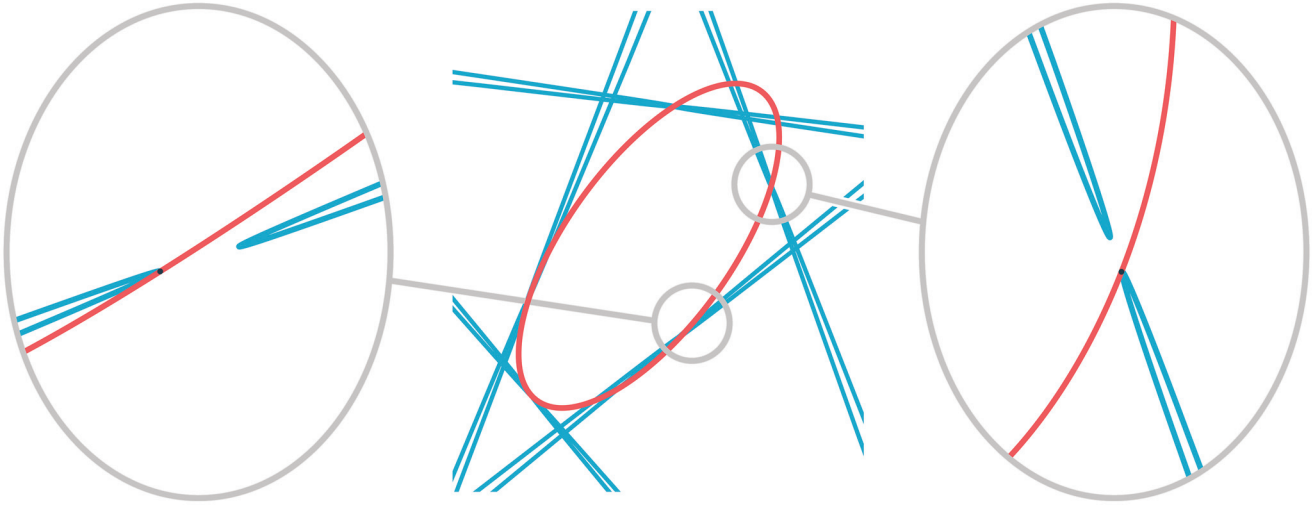
3264 Conics in a Second



Paul Breiding, Bernd Sturmfels, and Sascha Timme

This article and its accompanying web interface present *Steiner's conic problem* and a discussion on how *enumerative* and *numerical* algebraic geometry complement each other. The intended audience is students at an advanced undergrad level. Our readers can see current computational tools in action on a geometry problem that has inspired scholars for two centuries. The take-home message is that numerical methods in algebraic geometry are fast and reliable.

We begin by recalling the statement of Steiner's conic problem. A *conic* in the plane $\mathbb{R}^2$ is the set of solutions to a quadratic equation $A(x,y)=0$, where

$$A(x,y) = a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6. \qquad (1)$$

If there is a second conic

$$U(x,y) = u_1x^2 + u_2xy + u_3y^2 + u_4x + u_5y + u_6, \qquad (2)$$

then the two conics intersect in four points in $\mathbb{C}^2$, counting multiplicities and counting intersections at points at infinity, provided A and U are irreducible and not multiples of each other. This is the content of *Bézout's theorem*. To take into account the points of intersection at infinity, algebraic geometers like to replace the affine plane $\mathbb{C}^2$ with the complex projective plane $\mathbb{P}^2_{\mathbb{C}}$. In the following, when we write "count," we always mean counting solutions in projective space. Nevertheless, for our exposition we work with $\mathbb{C}^2$.

A solution (x,y) of the system $A=U=0$ has multiplicity ≥ 2 if it is a zero of the *Jacobian determinant*

$$\frac{\partial A}{\partial x}\cdot\frac{\partial U}{\partial y} - \frac{\partial A}{\partial y}\cdot\frac{\partial U}{\partial x} = 2(a_1u_2 - a_2u_1)x^2 + 4(a_1u_3 - a_3u_1)xy + \cdots + (a_4u_5 - a_5u_4). \qquad (3)$$

Geometrically, the conic U is *tangent* to the conic A if (1), (2), and (3) are zero for some $(x,y)\in\mathbb{C}^2$. For instance, Figure 1 shows a red ellipse and five other blue conics, which are tangent to the red ellipse. *Steiner's conic problem* asks the following question:

> How many conics in the plane are tangent to five given conics in general position?

The number is five, because each tangency condition removes one of the five degrees of freedom in a conic.

The present article concerns the following two subject areas and how they approach Steiner's problem:

Enumerative algebraic geometry:
How many conics are tangent to five conics?
Numerical algebraic geometry:
How do we find all conics tangent to five conics?

The first question is the original conic problem, first asked in 1848 by Steiner, who suggested the answer 7776. That

Paul Breiding is a postdoctoral researcher at Technische Universität Berlin. His email address is breiding@math.tu-berlin.de.
Bernd Sturmfels is a director of the Max Planck Institute for Mathematics in the Sciences, Leipzig, and a professor of mathematics, statistics, and computer science at University of California, Berkeley. His email address is bernd@mis.mpg.de.
Sascha Timme is a PhD candidate at Technische Universität Berlin. His email address is timme@math.tu-berlin.de.

