The American Mathematical Monthly

ISSN: 0002-9890 (Print) 1930-0972 (Online) Journal homepage: https://www.tandfonline.com/loi/uamm20

Polynomial Equations and Convex Polytopes

Bernd Sturmfels

Polynomial Equations and Convex Polytopes

Bernd Sturmfels

1. INTRODUCTION. One of the fundamental problems in mathematics is that of solving polynomial equations. We all learned in high school how to solve the quadratic polynomial equation

$$a_2x^2 + a_1x + a_0 = 0.$$

It has two distinct solutions

$$x = \frac{-a_1 + \sqrt{a_1^2 - 4a_0a_2}}{2a_2} \quad \text{and} \quad x = \frac{-a_1 - \sqrt{a_1^2 - 4a_0a_2}}{2a_2}, \tag{1.1}$$

provided the *discriminant* $a_1^2 - 4a_0a_2$ is nonzero. The two roots are real numbers if the discriminant is positive. If the discriminant is negative, then we are forced to consider the field of *complex numbers*, denoted $\mathbf{C}$. Indeed, every polynomial with real coefficients,

$$f(x) = a_nx^n + a_{n-1}x^{n-1} + \cdots + a_2x^2 + a_1x + a_0$$

has at least one complex zero x (the *Fundamental Theorem of Algebra*), and, typically, $f(x)$ has n distinct zeros in $\mathbf{C}$. *Galois theory* tells us that for $n \geq 5$ there is no general formula such as (1.1) that expresses the zeros in terms of radicals. A natural question is how many of the roots of $f(x)$ lie in the field $\mathbf{R}$ of real numbers. The answer is severely constrained by the number of terms and the signs of its coefficients. Consider a polynomial with m terms:

$$g(x) = c_1x^{i_1} + c_2x^{i_2} + c_3x^{i_3} + \cdots + c_mx^{i_m} \quad (i_1 > i_2 > i_3 > \cdots > i_m).$$

An index $j \in \{1, \dots, m-1\}$ marks a *sign alternation* if $c_jc_{j+1} < 0$.

Theorem 1.1. (Descartes' Rule of Signs) *The number of positive real roots of a polynomial $g(x)$ is bounded above by the number of its sign alternations.*

This implies that $g(x)$ has at most $2m + 1$ distinct real roots. As an example consider

$$g(x) = x^{53} + x^{48} - x^{36} - x^{17} + x^{11} + x^4 + 1. \tag{1.2}$$

It has 53 distinct complex roots. By Descartes' Rule of Signs, there are at most two positive real roots and (by replacing x by $-x$) at most five negative real roots. In truth, even these bounds are too high: the polynomial in (1.2) has exactly one real root $x \sim -1.019155918$.

In this article we consider the common zeros of d polynomials in d unknowns. Most of our discussion is restricted to the case $d = 2$, to keep things simple and easier to visualize (but it works just as well for $d > 2$). A polynomial in two unknowns looks like

$$f(x, y) = a_1x^{u_1}y^{v_1} + a_2x^{u_2}y^{v_2} + \cdots + a_nx^{u_n}y^{v_n}, \tag{1.3}$$

where the exponents u_i and v_i are nonnegative integers and the coefficients a_i are nonzero real numbers. Its *degree* $deg(f)$ is the maximum of the n numbers $u_1 + v_1, \ldots, u_n + v_n$. The following theorem gives an upper bound on the number of zeros.

Theorem 1.2. (Bezout's Theorem) *Let $f(x, y) = g(x, y) = 0$ be a system of two polynomial equations in two unknowns. If it has only finitely many common complex zeros $(x, y) \in \mathbf{C}^2$, then the number of those zeros is at most $deg(f) \cdot deg(g)$.*

An elementary proof of Bezout's Theorem and related introductory material can be found in the beautiful new text book **[5]** by Cox, Little and O'Shea.

Convex polytopes have been studied since the earliest days of mathematics. We shall see that they are very useful for analyzing and solving polynomial equations. The interplay between polynomials and polytopes can be traced back to work of Isaac Newton on plane curve singularities. Many new results about *Newton polytopes* of polynomial equations, however, have been discovered just in the last ten years or so. It is the aim of this article to convey some of the excitement and the mystery of these interconnections.

A *polytope* is a subset of $\mathbf{R}^d$ that is the convex hull of a finite set of points. A familiar example is the convex hull of $\{(0,0,0), (0,1,0), (0,0,1), (0,1,1), (1,0,0), (1,1,0), (1,0,1), (1,1,1)\}$ in $\mathbf{R}^3$; this is the regular 3-cube. A d-dimensional polytope has many *faces*, which are again polytopes of various dimensions between 0 and $d - 1$. The 0-dimensional faces are called *vertices*, the 1-dimensional faces are called *edges*, and the $(d - 1)$-dimensional faces are called *facets*. For instance, the cube has 8 vertices, 12 edges, and 6 facets. If $d = 2$ then the edges coincide with the facets. A 2-dimensional polytope is called a *polygon*.

Consider the polynomial $f(x, y)$ in (1.3). Each term $x^{u_i}y^{v_i}$ appearing in $f(x, y)$ corresponds to a lattice point (u_i, v_i) in the plane $\mathbf{R}^2$. The convex hull of all these points is called the *Newton polygon* of $f(x, y)$. In symbols we write the Newton polygon as follows:

$$New(f) := conv\{(u_1, v_1), (u_2, v_2), \ldots, (u_n, v_n)\}$$

This is a polygon in $\mathbf{R}^2$ having at most n vertices. More generally, every polynomial in d unknowns gives rise to a *Newton polytope* in $\mathbf{R}^d$.

Our reference example in this paper will be the pair of polynomials

$$\begin{aligned} g(x, y) &= a_1 + a_2 x + a_3 xy + a_4 y, \\ h(x, y) &= b_1 + b_2 x^2 y + b_3 xy^2, \end{aligned} \tag{1.4}$$

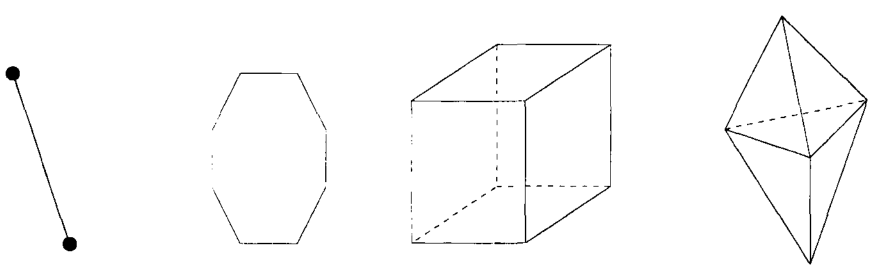

Figure 1.1. Polytopes in dimension one, two, and three

where the coefficients a_i and b_j are nonzero real numbers. The Newton polygon of the polynomial $g(x, y)$ is a quadrangle, and the Newton polygon of $h(x, y)$ is a triangle, as illustrated in Figure 1.2.

