1 Introduction

At first sight solutions to systems of polynomial equations and convex polytopes do not seem to be related at all. However, there is a very direct connection between the two. Newton pioneered the study, but it was in recent years that many important (and fascinating) results have been discovered. Bernd Sturmfels presents one such result and provides an algorithmic proof in [7]. It suggests an approach to both counting exactly and approximating solutions to $d$ equations in $d$ variables over $(\mathbb{C}^*)^d$. Here, the approach will be presented in general terms, i.e. the emphasis is on intuition rather than overwhelming detail.

We shall start by reviewing the basic concepts, which will allow the subsequent introduction of Bernstein’s Theorem. This result shall then be discussed in some detail. Two main consequences will be presented: complex root count and real root count. The latter is in a section of its own as it is here merely to tease (but you’ll have to read it all to see why!).

2 The Newton Polytope

Let $f \in \mathbb{C}[x_1, \ldots, x_d]$ be a polynomial in $d$ variables. It can be written as

$$f(x_1, \ldots, x_d) = \sum_{i=1}^{m} c_i x_1^{e_{1,i}} x_2^{e_{2,i}} \cdots x_d^{e_{d,i}}$$

where $m$ is smallest possible. A Laurent polynomial is one that is written exactly as above, but where negative powers are allowed. We define the Newton polytope of $f$ as

$$\text{New}(f) = \text{conv} \ \{(e_{1,1}, e_{2,1}, \ldots, e_{d,1}), \ldots, (e_{1,m}, e_{2,m}, \ldots, e_{d,m})\}$$

Before going any further, consider the following example. Let

$$g(x, y) = a_1 + a_2 x + a_3 xy + a_4 y$$
$$h(x, y) = b_1 + b_2 x^2 y + b_3 xy^2$$

(2)
Figure 1: The Newton polytopes of the polynomials in Eq. 2.

Thus,

\[ \text{New}(g) = \text{conv}\{ (0, 0), (1, 0), (1, 1), (0, 1) \} \]
\[ \text{New}(h) = \text{conv}\{ (0, 0), (2, 1), (1, 2) \} \].

See figure 1.

The next result shows a nice relation among Newton polytopes.

**Proposition 1.** The Newton polytope of a product of two polynomials is the Minkowski sum of the two given Newton polytopes, i.e.

\[ \text{New}(g \cdot h) = \text{New}(g) + \text{New}(h) \]

where the right-hand side denotes the Minkowski sum of the two polytopes.

**Proof.** (See Figure 2b for the Minkowski sum of the example in Figure 1) To see this, we first prove the following claim.

**Claim 1.** Let \( P \) and \( Q \) be polytopes. Then, their Minkowski sum \( P + Q \) is given by the convex hull of all points of the form

\[ p + q \quad p \in \text{ext}(P), q \in \text{ext}(Q) \]

where \( \text{ext}(P) \) is the set of extreme points of \( P \).

**Proof.** If \( p + q \) is an extreme point of \( P + Q \), suppose w.l.o.g. that \( p = \frac{p' + p''}{2} \). Then, we have

\[ p + q = \frac{p' + q + p'' + q}{2} \]

and, therefore, \( p' + q = p'' + q \), which means that \( p' = p'' \). Hence, \( p \) is extreme and, by the same argument, \( q \) is extreme. \( \square \)

Now, to show that two polytopes are the same, it suffices to prove that their extreme points are the same. Look at (1). The Newton polytope of \( g \cdot h \) is the convex hull of the points given by multiplying each monomial of \( g \) by each
monomial of $h$. (So, a point may be given by more than one such multiplication.) However, only the extreme points are important and the claim shows that we need only look at extreme points/monomials in either polytope/polynomial. Thus, by this observation, and the fact that exponents add up when terms multiply, we have the result.

3 Number of Common Solutions

In this section, we look at a result that bounds the number of nontrivial solutions of a system of polynomial equations. In particular, we are interested in counting the number of points $x \in (\mathbb{C}^*)^d$, where $\mathbb{C}^* = \mathbb{C} \setminus \{0\}$, that satisfy

$$\begin{cases}
    f_1(x) = 0 \\
    f_2(x) = 0 \\
    \vdots \\
    f_d(x) = 0
    \end{cases}$$

We shall refer to such $x$ as nontrivial solutions to the system $\mathcal{S}$. It turns out that there is a very direct relation between the Newton polynomials of the $f_i$ in $\mathcal{S}$ and the number of $x$ there are. In fact, it is the mixed area of all $\text{New}(f_i)$. This is known as Bernstein’s Theorem and we shall look at the ideas involved in proving it. To simplify matter, however, the discussion will be limited to $d = 2$, but the treatment will suggest a generalization.

