Algebraic Geometry

Frank Sottile

Texas A&M University, College Station, Texas 77843, USA

Physical objects and constraints may be modeled by polynomial equations and inequalities. For this reason algebraic geometry, the study of solutions to systems of polynomial equations, is a tool for scientists and engineers. Moreover, relations between concepts arising in science and engineering are often described by polynomials. Whatever their source, once polynomials enter the picture, notions from algebraic geometry—its theoretical base, trove of classical examples, and modern computational tools—may all be brought to bear on the problem at hand.

As a part of applied mathematics, algebraic geometry has two faces. One is an expanding list of recurring techniques and examples which are common to many applications, and the other consists of topics from the applied sciences which involve polynomials. Linking these two aspects are algorithms and software for algebraic geometry.

1 Algebraic Geometry for Applications

We present some concepts and objects common in applications of algebraic geometry.

1.1 Varieties and their ideals

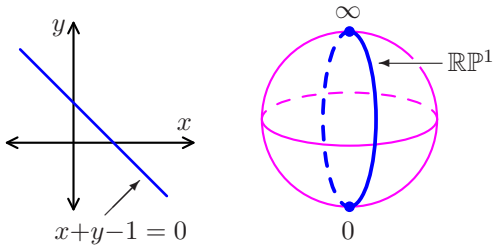
The fundamental object in algebraic geometry is a(n affine) *variety*, which is a set in the vector space $\mathbb{C}^n$ (perhaps restricted to $\mathbb{R}^n$ for an application) defined by polynomials,

$$V(S) := \{x \in \mathbb{C}^n \mid f(x) = 0 \ \forall f \in S\},$$

where $S \subset \mathbb{C}[x] = \mathbb{C}[x_1, \ldots, x_n]$ is a set of polynomials. Common geometric figures—points, lines, planes, circles, conics, spheres, *etc.*—are all algebraic varieties. Thus questions about everyday objects may be treated with algebraic geometry.

This material is based upon work supported by the National Science Foundation under Grant No. 0932078 000, while Sottile was in residence at the Mathematical Science Research Institute (MSRI) in Berkeley, California, during the winter semester of 2013.

On the left below are the real points of the line $x + y - 1 = 0$. Its complex points are the Argand plane $\mathbb{C}$ embedded obliquely in $\mathbb{C}^2$.



We may compactify algebraic varieties by adding points at infinity. This is done in projective space $\mathbb{P}^n$—the set of lines through the origin in $\mathbb{C}^{n+1}$ (or $\mathbb{RP}^n$ for $\mathbb{R}^{n+1}$), which may be thought of as $\mathbb{C}^n$ with a $\mathbb{P}^{n-1}$ at infinity, giving directions of lines in $\mathbb{C}^n$. The projective line $\mathbb{P}^1$ is the Riemann sphere (on the right above).

Points of $\mathbb{P}^n$ are represented by $(n+1)$-tuples of homogeneous coordinates where $[x_0, \ldots, x_n] = [\lambda x_0, \ldots, \lambda x_n]$ if $\lambda \neq 0$ and at least one x_i is nonzero. Projective varieties are subsets of $\mathbb{P}^n$ defined by homogeneous polynomials in $x_0, \ldots, x_n$.

To a subset Z of a vector space we associate the set of polynomials which vanish on Z,

$$I(Z) := \{f \in \mathbb{C}[x] \mid f(z) = 0 \ \forall z \in Z\}.$$

Let $f, g, h \in \mathbb{C}[x]$ with f, g vanishing on Z. Then both $f + g$ and $h \cdot f$ vanish on Z, which implies that $I(Z)$ is an *ideal* of the polynomial ring $\mathbb{C}[x_1, \ldots, x_n]$. Similarly, if I is the ideal generated by a set S of polynomials, then $V(S) = V(I)$.

Both V and I reverse inclusions with $S \subset I(V(S))$ and $Z \subset V(I(Z))$, with equality when Z is a variety. Thus we have the correspondence

$$\{\text{ideals}\} \underset{I}{\overset{V}{\rightleftarrows}} \{\text{varieties}\}$$

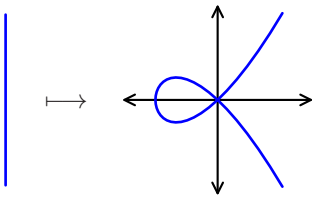
linking algebra and geometry. By Hilbert's Nullstellensatz, this correspondence is bijective when restricted to radical ideals ($f^N \in I \Rightarrow f \in I$). This allows ideas and techniques to flow in both directions and is the source of the power and the depth of algebraic geometry.

The Fundamental Theorem of Algebra asserts that a nonconstant univariate polynomial has a complex root. The Nullstellensatz is a multivariate version, for it is equivalent to the statement that if $I \subsetneq \mathbb{C}[x]$ is a proper ideal, then $V(I) \neq \emptyset$.

It is essentially for this reason that algebraic geometry works best over the complex numbers. Many applications require answers whose coordinates are real numbers, so results from algebraic geometry are often filtered through the lens of the real numbers when used in applications. While this restriction to $\mathbb{R}$ poses significant challenges for algebraic geometers, the generalization from $\mathbb{R}$ to $\mathbb{C}$ and then on to projective space often makes the problems easier to solve. The solution to this useful *algebraic relaxation* is often helpful in treating the original application.

1.2 Parameterization and rationality

Varieties also occur as images of polynomial maps. For example, the map $t \mapsto (t^2-1, t^3-t) = (x, y)$ has image the plane cubic $y^2 = x^3 + x^2$.



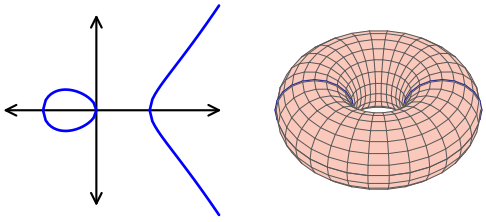
Given such a parametric representation of a variety (or any other explicit description), the *implicitization problem* asks for its ideal.