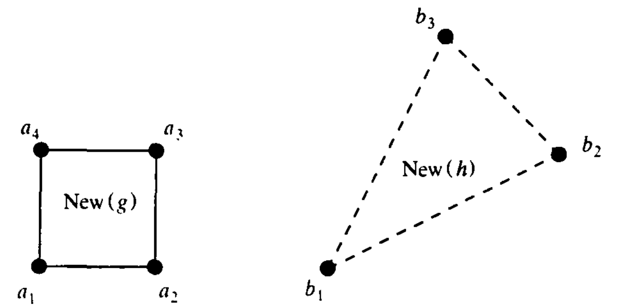

Figure 1.2. Two Newton polygons

If P and Q are any polygons in the plane, then their *Minkowski sum* is the polygon

$$P + Q := \{p + q : p \in P, q \in Q\}. \tag{1.5}$$

Note that each edge of $P + Q$ is parallel to an edge of P or an edge of Q.

The geometric operation of taking the Minkowski sum of polytopes mirrors the algebraic operation of multiplying polynomials. More precisely, the Newton polytope of a product of two polynomials equals the Minkowski sum of two given Newton polytopes:

$$New(g \cdot h) = New(g) + New(h). \tag{1.6}$$

In our example (1.4) the product of the two polynomials equals

$$\begin{aligned} g \cdot h = a_1b_1 &+ a_1b_2x^2y + a_1b_3xy^2 + a_2b_1x + a_2b_2x^3y + a_2b_3x^2y^2 \\ &+ a_3b_1xy + a_3b_2x^3y^2 + a_3b_3x^2y^3 + a_4b_1y + a_4b_2x^2y^2 + a_4b_3xy^3. \end{aligned}$$

The Newton polygon of $g \cdot h$ is exactly the heptagon $P + Q$ in Figure 1.3. As a higher-dimensional example consider the following formula for the regular 3-dimensional cube:

$$\begin{aligned} New((x-1)(y-1)(z-1)) &= New(x-1) + New(y-1) + New(z-1) \\ &= \text{the 3-cube}. \end{aligned}$$

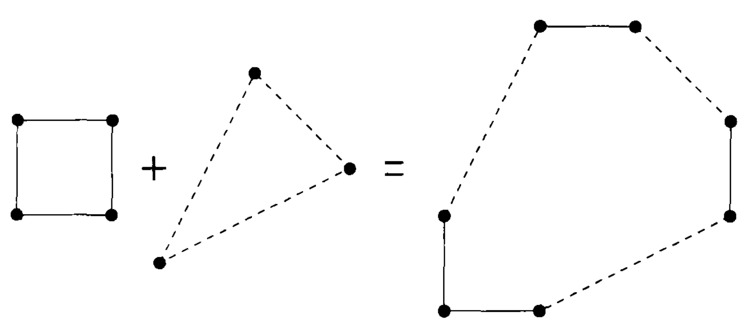

Figure 1.3. Minkowski sum of two polygons

2. COMPLEX ROOTS. We now come to the problem of finding all complex solutions $(x, y) \in \mathbf{C}^2$ of a system of two equations in two unknowns

$$g(x,y) = h(x,y) = 0. \tag{2.1}$$

The two polynomials in our example (1.4) have precisely *four* distinct roots $(x, y) \in \mathbf{C}^2$, provided the coefficients a_i and b_j are "sufficiently generic". The genericity condition takes the form that a certain polynomial in the coefficients, called the *discriminant*, should be nonzero. The discriminant of the system (1.4) is the following expression of degree 14:

$$\begin{aligned}
&4a_1^7a_3b_2^3b_3^3 + a_1^6a_2^2b_2^2b_3^4 - 2a_1^6a_2a_4b_2^3b_3^3 + a_1^6a_4^2b_2^4b_3^2 + 22a_1^5a_2a_3^2b_1b_2^2b_3^3 \\
&+22a_1^5a_3^2a_4b_1b_2^3b_3^2 + 22a_1^4a_2^3a_3b_1b_2b_3^4 + 18a_1a_2a_3a_4^5b_1^2b_2^4 - 30a_1^4a_2a_3a_4^2b_1b_2^3b_3^2 \\
&+a_1^4a_3^4b_1^2b_2^2b_3^2 + 22a_1^4a_3a_4^3b_1b_2^4b_3 + 4a_1^3a_2^5b_1b_3^5 - 14a_1^3a_2^4a_4b_1b_2b_3^4 \\
&+10a_1^3a_2^3a_4^2b_1b_2^2b_3^3 + 22a_1^3a_2^2a_3^3b_1^2b_2b_3^3 + 10a_1^3a_2^2a_4^3b_1b_2^3b_3^2 + 116a_1^3a_2a_3^3a_4b_1^2b_2^2b_3^2 \\
&-14a_1^3a_2a_4^4b_1b_2^4b_3 + 22a_1^3a_3^3a_4^2b_1^2b_2^3b_3 + 4a_1^3a_4^5b_1b_2^5 + a_1^2a_2^4a_3^2b_1^2b_3^4 \\
&+94a_1^2a_2^3a_3^2a_4b_1^2b_2b_3^3 - 318a_1^2a_2^2a_3^2a_4^2b_1^2b_2^2b_3^2 + 396a_1a_2^3a_3a_4^3b_1^2b_2^2b_3^2 \\
&+94a_1^2a_2a_3^2a_4^3b_1^2b_2^3b_3 + 4a_1^2a_2a_3^5b_1^3b_2b_3^2 + 4a_1^2a_3^5a_4b_1^3b_2^2b_3 + a_1^2a_3^2a_4^4b_1^2b_2^4 \\
&+18a_1a_2^5a_3a_4b_1^2b_3^4 - 216a_1a_2^4a_3a_4^2b_1^2b_2b_3^3 + 96a_1a_2^2a_3^4a_4b_1^3b_2b_3^2 \\
&-216a_1a_2^2a_3a_4^4b_1^2b_2^3b_3 + 96a_1a_2a_3^4a_4^2b_1^3b_2^2b_3 - 30a_1^4a_2^2a_3a_4b_1b_2^2b_3^3 \\
&-27a_2^6a_4^2b_1^2b_3^4 + 108a_2^5a_4^3b_1^2b_2b_3^3 + 4a_2^4a_3^3a_4b_1^3b_3^3 - 162a_2^4a_4^4b_1^2b_2^2b_3^2 \\
&- 132a_2^3a_3^3a_4^2b_1^3b_2b_3^2 + 108a_2^3a_4^5b_1^2b_2^3b_3 - 132a_2^2a_3^3a_4^3b_1^3b_2^2b_3 - 27a_2^2a_4^6b_1^2b_2^4 \\
&+16a_2a_3^6a_4b_1^4b_2b_3 + 4a_2a_3^3a_4^4b_1^3b_2^3
\end{aligned}$$

If this expression is nonzero, then the system (1.4) has four distinct complex roots. When the given polynomial system is just a little bigger than (1.4), the discriminant becomes very large and virtually impossible to compute explicitly. Techniques for dealing with this combinatorial explosion and for studying discriminants in general are explained in [**8**].

For computer algebra users we mention that the above expression has been computed in the system Maple by the following sequence of five commands:

```
g:= a1 + a2*x + a3*x*y + a4*y;
h:= b1 + b2*x^2*y + b3*x*y^2;
R:= resultant(g, h, x):
S:= factor(resultant(R, diff(R, y), y)):
discriminant:= op(nops(S), S);
```

This states that the discriminant of g and h of (2.1) is the main factor in the resultant of R and $\partial R/\partial y$, where R is itself the resultant of g and h with respect to the variable x.