Communicated by Notices *Associate Editor Daniel Krashen.*

DOI: https://doi.org/10.1090/noti2010

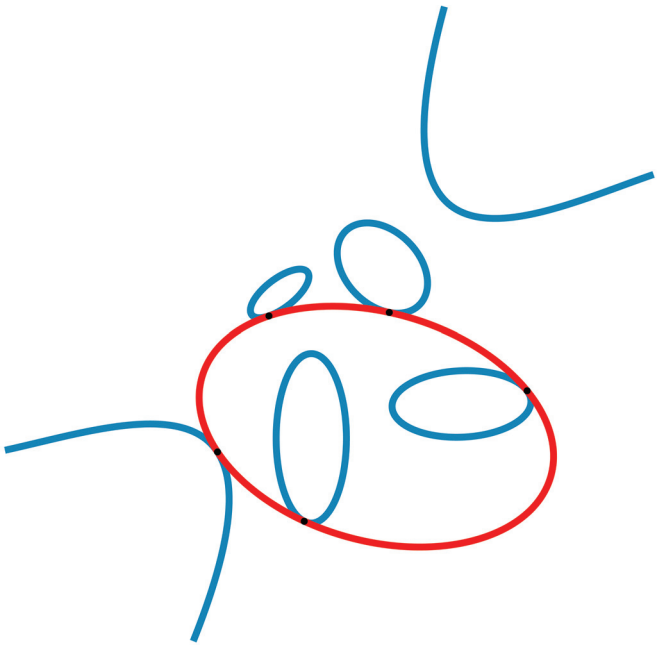


Figure 1. The red ellipse is tangent to four blue ellipses and one blue hyperbola.

number turned out to be incorrect. In the year 1864 Chasles gave the correct answer of 3264. This was further developed by Schubert, whose 1879 book led to Hilbert's 15th problem and thus to the twentieth-century development of enumerative algebraic geometry. The number 3264 appears prominently in the title of the textbook by Eisenbud and Harris [EH16]. A delightful introduction to Steiner's problem was presented by Bashelor, Ksir, and Traves in [BKT08].

Numerical algebraic geometry is a younger subject. It started about forty years ago, going back at least to [GZ79]. The textbook by Sommese and Wampler [SW05] is a standard reference. It focuses on numerical solutions to polynomial equations. The field is now often seen as a branch of applied mathematics. But, as we demonstrate in this article, its methodology can be used in pure mathematics too.

An instance of our problem is given by a list of $30 = 5{\times}6$ coefficients in $\mathbb{R}$ or $\mathbb{C}$:

$$\begin{aligned} A(x,y) &= a_1x^2 + a_2xy + a_3y^2 + a_4x + a_5y + a_6, \\ B(x,y) &= b_1x^2 + b_2xy + b_3y^2 + b_4x + b_5y + b_6, \\ C(x,y) &= c_1x^2 + c_2xy + c_3y^2 + c_4x + c_5y + c_6, \\ D(x,y) &= d_1x^2 + d_2xy + d_3y^2 + d_4x + d_5y + d_6, \\ E(x,y) &= e_1x^2 + e_2xy + e_3y^2 + e_4x + e_5y + e_6. \end{aligned} \tag{4}$$

By eliminating the two unknowns x and y from the three equations (1), (2), and (3), we can write the tangency condition directly in terms of the $12 = 6 + 6$ coefficients $a_1, \dots, a_6, u_1, \dots, u_6$ of A and U:

$$\begin{aligned} \mathcal{T}(A,U) \;=\; & 256a_1^4a_3^2u_3^2u_6^4 - 128a_1^4a_3^2u_3u_5^2u_6^3 \\ & + 16a_1^4a_3^2u_5^4u_6^2 + \cdots + a_5^4a_6^2u_1^2u_2^4. \end{aligned} \tag{5}$$

The polynomial $\mathcal{T}$ is a sum of 3210 terms. It is of degree six in the variables $a_1, \dots, a_6$ and of degree six in $u_1, \dots, u_6$. Known classically as the *tact invariant*, it vanishes precisely when the two conics are tangent.

If the coefficients are general, we can assume that each conic U that is tangent to A, B, C, D, E has nonzero constant term u_6. We can then set $u_6 = 1$. Steiner's problem for the conics A, B, C, D, E now translates into a system of five polynomial equations in five unknowns, u_1, u_2, u_3, u_4, u_5. Each of the five tangency constraints is an equation of degree six:

$$\mathcal{T}(A,U) \;=\; \mathcal{T}(B,U) \;=\; \cdots \;=\; \mathcal{T}(E,U) \;=\; 0. \tag{6}$$

Steiner used Bézout's theorem to argue that these equations have $6^5 = 7776$ solutions. However, this number overcounts, because there is a *Veronese surface* of extraneous solutions U, namely, the squares of linear forms. These degenerate conics have the form

$$U(x,y) \;=\; (x,y,1) \cdot \ell^T\ell \cdot (x,y,1)^T,$$

where $\ell = (\ell_1, \ell_2, \ell_3)$ is a row vector in $\mathbb{C}^3$. Since $U(x,y) = (x,y,1)\begin{pmatrix} 2u_1 & u_2 & u_4 \\ u_2 & 2u_3 & u_5 \\ u_4 & u_5 & 2u_6 \end{pmatrix}(x,y,1)^T$, the condition for U to be a square is equivalent to

$$\operatorname{rank}\begin{pmatrix} 2u_1 & u_2 & u_4 \\ u_2 & 2u_3 & u_5 \\ u_4 & u_5 & 2u_6 \end{pmatrix} \;\le\; 1. \tag{7}$$