The converse problem is more subtle: Can a given variety be parameterized? Euclid and Diophantus discovered the rational parameterization of the unit circle $x^2 + y^2 = 1$, $t \mapsto (x, y)$, where

$$x = \frac{2t}{1+t^2} \quad \text{and} \quad y = \frac{1-t^2}{1+t^2}. \tag{1}$$

This is the source of both Pythagorean triples and the rationalizing substitution $z = \tan(\frac{\theta}{2})$ of integral calculus. Homogenizing by setting $t = \frac{a}{b}$, (1) gives an isomorphism between $\mathbb{P}^1$ (with coordinates $[a, b]$) and the unit circle, and then by translation and scaling to any circle.

On the other hand, the cubic $y^2 = x^3 - x$ (at left) has no rational parameterization. This is because the corresponding cubic in $\mathbb{P}^2$ is a curve of genus one (an elliptic curve), which is a torus (see above), and there is no nonconstant map from the Riemann sphere $\mathbb{P}^1$ to the torus. However, $(x, y) \mapsto x$ sends the cubic curve to $\mathbb{P}^1$ and is two-to-one except at the branch points $\{-1, 0, 1, \infty\}$. In fact, any curve with a two-to-one map to $\mathbb{P}^1$ having four branch points has genus one.



A smooth biquadratic curve also has genus one. The product $\mathbb{P}^1 \times \mathbb{P}^1$ is a compactification of $\mathbb{C}^2$ different from $\mathbb{P}^2$. Suppose that $C \subset \mathbb{P}^1 \times \mathbb{P}^1$ is defined by an equation that is separately quadratic in the two variables s and t,

$$a_{00} + a_{10}s + a_{01}t + \cdots + a_{22}s^2t^2 = 0,$$

where s and t are coordinates for the $\mathbb{P}^1$ factors. Analyzing the projection onto the second factor, one can show that the map is two-to-one, except at four branch points, and so C has genus one.

1.3 Toric varieties

Varieties parameterized by monomials (toric varieties) often arise in applications, and may be completely understood in terms of the geometry and combinatorics of the monomials.

Let $\mathbb{C}^*$ be the nonzero complex numbers. An integer vector $\alpha = (a_1, \ldots, a_d) \in \mathbb{Z}^d$ is the exponent vector of a Laurent monomial $t^\alpha := t_1^{a_1} \cdots t_d^{a_d}$, where $t = (t_1, \ldots, t_d) \in (\mathbb{C}^*)^d$ is a d-tuple of nonzero complex numbers. Let $\mathcal{A} = \{\alpha_0, \ldots, \alpha_n\} \subset \mathbb{Z}^d$ be a finite set of integer vectors. Then the *toric variety* $X_\mathcal{A}$ is the closure of the image of the map

$$\varphi_\mathcal{A} \colon (\mathbb{C}^*)^d \ni t \longmapsto [t^{\alpha_0}, t^{\alpha_1}, \ldots, t^{\alpha_n}] \in \mathbb{P}^n.$$

The toric variety $X_\mathcal{A}$ has dimension equal to the dimension of the affine span of $\mathcal{A}$ and it has an action of $(\mathbb{C}^*)^d$ (via the map $\varphi_\mathcal{A}$) with a dense orbit (the image of $\varphi_\mathcal{A}$).

The implicitization problem for toric varieties is elegantly solved. Assume that $\mathcal{A}$ lies on an affine hyperplane, so that there is a vector $w \in \mathbb{R}^d$ with $w \cdot \alpha_i = w \cdot \alpha_j (\neq 0)$ for all i, j, where $\cdot$ is the dot product. For $v \in \mathbb{R}^{n+1}$, write $\mathcal{A}v$ for $\sum_i \alpha_i v_i$.

Theorem 1. *The homogeneous ideal of $X_\mathcal{A}$ is spanned by binomials $x^u - x^v$ where $\mathcal{A}u = \mathcal{A}v$.*

The assumption that we have w with $w \cdot \alpha_i = w \cdot \alpha_j$ for all i, j may be arranged by appending a new $(d{+}1)$st coordinate of 1 to each α_i and setting $w = (0, \dots, 0, 1) \in \mathbb{R}^{d+1}$. This does not change the projective variety $X_{\mathcal{A}}$.

Applications also use the tight relation between $X_{\mathcal{A}}$ and the convex hull $\Delta_{\mathcal{A}}$ of $\mathcal{A}$, which is a polytope with integer vertices. The points of $X_{\mathcal{A}}$ with nonnegative coordinates form its *nonnegative part* $X_{\mathcal{A}}^+$. This is identified with $\Delta_{\mathcal{A}}$ through the *algebraic moment map*, $\pi_{\mathcal{A}} \colon \mathbb{P}^n \dashrightarrow \mathbb{P}^d$ which sends a point x to $\mathcal{A}x$. (The broken arrow $\dashrightarrow$ means that the map is not defined everywhere.) By Birch's Theorem from statistics, $\pi_{\mathcal{A}}$ maps $X_{\mathcal{A}}^+$ homeomorphically to $\Delta_{\mathcal{A}}$.

There is a second homeomorphism $\beta_{\mathcal{A}} \colon \Delta_{\mathcal{A}} \xrightarrow{\sim} X_{\mathcal{A}}^+$ given by polynomials. The polytope $\Delta_{\mathcal{A}}$ is defined by linear inequalities,

$$\Delta_{\mathcal{A}} := \{x \in \mathbb{R}^d \mid \ell_F(x) \geq 0\},$$

where F ranges over the codimension one faces of $\Delta_{\mathcal{A}}$ and $\ell_F(F) \equiv 0$ with the coefficients of ℓ_F coprime integers. For each $\alpha \in \mathcal{A}$, set

$$\beta_\alpha(x) := \prod_F \ell_F(x)^{\ell_F(\alpha)}, \tag{2}$$

which is nonnegative on $\Delta_{\mathcal{A}}$. For $x \in \Delta_{\mathcal{A}}$, set

$$\beta_{\mathcal{A}}(x) := [\beta_{\alpha_0}(x), \dots, \beta_{\alpha_n}(x)] \in X_{\mathcal{A}}^+.$$

While $\pi_{\mathcal{A}}$ and $\beta_{\mathcal{A}}$ are homeomorphisms between the same spaces, they are typically not inverses.

