

This new system with absolute values will still have the same number of solutions in the positive chamber Δ_p (since $|p_i(y)| = p_i(y)$ for $i = 1, \ldots, l{+}n$ and $y \in \Delta_p$) as the original system. Its solutions in the hyperplane complement $M_{\mathcal{H}}^{\mathbb{R}}$ will include the solutions to the system of master functions from Theorem 5.6.

We illustrate this for the system (5.14) in Figure 5.3, which we may compare to Figure 5.1 as the system of master functions (5.12) is equivalent to the system (5.3) in the complement

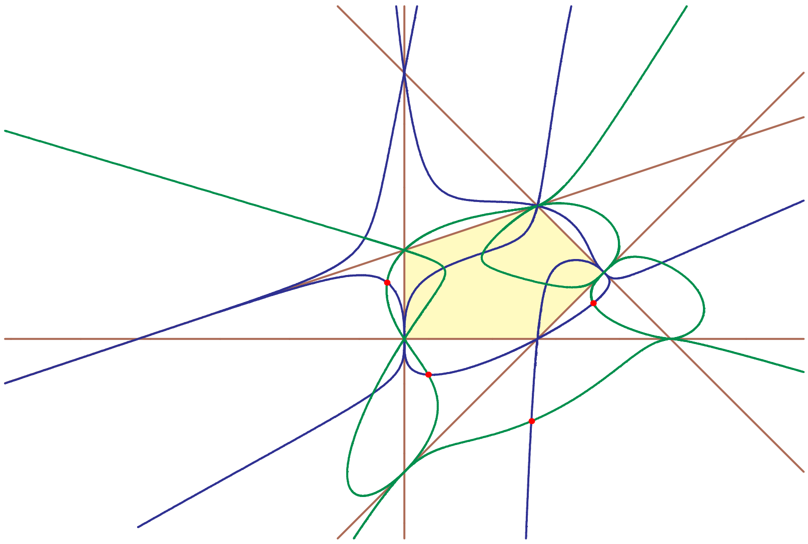


Figure 5.3: Curves from absolute values

of the lines. In particular, among the solutions to the system (5.14) are the three solutions to (5.3) in the positive chamber (which is shaded) as well as the three solutions to (5.3) outside the positive chamber. The system (5.14) has four additional solutions outside the positive chamber, which are marked in Figure 5.3.

We give a proof of Theorem 5.6 for systems of the form $|p(y)|^{\beta_j} = 1$ for $j = 1, \ldots, l$. This will imply the bound for systems of master functions. Taking absolute values allows nonintegral (real number) exponents in $|p(y)|^{\beta_j} = 1$, and so we neet not require that exponents are integral.

We promised two reductions.

1. The degree 1 polynomials $p_i(y)$ are in general position in that the hyperplanes in the arrangement $\mathcal{H}$ are in linear general position. That is, any j of them meet in a linear subspace of codimension j, if $j \leq l$, and their intersection is empty if $j > l$. We may do this, as we are bounding nondegenerate solutions, which cannot be destroyed if the $p_i(y)$ are perturbed to put the hyperplanes into this general position.

2. Let B be the matrix whose rows are $\beta_1, \ldots, \beta_l$. We may assume that every minor of B is non-zero. This may be done by perturbing the real-number exponents in the functions $|p(y)|^{\beta_j}$. This will not reduce the number of nondegenerate solutions.

   Perturbing exponents is not as drastic of a measure as it first seems. Note that in the hyperplane complement, $|p(y)|^\beta = 1$ defines the same set as $\log(|p(y)|^\beta) = 0$. If $\beta = (b_1, \ldots, b_{l+n})$, then this is simply

$$b_1 \log|p_1(y)| \;+\; b_2 \log|p_2(y)| \;+\; \cdots \;+\; b_{l+n} \log|p_{l+n}(y)| \;=\; 0\,. \tag{5.15}$$

   Expressing the equations in this form shows that we may perturb the exponents.

We first look at these reductions in the context of the system of master functions (5.12). The hyperplane arrangement $\mathcal{H}$ is an arrangement of lines in which no three meet and no two are parallel, and thus they are in general position. The matrix of exponents is

$$B \;=\; \begin{pmatrix} 2 & -2 & 3 & -1 & -2 \\ 1 & -3 & -1 & 3 & 1 \end{pmatrix}.$$

No entry and no minor of B vanishes.

Let now see how that Khovanskii-Rolle Theorem applies to the system (5.14) of Figure 5.3, restricted to the positive chamber. First, take logarithms and rearrange to obtain

$$\begin{aligned} 2\log|x| - 2\log|y| + 3\log|1{-}x{-}y| - \log|\tfrac{1}{2}{-}x{+}y| - 2\log|\tfrac{10}{11}(1{+}x{-}3y)| &= 0 \\ \log|x| - 3\log|y| + 1\log|1{-}x{-}y| + 3\log|\tfrac{1}{2}{-}x{+}y| + \log|\tfrac{10}{11}(1{+}x{-}3y)| &= 0 \end{aligned}$$

Call these functions f_1 and f_2, respectively. Their Jacobian is the rational function

$$\frac{2x^3 - 16x^2y + 12xy^2 + 6y^3 - \frac{31}{2}x^2 + 26xy - \frac{53}{2}y^2 + \frac{9}{2}x + \frac{15}{2}y - 2}{xy(1-x-y)(\frac{1}{2}{-}x{+}y)(1+x-3y)}$$

whose denominator is the product of the linear factors defining the lines in Figure 5.3. Clearing the denominator and multiplying by 2, we obtain a cubic polynomial

$$J_2 \;:=\; 4x^3 - 32x^2y + 24y^2x + 12y^3 - 31x^2 + 52xy - 53y^2 + 9x + 15y - 4\,.$$

This meets the curve C_1 (which is defined by $f_1 = 0$) in 6 points, five of which we display in Figure 5.4—the sixth is at $(3.69, -0.77)$. By the Khovanskii-Rolle Theorem, the number of solutions to $f_1 = f_2 = 0$ is at most this intersection number, $|V(f_1, J_2)|$, plus the number of unbounded components of C_1.

While we can see that C_1 has 14 unbounded components, a more universal bound can be found. Observe that C_1 meets the arrangement of lines only in some of its vertices, and at each such vertex it has four branches, one for each chamber incident to the vertex. This is due to the form of the master function defining C_1. In general

This gives the inequality

$$10 \;=\; |V(f_1, f_2)| \;\leq\; \mathrm{ubc}(C_1) + |V(f_1, J_2)| \;=\; 14 + 6 \;=\; 18\,.$$

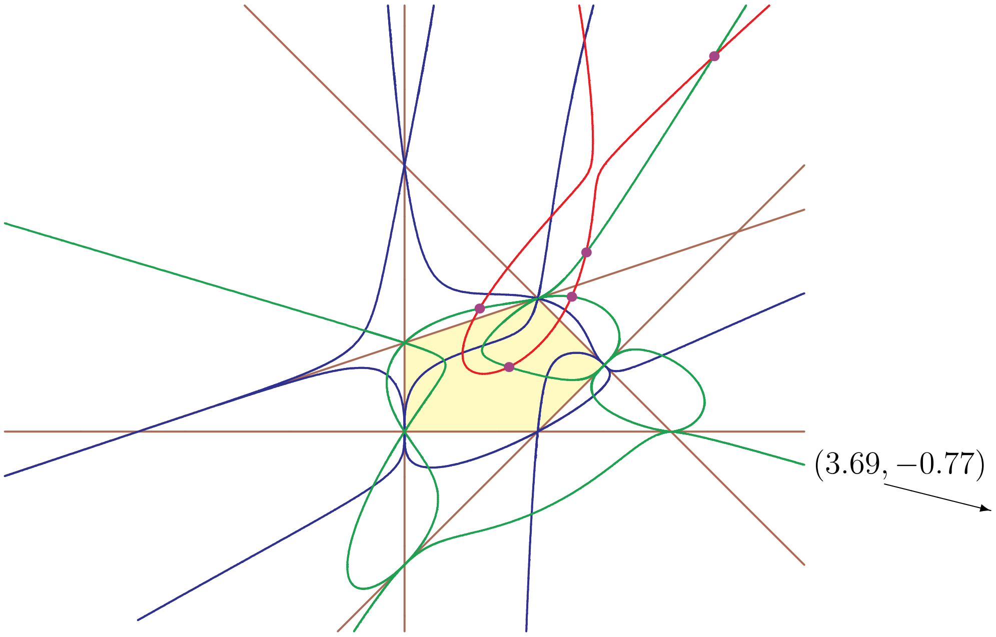


Figure 5.4: Gale system and Jacbian J_2.

5.2.3 Three lemmata

We follow the suggestion in the second reduction in Section 5.2.2 and replace the master functions by the logarithms of their absolute values. For each $j = 1, \dots, l$, define

$$g_j(y) \; := \; \log |p(y)|^{\beta_j} \; = \; \sum_{i=1}^{l+n} \beta_{i,j} \log |p_i(y)| \; ,$$

where $\beta_{i,j}$ is the ith component of β_j. Observe that both $f_j = 0$ and $|p(y)|^{\beta_j} = 1$ have the same solutions in the hyperplane complement $M^{\mathbb{R}}_{\mathcal{H}}$. For each $j = 1, \dots, l$, define $\mu_j := V(f_1, \dots, f_j)$.

Recall that the connected components of the complement $M^{\mathbb{R}}_{\mathcal{H}}$ are called chambers. A *flat* of the arrangement $\mathcal{H}$ is an affine subspace which is an intersection of some hyperplanes in $\mathcal{H}$. By our assumption that the hyperplane arrangement $\mathcal{H}$ is in general position, a flat of $\mathcal{H}$ has codimension j exactly when it is the intersection of j hyperplanes in $\mathcal{H}$.

Lemma 5.8 *For each $j = 1, \dots, l{-}1$, μ_j is a smooth submanifold of $M^{\mathbb{R}}_{\mathcal{H}}$ of codimension j. The closure $\overline{\mu_j}$ of μ_j in $\mathbb{R}^l$ meets the arrangement in a union of codimension $j{+}1$ flats. In the neighbourhood of point on a codimension $j{+}1$–flat meeting $\overline{\mu_j}$, μ_j has at most one branch in each chamber incident on that point.*

Define functions $J_l, J_{l-1}, \dots, J_1$ by recursion,

$$J_j \; := \; \text{Jacobian of } g_1, \dots, g_j, \; J_{j+1}, \dots, J_l \, .$$

The following is proven in [14, § 3.1] using the Cauchy-Binet Theorem.

Lemma 5.9 $J_j \cdot \Big(\prod_{i=1}^{l+n} p_i(y)\Big)^{2^{l-j}}$ *is a polynomial of degree* $2^{l-j} \cdot n$.

For each $j = 1, \ldots, l$, define $C_j := \mu_{j-1} \cap V(J_{j+1}, \ldots, J_l)$. By our assumptions that the polynomials g_i and the exponents β_j are general, this will be a smooth curve in the hyperplane complement $M_{\mathcal{H}}^{\mathbb{R}}$.

We now iterate the Khovanskii-Rolle Theorem 5.7 to estimate the number of solutions to a system of master functions as in Theorem 5.6.

$$\begin{aligned} |V(g_1, g_2, \ldots, g_l)| &\leq \text{ubc}(C_k) + |V(g_1, g_2, \ldots, g_{l-1}, J_l)| \\ &\leq \text{ubc}(C_l) + \text{ubc}(C_{l-1}) + |V(g_1, \ldots, g_{l-2}, J_{l-1}, J_l)| \\ &\leq \text{ubc}(C_l) + \cdots + \text{ubc}(C_1) + |V(J_1, J_2, \ldots, J_l)|\,. \end{aligned} \tag{5.16}$$

Here, $V(\cdots)$ may be taken to mean either the common zeroes in the hyperplane complement $M_{\mathcal{H}}^{\mathbb{R}}$ or in the positive chamber Δ_p. Let $\text{ubc}_\Delta(C)$ count the number of unbounded components of the curve C in Δ_p.

Lemma 5.10 *With these definitions, we have the estimates*

(1) $|V(J_1, \ldots, J_l)| \leq 2^{\binom{l}{2}} n^l$.

(2) $\text{ubc}_\Delta(C_j) \leq \frac{1}{2}\binom{1+l+n}{j} \cdot 2^{\binom{l-j}{2}} n^{l-j}$.

(3) $\text{ubc}(C_j) \leq \frac{1}{2}\binom{1+l+n}{j} \cdot 2^{\binom{l-j}{2}} n^{l-j} 2^j$.

The first statement follows from Lemma 5.9 and Bézout's Theorem. For the second, $C_j = \mu_{j-1} \cap V(J_{j+1}, \ldots, J_l)$. Since each unbounded component has two ends, we estimate the number of unbounded components by $\frac{1}{2}$ of the number of points in the boundary of Δ that are limits of points of C_j. By Lemma 5.8, these will be points in the codimension skeleta where $J_{j+1}, \ldots, J_l$ vanish. The bound in the second statement of 5.10 is simple $\frac{1}{2}$ multiplied by the product of $\binom{1+l+n}{j}$ and $2^{\binom{l-j}{2}} n^{l-j}$. That is by the number of codimension J flats in $\mathcal{H}$ (some of which meet the boundary of Δ) multiplied by the Bézout number of the system $J_{j+1} = \cdots J_l = 0$.

Note that the bound in (2) holds for any chamber of $M_{\mathcal{H}}^{\mathbb{R}}$. We get thebound in (3) by noting that in the neighborhood of any codimension j strata of $\mathcal{H}$, the complement has at most 2^j chambers, and so each point of C_j at the coundary of $M_{\mathcal{H}}^{\mathbb{R}}$ can contribute at most 2^j such ends.

The complement $M_{\mathcal{H}}^{\mathbb{R}}$ of the hyperplane arrangement consists of many chambers. The first bound of Theorem 5.6 is in fact a bound for the number of solutions in any chamber, while the second bound is for the number of solutions in all chambers. Accordingly

Give the extended example showing how the Theorem works in pentagon. (or find one with a hexagon ?)

Describe the buond for all chambers

Do the estimate for $(e^2 + 3)/4$

DO an example completely ?

Explain the sharpness. in a Theorem

Chapter 6: Lower Bounds for Sparse Polynomial Systems

Introduction

In Chapter 1, we mentioned how work of Welschinger [109], Mikhalkin [70], and of Kharlamov, Itenberg, and Shustin [54, 55] combined to show that there is a nontrivial lower bound W_d for the number of real rational curves of degree d interpolating $3d-1$ points in $\mathbb{RP}^2$. If N_d is the Kontsevich number (1.6) of complex rational curves of degree d interpolating $3d-1$ points in $\mathbb{CP}^2$, then we have

1. $W_d \geq \frac{d!}{3}$, and

2. $\lim_{d\to\infty} \frac{\log(N_d)}{\log(W_d)} = 1$. (Note that $\log(N_d) \sim 3d\log(d)$.)

An exposition of the beginning of this story written for a general mathematical audience is given in [94]. It continues. Recently, using the moduli spaces of open strings, J. Solomon [85] gave an interpretation of Welschinger invariants along the lines of the bounds of this section, as the degree of a map. He also gave recursions for these numbers, along the lines of the ???? recursions for Gromov-Witten invariants. Similar results were found by ??? for rational curves on real Calabi-Yau varieties, for example, there are at least $(2n-3)!!$ real lines on a smooth real hyperurface of degree $2n-3$ in $\mathbb{P}^n$.

Eremenko and Gabrielov [32] have a similar result for the number of real solutions to the inverse Wronski problem. They gave numbers $\sigma_{d,k} > 0$ for $0 < k \leq d$ with d even, and proved that if Φ is a real polynomial of degree $k(d{+}1{-}k)$ then there are at least $\sigma_{d,k}$ different k-dimensional subspaces of the form $\mathrm{Span}\{f_1, f_2, \ldots, f_k\}$ where each f_i is a real polynomial of degree d, and their Wronskian $\mathrm{Wr}\,(f_1, f_2, \ldots, f_k)$ is a scalar multiple of Φ. We will explain some more of this in Chapter ?????

While these results are spectacular, they are but the beginning of what we believe will be a bigger and more important story (at least for the applications of mathematics). These are examples of natural geometric problems possessing a *lower bound* on their numbers of real solutions, which may be thought of as existence proofs of solutions. It would be a very important development for some applications if this phenomenon were widespread, if there were methods to detect when such lower bounds existed, and if there were also methods to compute or estimate these lower bounds. The point is that nontrivial lower bounds imply the existence of real solutions to systems of equations, or to interesting geometric problems. A beginning of the interaction between applications and this new theory of lower bounds is found in work of Fiedler-Le Touzé [35] and discussed in Section 1.3 of Chapter 1.

This chapter will report on the first steps toward a theory of lower bounds for sparse polynomial systems as given in [86]. There are three papers where one may read more about this subject.

[32] A. Eremenko and A. Gabrielov, *Degrees of real Wronski maps*, Discrete Comput. Geom. **28** (2002), no. 3, 331–347.

- Establishes a lower bound for the Wronski map, realized as a degree of its lift to oriented double covers.

[86] E. Soprunova and F. Sottile, *Lower bounds for real solutions to sparse polynomial systems*, Advances in Math., **204** (2006), no. 1, 116–151.

- Begins the theory of lower bounds to sparse polynomial systems.

[57] M. Joswig and N. Witte, *Products of foldable triangulations*, Advances in Math., **210**, no. 2, 769–796.

- Uses geometric combinatorics to give many more examples of sparse polynomial systems with a lower bound on their number of real solutions.

The last two papers study lower bounds for unmixed systems, such as those covered by Kouchnirenko's Theorem. It remains an important open problem to develop a theory for unmixed systems such as those which appear in Bernstein's Theorem.† Example 6.7 at the end of this chapter may be the first step in this direction.

6.1 Polynomial systems from posets

Let P be a finite partially ordered set (*poset*), whose elements we take to be our variables. For convenience, assume that P has n elements. The order ideals $I \subset P$ are its up sets—subsets that are closed upwards. Specifically, a subset $I \subset P$ is an *order ideal* if whenever $x < y$ in P with $x \in I$, then we must have $y \in I$. Our polynomials will have monomials indexed by these order ideals. For $I \subset P$ set

$$x^I \;:=\; \prod_{x \in I} x\,.$$

A *linear extension* w of P is a permutation $w\colon x_1, x_2, \ldots, x_n$ of the elements of P such that $x_i < x_j$ in P implies that $i < j$. Let $\lambda(P)$ denote the number of linear extensions of a poset P. The *sign-imbalance $\sigma(P)$* of P is

$$\sigma(P) \;:=\; \left|\sum \mathrm{sgn}(w)\right|\,, \tag{6.1}$$

the sum over all linear extensions w of P. Here, $\mathrm{sgn}(w)$ is the sign of the permutation w. (To be precise, this makes sense only if we first fix one linear extension π of P, and measure the

†Put a statement like this in the chapter on Gale duality!

sign of $w\pi^{-1}$. Taking the absolute value in (6.1) removes the effect of that choice.) Lastly, a *Wronski polynomial* for P is a polynomial of the form

$$\sum_I c_{|I|} x^I \,, \tag{6.2}$$

the sum over all order ideals I of P, where $c_0, c_1, \dots, c_n \in \mathbb{R}$. Note that the coefficient of the monomial x^I depends only upon its degree, $|I|$. The reason for the term Wronski polynomial is that this comes from a linear projection with a special form that is also shared by the Wronski map. (This is discussed just before Section 8.2.1 in Chapter 8.)

Theorem 6.1 *Suppose that P is a finite poset in which every maximal chain has the same parity. Then a generic system of Wronski polynomials for P has $\lambda(P)$ complex solutions, at least $\sigma(P)$ of which will be real.*

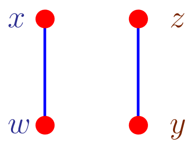
Remark 6.2 A variant of this is to first fix real numbers α_I for each order ideal I and then consider polynomials of the form

$$\sum_I c_{|I|} \alpha_I x^I \,. \tag{6.3}$$

That is, the coefficients $c_0, c_1, \dots, c_n$ vary in $\mathbb{R}^n$, but the numbers α_I are fixed. The same statement holds about such systems. In fact, this variant is closest to the Wronski map in Schubert calculus.

Example 6.3 Let P be the incomparable union of two chains, each of length 2,

P :=



.

The *order ideals* of P are its up-sets: subsets which are closed upwards. We write the corresponding monomials,

$$\{\emptyset,\ x,\ z,\ wx,\ xz,\ yz,\ xyz,\ wxz,\ wxyz\}\,.$$

A Wronski polynomial for P has the form

$$\begin{aligned} & c_4\, wxyz \\ & + c_3(xyz \;+\; wxz) \\ & + c_2(wx \;+\; xz \;+\; yz) \\ & + c_1(x \;+\; z) \\ & + c_0\,, \end{aligned} \tag{6.4}$$

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

There are six linear extensions of P, as each is a permutation of the word $wxyz$ where w precedes x and y precedes z. The sign-imbalance of P is seen to be 2, as computed in the following table.

permutation	$wxyz$	$wyxz$	$ywxz$	$wyzx$	$ywzx$	$yzwx$	$\sigma(P)$
sign	$+$	$-$	$+$	$+$	$-$	$+$	2

Since P has two maximal chains and each contains 2 elements, it satisfies the hypotheses of Theorem 6.1, and so we conclude that:

A system of four equations involving polynomials of the form (6.4) has six solutions, at least two of which are real.

The rest of this chapter will discuss the steps in the proof of Theorem 6.1, as well as the more general results contained in [86].

6.2 Orientability of real toric varieties

Recall from Chapter 3 that a sparse polynomial system with support $\mathcal{A} \subset \mathbb{Z}^n$ is equivalent to a linear section $X_\mathcal{A} \cap L$ of the toric variety $X_\mathcal{A}$. In fact, its solutions are

$$\varphi_\mathcal{A}^{-1}(X_\mathcal{A} \cap L)\,,$$

where $\varphi_\mathcal{A}\colon (\mathbb{C}^\times)^n \ni x \mapsto [x^a \mid a \in \mathcal{A}] \in \mathbb{P}^\mathcal{A}$ is the parametrization map and $X_\mathcal{A}$ is the closure of its image.

Issue 1. In order for there to be a bijective correspondence between real solutions of the original system and real points in the linear section, we need that the map $\varphi_\mathcal{A}$ be injective on $(\mathbb{R}^\times)^n$, which is the condition that the index $[\mathbb{Z}\mathcal{A}\colon \mathbb{Z}^n]$ be odd. This should be regarded as a minor issue and a mild assumption.

The key idea at the beginning of this theory is to realize the intersection $X_\mathcal{A} \cap L$ as the fiber of a map. To that end, let $E \subset L$ be a hyperplane in L that does not meet $X_\mathcal{A}$ and $M \simeq \mathbb{P}^n$ a linear space that is disjoint from E. Then E has codimension $n{+}1$ and the set of codimension n planes containing E is naturally identified with $M \simeq \mathbb{P}^n$, as each codimension n plane containing E meets M in a unique point. Define the *linear projection*

$$\pi = \pi_E \;:\; \mathbb{P}^\mathcal{A} - E \;\longrightarrow\; M \simeq \mathbb{P}^n \tag{6.5}$$

by sending a point $p \in \mathbb{P}^\mathcal{A} - E$ to the intersection of M with span of E and p. Figure 6.1 illustrates this in $\mathbb{P}^3$, where E and M are lines. We sometimes write $\pi\colon \mathbb{P}^\mathcal{A} - \to \mathbb{P}^n$, using the broken arrow $- \to$ to indicate that the map is not defined on all of $\mathbb{P}^\mathcal{A}$.