This discussion leads us to the following algebraic reformulation of Steiner's conic problem:

$$\begin{aligned} &\text{Find all solutions } U \text{ of the equations (6)} \\ &\text{such that the matrix in (7) has rank} \geq 2. \end{aligned} \tag{8}$$

Ronga, Tognoli, and Vust [RTV97] proved the existence of five real conics whose 3264 conics all have real coefficients. In their argument they do not give an explicit instance but rather show that in the neighborhood of some particular conic arrangement there must be an instance that has all of the 3264 conics real. Hence, this raises the following problem:

$$\begin{aligned} &\text{Find an explicit instance } A, B, C, D, E \text{ such} \\ &\text{that the 3264 solutions } U \text{ to (8) are all real.} \end{aligned} \tag{9}$$

Using numerical algebraic geometry we discovered the solution in Figure 2. We claim that all the 3264 conics that are tangent to those five conics are real.

Proposition 1. *There are* 3264 *real conics tangent to those given by the* 5×6 *matrix in Figure 2.*

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 & a_5 & a_6 \\ b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ c_1 & c_2 & c_3 & c_4 & c_5 & c_6 \\ d_1 & d_2 & d_3 & d_4 & d_5 & d_6 \\ e_1 & e_2 & e_3 & e_4 & e_5 & e_6 \end{bmatrix} = \begin{bmatrix} \frac{10124547}{662488724} & \frac{8554609}{755781377} & \frac{5860508}{2798943247} & \frac{-251402893}{1016797750} & \frac{-25443962}{277938473} & 1 \\ \frac{520811}{1788018449} & \frac{2183697}{542440933} & \frac{9030222}{652429049} & \frac{-12680955}{370629407} & \frac{-24872323}{105706890} & 1 \\ \frac{6537193}{241535591} & \frac{-7424602}{363844915} & \frac{6264373}{1630169777} & \frac{13097677}{39806827} & \frac{-29825861}{240478169} & 1 \\ \frac{13173269}{2284890206} & \frac{4510030}{483147459} & \frac{2224435}{588965799} & \frac{33318719}{219393000} & \frac{92891037}{755709662} & 1 \\ \frac{8275097}{452566634} & \frac{-19174153}{408565940} & \frac{5184916}{172253855} & \frac{-23713234}{87670601} & \frac{28246737}{81404569} & 1 \end{bmatrix}.$$

Figure 2. The five conics from Proposition 1.

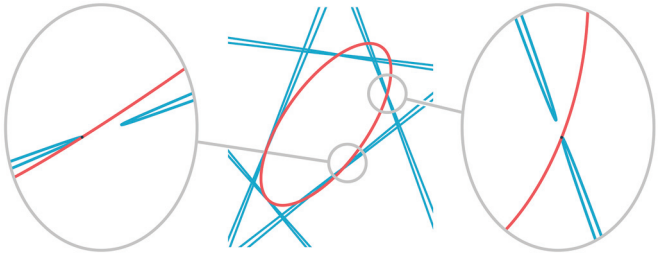


Figure 3. The five blue conics in the central picture are those in Proposition 1. Shown in red is one of the 3264 real conics that are tangent to the blue conics. Each blue conic looks like a pair of lines, but it is a thin hyperbola whose branches are close to each other. The two pictures on the sides show closeups around two of the five points of tangency. The red conic is tangent to one of the two branches of the blue hyperbola.

We provide an animation showing all the 3264 real conics of this instance at this URL:

`www.juliahomotopycontinuation.org/3264/`.

The construction of our example originates from an arrangement of double lines, which we call the *pentagon construction*. One can see the pentagon in the middle of Figure 3. There are points where the red conic seems to intersect a blue line, but they are actually points where the red conic touches one branch of a blue hyperbola. See [Sot] for further details.

Later we shall discuss the algebro-geometric meaning of the pentagon construction, and we present a rigorous computer-assisted proof that indeed all of the 3264 conics tangent to our five conics are real. But, first, let us introduce our web browser interface.

Do It Yourself

In this section we invite you, the reader, to choose your own instance of five conics. We offer a convenient way for you to compute the 3264 complex conics that are tangent to your chosen conics. Our web interface for solving instances of Steiner's problem is found at

`juliahomotopycontinuation.org/diy/`. (10)

Here you can type in your own $30 = 5 \times 6$ coefficients for the conics in (4). After specifying five conics, you press a button and this calls in the numerical algebraic geometry software `HomotopyContinuation.jl`. This is the open source `Julia` package described in [BT18]. Those playing with the web interface need not worry about the inner workings. But, if you are curious, please read our section titled "How Does This Work?"

Shortly after the user submits their instance, by entering real coefficients, the web interface reports whether the instance was generic enough to yield 3264 distinct complex solutions. These solutions are computed numerically. The browser displays the number of real solutions, along with a picture of the instance and a rotating sample of real solutions. As promised in our title, the computation of all solutions takes only around one second.