A useful variant is to translate $X_{\mathcal{A}}$ by a nonzero weight, $\omega = (\omega_0, \dots, \omega_n) \in (\mathbb{C}^*)^{n+1}$,

$$X_{\mathcal{A},\omega} := \{[\omega_0 x_0, \dots, \omega_n x_n] \mid x \in X_{\mathcal{A}}\}.$$

This translated toric variety is spanned by binomials $\omega^v x^u - \omega^u x^v$ with $\mathcal{A}u = \mathcal{A}v$ as in Theorem 1, and it is parameterized by monomials via

$$\varphi_{\mathcal{A},\omega}(t) = (\omega_0 t^{\alpha_0}, \dots, \omega_n t^{\alpha_n}).$$

When the weights ω_i are positive real numbers, Birch's Theorem holds, $\pi_{\mathcal{A}} \colon X_{\mathcal{A},\omega}^+ \xrightarrow{\sim} \Delta_{\mathcal{A}}$, and we have the parameterization $\beta_{\mathcal{A},\omega} \colon \Delta_{\mathcal{A}} \to X_{\mathcal{A},\omega}^+$ where the components of $\beta_{\mathcal{A},\omega}$ are $\omega_i \beta_{\alpha_i}$.

Example 1. When $\mathcal{A}$ consists of the standard unit vectors $(1, 0, \dots, 0), \dots, (0, \dots, 0, 1)$ in $\mathbb{R}^{n+1}$, the toric variety is projective space $\mathbb{P}^n$ and $\varphi_{\mathcal{A}}$ gives the usual homogeneous coordinates $[x_0, \dots, x_n]$ for $\mathbb{P}^n$. The nonnegative part of $\mathbb{P}^n$ is the convex hull of $\mathcal{A}$, which is the *standard n-simplex*, $\triangle^n$, and $\pi_{\mathcal{A}} = \beta_{\mathcal{A}}$ is the identity map.

Example 2. Let $\mathcal{A} = \{0, 1, \dots, n\}$ so that $\Delta_{\mathcal{A}} = [0, n]$, and choose weights $\omega_i = \binom{n}{i}$. Then $X_{\mathcal{A},\omega}$ is the closure of the image of the map

$$t \mapsto [1, nt, \tbinom{n}{2} t^2, \dots, nt^{n-1}, t^n] \in \mathbb{P}^n,$$

which is the (translated) moment curve. Its nonnegative part $X_{\mathcal{A},\omega}^+$ is the image of $[0, n]$ under the map $\beta_{\mathcal{A},\omega}$ whose components are

$$\beta_i(x) = \tfrac{1}{n^n} \tbinom{n}{i} x^i (n - x)^{n-i}.$$

Replacing x by ny gives the Bernstein polynomial

$$\beta_{i,n}(y) = \tbinom{n}{i} y^i (1 - y)^{n-i}, \tag{3}$$

and thus the moment curve is parameterized by the Bernstein polynomials. Because of this, we call the functions $\omega_i \beta_{\alpha_i}$ (2) *generalized Bernstein polynomials*.

The composition $\pi_{\mathcal{A}} \circ \beta_{\mathcal{A},\omega}(x)$ is

$$\frac{1}{n^n} \sum_{i=0}^{n} i \binom{n}{i} x^i (n - x)^{n-i} = \frac{nx}{n^n} \sum_{i=1}^{n} \binom{n-1}{i-1} x^{i-1} (n - x)^{n-i} = x,$$

as the last sum is $(x{+}(n{-}x))^{n-1}$. Similarly, $\frac{1}{n} \pi_{\mathcal{A}} \circ \beta(y) = y$, where β is the parameterization by the Bernstein polynomials. The weights $\omega_i = \binom{n}{i}$ are essentially the unique weights for which $\pi_{\mathcal{A}} \circ \beta_{\mathcal{A},\omega}(x) = x$.

Example 3. For positive integers m, n consider the map $\varphi \colon \mathbb{C}^m \times \mathbb{C}^n \to \mathbb{P}(\mathbb{C}^{m \times n})$ defined by

$$(x, y) \mapsto [x_i y_j \mid i = 1, \dots, m,\ j = 1, \dots, n].$$

Its image is the *Segre variety*, which is a toric variety, as the map φ is $\varphi_{\mathcal{A}}$ where $\mathcal{A}$ is

$$\{\mathbf{e}_i + \mathbf{f}_j \mid i = 1, \dots, m,\ j = 1, \dots, n\} \subset \mathbb{Z}^m \oplus \mathbb{Z}^n.$$

Here, $\{\mathbf{e}_i\}$ and $\{\mathbf{f}_j\}$ are the standard bases for $\mathbb{Z}^m$ and $\mathbb{Z}^n$, respectively.

If z_{ij} are the coordinates of $\mathbb{C}^{m \times n}$, then the Segre variety is defined by the binomial equations

$$z_{ij} z_{kl} - z_{il} z_{kj} = \begin{vmatrix} z_{ij} & z_{il} \\ z_{kj} & z_{kl} \end{vmatrix}.$$

Identifying $\mathbb{C}^{m\times n}$ with $m\times n$ matrices shows that the Segre variety is the set of rank one matrices.

