From Pascal's Theorem to d-Constructible Curves

Will Traves

Abstract. We prove a generalization of both Pascal's Theorem and its converse, the Braikenridge–Maclaurin Theorem: If two sets of k lines meet in k^2 distinct points, and if dk of those points lie on an irreducible curve C of degree d, then the remaining $k(k-d)$ points lie on a unique curve S of degree $k-d$. If S is a curve of degree $k-d$ produced in this manner using a curve C of degree d, we say that S is d-constructible. For fixed degree d, we show that almost every curve of high degree is not d-constructible. In contrast, almost all curves of degree 3 or less are d-constructible. The proof of this last result uses the group structure on an elliptic curve and is inspired by a construction due to Möbius. The exposition is embellished with several exercises designed to amuse the reader.

Dedicated to H.S.M. Coxeter, who demonstrated a heavenly *syzygy*: the sun and moon aligned with the Earth, through a pinhole. (Toronto, May 10, 1994, 12:24:14)

1. INTRODUCTION. In astronomy, the word *syzygy* refers to three celestial bodies that lie on a common line. Other interesting patterns are also sometimes called syzygies. For example, in a triangle, the three median lines that join vertices to the midpoints of opposite sides meet in a common point, the centroid, as illustrated in the left diagram of Figure 1. Choosing coordinates, this fact can be viewed as saying that three objects lie on a line: There is a linear dependence among the equations defining the three median lines. In commutative algebra and algebraic geometry, a syzygy refers to any equation relating the generators of a module.

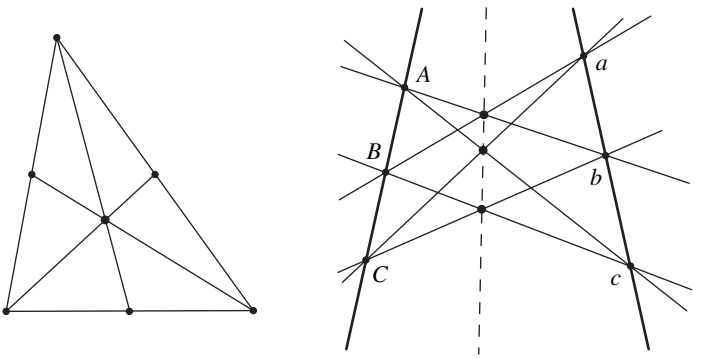

Figure 1. Two syzygies: the centroid (left) and Pappus's configuration (right)

Pappus's Theorem, which dates from the fourth century A.D., describes another syzygy. It is one of the inspirations of modern projective geometry.

http://dx.doi.org/10.4169/amer.math.monthly.120.10.901
MSC: Primary 14H50, Secondary 14H52

Theorem 1 (Pappus). *If three points A, B, and C lie on one line, and three points a, b, and c lie on another, then the lines Ab, Bc, and Ca meet the lines aB, bC, and cA in three new points and these new points are collinear, as illustrated in the right diagram of Figure 1.*

Pappus's Theorem has inspired a lot of amazing mathematics. The first chapter of a fascinating new book by Richter-Gebert [**18**] describes the connections between Pappus's Theorem and many areas of mathematics, including cross-ratios and the Grassmann–Plücker relations among determinants.

Pappus's Theorem appears in his text *Synagogue* [**17**], a collection of classical Greek geometry with insightful commentary. David Hilbert observed that Pappus's Theorem is equivalent to the claim that the multiplication of lengths is commutative (see, e.g., Coxeter [**3**, p. 152]). Thomas Heath believed that Pappus's intention was to revive the geometry of the Hellenic period [**11**, p. 355], but it wasn't until 1639 that the sixteen-year-old Blaise Pascal generalized Pappus's Theorem [**4**, Section 3.8], replacing the two lines with a more general conic section.

Theorem 2 (Pascal). *If six distinct points A, B, C, a, b, and c lie on a conic section, then the lines Ab, Bc, and Ca meet the lines aB, bC, and cA in three new points, and these new points are collinear.*

Pascal's Theorem is sometimes formulated as the Mystic Hexagon Theorem: If a hexagon is inscribed in a conic, then the three points lying on lines extending from pairs of opposite edges of the hexagon are collinear, as in Figure 2. It is not clear why the theorem deserves the adjective *mystic*. Perhaps it refers to the case where a regular hexagon is inscribed in a circle. In that case, the three pairs of opposite edges are parallel and the theorem then predicts that the parallel lines should meet (at infinity), and that all three points of intersection should be collinear. Thus, a full understanding of Pascal's Theorem requires knowledge of the projective plane, a geometric object described in Section 2.

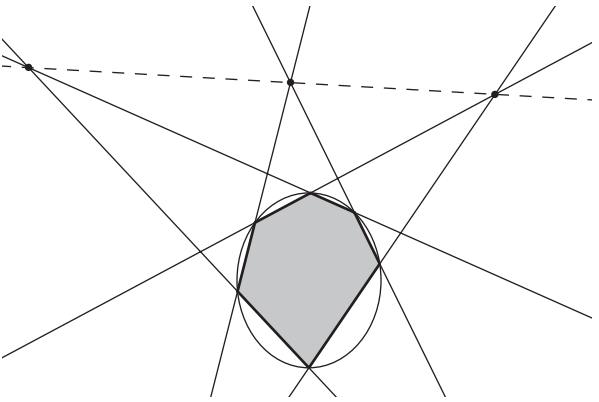

Figure 2. The Mystic Hexagon Theorem

Pascal's Theorem has an interesting converse named after the British mathematicians William Braikenridge and Colin Maclaurin. Braikenridge and Maclaurin seem to have arrived at the result independently, though they knew each other and their correspondence includes a dispute over priority.

Theorem 3 (Braikenridge–Maclaurin). *If three lines meet three other lines in nine points, and if three of these points lie on a line, then the remaining six points lie on a conic.*

In 1848, the astronomer and mathematician August Ferdinand Möbius generalized Pascal's Theorem. Suppose that a polygon with $4n + 2$ sides is inscribed in a nondegenerate conic and we determine $2n + 1$ points by extending opposite edges until they meet. If $2n$ of these $2n + 1$ points of intersection lie on a line, then the last point also lies on the line. Möbius had already developed a system of coordinates for projective figures, but surprisingly his proof relies on solid geometry. In Section 3, we prove an extension of Möbius's result using a significant generalization of Pascal's Theorem and its converse (Theorem 6): When two sets of k lines meet in k^2 distinct points and dk of these points lie on an irreducible curve C of degree $d < k$, then the remaining $k(k - d)$ points lie on a unique curve S of degree $k - d$. If S is a curve of degree $k - d$ produced in this manner using a curve C of degree d, we say that S is d-constructible. In a very interesting article [**12**] in this MONTHLY, Katz asks which curves are 2-constructible. In Section 4, following Katz's arguments, we give a dimension-counting argument to show that most curves of high degree are not d-constructible. In contrast, we show that most curves of degree 3 or less are d-constructible for all $d > 0$. The proof of this last result for cubics involves inscribing polygons in an elliptic curve in a surprising manner.