Write π for the restriction of the linear projection to the toric variety $X_\mathcal{A}$. If $y := L \cap M$ is the point where L meets M, then

$$X_\mathcal{A} \cap L \;=\; \pi^{-1}(y)\,.$$

This is also illustrated in Figure 6.1, where $X_\mathcal{A}$ is a cubic curve. The reason that we reformulate our system of polynomials as the fiber of a map is so that we may use topological methods to study its solutions.

That is, we work over $\mathbb{R}$, define $Y_\mathcal{A} := X_\mathcal{A} \cap \mathbb{RP}^\mathcal{A}$, and consider the real linear section $Y_\mathcal{A} \cap L_\mathbb{R}$. Equivalently, we restrict the projection π further to a map

$$f \;:\; Y_\mathcal{A} \;\to\; \mathbb{RP}^n\,,$$

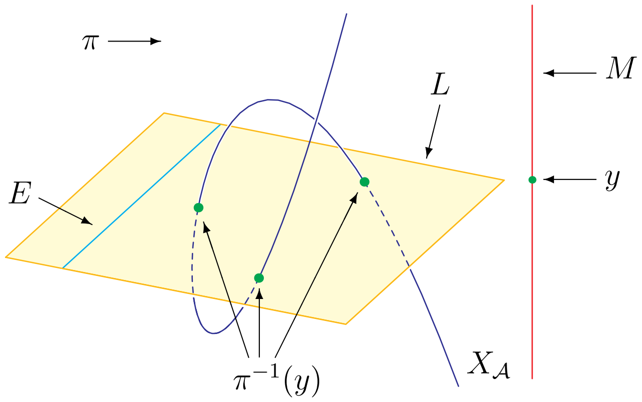


Figure 6.1: A linear projection π with center E.

and consider points in the fiber $f^{-1}(p)$, where $p \in \mathbb{RP}^n$. Since $Y_{\mathcal{A}}$ and $\mathbb{RP}^n$ have the same dimension, the map f has a topological degree. For this, we must fix orientations of $Y_{\mathcal{A}}$ and of $\mathbb{RP}^n$. Then, for every regular regular value $y \in \mathbb{RP}^n$ of f, the differential map df is a bijection on the tangent spaces of every point $x \in f^{-1}(y)$. Define $\operatorname{sign}(x) := 1$ if the differential preserves the orientations at x and y and $\operatorname{sign}(x) := -1$ is the orientations are reversed. Then the topological degree of f is the sum

$$\sum_{x \in f^{-1}(y)} \operatorname{sign}(x)\,.$$

This definition does not depend upon the choice of a regular value y, as $\mathbb{RP}^n$ is connected.[†] The value of the notion of topological degree is the following.[§]

Theorem 6.4 *The number of points in a fiber $f^{-1}(y)$, for $y \in \mathbb{RP}^n$ a regular value of f, is at least the absolute value of the topological degree of f.*

Since both $\mathbb{RP}^n$ and $Y_{\mathcal{A}}$ are not always orientable, the topological degree of f is not always defined. To remedy this, we consider also the lift of the linear projection π to the double covers coming from spheres:

$$\begin{array}{rcccccc} f^+ : & Y^+_{\mathcal{A}} & \subset & S^{\mathcal{A}} & \overset{\pi^+}{- - \to} & S^n \\ & \downarrow & & \downarrow & & \downarrow \\ f \ : & Y_{\mathcal{A}} & \subset & \mathbb{RP}^{\mathcal{A}} & \overset{\pi}{- - \to} & \mathbb{RP}^n \end{array} \qquad (6.6)$$

Then the topological degree of the map f^+ provides a lower bound for the number of solutions, as in Theorem 6.4. We seek criteria which imply that the map f or its lift f^+ to the double covers has a well-defined degree. That is, either $Y_{\mathcal{A}}$ or $Y^+_{\mathcal{A}}$ is orientable.

For this, we use Cox's construction of X_Δ as a quotient of a torus acting on affine space, as detailed in [20, Theorem 2.1]. (In the symplectic category, this is due to Delzant.) The

† Give a reference for this? How does it go back to Kronecker?

§ Reconcile this with the use of degree in Chapter 2.

convex hull of $\mathcal{A}$ is the polytope $\Delta := \Delta_{\mathcal{A}}$. We need the dual description of Δ in terms of intersections of half-spaces, or facet inequalities, which have the form

$$\Delta \;=\; \{z \in \mathbb{R}^n \;\mid\; B \cdot z \;\geq\; -b\}\,,$$

where B is in integer N by n matrix (N is the number of facets of Δ) whose rows are the inward-pointing normals to the facets of Δ, and $b \in \mathbb{Z}^N$ measures the (signed) lattice distance of each facet from the origin. Then there is a natural parametrization $\psi_\Delta\colon \mathbb{C}^N - Z_\Delta \twoheadrightarrow X_{\mathcal{A}}$, realizing $X_{\mathcal{A}}$ as a geometric invariant theory quotient of $\mathbb{C}^N$ by a subtorus of $\mathbb{T}^N$. (We do not give the details here, see [20] or [86].)

Issue 2. While ψ_Δ restricts to a map $\psi_\Delta\colon \mathbb{R}^N - Z_\Delta \to Y_{\mathcal{A}}$ (and also to Y^+_Δ), these maps are not necessarily surjective. This map ψ_Δ is surjective if and only if the column span $\mathrm{col}(B)$ of the matrix B has odd index in its saturation, $\mathbb{Z}^N \cap (\mathbb{Q} \otimes_{\mathbb{Z}} \mathrm{col}(B))$.

When the standard orientation of $\mathbb{R}^N$ drops to an orientation either of $Y_{\mathcal{A}}$ or of $Y^+_{\mathcal{A}}$ under the Cox quotient map ψ_Δ, we say that $Y_{\mathcal{A}}$ is *Cox-orientable*.

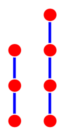
Theorem 6.5 *With the assumptions outlined in Issue 1 and Issue 2, if* $\mathrm{col}(B) + \mathbb{Z}\cdot b$ *contains a vector, all of whose coordinates are odd, then* $Y_{\mathcal{A}}$ *is Cox-orientable. If* $\mathrm{col}(B)$ *has such a vector, then* $Y_{\mathcal{A}}$ *receives the orientation, otherwise* $Y^+_{\mathcal{A}}$ *receives the orientatation.*

Remarks.

1. Given a projection map π (6.5) whose center E is disjoint from $Y_{\mathcal{A}}$, write f for its restriction to $Y_{\mathcal{A}}$. Suppose that $Y_{\mathcal{A}}$ is Cox-orientable. Then lift $f^+\colon Y^+_{\mathcal{A}} \to S^n$ of f to $Y^+_{\mathcal{A}}$ has a well-defined degree, which is a lower bound for the number of real solutions to polynomial systems arising as fibers of the map f. Call this number the *real degree* of the map f.

2. If P is a poset and $\mathcal{A} \subset \{0,1\}^P$ consists of the indicator functions of its order ideals, so that a Wronski polynomial for P (6.2) has support $\mathcal{A}$, then it is not hard to show that $\mathbb{Z}\mathcal{A} = \mathbb{Z}^n$ and also $Y_{\mathcal{A}}$ is Cox-orientable if all maximal chains of P have the same parity.

   This parity condition is sufficient, but not necessary for the Wronski polynomial system on $Y_{\mathcal{A}}$ to have a lower bound. The poset

   P =  ,

   has two maximal chains of lengths 3 and 4, and so it is not necessarily Cox-oriented. Nevertheless, its Wronski polynomial systems have $\binom{7}{3} = 35$ solutions, at least $\binom{3}{1} = 3$ of which are real.

6.3 Degree from foldable triangulations

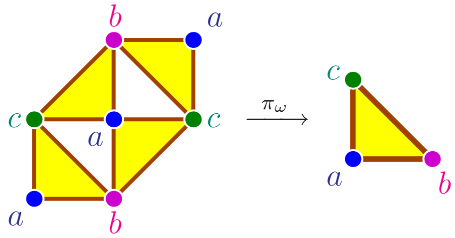
These results provide us with a challenge: compute the real degree of a (or any) map f arising as a linear projection of a Cox-orientable toric variety $Y_{\mathcal{A}}$. We give a methods that uses the toric degenerations of Lecture 2 to provide an answer to this question. It is by no means the only answer.

Recall that if $\mathcal{A} = \Delta \cap \mathbb{Z}^n$ for a lattice polytope Δ which has a regular unimodular triangulation Δ_ω, then there exists a flat deformation of $X_{\mathcal{A}}$ within the projective space $\mathbb{P}^{\mathcal{A}}$ into a limit scheme which is the union of coordinate n-planes

$$\lim_{t\to 0} t.X_{\mathcal{A}} \;=\; \bigcup_{\tau\in\Delta_\omega} \mathbb{P}^\tau\,.$$

The idea here is to find conditions on the triangulation that allow us to compute the real degree of some map f from this deformation. Some triangulations Δ_ω have a naturally defined linear projection with a nicely defined degree.

Example 6.6 Consider the triangulation of the hexagon (the HSBC Bank symbol rotated 45° anti-clockwise) shown below with the vertices labeled a, b, and c. Mapping these vertices to the corresponding vertices of the simplex defines a piecewise-linear 'folding' map π_ω whose degree is $4-2=2$.



Suppose now that $\mathcal{A} = \Delta \cap \mathbb{Z}^n$, where Δ is a lattice polytope. A triangulation Δ_ω is *foldable* if its facet simplices may be properly 2-colored, which is equivalent to there being a coloring of the vertices $\mathcal{A}$ of the triangulation with $n+1$ labels, where each simplex receives all $n+1$ labels [56]. Both the 2-coloring and the vertex-labeling are unique up to permuting the colors and labels. The difference in the number of simplices of different colors is the *signature $\sigma(\omega)$* of the balanced triangulation. Up to a sign, it is the degree of the combinatorial folding map from Δ to an n-simplex given by the labeling of $\mathcal{A}$.

The vertex labels define a linear projection $\pi_\omega\colon \mathbb{P}^{\mathcal{A}} - \to \mathbb{P}^n$. Let $\mathbf{e}_a$ be the standard basis vector of $\mathbb{P}^{\mathcal{A}}$ corresponding to the monomial $a\in\mathcal{A}$ and $\mathbf{e}_j$ for $j=0,1,\ldots,n$ the standard basis vector of $\mathbb{P}^n$. We assume that the labels in the balanced triangulation take values $0,1,\ldots,n$. Then $\pi_\omega(\mathbf{e}_a)=\mathbf{e}_j$, where the vertex a of has label j. The restriction of this linear projection to the coordinate spheres $\pi_\omega\colon \cup_\tau S^\tau \to S^n$ is the geometric counterpart of this combinatorial folding map.

From the geometry of this map, we can show that for $t>0$ and small, the restriction f of π_ω to $t.Y^+_{\mathcal{A}}$ has degree $\sigma(\omega)$. This can be deduced from the real version of the arguments used in the algorithmic proof of Koushnirenko's Theorem in Lecture 2. In fact, Figure 3.1 from Lecture 2 is reproduced from this argument in [86]. expand on this??

The restrictions of π_ω to $t.Y^+_{\mathcal{A}}$ for t near zero and $t=1$ are homotopic, **if** $t.Y_{\mathcal{A}}\cap\ker(\pi_\omega)=\emptyset$, for $t\in(0,1)$. This condition is not hard to check on specific examples.

6.4 Reprise: polynomial systems from posets

Suppose now that P is a poset and $\mathcal{A} \subset \{0,1\}^P$ is the set of indicator functions of order ideals I of P, so that a Wronski polynomial for P has support $\mathcal{A}$. Then $\Delta_P := \mathrm{conv}(\mathcal{A})$ is the *order polytope* of P. This was studied by Stanley [96]. Its facet inequalities are

$$B \,:\, \left\{\begin{array}{rcrl} y & \geq & 0 & \text{if } y \,\in\, P \text{ is mininmal} \\ -y & \geq & -1 & \text{if } y \,\in\, P \text{ is maximal} \\ x-y & \geq & 0 & \text{if } y \,\lessdot\, x \text{ is a cover in } P \end{array}\right. .$$

Give more details in these breif paragraphs It has a unimodular triangulation Δ_ω whose simplices correspond to linear extensions of P, and it is easy to see that it is regular. Unimodular implies that $\mathbb{Z}\mathcal{A} = \mathbb{Z}^n$. Recently, Stanley [97] and others introduced the notion of *sign-(im)balance* of posets, $\sigma(P)$, and it is tautological that $\sigma(\omega) = \sigma(P)$. It is also easy to use the description of the facet inequalities to show that $\mathrm{col}(B)$ is saturated, and if every maximal chain in P has the same parity, then $Y_\mathcal{A}$ is Cox-oriented. This explains Theorem 6.1.

6.5 Sagbi degenerations

This method of computing a degree through a limiting process applies to the original result of Eremenko and Gabrielov on the degree of the Wronski map in Schubert calculus [32]. The Grassmannian $\mathrm{Gr}(p, m+p)$ of p-planes in $(m+p)$-space admits a *sagbi degeneration* to the toric variety associated to the poset $P_{p,m}$ which is the product of a chain with p elements and a chain with m elements, and one may show that the degree of the Wronski map for the Grassmannian $\mathrm{Gr}(p, m+p)$ is equal to the degree of the Wronski map f (the restriction of π_ω) on $Y^+_\mathcal{A}$, where $\mathcal{A}$ is the set of monomials coming from order ideals of $P_{p,m}$. (For this, we use a more general version of the Wronski projection π_ω which corresponds to the more general Wronski polynomial of (6.3).)

Using the sagbi degeneration, we can recover the results of Eremenko and Gabrielov. In fact, this connection between the Grassmannian and its toric degeneration, as well as between Eremenko and Gabrielov's formula for the degree of the Wronski map, the sign-imbalance of $P_{p,m}$, and the geometric folding map was the genesis of this work [86].

6.6 Future directions

There is much more to be done in this area. Here are some suggestions.

1. Give more comprehensive conditions which imply that f, or its lift to some (not necessarily the one given above) double cover of $Y_\mathcal{A}$ is orientable.

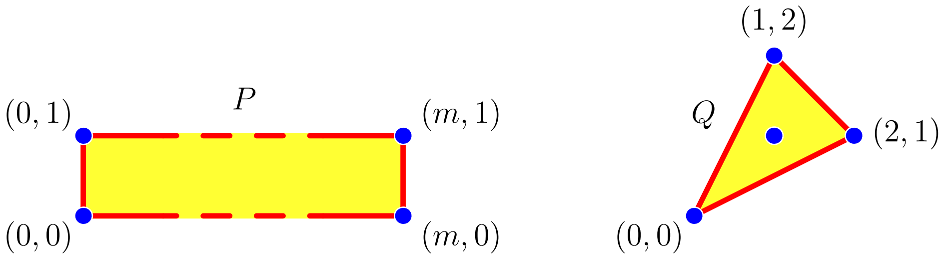
2. Find other methods to give polynomial systems whose degree may be computed or estimated.

3. Find more balanced triangulations (see [57]).

4. Apply these ideas to specific problems from the applied sciences.

5. Extend any of this from unmixed systems (all polynomials have the same Newton polytope) to more general mixed systems (those whose polynomials have different Newton polytopes). We end with an example in this direction which is due to Chris Hillar.

Example 6.7 Let P and Q be the two lattice polytopes given below



A polynomial with support P has the form

$$g \ := \ A(x) \ + \ yB(X)\,,$$

where A and B are univariate polynomials in x with degree m. Suppose that their coefficients are $a_0, \ldots, a_m$ and $b_0, \ldots, b_m$, and let h be the polynomial with support Q,

$$h \ := \ c \ + \ dxy \ + \ ex^2y \ + \ fxy^2\,.$$

By Bernstein's Theorem (Theorem 1.2), the mixed system $g(x,y) = h(x,y) = 0$ will have $2m+2$ solutions in $(\mathbb{C}^\times)^2$, as $2m+2$ is the mixed volume of P and Q. (For polygons P, Q, the mixed volume is $\mathrm{vol}(P+Q) - \mathrm{vol}(P) - \mathrm{vol}(Q)$.) We can compute an eliminant for this system by substituting $-A(x)$ for $yB(x)$ in $h(B(x))^2$, to obtain

$$cB(x)^2 \ - \ dxA(x)B(x) \ - \ ex^2A(x)B(x) + fxA(x)^2\,.$$

This has constant term cb_0^2 and leading term $-ea_mb_m$. If $ce > 0$ and $a_mb_m > 0$, then these have different signs, which implies that the mixed system has at least one positive root (and hence at least two real roots). This may be ensured by the condition that none of the coefficients vanish and the two linear equations, $c + e = a_m + b_m = 0$.

Chapter 7: Enumerative Real Algebraic Geometry

Introduction

Enumerative geometry is the art of counting geometric figures satisfying conditions imposed by other, fixed, geometric figures. For example, in 1848, Steiner [98] asked how many plane conics are tangent to five given conics? His answer, $6^5 = 7776$, turned out to be incorrect, and in 1864 Chasles [19] gave the correct answer of 3264. These methods were later systematized and used to great effect by Schubert [79], who codified the field of enumerative geometry.

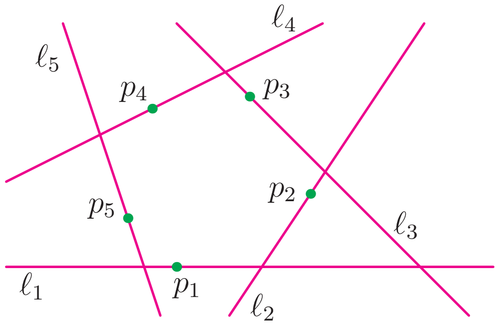
This classical work always concerned *complex* figures. It was only in 1984 that the question of reality was posed by Fulton [39, p. 55]: "The question of how many solutions of real equations can be real is still very much open, particularly for enumerative problems." He goes on to ask: "For example, how many of the 3264 conics tangent to five general conics can be real?" He later determined that all can be real, but did not publish that result. Ronga, Tognoli, and Vust [75] later gave a careful argument that all 3264 can be real.

Since this work of Ronga, Tognoli, and Vust, there have been many geometric problems for which it was shown that all solutions may be real. This means that the upper bound of d (= number of complex solutions) is sharp for these problems. In this chapter, we describe some of these problems, beginning with the problem of conics, and concentrating on the Schubert calculus. A survey of these questions (circa 2002) is given in [93].

7.1 3264 real conics

The basic idea of the arguments of Fulton and of Ronga, Tognoli, and Vust is to deform the same special configuration. We will sketch the idea in Fulton's construction.

Suppose that $\ell_1, \ldots, \ell_5$ are the lines supporting the edges of a convex pentagon and $p_i \in \ell_i$, $i = 1, \ldots, 5$ are points in the interior of the corresponding edge.



The points in this example are $\{(0,0), (\frac{11}{4}, \frac{3}{2}), (\frac{3}{2}, \frac{7}{2}), (-\frac{1}{2}, \frac{13}{4}), (-1, 1)\}$, and the corresponding slopes of the lines are $0, \frac{3}{2}, -1, \frac{1}{2}, -3$.

For every subset S of the lines, there are $2^{\min\{|S|,5-|S|\}}$ conics that are tangent to the lines in S and that meet the $5-|S|$ points not on the lines of S. This is the number of complex conics, and it does not depend upon the configuration of (generic) points and lines. However, when the points and lines are chosen in convex position, then all such conics will be real. Altogether, this gives

$$2^0\binom{5}{0} \;+\; 2^1\binom{5}{1} \;+\; 2^2\binom{5}{2} \;+\; 2^2\binom{5}{3} \;+\; 2^1\binom{5}{4} \;+\; 2^0\binom{5}{5} \;=\; 102$$

real conics, that, for each $i=1,\ldots,5$ either meet p_i or are tangent to ℓ_i. We draw these in Figure 7.1. Since our pentagon was asymmetric, exactly 51 of these conics meet each point

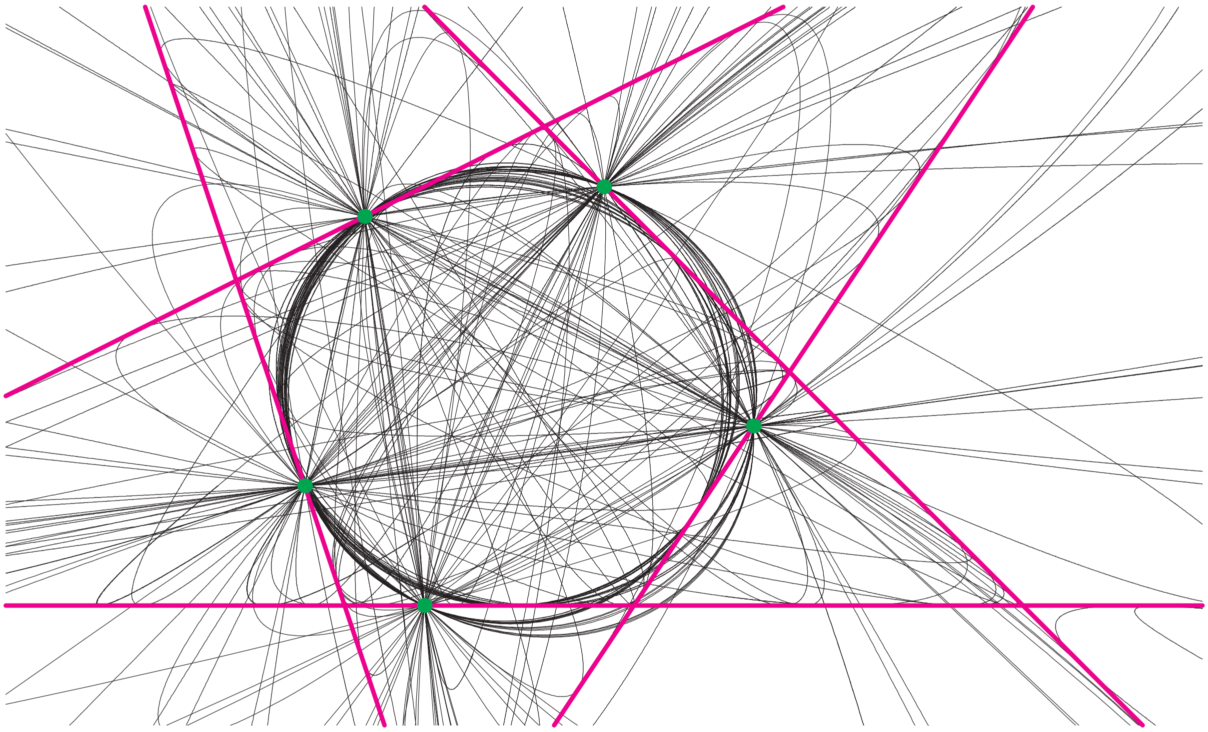
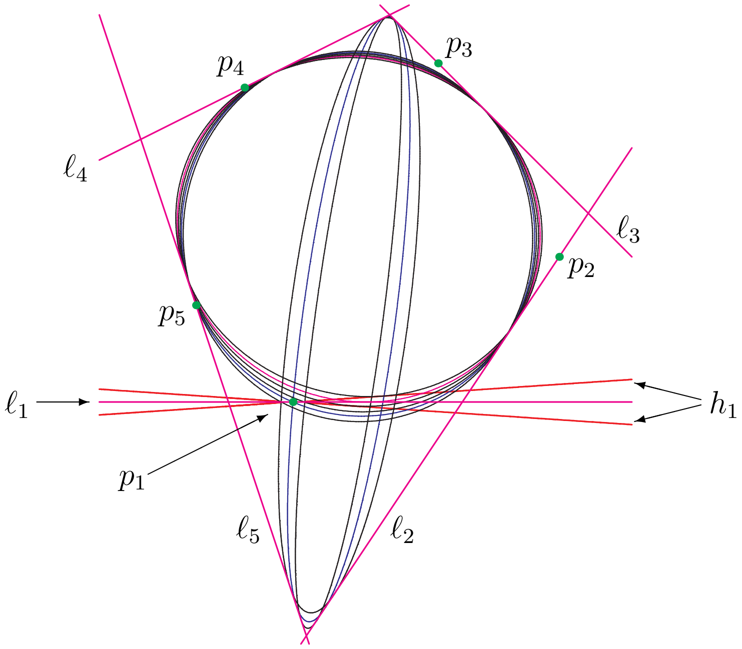


Figure 7.1: 102 conics

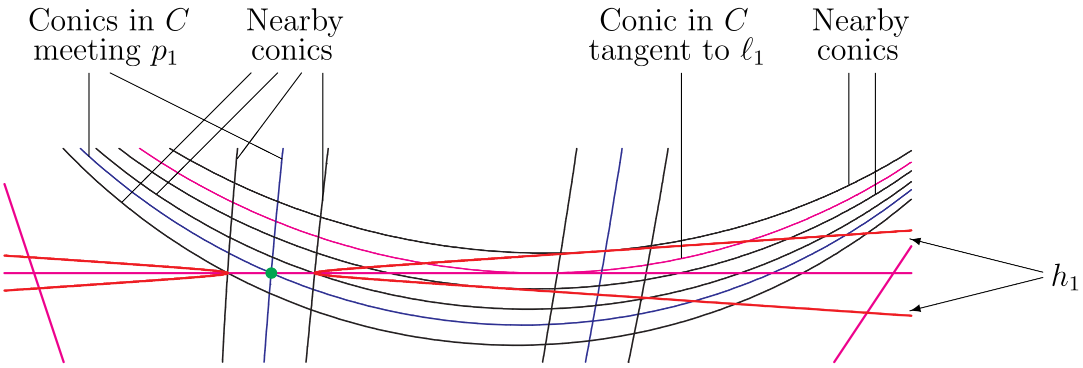
p_i and none of the 51 conics tangent to ℓ_i are tangent at p_i.