**Definition 1.** If $P$ and $Q$ are two polytopes, then their **mixed area** is given by

$$\mathcal{M}(P, Q) = \text{area}(P + Q) - \text{area}(P) - \text{area}(Q)$$

where $P + Q$ denotes the Minkowski sum of $P$ and $Q$.

Here, taking the area of a polytope means taking the volume if the dimension is more than 2. In general this definition takes the form of Definition 2 in the appendix.

Now we can formulate a simplified version of Bernstein’s Theorem

**Theorem 2 (Bernstein’s Theorem).** If $g$ and $h$ are two generic bivariate polynomials, then the number of common zeroes in $(\mathbb{C}^*)^2$ equals the mixed area $\mathcal{M}(\text{New}(g), \text{New}(h))$.

For the full version of the result, see Theorem 7 in the appendix. One consequence of is that if $P$ and $Q$ are polytopes with vertices of integer coordinates, then their mixed area is a positive integer.
3.1 The Proof

One way of proving Theorem 2 gives an explicit algorithm for computing the number of common solutions. It is presented here in three steps. First, the result is shown for the case of two binomials. Then, it is explained how to lift the system in one dimension higher (by adding an indeterminate). This allows to compute an analog to triangulation of polygons each cell of which corresponds to solving for a pair of binomial equations. The result follows by adding the intermediate ones for each cell.

3.1.1 Binomial system

Consider the following system:

\[
\begin{align*}
S & \begin{cases}
  g : \quad x^{a_1}y^{b_1} - c_1 = 0 \\
  h : \quad x^{a_2}y^{b_2} - c_2 = 0
\end{cases}
\end{align*}
\]

(3)

Note, since we allow the \(a_i\) and \(b_j\) to be negative integers as well as positive, the above is a general system of (Laurent) binomial equations. Further, if both \(g\) and \(h\) are multiplied by a Laurent monomial (i.e. of the form \(c'x^ay^b\)), then the number of nontrivial solutions does not change — at most we add a trivial solution. Also, once this only translates the Minkowski sum of \(g\) and \(h\), the mixed area of \(New(g)\) and \(New(h)\) remains the same.

We can therefore define the following action of \(SL_2(\mathbb{Z})\) on the above system: for \(U \in SL_2(\mathbb{Z})\) transform \(S\) by

\[
U.S = \begin{cases}
  \left( x^{a_1}y^{b_1} \right)^{u_{11}} \left( x^{a_2}y^{b_2} \right)^{u_{12}} = c_1^{u_{11}}c_2^{u_{12}} \\
  \left( x^{a_1}y^{b_1} \right)^{v_{21}} \left( x^{a_2}y^{b_2} \right)^{v_{22}} = c_1^{v_{21}}c_2^{v_{22}}
\end{cases}
\]

Now, take \(U\) to be such that

\[
\begin{pmatrix} u_{11} & u_{12} \\ u_{21} & u_{22} \end{pmatrix} \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \begin{pmatrix} r_1 & r_3 \\ r_2 & r_4 \end{pmatrix}
\]

Such an \(U\) is found using the Hermite normal form algorithm of integer linear algebra. Note that \(U^{-1}\) exists and that \(S = U^{-1}S\). In other words,

\[
x^{a_1}y^{b_1} = c_1 \quad \text{and} \quad x^{a_2}y^{b_2} = c_2
\]

\[\iff \begin{align*}
  \left( x^{a_1}y^{b_1} \right)^{u_{11}} \left( x^{a_2}y^{b_2} \right)^{u_{12}} & = c_1^{u_{11}}c_2^{u_{12}} \\
  \left( x^{a_1}y^{b_1} \right)^{v_{21}} \left( x^{a_2}y^{b_2} \right)^{v_{22}} & = c_1^{v_{21}}c_2^{v_{22}}
\end{align*}
\]

\[\iff \begin{align*}
  x^{r_1}y^{r_1} & = c_1^{u_{11}}c_2^{u_{12}} \quad \text{and} \quad y^{r_2} = c_1^{v_{21}}c_2^{v_{22}}
\end{align*}
\]

