Milan j. math. 75 (2007), 225–271

1424-9286/010225-47, *published online* 14.11.2007
DOI 10.1007/s00032-007-0078-4

Milan Journal of Mathematics

Foci and Foliations of Real Algebraic Curves

Joel C. Langer and David A. Singer

Abstract. We discuss diverse results whose common thread is the notion of *focus* of an algebraic curve. In a unified setting, which combines elements of projective geometry, complex analysis and Riemann surface theory, we explain the roles of ordinary and singular foci in results on numerical ranges of matrices, quadrature domains, Schwarzian reflection, and other topics. We introduce the notion of *canonical foliation* of a real algebraic curve, which places foci into the context of continuous families of plane curves and provides a useful method of visualization of all relevant structures in a planar graphical image.

Keywords. Real algebraic curve, quadrature domain, Schwarzian reflection, canonical foliation.

1. Introduction

In the nineteenth century, foci of ellipses were generalized to foci of higher degree curves: *The* (*real*) *foci of a real algebraic curve* $\Gamma \subset \mathbb{C}P^2$ *are the real points on isotropic tangent lines to* Γ [22]. In traditional terminology associated with an algebraic curve $P(X, Y, Z) = 0$ in the complex projective plane $\mathbb{C}P^2 = \{[X, Y, Z]\}$, an *isotropic line* is one containing either of the *circular points* $c_\pm = [1, \pm i, 0]$ (the pair of points common to all circles). The above definition does not extend the best-loved properties of conic sections, involving distances or angles in the Euclidean plane, and the reader could hardly be expected to appreciate the general notion of foci until the relevant background is provided.

In the meantime, here are some facts about ellipses, viewed as curves in the complex plane $\gamma \subset \mathbb{C} = \{x + iy\}$, which generalize directly to higher degree curves:

- Let $\mathcal{T}$ be the triangle with vertices $z_1, z_2, z_3 \in \mathbb{C}$. Then the critical points of the complex polynomial $p(z) = (z - z_1)(z - z_2)(z - z_3)$ are located at the foci of the ellipse γ inscribed in $\mathcal{T}$, tangent at the midpoints of sides $\overline{z_i z_j}$ ([2], [25]; see [19, pp. 7–11]).
- Let $R_\gamma : \mathbb{C} \to \mathbb{C}$ be *Schwarzian reflection in the ellipse* γ, extended to a two-valued, antiholomorphic function (one branch of which fixes points of γ). The two branch points of R_γ are the foci of γ ([24, pp. 21, 85–87]).
- Let $F(A) = \{w^* A w : w \in \mathbb{C}^2, w^* w = 1\}$ be the *field of values* (*numerical range*) of a 2×2 complex matrix A; here, w^* is the conjugate transpose of w. Then $F(A)$ is bounded by an ellipse γ—non-degenerate when A is non-normal—and the eigenvalues of A are located at the foci of γ [17].
- Suppose γ is an ellipse whose interior contains the origin $0 \in \mathbb{C}$. Let γ^{-1} be the curve obtained from γ by inversion $x + iy \mapsto \frac{1}{x+iy}$ of each of its points. Then γ^{-1} bounds a *quadrature domain* $\mathcal{D}$; that is, there exist points $z_k = x_k + iy_k \in \mathcal{D}$ and constants $C_k, k = 1, 2$, such that the following (two-point) quadrature identity holds for functions $h(x, y)$ which are harmonic on the closure of $\mathcal{D}$: $\int\int_{\mathcal{D}} h dx dy = \sum_{k=1,2} C_k h(x_k, y_k)$ ([24, pp. 19–20]).

In their general form, such results represent diverse topics in pure and applied mathematics, involving some substantial applications of Riemann surfaces, classical algebraic geometry, and other techniques of analysis, with implications for approximation theory, numerical analysis, and fluid models, to name a few. By now, a number of these subjects are highly developed, and many of the results have been surveyed in books and expository articles (see [7, 8, 10, 11, 13, 19, 24, 29]).

Our purpose is not to recount any of these particular subjects in depth, but we will explain how a variety of results on foci may be understood by direct and elementary means. We thus intend to clarify, and provide a unified context for, a number of points which are not well covered, if at all, in more specialized treatments. For instance, consider the relationship between numerical ranges and quadrature domains suggested by the last two facts on ellipses:

Theorem 1.1. *Let $\mathcal{N} \subset \mathbb{C}$ be the numerical range of an $n \times n$ complex matrix A. For simplicity, assume the generic case: All eigenvalues of A lie in the interior of $\mathcal{N}$, and the boundary of $\mathcal{N}$ is a nonsingular, algebraic curve γ. Then γ^{-1}, obtained by inversion as above, bounds a quadrature domain $\mathcal{D}$.*

This result will follow from our discussion as a simple observation, but is related to some interesting subtleties regarding the general notion of foci. We note that complex inversion $z \mapsto 1/z$ takes foci of a generic algebraic curve γ to foci of γ^{-1} (another "obscure" fact for which we will provide a simple proof). In the above example, the resulting foci of γ^{-1} thus lie outside of $\mathcal{D}$, and the points z_k might seem to have nothing at all to do with foci! (Which would explain the absence of the term *foci* in the literature on quadrature domains.)