2. PROJECTIVE GEOMETRY. Applying Pascal's Mystic Hexagon Theorem (Theorem 2) in the case where opposite sides of the hexagon are parallel, suggests that parallel lines should meet in a point and that the collection of such intersection points should lie on a line as we vary the pairs of parallel lines. This is manifestly false in the Euclidean plane, but the plane can be augmented by adding *points at infinity*, after which Pascal's Theorem holds. The resulting projective plane $\mathbb{P}^2$ is a fascinating object with many nice properties.

One powerful model of the projective plane identifies points in $\mathbb{P}^2$ with lines through the origin in 3-dimensional space. To see how this relates to the Euclidean plane, consider the plane $z = 1$ in 3-dimensional space as a model for $\mathbb{R}^2$, and note that most lines through $(0, 0, 0)$ meet this plane. The line passing through $(x, y, 1)$ is identified with the point $(x, y) \in \mathbb{R}^2$. But what about the lines that don't meet this plane? These are parallel to $z = 1$ and pass through $(0, 0, 0)$ so they are lines in the xy-plane. Each of these lines can be viewed as a different point at infinity, since they've been attached to our copy of $\mathbb{R}^2$.

In 1827, Möbius developed a useful system of coordinates for points in projective space [**16**], later extended by Grassmann. If we consider the punctured 3-space $\mathbb{R}^3 \setminus \{(0, 0, 0)\}$ and the equivalence relation

$$(x, y, z) \sim (\lambda x, \lambda y, \lambda z) \Leftrightarrow \lambda \neq 0,$$

then each equivalence class corresponds to a line in $\mathbb{R}^3$ through the origin. We denote the equivalence class of points on the line through (x, y, z) by $[x : y : z]$. This is a sensible notation, since the ratios between the coordinates determine the direction of the line. Returning to our earlier model of $\mathbb{P}^2$, the points with $z \neq 0$ correspond to points in our usual copy of $\mathbb{R}^2$, while the points with $z = 0$ correspond to points at infinity.

If points in $\mathbb{P}^2$ correspond to lines through the origin, then what do *lines* in $\mathbb{P}^2$ look like? If we once again identify the plane $z = 1$ with $\mathbb{R}^2$, we see that the points

making up a line on $z = 1$ correspond to lines through the origin that, together, form a plane. Any line in $\mathbb{R}^2$ can be described by an equation of the form $ax + by + c = 0$; the reader should check that this determines the plane $ax + by + cz = 0$. Thus, lines in $\mathbb{P}^2$ correspond to dimension-2 subspaces of $\mathbb{R}^3$. In particular, the line in $\mathbb{P}^2$ whose equation is $z = 0$ is the *line at infinity*.

Each curve C in the projective plane can be described as the zero-set of a homogeneous polynomial $F(x, y, z)$:

$$C = \{[x : y : z] : F(x, y, z) = 0\}.$$

The polynomial needs to be homogeneous (all terms in the polynomial have the same degree) in order for the curve to be well-defined (see Exercise 4.2 below). It is traditional to call degree-d homogeneous polynomials *degree-d forms*. The curve C is said to be a degree-d curve when $F(x, y, z)$ is a degree-d polynomial. We say that C is an irreducible curve when $F(x, y, z)$ is an irreducible polynomial. When $F(x, y, z)$ factors, then the set C is actually the union of several component curves, each determined by the vanishing of one of the irreducible factors of $F(x, y, z)$.

Exercise 4. If this is the first time you've met projective space, you might try these exercises to get a feel for projective space.

1. Show that the line $ax + by + cz = 0$ in $\mathbb{P}^2$ consists of all the points of the form $[x : y : 1]$ such that $ax + by + c = 0$, together with a single point at infinity (the point $[b : -a : 0]$). We say that the line $ax + by + cz = 0$ is the *projectivization* of the line $ax + by + c = 0$. Now show that the projectivizations of two parallel lines $ax + by + c = 0$ and $ax + by + d = 0$ in $\mathbb{R}^2$ meet at a point at infinity.
2. The projectivization of the hyperbola $xy = 1$ in $\mathbb{R}^2$ is the set of points in $\mathbb{P}^2$ that satisfy $xy - z^2 = 0$. Show that whether a point $[x : y : z]$ lies on the projectivization of the hyperbola or not is a well-defined property (i.e., the answer doesn't depend on which representative of the equivalence class $[x : y : z]$ we use). Where does the projectivization meet the line at infinity?
3. Show that if $a_1x + b_1y + c_1z = 0$ and $a_2x + b_2y + c_2z = 0$ are two distinct lines in $\mathbb{P}^2$, then they meet in a point $P = [a_3 : b_3 : c_3]$ whose coordinates are given by the cross product,

   $$\langle a_3, b_3, c_3\rangle = \langle a_1, b_1, c_1\rangle \times \langle a_2, b_2, c_2\rangle.$$

   Interpret the result in terms of the geometry of 3-dimensional space. Similarly, show that if $P_1 = [a_1 : b_1 : c_1]$ and $P_2 = [a_2 : b_2 : c_2]$ are two points in $\mathbb{P}^2$, then the line through P_1 and P_2 has equation $a_3x + b_3y + c_3z = 0$, with a_3, b_3, and c_3 as above. Describe how to phrase these results for lines and points in $\mathbb{R}^2$.
4. Pascal's Theorem predicts that if a regular hexagon is inscribed in a circle, then the three pairs of opposite edges intersect in three collinear points. Which line do the three points lie on? Is it surprising that it doesn't matter where in the plane the circle is centered?
5. Show that if $F(x, y, z) = 0$ is a homogeneous polynomial equation defining a curve $C \subset \mathbb{P}^2$, then a point $P \in C$ is smooth (that is, there is a uniquely-defined tangent line to C at P) if and only if $\nabla F(P) \neq 0$. (Hint: For which points P is there a tangent plane to the level surface in $\mathbb{R}^3$ given by $F = 0$?)
6. The polynomial $P(x_0, x_1, x_2, y_0, y_1, y_2)$ is said to be bihomogeneous in the variables x_i and y_i if P is homogeneous in the remaining variables when considering

all the x-variables or all the y-variables as constants. Show that if P is bihomogeneous, then the set of points $([x_0 : x_1 : x_2], [y_0 : y_1 : y_2]) \in \mathbb{P}^2 \times \mathbb{P}^2$ such that $P(x_0, x_1, x_2, y_0, y_1, y_2) = 0$ is well-defined.