Other common toric varieties include the *Veronese variety*, where $\mathcal{A}$ is $\mathcal{A}_{n,d} := n\Delta^d \cap \mathbb{Z}^{d+1}$, and the *Segre-Veronese variety*, where $\mathcal{A}$ is $\mathcal{A}_{m,d} \times \mathcal{A}_{n,e}$. When $d = e = 1$, $\mathcal{A}$ consists of the integer vectors in the $m \times n$ rectangle

$$\mathcal{A} = \{(i, j) \mid 0 \le i \le m,\ 0 \le j \le n\}.$$

2 Algorithms for Algebraic Geometry

Mediating between theory and examples and facilitating applications are algorithms developed to study, manipulate, and compute algebraic varieties. These come in two types, exact symbolic methods and approximate numerical methods.

2.1 Symbolic algorithms

The words algebra and algorithm share an Arabic root, but their relation is more than just history. When we write a polynomial, say as a sum of monomials or as an expression such as a determinant of polynomials, that symbolic representation is an algorithm for evaluating the polynomial.

Expressions for polynomials lend themselves to algorithmic manipulation. While these representations and manipulations have their origin in antiquity, and methods such as Gröbner bases predate the Computer Age, the rise of computers has elevated symbolic computation to a key tool for algebraic geometry and its applications.

Euclid's algorithm, Gaussian elimination, and Sylvester's resultants are important symbolic algorithms which are supplemented by universal symbolic algorithms based on Gröbner bases. They begin with a *term order* $\prec$, which is a well-ordering of all monomials that is consistent with multiplication. For example, $\prec$ could be the lexicographic order in which $x^u \prec x^v$ if the first nonzero entry of the vector $v - u$ is positive. A term order organizes the algorithmic representation and manipulation of polynomials and it is the basis for the termination of algorithms.

The initial term $\mathrm{in}_\prec f$ of a polynomial f is its term $c_\alpha x^\alpha$ with the $\prec$-largest monomial in f. The initial ideal $\mathrm{in}_\prec I$ of an ideal I is the ideal generated by initial terms of polynomials in I. This monomial ideal is a well-understood combinatorial object and the passage to an initial ideal preserves much information about I and its variety.

A *Gröbner basis* for I is a finite set $G \subset I$ of polynomials whose initial terms generate $\mathrm{in}_\prec I$. This set G generates I and facilitates the transfer of information from $\mathrm{in}_\prec I$ back to I. This information may typically be extracted using linear algebra, so a Gröbner basis essentially contains all the information about I and its variety.

Consequently, a bottleneck in this approach to symbolic computation is the computation of a Gröbner basis (which has high complexity due to its information content). Gröbner basis calculation also appears to be essentially serial—no efficient parallel algorithm is known.

The subject began in 1965 when Buchberger gave an algorithm to compute a Gröbner basis. Decades of development, including sophisticated heuristics and completely new algorithms, have led to reasonably efficient implementations of Gröbner basis computation. Many algorithms have been devised and implemented to use a Gröbner basis to study a variety. All of this is embedded in freely available software packages which are revolutionizing the practice of algebraic geometry and its applications.

2.2 Numerical algebraic geometry

While symbolic algorithms lie on the algebraic side of algebraic geometry, numerical algorithms, which compute and manipulate points on varieties, have a strongly geometric flavor.

These numerical algorithms rest upon Newton's method for refining an approximate solution to a system of polynomial equations. A system $F = (f_1, \ldots, f_n)$ of polynomials in n variables is a map $F\colon \mathbb{C}^n \to \mathbb{C}^n$ with solutions $F^{-1}(0)$. We focus on systems with finitely many solutions. A *Newton iteration* is the map $N_F\colon \mathbb{C}^n \to \mathbb{C}^n$ where

$$N_F(x) = x - DF_x^{-1}(F(x)),$$

with DF_x the Jacobian matrix of partial derivatives of F at x. If $\xi \in F^{-1}(0)$ is a solution to F with DF_ξ invertible, then when x is sufficiently close to ξ, $N_F(x)$ is closer still in that it has twice as many digits in common with ξ as does x. Smale showed that sufficiently close may be decided al-

gorithmically, which can allow the certification of output from numerical algorithms.

Newton iterations are used in numerical continuation. For a polynomial system H_t depending on a parameter t, the solutions $H_t^{-1}(0)$ for $t \in [0,1]$ form a collection of arcs. Given a point (x_t, t) of some arc and a step δ_t, a predictor is called to give a point $(x', t+\delta_t)$ that is near to the same arc. Then Newton iterations are used to refine this to a point $(x_{t+\delta_t}, t+\delta_t)$ on the arc. This numerical continuation algorithm can be used to trace arcs from $t=0$ to $t=1$.

We may use continuation to find all solutions to a system F consisting of polynomials f_i of degree d. Define a new system $H_t = (h_1, \ldots, h_n)$ by

$$h_i := tf_i + (1-t)(x_i^d - 1).$$

At $t=0$, this is $x_i^d - 1$ whose solutions are the dth roots of unity. When F is general, $H_t^{-1}(0)$ consists of d^n arcs connecting these known solutions at $t=0$ to the solutions of $F^{-1}(0)$ at $t=1$. These may be found by continuation.

While this *Bézout homotopy* illustrates the basic idea, it has exponential complexity and may not be efficient. In practice, other more elegant and efficient homotopy algorithms are used for numerically solving systems of polynomials.

These numerical methods underlie *numerical algebraic geometry* which uses them to manipulate and study algebraic varieties on a computer. The subject began when Sommese, Verschelde and Wampler introduced its fundamental data structure of a witness set, as well as algorithms to generate and manipulate witness sets.

Suppose we have a variety $V \subset \mathbb{C}^n$ of dimension $n-d$ that is a component of the zero set $F^{-1}(0)$ of d polynomials $F = (f_1, \ldots, f_d)$. A *witness set* for V consists of a general affine subspace $L \subset \mathbb{C}^n$ of dimension d (given by d affine equations) and (approximations to) the points of $V \cap L$. The points of $V \cap L$ may be numerically continued as L moves to sample points from V.