It is easy to count the number of nontrivial solutions of \(U.S\) — there are \(r_2\) possible \(y \in \mathbb{C}^\ast\) values, and (consequently) \(r_1\) possible \(x \in \mathbb{C}^\ast\). Finally,

\[
r_1r_2 = \det \begin{pmatrix} r_1 & r_3 \\ 0 & r_2 \end{pmatrix} = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}
\]
and \( \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} = \text{area}(\text{New}(g) + \text{New}(h)) = \mathcal{M}(\text{New}(g) + \text{New}(h)) \) since an edge has no area.

### 3.1.2 Toric Deformations

For general systems of polynomials in two variables, we lift the system from \( \mathbb{C}[x, y] \) to one in \( \mathbb{C}[x, y, t] \).

Hence, we have the following transformation

\[
\mathcal{J} \quad \begin{cases} 
    f_1 : \sum_{i=0}^{m} c_i x^{a_i} y^{b_i} = 0 \\
    f_2 : \sum_{i=0}^{m'} c'_i x^{a'_i} y^{b'_i} = 0 
\end{cases} \quad \Rightarrow \quad \mathcal{J}' \quad \begin{cases} 
    \sum_{i=0}^{m} c_i x^{a_i} y^{b_i} t^{\nu_i} = 0 \\
    \sum_{i=0}^{m'} c'_i x^{a'_i} y^{b'_i} t^{\omega_i} = 0 
\end{cases} (4)
\]

Hence, the Newton polytopes from the new system are given by

\[
P = \text{conv} \{(a_i, b_i, \nu_i) : 0 \leq i \leq m\} \\
Q = \text{conv} \{(a'_i, b'_i, \omega_i) : 0 \leq i \leq m'\}
\]

Now, the way \( \nu_i \) and \( \omega_j \) are picked is not arbitrary. (See Figure 2a for an example.) We require them to satisfy certain conditions on lower facets of \( P + Q \). Such a facet \( F \) is one for which there is \((u, v) \in \mathbb{Q}^2\) where \((u, v, 1)\) is an inner normal to \( F \). The conditions are

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 \( F_2 \) a facet of \( Q \), or
   b. \( F_1 \) is an edge of \( P \) and \( F_2 \) an edge of \( Q \), or
   c. \( F_2 \) is a vertex of \( Q \) and \( F_1 \) a facet of \( P \).

If we put \( \pi : \mathbb{C}^3 \rightarrow \mathbb{C}^2 \) to be the canonical projection (simply ignore the last coordinate) it is clear that \( \pi(P) = \text{New}(f_1), \pi(Q) = \text{New}(f_2) \) and that \( \pi(P + Q) = \text{New}(f_1) + \text{New}(f_2) \). Further, \( \pi \) is a bijection if we only take the lower facets (by convexity). Thus, we can think of this as defining a subdivision of \( \text{New}(f_1) + \text{New}(f_2) \). The cells coming from type (b) facets are called mixed cells.

**Proposition 3.** The mixed area of the Newton polytopes corresponding to \( f_1 \) and \( f_2 \) is equal to the sum of the areas of mixed cells.

\[\text{Technically, we may even assume Laurent polynomials. In that case, the lifting is from } \mathbb{C}[x, x^{-1}, y, y^{-1}] \text{ to } \mathbb{C}[x, x^{-1}, y, y^{-1}, t, t^{-1}]. \] However, this is a minor abuse of notation.
Figure 2: (a) Lifting of $\text{New}(g) + \text{New}(h)$ to $P + Q$ in 3D. The parameters are $\nu_1 = \nu_2 = \nu_3 = \nu_4 = \omega_3 = 0$ and $\omega_1 = \omega_2 = 1$. (b) The projection of the lower facets. Note that this gives the Minkowski sum of $\text{New}(g)$ and $\text{New}(h)$. The shaded areas are the mixed cells. Observe that those make up the mixed are as $\{4\}$ corresponds to $\text{New}(g)$ and $\{1\}$ to $\text{New}(h)$. (See Definition 1).
To see this, it suffices to show that the collection of lower facets of type (a) project to an object of area equal to the area of $Q$. That comes from the fact that the Minkowski sum of a vertex (point) with a facet simply yields a translated facet. Further, no two facets intersect at more than an edge (by convexity).