7. There is an interesting duality between points and lines in $\mathbb{P}^2$. The dual of the point $P = [a : b : c] \in \mathbb{P}^2$ is the line $\check{P} \subset \mathbb{P}^2$, with equation $ax + by + cz = 0$, and vice-versa.
   (a) Show that a line $L : ax + by + cz = 0$ in $\mathbb{P}^2$ goes through two points $P_1 \neq P_2$, if and only if the dual point $\check{L} = [a : b : c]$ lies on the intersection of the two dual lines $\check{P}_1$ and $\check{P}_2$.
   (b) It turns out that the duals of all the tangent lines to an irreducible conic C form a collection of points lying on a dual irreducible conic $\check{C}$, and vice-versa (see Bashelor, Ksir, and Traves [**1**] for details). Show that dualizing Pascal's Theorem gives Brianchon's Theorem: If an irreducible conic is inscribed in a hexagon, then the three lines joining pairs of opposite vertices intersect at a single point.

Projective space $\mathbb{P}^2$ enjoys many nice properties that Euclidean space $\mathbb{R}^2$ lacks. Many results are easier to state and more elegant in projective space than in Euclidean space. For instance, in Euclidean space two distinct lines meet in either one point or in no points (in the case where the two lines are parallel). By adding points at infinity to Euclidean space, we've ensured that *any* two distinct lines meet in a point. This is just the first of a whole sequence of results encapsulated in Bézout's Theorem.

Theorem 5 (Bézout's Theorem). *If C_1 and C_2 are curves of degrees d_1 and d_2 in the complex projective plane $\mathbb{P}^2_{\mathbb{C}}$ sharing no common components, then they meet in $d_1 d_2$ points, counted appropriately.*

Bézout's Theorem requires that we work in *complex* projective space; in $\mathbb{P}^2_{\mathbb{R}}$, two curves may not meet at all. For instance, the line $y - 2z = 0$ misses the circle $x^2 + y^2 - z^2 = 0$ in $\mathbb{P}^2_{\mathbb{R}}$; the points of intersection have complex coordinates. In the rest of the paper, we'll work in complex projective space (denoted $\mathbb{P}^2$) so that we can take advantage of Bézout's Theorem. The points of $\mathbb{P}^2$ correspond to one-dimensional subspaces of $\mathbb{C}^3$.

To say what it means to count points appropriately, requires a discussion of intersection multiplicity. This can be defined in terms of the length of certain modules [**8**], but an intuitive description will be sufficient for our purposes. When two curves meet transversally at a point P (there is no containment relation between their tangent spaces), then P counts as 1 point in Bézout's Theorem. If the curves are tangent at P or if one curve has several branches passing through P, then P counts as a point with multiplicity. One way to determine the multiplicity of P is to look at well-chosen families of curves $C_1(t)$ and $C_2(t)$ so that $C_1(0) = C_1$ and $C_2(0) = C_2$, and to count how many points in $C_1(t) \cap C_2(t)$ approach P as t goes to 0. For instance, the line $y = 0$ meets the parabola $yz = x^2$ in one point $P = [0 : 0 : 1]$. Letting $C_1(t)$ be the family of curves $y - t^2 z = 0$ and letting $C_2(t)$ be the family consisting only of the parabola, we find that if $t \neq 0$, then $C_1(t) \cap C_2(t) = \{[t : t^2 : 1], [-t : t^2 : 1]\}$; so two points converge to P as t goes to 0. In this case, P counts as two points. The reader interested in testing their understanding could check that the two concentric circles $x^2 + y^2 - z^2 = 0$ and $x^2 + y^2 - 4z^2 = 0$ meet in two points, each of multiplicity two. More details can be found in Fulton's lecture notes [**8**, Chapter 1].

It is traditional to call this result Bézout's Theorem because it appeared in his widely-circulated and highly-praised book.[1] However, Isaac Newton proved the result over 80 years before Bézout's book appeared! Both Etienne Bézout (1730–1738) and Charles Julien Brianchon (1783–1864) had positions with the French military. The 18th- and 19th-century French military played an interesting role in supporting the development and teaching of mathematics. As Examiner of the Guards of the Navy in France, Étienne Bézout was responsible for creating new textbooks for teaching mathematics to the students at the French Naval Academy. Kirwan [**13**] gives a nice proof of Bézout's Theorem.

We can also construct higher-dimensional projective spaces. Naturally, we add points at infinity to $\mathbb{C}^n$ to create n-dimensional projective space $\mathbb{P}^n$. As in the two-dimensional case, points in $\mathbb{P}^n$ can be identified with one-dimensional subspaces of $\mathbb{C}^{n+1}$ and each point is denoted using homogeneous coordinates $[x_0 : x_1 : \ldots : x_n]$.

Higher projective spaces arise naturally when considering *moduli spaces* of curves in the projective plane. For instance, consider a degree-2 curve C given by the formula

$$a_0x^2 + a_1xy + a_2xz + a_3y^2 + a_4yz + a_5z^2 = 0. \tag{1}$$

Multiplying the formula by a nonzero constant gives the same curve, so the curve C can be identified with the point $[a_0 : a_1 : a_2 : a_3 : a_4 : a_5]$ in $\mathbb{P}^5$. More generally, letting R_d be the vector space of degree-d homogeneous polynomials in three variables, the degree-d curves in $\mathbb{P}^2$ are identified with points in the projective space $\mathbb{P}(R_d)$, where we identify polynomials if they are nonzero scalar multiples of one another. A basis of R_d is given by the $D = \binom{d+2}{2}$ monomials of degree d in three variables, so the degree-d curves in $\mathbb{P}^2$ are identified with points in the projective space $\mathbb{P}(R_d) \cong \mathbb{P}^{D-1}$.

Returning to the case of degree-2 curves in $\mathbb{P}^2$, if we require C to pass through a given point, then the coefficients $a_0, \ldots, a_5$ of C must satisfy the linear equation produced by substituting the coordinates of the point into (1). Now, if we require C to pass through five points in $\mathbb{P}^2$, then the coefficients must satisfy a homogeneous system of five linear equations in six variables. The Rank-Nullity Theorem shows that such a system always has a non-trivial solution: There is a conic through any five points in $\mathbb{P}^2$. If the points are in general position (so that the resulting system has full rank), then the system has a one-dimensional solution space and so there is a unique conic passing through all five points (see [**1**] for details).

The Zariski topology is the coarsest topology that makes polynomial maps from $\mathbb{P}^m$ to $\mathbb{P}^n$ continuous. More concretely, every homogeneous polynomial F in $n+1$ variables determines a closed set in $\mathbb{P}^n$

$$\mathbb{V}(F) = \{P \in \mathbb{P}^n : F(P) = 0\},$$

and every closed set is built by taking finite unions and arbitrary intersections of such sets. Closed sets in the Zariski topology are called varieties. The nonempty open sets in this topology are dense; such a set is not contained in a proper subset of the form $\mathbb{V}(F)$. We'll say that a property holds for almost every point in $\mathbb{P}^n$ if it holds on a dense Zariski-open set in $\mathbb{P}^n$. More details on the Zariski topology can be found in Shafarevich [**19**, Section 4.1].