Remark 2. We always assume that the five given conics are real and generic. This ensures that there are 3264 complex solutions, and these conics are tangent to the given conics at 5×3264 distinct points. The number of real solutions is even, and our web interface displays them sequentially. *For every real solution, the points of tangency on the given conics are also real.* This fact uses the genericity assumption, since two particular real conics can be tangent at two complex conjugate points. For instance, the conics defined by $x^2 - y^2 + 1$ and $x^2 - 4y^2 + 1$ are tangent at the points $(i : 0)$ and $(-i : 0)$ where $i = \sqrt{-1}$.

Figure 4 shows what the input and the visual output of our web interface look like. The user inputs five conics, and the system shows these in blue. After the user clicks the "compute" button, it responds with the number of complex and real conics that were found. The 3264 conics, along with all points of tangency, are available to the user upon request. The real conics are shown in red, as seen on the right in Figure 4.

When seeing this output, the user might ask a number of questions. For instance, among the real conics, how many are ellipses and how many are hyperbolas? Our web interface answers this question. The distinction between ellipses and hyperbolas is characterized by the eigenvalues of the real symmetric matrix

$$\begin{pmatrix} 2u_1 & u_2 \\ u_2 & 2u_3 \end{pmatrix}.$$

If the two eigenvalues of this matrix have opposite signs, then the conic is a hyperbola. If they have the same sign, then the conic is an ellipse. Among the ellipses, we might

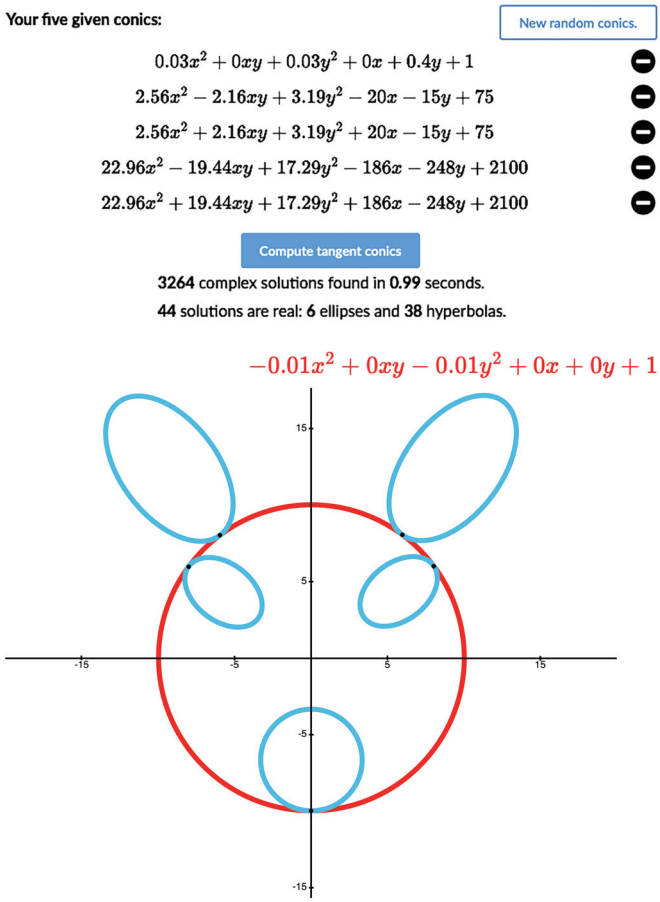


Figure 4. Input and output of the web interface (10).

ask for the solution that looks most like a circle. Our program does this by minimizing the expression $(u_1-u_3)^2+u_2^2$. Users with a numerical analysis background might be interested in maximizing the distance to the degenerate conics. Equivalently, we ask: Among all 3264 solutions, which 3×3 matrix in (7) has the smallest *condition number*?

You can adapt all of this for your favorite geometry problems. As pointed out above, the `Julia` package `HomotopyContinuation.jl` is available to everyone—follow the link at [BT18]. This may enable you to solve your own polynomial systems in record time.

Chow Rings and Pentagons

We next present the approach to deriving the number 3264 that would be taught in an algebraic geometry class, along the lines of the article [BKT08]. Thereafter we explain the geometric degeneration we used to construct the fully real instance in Proposition 1.

Steiner phrased his problem as that of solving five equations of degree six on the five-dimensional space $\mathbb{P}^5_{\mathbb{C}}$. The incorrect count occurred because of the locus of double conics in $\mathbb{P}^5_{\mathbb{C}}$. This is a surface of extraneous solutions. One fixes the problem by replacing $\mathbb{P}^5_{\mathbb{C}}$ with another five-dimensional manifold, namely, the *space of complete conics*. This space is the blow-up of $\mathbb{P}^5_{\mathbb{C}}$ at the locus of double lines. It is a compactification of the space of nonsingular conics that has desirable geometric properties. A detailed description of this construction, suitable for a first course in algebraic geometry, can be found in [BKT08, §5.1].

In order to answer enumerative geometry questions about the space of complete conics, one considers its Chow ring, as explained in [BKT08, §5.2]. Elements in the Chow ring of the space of complete conics correspond to subvarieties of this space—more precisely, to *classes* of subvarieties. Two subvarieties belong to the same class if and only if they are *rationally equivalent*. Rational equivalence is a technical concept. We refer interested readers to the textbook by Eisenbud and Harris [EH16]. The Chow ring for the space of complete conics is worked out in [EH16, §8.2.4]. Nevertheless, the idea behind studying Chow rings is crystal clear: taking intersections of varieties is translated to multiplication in the Chow ring. In the remainder of this section we will see this in action.