3.1.3 Puiseux series

Using Puiseux series, it can be established that each mixed cell corresponds to a system of two binomial equations. Thus, the problem is reduced to the first part of the discussion.

However, not only is it possible to count the number of distinct nontrivial roots, but the method provides a Puiseux series in $t$ expression for each of the roots in a neighborhood of zero. It is suggested in [7] that numerical continuation methods can be used to trace it out all the way to $t = 1$ which is how roots of the original system are found. That is, approximate the series better to find a solution closer to the actual one.

4 Real Solutions

In the previous section, a method for counting and, to a lesser extent, determining complex roots was outlined. But what about real roots? Very often these are at the heart of the problem and even just a count is very useful. Sadly, it is still largely unknown how to answer those questions for general polynomial systems. However, some results in constrained situations do exist.

Consider the following

**Problem 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 careful reader may wonder why this question even makes sense. Indeed, there does not seem to be any a priori evidence for that! Not to worry, though – not only is this a valid inquiry, but we’ll even be able to give a partial answer.

First, notice that we are trying to say something about the system $f(x, y) = g(x, y) = 0$ that can be written as:

$$f(x, y) = a_1x^{u_1}y^{v_1} + a_2x^{u_2}y^{v_2} + a_3x^{u_3}y^{v_3} + a_4x^{u_4}y^{v_4}$$

$$g(x, y) = b_1x^{\bar{u}_1}y^{\bar{v}_1} + b_2x^{\bar{u}_2}y^{\bar{v}_2} + b_3x^{\bar{u}_3}y^{\bar{v}_3} + b_4x^{\bar{u}_4}y^{\bar{v}_4}$$

where $a_i, b_j$ are arbitrary reals and $u_i, v_j, \bar{u}_i, \bar{v}_j$ are arbitrary integers. Further, if we define

$$f(x, y) = x^6 - 6x^4 + 11x^2 - 6 = (x^2 - 1)(x^2 - 2)(x^2 - 3)$$

$$g(x, y) = y^6 - 6y^4 + 11y^2 - 6 = (y^2 - 1)(y^2 - 2)(y^2 - 3)$$

(5)
we see that the system has exactly 36 common roots. Interestingly, many believe that the answer to Problem 1 is precisely this number: thirty six. Unfortunately, no-one can provide a rigorous reason for the belief. All is not lost, however, we do have a proof of the following

**Theorem 4 (Khovanskii’s Theorem on Fewnomials).** Consider $d$ polynomials in $d$ variables involving $n$ distinct monomials in all. The number of isolated roots in $\mathbb{R}^d$ of any such system is no more than $2\binom{n}{2} \cdot (d + 1)^n$.

**Proof.** For a proof see [1] or [4].

Great, Problem 1 does make sense – we can even get a bound on the number of real roots for (5). For this, note that the theorem gives information for all $2^d$ hyper-quadrants – simply rewrite the system by replacing variable $x_j$ for $-x_j$ where appropriate. Hence, for our problem, we get that the number of real roots is bounded by

$$(2^d \cdot 2\binom{n}{2} \cdot (d + 1)^n)(2, 7) = 18, 345, 885, 696$$

In other words, way more than 36! Tough luck. Well, so much for the “direct” way. Let’s see what we can do with Newton polynomials.

First, let us consider the simple case when the system is, in fact, made up of binomial equations

$$c_1 x^{a_1} y^{b_1} + c_2 x^{a_2} y^{b_2} = 0$$

$$c_3 x^{a_3} y^{b_3} + c_4 x^{a_4} y^{b_4} = 0$$ (6)

**Lemma 5.** The system (6) has precisely one solution in $\mathbb{R}^2_+$ if and only if $c_1 c_2 < 0$ and $c_3 c_4 < 0$. In all other cases it has no solution there.