[1] Both the MathSciNet and Zentralblatt reviews of the English translation [**2**] are entertaining. The assessment in the MathSciNet review is atypically colorful: "This is not a book to be taken to the office, but to be left at home, and to be read on weekends, as a romance", while the review in Zentralblatt Math calls it "an immortal evergreen of astonishing actual relevance".

3. AN EXTENSION OF PASCAL'S THEOREM. We start this section by establishing a significant generalization of both Pascal's Theorem and the Braikenridge–Maclaurin Theorem.

Theorem 6. *Let $F(x, y, z) = 0$ and $G(x, y, z) = 0$ define two curves of degree k and assume that these curves meet in a set Γ of k^2 distinct points. If $H(x, y, z) = 0$ defines an irreducible curve C of degree $d > 0$ that passes through kd of the points in Γ, then there is a curve $S = 0$ of degree $k - d$ that passes through the remaining $k(k - d)$ points in $\Gamma \setminus C$. Moreover, $S = 0$ is the unique curve of degree $k - d$ containing $\Gamma \setminus C$, if G factors into distinct linear forms $G = G_1 \cdots G_k$.*

The first part of the following proof is due to Kirwan [**13**, Theorem 3.14]. Katz [**12**, Theorem 3.1] gives the proof of the case of Max Noether's Theorem discussed in the second paragraph.

Proof. Let $[a : b : c]$ be a point on C but not in Γ, and define a curve X of degree k by the equation $M(x, y, z) = 0$, where

$$M(x, y, z) = G(a, b, c)F(x, y, z) - F(a, b, c)G(x, y, z).$$

The degree-k curve X meets the degree-d curve C in at least $kd + 1$ points, namely the kd points of $\Gamma \cap C$ and the point $[a : b : c]$, so by Bézout's Theorem, C and X must share a common component. Since C is irreducible, $M(x, y, z) = H(x, y, z)S(x, y, z)$ for some degree $k - d$ form S. Since M vanishes on Γ, the curve defined by $S = 0$ must contain all the $k(k - d)$ points of Γ off C.

Now, we assume that G has k distinct linear factors $G_1, \ldots, G_k$, and show that any degree-k form N defining a curve that contains Γ must satisfy $N = aF + bG$ for suitable constants a and b. This is a special case of a result that Max Noether called the Fundamental Theorem of Algebraic Functions, though today it is known by a more technical name, the $AF + BG$ Theorem. Both N and F have the same zeros when restricted to $G_1 = 0$, so for some constant a, $N - aF$ vanishes identically on the first line $G_1 = 0$, and $N - aF = G_1 Q_{k-1}$ for some degree $k - 1$ form Q_{k-1}. Now, Q_{k-1} vanishes at all points of Γ off the first line $G_1 = 0$. In particular, for each of the remaining lines $G_i = 0$, Q_{k-1} vanishes at k points on the line, hence G_i divides Q_{k-1}. Since the forms $G_2, \ldots, G_k$ are relatively prime, $Q_{k-1} = bG_2 \cdots G_k$ for some constant b and $N = aF + bG$.

Now, if $S_1(x, y, z)$ and $S_2(x, y, z)$ are two forms of degree $k - d$ vanishing on $\Gamma \setminus C$, then $HS_1 = a_1F + b_1G$ and $HS_2 = a_2F + b_2G$ for constants a_1, a_2, b_1, and b_2. Then H must divide both $b_2HS_1 - b_1HS_2 = (b_2a_1 - b_1a_2)F$ and $a_2HS_1 - a_1HS_2 = (a_2b_1 - a_1b_2)G$. Now if $a_2b_1 - a_1b_2 \neq 0$, then the curve C given by $H = 0$ must be contained in the set of points $\Gamma = (F = 0) \cap (G = 0)$, a contradiction. So $a_2b_1 - a_1b_2 = 0$ and the forms HS_1 and HS_2 are scalar multiples of one another. It follows that the curves defined by $S_1 = 0$ and $S_2 = 0$ are identical. ∎

In the rest of the paper we will be interested in the case where *both* of the forms F and G factor completely into linear forms, in which case their zero-sets determine collections of k blue and k red lines, respectively. If the polynomial HS in Theorem 6 also factors completely into linear forms, then the resulting arrangement of lines—in which each point of intersection lies on a line from $HS = 0$, a red line, and a blue line—is called a multinet, an intriguing combinatorial and geometric object in the theory of hyperplane arrangements (see Falk and Yuzvinsky [**7**] for details on the connection between multinets and resonance varieties).

Pascal's Mystic Hexagon Theorem, Theorem 2, follows from an easy application of Theorem 6. Color the lines Ab, Bc, and Ca red and the lines aB, bC, and cA blue. A degree-2 curve passes through six of the intersection points of the blue and red lines, so the remaining three points must lie on a degree-1 curve—they are collinear. To prove the Braikenridge–Maclaurin Theorem, Theorem 3, just color one collection of three lines blue and the other collection red, and apply Theorem 6.

Theorem 6 can be extended to cover the case where the points in Γ are not distinct; however, this would divert us to a discussion of scheme theory. Details can be found in David Eisenbud, Mark Green, and Joe Harris's amazing survey paper [**6**, Section 1.3] on the Cayley–Bacharach Theorem, a vast generalization of Theorem 6. They connect the result to a host of interesting mathematics, including the Riemann–Roch Theorem, residues, and homological algebra. Their exposition culminates in the assertion that the theorem is equivalent to the statement that polynomial rings are Gorenstein. The Cayley–Bacharach Theorem has many practical applications; see, for example, Gold, Little, and Schenck [**9**] for an application in algebraic coding theory.

Rather than state the full Cayley–Bacharach Theorem, we recall an early version of the theorem, first proved by Michel Chasles: If two plane cubic curves meet in nine distinct points, then any other cubic passing through eight of these points must also pass through the ninth. Because of its content, the result is often called the $8 \Rightarrow 9$ Theorem. Chasles used the $8 \Rightarrow 9$ Theorem to prove Pascal's Mystic Hexagon Theorem. The theorem can also be used to prove that the group law on an elliptic curve is associative. Terrence Tao recently gave a simple, elementary proof of the $8 \Rightarrow 9$ Theorem in his blog.[2]

Now we turn to Möbius's generalization of Pascal's Theorem [**15**]. Möbius proved two results in this direction. In the first, a polygon with $4n + 2$ sides is inscribed in a irreducible conic, and we determine $2n + 1$ points by extending opposite edges until they meet. If $2n$ of these $2n + 1$ points of intersection lie on a line, then the last point also lies on the line. Using Theorem 6 allows us to extend Möbius's result, replacing the constraint on the number of sides of the polygon by the constraint that the intersection points are distinct.