The Chow ring of the space of complete conics contains two special classes P and L. The class P encodes the conics passing through a fixed point, while the class L encodes the conics tangent to a fixed line. The following relations hold in the Chow ring:

$$P^5 = L^5 = 1, \quad P^4L = PL^4 = 2, \quad P^3L^2 = P^2L^3 = 4.$$

These relations are derived in [BKT08, §§4.4–5.3]. For instance, the first equation means that if we take five general conics passing through a fixed point, then the intersection contains one point (namely, the point we fixed in the first place). See [BKT08, Table 3] for the geometric meaning of the other equations.

We write C for the class of conics that are tangent to a given conic. In the Chow ring, we have

$$C \;=\; 2P + 2L.$$

This identity is derived in [BKT08, equation (8)]. Our preferred proof is to inspect the first three terms in the expression (5) for the tact invariant $\mathcal{T}(A, U)$:

$$\mathcal{T} = 16 \cdot u_6^2(4u_3u_6 - u_5^2)^2 \cdot a_1^4a_2^3 \mod\langle a_2, a_3^3, a_4, a_5, a_6\rangle.$$

This has the following intuitive interpretation. We assume that the given fixed conic A satisfies

$$|a_1| \gg |a_3| \gg \max\{|a_2|, |a_4|, |a_5|, |a_6|\}. \tag{11}$$

Thus the conic A is close to $x^2-\epsilon y^2$, where ϵ is a small quantity. The process of letting ϵ tend to zero is understood as a degeneration in the sense of algebraic geometry. With this, the condition for U to be tangent to A degenerates to $u_6^2 \cdot (4u_3u_6 - u_5^2)^2 = 0$.

The first factor u_6 represents all conics that pass through the point $(0,0)$. The second factor $4u_3u_6 - u_5^2$ represents all conics tangent to the line $\{x = 0\}$. The Chow ring classes of these factors are P and L. Each of these arises with multiplicity 2, as seen from the exponents. The desired intersection number is now obtained from the Binomial Theorem:

$$\begin{aligned} C^5 &= 32(L+P)^5 \\ &= 32(L^5 + 5L^4P + 10L^3P^2 + 10L^2P^3 + 5LP^4 + P^5) \\ &= 32(1 + 5\cdot 2 + 10\cdot 4 + 10\cdot 4 + 5\cdot 2 + 1) \\ &= 32\cdot 102 = 3264\,. \end{aligned}$$

The final step in turning this into a rigorous proof of Chasles' result is carried out in [BKT08, §7].

The degeneration idea in (11) can be used to construct real instances of Steiner's problem whose 3264 solutions are all real. Fulton first observed this and communicated it to Sottile, who then wrote down Fulton's proof in detail [Sot95, Sot]. Ronga, Tognoli, and Vust [RTV97] independently gave a proof. Apparently, they did not know about Fulton's ideas.

Fix a convex pentagon in $\mathbb{R}^2$ and one special point somewhere in the relative interior of each edge. Consider all conics C such that, for each edge of the pentagon, C either passes through the special point or is tangent to the line spanned by the edge. By the count above, there are $(L+P)^5 = 102$ such conics C. If the pentagon is chosen sufficiently asymmetric, then the 102 conics are all real. We now replace each pointed edge by a nearby hyperbola, satisfying (11). For instance, if the edge has equation $x = 0$ and $(0,0)$ is its special point, then we take the hyperbola $x^2 - \epsilon y^2 + \delta$, where $\epsilon > \delta > 0$ are very small. After making appropriate choices of these parameters along all edges of the pentagon, each of the 102 conics splits into 32 conics, each tangent to the five hyperbolas. Here "splits" means, if the process is reversed, then the 32 different conics collapse into one solution of multiplicity 32. By construction, all 3264 conics are real.

The argument shows that there *exists* an instance in the neighborhood of the pentagon whose 3264 conics are all real, but it does not say *how close* they should be. Serious hands-on experimentation was necessary for finding the instance in Proposition 1.

We next present an **alternative formulation** of Steiner's conic problem. The idea is to remember the five points of tangency on each solution conic. The five sextics in (6) did not involve these points. They were obtained directly from the tact invariant. The next system of equations avoids the use of the tact invariant. It uses five copies of the equations (1)–(3), each with a different point of tangency (x_i, y_i), for $i = 1,2,3,4,5$. The ten equations from (1) and (2) are quadrics. The five equations from (3) are cubics. Altogether, we get the following system of fifteen equations, which we display as a 5×3 matrix $F_{(A,B,C,D,E)}$:

$$\begin{bmatrix} A(x_1,y_1) & U(x_1,y_1) & (\frac{\partial A}{\partial x}\frac{\partial U}{\partial y} - \frac{\partial A}{\partial y}\frac{\partial U}{\partial x})(x_1,y_1) \\ B(x_2,y_2) & U(x_2,y_2) & (\frac{\partial B}{\partial x}\frac{\partial U}{\partial y} - \frac{\partial B}{\partial y}\frac{\partial U}{\partial x})(x_2,y_2) \\ C(x_3,y_3) & U(x_3,y_3) & (\frac{\partial C}{\partial x}\frac{\partial U}{\partial y} - \frac{\partial C}{\partial y}\frac{\partial U}{\partial x})(x_3,y_3) \\ D(x_4,y_4) & U(x_4,y_4) & (\frac{\partial D}{\partial x}\frac{\partial U}{\partial y} - \frac{\partial D}{\partial y}\frac{\partial U}{\partial x})(x_4,y_4) \\ E(x_5,y_5) & U(x_5,y_5) & (\frac{\partial E}{\partial x}\frac{\partial U}{\partial y} - \frac{\partial E}{\partial y}\frac{\partial U}{\partial x})(x_5,y_5) \end{bmatrix}. \quad (12)$$