The idea now is to replace each pair $p_i \in \ell_i$ by a hyperbola h_i that is close to the pair $p_i \in \ell_i$, in that it lies close to its asymptotes, which are two lines close to ℓ_i that meet at p_i. If we do this for one pair $p_i \in \ell_i$, then, for every conic in our configuration, there will be two nearby conics tangent to h_i. To see this, suppose that $i=1$. Then the set C of conics which satisfy one of the conditions "meet p_j" or "tangent to ℓ_j" for each $j=2,3,4,5$ will form an irreducible curve C. For each conic in C that meets p_1, there will be two nearby conics in C tangent to h_1 near p_1, and for each conic in C tangent to ℓ_1, there will be two nearby conics in C tangent to each of the two nearby branches of h_1. We ilustrate this when

C is the curve of conics tangent to ℓ_2, ℓ_3, ℓ_4, and ℓ_5, showing the conics in C,



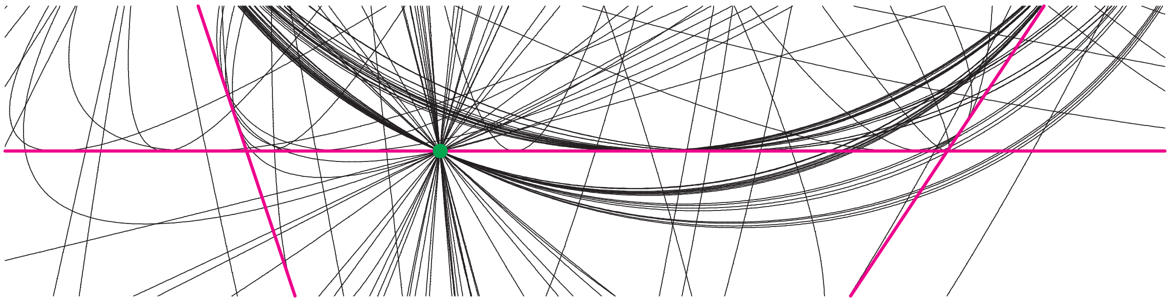
and then a closeup near l_1.



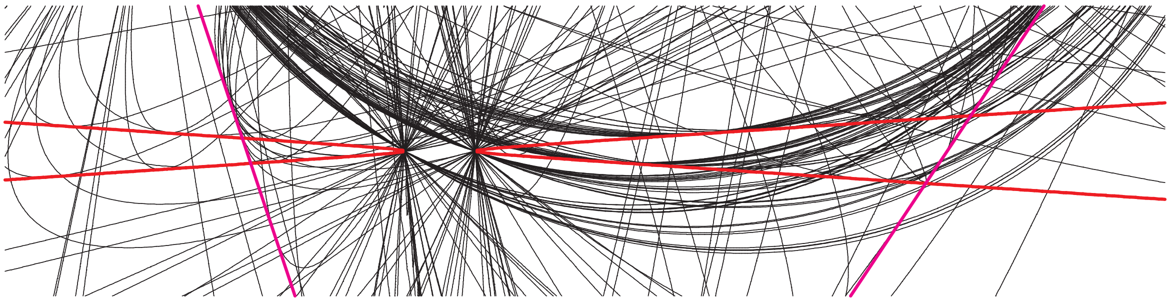
For the configuration of 102 conics of figure 7.1, the hyperbola

$$h_1 : \qquad (y - \tfrac{1}{15}x)(y + \tfrac{1}{15}x) + \tfrac{1}{15000} = 0,$$

is sufficiently close to its asymptotes, which meet at p_1 and are sufficiently close to ℓ_1, and so each of our 102 conics that meet p_1 or are tangent to ℓ_1 becomes two conics tangent to h_1. We first show the configuration of 102 conics of Figure 7.1 in the neighborhood of ℓ_1,



and then the resulting 204 conics in the same region.



The key to the existence of this construction was that no tangent direction to a conic through p_1 met h_1, which is possible as no conic was tangent to ℓ_1 at p_1. Figure 7.2 shows the resulting 204 conics that are tangent to h_1 and, for each $i = 2, 3, 4, 5$ either contain p_i or are tangent to ℓ_i.

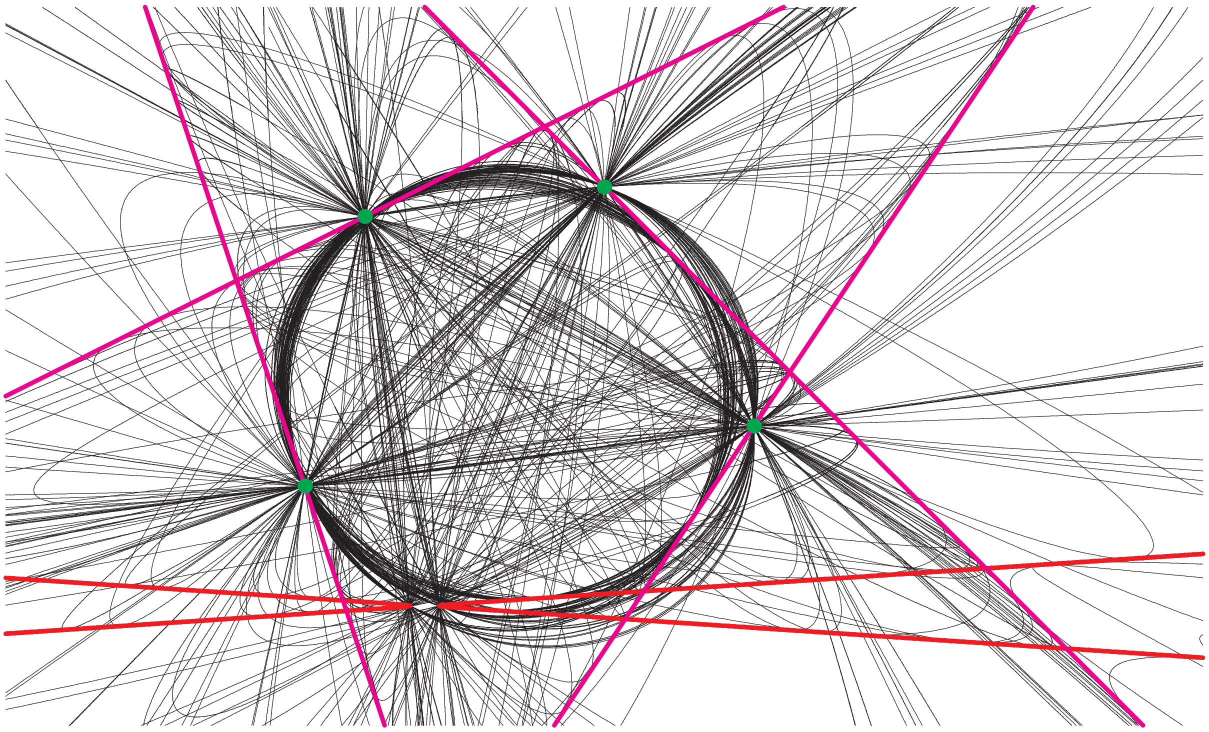


Figure 7.2: 204 conics

If we now replace $p_2 \in \ell_2$ by a similar nearby hyperbola, then the 204 conics become 408. Replacing $p_3 \in \ell_3$ by a nearby hyperbola, will give 816 conics. Continuing with $p_4 \in \ell_4$ gives 1632, and finally replacing $p_5 \in \ell_5$ with a hyperboloa gives five hyperbolae, $h_1, \ldots, h_5$ for which there are $2^5 \cdot 102 = 3264$ real conics tangent to each h_i. In this way, the classical problem of 3264 conics can have all of its solutions be real. Observe that this discussion also gives a derivation of the number 3264 without reference to intersection theory [37]

7.2 Some geometric problems

We discuss some other geometric problems that can have all their solutions be real.

7.2.1 The Stewart-Gough platform

The position of a rigid body in $\mathbb{R}^3$ has 6 degrees of freedom (3 rotations and 3 translations). This is exploited in robotics, giving rise to the *Stewart-Gough platform* [44, 99]: Suppose that we have 6 fixed points $A_1, A_2, \ldots, A_6$ in space and 6 points $B_1, B_2, \ldots, B_6$ on a rigid body B (the framework of Figure 7.3). The body is controlled by varying each distance l_i

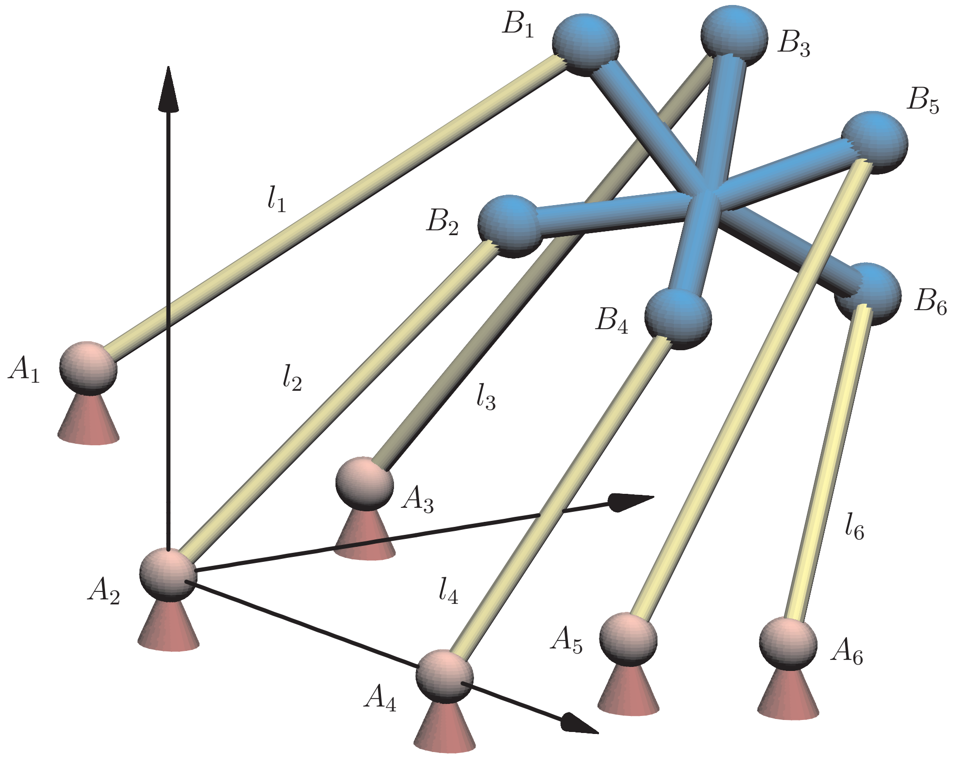


Figure 7.3: A Stewart platform.

between the fixed point A_i and the point B_i on B. This may be accomplished by attaching rigid actuators between spherical joints located at the points A_i and B_i, or by suspending the platform from a ceiling with cables. (Apparently, this configuration is often used in factories.)

Given a position of the body B, the distances $l_1, l_2, \ldots, l_6$ are uniquely determined. A fundamental problem is the inverse question (in robotics, this is called the forward problem): Given a platform (positions of the A_i fixed and the relative positions of the B_i specified) and a sextuple of distances $l_1, l_2, \ldots, l_6$, what is the position of the platform?

It had long been understood that several positions were possible for a given sextuple of lengths. This led to the following enumerative problem.

For a given (or general) Stewart platform, how many (complex) positions are there for a generic choice of the distances $l_1, l_2, \dots, l_6$? How many of these can be real?

In the early 1990's, several approaches (including a nice interaction between theory and computer experimentation [65, 76]) showed that there are 40 complex positions of a general Stewart platform. The obviously practical question of how many positions could be real remained open until 1998, when Dietmaier introduced a novel method involving numerical homotopy to find a platform and value of the distances $l_1, l_2, \dots, l_6$ with all 40 positions real.

Theorem 7.1 (Dietmaier [26]) *All 40 positions can be real.*

7.2.2 Real rational cubics through 8 points in $\mathbb{P}^2_{\mathbb{R}}$

Recall from Section 1.3 that there are 12 singular (rational) cubic curves containing 8 general points in the plane. Kharlamov studied this over the real numbers.

Theorem 7.2 ([22, Proposition 4.7.3]) *Given 8 general points in $\mathbb{P}^2_{\mathbb{R}}$, there are at least 8 real rational cubics containing them, and there are choices of the 8 points for which all 12 rational cubics are real.*

The bound is a nice exercise in Euler characteristic. A homogeneous cubic has 10 coefficients, and so the set of plane cubics is naturally identified with 9-dimensional projective space. Let $p_1, \dots, p_8$ be general points in $\mathbb{RP}^2$. As the condition for a cubic to contain a point p_i is linear in the coefficients of the cubic, there is a pencil ($\mathbb{P}^1$) of cubics through these 8 points. Let P, Q be two distinct cubics in this pencil, which is then parametrized by $sP + tQ$ for $[s,t] \in \mathbb{P}^1$. By Bézout's Theorem, the cubics P and Q, and hence every cubic in the pencil, vanish at a ninth point, p_9.

It is not hard to see that there is a unique cubic in the pencil that vanishes at any point $p \in \mathbb{RP}^2 - \{p_1, \dots, p_9\}$. A little harder, but still true, is that there is a unique cubic in the pencil with any given tangent direction at some point p_i. In this way, we have maps

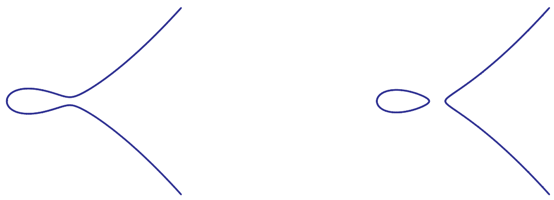
$$\begin{array}{ccc} Z & := & \mathrm{Bl}_{\{p_1,\dots,p_9\}}\mathbb{RP}^2 \\ {\scriptstyle C}\big\downarrow & & \big\downarrow{\scriptstyle \pi} \\ \mathbb{RP}^1 & & \mathbb{RP}^2 \end{array}$$

where $\mathrm{Bl}_{\{p_1,\dots,p_9\}}\mathbb{RP}^2$ is the blow-up of $\mathbb{RP}^2$ in the 9 points, which is obtained by removing each point p_i and replacing it with the tangent directions $\mathbb{RP}^1 \simeq S^1$ to $\mathbb{RP}^2$ at p_i. The map π is the blow-down, and the map C associates a point of Z to the unique curve in the pencil which contains that point.

Because Z is a blow-up, we may compute its Euler characteristic to obtain

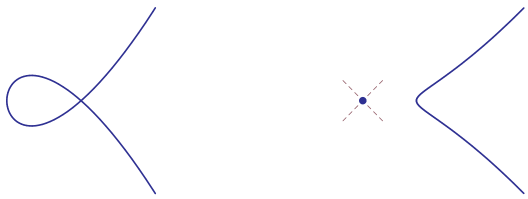
$$\begin{aligned} \chi(Z) &= \chi(\mathbb{RP}^2) - 9 \cdot \chi(\mathrm{pt}) + 9 \cdot \chi(S^1) \\ &= 1 - 9 + 0 = -8\,. \end{aligned}$$

The key to Theorem 7.2 is to compute the Euler characteristic of Z using the map $C\colon Z \to S^1$. The fibers of this map are the cubic curves in the pencil. Smooth real cubics either have one or two topological components,



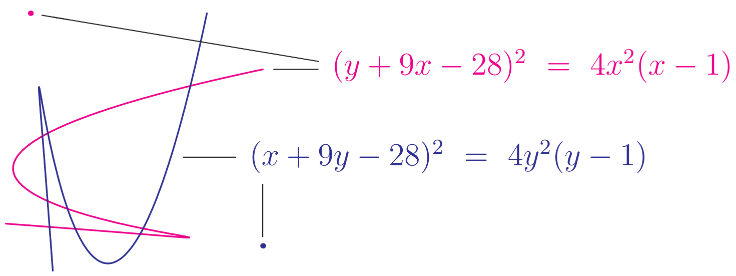
and hence are homeomorphic to one or two copies of S^1. In either case, their Euler characteristic is zero, and so the Euler characteristics of singular fibers of the map C sum to -8.

Because the points $p_1, \ldots, p_8$ were general, the only possible singular fibers are nodal cubics, and there are two types of real nodal cubics.



The first is the topological join of two circles and has Euler characteristic -1, while the second is the disjoint union of a circle with a point and therefore has Euler characteristic 1. Thus, if n is the number of nodal cubics in the pencil and s is the numer with a solitary point, then we have $n + s \leq 12$ (as there are 12 complex rational cubics), and $s - n = -8$.

There are three solutions to this system, $(n, s) \in \{(8, 0), (9, 1), (10, 2)\}$. Thus there are at least 8 real cubics through the 8 points, and this is the derivation of the Welschinger number $W_3 = 8$. Moreover, if there are 2 cubics in the pencil with solitary points, then all 12 rational cubics will be real. Such a pencil is generated by the two cubics given below, whose curves we also display.



We conclude that there will be 12 real rational cubics interpolating any subset of 8 of the 9 points where these cubics meet.

The question of how many of the N_d (1.6) rational curves of degree d which interpolate $3d-1$ points in $\mathbb{RP}^2$ can be real remains open. (This was posed in [87].)

7.2.3 Common tangent lines to $2n{-}2$ spheres in $\mathbb{R}^n$

How many common tangent lines are there to $2n{-}2$ spheres in $\mathbb{R}^n$? For example, when $n = 3$, how many common tangent lines are there to four spheres in $\mathbb{R}^3$? (The number $2n{-}2$ is the dimension of the space of lines in $\mathbb{R}^n$ and is necessary for there to be finitely many common tangents.) Despite its simplicity, this question does not seem to have been asked classically, but rather arose in computational geometry. Macdonald, Pach, and Theobald [68] gave an elementary argument that four spheres with the same radius in $\mathbb{R}^3$ can have at most 12 common tangents. Then they considered the symmetric configuration where the spheres are centered at the vertices of a regular tetrahedron. If the spheres overlap pairwise, but no three have a common point, then there will be exactly 12 common real tangents, as illustrated in Figure 7.4.

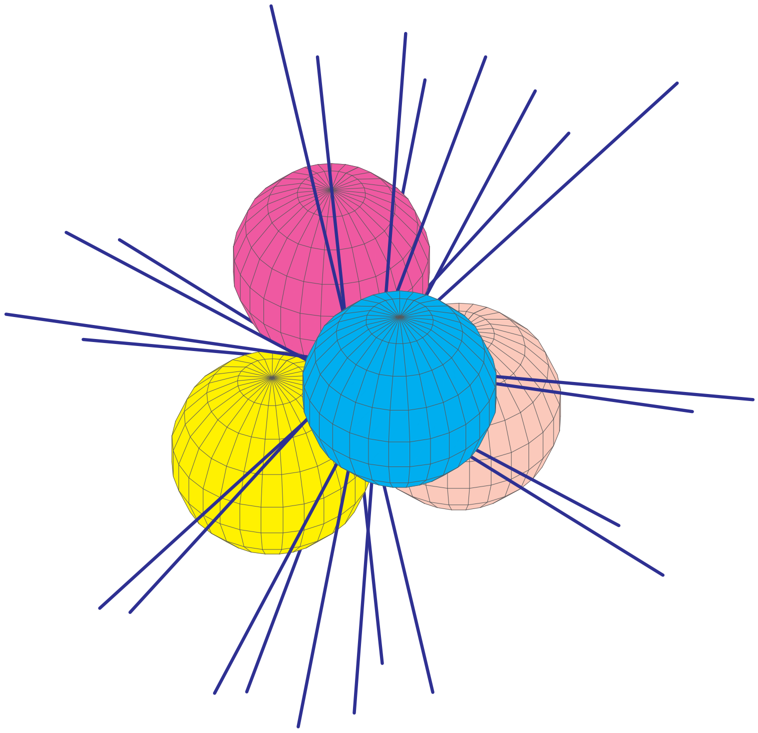


Figure 7.4: Four spheres with 12 common tangents.

The general case was established soon after [95].

Theorem 7.3 *$2n-2$ general spheres in $\mathbb{R}^n$ $(n \geq 3)$ have $3 \cdot 2^{n-1}$ complex common tangent lines, and there are $2n-2$ such spheres with all common tangent lines real.*

The same elementary arguments of Macdonald, Pach, and Theobald give a bound valid for all n and spheres of any radius, and a generalization of the symmetric configuration of Figure 7.4 gives a configuration of $2n-2$ spheres having $3 \cdot 2^{n-1}$ common real tangents.

Megyesi [69] showed that this result for $n = 3$ remains true if the spheres have coplanar centers (Figure 7.5), but that there can only be 8 common real tangents (out of 12 complex ones) if the spheres have the same radii (Figure 7.6).