As it turns out, z_1 and z_2 are *singular foci* of γ^{-1} (classical nomenclature), resulting from isotropic tangency at one of the circular points $c_\pm$, both of which ideal points lie on Γ^{-1}. Here we return to the setting of algebraic curves $\Gamma \subset \mathbb{C}P^2$, we identify the complex plane $\mathbb{C}$ with the real plane $\mathbb{R}^2 \subset \mathbb{C}P^2$, and we sometimes refer to Γ and its real part $\gamma = \Gamma \cap \mathbb{R}^2 \subset \mathbb{C}$ interchangeably. Then complex inversion $z \mapsto 1/z$ extends to (most of) $\mathbb{C}P^2$, where it may be understood as a Cremona transformation which turns generic curves Γ into *circular curves* Γ^{-1} (i.e., curves containing $c_\pm$), taking (ordinary) foci to (ordinary) foci, while also producing the singular foci "from nowhere." We remark that, for many purposes, ideal points on an algebraic curve $\Gamma \subset \mathbb{C}P^2$ should be regarded just like other points on Γ. But the circular points and the real plane play distinguished roles in the geometry of foci, whose underlying symmetries are consequently not all projective transformations but *real similarities*. To summarize, thus far: *Singular foci are like ordinary foci . . . except when they're not!*

The embedding $\mathbb{C} \subset \mathbb{C}P^2$ also brings analytic function theory into the foreground, as a convenient handle on many other elements of the subject. *Isotropic projections* onto the real plane (from c_+ and c_-) determine meromorphic functions $\rho, \beta : \Gamma \to \mathbb{C} \cup \{\infty\}$, yielding an algebraic function $S(z) = \beta \circ \rho^{-1}(z)$, the *Schwarz function of* γ. The branch points and finite poles of $S(z)$ are, respectively, ordinary and singular foci of γ; in special cases, application of the residue theorem to $S(z)$ thus immediately produces a quadrature identity. Schwarzian reflection in γ is given by $R_\gamma(z) = \overline{S(z)}$. The zeros of $S'(z)$ are the *defoci* of γ—our name for the mirror images of foci under R_γ.

In this connection, defoci seem to have been generally ignored, even though they are merely the "other projections" of points of isotropic tangency $a \in \Gamma$. That is, if a is a point of higher multiplicity for ρ, then $\rho(a)$ is a focus and $\beta(a)$ is a defocus; because of the underlying involution (real structure) on Γ, the roles of ρ, β may also be interchanged here. In general, it is useful to make a clear distinction between features of the embedded algebraic curve $\Gamma \subset \mathbb{C}P^2$ itself (isotropic points, ideal points, etc), and byproducts of projection onto $\mathbb{C}$ (foci *vs.* defoci, etc.).

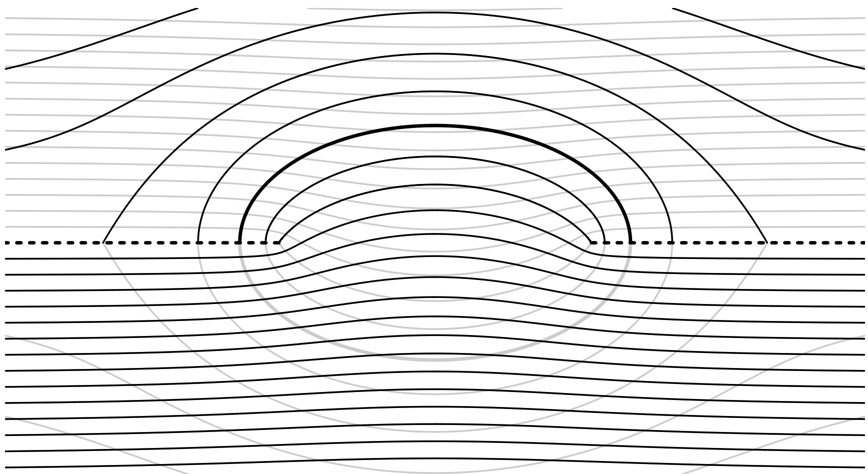

FIGURE 1.1. Canonical foliation of the ellipse $\frac{x^2}{25} + \frac{y^2}{9} = 1$.

The previous paragraph leads us from *foci*, right to the doorstep of our other main topic: *The canonical foliation of a real algebraic curve.* Here we observe—and this is undoubtedly the main new element we introduce into the established circle of ideas—that there is a special quadratic differential on $\Gamma \subset \mathbb{C}P^2$ whose features mesh perfectly with the preceding discussion. Namely, the *fundamental form* $Q = dX^2 + dY^2$ on $\mathbb{C}^2 \simeq \{[X, Y, 1]\}$ pulls back by inclusion to a holomorphic quadratic differential $q = \iota^*Q = d\rho d\beta$ on $\Gamma \cap \mathbb{C}^2$ and extends meromorphically to Γ. With respect to $\mathbb{C}$, one has the local representation $q = S'(z)dz^2$, valid away from foci. One finds that q has poles at the ideal points of Γ and zeroes at finite isotropic points of Γ, while $\gamma = \Gamma \cap \mathbb{R}^2$ itself will be seen as a leaf (or union of finitely many leaves) in a foliation determined by q, which "interpolates" between the latter curve(s) and the former points. The canonical foliation is the result