Theorem 7. *Suppose that k red lines and k blue lines meet in a set Γ of k^2 distinct points, with $2k$ points of Γ lying on a conic Q. If a line L contains $k - 1$ of the $k^2 - 2k$ points of Γ off Q, then L contains one other point of Γ off Q as well.*

To see the connection with Möbius's result, suppose that a polygon with $2k = 4n + 2$ sides is inscribed in an irreducible conic. Working around the perimeter of the polygon, color the edges alternately red and blue. Since there are $4n + 2$ sides, opposite sides have opposite colors. Extending the edges to lines, consider the $k = 2n + 1$ points of intersection of the pairs of opposite sides. If $k - 1 = 2n$ of these points lie on a line L, then Theorem 7 shows that another of the points in Γ lies on L as well. Since the points of Γ are distinct, the only possibility is that the remaining pair of corresponding edges intersect on the line L.

Proof of Theorem 7. Let $F = F_1 \cdots F_k$ and $G = G_1 \cdots G_k$ be completely reducible forms of degree k whose zero-sets determine the union of the red lines and the union of the blue lines, respectively. Since there are $2k$ points of Γ on the conic Q, Theorem 6 guarantees the existence of a degree-$(k - 2)$ curve C_1, so that C_1 passes through the remaining $k^2 - 2k$ points of Γ off Q. Now, C_1 meets the line L in at least $k - 1$ points, so Bézout's Theorem forces L to be a component of C_1. Write $C_1 = C_2 \cup L$, where

[2]See Tao's July 15, 2011 post at http://terrytao.wordpress.com.

C_2 is a curve of degree $k-3$. Now, $D = C_2 \cup Q$ is a curve of degree $k-1$ that passes through all points of Γ off L. We'll assume that L contains only $k-1$ points of Γ off Q, and derive a contradiction. Under this assumption, D contains $k^2-(k-1) = k(k-1)+1$ points of Γ. So the degree-$(k-1)$ curve D meets the degree-k curve $F=0$ in more than $k(k-1)$ points. It follows that D and $(F_1=0)\cup\cdots\cup(F_k=0)$ share a common component. Relabeling the linear forms in F if necessary, assume that the common component is $F_k=0$. Note that each red line $F_i=0$ only contains k points of Γ: the intersections of $F_i=0$ with the k blue lines. Removing $F_k=0$ from D produces a curve D_1, which meets $(F_1=0)\cup\cdots\cup(F_{k-1}=0)$ in at least $k(k-2)+1$ points. Carrying on in this fashion produces a curve D_{k-2} of degree 1 that meets the line $(F_k=0)$ in precisely $k+1$ points of Γ. This is impossible—the red line $F_k=0$ only contains k points of Γ—so L must contain another point of Γ off Q. If L contains more than k points of Γ, then it must share a common component with both F and G, and hence one of the blue lines must equal one of the red lines. Since the blue and red lines intersect in a finite set of points, this is impossible. So there are precisely k points of Γ on L. ∎

Möbius also proved a result involving two polygons inscribed in a conic. Consider two polygons P_1 and P_2, each with $2k$ edges, inscribed in a conic, and associate one edge from P_1 with one edge from P_2. Working counterclockwise in each polygon, associate the other edges of P_1 with the edges of P_2. Extending these edges to lines, Möbius proved that if $2k-1$ of the intersections of pairs of corresponding edges lie on a line, then the last pair of corresponding edges also meet in a point on this line. Assuming that the points of intersection of the lines are distinct, this result also follows from Theorem 6. A similar construction using a pair of inscribed polygons will reappear when we consider constructible cubics in the next section.

4. CONSTRUCTIBLE CURVES. Let's take a constructive view of Theorem 6.

Definition 8. A curve S of degree t is *d-constructible* if there exist $d+t$ red lines $\ell_1,\ldots,\ell_{d+t}$ and $d+t$ blue lines $L_1,\ldots,L_{d+t}$, so that: (a) $\Gamma = \{\ell_i \cap L_j : 1 \le i, j \le d+t\}$ consists of $(d+t)^2$ distinct points; (b) $d(d+t)$ of the points in Γ lie on a degree-d curve C; and (c) the remaining $t(d+t)$ points in Γ lie on S. Setting $R = \mathbb{C}[x,y,z]$, we say that the d-construction is *dense in degree t* if there is a nonempty Zariski-open set $U \subset \mathbb{P}(R_t)$ so that every degree-t curve in U is d-constructible.

The two curves S and C are said to be *directly linked* via the set Γ. The notion of linkage has important applications in the study of curves. Indeed, special properties of one curve are reflected in special properties of the other curve. This point of view leads to the beautiful subject of liaison theory. The last chapter of Eisenbud [**5**] introduces this advanced topic in commutative algebra; more details can be found in Migliore and Nagel's notes [**14**].

We'll restrict our attention in the rest of the paper to the question of which curves are d-constructible. A simple dimension count shows that most curves of high degree are not d-constructible, so the d-construction is not dense in high degrees.

Theorem 9. *If $d \ge 3$, then the d-construction is not dense in degrees $d+4$ or higher. The 2-construction is not dense in degrees five or higher. The 1-construction is not dense in degrees six or higher.*

Proof. The curves of degree t are parameterized by a projective space of dimension $\binom{t+2}{2} - 1 = (t^2+3t)/2$. Let's try to parameterize the set of d-constructible curves of

degree t. For each such curve, there is a curve C of degree d, $d+t$ blue lines, and $d+t$ red lines, as in Definition 8. There are $\binom{d+2}{2}-1$ degrees of freedom in choosing the curve C and $2(d+t)$ degrees of freedom in choosing the blue lines. Since the red lines must meet C in the $d(d+t)$ points of intersection of the blue lines with C, there are finitely many choices for the red lines (when $d \geq 2$) and so these do not add anything to our dimension count. Altogether, the parameterizing set has dimension $(d^2+3d+4(d+t))/2$. Since this quantity is smaller than $(t^2+3t)/2$ when $t \geq d+4$, the first statement must hold. In fact, the dimension of the 2-constructible curves of degree t is $9+2t$, which is less than $(t^2+3t)/2$ when $t \geq 5$, proving the second claim. When $d=1$, there are not just finitely many choices for the red lines—each red line must pass through one point where the $t+1$ blue lines meet the line C. In this case, two parameters determine the line C, $t+1$ parameters determine the points in Γ on C, and $2(t+1)$ parameters determine the slopes of the blue and red lines. So the 1-constructible curves can be parameterized by a space of dimension $3t+5$. This is smaller than $(t^2+3t)/2$ when $t \geq 6$, proving the last claim. ∎

It is easy to see that all lines are d-constructible. For instance, if L is a line, then choose any set of $d+1$ points on L and pick red and blue lines that pass through these points and meet in $(d+1)^2$ distinct points. Then Theorem 6 shows that there exists a curve C of degree d passing through the remaining points, showing that L is d-constructible.

As well, for all $d>0$, the d-construction is dense in degree 2. The defining polynomial of any conic can be expressed in the form $[x,y,z]A[x,y,z]^T$, where A is a symmetric matrix. The conic is irreducible if and only if $\operatorname{rank}(A)=3$. So the set of irreducible conics is Zariski-open; it is the complement of the hypersurface $\det(A)=0$ in $\mathbb{P}(R_2) \cong \mathbb{P}^5$. Now it is easy to show that any irreducible conic Q is d-constructible. Just inscribe a polygon with $2(d+2)$ edges in Q, color the edges alternately red and blue and, if necessary, move the vertices so that the extensions of the red and blue edges meet in distinct points, Γ. Theorem 6 shows that the points in Γ that lie off Q form a degree-d curve, so Q is d-constructible.