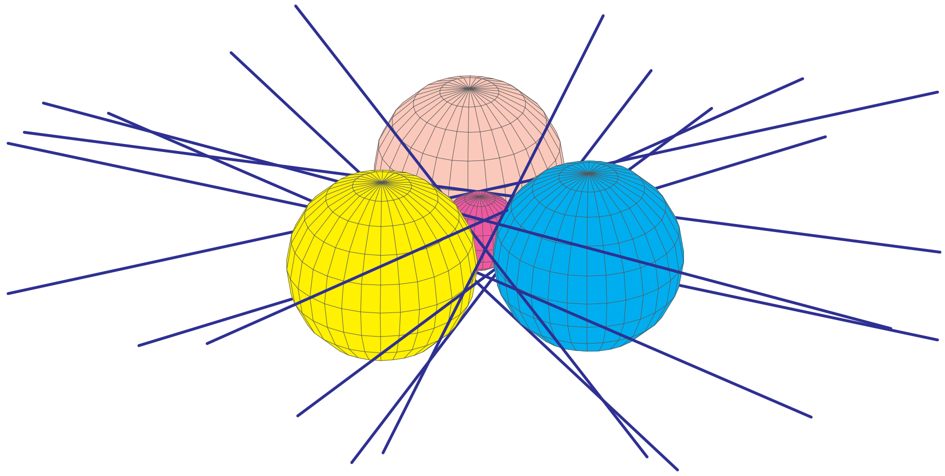


Figure 7.5: Four spheres with coplanar centers and 12 common tangents.

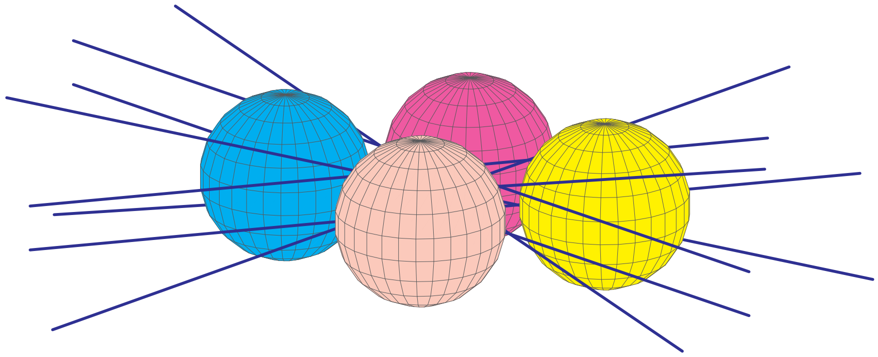
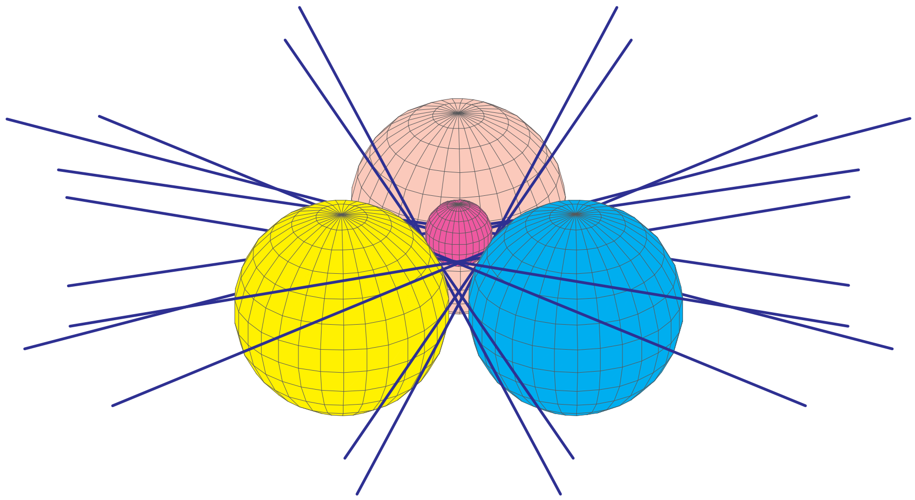


Figure 7.6: Four equal spheres with coplanar centers and 8 common tangents.

The spheres in Figures 7.4 and 7.5 are not disjoint, in fact their union is connected. Fulton asked if it were possible for 4 disjoint spheres to have 12 common real tangents. It turns out that a pertubation of the configuration of Figure 7.5 gives four pairwise disjoint spheres with 12 common tangents, as we show below.



The three large spheres have radius 4/5 and are centered at the vertices of an equilateral triangle of side length $\sqrt{3}$, while the smaller sphere has radius 1/4 and is centered on the axis of symmetry of the triangle, but at a distance of 35/100 from the plane of the triangle. It remains an open question whether it is possible for four disjoint *unit* spheres to have 12 common tangents.

7.3 Schubert calculus

The largest class of problems which have been studied from the perspective of having all solutions real come from the classical Schubert calculus of enumerative geometry, which involves linear spaces meeting other linear spaces. The simplest nontrivial example illustrates some of the vivid geometry behind this class of problems. Consider the following question:

How many line transversals are there to four given lines in space?

To answer this, first consider three lines. They lie on a unique hyperboloid. (See Figure 7.7.) This hyperboloid has two rulings by lines. The three lines are in one ruling, and

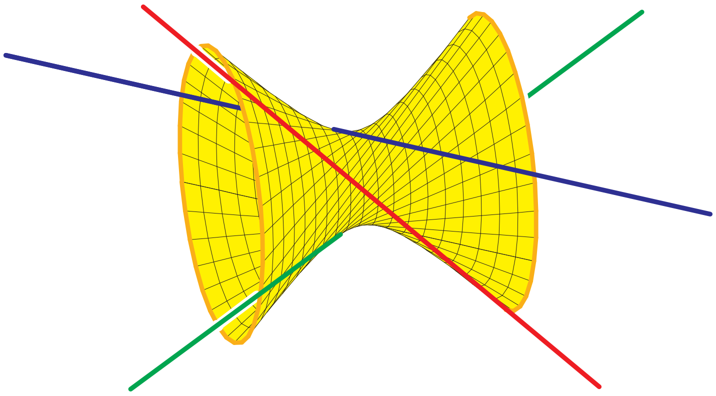


Figure 7.7: Hyperboloid containing three lines.

the other ruling (which is drawn on the hyperboloid in Figure 7.7) consists of the lines which meet the three given lines.

The fourth line will meet the hyperboloid in two points (the hyperboloid is defined by a quadratic polynomial). Through each point of intersection there will be one line in the second family, and that line will meet our four given lines. In this way, we see that the answer to the question is 2. Note that the fourth line may be drawn so that it meets the hyperboloid in two real points, and both solution lines will be real when this happens.

Let $\mathrm{Gr}(p, m{+}p)$ be the *Grassmannian* of p-dimensional linear subspaces (*p-planes* planes in an $m{+}p$-dimensional vector space, which is an algebraic manifold of dimension mp. The Schubert calculus involves fairly general incidence conditions imposed on p-planes H. These general conditions are imposed by *flags*, which are sequences of linear subspaces, one contained in the next. More specifically, a flag $F_\bullet$ is a sequence

$$F_\bullet \;:\; F_1 \;\subset\; F_2 \;\subset\; \cdots \;\subset\; F_{m+p-1} \;\subset\; F_{m+p}\,,$$

where F_i is a linear subspace of $m+p$-space having dimension i.

These general conditions are succinctly indexed by sequences $\alpha\colon 1 \leq \alpha_1 < \alpha_2 < \cdots < \alpha_p < m+p$ of integers, written $\alpha \in \binom{[m+p]}{p}$. The set of all p-planes which satisfy the condition α imposed by the flag $F_\bullet$ is a *Schubert variety*, defined by

$$X_\alpha F_\bullet \;:=\; \{H \in \mathrm{Gr}(p, m+p) \;\mid\; \dim H \cap F_{\alpha_j} \geq j \quad \text{for} \quad j = 1, \ldots, p\}\,.$$

This is a subvariety of the Grassmannian of dimension $|\alpha| := \alpha_1 - 1 + \alpha_2 - 2 + \cdots + \alpha_p - p$. Its codimension is $mp - |\alpha|$.

If K is a linear subspace of codimension p and $F_\bullet$ is a flag with $F_m = K$, then

$$X_{\square} F_\bullet \;:=\; X_{(m,m+2,\ldots,m+p)} F_\bullet \;=\; \{H \;\mid\; H \cap K \neq \{0\}\}\,.$$

The condition that H meets K is is called a *simple Schubert condition* and $X_{\square} F_\bullet$ is a *simple Schubert variety*. More generally, a *special Schubert variety* of codimension a is

$$X_a F_\bullet \;:=\; X_{(m+1-a,m+2,\ldots,m+p)} F_\bullet \;=\; \{H \;\mid\; H \cap F_{m+1-a} \neq \{0\}\}\,.$$

Theorem 7.4 *Given sequences $\alpha^1, \alpha^2, \ldots, \alpha^n$ with $\sum_i (mp - |\alpha^i|) = mp$, there exist real flags $F^1_\bullet, F^2_\bullet, \ldots, F^n_\bullet$ such that the intersection*

$$\bigcap_{i=1}^{n} X_{\alpha^i} F^i_\bullet$$

is transverse with all points real.

The condition that $\sum_i (mp - |\alpha^i|) = mp$ implies that the expected dimension of the intersection is zero. If the flags $F^i_\bullet$ are in general position, then the Kleiman-Bertini Theorem [61] implies that the intersection is in fact transverse and zero-dimensional, so there will be finitely many p-planes satisfying the incidence conditions α^i imposed by general flags, and these p-planes are exactly the points in the intersection.

Theorem 7.4 was proved in several stages. First, when $p = 2$ [88], and then for any p, but only for special Schubert conditions [89], and then finally for general Schubert conditions by Vakil [106].

7.4 Quantum Schubert calculus

Given points $s_1, s_2, \ldots, s_{d(m+p)+mp} \in \mathbb{P}^1$ and m-planes $K_1, K_2, \ldots, K_{d(m+p)+mp}$ in $m+p$-space, there are finitely many rational curves $\gamma\colon \mathbb{P}^1 \to \mathrm{Gr}(p, m+p)$ of degree d so that

$$\gamma(s_i) \;\cap\; K_i \;\neq\; \{0\} \qquad i = 1, 2, \ldots, d(m+p) + mp\,. \tag{7.1}$$

These are *simple quantum Schubert conditions*. More generally, one could (but we will not) impose the condition that the p-plane $\gamma(s_i)$ lie in some predetermined Schubert variety. The number of solutions to such problems are certain Gromov-Witten invariants of the Grassmannian, and may be computed by the quantum Schubert calculus [8, 52, 84, 105].

Theorem 7.5 ([90]) *There exist real points $s_1, s_2, \ldots, s_{d(m+p)+mp} \in \mathbb{RP}^1$ and real m-planes $K_1, K_2, \ldots, K_{d(m+p)+mp}$ in $m+p$-space so that every rational curve $\gamma\colon \mathbb{P}^1 \to \mathrm{Gr}(p, m+P)$ of degree d satisfying* (7.1) *is real.*

7.5 Theorem of Mukhin, Tarasov, and Varchenko

In May of 1995, Boris Shapiro communicated to Frank Sottile a remarkable conjecture that he and his brother Michael had made concerning reality in the Schubert calculus. They conjectured that there would only be real points in a zero-dimensional intersection of Schubert varieties given by flags osculating the rational normal curve. Subsequent computation [77, 91] gave strong evidence for the conjecture and revealed that the intersection should be transverse. Partial results were obtained [89, 33], and the full conjecture was proven by Mukhin, Tarasov, and Varchenko [71]. Subsequently, two other proofs have appeared. The first uses completely different methods [43] and the second is related to the original proof, it includes a proof of transversality [72].

This *Shapiro Conjecture* has been a motivating conjecture for the study of reality in the Schubert calculus with several interesting (and as-yet-unproven) generalizations that we will discuss in subsequent sections. Let γ be the rational normal (or moment) curve in $\mathbb{R}^{m+p}$, which we will take to be the image of the map

$$\gamma(t) \;=\; (1,\ t,\ t^2,\ \ldots,\ t^{m+p-1}) \;\in\; \mathbb{R}^{m+p}\,,$$

defined for $t \in \mathbb{R}$ (and in fact for $t \in \mathbb{RP}^1$). Given a point $t \in \mathbb{R}$, the *osculating flag* $F_\bullet(t)$ is the flag of subspaces whose i-plane is the linear span of the first i derivatives of γ, evaluated at t

$$F_i(t) \;:=\; \mathrm{Span}\{\gamma(t),\ \gamma'(t),\ \gamma''(t),\ \ldots,\ \gamma^{(i-1)}(t)\}\,.$$

Here is the strongest form of the Shapiro conjecture that has been proven.

Theorem 7.6 *If $\alpha^1, \alpha^2, \ldots, \alpha^n \in \binom{[m+p]}{p}$ satisfy $\sum_i (mp - |\alpha^i|) = mp$, then, for every choice of n distinct points $s_1, s_2, \ldots, s_n \in \mathbb{RP}^1$, the intersection*

$$\bigcap_{i=1}^{n} X_{\alpha^i} F_\bullet(s_i)$$

is transverse with all points real.

It is instructive to consider this for the problem of two lines that we saw in Section 7.3. An osculating line will be a tangent line to the rational normal curve γ. The Shapiro Conjecture asserts that given four lines which are tangent to the rational normal curve at real points, there will be two lines meeting all four, and they will be real. As before, first consider three lines tangent to the rational normal curve. They will lie in one ruling of a quadric surface, and the other ruling consists of the lines meeting all four.

For example, if we let $\gamma(t) = (6t^2 - 1, \frac{7}{2}t^3 + \frac{3}{2}t, -\frac{1}{2}t^3 + \frac{3}{2}t)$, and consider tangent lines $\ell(-1)$, $\ell(0)$, and $\ell(1)$ to γ,

$$\ell(\pm 1) \;:\; (5, \pm 5, 1) + u(\pm 1, 1, 0) \qquad \ell(0) \;:\; (-1, 0, 0) + u(0, 1, 1) \qquad u \in \mathbb{R}\,,$$

then the quadric is $x^2 - y^2 + z^2 = 1$. We display this in Figure 7.8, together with the rational normal curve. Since any three real points on any real rational normal curve may be carried to any three real points on any other real rational normal curve by an automorphism of projective 3-space, we can assume that we are in the situation of Figure 7.8.

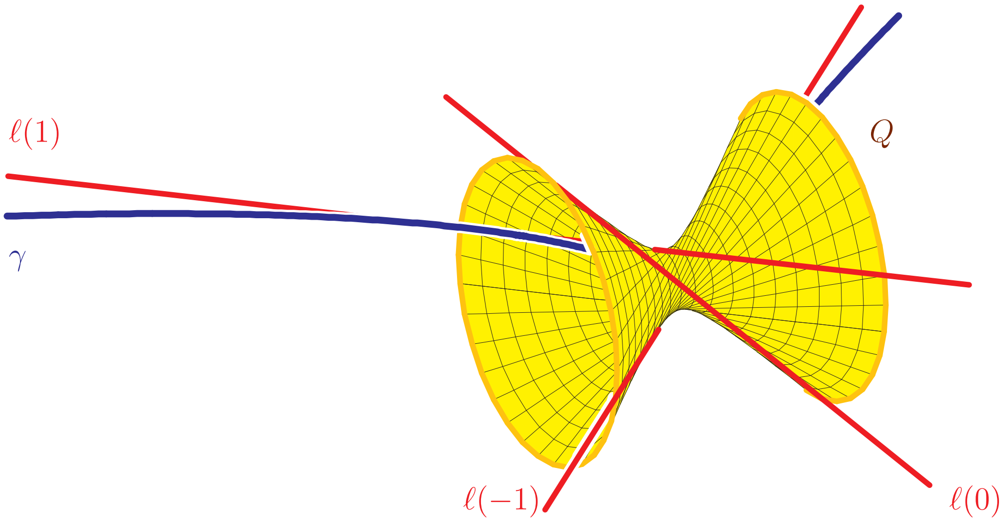


Figure 7.8: Quadric containing three lines tangent to the rational normal curve.

The rational normal curve γ (a cubic) meets the quadric tangentially at the points where it is tangent to the lines. Since the total multiplicity of its intersection with the quadric is six, these are its only points of contact. In particular, γ always lies on one side of the quadric, looping around on the inside. In this way we can see that a fourth tangent line to the rational normal cuirve must meet the quadric in two real points, giving the two real transversals asserted by Shapiro's Conjecture.

It is no loss to assume that the fourth tangent line is tangent at some point $\gamma(v)$ on the arc between points $\gamma(-1)$ and $\gamma(0)$. As illustrated in Figure 7.9, the tangent line $\ell(v)$ does meet meet the hyperboloid in two real points.

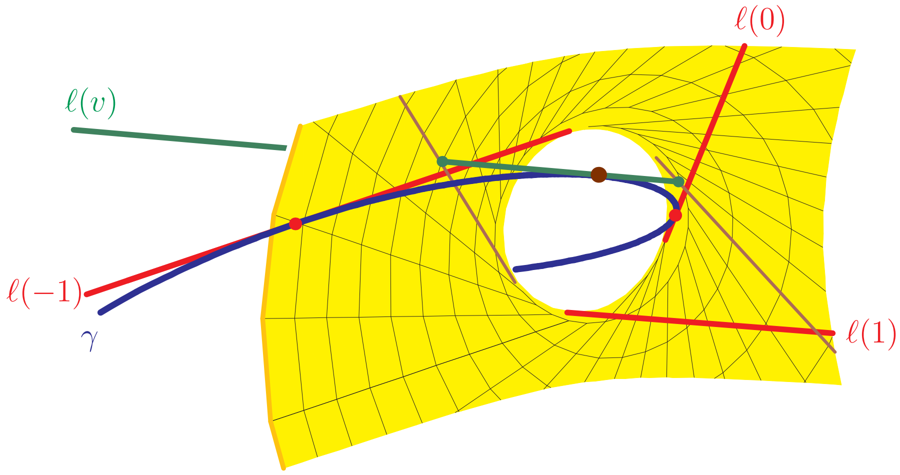


Figure 7.9: The configuration in the throat of quadric.

Chapter 8: The Shapiro Conjecture for Grassmannians

In Chapter 7, we considered the rational normal curve γ in $\mathbb{R}^{m+p}$, which we took to be the image of the map

$$\gamma(t) \;=\; (1,\ t,\ t^2,\ \ldots,\ t^{m+p-1}) \;\in\; \mathbb{R}^{m+p}\,,$$

defined for $t \in \mathbb{R}$ (in fact for $t \in \mathbb{RP}^1$). For a point $t \in \mathbb{R}$, the osculating flag $F_\bullet(t)$ is the flag of subspaces whose i-plane is the linear span of $\gamma(t)$ and the the first $i-1$ derivatives of γ, evaluated at t,

$$F_i(t) \;:=\; \mathrm{Span}\{\gamma(t),\ \gamma'(t),\ \gamma''(t),\ \ldots,\ \gamma^{(i-1)}(t)\}\,.$$

We work in the Grassmannian $\mathrm{Gr}(p, m+p)$ of p-planes in $(m+p)$-space. Recall from Section 7.3 that Schubert varieties are indexed by Schubert conditions, which are sequences $\alpha\colon 0 < \alpha_1 < \alpha_2 < \cdots < \alpha_p \leq m+p$ of integers. Given a flag $F_\bullet$, the corresponding Schubert variety is

$$X_\alpha F_\bullet \;=\; \{H \in \mathrm{Gr}(p, m+p) \;\mid\; \dim H \cap F_{\alpha_j} \geq j \quad \text{for} \quad j = 1, \ldots, p\}\,.$$

This has dimension $|\alpha| := \sum_j (\alpha_j - j)$.

Recall the Shapiro conjecture for Grassmannians (Theorem of Mukhin, Tarasov, and Varchanko), which we shall investigate in this chapter.

Theorem 7.6 *If $\alpha^1, \alpha^2, \ldots, \alpha^n$ are Schubert conditions with $\sum_i (mp - |\alpha^i|) = mp$, then, for every choice of n distinct points $s_1, s_2, \ldots, s_n \in \mathbb{RP}^1$, the intersection*

$$\bigcap_{i=1}^{n} X_{\alpha^i} F_\bullet(s_i)$$

is transverse with all points real.

8.1 The Wronski map

The Shapiro conjecture has another formulation in terms of the Wronski map. Let $f_1(t), \ldots, f_k(t)$ be polynomials of degree d. Their *Wronskian* is

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

which is a polynomial of degree $k(d+1-k)$, when the polynomials $f_1(t), f_2(t), \ldots, f_k(t)$ are generic among all polynomials of degree d.

Up to a scalar factor, this Wronskian depends only upon the linear span of the polynomials $f_1(t), f_2(t), \ldots, f_k(t)$. Removing these ambiguities, we get the *Wronski map*

$$\mathrm{Wr} \;:\; \mathrm{Gr}(k, d{+}1) \;\longrightarrow\; \mathbb{P}^{k(d+1-k)}\,, \tag{8.1}$$

where $\mathrm{Gr}(k, d{+}1)$ is the Grassmannian of k-planes in the space of polynomials of degree d, and $\mathbb{P}^{k(d+1-k)}$ is the space of polynomials of degree $k(d{+}1{-}k)$, modulo scalars.

8.1.1 Some linear algebra

Let us begin with the moment (rational normal) curve. For $t \in \mathbb{C}$, set

$$\gamma(t) \;=\; (1, t, t^2, \ldots, t^d) \;\in\; \mathbb{C}^{d+1}\,.$$

Let $\Gamma = \Gamma(t)\colon \mathbb{C}^k \to \mathbb{C}^{d+1}$ be the map such that

$$\Gamma(\mathbf{e}_i) \;=\; \gamma^{(i-1)}(t)\,, \tag{8.2}$$

the $(i{-}1)$-th derivative of γ. (We will take $\mathbf{e}_1, \mathbf{e}_2, \ldots$ to be the standard basis vectors of the vector space in which we are working.)

A polynomial f corresponds to a linear form (also written f):

$$f \;:\; \mathbb{C}^{d+1} \;\longrightarrow\; \mathbb{C} \qquad \text{so that} \qquad f\circ\gamma(t) \;=\; f(t)\;.$$

The matrix in the definition of the Wronskian is the matrix of the composition

$$\mathbb{C}^k \;\xrightarrow{\;\Gamma(t)\;}\; \mathbb{C}^{d+1} \;\xrightarrow{\;\Psi\;}\; \mathbb{C}^k\,, \tag{8.3}$$

where the columns of Ψ are the linear forms defining the polynomials $f_1, f_2, \ldots, f_k$. Now suppose that H is a basis for the kernel of the map Ψ. We consider H to be a map $\mathbb{C}^{d+1-k} \xrightarrow{\;H\;} \mathbb{C}^{d+1}$, which we sum with $\Gamma(t)$ to get a map

$$\mathbb{C}^{d+1} \;=\; \mathbb{C}^k \oplus \mathbb{C}^{d+1-k} \;\xrightarrow{\;[\Gamma(t):H]\;}\; \mathbb{C}^{d+1}\,.$$

This map is invertible if and only if the composition (8.3) is invertible. Thus, up to a constant, we have

$$\mathrm{Wr}\,(f_1, f_2, \ldots, f_k) \;=\; \det[\Gamma(t) \;:\; H]\,, \tag{8.4}$$

as both are polynomials of the same degree with the same roots. (Strictly speaking, we need the Wronskian to have distinct roots for this argument. The general case follows via a limiting argument.)