An advantage of numerical algebraic geometry is that path-tracking is inherently parallelizable, as each of the arcs in $H_t^{-1}(0)$ may be tracked independently. This parallelism is one reason why numerical algebraic geometry does not face the complexity affecting symbolic methods. Another reason is that by computing approximate solutions to equations, the full information of a variety is never computed.

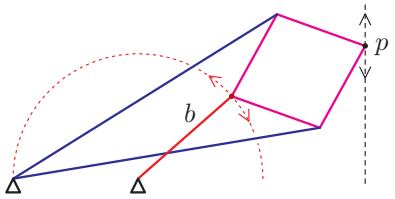
3 Algebraic Geometry in Applications

We illustrate some of the many ways that algebraic geometry arises in applications.

3.1 Kinematics

Kinematics is concerned with motions of *linkages* (rigid bodies connected by movable joints). While its origins were in the simple machines of antiquity, its importance grew with the age of steam and today it is fundamental to robotics. As the positions of a linkage are solutions to a system of polynomial equations, kinematics has long been an area of application of algebraic geometry.

An early challenge important to the development of the steam engine was to find a linkage with a motion along a straight line. Watt discovered a linkage in 1784 approximating straight line motion (tracing a curve near a flex) and in 1864 Peaucellier gave the first linkage with a true straight line motion (based on circle inversion).

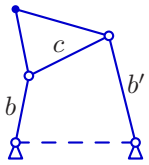


(When the bar b is rotated about its anchor point, the point p traces a straight line.)

Cayley, Chebyshev, Darboux, Roberts, and others contributed to kinematics in the 19th century. The French Academy of Sciences recognized the importance of kinematics, posing the problem of determining the nontrivial mechanisms with a motion constrained to a sphere for its 1904 Prix Vaillant, which was awarded to Borel and Bricard for their partial solutions.

The *four-bar linkage* consists of four bars in the plane connected by rotational joints with one bar fixed. A triangle is erected on the *coupler bar* opposite the fixed bar and we wish to describe the

coupler curve traced by the apex of the triangle.

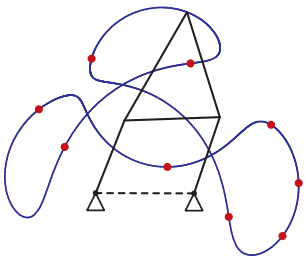


To understand the motion of this linkage, note that if we remove the coupler bar c, the bars b and b' swing freely, tracing two circles, each of which we parameterize by $\mathbb{P}^1$ as in (1). The coupler bar constrains the endpoints of the bars b, b' to lie at a fixed distance. In the parameters s, t of the circles and if b, b', c are lengths of the corresponding bars, the coupler constraint gives the equation

$$\begin{aligned} c^2 &= \left(b\tfrac{1-s^2}{1+s^2} - b'\tfrac{1-t^2}{1+t^2}\right)^2 + \left(b\tfrac{2s}{1+s^2} - b'\tfrac{2t}{1+t^2}\right)^2 \\ &= b^2 + b'^2 - 2bb'\frac{(1-s^2)(1-t^2)+4st}{(1+s^2)(1+t^2)}. \end{aligned}$$

Clearing denominators gives a biquadratic equation in the variety $\mathbb{P}^1 \times \mathbb{P}^1$ that parameterizes the rotations of the bars b, b'. Thus the coupler curve is a genus one curve and is irrational. The real points of a genus one curve have either one or two components, which corresponds to the linkage having one or two assembly modes—to reach all points of a coupler curve with two components requires disassembly of the mechanism.

Roberts and Chebyshev discovered that there are three linkages (Roberts cognates) with the same coupler curve, and they may be constructed from one another using straightedge and compass. The nine-point path synthesis problem asks for the four-bar linkages whose coupler curve contains nine given points. Morgan, Sommese and Wampler used numerical continuation to solve the equations, finding 4326 distinct linkages in 1442 triplets of Roberts cognates. Here is one linkage solving this problem for the indicated nine points.



Such applications in kinematics drove the early development of numerical algebraic geometry.

3.2 Geometric modeling

Geometric modeling uses curves and surfaces to represent objects on a computer for industrial design, manufacture, architecture, and entertainment. These applications of computer-aided geometric design and computer graphics are profoundly important to the world economy.

Geometric modeling began around 1960 in the work of de Casteljau at Citroën, who introduced what are now called Bézier curves (they were popularized by Bézier at Renault) for use in automobile manufacturing.

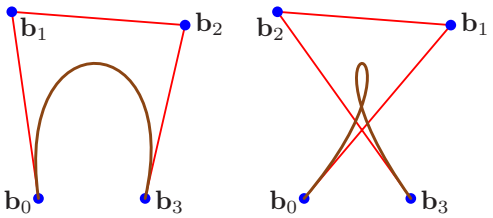
Bézier curves (and higher-dimensional analogs, rectangular tensor-product and triangular Bézier patches) are parametric curves (and surfaces) that have become widely used for many reasons, including ease of computation and the intuitive method to control shape by manipulating control points. They begin with Bernstein polynomials (3), which are nonnegative on $[0, 1]$. Expanding $1^n = (t + (1-t))^n$ shows that

$$1 = \sum_{i=1}^{n} \beta_{i,n}(t).$$

Given control points $\mathbf{b}_0, \ldots, \mathbf{b}_n$ in $\mathbb{R}^2$ (or $\mathbb{R}^3$), we have the Bézier curve

$$[0,1] \ni t \longmapsto \sum_{i=0}^{n} \mathbf{b}_i \beta_{i,n}(t). \tag{4}$$

Here are two cubic ($n = 3$) Bézier curves in $\mathbb{R}^2$.



By (4), a Bézier curve is the image of the nonnegative part of the translated moment curve of Example 2 under the map defined on projective space by

$$[x_0, \ldots, x_n] \mapsto \sum_{i=0}^{n} x_i \mathbf{b}_i.$$

On the standard simplex $\triangle^n$, this is the canonical map to the convex hull of the control points.