Using the group law on elliptic curves allows us to show that for each d, the d-construction is dense in degree 3. An elliptic curve is a smooth plane curve of degree 3. In particular, each elliptic curve is irreducible; it is not the union of other curves. The points on a fixed elliptic curve E form an abelian group; the sum of three distinct points is equal to the identity element in the elliptic curve group if and only if they are collinear.[3] Figure 3 illustrates the group law on the elliptic curve E given by $y^2z - x^3 + xz^2 = 0$. The point at infinity $[0:1:0] \in E$ serves as the identity element 0_E. The three points A, B, and C are collinear, so $A+B+C=0_E$ and $C=-(A+B)$. The vertical line through C and 0_E meets the curve in one more point D, and since $C+0_E+D=0_E$, D equals $-C=A+B$.

Theorem 10. *Every elliptic curve is d-constructible for each $d>0$.*

Proof. Given an elliptic curve E, we produce a set of $d+3$ blue lines and $d+3$ red lines meeting in a set of $(d+3)^2$ distinct points Γ with $3(d+3)$ of them on E. Then by Theorem 6, there exists a curve S of degree d through the points on $\Gamma \setminus E$, and so E is d-constructible.

To produce the red and blue lines, we start with $d+4$ properly selected points $A_0, B_0, P_1, P_2, \ldots, P_{d+2}$ on E (we'll say more on how to pick the points later). We'll

[3]Two of the points are the same if and only if the line is tangent to E at this point. All three points are equal (to Q, say) if and only if Q is a flex point—the tangent line to E at Q intersects E with multiplicity 3.

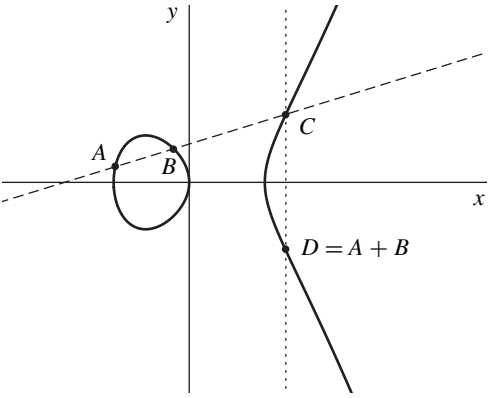

Figure 3. The group law on an elliptic curve

use these points to produce a sequence of auxiliary points $\{A_s, B_s\}(1 \le s \le d+2)$ and Γ will contain all the points A_s and B_s, together with the points $P_1, \ldots, P_{d+2}$. The group law on E helps determine the final point $Q \in \Gamma$. The precise construction breaks into two cases depending on the parity of d.

When d is even, let A_0 and B_0 be arbitrary points on E. For $0 \le s \le d+1$, let A_{s+1} be the third point of intersection of E with the line through A_s and P_{s+1}, and let B_{s+1} be the third point of intersection of E with the line through B_s and P_{s+1}. The points A_s and B_s are the vertices of a polygon inscribed in the elliptic curve, depicted in Figure 4. In the picture, each edge with endpoints labeled a and b and midpoint labeled c represents a line that passes through a, b, and c. The dotted edges correspond to blue lines and the solid edges correspond to red lines.

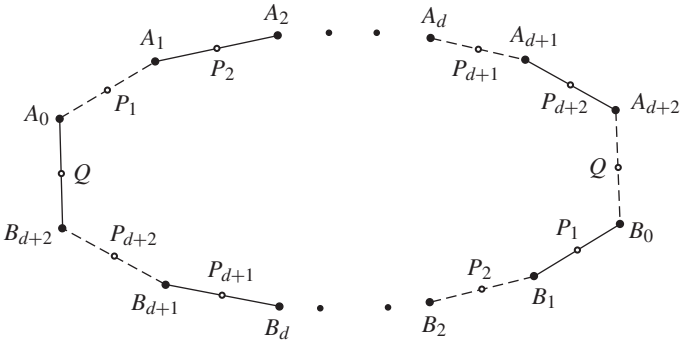

Figure 4. A polygon with $2(d+3)$ vertices inscribed in an elliptic curve

Figure 4 suggests that the line through A_{d+2} and B_0 meets E at the same point, Q, where the line through B_{d+2} and A_0 meets E. To see this, first note that

$$\begin{aligned} A_0 + A_1 + P_1 = 0_E &\Rightarrow A_1 = -A_0 - P_1, \\ A_1 + A_2 + P_2 = 0_E &\Rightarrow A_2 = -A_1 - P_2 = A_0 + P_1 - P_2, \\ &\vdots \\ A_{d+1} + A_{d+2} + P_{d+2} = 0_E &\Rightarrow A_{d+2} = A_0 + P_1 - P_2 + \cdots - P_{d+2}, \end{aligned}$$

and similarly, $B_{d+2} = B_0 + P_1 - P_2 + \cdots - P_{d+2}$. Then

$$\begin{aligned} Q &= -B_0 - A_{d+2} \\ &= -B_0 - (A_0 + P_1 - P_2 + \cdots - P_{d+2}) \\ &= -A_0 - (B_0 + P_1 - P_2 + \cdots - P_{d+2}) \\ &= -A_0 - B_{d+2}. \end{aligned}$$

It follows that $Q \in E$ is collinear with B_0 and A_{d+2}, as well as with A_0 and B_{d+2}.

Note that each point A_s and each point B_s lie on the intersection of one blue and one red line. Because d is even, the number of edges from A_0 to B_0 is odd, so that each point P_s (and the point Q) also lie on both a red and a blue line. So the red and blue lines intersect E in the subset Γ consisting of all the points A_s and B_s, together with the points $P_1, \ldots, P_{d+2}$, and Q.

Figure 4 is misleading when d is odd, because there are an even number of edges from A_0 to B_0, so each point P_s occurs on two lines of the same color. In this case, we adopt Möbius's approach: We inscribe *two* polygons in the elliptic curve, each with $d + 3$ vertices, as in Figure 5.