We obtain a useful formula for the Wronskian when we expand the determinant (8.4) along the columns of $\Gamma(t)$

$$\mathrm{Wr}\,(f_1, f_2, \ldots, f_k) \;=\; \det[\Gamma(t) \;:\; H] \;=\; \sum_\alpha (-1)^{|\alpha|} p_\alpha(\Gamma(t)) \cdot p_{\alpha^c}(H)\,. \tag{8.5}$$

Here, the sum is over all choices $\alpha\colon \{\alpha_1 < \alpha_2 < \cdots < \alpha_k\} \subset [d+1]$ of k distinct rows of the matrix $\Gamma(t)$, $\alpha^c = [d+1] - \alpha$ are the complimentary rows of H, and $p_\alpha(\Gamma(t))$ is the αth maximal minor of $\Gamma(t)$, which is the determinant of the submatrix of $\Gamma(t)$ formed by the rows in α, and similarly for $p_{\alpha^c}(H)$. Write $\binom{[d+1]}{k}$ for this set of indices α.

There is a similar expansion for the Wronskian using the composition (8.3). Take the top exterior power ($\wedge^k$) of this composition,

$$\mathbb{C} \;=\; \wedge^k\mathbb{C}^k \;\xrightarrow{\wedge^k\Gamma(t)}\; \wedge^k\mathbb{C}^{d+1} \;\xrightarrow{\wedge^k\Psi}\; \wedge^k\mathbb{C}^k \;=\; \mathbb{C}\,,$$

where we have used the ordered basis of $\mathbb{C}^k$ so that $\wedge^k\mathbb{C}^k = \mathbb{C}\cdot\mathbf{e}_1\wedge\mathbf{e}_2\wedge\ldots\wedge\mathbf{e}_k$, which identifies $\wedge^k\mathbb{C}^k$ with $\mathbb{C}$. Then $\wedge^k\Gamma(t)$ is simply a vector in $\wedge^k\mathbb{C}^{d+1}$ and $\wedge^k\Psi$ is a covector for $\wedge^k\mathbb{C}^{d+1}$. If we use the basis $\mathbf{e}_\alpha := \mathbf{e}_{\alpha_1}\wedge\mathbf{e}_{\alpha_2}\wedge\ldots\wedge\mathbf{e}_{\alpha_k}$ for $\wedge^k\mathbb{C}^{d+1}$, where $\alpha\in\binom{[d+1]}{k}$, then we see that the Wronskian has the form

$$\mathrm{Wr}\,(f_1, f_2, \ldots, f_k) \;=\; \sum_{\alpha\in\binom{[d+1]}{k}} p_\alpha(\Gamma(t))\cdot p_\alpha(\Psi)\,.$$

Here, $p_\alpha(\Gamma(t))$ and $p_\alpha(\Psi)$ are the αth coordinates of the corresponding vector/covector (which are the αth maximal minors of the corresponding matrices). Equating these two expressions for the Wronskian gives the equality (again up to a constant)

$$\sum_{\alpha\in\binom{[d+1]}{k}} (-1)^{|\alpha|} p_\alpha(\Gamma(t))\cdot p_{\alpha^c}(H) \;=\; \sum_{\alpha\in\binom{[d+1]}{k}} p_\alpha(\Gamma(t))\cdot p_\alpha(\Psi)\,.$$

This argument does not use much about the matrix $\Gamma(t)$, besides that it depends upon an indeterminate t. Letting $\Gamma(t)$ be a matrix of indeterminates proves an interesting (and well-known) matrix identity.

Proposition 8.1 *Suppose that Ψ and H are matrices of format n by k and $(n-k)$ by n such that $\Psi\circ H = 0$ (the image of H is the kernel of Ψ. Then there is a constant C so that*

$$p_\alpha(\Psi) \;=\; C\cdot(-1)^{|\alpha|} p_{\alpha^c}(H)\ ,$$

for all $\alpha\in\binom{[n]}{k}$.

8.1.2 Connection to Schubert calculus

We explore some geometric consequences of the determinantal formulas (8.4) and (8.5). Let H be the column space of the matrix H, which is the kernel of the map Ψ. Then H is a point in the Grassmannian $\mathrm{Gr}(p, m+p)$, where we have set $p := d+1-k$ and $m := k$. From the definition (8.2) of Γ, we see that the column space of the matrix $\Gamma(t)$ is the m-plane $F_m(t)$ osculating the rational normal curve γ at the point $\gamma(t)$. From (8.4) and (8.5), we see that s is a zero of the Wronskian $\Phi(t) := \mathrm{Wr}\,(f_1(t), f_2(t), \ldots, f_k(t))$ if and only if

$$0 \;=\; \det[\Gamma(s)\ :\ H]\,.$$

This implies that there is a linear dependence among the columns of this matrix and thus there is a nontrivial intersection between the subspaces $F_m(s)$ and H.

Suppose that a polynomial $\Phi(t)$ has distinct zeroes $s_1, s_2, \ldots, s_{mp}$ ($mp = k(d+1-k)$). Then the columns of the matrix H are linear forms which cut out the linear span of polynomials $f_1(t), f_2(t), \ldots, f_k(t)$ of degree d having Wronskian $\Phi(t)$

$\Leftrightarrow$ the p-plane H meets the m-plane $F_m(s_i)$ nontrivially for each $i = 1, 2, \ldots, mp$,

$\Leftrightarrow$ H lies in the Schubert variety $X_{\square}F_\bullet(s_i)$ for each $i = 1, 2, \ldots, mp$.

If the roots $s_1, \ldots, s_{mp}$ are all real, then the Shapiro Conjecture (Theorem 7.6) asserts that all such p-planes H are real, and there are the expected number of them. In this way, Theorem 1.9 is a consequence of Theorem 7.6. Recall that a univariate polynomial $\Phi(t)$ is *hyperbolic* if it has only real roots.

Theorem 1.9 *If $\Phi(t) \in \mathbb{P}^{k(d+1-k)}$ is a hyperbolic polynomial with no multiple roots, then*

1 *$\Phi(t)$ is a regular value of the Wronski map* (8.1), *and*

2 *if $f_1(t), f_2(t), \ldots, f_k(t)$ are polynomials of degree d with Wronskian $\Phi(t)$, then their linear span is real.*

Recall from Chapter 1 the formula for the degree of the Wronski map,

$$\#_{d,k} \;=\; \frac{1!2!\cdots(k-1)!\cdot[k(d+1-k)]!}{(d+1-k)!(d+2-k)!\cdots d!}\,, \tag{1.4}$$

which is the number of inverse images of a regular value of the Wronski map.

8.2 Asymptotic form of the Shapiro Conjecture is true

It is not too hard to show that the conclusion of the Shapiro conjecture holds for *some* $s_1, \ldots, s_{mp} \in \mathbb{R}$. We give a (by now) standard argument for the following asymptotic form of the Shapiro conjecture, which was given in [89].

Theorem 8.2 *There exist real numbers $s_1, s_2, \ldots, s_{mp}$ such that*

$$X_{\square}F_\bullet(s_1) \cap X_{\square}F_\bullet(s_2) \cap \cdots \cap X_{\square}F_\bullet(s_{mp})$$

is transverse with all points real.

The proof is a version of Schubert's principle of degeneration to special position and the same ideas can be used to establish similar results for other flag manifolds and related varieties, such as Theorem 7.5.

Set $p := d{+}1 - k$ and $m := k$, and then the expansion (8.5) becomes (up to a sign)

$$\det[\Gamma(t) \;:\; H] \;=\; \sum_{\alpha \in \binom{[m+p]}{p}} p_\alpha(H)\cdot(-1)^{|\alpha|} p_{\alpha^c}(\Gamma(t))\,.$$

(Note that we have interchanged α with α^c.) We first expand the minor $p_{\alpha^c}(\Gamma(t))$. Let $\alpha\colon \alpha_1 < \alpha_2 < \cdots < \alpha_m \in \binom{[m+p]}{m}$. Observe that the determinant

$$\det\begin{bmatrix} t^{\alpha_1-1} & (\alpha_1-1)t^{\alpha_1-2} & \ldots & \frac{(\alpha_1-1)!}{(\alpha_1-m)!}t^{\alpha_1-m} \\ t^{\alpha_2-1} & (\alpha_2-1)t^{\alpha_2-2} & \ldots & \frac{(\alpha_2-1)!}{(\alpha_2-m)!}t^{\alpha_2-m} \\ \vdots & \vdots & \ddots & \vdots \\ t^{\alpha_m-1} & (\alpha_m-1)t^{\alpha_m-2} & \ldots & \frac{(\alpha_m-1)!}{(\alpha_m-m)!}t^{\alpha_m-m} \end{bmatrix}$$

is equal to

$$t^{|\alpha|}\cdot\det\begin{bmatrix}1 & (\alpha_1-1) & \dots & \frac{(\alpha_1-1)!}{(\alpha_1-m)!}\\ 1 & (\alpha_2-1) & \dots & \frac{(\alpha_2-1)!}{(\alpha_2-m)!}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & (\alpha_m-1) & \dots & \frac{(\alpha_m-1)!}{(\alpha_m-m)!}\end{bmatrix} \quad=\quad t^{|\alpha|}\cdot\det\begin{bmatrix}1 & \alpha_1 & \dots & \alpha_1^{m-1}\\ 1 & \alpha_2 & \dots & \alpha_2^{m-1}\\ \vdots & \vdots & \ddots & \vdots\\ 1 & \alpha_m & \dots & \alpha_m^{m-1}\end{bmatrix}.$$

(The second equality is via column operations.) We recognize this last determinant as the Van der Monde, $\prod_{i<j}(\alpha_j-\alpha_i)$. Write V_α for the product

$$(-1)^{|\alpha|}\cdot\prod_{i<j}(\alpha_j-\alpha_i)\,.$$

Since $|\alpha^c| = mp-|\alpha|$, we obtain the expansion for the Wronskian (up to a global sign).

$$\det[\Gamma(t)\ :\ H] \;=\; \sum_{\alpha\in\binom{[m+p]}{p}} t^{mp-|\alpha|}\,V_{\alpha^c}\,p_\alpha(H)\,. \tag{8.6}$$

Observe that if we write $y_\alpha := V_{\alpha^c}p_\alpha(H)$, then (8.6) becomes

$$\sum_{\alpha\in\binom{[m+p]}{p}} t^{mp-|\alpha|}\,y_\alpha\,.$$

In particular, the coefficient of y_α depends only upon the rank, $|\alpha|$, of $\alpha\in\binom{[m+p]}{p}$. Compare this to the Wronski polynomial (6.2) of Section 6.1, and compare the original polynomial (8.6) to (6.3).

8.2.1 Schubert varieties

We take transposes of all matrices, replacing column vectors by row vectors. Let $H\in \mathrm{Gr}(p,m+p)$ be represented as the row space of a p by $(m+p)$-matrix, and apply Gaussian elimination with pivoting from top to bottom and right to left to H to obtain a unique representative matrix of the form

$$H \;=\; \begin{pmatrix} * & \cdots & * & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\ * & \cdots & * & 0 & * & \cdots & * & 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0\\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & & \ddots & & \vdots & \vdots & & \vdots\\ * & \cdots & * & 0 & * & \cdots & * & \vdots & * & \cdots & * & 1 & 0 & \cdots & 0\end{pmatrix}. \tag{8.7}$$

Here, the entries $*$ indicate an unspecified element of our field ($\mathbb{R}$ or $\mathbb{C}$).

The set of columns containing the leading 1s (pivots) is a discrete invariant of the linear subspace H. Let $\alpha\colon \alpha_1<\alpha_2<\cdots<\alpha_p$ be the positions of the pivots, that is, α_i is the column of the leading 1 in row i. Observe that $p_\beta(H)=0$ unless $\beta_i\le\alpha_i$ for every i. This coordinatewise comparison defines the *Bruhat order* on the indices $\alpha\in\binom{[m+p]}{p}$. The set of linear spaces whose row reduced echelon forms (8.7) have pivots in the columns of α forms a topological cell of dimension $|\alpha|$, called the *Schubert cell*. The undetermined entries $*$ in (8.7) show that it is isomorphic to $\mathbb{A}^{|\alpha|}$, the affine space of dimension $|\alpha|$.

We ask: Which linear spaces are in the closure of the Schubert cell? For the answer, let M_α be the set of matrices with full rank p where the entries in row i are undetermined up to column α_i, and then 0 thereafter. These matrices have the form

$$\begin{pmatrix} * & \cdots & * & 0 & \cdots & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ * & \cdots & * & * & \cdots & * & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots & & \ddots & & 0 & \cdots & 0 \\ * & \cdots & * & * & \cdots & * & * & \cdots & * & 0 & \cdots & 0 \end{pmatrix},$$

where the last undetermined entry $*$ in row i occurs in column α_i. This is a closed subset of the set of p by $(m+p)$-matrices of full rank p. The pivots β of a matrix M in M_α will occur weakly to the left of the columns indexed by α, so that $\beta \leq \alpha$, and all possibilities β can occur.

In particular, this shows that the set of p-planes H parametrized by matrices in M_α is the union of the Schubert cells indexed by β for $\beta \leq \alpha$ in the Bruhat order. This is a closed subset of the Grassmannian, in fact it is one of the *Schubert varieties* defined in Chapter 7.

Let $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{m+p}$ be basis vectors corresponding to the columns of our matrices. For each $i = 1, \ldots, m+p$ let F_i be the linear span of the vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_i$. From the form of matrices in M_α, we see that if H is the row space of a matrix in M_α, then we have

$$\dim H \cap F_{\alpha_j} \ \geq\ j \qquad \text{for } j = 1, \ldots, p\,.$$

These dimension inequalities define the Schubert variety $X_\alpha F_\bullet$. Note that if $H \in X_\alpha F_\bullet$, then $p_\beta(H) = 0$ unless $\beta \leq \alpha$.

The key lemma in our proof of Theorem 8.2 is due essentially to Schubert [81]:

$$X_\alpha F_\bullet \cap \{H \mid p_\alpha(H) = 0\} \ =\ \bigcup_{\beta \lessdot \alpha} X_\beta F_\bullet\,. \tag{8.8}$$

Here $\beta \lessdot \alpha$ means that $\beta < \alpha$, but there is no index μ in the Bruhat order with $\beta < \mu < \alpha$. This is easy to see set-theoretically, as for $H \in X_\alpha F_\bullet$ we have $p_\beta(H) = 0$ unless $\beta \leq \alpha$.

It is also easy to see that this is true on the generic point of each $X_\beta F_\bullet$ for $\beta \lessdot \alpha$. Fix some index β with $\beta \lessdot \alpha$. Then there is a unique index k with $\beta_k = \alpha_k - 1$, and for all other indices i, $\beta_i = \alpha_i$. Consider the subset of the matrices M_α, where we require the entries in row i and column β_i to be 1, and write x_{k,α_k} for the entry in row k and column α_k

$$\begin{pmatrix} * & \cdots & * & 1 & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & 0 & & \ddots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & * & \cdots & * & 1 & x_{k,\alpha_k} & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & 0 & 0 & & \ddots & & \vdots & \vdots & & \vdots \\ * & \cdots & * & 0 & * & \cdots & * & \vdots & \vdots & * & \cdots & * & 1 & 0 & \cdots & 0 \end{pmatrix}. \tag{8.9}$$

The row spans of these matrices form a dense subset of the Schubert variety $X_\alpha F_\bullet$, and therefore define a coordinate patch for $X_\alpha F_\bullet$. If we set $x_{k,\alpha_k} = 0$, then we get all matrices of the form (8.7), but for the index β.

If H is the row space of a matrix in this set (8.9), then $p_\alpha(H) = x_{k,\alpha_k}$. Thus, on this coordinate patch for $X_\alpha F_\bullet$, the vanishing of the Plücker coordinate p_α cuts out the Schubert variety $X_\beta F_\bullet$, scheme-theoretically. Repeating this local argument for each $\beta \lessdot \alpha$, proves (8.8), at least at the generic point of each component $X_\beta F_\bullet$ (which is sufficient for our purposes). More careful arguments show this is true even at the level of their homogeneous ideals.

We now have everything that we need to prove Theorem 8.2. We use induction on the Bruhat order (sometimes called *Schubert induction*) to prove that there exist real numbers $s_1, s_2, \ldots, s_{mp}$ such that for all $\alpha \in \binom{[m+p]}{p}$,

$$X_\alpha F_\bullet \ \cap \ \bigcap_{i=1}^{|\alpha|} X_{\square} F_\bullet(s_i) \tag{8.10}$$

is transverse with all points of intersection real. Then the statement of Theorem 8.2 is the case $\alpha = m{+}1, m{+}2, \ldots m{+}p$, so that $X_\alpha F_\bullet$ is the Grassmannian.

Observe that when $\alpha = 1, 2, \ldots, p$, then $|\alpha| = 0$ and the Schubert variety $X_\alpha F_\bullet$ consists of the single point $\{F_p\}$. Thus the base case of the induction is trivial, as there is no intersection to contend with.

Remark 8.3 In fact it is not hard to see (it is equivalent to the Plücker formula [73] for rational curves and was noted by Eisenbud and Harris [29, Theorem 2.3]) that the intersection lies in the Schubert cell for the index α. That is, every point H in the intersection (8.10) has row-reduced echelon form (8.7), for the index α.

Suppose that the statement (8.10) is true for all α with $|\alpha| = j$. Let $\alpha \in \binom{[m+p]}{p}$ with $|\alpha| = j{+}1$. Observe that by (8.6) the intersection $X_\alpha F_\bullet \cap X_{\square} F_\bullet(t)$ is defined by the single polynomial equation

$$\sum_\beta t^{mp-|\beta|} V_{\beta^c} p_\beta(H) \ = \ 0 \qquad \text{for } H \in X_\alpha F_\bullet\,.$$

Since $p_\beta(H) = 0$ unless $\beta \leq \alpha$, this becomes

$$\sum_{\beta \leq \alpha} t^{mp-|\beta|} V_{\beta^c} p_\beta(H) \ = \ 0\,.$$

Dividing by the lowest power $t^{mp-|\alpha|}$ of t, this becomes

$$\sum_{\beta \leq \alpha} t^{|\alpha|-|\beta|} V_{\beta^c} p_\beta(H) \ = \ V_{\alpha^c} p_\alpha(H) \ + \ t \cdot \sum_{\beta < \alpha} t^{|\alpha|-|\beta|-1} V_{\beta^c} p_\beta(H) \ = \ 0\,.$$

Since $V_{\alpha^c} \neq 0$, we see that in the limit as $t \to 0$, this equation becomes $p_\alpha(H) = 0$. Using the lemma of Schubert (8.8), we conclude that

$$\lim_{t \to 0} X_\alpha F_\bullet \cap X_{\square} F_\bullet(t) \ = \ X_\alpha F_\bullet \cap \{H \mid p_\alpha(H) = 0\} \ = \ \bigcup_{\beta \lessdot \alpha} X_\beta F_\bullet\,. \tag{8.11}$$

By our induction assumption on j, and the remark that the intersection is contained in the Schubert cell, we see that

$$\Bigl(\bigcup_{\alpha\lessdot\beta} X_\beta F_\bullet\Bigr) \ \cap\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i)$$

is transverse with all points real. By the computation of the limit (8.11), and the observation that transversality is preserved by small perturbations, we see that there is a number $0<\varepsilon_\alpha$ such that if $t\leq\varepsilon_\alpha$ then

$$X_\alpha F_\bullet \ \cap\ X_{\square}F_\bullet(t) \ \cap\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i)$$

is transverse with all points real.

We complete the induction by setting s_{j+1} to be the minimum of the numbers ε_α where $|\alpha|=j+1$. This proves Theorem 8.2.

Similar asymptotic arguments prove reality in other Schubert calculus settings. For example, the result concerning rational curves in Grassmannians of Theorem 7.5 [90], for the classical flag manifolds and for the orthogonal Grassmannian [92].

Remark 8.4 In the proof of Theorem 8.2, we not only showed that the intersection (8.10) is transverse with all points real, but we actually constructed all the points of intersection.

The induction began with $\alpha=1,2,\ldots,p$ so that $|\alpha|=0$ and the Schubert variety $X_\alpha F_\bullet$ consists of the single point $\{F_p\}$. When $|\alpha|=j+1$, the limit

$$\Bigl(\lim_{t\to 0} X_\alpha F_\bullet\cap X_{\square}F_\bullet(t)\Bigr) \ \cap\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i) \ =\ \bigcup_{\beta\lessdot\alpha} X_\beta F_\bullet \ \cap\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i)$$

shows that each point in the intersection (8.10) is connected to a point in

$$\bigcup_{\beta\lessdot\alpha} X_\beta F_\bullet \ \cap\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i) \ =\ \bigcup_{\beta\lessdot\alpha}\Bigl(X_\beta F_\bullet\ \bigcap_{i=1}^{j} X_{\square}F_\bullet(s_i)\Bigr)$$

by a path as t ranges from s_{j+1} to 0. This is a union of previously constructed intersections. Beginning with a previously constructed point, and following the corresponding path constructs points in the intersection (8.10). Following paths in $\mathbb{C}$, and not necessarily in $\mathbb{R}$, leads to the numerical *Pieri homotopy algorithm* for constructing all the points in a fiber of the Wronski map. This was proposed in [49] and implemented in [51].

If $\delta(\alpha)$ is the number of points in the intersection (8.10), then this limiting process also gives the following recursion for this number:

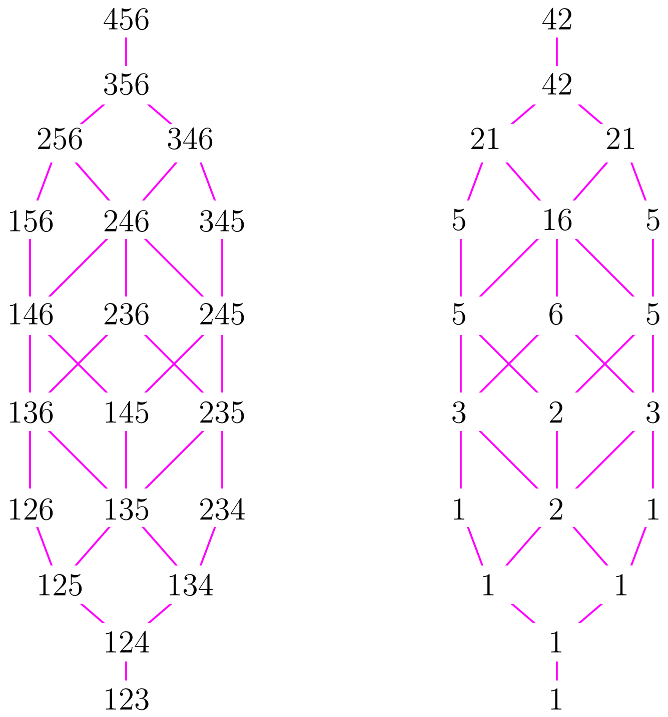
$$\begin{aligned}\delta(1,2,\ldots,p) &= 1\\ \delta(\alpha) &= \sum_{\beta\lessdot\alpha}\delta(\beta)\end{aligned}$$

In fact, Schubert discovered this recursion [80] and used it to compute the number $\delta(567) = 462$ when $d = 6$ and $k = 4$ ($m = 4$ and $p = 3$), which we invite the reader to check is the number given by the formula,

$$\frac{1!2!\cdots(k-1)!\cdot[k(d+1-k)]!}{(d+1-k)!(d+2-k)!\cdots d!}\ , \tag{1.4}$$

which is also due to Schubert, as he solved his recursion to obtain a closed formula. This recursion shows that the number $\delta(\alpha)$ may be interpreted combinatorially as the number of paths from the bottom of the Bruhat order to the element α.

We illustrate the Bruhat order for the case when $m = p = 3$ and the recursion for the numbers $\delta(\alpha)$.



Perhaps we should explain how this inductive proof & the MTV Theorem identifies sheets of the Wronskii cover of the hyperbolic polynomials with Young tableaux? What about the obvious questions concerning the actions of $\mathbb{Z} = \pi_1$ of the configuration space on these sheets?