The tensor product patch of bidegree (m, n) has basis functions

$$\beta_{i,m}(s)\beta_{j,m}(t),$$

for $i = 0, \ldots, m$ and $j = 0, \ldots, n$. These are functions on the unit square. Control points

$$\{\mathbf{b}_{i,j} \mid i = 0, \ldots, m,\ j = 0, \ldots, n\} \subset \mathbb{R}^3$$

determine the map

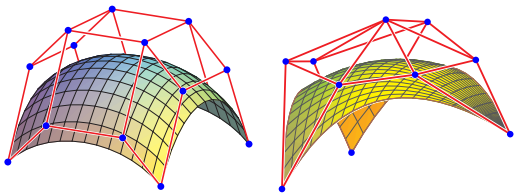
$$(s, t) \longmapsto \sum \mathbf{b}_{ij}\beta_{i,m}(s)\beta_{j,n}(t),$$

whose image is a rectangular patch.

Bézier triangular patches of degree d have basis functions

$$\beta_{i,j;d}(s, t) = \frac{d!}{i!j!(d-i-j)!}s^i t^j (1-s-t)^{d-i-j},$$

for $0 \leq i, j$ with $i+j \leq d$. Again, control points give a map from the triangle with image a Bézier triangular patch. Here are two surface patches.



These patches correspond to toric varieties, with tensor product patches coming from Segre-Veronese surfaces and Bézier triangles from Veronese surfaces. The basis functions are the generalized Bernstein polynomials $\omega_i\beta_{\alpha_i}$ of Subsection 1.3, and this explains their shape as they are images of $\Delta_{\mathcal{A}}$, which is a rectangle for the Segre-Veronese surfaces and a triangle for the Veronese surfaces.

An important question is to determine the intersection of two patches, given parametrically as $F(x)$ and $G(x)$ for x in some domain (a triangle or rectangle). This is used for trimming the patches or drawing the intersection curve. A common approach is to solve the implicitization problem for G, giving a polynomial g which vanishes on the patch G. Then $g(F(x))$ defines the intersection in the domain of F. This application has led to theoretical and practical advances in algebra concerning resultants and syzygies.

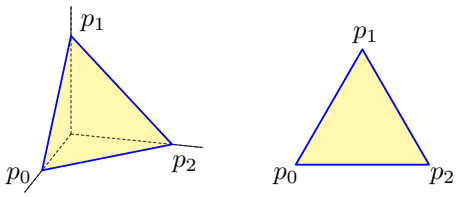
3.3 Algebraic Statistics

Algebraic statistics applies tools from algebraic geometry to questions of statistical inference. This is possible because many statistical models are (part of) algebraic varieties, or they have significant algebraic or geometric structures.

Suppose that X is a discrete random variable with $n+1$ possible states, $0, \ldots, n$. (E.g., the number of tails observed in n coin flips.) If p_i is the probability that X takes value i,

$$p_i := P(X = i),$$

then $p_0, \ldots, p_n$ are nonnegative and sum to 1. Thus p lies in the standard n-simplex, $\triangle^n$. Here are two views of it when $n = 2$.

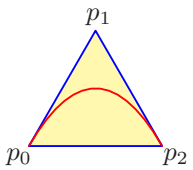


A *statistical model* M is a subset of $\triangle^n$. If $(p_0, \ldots, p_n) \in M$, then we may think of X as being explained by M.

Example 4. Let X be a discrete random variable whose states are the number of tails in n flips of a coin with a probability t of landing on tails and $1-t$ of heads. We may calculate that

$$P(X = i) = \tbinom{n}{i}t^i(1-t)^{n-i},$$

the Bernstein polynomial $\beta_{i,n}$ (3) evaluated at the parameter t. We call X a *binomial random variable* or *binomial distribution*. The set of binomial distributions as t varies gives the translated moment curve of Example 2 parameterized by Bernstein polynomials. This curve is the model for binomal distributions. Here is a picture of this curve when $n = 2$.



Example 5. Suppose that we have discrete random variables X and Y with m and n states, respectively. Their joint distribution has mn states

(cells in a table or a matrix) and lies in the simplex Δ^{mn-1}. The *model of independence* consists of all distributions $p \in \Delta^{mn-1}$ such that

$$P(X = i, Y = j) = P(X = i)P(Y = j). \quad (5)$$

It is parameterized by $\Delta^{m-1} \times \Delta^{n-1}$ (probability simplices for X and Y), and (5) shows it is the nonnegative part of the Segre variety of Example 3. Thus the model of independence consists of those joint probability distributions which are rank one matrices.

Other common statistical models, called discrete exponential families or *toric models*, are also the nonnegative part $X^+_{\mathcal{A},\omega}$ of some toric variety. For these, the algebraic moment map $\pi_{\mathcal{A}} \colon \Delta^n \to \Delta_{\mathcal{A}}$ (or $u \mapsto \mathcal{A}u$) is a *sufficient statistic*. For the model of independence, $\mathcal{A}u$ is the vector of row and column sums of the table u.

Suppose that we have data from N independent observations (or draws), each from the same distribution $p(t)$ from a model M, and we wish to estimate the parameter t best explaining the data. One method is to maximize the *likelihood* (the probability of observing the data given a parameter t). Suppose that the data are represented by a vector u of counts where u_i is how often state i was observed in the N trials. The likelihood function is

$$L(t|u) = \binom{N}{u} \prod_{i=0}^{n} p_i(t)^{u_i},$$

where $\binom{N}{u}$ is the multinomial coefficient.