If you are now in the vicinity of your computer screen, then try to run such a sequence of commands in Maple, Mathematica, or Reduce. Also, do verify (using the command `solve`) that the two polynomials g and h really have four common roots.

Bezout's Theorem predicts an upper bound of at most $deg(g) \cdot deg(h) = 6$ common complex roots for the equations in (1.4). Indeed, a geometer would expect the cubic curve $\{g = 0\}$ and the quadratic curve $\{h = 0\}$ to intersect in six points.

But these particular curves never intersect in more than four points. How come? To understand why the number is four and not six, we shall take a closer look at the Newton polygons.

If P and Q are any two polygons then we define their *mixed area* as follows:

$$\mathscr{M}(P,Q) := area(P+Q) - area(P) - area(Q), \tag{2.2}$$

where "*area*" denotes the usual area measure of plane Euclidean geometry. For instance, the mixed area of the two Newton polygons in our reference example equals

$$\mathscr{M}(P,Q) = \mathscr{M}(New(g), New(h)) = 4. \tag{2.3}$$

The correctness of this computation can be seen in Figure 2.1, which shows a subdivision of $P+Q$ into five pieces: a translate of P, a translate of Q, and three parallelograms. The mixed area is the sum of the areas of the three parallelograms. The parallelograms (1) and (3) both have area 1, while (2) has area 2. Hence the mixed area equals 4.

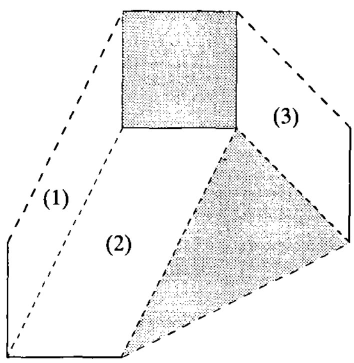

Figure 2.1. Mixed subdivision

In our example the mixed area coincides with the number of common zeros of g and h. This is not an accident, but it is an instance of a general theorem due to David Bernstein [3]. We abbreviate $\mathbf{C}^* := \mathbf{C}\setminus\{0\}$. The set $(\mathbf{C}^*)^2$ of all pairs (x, y) with $x \neq 0$ and $y \neq 0$ is called the *two-dimensional algebraic torus.*

Theorem 2.1. (Bernstein's Theorem) *If g and h are two generic bivariate polynomials, then the number of solutions of $g(x, y) = h(x, y) = 0$ in $(\mathbf{C}^*)^2$ equals the mixed area $\mathscr{M}(New(g), New(h))$.*

Actually, this assertion is valid for *Laurent polynomials,* which means that the exponents in (2.1) can be any integers, possibly negative. Bernstein's Theorem implies the following combinatorial fact: If P and Q are lattice polygons (i.e., the vertices of P and Q have integer coordinates), then $\mathscr{M}(P,Q)$ is a nonnegative integer.

Theorem 2.1 was proved in [3] for d equations in d variables: for generic coefficients the number of zeros in $(\mathbf{C}^*)^d$ is the mixed volume of the Newton polytopes; see also [4]. A much earlier paper on this subject is [16] where a statement equivalent to Theorem 2.1 was derived for $d = 2$. Danilov and Khovanskii [6] use Newton polytopes to compute geometric and topological invariants of the complex variety defined by k general equations in d variables where $k < d$.

To understand Bernstein's Theorem it is best to take a closer look at one of its proofs. We shall describe a proof that is algorithmic in the sense that it tells us how to approximate all the roots numerically. This algorithm appeared in [**9**]. We shall proceed in three steps. The first deals with an easy special case.

Step 1: Binomial systems
A *binomial* is defined to be a polynomial with two terms. We first prove Theorem 2.1 in the case when g and h are binomials. After multiplying or dividing both binomials by suitable scalars and powers of the variables, we may assume that our given equations are

$$g = x^{a_1}y^{b_1} - c_1 \quad \text{and} \quad h = x^{a_2}y^{b_2} - c_2, \tag{2.4}$$

where a_1, a_2, b_1, b_2 are integers (possibly negative) and c_1, c_2 are nonzero real numbers. Note that multiplying the given equations by a (Laurent) monomial changes neither the number of roots in $(\mathbf{C}^*)^2$ nor the mixed area of their Newton polygons.

We compute an invertible integer 2×2-matrix $U = (u_{ij}) \in SL_2(\mathbf{Z})$ such that

$$\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \cdot \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix}.$$

This is accomplished using the *Hermite normal form* algorithm of integer linear algebra. The invertible matrix U triangularizes our system of equations as follows:

$$\begin{aligned} g = h = 0 &\Leftrightarrow x^{a_1}y^{b_1} = c_1 \quad \text{and} \quad x^{a_2}y^{b_2} = c_2 \\ &\Leftrightarrow \left(x^{a_1}y^{b_1}\right)^{u_{11}}\left(x^{a_2}y^{b_2}\right)^{u_{12}} = c_1^{u_{11}}c_2^{u_{12}} \quad \text{and} \\ &\quad\ \left(x^{a_1}y^{b_1}\right)^{u_{21}}\left(x^{a_2}y^{b_2}\right)^{u_{22}} = c_1^{u_{21}}c_2^{u_{22}} \\ &\Leftrightarrow x^{r_1}y^{r_3} = c_1^{u_{11}}c_2^{u_{12}} \quad \text{and} \quad y^{r_2} = c_1^{u_{21}}c_2^{u_{22}}. \end{aligned}$$

This triangularized system has precisely $r_1 r_2$ distinct nonzero complex roots. These can be expressed in terms of radicals in the coefficients c_1 and c_2. The number of roots equals

$$r_1 r_2 = \det\begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix} = \det\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \mathit{area}(\mathit{New}(g) + \mathit{New}(h)).$$

This equals the mixed area $\mathcal{M}(\mathit{New}(g), \mathit{New}(h))$, since the two Newton polygons are just segments, so that $\mathit{area}(\mathit{New}(g)) = \mathit{area}(\mathit{New}(h)) = 0$. This proves Bernstein's Theorem for binomials. Moreover, it gives a simple algorithm for finding all roots in this case.

Step 2: Toric Deformations
We introduce a new indeterminate t, and we multiply each monomial of g and each monomial of h by a power of t. For instance, our system (1.4) gets replaced by

$$\begin{aligned} g_t(x,y) &= a_1 t^{\nu_1} + a_2 x t^{\nu_2} + a_3 xy t^{\nu_3} + a_4 y t^{\nu_4} \\ h_t(x,y) &= b_1 t^{\omega_1} + b_2 x^2 y t^{\omega_2} + b_3 x y^2 t^{\omega_3} \end{aligned} \tag{2.5}$$

We require that the integers ν_i and ω_j are "sufficiently generic" in a sense to be made precise in conditions (1) and (2) below. The system (2.5) can be interpreted as a bivariate system that depends on a parameter t. Its zeros $(x(t), y(t))$ depend on that parameter. They define the branches of an *algebraic function* $t \mapsto (x(t), y(t))$. Our goal is to count the number of branches.