[

Chapter 9: The Shapiro Conjecture for Rational Functions

Introduction

We continue our study of the Shapiro Conjecture, which asserts that if a polynomial $\Phi(t)$ has degree $k(d+1-k)$ with distinct real roots, and if $f_1(t), f_2(t), \ldots, f_k(t)$ are polynomials of degree d whose Wronskian is $\Phi(t)$, then the linear span of $f_1(t), f_2(t), \ldots, f_k(t)$ is real.

Such a list of polynomials defines a rational curve of degree d in projective space—the image of the map

$$\mathbb{R} \ \ni \ t \ \longmapsto [f_1(t) : f_2(t) : \cdots : f_k(t)] \ \in \ \mathbb{RP}^{k-1}\,.$$

In Chapter 11 we will interpret the Shapiro conjecture in terms of such curves, with particular applications to plane curves. For now, we consider an apparently degenerate case, that of rational functions when—$k = 2$.

9.1 The Shapiro Conjecture for rational functions

Eremenko and Gabrielov [33] originally gave a proof of the Shapiro Conjecture when $k = 2$ using essentially the uniformization theorem from complex analysis. This proof was also complicated and challenging for *this* reader. They subsequently found a second, significantly more elementary proof [31]. This lecture will present the main points of this new proof. Fortunately, the most difficult part was already accomplished in Chapter 8.

9.1.1 Preliminaries on rational functions

A point p in the Grassmannian $\mathrm{Gr}(2, d+1)$ defines an equivalence class of rational functions $r\colon \mathbb{P}^1 \to \mathbb{P}^1$ (modulo automorphisms of the target $\mathbb{P}^1$). Working with this equivalence is slightly awkward, so we work (mostly) instead on the *Stiefel manifold*, St_d, which is a $GL(2, \mathbb{R})$-fiber bundle over $\mathrm{Gr}(2, d+1)$.

The points of St_d are pairs of nonproportional polynomials $(f(t), g(t))$ of degree at most d. Hence St_d is an open subset of $\mathbb{R}^{2d+2}$, where the coordinates are the coefficients of f and g. The association $\mathrm{St}_d \ni (f, g) \longmapsto f/g$ defines a map (written ρ) from St_d to the space of rational functions. While this map is not continuous as a map of spaces, it does have the weak continuity property given in Proposition 9.1, whose proof is elementary.

Let $Z \subset \mathrm{St}_d$ be the locus of pairs (f, g) with either

$$\deg \gcd(f, g) > 0 \quad \text{ or } \quad \deg(f) < d \text{ and } \deg(g) < d.$$

That is, f and g either have a common root in $\mathbb{C}$ or else a common root at ∞, and thus define a rational function f/g of degree less than d.

Proposition 9.1 *Let $\{p_j \mid j \in \mathbb{N}\} \subset \mathrm{St}_d \setminus Z$ be a sequence of points that converges to some point $p \in Z$. Then $p = (f, g)$ and write q for the greatest common factor of f and g, $q := \mathrm{gcf}(f, g)$. Let $x_1, x_2, \ldots, x_k$ be the roots of q (including ∞ if $\deg(f)$ and $\deg(g)$ are both less than d).*

Then the sequence of functions $\{\rho(p_j) \mid j \in \mathbb{N}\}$ converges to $\rho(p)$ uniformly on compact subsets of $\mathbb{CP}^1 \setminus \{x_1, x_2, \ldots, x_k\}$.

Here, the topology on St_d is the subspace topology, coming from its realization as an open subset of $\mathbb{R}^{2d+2}$, that is, from the coefficients of f and g. This nearly trivial proposition is half of the engine of this new proof of Eremenko and Gabrielov. The other half is the asymptotic proof of Shapiro's conjecture (Theorem 8.2 from Chapter 8).

Let R_d be the set of nonconstant real rational functions of degree at most d, all of whose critical points are real. If $r \in R_d$, then $r^{-1}(\mathbb{RP}^1) \subset \mathbb{CP}^1$ defines an embedded (multi-) graph G with the following properties:

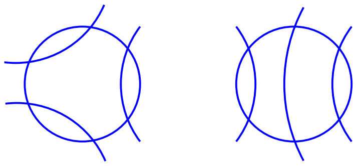
(i) G is stable under complex conjugation and $\mathbb{RP}^1 \subset G$.

(ii) The vertices of G lie on $\mathbb{RP}^1$ and are the critical points of the rational function r. The valence of a vertex is even and it equals twice the local degree of r at that critical point.

The set-theoretic difference $\mathbb{CP}^1 \setminus \Gamma$ is a union of $2d'$ cells, where d' is the degree of r. The closure of each cell is homeomorphic to a disc, and the boundary of each cell maps homeomorphically to $\mathbb{RP}^1$. This is because the cells (and their closures) are the inverse images of one of the two discs in $\mathbb{CP}^1 \setminus \mathbb{RP}^1$ (or their closures), and there are no critical points in the interior of any cell. We deduce the following additional property of these multi-graphs.

(iii) No edge of G can begin and end at the same vertex.

Here are two pictures of such embedded (multi-) graphs. We have drawn $\mathbb{RP}^1$ as a circle, the point $\sqrt{-1}$ is at the center, its complex conjugate cousin $-\sqrt{-1}$ is the point at infinity, and complex conjugation is inversion in the circle.

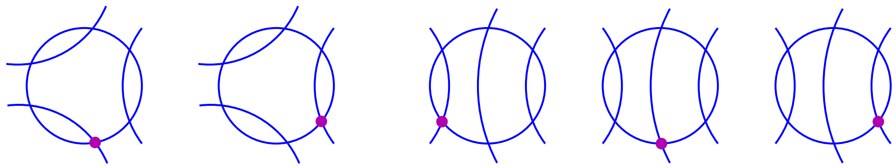


We will be seeking to analytically continue rational functions whose Wronskians lie in a curve of polynomials $\{\Phi_z(t) \mid z \in [0, 1]\}$ where each $\Phi_z(t)$ has degree $2d - 2$ with distinct real roots $s_1(z), s_2(z), \ldots, s_{2d-2}(z)$, and where each s_i is a continuous function of z. The vertices of the graphs will be labeled by these functions, or equivalently by the numbers $1, 2, \ldots, n$. Since the relative order of these roots $s_1(z), s_2(z), \ldots, s_{2d-2}(z)$ does not change as

z varies (because each polynomial $\Phi_z(t)$ has distinct roots), we can capture this information by labeling ony one root, say $s_1(z)$ (which is a vertex of the corresponding graph), and assuming that the roots are ordered in a manner consistent with a fixed orientation of $\mathbb{RP}^1$. It is these labeled graphs that we wish to consider up to isotopy (deformation in $\mathbb{CP}^1$).

Definition 9.2 A *net* is an (isotopy-) equivalence class of such embedded multi-graphs in $\mathbb{CP}^1$ satisfying (i), (ii), and (iii), with a distinguished vertex.

Here are the five nets with 6 vertices, each with local degree 2.



Proposition 9.1 implies the following continuity result about nets.

Proposition 9.3 *Let $\{p_j\} \subset \mathrm{St}_d$ be a convergent sequence with limit p. Then the sets $\{\rho(p_j)^{-1}(\mathbb{RP}^1)\}$ converge in the Hausdorff topology to the set $\{\rho(p)^{-1}(\mathbb{RP}^1)\}$.*

We deduce two corollaries from this proposition.

Corollary 9.4 *Suppose that $\{r_z \mid z \in [0,1]\}$ is a continuous path in R_d where each r_z has $2d-2$ critical points. Let $v_0(z)$ be a continuous function of z which is equal to a critical point of r_z, for each z. Then the net of the pair*

$$\left(r_z^{-1}(\mathbb{RP}^1),\, v_0(z)\right)$$

does not depend upon z.

Corollary 9.5 *Suppose that $\{p_z \mid z \in [0,1]\}$ is a continuous path in the Stiefel manifold St_d. Suppose that for $0 < z$ we have $\rho(p_z) \in R_d$ and $\rho(p_z)$ has distinct critical points $v_0(z), v_1(z), \ldots, v_n(z)$, where $v_i(z)$ is a continuous function of z, and also that, at $z = 0$, we have $v_0(0) = v_1(0)$, but all other critical points are distinct. Then the degree of $\rho(p_z)$ is constant for $z \in (0,1]$, and $\deg(\rho(p_0)) < \deg(\rho(p_1))$ if and only if the net of $\rho(p_1)$ has an edge between $v_0(1)$ and $v_1(1)$.*

Proof. The degree of a rational function $r \in R_d$ is one-half the number of two-cells in $\mathbb{CP}^1 \setminus r^{-1}(\mathbb{RP}^1)$. Set $r_z := \rho(p_z)$. The only way for the number of cells in the complement of the net of r_z to change at some $z_0 \in [0,1]$ would be if some edge of $r_z^{-1}(\mathbb{RP}^1)$ collapsed as $z \to z_0$. Since the vertices of $r_z^{-1}(\mathbb{RP}^1)$ are the critical points $v_0(z), v_1(z), \ldots, v_n(z)$, which are distinct for $z \in (0,1]$, we see that the degree of r_z is constant for $z \in (0,1]$.

If the degree of r_0 is less than that of r_1, then some edges must disappear in the limit as $z \to 0$. But this can only happen to an edge between $v_0(z)$ and $v_1(z)$, as they are the only critical points which collide in the limit as $z \to 0$. This shows the necessity of an edge between $v_0(1)$ and $v_1(1)$ for the degree to drop.

For sufficiency, note that if there is an edge between $v_0(1)$ and $v_1(1)$, then it must collapse in the limit as $z \to 0$ for otherwise condition (iii) for nets would be violated.

9.1.2 Schubert induction and the proof

In Chapter 8, we used Schubert induction to construct a sequence of points $s_1, s_2, \ldots, s_{\dim \mathrm{Gr}} \in \mathbb{R}$ and sufficiently many real points in each Schubert variety $X_\alpha F_\bullet(0)$ which also lie in $X_{\square} F_\bullet(s_i)$ for $i = 1, 2, \ldots, |\alpha|$. Without re-running that proof, we will describe what we got from that construction in the case of rational functions. We warn the reader that we will be adopting the pedagogically unsound practice of changing our notation for this lecture[†]. While this is a serious violation of the implicit contract between an author and potential readers, we will do it anyway.

Recall that Schubert varieties $X_\alpha F_\bullet(0) =: X_\alpha$ are indexed by pairs $\alpha\colon \alpha_1 < \alpha_2$. We work with polynomials (where the degree of a monomial is natural), rather than matrices (where the column index of a pivot is natural). This switch in emphasis is the reason for the shift in our indexing conventions. For each pair $\kappa = (k_1, k_2)$ with $0 \le k_1 < k_2 \le d$, we have a Schubert variety X_κ. (In the notation of Chapter 8, this corresponds to $X_\alpha F_\bullet$ for $\alpha = d+1-k_2 < d+1-k_1$.) A general point of X_κ is represented uniquely in St_d by a pair of polynomials of the form

$$\begin{array}{rcl} f & = & \qquad\quad t^{d-1} + a_{d-2}t^{d-2} + \quad \cdots \quad + a_{k_1}t^{k_1} \\ g & = & t^d + b_{d-1}t^{d-1} + b_{d-2}t^{d-2} + \cdots + b_{k_2}t^{k_2} \end{array} \tag{9.1}$$

Actually, this lies in X_κ°, which is the intersection of X_κ with the set of pairs (f, g) of polynomials where f/g does not have a critical point at ∞.

This set of indices is ordered by componentwise comparison, which gives the *Bruhat order*. Here, the order is the reverse of the geometric order given by closures of cells. Figure 9.1 shows this Bruhat order for $d = 4$, and on the right the recursion for the number $\delta(k)$ of chains from the top[†] to the point κ. In Chapter 8, we constructed points $0 < s_{2d-2} <$

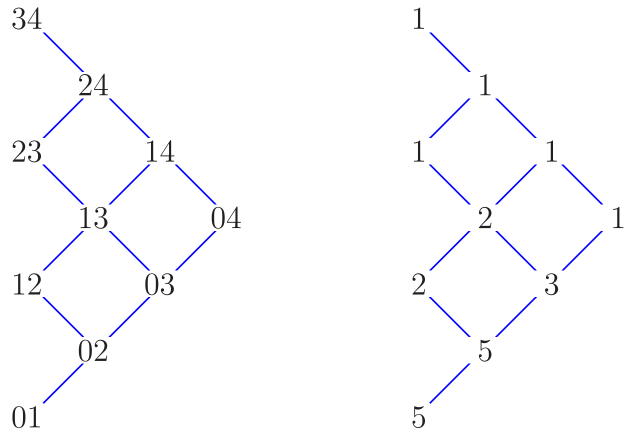


Figure 9.1: The Bruhat order and enumeration of chains for $d = 4$.

$s_{2d-4} < \cdots < s_2 < s_1 \in \mathbb{R}$, and, for each index κ, $\delta(\kappa)$ pairs of polynomials (f, g) of the form (9.1) with Wronskian

$$f'g - fg' \;=\; \text{constant} \cdot t^{k_1+k_2-1} \cdot \prod_{i=1}^{2d-1-k_1-k_2} (t - s_i)\,.$$

[†]The difference in our notations is simply reversing the top and bottom of the Bruhat order

As mentioned in Remark 8.4, we did even more in that proof. Set $\kappa^1 := (k_1 + 1 < k_2)$ and $\kappa^2 := (k_1 < k_2 + 1)$. (These are not always defined, but they are the indices lying above κ in this Bruhat order.) Since $\delta(\kappa)$ is the number of paths from the top, $(d-1, d)$, to κ, it satisfies the path-counting recursion $\delta(\kappa) = \delta(\kappa^1) + \delta(\kappa^2)$, with initial case $\delta(d-1 < d) = 1$ (as there is a unique pair of polynomials f, g of the form (9.1) for $k = (d-1, d)$). In the proof, we constructed $\delta(\kappa)$ families of pairs of polynomials $\{(f_z, g_z) \mid z \in [0,1]\}$ such that

1. For $z \neq 0$, $(f_z, g_z) \in X^\circ_\kappa$.
2. $f'_z g_z - f_z g'_z \;=\; \text{constant} \cdot t^{k_1+k_2-1} \cdot \left(\prod_{i=1}^{2d-2-k_1-k_2} (t - s_i) \right) \cdot (t - z s_{2d-1-k_1-k_2})$.
3. Exactly $\delta(\kappa^i)$ of these paths began in X_{κ^i}. That is, have $(f_0, g_0) \in X_{\kappa^i}$.

9.1.3 Main theorem and completion of the proof

Theorem 9.6 *The $\delta(\kappa)$ rational functions we constructed in X°_κ have different nets.*

Proof. Removing the common zeroes from f/g gives a rational function r of degree $d - k_1$ with one critical point at zero of high multiplicity $k_2 - k_1 - 1$ and simple critical points $s_1, s_2, \ldots, s_{2d-1-k_1-k_2}$. The pair (f, g) corresponds to a path π from $(d-1, d)$ to $\kappa = (k_1, k_2)$ in the Bruhat order. We claim that this path can be recovered from the net $r^{-1}(\mathbb{RP}^1)$ of r.

Indeed, by Corollary 9.5, there is an edge in the net between critical points s_i and s_{i+1} if and only the degree of the rational function increased in the ith step, when the critical point s_i was created. From the form (9.1) of (f, g), this means that the ith step in the corresponding path had the form

$$(l_1 < l_2) \xrightarrow{\;i\text{th step}\;} (l_1 - 1 < l_2)\,.$$

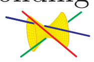
Similarly, if there is no edge between s_i and s_{i+1}, then the ith step in the corresponding path had the form $(l_1 < l_2) \to (l_1 < l_2 - 1)$. 

Figure 7.2 (compare this to Figure 9.1) illustrates the formation of the nets during the Schubert induction for $d = 4$ rational curves.

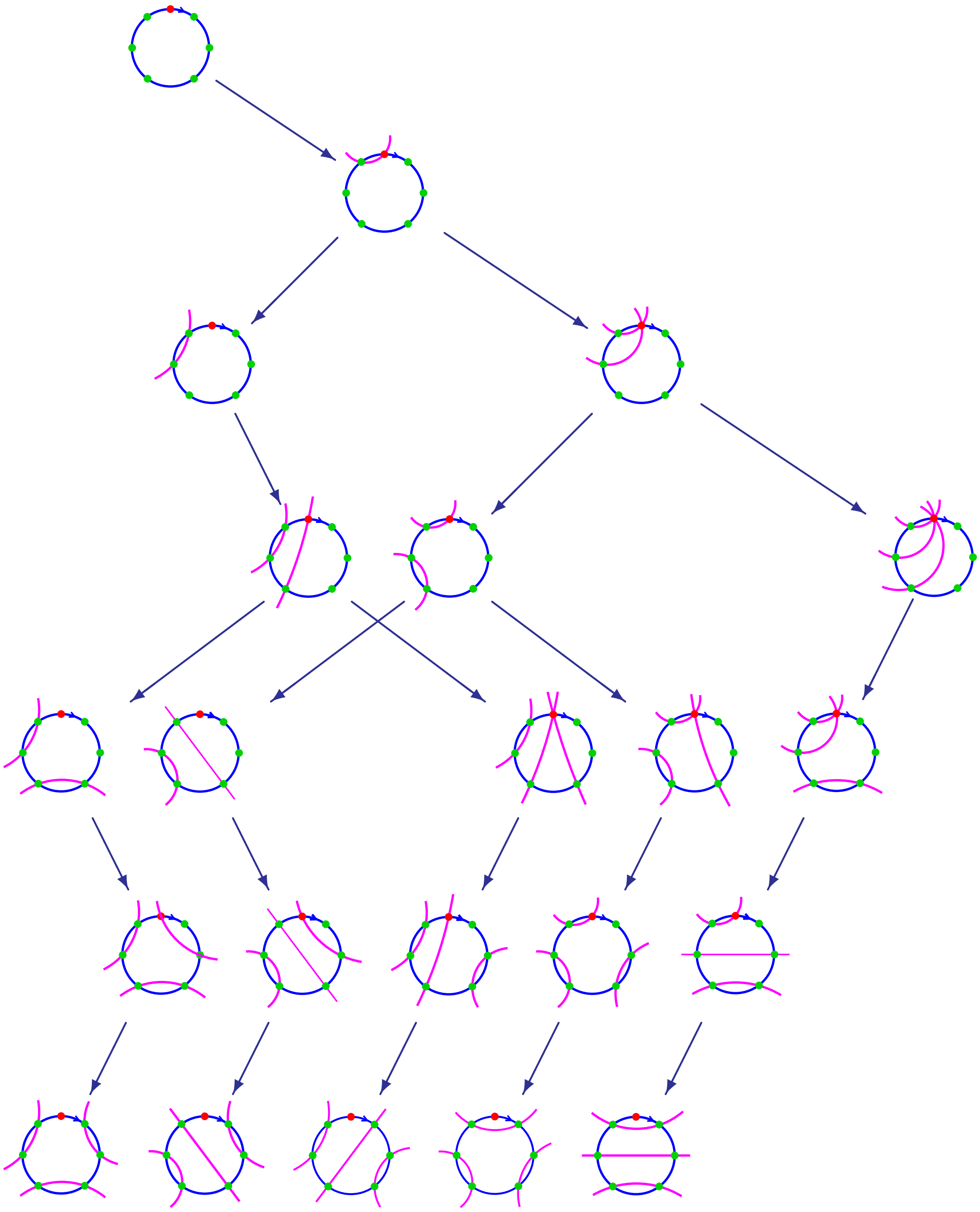


Figure 7.2: Formation of nets during Schubert induction.

We complete the proof of the Shapiro conjecture for rational curves.

Theorem 9.7 *Let $\Phi(t)$ be a hyperbolic polynomial of degree $2d-2$ with distinct roots. Then there are exactly $\delta(0,1)$ real rational functions $r = f/g$ with $f'g - fg' = \Phi(t)$, up to equivalence.*

Proof. Let $s_1, s_2, \ldots, s_{2d-2} \in \mathbb{R}$ be the numbers constructed in Chapter 8 such that every rational function with critical points s_i is real. Let $\Phi_0(t)$ be the polynomial $\prod_i (t - s_i)$ having these numbers as its roots. Then there is a continuous family $\{\Phi_z \mid z \in [0,1]\}$ of hyperbolic polynomials all having distinct roots with $\Phi_1(t) = \Phi(t)$. (For example, we can move one set of roots to the other, keeping them distinct and real at all times.)

We try to analytically continue each point in the fiber $\mathrm{Wr}^{-1}(\Phi_z)$ from $z = 0$ to $z = 1$. The only way this could fail would be if we encountered a fiber $\mathrm{Wr}^{-1}(\Phi_z)$ with a multiple point. But then two of the rational functions in this fiber would have to coincide. In particular two would have to have the same net, (where we have labeled the nets by the root of the $\Phi_z(t)$ corresponding to s_1). By Corollary 9.4, we must have that two of the original rational functions in $\mathrm{Wr}^{-1}(\Phi_0)$ would then have the same net. But this contradicts Theorem 9.6.

Chapter 10: Proof of the Shapiro conjecture

This chapter will be derived from Sottile's Current Events Bulletin Lecture at the AMS meeting in January 2009.

Chapter 11: Beyond the Shapiro Conjecture

Introduction

Here, we will touch on three further topics related to the Shapiro Conjecture.

I Transversality. Talk about the second proof of Mukhin, Tarasov, and Varchenko. Also discuss discriminants.

II Maximally Inflected Curves.

III The Shapiro Conjecture for flag manifolds.

IV The Secant Conjecture.

11.1 Transversality

In Chapter 9, we presented a proof of the Shapiro Conjecture for rational functions by Eremenko and Gabrielov [31]. Its main point was that there is no obstruction to analytically continuing the rational functions that were constructed in Chapter 8 to give rational functions with any given Wronskian having distinct real zeroes. The key to this was the association of a net to each rational function with only real critical points, which showed that the analytic continuation was possible.

A consequence of this proof is the statement that when $\min(m,p)=2$, the Wronski map

$$\mathrm{Wr} \;:\; \mathrm{Gr}(p,\mathbb{C}^{m+p}) \;\longrightarrow\; \mathbb{CP}^{mp}$$

is unramified over the locus of hyperbolic polynomials with distinct (real) roots. This is in fact true for all Grassmannians, as Mukhin, Tarasov, and Varchenko showed [72].

Theorem 11.1 *The Wronski map is unramified over the locus of hyperbolic polynomials with distinct roots, for any m and p.*

More generally, given called Schubert data, $\alpha_1,\alpha_2,\ldots,\alpha_n$ we can consider intersections of the form

$$X_{\alpha_1}F_\bullet(s_1) \;\cap\; X_{\alpha_2}F_\bullet(s_2) \;\cap\; \cdots \;\cap\; X_{\alpha_n}F_\bullet(s_n)\ , \qquad (11.1)$$

where $s_1,s_2,\ldots,s_n\in\mathbb{P}^1$. Recall that a collection of subvarieties meet *transversally* if at all points of their intersection, their tangent spaces meet properly (have the expected dimension of intersection). We rephrase Theorem 11.1.

Transversality Theorem. *If the points $s_1,s_2,\ldots,s_n$ are real and distinct, then the intersection* (11.1) *is transverse.*

Just as in Chapter 9, this Transversality Conjecture, together with the asymptotic existence result (Theorem 8.2 of Chapter 8) implies the full Shapiro Conjecture, as it asserts that there are no obstructions to analytic continuation.