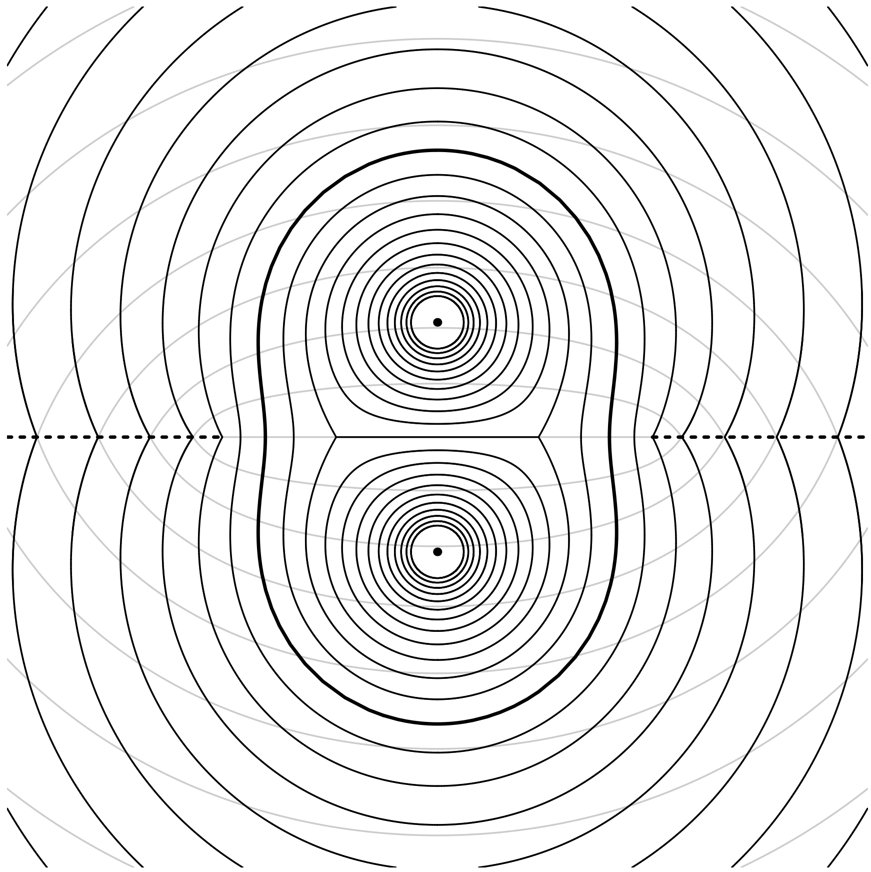

FIGURE 1.2. Canonical foliation of Neumann's quadrature domain.

of applying to q a standard construction for quadratic differentials, to be described shortly.

Isotropic projection of the canonical foliation onto $\mathbb{C}$ shows the distinguished roles of foci, defoci, and ideal points. Though one traditionally graphs a real algebraic curve $\Gamma \subset \mathbb{C}P^2$ in the real plane—where only the real part γ is "visible"—such a projected foliation thus provides a very natural planar visualization of the whole of Γ. For example, when γ is the circle $x^2+y^2=1$, one obtains the family of concentric circles $|z|=e^{-t}$; the only (finite) singularity is the singular focus at $z=0$, showing the standard behavior near the second order pole of $q=-dz^2/z^2$. Canonical foliations of other algebraic curves (not circles or lines) will be visualized in multiple sheets covering $\mathbb{C}$, glued together along branch cuts terminating in foci. Figure 1.1 shows the simplest such example, the canonical foliation of the ellipse $x^2/25 + y^2/9 = 1$; see also Figure 13.1. It consists of two sheets—a black one and a gray one—and shows two foci joined by the branch cut, and two defoci exhibiting the characteristic *triradiate* at a first order zero of q; poles of q lie out of view at (both copies of) $z = \infty$. The corresponding picture for the inverted ellipse, Figure 1.2, combines all three elements—foci, defoci, and singular foci. As far as we know, our images of these simple algebraic curves (which required surprisingly careful *Mathematica* computations) are the first of their kind to appear in print.

We attempt now a concise explanation of such figures in terms of the quadratic differential q. First, recall that any meromorphic quadratic differential $\varphi(z)dz^2$ on a Riemann surface determines a singular flat geometry on the surface, and singular foliation by geodesic arcs. A local description of the construction near a regular point z_0 of $\varphi(z)dz^2$ may be given as follows. For z near z_0, one defines an analytic function $\Phi(z) = \int_{z_0}^{z} \sqrt{\varphi(\zeta)}d\zeta$ by picking one of the two branches of $\sqrt{\varphi(z)}$; the local diffeomorphism Φ defines the *natural parameter* $w = \Phi(z)$ near z_0. The flat metric $g = |\varphi(z)|dzd\bar{z}$ is the pull-back of the standard Euclidean metric on the w-plane by Φ, while the foliation of the w-plane by horizontal lines $Im[w] = t_0$ pulls back to a local geodesic foliation by *horizontal arcs* $\Phi^{-1}(s + it_0)$. We refer the reader to [27] for a detailed discussion of behavior near a singular point of $\varphi(z)dz^2$ and for global issues.