Each matrix entry is a polynomial in the 15 variables $u_1, \ldots, u_5, x_1, y_1 \ldots, x_5, y_5$. The parameters of this system are the coefficients of the conics A, B, C, D, E. The system of five equations seen in (6) is obtained by eliminating the 10 variables $x_1, y_1, x_2, y_2, x_3, y_3, x_4, y_4, x_5, y_5$ from the new system $F_{(A,B,C,D,E)}(x)$ introduced in (12).

At first glance, it looks like the new formulation (12) is worse than the one in (6). Indeed, the number of variables has increased from 6 to 15, and the Bézout number has increased from $6^5 = 7776$ to $2^{10}3^5 = 248832$. However, the new formulation is better suited for the numerical solver that powers our website. We explain this in the last section.

Approximation and Certification

Steiner's conic problem amounts to solving a system of polynomial equations. Two formulations were given in (6) and (12). But what does "solving" actually mean? One answer is suggested in the textbook by Cox, Little, and O'Shea [CLO15]: Solving means computing a *Gröbner basis* $\mathcal{G}$. Indeed, crucial invariants, such as the dimension and degree of the solution variety, are encoded in $\mathcal{G}$. The number of real solutions is found by applying techniques like deriving *Sturm sequences* from the polynomials in $\mathcal{G}$. Yet Gröbner bases can take a very long time to compute. We found them impractical for Steiner's problem.

Computing 3264 conics in a second requires numerical methods. Our encodings of the solutions are not Gröbner bases but *numerical approximations*. How does one make this rigorous? This question can be phrased as follows. Suppose $u_1, \ldots, u_6$ are the true coordinates of a solution and $u_1 + \Delta u_1, \ldots, u_6 + \Delta u_6$ are approximations of those complex numbers. How small must the entries of $\Delta u_1, \ldots, \Delta u_6$ be before it is justified to call them approximations? This question is elegantly circumvented by using Smale's definition of *approximate zero* [BCSS98, Definition 1 in §8].

In short, an approximate zero of a system $F(x)$ of n polynomials in n variables is any point $z \in \mathbb{C}^n$ such that Newton's method when applied to z converges quadratically fast towards a zero of F. Here is the precise definition.

Definition 3 (Approximate zero). Let $J_F(x)$ be the $n \times n$ Jacobian matrix of $F(x)$. A point $z \in \mathbb{C}^n$ is an *approximate zero* of F if there exists a zero $\zeta \in \mathbb{C}^n$ of F such that the

sequence of Newton iterates

$$z_{k+1} = N_F(z_k), \quad \text{where } N_F(x) = x - J_F(x)^{-1}F(x),$$

starting at $z_0 = z$, satisfies for all $k = 1, 2, 3, \ldots$ that

$$\|z_{k+1} - \zeta\| \leq \frac{1}{2}\|z_k - \zeta\|^2.$$

If this holds, then we call ζ the *associated zero* of z. Here $\|x\| := (\sum_{i=1}^n x_i\overline{x_i})^{\frac{1}{2}}$ is the standard norm in $\mathbb{C}^n$, and the zero ζ is assumed to be nonsingular; i.e., $\det(J_F(\zeta)) \neq 0$.

The reader should think of approximate zeros as a *data structure* for representing solutions to polynomial systems, just as a Gröbner basis is a data structure. Different types of representations of data provide different levels of accessibility to the desired information. For instance, approximate zeros are not well suited for computing algebraic features of an ideal. But they are a powerful tool for answering geometric questions in a fast and reliable manner.

Suppose that z is a point in $\mathbb{C}^n$ whose real and imaginary parts are rational numbers. How can we tell whether z is an approximate zero of F? This is not clear from the definition.

It is possible to certify that z is an approximate zero without dealing with the infinitely many Newton iterates. We next explain how this works. This involves Smale's γ-number and Smale's α-number:

$$\begin{aligned} \gamma(F,z) &= \sup_{k\geq 2} \left\| \frac{1}{k!} J_F(z)^{-1} D^k F(z) \right\|^{\frac{1}{k-1}}, \\ \alpha(F,z) &= \|J_F(z)^{-1}F(z)\| \cdot \gamma(F,z). \end{aligned}$$

Here $D^kF(z)$ denotes the tensor of order-k derivatives at the point z, the tensor $J_F(z)^{-1}D^kF(z)$ is understood as a multilinear map $A : (\mathbb{C}^n)^k \to \mathbb{C}^n$, and the norm of this map is $\|A\| := \max_{\|v\|=1} \|A(v, \ldots, v)\|$.

Shub and Smale [SS93] derived an upper bound for $\gamma(F,z)$ that can be computed exactly. Based on the next theorem [BCSS98, Theorem 4 in Chapter 8], one can thus decide algorithmically if z is an approximate zero, *using only data of the point z itself*.