There is more evidence than the results of Eremenko and Gabrielov. A recent project [78] computed over 133 million different Schubert intersections of the form (11.1), where the points $s_1, s_2, \ldots, s_n$ were distinct rational numbers chosen at random. Every intersection was transverse with all points real, without exception. This involved 212 different collections of Schubert data $\alpha_1, \alpha_2, \ldots, \alpha_n$ on the Grassmannians $\mathrm{Gr}(3,6)$, $\mathrm{Gr}(3,7)$, and $\mathrm{Gr}(4,8)$. In contrast, that same project computed over 350 million instances of more general Schubert intersections of the form (11.1) on flag manifolds, where we know that transversality does not hold (on some subset of codimension 1). In that phase of the project, we found several hundred intersections which were not transverse. Either the search restricted to the Grassmannian was extremely unlucky to not chance upon a nontransverse intersection, or there are no choices of distinct real numbers $s_1, s_2, \ldots, s_n$ giving a nontransverse intersection in (11.1). We believe the latter.

This geometric Transversality Conjecture has a strengthening, which is stated in terms of real algebra.

Definition 11.2 The *discriminant* is the locus in $\mathbb{P}^{mp}$ of critical values of the Wronski map (points over which it is ramified). It is an algebraic hypersurface and defined by a single polynomial, also called the discriminant.

The Transversality Conjecture asserts that the discriminant does not meet the set of hyperbolic polynomials with distinct roots.

Equivalently (and more generally), if we consider a given family of intersections of the form (11.1), then the discriminant is the set of points $(s_1, s_2, \ldots, s_n)$ for which this intersection is not transverse. Again, this discriminant is a polynomial $\Delta(s_1, s_2, \ldots, s_n)$ in the parameters $(s_1, s_2, \ldots, s_n) \in (\mathbb{P}^1)^n$, and the Transversality Conjecture asserts that Δ does not vanish when the parameters are real and distinct, that is, when $s_i \neq s_j$, for all i, j. We conjecture something much stronger.

Conjecture 11.3 (Discriminant Conjecture) *The discriminant polynomial Δ is a sum of squares. Each term in the sum is a monomial in the differences $(s_i - s_j)$.*

There is some evidence for this conjecture. For the problem of two lines in 3-space meeting four lines $\ell(s)$, $\ell(t)$, $\ell(u)$, and $\ell(v)$, each tangent to the rational normal curve, the discriminant is a constant multiple of

$$(s-t)^2(u-v)^2 \ + \ (s-u)^2(t-v)^2 \ + \ (s-v)^2(t-u)^2\,.$$

For $\mathrm{Gr}(2,5)$, the discriminant of the Schubert intersection

$$X_{\square}F_\bullet(0) \ \cap \ X_{\square}F_\bullet(s) \ \cap \ X_{\square}F_\bullet(t) \ \cap \ X_{\square}F_\bullet(u) \ \cap \ X_{\square}F_\bullet(v) \ \cap \ X_{\square}F_\bullet(\infty)\,,$$

has degree 20 in the variables s, t, u, and v, and it has 711 different terms. Ad hoc methods [91] showed that it was a sum of squares.[†] This was quite surprising. Hilbert [48] showed that not every polynomial of even degree greater than 2 in four variables that is nonnegative can be written as a sum of squares.

We feel that it is a useful project to further investigate these discriminant polynomials.

[†] Give more details?

11.2 Maximally inflected curves

This section represents joint work with Kharlamov [58]. A list $f_1(t), f_2(t), \ldots, f_k(t)$ of polynomials with degree d defines a map $\varphi\colon \mathbb{P}^1 \to \mathbb{P}^{k-1}$ as follows,

$$\varphi \;:\; \mathbb{P}^1 \ni t \;\longrightarrow\; [f_1(t) : f_2(t) : \cdots : f_k(t)] \in \mathbb{P}^{k-1}\,.$$

The image of the curve is *convex* at a point $\varphi(t)$ if and only if the first $k{-}1$ derivatives of $\varphi(t)$ are linearly independent. The failure to be convex is measured exactly by the vanishing of the Wronskian of the polynomials $f_1(t), f_2(t), \ldots, f_k(t)$. We say that φ is *ramified* at a point t if this Wronskian vanishes at t. Another geometric term is that the curve φ has an *inflection point* (or flex) at t. This corresponds exactly to the usual notion of an inflection point for plane curves. A rational curve φ of degree d in $\mathbb{P}^{k-1}$ has $k(d{+}1{-}k)$ flexes, counted with multiplicity.

The connection between the Schubert calculus and rational curves in projective space (linear series on $\mathbb{P}^1$) originated in work of Castelnuovo [18] on g-nodal rational curves. This led to the use of Schubert calculus in Brill-Noether theory (see Chapter 5 of [47] for an elaboration). In turn, the theory of limit linear series of Eisenbud and Harris [29, 30] provides essential tools to show reality of the special Schubert calculus [89]. (That result on the special Schubert calculus was a generalization of Theorem 8.2 of Chapter 8.)

Mukhin, Tarasov, Varchenko Theorem for rational curves. *If a rational curve in $\mathbb{P}^{k-1}$ has all of its flexes real, then it must be real.*

Definition 11.4 A real rational curve with all of its flexes real is said to be *maximally inflected*.

The Shapiro conjecture asserts that there are lots (maximally many, in fact) of these maximally inflected curves.

Up to projective transformation and reparametrization, there are only three real rational plane cubic curves. They are represented by the equations

$$y^2 \;=\; x^3 + x^2\,, \qquad y^2 \;=\; x^3 - x^2\,, \qquad \text{and} \qquad y^2 \;=\; x^3\,,$$

and they have the shapes shown in Figure 11.1. All three have a real flex at infinity and are

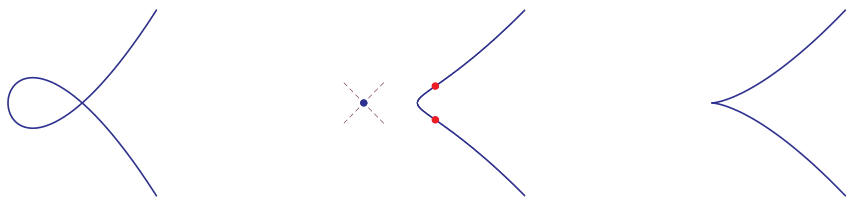


Figure 11.1: Real rational cubics.

singular at the origin. The first has a real node and no other real flexes, the second has a solitary point and two real flexes at $(\frac{4}{3}, \pm\frac{4}{3\sqrt{3}})$ (we indicate these with dots and the complex conjugate tangents at the solitary point with dashed lines), and the third has a real cusp. The last two are maximally inflected, while the first is not.

The Schubert calculus gives 5 rational quartics with 6 given points of inflection and Figure 11.2 shows 5 maximally inflected curves with flexes at $\{-3,-1,0,1,3,\infty\}$. (Each nodal curve has 2 flexes at its node, which is a consequance of the symmetry in the choice of flexes.) We indicate the differences in the parameterizations of these curves, labeling the

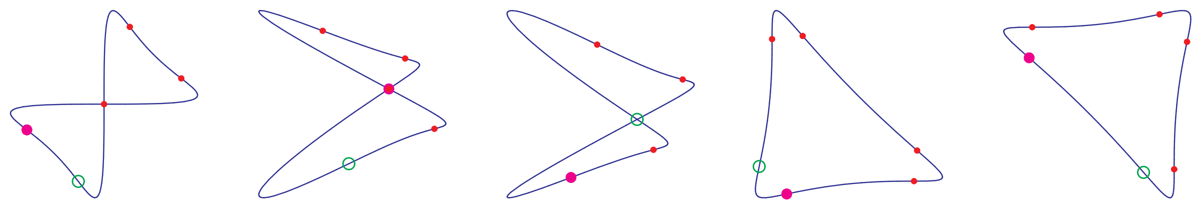
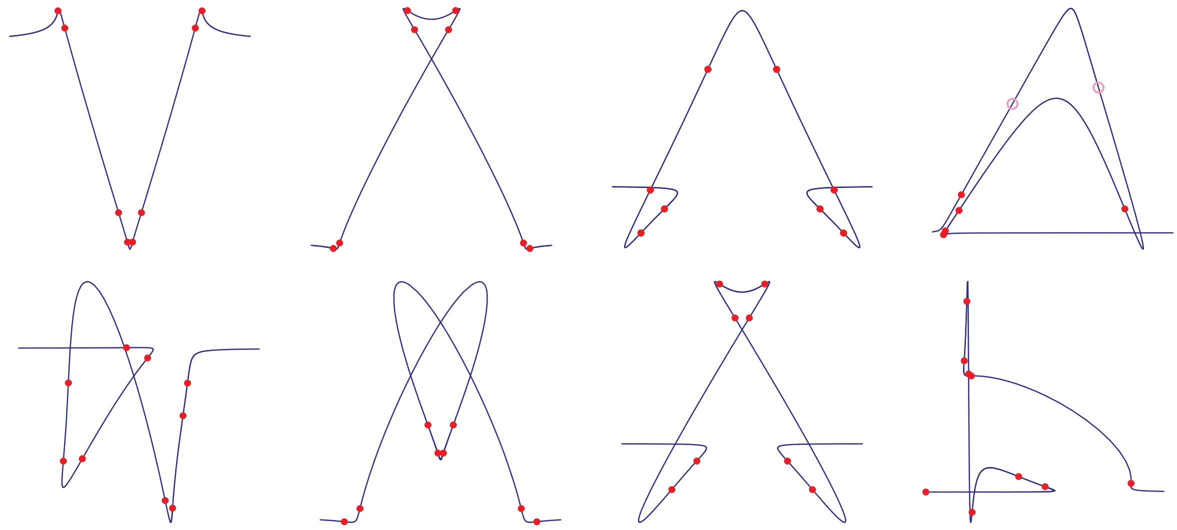


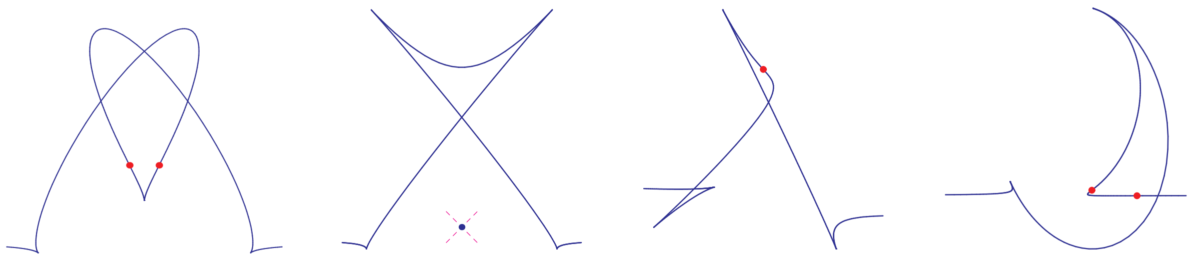
Figure 11.2: The 5 curves with flexes at $\{-3,-1,0,1,3,\infty\}$.

flex at -3 by the larger dot and the flex at -1 by the circle. The solitary points are not drawn. The first three curves have two solitary points, while the last 2 have three solitary points.

Here are a few quintics. The flexes are indicated (the symmetric curves have one additional flex at infinity), but we do not draw solitary points. Also, the open circles represent two flexes which have merged into a *planar point*.



Finally, here are four singular quintics. The solitary point is drawn on the second.

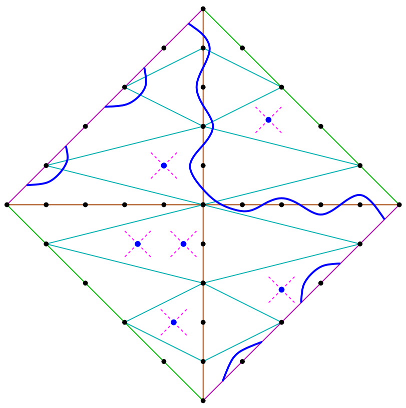


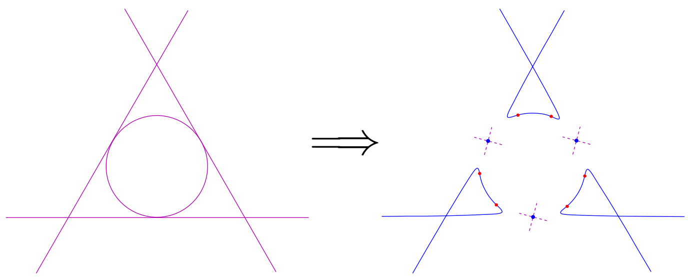
None of these maximally inflected curves had many nodes. Indeed, rational quartics typically have 3 double points and quintics have 6, yet we saw quartics with at most 1 node and quintics with at most 3 nodes. More generally, consider a maximally inflected curve with only flexes and cusps, and whose other singularities are ordinary double points. Let ι be its number of flexes and κ be its number of cusps. Then $\iota + 2\kappa = 3(d-2)$. By the genus formula, it has $\binom{d-1}{2} - \kappa$ double points. The following theorem is an easy consequence of the Klein [62] and Plücker [73] formulas.

Theorem 11.5 (Topological Restrictions [58]) *Suppose that a maximally inflected curve has only flexes, cusps, and ordinary double points, and let ι and κ be its number of flexes and cusps, respectively. Then it has at least $d-2-\kappa$ solitary points and at most $\binom{d-2}{2}$ nodes.*

Thus maximally inflected cubics have at most $\binom{3-2}{2} = 0$ nodes, quartics have at most $\binom{4-2}{2} = 1$ node, and quintics have at most $\binom{5-2}{2} = 3$ nodes.

The existence of curves satisfying the hypotheses of Theorem 11.5 is not guaranteed, even if we know the Shapiro conjecture. For example, the construction in Theorem 4 of Chapter 8 does not guarantee that the maximally inflected curve has only ordinary double points. There are, however two constructions which guarantee curves having only double points. The first uses Shustin's patchworking of singular curves [83] to obtain degenerate Harnack curves with $\binom{d-1}{2}$ solitary points (the figure on the left illustrates this patchworking). The second perturbs $d-2$ lines tangent to a conic to obtain maximally inflected curves of degree d with the minimal number of solitary points (and up to $d-2$ cusps).





The topological classification of maximally inflected plane quintics is open.[†] Also, maximally inflected curves in higher-dimensional space have not been investigated. For example, which knot types can occur for maximally inflected space curves? We speculate that there should be some restrictions on the number of real quadrisecants of a maximally inflected space curve.

11.3 The Shapiro Conjecture for flag manifolds

The Shapiro conjecture involved linear spaces satisfying incidence conditions imposed by flags osculating the rational normal curve at real points. One natural variant is to consider

[†]Explain a couple of open questions

flags satisfying incidence conditions imposed by osculating flags.

Example 11.6 Consider partial flags of subspaces $m \subset H$, where m is a line which lies on a plane H in $\mathbb{P}^3$. Here are two typical incidence conditions on these flags.

- m meets a fixed line $\ell(t)$ tangent to the rational normal curve γ at the point $\gamma(t)$.
- H contains a fixed point $\gamma(t)$ of the rational normal curve.

We consider the problem of partial flags $m \subset H$ where m meets three tangent lines $\ell(-1)$, $\ell(0)$, and $\ell(1)$ to γ, and H contains two points $\gamma(v)$ and $\gamma(w)$. First observe that H contains the line $\lambda(v, w)$ spanned by these two points. But then m must also meet this secant line $\lambda(v, w)$, as $m \subset H$. In this way, we see that this problem reduces to our old friend, the problem of four lines. Now, however, we seek lines m which meet three tangent lines $\ell(-1)$, $\ell(0)$, and $\ell(1)$ to the rational normal curve, and one secant line $\lambda(v, w)$. Given such a line m, we may recover H as the linear span of m and $\lambda(v, w)$.

As in Section 4 of Chapter 7, we begin with the quadric Q containing the three lines $\ell(-1)$, $\ell(0)$, and $\ell(1)$ tangent to the rational normal curve γ. The lines meeting $\ell(-1)$, $\ell(0)$,

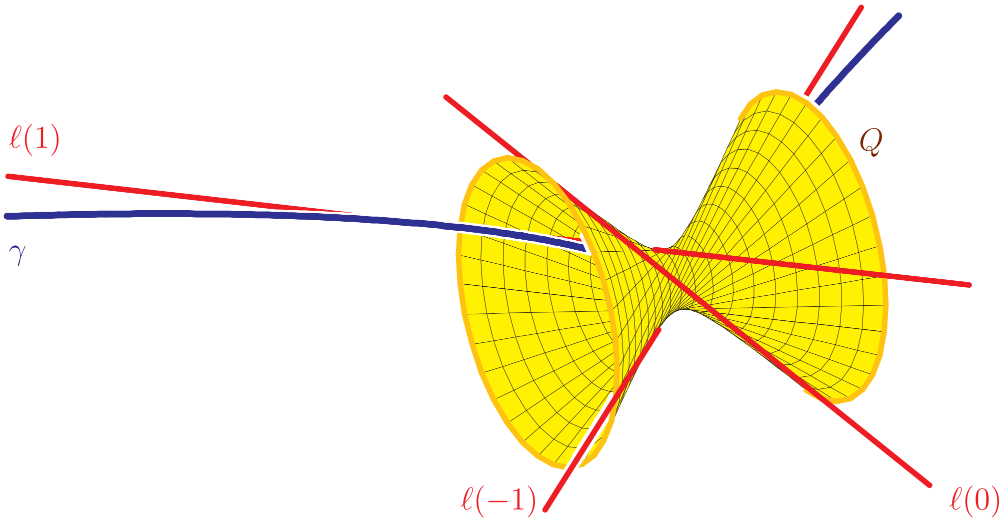


Figure 11.3: Quadric containing three lines tangent to the rational normal curve.

$\ell(1)$, and the secant line $\lambda(v, w)$ correspond to the points where $\lambda(v, w)$ meets the quadric Q. In Figure 11.4, we display a secant line $\lambda(v, w)$ which meets the hyperboloid in two real points, and therefore these choices for v and w give two real solutions to our geometric problem. There is also a secant line which meets the hyperboloid Q in no real points, and hence in two complex conjugate points. For this secant line, neither flag solving our problem is real. We show this configuration in Figure 11.5.

The original Shapiro conjecture concerned partial flags of a given type meeting flags osculating the rational normal curve (at real points), and it asserted that all such partial flags would be real. This example shows that the Shapiro conjecture fails for flags, but not too badly. It is useful to consider this failure schematically. In Figure 11.3 we represent the rational normal curve γ as a circle, and indicate the relative positions of three tangent lines

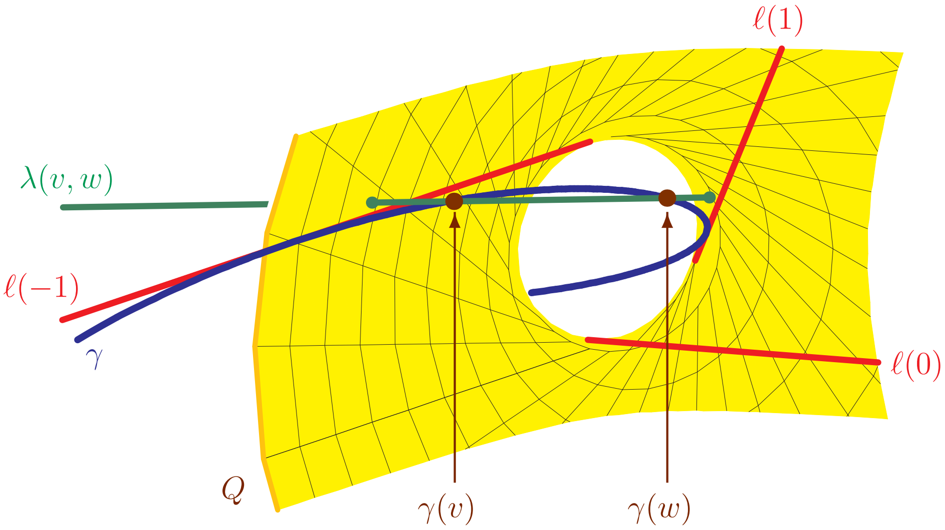


Figure 11.4: A secant line meeting Q.

and the secant line. Observe that in the diagram on the left, one may travel along the circle, first encountering the three points where the lines are tangent and then the two points where the line is secant. If we record the dimension of the piece of the flag $m \subset H$ which is affected by the tangent line or point on the rational normal curve as we travel along the circle, we get the sequence $(1,1,1,2,2)$, which is weakly increasing. The reading the diagram on the right either gives the sequence $(1,1,2,1,2)$, or its reverse $(2,1,2,1,1)$, or $(1,2,1,2,1)$, none of which are monotone.

11.3.1 The Monotone Conjecture

A *partial flag* $E_\bullet$ of *type* $a = (a_1, a_2, \ldots, a_k)$ is a sequence of linear subspaces

$$E_\bullet \;:\; E_{a_1} \;\subset\; E_{a_2} \;\subset\; \cdots \;\subset\; E_{a_k} \;\subset\; \mathbb{C}^n\,,$$

where $a_i = \dim E_i$. Its possible positions with respect to a fixed (complete) flag are encoded by certain permutations, α. The original Shapiro conjecture concerned partial flags of a given type meeting flags osculating the rational normal curve (at real points), and it asserted that all such partial flags would be real. The example that we just gave shows that it was wrong[†]. However, it can be repaired.

A position α for flags $E_\bullet$ of type a which may be expressed in terms of only one piece E_{a_i} of the flag $E_\bullet$ is called *Grassmannian* of *index* a_i. When this occurs, write $\iota(\alpha) = a_i$. In the counterexample to the Shapiro conjecture, both conditions are Grassmannian—one has index 2 (for the requirement that m meet a tangent line) and the other has index 3 (for the requirement that H meet a point on the rational normal curve.) These differ from the dimension of the linear spaces m and H, as we were working in projective 3-space, rather than a 4-dimensional vector space.

Conjecture 11.7 (Monotone Conjecture) *Let $\alpha_1, \alpha_2, \ldots, \alpha_m$ be Grassmannian conditions on flags of type a whose indices satisfy $\iota(\alpha_1) \le \iota(\alpha_2) \le \cdots \le \iota(\alpha_m)$. Then if*

[†]In fact, it fails for the first nontrivial problem on a flag manifold.