While any such quadratic differential carries information about the abstract Riemann surface, the choice $q = d\rho d\beta$ directly reflects the embedding $\Gamma \subset \mathbb{C}P^2$; the canonical foliation determined by q embeds $\gamma = \Gamma \cap \mathbb{R}^2$ in a geodesic foliation of $(\Gamma, |q|)$ with ideal points and isotropic points removed. Further, γ plays a distinguished role in this foliation, as explained by one additional observation about the natural parameter for $q = d\rho d\beta$. In general, for fixed t_0, $s \mapsto \Phi^{-1}(s + it_0)$ parametrizes a given horizontal arc by unit speed with respect to the metric $g = |\varphi(z)|dzd\bar{z}$. Since Q agrees with the standard Euclidean metric on the real plane, it follows that the natural parameter for q agrees with the usual arclength parameter for the plane curve $\gamma \subset \mathbb{C}$. To be more specific, let us take z_0 to be a point on γ, for convenience, and let $\Phi(z) = \int_{z_0}^{z} \sqrt{S'(\zeta)}d\zeta$. Then a parametrization of γ by Euclidean arclength may be defined by $s \mapsto \gamma(s) = \Phi^{-1}(s)$—not so for the other trajectories $s \mapsto \Phi^{-1}(s + it_0)$!

To globalize the above constructions one may consider the multivalued "complex arclength function" $\Phi(p) = \int_{p_0}^{p} \sqrt{q}$, defined via continuation of $\sqrt{q}$ along paths of integration joining points $p_0, p \in \Gamma$, avoiding the singular set of q. By uniqueness of analytic continuation, the forgoing remarks may then be given a simple interpretation, which we summarize as an informal theorem:

Theorem 1.2. *The isotropic points and ideal points of a real algebraic curve $\Gamma \subset \mathbb{C}P^2$ are the singularities of analytic continuation of the arclength function of $\gamma = \Gamma \cap \mathbb{R}^2$. Analytic continuation of the corresponding arclength parametrization $\gamma(s)$ yields a parametrization $\Gamma(s + it)$ of the canonical foliation of Γ.*

By considering properties of isotropic projection on an algebraic curve and invoking the normal form theory of quadratic differentials (see [27]), one could discuss the types of singularities thus arising in the arclength function; generically, isotropic points yield branch points (see the ellipse) and ideal points give simple poles, but circular points give logarithmic singularities (see the circle). A full discussion of the multivalued arclength function and its "inverse" $\Gamma(s + it)$ would incorporate such local data, along with global analytic continuation arguments, etc.

In lieu of such a discussion, we will develop a more local result, below, which can be visualized as a process in the complex plane relating an algebraic curve $\gamma \subset \mathbb{C}$ directly to its foci and defoci. Namely, we will consider non-singular analytic mappings $f : A \to \mathbb{C}$ defined on annuli, extending unit speed parametrizations of algebraic plane curves $\gamma = f(S^1)$; for maximal extensions of such mappings, the *nearest* foci and defoci of γ will be seen to lie in the closure $\overline{f(A)}$. We will also give a dynamical/geometrical interpretation to the process of filling out the region $f(A)$ by a family of analytic curves γ_t, $-\tau < t < \tau$. Since it was one of our original motivations for exploring the present topic, we conclude the introduction with a brief explanation of this interpretation.

As developed in [3], [4], an *infinite dimensional geometry of analytic curves* may be based on a formal symmetric space multiplication given by Schwarzian reflection of one analytic curve in another. The abstract theory of symmetric spaces then leads to the following geodesic equation: $\frac{\partial^2 S}{\partial t^2} - (\frac{\partial S}{\partial t} / \frac{\partial S}{\partial z}) \frac{\partial^2 S}{\partial t \partial z} = 0$. This second order partial differential equation for a time-dependent Schwarz function $S(t, z)$ describes a planar curve $\gamma_t \subset \mathbb{C}$ whose evolution may be regarded as a continuous limit of iterated Schwarzian reflection generated by a "colliding" pair of curves.

This "continuous reflection" process (defined only for a finite time) is uniquely determined by a curve's initial position $\gamma_0 = \overline{S(0,\gamma_0)}$ and suitable normal velocity field $\dot\gamma_0 = -\frac{1}{2}\frac{\partial S}{\partial t}(0,\gamma_0)/\frac{\partial S}{\partial z}(0,\gamma_0)$. In this context, our canonical foliations may be interpreted locally as a special class of solutions to the above geodesic equation, namely, those generated by taking as initial velocity the unit normal vector $\dot\gamma_0 = N = i/\sqrt{S'}$ along an algebraic curve $\gamma_0 = \gamma$. The examples of *canonical geodesics* thus obtained complement the local classification of geodesics given in [4], where the unit circle was fixed as initial curve $\gamma_0 = S^1$, but the (analytic) normal velocity field $\dot\gamma_0 = fN$ was arbitrary.

In the case of the ellipse as initial curve γ_0, the geodesic interpretation is best seen using a different branch cut from the one just considered (where the upper and lower halves of the ellipse appear on different sheets). Figure 13.1 shows how the initial ellipse $\gamma_0 \subset \mathbb{C}$ moves forward and backward in time to fill out a singular ring domain bounded by four of the singular trajectories of q. In the outward direction, the motion continues until γ_t encounters the pair of defoci; reversing time, γ_t moves inward for the same time interval until the foci are reached. Over larger time intervals the process would appear to require an interpretation allowing multi-valuedness.

2. Generalities on Projective Space

We will be studying the real projective plane, the complex projective plane, and the complex projective line. To see them all in a single context, we first establish some helpful notation. Real coordinates will be denoted by small letters, complex coordinates by capital letters. The complexification of a variable will be represented by capitalization.

$\mathbb{R}^3 = \{(x,y,z)|x,y,z \in \mathbb{R}\} \subset \mathbb{C}^3 = \{(X,Y,Z)|X,Y,Z \in \mathbb{C}\}$

$x = \Re(X), x' = \Im(X)$, etc.