Suppose that M is the binomial distribution of Example 4. It suffices to maximize the logarithm of $L(t|u)$, which is

$$C + \sum_{i=0}^{n} u_i(i \log t + (n-i) \log(1-t)),$$

where C is a constant. By calculus, we have

$$0 = \frac{1}{t} \sum_{i=0}^{n} i u_i + \frac{1}{1-t} \sum_{i=0}^{n} (n-i) u_i.$$

Solving, we obtain that

$$t := \frac{1}{n} \sum_{i=0}^{n} i \frac{u_i}{N} \quad (6)$$

maximizes the likelihood. If $\widehat{u} := \frac{u}{N} \in \Delta^n$ is the point corresponding to our data, then (6) is the normalized algebraic moment map $\frac{1}{n}\pi_{\mathcal{A}}$ of Example 2 applied to $\widehat{u}$. For a general toric model $X^+_{\mathcal{A},\omega} \subset \Delta^n$, likelihood is maximized at the parameter t satisfying $\pi_{\mathcal{A}} \circ \beta_{\mathcal{A},\omega}(t) = \pi_{\mathcal{A}}(\widehat{u})$. An algebraic formula exists for the parameter maximizing likelihood exactly when $\pi_{\mathcal{A}}$ and $\beta_{\mathcal{A},\omega}$ are inverses.

Suppose that we have data u as a vector of counts as before and a model $M \subset \Delta^n$, and we wish to test the null hypothesis that the data u come from a distribution in M. *Fisher's exact test* uses a score function $\Delta^n \to \mathbb{R}_{\geq}$ which is zero exactly on M, and computes how likely it is for data v to have a higher score than u, when v is generated from the same probability distribution as u. This requires that we sample from the probability distribution of such v.

For a toric model $X^+_{\mathcal{A},\omega}$, this is a probability distribution on the set of possible data with the same sufficient statistics,

$$\mathcal{F}_u := \{v \mid \mathcal{A}u = \mathcal{A}v\}.$$

For a parameter t, this distribution is

$$\begin{aligned} L(v \mid v \in \mathcal{F}_u, t) &= \frac{\binom{N}{v}\omega^v t^{\mathcal{A}v}}{\sum_{w \in \mathcal{F}_u} \binom{N}{w}\omega^w t^{\mathcal{A}w}} \\ &= \frac{\binom{N}{v}\omega^v}{\sum_{w \in \mathcal{F}_u} \binom{N}{w}\omega^w}, \end{aligned} \quad (7)$$

as $\mathcal{A}v = \mathcal{A}w$ for $v, w \in \mathcal{F}(u)$.

This sampling may be accomplished using a random walk on the fiber $\mathcal{F}_u$ with stationary distribution (7). This requires a connected graph on $\mathcal{F}_u$. Remarkably, any Gröbner basis for the ideal of the toric variety $X_{\mathcal{A}}$ gives such a graph.

3.4 Tensor rank

The fundamental invariant of a $m \times n$ matrix is its rank. The set of all matrices of rank at most r is defined by the vanishing of the determinants of all $(r+1) \times (r+1)$ submatrices. From this perspective, the simplest matrices are those of rank one, and the rank of a matrix A is the minimal number of rank one matrices that sum to A.

A $m \times n$ matrix A is a linear map $V_2 = \mathbb{C}^n \to V_1 = \mathbb{C}^m$. If it has rank one, it is the composition

$$V_2 \xrightarrow{v_2} \mathbb{C} \xhookrightarrow{v_1} V_1,$$

of a linear function v_2 on V_2 ($v_2 \in V_2^*$) and an inclusion given by $1 \mapsto v_1 \in V_1$. Thus $A = v_1 \otimes v_2 \in V_1 \otimes V_2^*$, as this tensor space is naturally the space of linear maps $V_2 \to V_1$. A tensor of the form $v_1 \otimes v_2$ has rank one, and the set of rank one tensors forms the Segre variety of Example 3.

Singular value decomposition writes a matrix A as a sum of rank one matrices of the form

$$A = \sum_{i=1}^{\operatorname{rank}(A)} \sigma_i v_{1,i} \otimes v_{2,i}, \tag{8}$$

where $\{v_{1,i}\}$ and $\{v_{2,i}\}$ are orthonormal and $\sigma_1 \geq \cdots \geq \sigma_{\operatorname{rank}(A)} > 0$ are the singular values of A. Often, only the relatively few terms of (8) with largest singular values are significant, and the rest are noise. Letting A_{lr} be the sum of terms with large singular values and A_{noise} be the sum of the rest, then A is the sum of the low-rank matrix A_{lr} plus noise.

A *k-way tensor* (k-way table) is an element of $V_1 \otimes \cdots \otimes V_k$, where each V_i is a finite dimensional vector space. A rank one tensor has the form $v_1 \otimes \cdots \otimes v_k$, where $v_i \in V_i$. These form a toric variety, and the rank of a tensor v is the minimal number of rank one tensors that sum to v.

The (closure of) the set of rank r tensors is the *rth secant variety*. When $k = 2$ (matrices), the set of determinants of all $(r+1) \times (r+1)$ submatrices solves the implicitization problem for the rth secant variety. For $k > 2$ there is not yet a solution to the implicitization problem for the rth secant variety.

Tensors are more complicated than matrices. Some tensors of rank $> r$ lie in the rth secant variety, and these may be approximated by low rank (rank r) tensors. Algorithms for tensor decomposition generalize singular value decomposition. Often their goal is an expression of the form $v = v_{\text{lr}} + v_{\text{noise}}$ for a tensor v as the sum of a low rank tensor v_{lr} plus noise v_{noise}.

Some mixture models in algebraic statistics are secant varieties. Consider an inhomogeneous population in which the fraction θ_i obeys a probability distribution $p^{(i)}$ from a model M. Then the distribution of data collected from this population behaves as the convex combination

$$\theta_1 p^{(1)} + \theta_2 p^{(2)} + \cdots + \theta_r p^{(r)},$$

which is a point on the rth secant variety of M.

Theoretical and practical problems in complexity may be reduced to knowing the rank of specific tensors. Matrix multiplication gives a nice example of this.

Let $A = (a_{ij})$ and $B = (b_{ij})$ be 2×2 matrices. In the usual multiplication algorithm, $C = AB$ is

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j}, \qquad i, j = 1, 2. \tag{9}$$

This involves eight multiplications. For $n \times n$ matrices, the algorithm uses n^3 multiplications.