Another interpretation of (2.5) is that of two polynomials in three variables, whose zero set is a curve in $(\mathbf{C}^*)^3$. The *Newton polytopes* of these trivariate polynomials are

$$P := conv\{(0,0,\nu_1),(1,0,\nu_2),(1,1,\nu_3),(0,1,\nu_4)\}$$
$$\text{and} \quad Q := conv\{(0,0,\omega_1),(2,1,\omega_2),(1,2,\omega_3)\} \quad \text{in } \mathbf{R}^3.$$

We consider their Minkowski sum $P + Q$, which is a polytope in $\mathbf{R}^3$. By a *facet* of $P + Q$ we mean a two-dimensional face. A facet F of $P + Q$ is called a *lower facet* if there exists a vector $(u, v) \in \mathbf{Q}^2$ such that $(u, v, 1)$ is an inward pointing normal vector to $P + Q$ at F.

We require the following genericity conditions for the integers ν_i and ω_j:

(1) The Minkowski sum $P + Q$ is a 3-dimensional polytope.
(2) Every lower facet of $P + Q$ has the form $F_1 + F_2$ where either
   (a) F_1 is a vertex of P and F_2 is a facet of Q, or
   (b) F_1 is an edge of P and F_2 is an edge of Q, or
   (c) F_1 is a facet of P and F_2 is a vertex of Q.

As an example consider $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \omega_3 = 0$ and $\omega_1 = \omega_2 = 1$. This doesn't look generic. Nevertheless, our two requirements (1) and (2) are met. Under this lifting P is still a quadrangle and Q is still triangle. But they lie in non-parallel planes in $\mathbf{R}^3$. Their Minkowski sum is a 3-dimensional polytope with 10 vertices, illustrated in Figure 2.2.

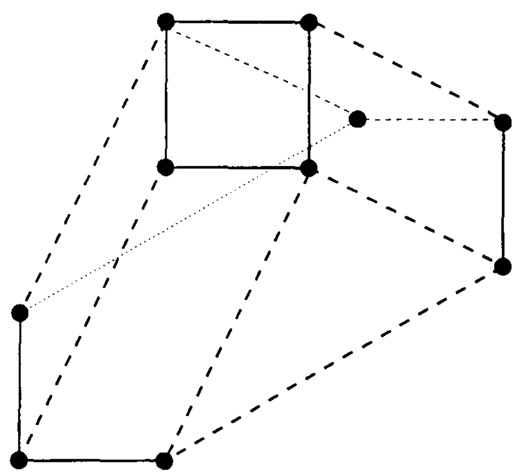

Figure 2.2. The 3-dimensional polytope $P + Q$

The union of all lower facets of $P + Q$ is called the *lower hull* of the polytope $P + Q$. Algebraically speaking, the lower hull is the subset of all points in $P + Q$ at which some linear functional of the form $(x_1, x_2, x_3) \mapsto ux_1 + vx_2 + x_3$ attains its minimum. Geometrically speaking, the lower hull is the part of the boundary of $P + Q$ that is visible from below. This is indicated by the thicker edges in Figure 2.2.

Let $\pi: \mathbf{R}^3 \to \mathbf{R}^2$ denote the projection onto the first two coordinates. Then

$$\pi(P) = New(g), \quad \pi(Q) = New(h), \quad \text{and} \quad \pi(P+Q) = New(g) + New(h).$$

The map π restricts to a bijection from the lower hull onto $New(g) + New(h)$. Let $\Delta := \{\pi(F) : F$ lower facet of $P + Q\}$. This collection of polygons defines a subdivision of $New(g) + New(h)$. A subdivision Δ constructed by this process, for some choice of ν_i and ω_j, is called a *mixed subdivision* of the given Newton polygons. We say that the polygons $\pi(F)$ are the *cells* of the mixed subdivision Δ.

Every cell of a mixed subdivision Δ has the form $F_1 + F_2$ where either

(a) $F_1 = \{(u_i, v_i)\}$ where $x^{u_i}y^{v_i}$ appears in g and F_2 is the projection of a facet of Q, or
(b) F_1 is the projection of an edge of P and F_2 is the projection of an edge of Q, or
(c) F_1 is the projection of a facet of P and $F_2 = \{(u_i, v_i)\}$ where $x^{u_i}y^{v_i}$ appears in h.

The cells of type (b) are called the *mixed cells* of Δ. We record the following fact.

Proposition 2.2. *If Δ is any mixed subdivision of the given Newton polygons, then the mixed area $\mathscr{M}(\mathrm{New(g), New(h)})$ equals the sum of the areas of the mixed cells in Δ.*

The proof of Proposition 2.2 is not hard, but is omitted. Let us now return to the proof of Theorem 2.1, entering the final step of our algorithm.

Step 3: Puiseux series

Consider the roots of (2.5) as an algebraic function $t \mapsto (x(t), y(t))$. In a neighborhood of the origin in the complex plane, each branch of this function can be written as follows:

$$\begin{aligned} x(t) &= x_0 \cdot t^u + \text{higher order terms in } t, \\ y(t) &= y_0 \cdot t^v + \text{higher order terms in } t, \end{aligned} \tag{2.6}$$

where x_0, y_0 are nonzero complex numbers and u, v are rational numbers. To determine the exponents u and v we substitute the tentative roots (2.6) into the equations $g_t = f_t = 0$. In our example (2.5) this gives

$$\begin{aligned} g_t(x(t), y(t)) &= a_1 t^{\nu_1} + a_2 x_0 t^{u+\nu_2} + a_3 x_0 y_0 t^{u+v+\nu_3} + a_4 y_0 t^{v+\nu_4} + \cdots, \\ h_t(x(t), y(t)) &= b_1 t^{\omega_1} + b_2 x_0^2 y_0 t^{2u+v+\omega_2} + b_3 x_0 y_0^2 t^{u+2v+\omega_3} + \cdots. \end{aligned} \tag{2.7}$$

In order for (2.6) to be a root the term of lowest order must vanish. Since x_0 and y_0 are chosen to be nonzero, this is possible only if the lowest order in t is attained by at least two different terms. This imposes certain restrictions on the vector (u, v), which can be expressed algebraically as systems of linear equations and linear inequalities.

In our example (2.7) the restrictions amount to a disjunction of conditions like

$$\begin{aligned} &\nu_1 = v + \nu_4 \leq min\{u + \nu_2, u + v + \nu_3\} \quad \text{and} \\ &\omega_1 = u + 2v + \omega_3 \leq 2u + v + \omega_2. \end{aligned} \tag{2.8}$$

The geometric meaning of each such linear system is as follows.

Lemma 2.3. *A pair of rational numbers (u, v) determines the orders of the lowest terms in a series solution $(x(t), y(t))$ of the equations $g_t(x, y) = h_t(x, y) = 0$ if and only if $(u, v, 1)$ is the normal vector to a mixed lower facet of the polytope $P + Q$.*

The "only if" direction in Lemma 2.3 is easily derived from the considerations above. However, the "if"-direction does need further justification. In [9] the

Implicit Function Theorem is invoked for this purpose. Algebraic proofs that work over any algebraically closed field are based on the technique of *Puiseux series*; see [14], [15]. Puiseux series are elements in the field $\mathbf{C}((t))$ of fractional power series. The expressions (2.6) represent such series.