Points in complex projective space $\mathbb{C}P^2$ are complex lines through the origin in $\mathbb{C}^3$; they have homogeneous coordinates $[X,Y,Z]$, where (X,Y,Z) are the coordinates of any non-zero point on the line. A line in $\mathbb{C}P^2$ is given by an equation of the form $AX + BY + CZ = 0$; the coordinates of this line will be denoted $[A\ B\ C]$. The space of lines in $\mathbb{C}P^2$ is the dual space $\mathbb{C}P^{2*}$.

A point in $\mathbb{C}^2$ with (complex) coordinates (X, Y) may be thought of as a point in $\mathbb{R}^4$ whose real coordinates are (x, y, x', y'). It may also be viewed as a point in $\mathbb{C}P^2$ with homogeneous coordinates $[X, Y, 1]$. Another way of representing this point is with *isotropic* coordinates

$$(R, B), \quad R = X + iY, \quad B = X - iY.$$

A point (X, Y) in $\mathbb{C}^2$ is real if and only if in isotropic coordinates $B = \overline{R}$. (For, if $\overline{X} - i\overline{Y} = X - iY$, then $\overline{X} - X = i(\overline{Y} - Y)$. The left hand side is imaginary and the right hand side is real.)

There are two natural embeddings (the "red" one and the "blue" one) of the extended complex numbers $\mathbb{C}^*$ (that is, the Riemann sphere, which we will avoid calling the "complex plane" to avoid obvious confusions) into $\mathbb{C}P^{2*}$([24, page 56]). If $Z_0 = x_0 + iy_0$ is a complex number, then the equation $X + iY - Z_0 Z = 0$ represents a complex line whose homogeneous coordinates are $[1\ i\ Z_0]$. This is the unique line in $\mathbb{C}P^{2*}$ which passes through the *circular point* $[1, i, 0]$ and the real point $[x_0, y_0, 1]$. This gives the "red" embedding

$$R : \mathbb{C} \longrightarrow \mathbb{C}P^{2*},\ R[Z_0] = [1\ i\ \ -Z_0].$$

Similarly, the "blue" embedding is given by

$$B : \mathbb{C} \longrightarrow \mathbb{C}P^{2*},\ R[Z_0] = [1\ \ -i\ \ -\overline{Z_0}].$$

Both maps are extended to the Riemann sphere by taking the point at infinity to the ideal line $[0\ 0\ 1]$, which is the unique line through the two circular points $[1, i, 0]$ and $[1, -i, 0]$.

If H is a holomorphic function on some domain Ω in $\mathbb{C}$, then we can extend it to a map from (a part of) $\mathbb{C}^2$ to itself by the formula

$$TH : \mathbb{C} \times \mathbb{C} \longrightarrow \mathbb{C} \times \mathbb{C}, \quad TH(R, B) = (H(R), \overline{H}(B))$$

where the holomorphic function $\overline{H}$ is defined by $\overline{H}(Z) = \overline{H(\overline{Z})}$.

3. Circular points and geometry

The circular points $[1, \pm i, 0]$ described in the last section will play a vital role in what follows. They are so named because of the property that they lie on all circles. That is, they distinguish circles among conics. In general, a line through the origin is parallel to an asymptote of a conic if it is a linear factor of the polynomial given by the second order terms in x and y. Thus a hyperbola has two real asymptotes and passes through two real ideal points, while an ellipse passes through two imaginary ideal points. In the case of a circle, the asymptotes are the *isotropic lines*

$$X \pm iY = 0,$$

and we conclude that all circles pass through the circular points. We may interpret this group-theoretically, as follows.

The group of projective transformations of $\mathbb{C}P^2$ consists of (equivalence classes of) three-by-three invertible matrices with complex entries, where A is equivalent to ρA for $\rho \neq 0$. If $A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix}$, then the condition that A preserves both circular points is that $\begin{pmatrix} a_1 + b_1 i \\ a_2 + b_2 i \\ a_3 + b_3 i \end{pmatrix} = \begin{pmatrix} u \\ iu \\ 0 \end{pmatrix}$ and $\begin{pmatrix} a_1 - b_1 i \\ a_2 - b_2 i \\ a_3 - b_3 i \end{pmatrix} = \begin{pmatrix} v \\ -iv \\ 0 \end{pmatrix}$.

It is immediate that $a_3 = 0 = b_3$. Furthermore, since $a_2 + b_2 i = -b_1 + a_1 i$ and $a_2 - b_2 i = -b_1 - a_1 i$, it follows that $a_1 = b_2$ and $a_2 = -b_1$. Since A is invertible, we may define normalize $c_3 = 1$ and rewrite the matrix as

$$A = \begin{pmatrix} a_1 & -a_2 & c_1 \\ a_2 & a_1 & c_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is the (matrix) group of projective transformations preserving the circular points. Assume further that real points go to real points. The origin $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$ goes to $\begin{pmatrix} c_1 \\ c_2 \\ 1 \end{pmatrix}$, which is therefore a real column vector. The point $\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ goes to $\begin{pmatrix} a_1 + c_1 \\ a_2 + c_2 \\ 1 \end{pmatrix}$, so a_1 and a_2 must also be real. This shows that the matrix is real. Writing $a_1 = r\cos\theta$ and $a_2 = r\sin\theta$, where $r > 0$, we have

$$A = \begin{pmatrix} r\cos\theta & -r\sin\theta & c_1 \\ r\sin\theta & r\cos\theta & c_2 \\ 0 & 0 & 1 \end{pmatrix}.$$

This is exactly the matrix representation of the group of real similarities.