Strassen discovered an alternative. Set

$$\begin{aligned}
s_1 &:= (a_{11} + a_{22})(b_{11} + b_{22}) \\
s_2 &:= (a_{21} + a_{22})b_{11} \\
s_3 &:= (a_{11} + a_{12})b_{22} \\
s_4 &:= a_{11}(b_{12} - b_{22}) \\
s_5 &:= a_{22}(b_{21} - b_{11}) \\
s_6 &:= (a_{21} - a_{11})(b_{11} + b_{12}) \\
s_7 &:= (a_{12} - a_{22})(b_{21} + b_{22})
\end{aligned}$$

Then $C = AB$ is

$$\begin{pmatrix} s_1 - s_3 + s_5 + s_7 & s_2 + s_5 \\ s_3 + s_4 & s_1 - s_2 + s_4 + s_6 \end{pmatrix},$$

which only involves seven multiplications.

If A and B are $2k \times 2k$ matrices with $k \times k$ blocks a_{ij} and b_{ij}, then these formulas apply and enable the computation of AB using $7k^3$ multiplications. Recursive application of this idea enables the multiplication of $n \times n$ matrices using only $n^{\log_2 7} \simeq n^{2.81}$ multiplications. This method is used in practice to multiply large matrices.

We interpret Strassen's algorithm in terms of tensor rank. The formula (9) for $C = AB$ is a tensor $\mu \in V \otimes V^* \otimes V^*$, where $V = M_{2\times 2}(\mathbb{C})$. Each multiplication is a rank one tensor, and (9) exhibits μ as a sum of eight rank one tensors, so μ has rank at most eight. Strassen's algorithm shows that μ has rank at most seven. We now know that the rank of any tensor in $V \otimes V^* \otimes V^*$ is at most seven, which shows how Strassen's algorithm could have been anticipated.

The fundamental open question about the complexity of multiplying $n \times n$ matrices is to determine the rank r_n of the multiplication tensor. We only have bounds for r_n. We know that $r_n \geq o(n^2)$, as matrices have n^2 entries and improvements to the idea behind Strassen's algorithm show that $r_n < O(n^{2.38})$.

3.5 Hardy-Weinberg equilibrium

We close with a simple application to Mendelian genetics.

Suppose that a gene exists in a population in two variants (alleles), a and b. Individuals will have one of three genotypes, aa, ab or bb, and their distribution $p = (p_{aa}, p_{ab}, p_{bb})$ is a point in the probability 2-simplex. A fundamental and originally controversial question in the early days of Mendelian genetics was: Which distributions are possible in a population at equilibrium? (Assuming no evolutionary pressures, equidistribution of the alleles among the sexes, etc.)

The proportions q_a and q_b of alleles a and b in the population are

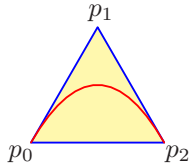
$$q_a = p_{aa} + \tfrac{1}{2}p_{ab}, \qquad q_b = \tfrac{1}{2}p_{ab} + p_{bb}, \qquad (10)$$

and the assumption of equilibrium is

$$p_{aa} = q_a^2, \quad p_{ab} = 2q_a q_b, \quad p_{bb} = q_b^2. \qquad (11)$$

If $\mathcal{A} = \left(\begin{smallmatrix} 2 & 1 & 0 \\ 0 & 1 & 2 \end{smallmatrix}\right)$ with $\left(\begin{smallmatrix} 2 \\ 0 \end{smallmatrix}\right) \leftrightarrow aa$, $\left(\begin{smallmatrix} 1 \\ 1 \end{smallmatrix}\right) \leftrightarrow ab$, and $\left(\begin{smallmatrix} 0 \\ 2 \end{smallmatrix}\right) \leftrightarrow bb$, then (10) is $(q_a, q_b) = \frac{1}{2}\pi_{\mathcal{A}}(p_{aa}, p_{ab}, p_{bb})$, the normalized algebraic moment map of Examples 2 and 4 applied to (p_{aa}, p_{ab}, p_{bb}). Similarly, the assignment $q \to p$ of (11) is the parametrization β of the translated quadratic moment curve of Example 2 given by the Bernstein polynomials.

Since $\frac{1}{2}\pi_{\mathcal{A}} \circ \beta(q) = q$, the population is at equilibrium if and only if the distribution (p_{aa}, p_{ab}, p_{bb}) of alleles lies on the translated quadratic moment curve, that is, if and only if it is a point in the binomial distribution, which we reproduce.



This is called the *Hardy-Weinberg equilibrium* after its two independent discoverers.

Hardy here is the great English mathematician G.H. Hardy, who was known for his disdain for applied mathematics, and this contribution came early in his career, in 1908. He was later famous for his work in number theory, a subject that he extolled for its purity and uselessness. As we all now know, Hardy was mistaken on this last point for number theory underlies our modern digital world, from the security of financial transactions via cryptography to using error correcting codes to ensure the integrity of digitally transmitted documents, such as the one you have now finished reading.

Further Reading

1. D. Cox, J. Little, and D. O'Shea, *Using algebraic geometry*, second ed., Springer, 2005.
2. ———, *Ideals, varieties, and algorithms*, third ed., Springer, 2007.
3. D. Cox, J. Little, and H. Schenck, *Toric varieties*, AMS, 2011.
4. J. M. Landsberg, *Tensors: geometry and applications*, AMS, 2012.
5. A. Sommese and C. Wampler, *The numerical solution of systems of polynomials*, World Scientific, 2005.

Biography of contributor

Frank Sottile is a Professor of Mathematics at Texas A&M University in College Station, a former Churchill Scholar, and a Fellow of the American Mathematical Society. His research interests range from Algebraic Combinatorics to Applications of Algebraic Geometry. He helped to found the SIAM Activity Group on Algebraic Geometry and served as its first chair.