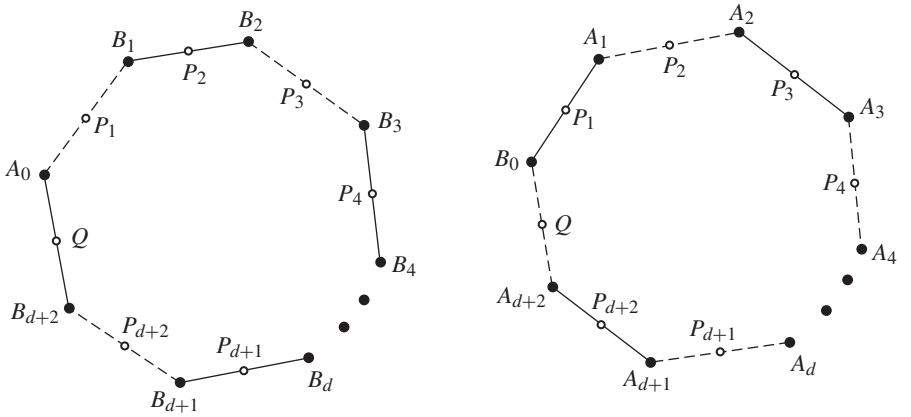

Figure 5. Two polygons, each with $d + 3$ vertices, inscribed in an elliptic curve

To be precise, let A_0 and B_0 be arbitrary points on E. Let B_1 be the third point of intersection of E with the line through A_0 and P_1, and let A_1 be the third point of intersection of E with the line through B_0 and P_1. As in the d even case, for $1 \le s \le d + 1$, let A_{s+1} be the third point of intersection of E with the line through A_s and P_{s+1}, and let B_{s+1} be the third point of intersection of E with the line through B_s and P_{s+1}.

Once again, the line through A_{d+2} and B_0 meets E at the same point, Q, where the line through B_{d+2} and A_0 meets E, as depicted in Figure 5. To see this, we use the same method as in the case where d is even, though the computations are slightly different. If we set $T = -P_1 + P_2 - P_3 + \cdots - P_{d+2}$, then $A_{d+2} = -B_0 + T$ and $B_{d+2} = -A_0 + T$, so

$$Q = -B_0 - A_{d+2} = -B_0 + (B_0 - T) = -A_0 + (A_0 - T) = -A_0 - B_{d+2},$$

from which we conclude the collinearity claims. It follows that when d is odd, the red and blue lines intersect the elliptic curve E in the subset Γ consisting of all the points A_s and B_s, together with the points $P_1, \ldots, P_{d+2}$, and Q.

It remains to show that, no matter what the parity of d, if we pick the points $A_0, B_0, P_1, P_2, \ldots, P_{d+2}$ carefully on E, we can ensure that the intersection points of the red and blue lines are *distinct*. When we dealt with conics, it was possible to move the vertices of the polygons independently to ensure that the points of intersection were distinct, but in the case of cubics the position of one vertex affects all the others. Instead, we give an algorithm that produces a set of lines satisfying a stronger, color blind, statement: We can arrange to make the points of intersection of any two lines distinct, irrespective of their color. At step 0, pick distinct points A_0 and B_0 on E. For step s ($s = 1, \ldots, d+1$) we choose the next point P_{s+1} and form the red and blue lines A_sP_{s+1} and B_sP_{s+1}; these in turn determine the points A_{s+1} and B_{s+1} on E. Let Γ_s be the collection of points A_t, B_t, P_t with $t \le s$, together with all the points of intersection of the (red and blue) lines already constructed. We may assume that the points in Γ_s are distinct and we aim to pick P_{s+1} so that the points in Γ_{s+1} are distinct as well. To do this, we will show that if the points in Γ_{s+1} are not distinct, then $P_{s+1} \in E$ must lie on a (finite) union of lines. Since E is irreducible, these lines meet E in finitely many points, and so it suffices to pick $P_{s+1} \in E$ outside of this finite set.

The new points in Γ_{s+1} that are not obviously in Γ_s are P_{s+1}, A_{s+1}, B_{s+1}, and all the points of intersection of the lines A_sP_{s+1} and B_sP_{s+1} with the previously constructed red and blue lines. If any of these new points equal a point in Γ_s, then P_{s+1} lies on a line joining a point in Γ_s to either A_s or B_s (here, the line joining A_s or B_s to itself should be interpreted as the tangent line to E). As well, if one of the points P_{s+1}, A_{s+1}, or B_{s+1} lies on a previously constructed red or blue line, then, since that point is also on E, it must equal one of P_t, A_t, or B_t for $t \le s$, and so P_{s+1} must again lie on a line joining A_s or B_s to a point of Γ_s. This last case includes the situation where the two new lines are coincident with a previously constructed line, since P_{s+1} is the only point on both new lines. So it remains to determine when P_{s+1}, A_{s+1}, and B_{s+1} are distinct. If $P_{s+1} = A_{s+1}$, then $A_s + 2P_{s+1} = A_s + P_{s+1} + A_{s+1} = 0_E$, so P_{s+1} lies on a line tangent to E that passes through A_s. The number of lines tangent to a nonsingular curve that passes through a point is independent of the point and is called the class of the curve; the class of E is six (see Fulton [**8**]). Similarly, if $P_{s+1} = B_{s+1}$, then P_{s+1} lies on a line tangent to E through B_s. It turns out that A_{s+1} cannot be equal to B_{s+1}; if $A_{s+1} = B_{s+1}$, then canceling terms in $A_s + P_{s+1} + A_{s+1} = 0_E = B_s + P_{s+1} + B_{s+1}$ forces $A_s = B_s$, which contradicts our assumption that the points in Γ_s are distinct.

For $s = 1, \ldots, d+1$, the points in Γ_{s+1} are distinct as long as P_{s+1} does not lie on a line joining A_s or B_s to a point in Γ_s, or on a line through A_s or B_s that is tangent to E. In the final stage of the construction (step $s = d+1$) we need to take additional precautions with the choice of P_{s+1}. In choosing P_{d+2}, we not only determine lines $A_{d+1}P_{d+2}$ and $B_{d+1}P_{d+2}$, but also determine lines A_0B_{d+2} and B_0A_{d+2} and the point $Q \in E$ on their intersection, as illustrated in Figure 6.

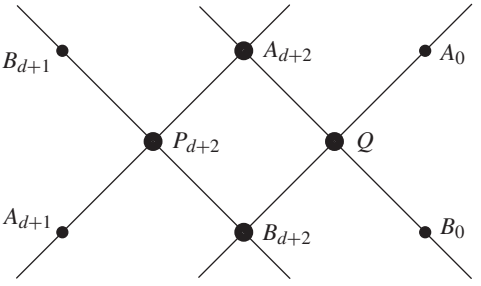

Figure 6. P_{d+2} determines A_{d+2}, B_{d+2}, and Q

If $Q \in \Gamma_{d+1}$, then

$$B_{d+2} \in \Delta = \{p \in E : p \text{ lies on a line joining } A_0 \text{ or } B_0 \text{ to a point of } \Gamma_{d+1}\}.$$

Thus, P_{d+2} lies on a line joining A_{d+1} or B_{d+1} to a point of Δ. This case includes the situation where $Q \in E$ lies on a previously constructed line, because then the point of intersection is in E and so is in Γ_{d+1}. If one of the new lines passes through a point of intersection of the previous lines, then again P_{d+2} must lie on a line joining A_{d+1} or B_{d+1} to a point of Δ. If $Q = A_{d+2}$, then $2A_{d+2} + B_0 = 0_E$, so

$$A_{d+2} \in \Lambda = \{p \in E : p \text{ lies on a tangent line to } E \text{ that passes through } A_0 \text{ or } B_0\}.$$

It follows that P_{d+2} lies on a line joining either B_{d+1} or A_{d+1} to a point of Λ (similarly, if $Q = B_{d+2}$, then the same conclusion holds). Finally, if $Q = P_{d+2}$, then $A_0 + Q + B_{d+2} = 0_E = B_{d+1} + P_{d+2} + B_{d+2}$ and canceling terms gives $A_0 = B_{d+1}$, which is impossible by our construction, since all the points of Γ_{d+1} are distinct.