To summarize,

Theorem 3.1. *The subgroup of the projective group consisting of real similarities is characterized as those transformations which preserve the real projective plane and the circular points.*

4. Analytic curves

A real-analytic curve in the complex plane may be represented in several ways. Here we give three such representations and relate them.

1. If Γ is a parametrized real-analytic curve in $\mathbb{C}$, it is given by a pair of real-analytic functions $x(t)$ and $y(t)$:
$$\Gamma = \{Z(t) = x(t) + iy(t) | a \leq t \leq b\}.$$
We can extend these functions to holomorphic functions to get a (complex) curve in $\mathbb{C}^2$ given by
$$\mathbf{\Gamma} = \{(R(T), B(T)) = (X(T) + iY(T), X(T) - iY(T)) | \\ a \leq \Re T \leq b, |\Im T| \leq \epsilon > 0\}.$$
2. Let $\Phi : \mathbb{R}^2 \longrightarrow \mathbb{R}$ be a real-analytic function of two variables. This defines a curve in $\mathbb{R}^2$ by $\Gamma = \{(x, y) | \Phi(x, y) = 0\}$, assuming 0 is a regular value of Φ. We may extend this to the complex locus $\mathbf{\Gamma} = \{(X, Y) | \Phi(X, Y) = 0\}$.
3. Suppose an analytic arc Γ in $\mathbb{C}$ is defined by the equation:
$$\overline{Z} = S(Z)$$
where $S(Z)$ is a complex analytic function (the *Schwarz function* of Γ). Let $CC : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$ be defined by $CC(R, B) = (B, R)$. Equivalently, $CC(X, Y) = (X, -Y)$. Then we can extend the curve Γ by the equation
$$CC(R, B) = TS(R, B) \text{ or } B = S(R), R = \overline{S}(B).$$
The last two equations are redundant. For if $X - iY = S(X + iY)$, then $\overline{X} + i\overline{Y} = \overline{S(X + iY)} = \overline{S}(\overline{X} - i\overline{Y})$. Then if $(R, B) = (\overline{X} + i\overline{Y}, \overline{X} - i\overline{Y})$, we have $R = \overline{S}(B)$.

The first and third representations regard the curve as lying in $\mathbb{C}$. To extend these curves to $\mathbb{C}^2$ we use the embedding given by $Z \longmapsto (Z, \overline{Z})$. Equivalently, $(x + iy)$ is taken to (x, y).

Example 4.1. The unit circle.
The circle may be represented as the rational parametrized curve

$$Z(t) = \frac{1-t^2}{1+t^2} + i\frac{2t}{1+t^2} = \frac{(1+it)^2}{1+t^2} = \frac{1+it}{1-it}$$

or as the solutions to

$$0 = \Phi(x,y) = x^2 + y^2 - 1$$

or as the locus of the equation

$$\overline{z} = \frac{1}{z}.$$

The first representation leads to the complex curve

$$\mathbf{\Gamma}: \quad (R(T), B(T)) = (\frac{1+iT}{1-iT}, \frac{1-iT}{1+iT}).$$

This curve satisfies $1 \equiv RB = X^2 + Y^2$.

The second representation gives:

$$\mathbf{\Gamma}: \quad X^2 + Y^2 - 1 = 0.$$

The third representation leads to:

$$\mathbf{\Gamma}: \quad (B, R) = (\frac{1}{R}, \frac{1}{B}).$$

That is,

$$(X - iY) = \frac{1}{(X + iY)} \text{ or } X^2 + Y^2 = 1.$$

5. Study's Construction of Schwarz reflection, and Foci

Let Γ be a curve in $\mathbb{R}^2$ given by the algebraic equation $\phi(x, y) = 0$. Study defines reflection of points across Γ as follows: If $P = (x_1, y_1)$ is any point in the plane, we identify it with the point in $\mathbb{C}P^2$ with homogeneous coordinates $[x_1, y_1, 1]$. There are two lines, R_1 and B_1 (a "red" one and a "blue" one), given in homogeneous coordinates $[X, Y, Z]$ by

$$R_1 : X + iY = (x_1 + iy_1)Z \qquad B_1 : X - iY = (x_1 - iy_1)Z.$$

These are the *isotropic lines* through P, which pass through P and the circular points $[1, \pm i, 0]$.

Let (X_2, Y_2) be a point in $\mathbb{C}^2$ lying on the line R_1 and satisfying the equation $\phi(X_2, Y_2) = 0$; such a point lies on the complex extension of Γ. Generically, there will be n choices of such a point if Γ is an algebraic curve of order n. This reveals the fact that Schwarz reflection is only locally defined in a $1-1$ manner. Now take the line

$$B_2 : X - iY = (X_2 - iY_2)Z.$$

Let $Q = (x_3, y_3)$ be the point of intersection of B_2 with the real plane. $Q = \mathcal{R}(P)$ is the Schwarz reflection of P.