Lemma 2.3 implies that the valid choices of (u, v) are in bijection with the mixed cells in the mixed subdivision Δ. Each mixed cell of Δ is expressed uniquely as the Minkowski sum of a Newton segment $New(g')$ and a Newton segment $New(h')$, where g' is a binomial consisting of two terms of g and h' is a binomial consisting of two terms of h. Thus each mixed cell in Δ can be identified with a system of two binomial equations $g'(x, y) = h'(x, y) = 0$. In this situation we can rewrite (2.7) as follows:

$$g_t(x(t), y(t)) = g'(x_0, y_0) \cdot t^a + \text{higher order terms in } t,$$

$$h_t(x(t), y(t)) = h'(x_0, y_0) \cdot t^b + \text{higher order terms in } t,$$

where a and b suitable rational numbers. This implies the following lemma.

Lemma 2.4. *Let (u, v) as in Lemma 2.3. Then the corresponding choices of $(x_0, y_0) \in (\mathbf{C}^*)^2$ are precisely the nonzero roots of the binomial system $g'(x_0, y_0) = h'(x_0, y_0) = 0$.*

We are now prepared to complete the proof of Theorem 2.1. By Step 1, the number of roots (x_0, y_0) of the binomial system in Lemma 2.4 coincides with the area of the mixed cell $New(g') + New(h')$. Each of these roots provides the leading coefficients in a series solution (2.6) of the equations (2.5). Conversely, Lemma 2.3 ensures that every series solution arises from some mixed cell of Δ. We conclude that the number of series solutions equals the sum of these areas over all mixed cells in Δ. By Proposition 2.2, this quantity coincides with the mixed area $\mathscr{M}(New(f), New(g))$. General facts from algebraic geometry guarantee that the same number of roots is attained for almost all choices of coefficients. ∎

We illustrate Step 3 in the proof of Theorem 2.1 for the following toric deformation:

$$g_t(x, y) = a_1 + a_2 x + a_3 xy + a_4 y, \quad h_t(x, y) = b_1 t + b_2 x^2 y t + b_3 x y^2. \quad (2.9)$$

Figure 2.2 depicts the polytope $P + Q$ for this choice of parameters ν_i and ω_j. The polytope $P + Q$ has three mixed lower facets. Their normal vectors are

(1) $(u, v, 1) = (1, 0, 1)$,
(2) $(u, v, 1) = (0, 1/2, 1)$,
(3) $(u, v, 1) = (-1, 0, 1)$.

Here the labeling (1)–(3) is as in Figure 2.1, under the identification of mixed lower facets of $P + Q$ with mixed cells in the mixed subdivision Δ. Note that $u = 1, v = 0$ in (1) is the solution to (2.8). The binomial systems corresponding to these three mixed cells are

(1) $f'(x_0, y_0) = a_1 + a_4 y_0, \quad g'(x_0, y_0) = b_1 + b_3 x_0 y_0^2$.
(2) $f'(x_0, y_0) = a_1 + a_2 x_0, \quad g'(x_0, y_0) = b_1 + b_3 x_0 y_0^2$.
(3) $f'(x_0, y_0) = a_2 x_0 + a_3 x_0 y_0, \quad g'(x_0, y_0) = b_2 x_0^2 y_0 + b_3 x_0 y_0^2$.

These binomial systems have one, two, and one root respectively. These numbers are the areas of *New*(f') + *New*(g'), as explained in Step 1. For instance, the root of (1) equals

$$x_0 = -\frac{a_4^2 b_1}{a_1^2 b_3} \quad \text{and} \quad y_0 = -\frac{a_1}{a_4}. \tag{2.10}$$

Hence our system has four branches. If one wishes more information about the four branches, one can now compute further terms in the Puiseux expansions of these branches. A general algorithm for doing this is given in **[14]**. For instance, the Puiseux series contributed by the mixed cell (1) starts out like this:

$$\begin{aligned} x(t) &= -\frac{a_4^2 b_1}{a_1^2 b_3} \cdot t + 2 \cdot \frac{a_4^3 b_1^2 (a_1 a_3 - a_2 a_4)}{a_1^5 b_3^2} \cdot t^2 \\ &\quad + \frac{a_4^4 b_1^2 \left(a_1^3 a_4 b_2 - 5a_1^2 a_3^2 b_1 + 12 a_1 a_2 a_3 a_4 b_1 - 7 a_2^2 a_4^2 b_1\right)}{a_1^8 b_3^8} \cdot t^3 + \cdots \qquad (2.11) \\ y(t) &= -\frac{a_1}{a_4} + \frac{b_1 (a_1 a_3 - a_2 a_4)}{a_1^2 b_3} \cdot t \\ &\quad + \frac{a_4 b_1^2 (a_1 a_3 - a_2 a_4)(a_1 a_3 - 2 a_2 a_4)}{a_1^5 b_3^2} \cdot t^2 + \cdots \end{aligned}$$

Our proof of Bernstein's Theorem gives rise to a numerical algorithm for finding of all roots of a sparse system of polynomial equations. This algorithm belongs to the general class of *numerical continuation* methods **[1]**, which are sometimes also called *homotopy methods* **[7]**. The idea is to trace each of the branches of the algebraic curve $(x(t), y(t))$ between $t = 0$ and $t = 1$. We have shown that the number of branches equals the mixed area. Our constructions give sufficient information about the Puiseux series so that we can approximate $(x(t), y(t))$ for any t in a small neighborhood of zero. Using numerical continuation, it is now possible to approximate $(x(1), y(1))$. For details see **[9]** or **[19]**.

We remark that Bezout's Theorem 1.2 can be derived as a special case from Bernstein's Theorem 2.1. We illustrate this for a pair of general (dense) bivariate polynomials $g(x, y)$ and $h(x, y)$ where g has degree 2 and h has degree 3. Their Newton polygons are triangles:

$$\begin{aligned} New(g) &= conv\{(0,0),(2,0),(0,2)\} \quad \text{and} \\ New(h) &= conv\{(0,0),(3,0),(0,3)\}. \end{aligned}$$

We must show that their mixed area equals six.

The identity $\mathcal{M}(New(g), New(h)) = 6$ is proved by the picture in Figure 2.3. Indeed, there are precisely six mixed cells in Figure 2.3, and each of them has unit

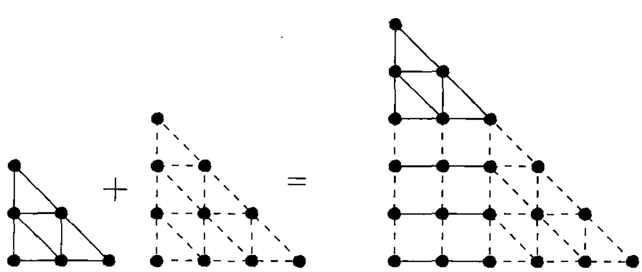

Figure 2.3. Mixed subdivision proving Bezout's Theorem

area. Hence the triangles $New(g)$ and $New(h)$ have mixed area 6. Using Bernstein's Theorem, we conclude that the expected number of complex zeros of $g(x, y) = h(x, y) = 0$ equals 6.