So the points in $\Gamma = \Gamma_{d+2} \cup \{Q\}$ are distinct as long as P_{d+2} does not lie on a line joining A_{d+1} or B_{d+1} to a point in $\Gamma_{d+1} \cup \Delta \cup \Lambda$ or on a line through A_{d+1} or B_{d+1} that is tangent to E. By avoiding poor choices of the points P_s, we can ensure that all the red lines intersect the blue lines in distinct points. This completes the proof that each elliptic curve is d-constructible. ∎

Corollary 11. *The d-construction is dense in degree* 3.

Proof. To show that the d-construction is dense in degree 3, it is enough to show that the elliptic curves form a dense open set in the set $\mathbb{P}(R_3) \cong \mathbb{P}^9$ parameterizing all degree-3 curves. This is well known, but we sketch the proof. Consider the set

$$\mathcal{C} = \{(F, P) \in \mathbb{P}(R_3) \times \mathbb{P}^2 : P \text{ is a singular point of the curve } F = 0\}.$$

One way to check whether a point P is singular on the level curve $F = 0$ is to check whether $\nabla F(P)$ is zero—in this case, there is no well-defined tangent line (see Exercise 4.5). Since $\nabla F(P)$ is a bihomogeneous polynomial, the set $\mathcal{C}$ is Zariski-closed in the product of projective spaces $\mathbb{P}(R_3) \times \mathbb{P}^2$ (see Exercise 4.6 and Shafarevich [**19**, Section 5.1] for details). Now, the image of a projective variety under the projection $\pi_1 : \mathbb{P}(R_3) \times \mathbb{P}^2 \to \mathbb{P}(R_3)$ is also Zariski-closed (see Shafarevich [**19**, Section 5.2] for details), so the set $\pi_1(\mathcal{C})$ of singular degree-3 curves is Zariski-closed. It follows that the set of smooth (nonsingular) curves is Zariski-open, and hence dense, in $\mathbb{P}(R_3)$. ∎

Theorem 9 suggests that for all d, the d-construction is also dense in degree 4. Proving this result seems to require that we inscribe polygons in degree-4 curves, but this appears to be difficult to do in general.

ACKNOWLEDGMENTS. I am grateful for conversations with my colleagues Mark Kidwell, Mark Meyerson, Thomas Paul, and Max Wakefield, and with my friends Keith Pardue and Tony Geramita. Amy Ksir and Jessica Sidman provided very useful comments on an early draft, and several referees gave excellent feedback that considerably improved the exposition. Many computations and insights were made possible using the excellent software packages Macaulay2, GeoGebra, Sage, and Maple.

REFERENCES

1. A. Bashelor, A. Ksir, W. Traves, Enumerative algebraic geometry of conics, *Amer. Math. Monthly* **115** (2008) 701–728.

2. E. Bézout, *General Theory of Algebraic Equations*. Translated from the 1779 French original by Eric Feron. Princeton University Press, Princeton, NJ, 2006.
3. H. S. M. Coxeter, *Projective Geometry*. Blaisdell Publishing Co. Ginn and Co., New York-London-Toronto, 1964.
4. H. S. M. Coxeter, S. L. Greitzer, *Geometry Revisited*, The Mathematical Association of America, Washington, DC, 1967.
5. D. Eisenbud, *Commutative Algebra with a View Toward Algebraic Geometry*. Springer-Verlag, New York, 1995.
6. D. Eisenbud, M. Green, J. Harris, Cayley–Bacharach theorems and conjectures, *Bull. Amer. Math. Soc. (N.S.)* **33** (1996) 295–324, available at `http://dx.doi.org/10.1090/S0273-0979-96-00666-0`.
7. M. Falk, S. Yuzvinsky, Multinets, resonance varieties, and pencils of plane curves, *Compos. Math.* **143** (2007) 1069–1088.
8. W. Fulton, *Introduction to Intersection Theory in Algebraic Geometry*. American Mathematical Society, Providence, RI, 1994.
9. L. Gold, J. Little, H. Schenck, Cayley–Bacharach and evaluation codes on complete intersections, *J. Pure Appl. Algebra* **196** (2005) 91–99, available at `http://dx.doi.org/10.1016/j.jpaa.2004.08.015`.
10. D. R. Grayson, M. E. Stillman, Macaulay2, a software system for research in algebraic geometry (2010), available at `http://www.math.uiuc.edu/Macaulay2/`.
11. T. Heath, *A History of Greek Mathematics*, Vol. II. Dover, New York, 1981.
12. G. Katz, Curves in cages: an algebro-geometric zoo, *Amer. Math. Monthly* **113** (2006) 777–791.
13. F. Kirwan, *Complex Algebraic Curves*. Cambridge University Press, Cambridge, 1992.
14. J. C. Migliore, U. Nagel, Liaison and related topics: notes from the Torino workshop-school, *Rend. Sem. Mat. Univ. Politec. Torino* **59** (2003) 59–126.
15. A. F. Möbius, Verallgemeinerung des pascalschen theorems, das in einen kegelschnitt beschriebene sechseck betreffend, *J. Reine Angew. Math.* (*Crelle's Journal*) **36** (1848) 216–220.
16. A. F. Möbius, *Der Barycentrische Calcul*. Georg Olms Verlag, Hildesheim, 1976.
17. Pappus of Alexandria, *Book 7 of the Collection*, Part 1. Introduction, text, and translation. Part 2. Commentary, index, and figures. Edited and with translation and commentary by A. Jones. Springer, Berlin, 1986.
18. J. Richter-Gebert, *Perspectives on Projective Geometry. A Guided Tour through Real and Complex Geometry*. Springer, Heidelberg, 2011.
19. I. R. Shafarevich, *Basic Algebraic Geometry*, Vol. 1, Varieties in Projective Space, second edition. Translated from the 1988 Russian edition and with notes by Miles Reid. Springer, Berlin, 1994.

WILL TRAVES grew up in Toronto, Canada and moved to the United States during graduate school. His paper [**1**] in this MONTHLY, co-authored with Andy Bashelor and Amy Ksir, won both the Lester R. Ford and Merten M. Hasse prizes. He joined the faculty of the United States Naval Academy in 1999 and is a "brown-dot" Project NExT fellow.
U.S. Naval Academy, Math Department, Mail Stop 9E, Annapolis, MD, 21402
traves@usna.edu