Example 5.1. Let $c > 0, \rho_0 > 1$, $a = c(\rho_0 + \rho_0^{-1})/2$, and $b = c(\rho_0 - \rho_0^{-1})/2$. The curve Γ is the ellipse

$$f(x, y) = \frac{x^2}{a^2} + \frac{y^2}{b^2} - 1 = 0.$$

We compute the reflection of a point

$$(x_1, y_1) = (c(\rho + \rho^{-1}) \cos\theta/2, c(\rho - \rho^{-1}) \sin\theta/2).$$

For fixed ρ and varying θ this point sweeps out one of the confocal family of ellipses; we want to see that reflection takes each such ellipse to another such ellipse.

We have to solve the simultaneous equations:

$$X_2 + iY_2 = x_1 + iy_1 = Z_1 \qquad b^2X_2^2 + a^2Y_2^2 = a^2b^2.$$

Eliminating X_2 yields the quadratic equation

$$a^2b^2 = b^2Z_1^2 - 2ib^2Z_1Y_2 - b^2Y_2^2 + a^2Y_2^2 = b_2Z_1^2 - 2ib^2Z_1Y_2 + c^2Y_2^2.$$

This generically has two solutions:

$$Y_2 = i\frac{b^2}{c^2}Z_1 \pm \frac{ab}{c^2}\sqrt{c^2 - Z_1^2} \tag{5.1}$$

$$X_2 = Z_1 - iY_2 = \frac{a^2}{c^2}Z_1 \mp \frac{abi}{c^2}\sqrt{c^2 - Z_1^2}. \tag{5.2}$$

Combining these, we get

$$x_3 - iy_3 = X_2 - iY_2 = \frac{a^2 + b^2}{c^2}Z_1 \mp \frac{2ab}{c^2}\sqrt{Z_1^2 - c^2} \tag{5.3}$$

$$Z_1^2 = c^2\left(\frac{\rho^2 + \rho^{-2}}{4}\cos 2\theta + \frac{1}{2} + i\frac{\rho^2 - \rho^{-2}}{4}\sin 2\theta\right).$$

Plugging into equation (5.3) yields

$$x_3 - iy_3 = \frac{a^2 + b^2}{c}\left(\frac{\rho + \rho^{-1}}{2}\cos\theta + i\frac{\rho - \rho^{-1}}{2}\sin\theta\right) \mp \frac{2ab}{c}\left(\frac{\rho - \rho^{-1}}{2}\cos\theta + i\frac{\rho + \rho^{-1}}{2}\sin\theta\right) \tag{5.4}$$

$$x_3 - iy_3 = \frac{(a \mp b)^2\rho + (a \pm b)^2\rho^{-1}}{2c}\cos\theta + i\frac{(a \mp b)^2\rho - (a \pm b)^2\rho^{-1}}{2c}\sin\theta. \tag{5.5}$$

Now using the definition of a and b this gives

$$x_3 - iy_3 = c\left(\frac{\rho_0^{\mp 2}\rho + \rho_0^{\pm 2}\rho^{-1}}{2}\cos\theta + i\frac{\rho_0^{\mp 2}\rho - \rho_0^{\pm 2}\rho^{-1}}{2}\sin\theta\right). \tag{5.6}$$

The images of the point with coordinates (ρ, θ) are the points with coordinates and $(\rho_0^2\rho, \theta)$ and $(\max\{\frac{\rho_0^2}{\rho}, \frac{\rho}{\rho_0^2}\}, \theta)$. These points are always distinct if $\sin\theta \neq 0$. If $\theta = 0 \mod \pi$, then these points will coincide if and only if $\rho_0^2\rho = \frac{\rho_0^2}{\rho}$, i.e., when $\rho = 1$. In that case $x_1 + iy_1 = \pm c$; this gives the foci of the ellipse.

This example suggests the idea of a *focus* of a real analytic curve Γ as a point of ramification of the mapping defined by Schwarzian reflection in Γ. More precisely, we adopt the following

Definition 5.2. An isotropic tangent line to a real analytic curve Γ meets the real plane in a *focus* of Γ; the point of tangency $(u, v) \in \Gamma \subset \mathbb{C}^2$ itself is an *isotropic point* of Γ. The focus is *ordinary* if the point of tangency is not the circular point itself; otherwise it is *singular*.

The number of foci of an algebraic curve is related to its *class*, which we now describe. Given an algebraic curve Γ defined by a homogeneous polynomial of degree n by the equation $P(x, y, z) = 0$, there is another algebraic curve Γ^* defined by a homogeneous polynomial of degree m, $\phi(u, v, w) = 0$, called the *dual curve*, which is defined by the condition that $\phi(u, v, w) = 0$ if and only if the line $ux + vy + wz = 0$ is tangent to Γ. The number m is defined to be the class of Γ. It is well-known that $\Gamma^{**} = \Gamma$; thus the dual curve has degree m and class n. For a curve with δ nodes and κ cusps (and no more complicated singularities), one of Plücker's equations is

$$m = n(n-1) - 2\delta - 3\kappa. \tag{5.7}$$

Thus, for example, a nonsingular conic has degree 2 and class 2; a nonsingular cubic curve has degree 3 and class 6. (It follows, however, that the dual of a nonsingular curve of degree greater than 2 always has singularities.)

Given any point X not lying on a curve Γ of class m there will be exactly m lines through X (counting multiplicity) tangent to Γ. In particular, letting X be a circular point, we see that a curve of class m will have (at most) m foci.