3. REAL ROOTS. Polynomial equations arise in many mathematical models in science and engineering. In such applications one is typically interested in solutions over the real numbers $\mathbf{R}$ instead of the complex numbers $\mathbf{C}$. This study of real roots of polynomial systems is considerably more difficult than the study of complex roots. Even the most basic questions remain unanswered to-date. Let us start out with a very concrete question:

Problem 3.1. *What is the maximum number of isolated real roots of any system of two polynomial equations in two variables each having four terms?*

The polynomial equations considered here look like $f(x, y) = g(x, y) = 0$, where

$$\begin{aligned} f(x,y) &= a_1 x^{u_1} y^{v_1} + a_2 x^{u_2} y^{v_2} + a_3 x^{u_3} y^{v_3} + a_4 x^{u_4} y^{v_4}, \\ g(x,y) &= b_1 x^{\tilde{u}_1} y^{\tilde{v}_1} + b_2 x^{\tilde{u}_2} y^{\tilde{v}_2} + b_3 x^{\tilde{u}_3} y^{\tilde{v}_3} + b_4 x^{\tilde{u}_4} y^{\tilde{v}_4}, \end{aligned} \tag{3.1}$$

a_i, b_j are arbitrary real numbers, and $u_i, v_j, \tilde{u}_i, \tilde{v}_j$ are arbitrary integers. To stay consistent with our discussion in Section 2, we shall count only solutions (x, y) in $(\mathbf{R}^*)^2$, that is, we require that both x and y are nonzero real numbers.

A priori it is not clear whether Problem 3.1 even makes sense: why should such a maximum exist? It certainly does not exist if we consider complex zeros, because one can get arbitrarily many complex zeros by increasing the degrees of the equations. The point is that such an unbounded increase of roots is impossible over the real numbers. In fact, there is a precise conjecture what the solution to Problem 3.1 should be: *thirty-six.*

It is easy to write down a system of the form (3.1) that has 36 real roots:

$$f(x) = (x^2-1)(x^2-2)(x^2-3) \quad \text{and} \quad g(y) = (y^2-1)(y^2-2)(y^2-3). \tag{3.2}$$

Each of the polynomials f and g depends on one variable only, and it has 6 nonzero real roots in that variable. Therefore the system $f(x) = g(y) = 0$ has 36 distinct isolated roots in $(\mathbf{R}^*)^2$. Note also that the expansions of f and g have exactly four terms each, as required. Embarrassingly, the system (3.2) is the best example we know at present. Writing $\mathbf{R}_+$ for the set of positive real numbers, we have the following general conjecture:

Conjecture 3.2. (Kouchnirenko's Conjecture) *Consider any system of d polynomial equations in d unknown, where the i-th equation has at most m_i terms. The number of isolated real roots in $(\mathbf{R}_+)^d$ of such a system is at most $(m_1 - 1)(m_2 - 1)\cdots(m_d - 1)$.*

This number is attained by equations in distinct variables, as was demonstrated in (3.2). If we wish to consider roots in $(\mathbf{R}^*)^d$ instead of $(\mathbf{R}_+)^d$ in Conjecture 3.2, then we simply multiply the asserted bound by 2^d.

There is only one case when Kouchnirenko's Conjecture is known to be true: the case $d = 1$ of a single polynomial in one variable. The number of real positive roots is bounded above by the number of terms minus 1. This is a consequence of Descartes' Rule of Signs (Theorem 1.1). Kouchnirenko's Conjecture remains open in all higher dimensions. However, a major breakthrough was accomplished by Khovanskii [11]. He found a bound on the number of real roots that does not depend on the degrees of the given equations.

Theorem 3.3. (Khovanskii's Theorem on Fewnomials) *Consider d polynomials in d variables involving n distinct monomials in total. The number of isolated roots in the positive orthant $(\mathbf{R}_+)^d$ of any such system is at most $2^{\binom{n}{2}} \cdot (d+1)^n$.*

The basic idea behind the proof of Khovanskii's Theorem is to establish the following more general result. We consider systems of d equations that can be expressed as polynomial functions in at most n monomials in $\mathbf{x} = (x_1, \dots, x_d)$. If we abbreviate the i-th such monomial by $\mathbf{x}^{\mathbf{a}_i} := x_1^{a_{i1}} x_2^{a_{i2}} \cdots x_n^{a_{in}}$, then we can write our d polynomials as

$$F_i(\mathbf{x}^{\mathbf{a}_1}, \mathbf{x}^{\mathbf{a}_2}, \dots, \mathbf{x}^{\mathbf{a}_n}) = 0 \quad (i = 1, 2, \dots, d) \tag{3.3}$$

We claim that the number of real roots of (3.3) in the positive orthant is at most

$$2^{\binom{n}{2}} \cdot \left(1 + \sum_{i=1}^{d} deg(F_i)\right)^n \cdot \prod_{i=1}^{d} deg(F_i).$$

Theorem 3.3 concerns the case where $deg(F_i) = 1$ for all i.

We proceed by induction on $n - d$. If $n = d$ then (3.3) is expressed in d monomials in d unknowns. By a multiplicative change of variables

$$x_i \mapsto z_1^{u_{i1}} z_2^{u_{i2}} \cdots z_d^{u_{id}}$$

we can transform our d monomials into the d coordinate functions $z_1, z_2, \dots, z_d$. (Here the u_{ij} can be rational numbers, since all roots under consideration are positive reals.) Our assertion follows from Bezout's Theorem, which states that the number of isolated complex roots is at most the product of the degrees of the equations.

Now suppose $n > d$. We introduce a new variable t, and we multiply one of the given monomials by t. For instance, we may do this to the first monomial and set

$$G_i(t, x_1, \dots, x_d) := F_i(\mathbf{x}^{\mathbf{a}_1} \cdot t, \mathbf{x}^{\mathbf{a}_2}, \dots, \mathbf{x}^{\mathbf{a}_n}) \quad (i = 1, 2, \dots, d)$$

This is a system of equations in $\mathbf{x}$ depending on the parameter t. We study the behavior of its positive real roots as t moves from 0 to 1. At $t = 0$ we have a system involving one monomial less, so the induction hypothesis provides a bound on the number of roots. Along our trail from 0 to 1 we encounter some bifurcation points at which two new roots are born. Hence the number of roots at $t = 1$ is at most twice the number of bifurcation points plus the number of roots of $t = 0$.

Each bifurcation point corresponds to a root $(\mathbf{x}, t)$ of the augmented system

$$J(t, \mathbf{x}) = G_1(t, \mathbf{x}) = \cdots = G_d(t, \mathbf{x}) = 0, \tag{3.4}$$

where $J(t, \mathbf{x})$ denotes the *toric Jacobian*:

$$J(t, x_1, \dots, x_d) = det\left(x_i \cdot \frac{\partial}{\partial x_j} G_j(t, \mathbf{x}) \right)_{1 \le i, j \le d}.$$

Now, the punch line is that each of the $d + 1$ equations in (3.4), including the Jacobian, can be expressed in terms of only n monomials $\mathbf{x}_1^{\mathbf{a}} \cdot t, \mathbf{x}_2^{\mathbf{a}}, \dots, \mathbf{x}_n^{\mathbf{a}}$. Therefore we can bound the number of bifurcation points by the induction hypothesis, and we are done.