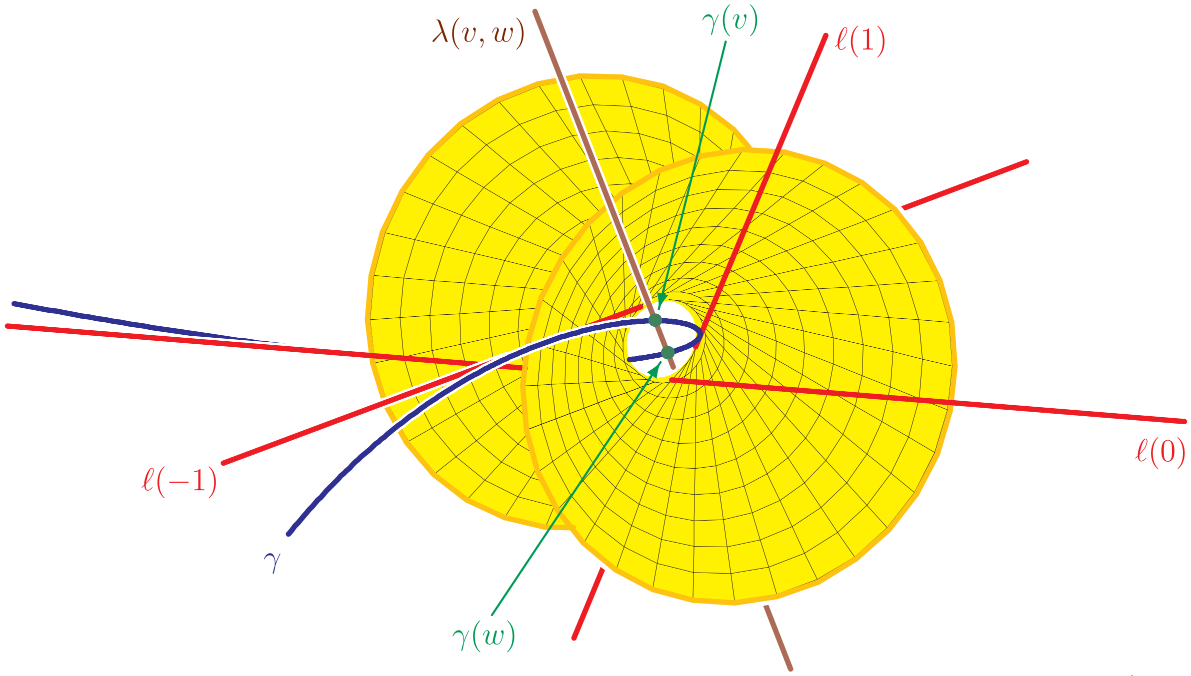
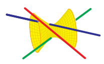


Figure 11.5: A secant line not meeting Q.



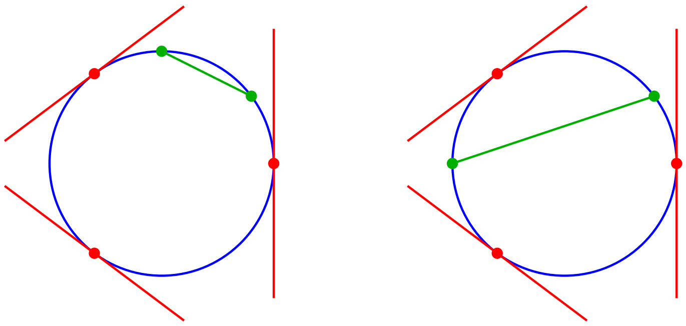


Figure 11.6: Always real. Not always real.

$s_1 < s_2 < \cdots < s_m$ are real numbers, the interection

$$\bigcap_{i=1}^{m} X_{\alpha_i} F_\bullet(s_i)$$

is transverse with all points real.

As with the Shapiro conjecture for Grassmannians, it is known that transversality will imply that the points of intersection are real. There is also a similar discriminant conjecture. The *preprime* generated by polynomiald $g_1, g_2, \ldots, g_m$ is the collection of polynomials f of the form

$$f \;=\; \sigma_0 \;+\; \sigma_1 g_1 \;+\; \sigma_2 g_2 \;+\; \cdots \;+\; \sigma_m g_m\,,$$

where the polynomials σ are sums of squares. Any polynomial f in this preprime is positive on the set

$$K \;:=\; \{x \mid g_i(x) > 0\}$$

and a representation of f as an element of this preprime is a certificate of its positivity on K.

Conjecture 11.8 (Discriminant Conjecture) *The discriminant of the polynomial system associated to an intersection of Schubert varieties defined by Grassmannian conditions as in the monotone conjecture lies in the preprime generated by the differences $(s_i - s_j)$ for $i > j$.*

The point of this conjecture is that not only is the discriminant positive on the set $\{(s_1, s_2, \ldots, s_m) \mid s_1 < s_2 < \cdots < s_m\}$ of monotone parameters, but that it has a special form from which this positivity is transparent. It is known that not all polynomials which are positive on a set of this form can lie in the preprime generated by the differences $s_i - s_j$ for $i > j$.

There is a continually unfolding story of the Shapiro conjecture and its generalizations, both to secant flags (by Eremenko, Gabrielov, Shapiro, and Vainshtein [34]) and for other flag manifolds, for other groups. Very few computations have been done (but see [92] for the beginnings for the Lagrangian Grassmannian and orthogonal Grassmannian.)

We end these notes with the remark that, just as they were being completed, Mukhin, Tarasov, and Varchenko [71] gave a proof the the Shapiro conjecture for Grassmannians, using the Bethe Ansatz and the Gaudin model (a connection between differential equations with polynomial solutions, representation theory of the special linear group, and intersections of Schubert varieties given by flags osculating the rational normal curve). The discriminant conjecture remain open, however.

Topics to work on for next rewrite:

1. Discuss some of the evidence for the Monotone Conjecture
2. Explain the result of Eremenko, Gabrielov, Shapiro , and Vainshtein and how it establishes the monotone conjecture for certain flag manifolds.
3. Pose the secant conjecture and explain the proof for evenly spaced points (it you can).
4. Make the monotone secant conjecture and give a few words about current computations.

Bibliography

[1] Martin Avendaño, *The number of real roots of a bivariate polynomial on a line*, arXiv:math/0702891v1.

[2] S. Basu, R. Pollack, and M.-F. Roy, *Algorithms in real algebraic geometry*, Algorithms and Computation in Mathematics, vol. 10, Springer, 2003.

[3] D.J. Bates, F. Bihan, and F. Sottile, *Bounds on real solutions to polynomial equations*, Int. Math. Res. Notes (2007), 2007:rnm114–7.

[4] E. Becker, M.G. Marinari, T. Mora, and C. Traverso, *The shape of the Shape Lemma*, Proceedings ISSAC-94, 1993, pp. 129–133.

[5] R. Benedetti and J.-J. Risler, *Real algebraic and semi-algebraic sets*, Hermann, 1990.

[6] D. Bernstein, A. Kushnirenko, and A. Khovanskii, *Newton polytopes*, Usp. Math. Nauk. **31** (1976), 201–202.

[7] D. N. Bernstein, *The number of roots of a system of equations*, Funct. Anal. Appl. **9** (1975), 183–185.

[8] A. Bertram, *Quantum Schubert calculus*, Adv. Math. **128** (1997), no. 2, 289–305.

[9] B. Bertrand, F. Bihan, and F. Sottile, *Polynomial systems with few real zeroes*, Math. Z. **253** (2006), no. 2, 361–385.

[10] É. Bézout, *Théorie générale des équations algébriques*, Ph.-D. Pierres, Paris, 1779.

[11] F. Bihan, *Polynomial systems supported on circuits and dessins d'enfants*, J. Lond. Math. Soc. (2) **75** (2007), no. 1, 116–132.

[12] F. Bihan, J.M. Rojas, and F. Sottile, *Sharpness of fewnomial bounds and the number of components of a fewnomial hypersurface*, Algorithms in Algebraic Geometry (A. Dickenstein, F. Schreyer, and A. Sommese, eds.), IMA Volumes in Mathematics and its Applications, vol. 146, Springer, 2007, pp. 15–20.

[13] F. Bihan and F. Sottile, *Gale duality for complete intersections*, 2007, Ann. de l'Institut Fourier, to appear. `ArXiV:0706.3745`.

[14] ———, *New fewnomial upper bounds from Gale dual polynomial systems*, Moscow Mathematical Journal **7** (2007), no. 3, 387–407.

[15] F. Budan de Boislaurent, *Nouvelle méthode pour la résolution des équations numériques d'un degré quelconque*, Paris, 1807.

[16] Y. Burda and A. Khovanskii, *Degree of a rational mapping, theorems of sturm and tarski*, 2008.

[17] L. Caporaso and J. Harris, *Counting plane curves of any genus*, Invent. Math. **131** (1998), no. 2, 345–392.

[18] G. Castelnuovo, *Numero delle involuzioni razionali gaicenti sopra una curva di dato genere*, Rendi. R. Accad. Lincei **4** (1889), no. 5, 130–133.

[19] M. Chasles, *Construction des coniques qui satisfont à cinque conditions*, C. R. Acad. Sci. Paris **58** (1864), 297–308.

[20] D. Cox, *The homogeneous coordinate ring of a toric variety*, J. Alg. Geom. **4** (1995), 17–50.

[21] David Cox, John Little, and Donal O'Shea, *Ideals, varieties, and algorithms*, third ed., Undergraduate Texts in Mathematics, Springer, New York, 2007, An introduction to computational algebraic geometry and commutative algebra.

[22] A. I. Degtyarev and V. M. Kharlamov, *Topological properties of real algebraic varieties: Rokhlin's way*, Uspekhi Mat. Nauk **55** (2000), no. 4(334), 129–212.

[23] R. Descartes, *La géométrie*, 1637.

[24] A. Dickenstein, E.-M. Feichtner, and B. Sturmfels, *Tropical discriminants*, ArXiv.org:math.AG/0510126.

[25] A. Dickenstein, J. M. Rojas, K. Rusek, and J. Shih, *Extremal real algebraic geometry and $\mathcal{A}$-discriminants*, Mosc. Math. J. **7** (2007), no. 3, 425–452, 574.

[26] P. Dietmaier, *The Stewart-Gough platform of general geometry can have 40 real postures*, Advances in Robot Kinematics: Analysis and Control, Kluwer Academic Publishers, 1998, pp. 1–10.

[27] E. Ehrhart, *Sur les polyèdres rationnels homothétiques à n dimensions*, C. R. Acad. Sci. Paris **254** (1962), 616–618.

[28] D. Eisenbud, *Commutative algebra with a view towards algebraic geometry*, GTM, no. 150, Springer-Verlag, 1995.

[29] D. Eisenbud and J. Harris, *Divisors on general curves and cuspidal rational curves*, Invent. Math. **74** (1983), 371–418.

[30] ———, *When ramification points meet*, Invent. Math. **87** (1987), 485–493.

[31] A. Eremenko and A. Gabrielov, *Elementary proof of the B. and M. Shapiro conjecture for rational functions*, `math.AG/0512370`.

[32] ———, *Degrees of real Wronski maps*, Discrete Comput. Geom. **28** (2002), no. 3, 331–347.

[33] ———, *Rational functions with real critical points and the B. and M. Shapiro conjecture in real enumerative geometry*, Ann. of Math. (2) **155** (2002), no. 1, 105–129.

[34] A. Eremenko, A. Gabrielov, M. Shapiro, and A. Vainshtein, *Rational functions and real Schubert calculus*, Proc. Amer. Math. Soc. **134** (2006), no. 4, 949–957 (electronic).

[35] S. Fiedler-Le Touzé, *Pencils of cubics as tools to solve an interpolation problem*, Appl. Algebra Engrg. Comm. Comput. **18** (2007), no. 1-2, 53–70.

[36] J. Fourier, *Analyse des équations déterminées*, F. Didot, Paris, 1831.

[37] William Fulton, *Intersection theory*, second ed., Ergebnisse der Mathematik und ihrer Grenzgebiete. 3. Folge. A Series of Modern Surveys in Mathematics [Results in Mathematics and Related Areas. 3rd Series. A Series of Modern Surveys in Mathematics], vol. 2, Springer-Verlag, Berlin, 1998.

[38] Wm. Fulton, *Introduction to toric varieties*, Annals of Mathematics Studies, vol. 131, Princeton University Press, Princeton, NJ, 1993, The William H. Roever Lectures in Geometry.

[39] ———, *Introduction to intersection theory in algebraic geometry*, CBMS 54, AMS, 1996, second edition.

[40] A. Gabrielov, D. Novikov, and B. Shapiro, *Mystery of point charges*, Proc. Lond. Math. Soc. (3) **95** (2007), no. 2, 443–472.

[41] J. Gahleitner, B. Jüttler, and J. Schicho, *Approximate parameterization of planar cubics*, Curve and surface design (Saint-Malo, 2002), Mod. Methods Math., Nashboro Press, Brentwood, TN, 2003, pp. 167–176.

[42] A. Gathmann and H. Markwig, *The Caporaso-Harris formula and plane relative Gromov-Witten invariants in tropical geometry*, Math. Ann. **338** (2007), no. 4, 845–868.

[43] I. Gordon, E. Horozov, and M. Yakimov, *The real loci of Calogero-Moser spaces, representations of rational Cherednik algebras and the Shapiro conjecture*, arXiv:0711.4336.

[44] V. Gough, *Contribution to discussion papers on research in automobile stability and control in tyre performance*, 1957.

[45] B. Grünbaum, *Convex polytopes*, second ed., Graduate Texts in Mathematics, vol. 221, Springer-Verlag, New York, 2003, Prepared and with a preface by Volker Kaibel, Victor Klee and Günter M. Ziegler.

[46] B. Haas, *A simple counterexample to Kouchnirenko's conjecture*, Beiträge Algebra Geom. **43** (2002), no. 1, 1–8.

[47] J. Harris and I. Morrison, *Moduli of curves*, Graduate Texts in Mathematics 187, Springer-Verlag, 1998.

[48] D. Hilbert, *Über die Darstellung definiter Formen als Summe von Formenquadraten*, Math. Ann. **32** (1888), 342–350.

[49] B. Huber, F. Sottile, and B. Sturmfels, *Numerical Schubert calculus*, J. Symb. Comp. **26** (1998), no. 6, 767–788.

[50] B. Huber and B. Sturmfels, *A polyhedral method for solving sparse polynomial systems*, Math. Comp. **64** (1995), no. 212, 1541–1555.

[51] B. Huber and J. Verschelde, *Pieri homotopies for problems in enumerative geometry applied to pole placement in linear systems control*, SIAM J. Control and Optim. **38** (2000), 1265–1287.

[52] K. Intriligator, *Fusion residues*, Mod. Phys. Lett. A **6** (1991), 3543–3556.

[53] I. Itenberg and M.-F. Roy, *Multivariate Descartes' rule*, Beiträge zur Algebra und Geometrie **37** (1996), no. 2, 337–346.

[54] I. V. Itenberg, V. M. Kharlamov, and E. I. Shustin, *Welschinger invariant and enumeration of real rational curves*, Int. Math. Res. Not. (2003), no. 49, 2639–2653.

[55] ———, *Logarithmic equivalence of the Welschinger and the Gromov-Witten invariants*, Uspekhi Mat. Nauk **59** (2004), no. 6(360), 85–110.

[56] M. Joswig, *Projectivities in simplicial complexes and colorings of simple polytopes*, Math. Z. **240** (2002), no. 2, 243–259.

[57] M. Joswig and N. Witte, *Products of foldable triangulations*, Adv. Math. **210** (2007), no. 2, 769–796.

[58] V. Kharlamov and F. Sottile, *Maximally inflected real rational curves*, Moscow Math. J. **3** (2003), 947–987, 1199–1200.

[59] A.G. Khovanskii, *A class of systems of transcendental equations*, Dokl. Akad. Nauk. SSSR **255** (1980), no. 4, 804–807.

[60] ———, *Fewnomials*, Trans. of Math. Monographs, 88, AMS, 1991.

[61] S. L. Kleiman, *The transversality of a general translate*, Compositio Math. **28** (1974), 287–297.

[62] F. Klein, *Eine neue Relation zwischen den Singularitäten einer algebraischen Kurve*, Math. Ann. **10** (1876), 199–209.

[63] M. Kontsevich and Yu. Manin, *Gromov-Witten classes, quantum cohomology, and enumerative geometry*, Comm. Math. Phys. **164** (1994), no. 3, 525–562.

[64] J. Lagarias and T. Richardson, *Multivariate Descartes rule of signs and Sturmfels's challenge problem*, Math. Intelligencer **19** (1997), 9–15.

[65] D. Lazard, *Generalized Stewart platform: how to compute with rigid motions?*, DIMACS-SC'93, 1993.

[66] T.-Y. Li, J.M. Rojas, and X. Wang, *Counting real connected components of trinomial curve intersections and m-nomial hypersurfaces*, Discrete Comput. Geom. **30** (2003), no. 3, 379–414.

[67] T.-Y. Li and X. Wang, *On multivariate Descartes' rule—a counterexample*, Beiträge Algebra Geom. **39** (1998), no. 1, 1–5.

[68] I.G. Macdonald, J. Pach, and T. Theobald, *Common tangents to four unit balls in $\mathbb{R}^3$*, Discrete Comput. Geom. **26** (2001), no. 1, 1–17.

[69] G. Megyesi, *Lines tangent to four unit spheres with coplanar centres*, Discrete Comput. Geom. **26** (2001), no. 4, 493–497.

[70] G. Mikhalkin, *Enumerative tropical algebraic geometry in $\mathbb{R}^2$*, J. Amer. Math. Soc. **18** (2005), no. 2, 313–377 (electronic).

[71] E. Mukhin, V. Tarasov, and A. Varchenko, *The B. and M. Shapiro conjecture in real algebraic geometry and the Bethe ansatz*, 2005, Annals of Mathematics, to appear.

[72] ———, *Schubert calculus and representations of general linear group*, 2007, `arXiv:0711.4079`.

[73] J. Plücker, *Solution d'une question fondamentale concernant la théorie générale des courbes*, J. reine agnew. Math. **12** (1834), 105–108.

[74] Z. Ran, *Enumerative geometry of singular plane curves*, Invent. Math. **97** (1989), no. 3, 447–465.

[75] F. Ronga, A. Tognoli, and Th. Vust, *The number of conics tangent to 5 given conics: the real case*, Rev. Mat. Univ. Complut. Madrid **10** (1997), 391–421.

[76] F. Ronga and Th. Vust, *Stewart platforms without computer?*, Real Analytic and Algebraic Geometry, Proceedings of the International Conference, (Trento, 1992), Walter de Gruyter, 1995, pp. 196–212.

[77] Joachim Rosenthal and Frank Sottile, *Some remarks on real and complex output feedback*, Systems Control Lett. **33** (1998), no. 2, 73–80.

[78] Y. Ruffo, J.and Sivan, E. Soprunova, and F. Sottile, *Experimentation and conjectures in the real Schubert calculus for flag manifolds*, Experiment. Math. **15** (2006), no. 2, 199–221.

[79] H. Schubert, *Kalkul der abzählenden Geometrie*, Springer-Verlag, 1879, reprinted with an introduction by S. Kleiman, 1979.

[80] ———, *Anzahl-Bestimmungen für lineare Räume beliebiger Dimension*, Acta. Math. **8** (1886), 97–118.

[81] ———, *Losüng des Charakteritiken-Problems für lineare Räume beliebiger Dimension*, Mittheil. Math. Ges. Hamburg (1886), 135–155, (dated 1885).

[82] M. Shub and S. Smale, *Complexity of Bezout's theorem. II. Volumes and probabilities*, Computational algebraic geometry (Nice, 1992), Progr. Math., vol. 109, Birkhäuser Boston, Boston, MA, 1993, pp. 267–285.

[83] E. Shustin, *Glueing of singular algebraic curves*, Methods of Qualitative Theory, Gorky University Press, 1985, (Russian), pp. 116–128.

[84] B. Siebert and G. Tian, *On quantum cohomology rings of Fano manifolds and a formula of Vafa and Intrilligator*, Asian J. Math. **1** (1997), 679–695.

[85] J. Solomon, *Intersection theory on the moduli space of holomorphic curves with Lagrangian boundary conditions*, `math.SG/0606429`.

[86] E. Soprunova and F. Sottile, *Lower bounds for real solutions to sparse polynomial systems*, Adv. Math. **204** (2006), no. 1, 116–151.

[87] F. Sottile, *Enumerative geometry for real varieties*, Algebraic Geometry, Santa Cruz 1995 (J. Kollár, R. Lazarsfeld, and D. Morrison, eds.), Proc. Sympos. Pure Math., vol. 62, Part 1, Amer. Math. Soc., 1997, pp. 435–447.

[88] ———, *Enumerative geometry for the real Grassmannian of lines in projective space*, Duke Math. J. **87** (1997), no. 1, 59–85.

[89] ———, *The special Schubert calculus is real*, ERA of the AMS **5** (1999), 35–39.

[90] ———, *Real rational curves in Grassmannians*, J. Amer. Math. Soc. **13** (2000), 333–341.

[91] ———, *Real Schubert calculus: Polynomial systems and a conjecture of Shapiro and Shapiro*, Exper. Math. **9** (2000), 161–182.

[92] ———, *Some real and unreal enumerative geometry for flag manifolds*, Mich. Math. J. **48** (2000), 573–592, Special Issue in Honor of Wm. Fulton.

[93] ———, *Enumerative real algebraic geometry*, Algorithmic and quantitative real algebraic geometry (Piscataway, NJ, 2001), DIMACS Ser. Discrete Math. Theoret. Comput. Sci., vol. 60, Amer. Math. Soc., Providence, RI, 2003, on-line version: `www.math.tamu.edu/~sottile/pages/ERAG/index.html`, pp. 139–179.

[94] ———, *Tropical interpolation*, Emissary (MSRI Newletter) (Autumn 2004), On-line version: `www.math.tamu.edu/~sottile/stories/MSRI04/`.

[95] F. Sottile and Th. Theobald, *Lines tangent to $2n-2$ spheres in $\mathbb{R}^n$*, Trans. Amer. Math. Soc. **354** (2002), 4815–4829.

[96] R. Stanley, *Two poset polytopes*, Discrete and Comput. Geom. **1** (1986), 9–23.

[97] R. Stanley, *Some remarks on sign-balanced and maj-balanced posets*, Adv. in Appl. Math. **34** (2005), no. 4, 880–902.

[98] J. Steiner, *Elementare Lösung einer geometrischen Aufgabe, und über einige damit in Beziehung stehende Eigenschaften der Kegelschnitte*, J. Reine Angew. Math. **37** (1848), 161–192.

[99] D. Stewart, *A platform with 6 degree of freedom*, Proc. of the Institution of Mechanical Engineers **180** (1965-66), 371–386.

[100] B. Sturmfels, *On the number of real roots of a sparse polynomial system*, Hamiltonian and gradient flows, algorithms and control, Fields Inst. Commun., vol. 3, American Mathematical Society, Providence, 1994, pp. 137–143.

[101] ______, *Viro's theorem for complete intersections*, Ann. Scuola Norm. Sup. Pisa Cl. Sci. (4) **21** (1994), no. 3, 377–386.

[102] ______, *Gröbner bases and convex polytopes*, American Mathematical Society, Providence, RI, 1996.

[103] ______, *Polynomial equations and convex polytopes*, Amer. Math. Monthly **105** (1998), no. 10, 907–922.

[104] ______, *Solving systems of polynomial equations*, CBMS, vol. 97, American Mathematical Society, Providence, RI, 2002.

[105] C. Vafa, *Topological mirrors and quantum rings*, Essays on Mirror Manifolds, International Press, 1992, ed. by S.-T. Yau, pp. 96–119.

[106] R. Vakil, *Schubert induction*, Ann. of Math. (2) **164** (2006), no. 2, 489–512.

[107] J. Verschelde, P. Verlinden, and R. Cools, *Homotopies exploitating Newton polytopes for solving sparse polynomial systems*, SIAM J. Num. Anal. **31** (1994), 915–930.

[108] O. Ya. Viro, *Gluing of plane real algebraic curves and constructions of curves of degrees* 6 *and* 7, Topology (Leningrad, 1982), Lecture Notes in Math., vol. 1060, Springer, Berlin, 1984, pp. 187–200.

[109] J.-Y. Welschinger, *Invariants of real rational symplectic 4-manifolds and lower bounds in real enumerative geometry*, C. R. Math. Acad. Sci. Paris **336** (2003), no. 4, 341–344.

[110] H. G. Zeuthen, *Almindelige Egenskaber ved Systemer af plane Kurver*, Danske Videnskabernes Selskabs Skrifter, Naturvidenskabelig og Mathematisk, Afd. **10** Bd. IV (1873), 286–393.