**Proof.** Notice that is the signs of $c_1$ and $c_2$ are the same, then there is no real solution to the first binomial equation. If there were one, say $(x_0, y_0)$, then (if w.l.o.g. $a_1 \neq 0$)

$$x_0 = \left( -\frac{c_2 x_0^{a_1} y_0^{b_1}}{c_1 y_0} \right)^{\frac{1}{c_1}}$$

which is a contradiction since $\frac{c_2 x_0^{a_1} y_0^{b_1}}{c_1 y_0} > 0$ by the assumption. Similarly, if $c_3 c_4 > 0$, the second binomial has no positive real roots.

On the other hand, if the $c_1 c_2 < 0$ and $c_3 c_4 < 0$, then (6) can be written in the form of system (3) on page 4. That is, we can write

$$x^{a_1-2} y^{b_1-b_2} + \frac{c_2}{c_1} = 0 \Rightarrow x^{a_1-2} y^{b_1-b_2} = -\frac{c_2}{c_1}$$

$$x^{a_3-4} y^{b_3-b_4} + \frac{c_4}{c_3} = 0 \Rightarrow x^{a_3-4} y^{b_3-b_4} = -\frac{c_4}{c_3}$$

Now, refer to Section 3.1.1 and replace $c_1$ for $-\frac{c_2}{c_1}$ and $c_2$ for $-\frac{c_4}{c_3}$. One then sees that system has exactly one real root as $-\frac{c_2}{c_1} > 0$ and $-\frac{c_4}{c_3} > 0$ by assumption. \qed
Again, to see if there is a root in an other quadrant, simply replace \( x \) for \(-x\) and \( y \) for \(-y\) where appropriate. Hence, the lemma shows that there are at most 4 real roots in the case of system (3) — one in each quadrant (or \(2^d\) for \(d\) variables).

Nice, but what about Problem 1? Look at the transformation in (4) on page 5. It turns out that the mixed cells (see page 5) correspond to pairs of binomials \( h_1 \) and \( h_2 \) where \( h_1 \) has two monomials from \( f_1 \) in \( \mathcal{S} \) and \( h_2 \) has two monomials from \( f_2 \) in \( \mathcal{S} \). Hence, we can use Lemma 3 on \( h_1 = h_2 = 0 \). In particular, we shall call a mixed cell an alternating mixed cell whenever the system \( h_1 = h_2 = 0 \) has a root in \( \mathbb{R}_+^2 \). The following result is known

**Theorem 6.** There exists \( \varepsilon > 0 \) such that, for all \( 0 < t < \varepsilon \), the number of zeros in \( \mathbb{R}_+^2 \) of the system (4) for valid choices of \( \nu_i \) and \( \omega_j \) equals the number of alternating mixed cells.

In other words, for sufficiently small \( t > 0 \), the system \( \mathcal{S}_t \) has as many roots in \( \mathbb{R}_+^2 \) as there are alternating mixed cells. The real problem is that “sufficiently small” is yet to be made precise — and this is a result about the toric deformation not about the actual problem.

Still, people have conjectured (see [3]) that the combinatorial bound on the number of possible alternating mixed cells provides a bound on the number of roots in \( \mathbb{R}_+^2 \). Much to their chagrin, Li and Wang [5] gave a counterexample.

Oh well, at least now you known why this is a teasing section. The question remains largely unanswered. However, we do know that it makes sense.

**Appendix**

**Definition 2.** The mixed volumes of the polytopes \( P_1, \ldots, P_d \) given by:

\[
\mathcal{M}(P_1, \ldots, P_d) = \sum_{k=1}^{d} (-1)^{d-k} \sum_{I \subseteq \{1, \ldots, d\}, |I| = k} Vol_d \left( \sum_{i \in I} P_i \right)
\]

where \( Vol_d \) is Euclidean volume in \(d\) dimensions.

**Theorem 7 (Bernstein’s Theorem).** For generic coefficient, the number of common nontrivial zeroes of a system of Laurent polynomials such as \( \mathcal{S} \) in Section 3 equals the mixed volume \(\mathcal{M}(\text{New}(f_1), \ldots, \text{New}(f_d))\).

**Theorem 8 (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. Similarly, the number of negative real roots is bounded above by the number of sign alternations of \(g(-x)\).
Theorem 9 (Bezout’s Theorem). Let \( f(x, y) = g(x, y) = 0 \) be a system of two polynomial equations in two unknowns. It is has finitely many common solutions \((x, y) \in \mathbb{C}^2\), then there are no more than \( \text{deg}(f) \cdot \text{deg}(g) \) such pairs.


References