This was only to give the flavor of how Theorem 3.3 is proved; several combinatorial and topological fine points need careful attention. The complete proof is in [**2**] and [**12**].

Khovanskii's Theorem implies an upper bound for the root count suggested in Problem 3.1. After multiplying one of the given equations by a suitable monomial, we may assume that the system (3.1) has seven distinct monomials. Substituting $d = 2$ and $n = 7$ into Khovanskii's formula, we see that (3.1) has at most $2^{\binom{7}{2}} \cdot (2+1)^7 = 4{,}586{,}471{,}424$ roots in the positive quadrant. By summing over all four quadrants, we conclude that the maximum in Problem 3.1 lies between 36 and $18{,}345{,}885{,}696 = 2^2 \cdot 2^{\binom{7}{2}} \cdot (2+1)^7$. The gap between 36 and 18,345,885,696 is frustratingly large. Experts agree that the truth should be closer to the lower bound than to the upper bound, but at the moment no one seems to have a clue how to decide whether 36 is correct.

One approach pursued in the recent literature is to refine Conjecture 3.2 by taking the Newton polytopes of the given equations into account. We shall analyze the behavior of the number of real roots under a toric deformation of the form (2.5). In analogy to our discussion in Section 2, we must first understand the special case of a binomial system

$$c_1 x^{a_1} y^{b_1} + c_2 x^{a_2} y^{b_2} = c_3 x^{a_3} y^{b_3} + c_4 x^{a_4} y^{b_4} = 0, \tag{3.5}$$

where the c_i are nonzero reals. We consider the roots in the open positive quadrant $(\mathbf{R}_+)^2$.

Lemma 3.4. *The system* (3.5) *has precisely one solution in* $(\mathbf{R}_+)^2$ *if and only if* $c_1 c_2 < 0$ *and* $c_3 c_4 < 0$. *In all other cases it has no solution in* $(\mathbf{R}_+)^2$.

Proof: If $c_1 c_2 > 0$ or $c_3 c_4 > 0$ then there are no positive numbers x and y satisfying (3.5). If $c_1 c_2 < 0$ and $c_3 c_4 < 0$ then we write the binomial system (3.5) in the form (2.4) where c_1 and c_2 are positive. The subsequent transformations in Step 1 of Section 2 show that such a system (2.4) has precisely one positive real root (x, y). ■

Using Lemma 3.4 we can easily determine the number of roots of (3.5) in the other three quadrants and hence in all of $(\mathbf{R}^*)^2$. To this end we simply replace x by $-x$ and/or y by $-y$ in (3.5) and look at the signs of the coefficients of the resulting binomials.

We now consider an arbitrary system of two equations in two unknowns, and we fix a toric deformation $g_t(x, y) = h_t(x, y) = 0$ that is sufficiently generic in the sense of Step 2 in Section 2. All we need to know from the coefficients of g and h are their signs.

Each mixed cell in the mixed subdivision Δ corresponds to a pair of binomials $g'(x, y) = h'(x, y) = 0$. Depending on the signs of their coefficients, these binomials have either one common root or no common root in $(\mathbf{R}_+)^2$, by Lemma 3.4. If they have one common root, then we call that mixed cell *alternating*. This has the following pictorial explanation. Each node of Δ is expressed uniquely as a sum $(u_i, v_i) + (u_j, v_j)$, where $c_i x^{u_i} y^{v_i}$ is a term of g and $c_j x^{u_j} y^{v_j}$ is a term of h. We label that node by the sign vector $(sign(c_i), sign(c_j))$ when drawing the mixed subdivision

Δ. A mixed cell is alternating if and only if all four possible sign vectors $(+,+),(-,+),(+,-),(+,+)$ appear as the labels of its four vertices.

Theorem 3.5. *There exists $\epsilon > 0$ such that, for all $0 < t < \epsilon$, the number of zeros of system $g_t(x, y) = h_t(x, y) = 0$ in $(\mathbf{R}_+)^2$ equals the number of alternating mixed cells in Δ.*

We illustrate Theorem 3.5 for our reference example (2.5). Let us assume that the signs are distributed as follows:

$$a_1, a_3, b_1 > 0 \quad \text{and} \quad a_2, a_4, b_2, b_3 < 0. \tag{3.6}$$

Then the mixed subdivision Δ in Figure 2.1 gets its nodes labeled as shown in Figure 3.1.

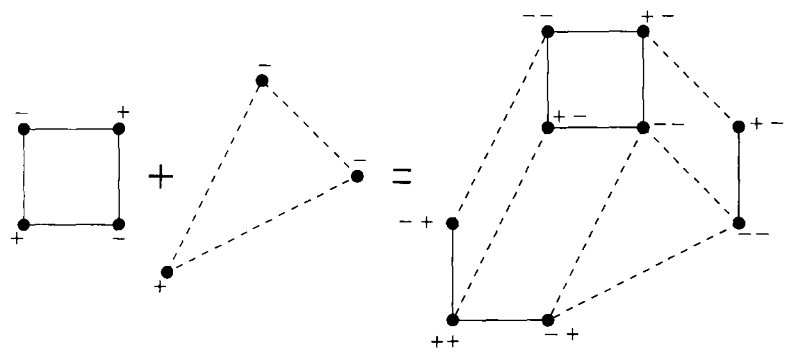

Figure 3.1. A signed mixed subdivision

Two of the three mixed cells are alternating. Hence for $t > 0$ sufficiently small the system (2.5) has precisely two positive real roots. For instance, consider the leftmost mixed cell, which was labeled (1) in Figure 2.1. Under the hypotheses (3.6), the expressions for x_0 and y_0 in (2.10) are both positive. The series $x(t)$ and $y(t)$ in (2.11) converge to positive real numbers for $t > 0$ sufficiently small.

Proofs of Theorem 3.5 can be found in **[10]** and **[17]**. There is a close connection to the *Viro construction* for studying the topology of real algebraic hypersurfaces; see **[20]** and **[8**, Section 11.5**]**. A common generalization of Viro's construction and Theorem 3.5 appears in **[18]**. That paper gives a combinatorial construction for the asymptotic topology of the real solution set of a system of $k \leq d$ equations in d unknowns. Here "asymptotic" means for $t \to 0+$ in a toric deformation of the k equations.

A difficult issue in Viro's construction and in Theorem 3.5 is how to precisely quantify "sufficiently small". We shall explain how the value of ϵ can be computed for our reference example (2.5). Consider the discriminant of the original system (1.4), which is the big polynomial of degree 14 displayed in the beginning of Section 2. In the discriminant replace a_i by $a_i t^{\nu_i}$ for $i = 1, 2, 3, 4$ and replace b_j by $b_j t^{\omega_j}$ for $j = 1, 2, 3$. Consider the resulting expression as a polynomial in one variable t. Let $t = \epsilon$ be the smallest positive root of that polynomial. This choice of $\epsilon > 0$ has the property asserted in Theorem 3.5.