Theorem 4 (Smale's α-theorem). *If $\alpha(F,z) < 0.03$, then z is an approximate zero of $F(x)$. Furthermore, if $y \in \mathbb{C}^n$ is any point with $\|y - z\| < (20\,\gamma(F,z))^{-1}$, then y is also an approximate zero of F with the same associated zero ζ as z.*

Actually, Smale's α-theorem is more general in the sense that $\alpha_0 = 0.03$ and $t_0 = 20$ can be replaced by any two positive numbers α_0 and t_0 that satisfy a certain list of inequalities.

Hauenstein and Sottile [HS12] use Theorem 4 in an algorithm, called `alphaCertified`, that decides if a point $z \in \mathbb{C}^n$ is an approximate zero and if two approximate zeros have distinct associated solutions. An implementation

```
Analyzing 3264 points using exact arithmetic.

Isolating 3264 approximate solutions.

Classifying 3264 distinct approximate solutions.

Rational certification results:

Number of points tested:                3264
Certified approximate solutions:        3264
Certified distinct solutions:           3264
Certified real distinct solutions:      3264
```

Figure 5. A proof for Proposition 1 given by the software `alphaCertified`.

is publicly available. Furthermore, if the polymomial system F has only real coefficients, then `alphaCertified` can decide if an associated zero is real. The idea behind this is as follows: Let $z \in \mathbb{C}^n$ be an approximate zero of F with associated zero ζ. If the coefficients of F are all real, then the Newton operator $N_F(x)$ from Definition 3 satisfies $N_F(\overline{x}) = \overline{N_F(x)}$. Hence $\overline{z}$ is an approximate zero of F with associated zero $\overline{\zeta}$. If $\|z - \overline{z}\| < (20\,\gamma(F,z))^{-1}$, then, by Theorem 4, the associated zeros of z and $\overline{z}$ are equal. This means $\zeta = \overline{\zeta}$.

A fundamental insight is that Theorem 4 allows us to certify candidates for approximate zeros regardless of how they were obtained. Typically, candidates are found by inexact computations using floating point arithmetic. We do not need to know what happens in that computation, because we can certify the result a posteriori. Certification constitutes a rigorous proof of a mathematical result. Let us see this in action.

Proof of Proposition 1. Fix five nondegenerate conics with rational coefficients listed after equation (9). We apply `HomotopyContinuation.jl` [BT18] to compute 3264 solutions in a second in 64-bit floating point arithmetic. The output is inexact. Each coefficient u_i of each true solution U is a complex number that is algebraic of degree 3264 over $\mathbb{Q}$. The floating point numbers that represent these coefficients are rational numbers, and we treat them as elements of $\mathbb{Q}$.

Our proof starts with the resulting list of 3264 vectors $x \in \mathbb{Q}^{15}$ corresponding to the 15 variables of (12). The computation was mentioned to make the exposition more friendly. It is *not* part of the proof.

We are now given 3264 candidates for approximate zeros of the polynomial system in (12). These candidates have rational coordinates. We use them as input to the software `alphaCertified` from [HS12]. That software performs *exact computations in rational arithmetic*. Its output shows that the 3264 vectors x are approximate zeros, that their associated zeros ζ are distinct, and that they all have real coordinates. This is shown in Figure 5. The

output data of `alphaCertified` is available for download through the arXiv version of this article. □

This was a rigorous proof of Proposition 1, just as trustworthy as a computer-assisted proof by symbolic computation (e.g. Gröbner bases and Sturm sequences) might have been. Readers who are experts in algebra should not get distracted by the appearance of floating point arithmetic: it is *not* part of the proof! Floating point numbers are only a tool for obtaining the 3264 candidates. The actual proof is carried out by exact symbolic computations.

How Does This Work?

In this section we discuss the methodology and software that powers the web interface (10).

We use the software `HomotopyContinuation.jl` that was developed by two of us [BT18]. This is a `Julia` [BEKV17] implementation of a computational paradigm called *homotopy continuation*. The reasons we chose `Julia` as the programming language are threefold: the first is that `Julia` is open source and free for anyone to use. The second is that `Julia` can be as fast as well-written `C`. For instance, we use `Julia`'s JIT compiler for fast evaluation of polynomials. Finally, the third reason is that, despite its high performance, `Julia` still provides an easy high-level syntax. This makes our software accessible for users from many backgrounds.

Homotopy continuation works as follows: We wish to find a zero in $\mathbb{C}^n$ of a system $F(x)$ of n polynomials in n variables. Let $G(x)$ be another such system with a known zero $G(\zeta) = 0$. We connect F and G in the space of polynomial systems by a path $t \mapsto H(x,t)$ with $H(x,0) = G(x)$ and $H(x,1) = F(x)$.

The aim is to approximately follow the *solution path* $x(t)$ defined by $H(x(t),t) = 0$. For this, the path is discretized into steps $t_0 = 0 < t_1 < \cdots < t_k = 1$. If the discretization is fine enough, then ζ is also an approximate zero of $H(x,t_1)$. Hence, by Definition 3, applying the Newton operator $N_{H(x,t_1)}(x)$ to ζ, we get a sequence $\zeta_0, \zeta_1, \zeta_2, \ldots$ of points that converges towards a zero ξ of $H(x,t_1)$. If $t_2 - t_1$ and $\|\zeta_i - \xi\|$ are small enough, for some $i \geq 0$, then the iterate ζ_i is an approximate zero of $H(x,t_2)$.