Itenberg and Roy **[10]** asked whether the combinatorial construction of Theorem 3.5 yields an upper bound for *all* choices of coefficients, not just those appearing in the limit of a toric deformation. Let $g(x, y)$ and $h(x, y)$ be arbitrary polynomials with fixed support. For each monomial in the support of g we record only the sign of the corresponding coefficient, and similarly for h. For any mixed subdivision Δ this induces a labeling of the nodes of Δ with sign vectors $(+,+),(+,-),(-,+),(-,-)$, and we can count the number of alternating mixed cells. We define the *combinatorial upper bound* for the system $g(x, y) = h(x, y) = 0$ to be the maximum number of alternating mixed cells in any mixed subdivision Δ. Itenberg & Roy **[10]** conjectured that the number of real positive roots of any polynomial system $g(x, y) = h(x, y) = 0$ is bounded above by the combinatorial upper bound.

This conjecture was disproved very recently by Li and Wang **[13]**. Their example is the following strikingly simple system of two equations in two unknowns:

$$g(x,y) = y - x - 1 \quad \text{and} \quad h(x,y) = x^3y^3 + 100y^3 - 900x^3 - 200. \quad (3.7)$$

Here the combinatorial upper bound is only two, but there are three positive real roots:

$$(0.317659, 1.317659), \quad (0.659995, 1.659995), \quad (8.120576, 9.120576).$$

See [13] for the derivation of the combinatorial upper bound for (3.7).

The Li-Wang counterexample suggests that also Kouchnirenko's Conjecture 3.2 may be too optimistic. Perhaps the true bound is closer to the one given in Khovanskii's Theorem 3.3? Clearly we are still far from fully understanding the real solutions of systems of polynomial equations. Much work remains to be done in this fascinating area, which is full of challenges for combinatorialists, algebraic geometers, and applied mathematicians.

Note added in proof: Bill Fulton pointed out that Kouchnirenko's conjecture is false if we do not require the real roots to be isolated over the complex numbers. A counterexample appears in Example 13.6 of W. Fulton, *Intersection Theory*, Springer, Berlin, 1984: Take $d = 3$, m at least 5,

$$f = \prod_{i=1}^{m} (x - i)^2 + \prod_{i=1}^{m} (y - i)^2,$$

$g = x(z - 1)$, and $h = y(z - 1)$. Then Kouchnirenko's bound is

$$(4m + 1 - 1)(2 - 1)(2 - 1) = 4m,$$

but there are m^2 real roots of the form $(i, j, 1)$, for i and j between 1 and m.

REFERENCES

1. Allgower, E. and Georg, K., *Numerical Continuation Methods*, Springer Verlag, 1990.
2. Benedetti, R. and Risler, J. -J., *Real algebraic and semi-algebraic sets*, Actualités Mathematiques, Hermann, Paris, 1990.
3. Bernstein, D. N., The number of roots of a system of equations, *Functional Analysis Appl.* **9** (1975) 1–4.
4. Bernstein, D. N., Kouchnirenko, A. G., and Khovanskii, A. G., Newton polyhedra. *Uspekhi Mat. Nauk.* **31** (1976) 201–202.
5. Cox, D., Little, J., and O'Shea, D., *Using Algebraic Geometry*, Springer Verlag, 1997.
6. Danilov, V. I. and Khovanskii, A. G., Newton polyhedra and an algorithm for computing Hodge-Deligne numbers, *Math. USSR Izvestiya* **29** (1987) 279–298.
7. Drexler, F. J., A homotopy method for the calculation of zeros of zero dimensional ideals, in *Continuation Methods*, (H.G. Wacker, Ed.), Academic Press, New York, 1978.

8. Gel'fand, I. M., Kapranov, M. M., and Zelevinsky, A. V., *Discriminants, Resultants and Multi-Dimensional Determinants*, Birkhäuser, Boston. 1994.
9. Huber, B. and Sturmfels, B., A polyhedral method for solving sparse polynomial systems, *Mathematics of Computation* **64** (1995) 1541–1555.
10. Itenberg, I. and Roy, M. -F., Multivariate Descartes' rule, *Beiträge Algebra Geom.* **37** (1996) 337–346.
11. Khovanskii, A. G., On a class of systems of transcendental equations, *Soviet Math. Doklady* **22** (1980) 762–765.
12. Khovanskii, A. G., *Fewnomials*, American Mathematical Society, Translations of Mathematical Monographs, Vol. **88**, 1991.
13. Li, T. Y. and Wang, X., On multivariate Descartes' rule—a counterexample, *Beiträge Algebra Geom.* **39** (1998) 1–5.
14. McDonald, J., Fiber polytopes and fractional power series, *J. Pure Appl. Algebra* **104** (1995) 213–233.
15. Maurer, J., Puiseux expansions for space curves. *Manuscripta Math.* **32** (1980) 91–100.
16. Minding, F., Über die Bestimmung des Grades der durch Elimination hervorgehenden Gleichung, *J. Reine Angew. Math.* **22** (1841) 178–183.
17. Pedersen, P. and Sturmfels, B., Mixed monomial bases, in *Algorithms in Algebraic Geometry and Applications*, (eds. L. Gonzales-Vega and T. Recio), Progress in Mathematics, Vol. 143, Birkhäuser, Basel, 1994, pp. 307–316.
18. Sturmfels, B., Viro's theorem for complete intersections, *Ann. Scuola Norm. Sup. Pisa* **21** (1994), 377–386.
19. Verschelde, J., Verlinden, P., and Cools, R., Homotopies exploiting Newton polytopes for solving sparse polynomial systems, *SIAM J. Numer. Anal.* **31** (1994) 915–930.
20. Viro, O., Gluing of plane real algebraic curves and constructions of curves of degrees 6 and 7, Lecture Notes in Math. **1060**, Springer Verlag, 1984, pp. 187–200.

BERND STURMFELS received his Ph.D. in 1987 at the University of Washington under the supervision of Victor Klee. After postdoctoral years in Minneapolis and Linz, Austria, he taught at Cornell University for six years, before moving permanently to UC Berkeley. Sturmfels has been a Sloan Fellow, an NSF National Young Investigator, and a David and Lucile Packard Fellow. He has authored five books and 90 research articles in combinatorics, computational algebra, and algebraic geometry. With his wife, Hyungsook, and two young children, Nina and Pascal, he spent the academic year 1997–98 in Kyoto, Japan.
University of California, Berkeley, CA 94720
bernd@math.berkeley.edu

Nowadays many mathematical books do not seem to be written by living men who not only know, but doubt and ask and guess, who see details in their true perspective—light surrounded by darkness —who, endowed with a limited memory, in the twilight of questioning, discovery, and resignation, weave a connected pattern, imperfect but growing, and colored by infinite gradations of significance. The books of the type I refer to are rather like slot machines which fire at you for the price you pay a medley of axioms, definitions, lemmas, and theorems, and then remain numb and dead however you shake them.

From Hermann Weyl's book review of volume 2 of Courant and Hilbert's *Methoden der Mathematischen Physik*. The review appeared in *Bull. Amer. Math. Soc.* 44 (1938) 602–604. The quotation is from page 602.

Contributed by Harold Boas