We may repeat the procedure for $H(x,t_2)$ and starting with ζ_i. Inductively, we find an approximate zero of $H(x,t_j)$ for all j. In the end, we obtain an approximate zero for the system $F(x) = H(x,1)$. Most implementations of homotopy continuation, including `Bertini` [BHSW06] and `HomotopyContinuation.jl` [BT18], use heuristics for setting both the step sizes $t_{j+1} - t_j$ and the number of Newton iterations.

Our homotopy for Steiner's conic problem computes zeros of the system $F_{(A,B,C,D,E)}(x)$ from (12). We prefer formulation (12) over (6), because the equations in the former formulation have lower degrees and fewer terms. It is known that high degrees and many terms introduce numerical instability in the evaluation of polynomials. We use the homotopy

$$H(x,t) = F_{t\cdot(A,B,C,D,E)+(1-t)\cdot(A',B',C',D',E')}(x). \tag{13}$$

The conic $tA + (1-t)A'$ is defined by the coefficients $ta_i + (1-t)a_i'$, where a_i and a_i' are the coefficients of A and A'. This is called a *parameter homotopy* in the literature. Geometrically, (13) is a straight line in the space of quintuples of conics. An alternative would have been

$$\tilde{H}(x,t) = tF_{(A,B,C,D,E)} + (1-t)F_{(A',B',C',D',E')}. \tag{14}$$

The advantage of (13) over (14) is that the path stays within the space of *structured systems*

$$\{F_{(A,B,C,D,E)}(x) \mid (A,B,C,D,E) \text{ are conics}\}.$$

The structure of the equations is preserved. The system in (13) has 3264 solutions for almost all t, whereas (14) has 7776 solutions for random t. In the language of algebraic geometry, we prefer the *flat family* (13) over (14), which is not flat.

The last missing piece in our *Steiner homotopy* is a start system. That is, we need five explicit conics together with all 3264 solutions. For this, we construct a generic instance (A,B,C,D,E) by randomly sampling complex coefficients for the conics. Then, we compute the 3264 solutions using standard homotopy continuation techniques [SW05]. Those 3264 solutions are saved and used for further computations. This initial computation is significantly more expensive than tracking 3264 solutions along the homotopy (13), but it only has to be done once.

In closing, we emphasize the important role played by enumerative geometry for solving polynomial systems. It gives a criterion for deciding if the initial numerical computation found the correct number of solutions. This is why numbers like 3264 are so important and why a numerical analyst might care about Chow rings and the pentagon construction. We conclude that enumerative algebraic geometry and numerical algebraic geometry complement each other.

References

[BKT08] Bashelor A, Ksir A, Traves W. Enumerative algebraic geometry of conics, *Amer. Math. Monthly* **115** (2008), 701–728. MR2456094

[BHSW06] Bates D, Hauenstein J, Sommese A, Wampler, C. Bertini: Software for Numerical Algebraic Geometry. Available at `bertini.nd.edu`.

[BEKV17] Bezanson J, Edelman A, Karpinski S, Shah V. Julia: A fresh approach to numerical computing, *SIAM Review* **59** (2017), 65–98. MR3605826

[BCSS98] Blum L, Cucker F, Shub M, Smale S. *Complexity and Real Computation*, Springer, 1998. MR1479636

[BT18] Breiding P, Timme S. *HomotopyContinuation.jl - a package for solving systems of polynomial equations in Julia*, Mathematical Software – ICMS 2018. Lecture Notes in Computer Science. Software available at `www.juliahomotopycontinuation.org`.
[CLO15] Cox D, Little J, O'Shea D. *Ideals, Varieties, and Algorithms*: *An Introduction to Computational Algebraic Geometry and Commutative Algebra, 4th ed.*, Undergraduate Texts in Mathematics, Springer, 2015. MR3330490
[EH16] Eisenbud D, Harris J. *3264 and All That – a Second Course in Algebraic Geometry*, Cambridge University Press, 2016. MR3617981
[GZ79] Garcia CB, Zangwill WL. Finding all solutions to polynomial systems and other systems of equations, *Math. Programming* **16** (1979), 159–176. MR527572
[HS12] Hauenstein J, Sottile F. alphaCertified: Certifying solutions to polynomial systems, *ACM Trans. Math. Software* **48** (2012), no. 4. MR2972672
[RTV97] Ronga F, Tognoli A, Vust T. The number of conics tangent to five given conics: the real case, *Rev. Mat. Univ. Complut. Madrid* **10** (1997), 391–421. MR1605670
[SS93] Shub M, Smale S. Complexity of Bézout's theorem I. Geometric aspects, *J. Amer. Math. Soc.* **6** (1993), 459–501. MR1175980
[Sot95] Sottile F. Enumerative geometry for real varieties, *Algebraic geometry – Santa Cruz 1995*, Proc. Sympos. Pure Math., 62, Part 1, Amer. Math. Soc., Providence, RI, 1997, 435–447. MR1492531
[Sot] Sottile F. 3264 real conics, `www.math.tamu.edu/~sottile/research/stories/3264/`.
[SW05] Sommese A, Wampler C. *The Numerical Solution of Systems of Polynomials. Arising in Engineering and Science*, World Scientific, 2005. MR2160078


Paul Breiding


Bernd Sturmfels


Sascha Timme

Credits

Figures 1 and 3 are courtesy of Thomas Endler.
Figures 2, 4, 5, and author photos are courtesy of the authors.