Remark 5.3. Early references on foci, e.g. [22], do not distinguish between ordinary and singular foci, while later references, e.g. [6], [14], do. Theorems such as Theorem 6.1 require singular foci to count as foci, while Theorem 8.1 fails for singular foci. To be consistent with classical terminology, we would properly speak of *real* foci, a general "focus" being any point in $\mathbb{C}^2$ obtained as the intersection of a pair of red and blue isotropic tangent lines to Γ. While a *curve of class* m typically has m^2 such "foci", there are at most m foci in the present sense of the term. (See [14, Chapter V]) When Γ happens to be *circular*—meaning it contains the circular points—it will be necessary to consider singular foci, though they will be seen to have very different properties from ordinary foci.

6. Foci and numerical range of a matrix

While the notion of foci of an algebraic curve has been around since the nineteenth century as a generalization of the foci of a conic, it remains an intriguingly elusive concept. We saw in the last section that foci can be viewed as the branch points of Schwarz reflection; in this section we describe another important way in which they arise.

Let A be an $n \times n$ matrix with complex coefficients. The *numerical range* of A (also called the *field of values*) is the subset $F(A)$ of the complex plane defined by

$$F(A) = \{w^*Aw | w \in \mathbb{C}^n, w^*w = 1\} = \left\{ \frac{w^*Aw}{w^*w} | w \neq 0 \right\}$$

where * is the transpose conjugate.

Theorem 6.1. *$F(A)$ is a compact convex set containing the eigenvalues of A. There is an algebraic curve Γ of class n whose foci are the eigenvalues of A, such that $F(A)$ is the convex hull of Γ.*

Proof. The fact that $F(A)$ is compact and convex is the Toeplitz-Hausdorff Theorem [12],[28]. Kippenhahn ([17]; see also [21]) showed that $F(A)$ is the convex hull of an algebraic curve whose foci are the eigenvalues of A. We demonstrate this fact below.

Define the Hermitian matrices H and K by $H = \frac{A+A^*}{2}$ and $K = \frac{A-A^*}{2i}$; then $A = H + iK$. H and K are called, respectively, the real and imaginary parts of A. (But note that they are complex matrices in general!) It follows immediately that for any vector w, $w^*Aw = w^*Hw + iw^*Kw$ is the decomposition of w^*Aw into its real and imaginary parts. Thus projection of $F(A)$ to the real axis takes $F(A)$ to $F(H) = [\lambda_{min}, \lambda_{max}]$, which is the interval bounded by the minimum and maximum eigenvalues of H. The lines $x = \lambda_{min}$ and $x = \lambda_{max}$ are support lines to $F(A)$.

We may rotate $F(A)$ in the plane using the relation $F(e^{-i\theta}A) = e^{-i\theta}F(A)$. Since

$$e^{-i\theta}A = (\cos\theta H + \sin\theta K) + i(-\sin\theta H + \cos\theta K) = H_\theta + iK_\theta$$

gives the real and imaginary parts of $e^{-i\theta}A$, the support lines of $F(A)$ are given by

$$\cos\theta x + \sin\theta y = \lambda(\theta)$$

for $\lambda(\theta)$ the maximum eigenvalue of H_θ.

The polynomial equation

$$f(u, v, w) = \det(uH + vK - wI) = 0$$

defines a curve Γ of class n as the envelope of lines $ux + vy - wz = 0$. In particular, $f(\cos\theta, \sin\theta, \lambda(\theta)) = \det(H_\theta - \lambda(\theta)I) = 0$. This shows that the boundary curve of $F(A)$ includes (part of) the algebraic curve Γ.

The real foci of Γ are the points $w_0 = x_0 + iy_0$ (x_0 and y_0 real) such that the isotropic lines through $(x_0, y_0, 1)$ are tangent to Γ. The isotropic lines are given by the equations $x + iy - w_0 = 0$ and $x - iy - \overline{w_0} = 0$; that is, they are the lines given by the coordinates $[u, v, w] = [1, i, w_0]$ and $[u, v, w] = [1, -i, \overline{w_0}]$.

Since $f(1, i, w) = \det(H + iK - wI) = det(A - wI) = 0$ if and only if w is an eigenvalue of A, it follows that the eigenvalues of A are precisely the foci of Γ. (Note that $\det(H - iK - \overline{w_0}I) = \det(A^* - \overline{w_0}I) = 0$ also holds). □

When A is decomposable, i.e., when there are orthogonal invariant subspaces, the boundary of $F(A)$ is the convex hull of the union of numerical ranges corresponding to invariant subspaces. Even when it is indecomposable the algebraic curve defining the boundary may fail to be convex, in which case the boundary of $F(A)$ will have flat pieces. When the boundary curve Γ of $F(A)$ is a smooth closed curve in the complex plane, then the Schwarz function has its branch points (i.e. the foci) in the interior.

7. Dual curves and the RZ property

The boundary Γ of the numerical range of an $n \times n$ matrix A is the convex hull of the curve whose equation in line coordinates (u, v, w) is given by

$$\Phi(u, v, w) = \det(uH + vK + wI) = 0$$

where $A = H + iK$ and H and K are Hermitian matrices. Thus, Γ is the dual curve of the curve $\Phi(u, v, w) = 0$. What homogeneous polynomials Φ arise in this way? This is the content of the Lax conjecture, which was resolved by the work of Helton and Vinnikov [13], as observed by Lewis, Parrilo, and Ramana [18].

If u and v are real, then the roots of $p(w) = \det(uH + vK + wI) = 0$ are the eigenvalues of a Hermitian matrix, and therefore they are real. This implies a result of Kippenhahn, that in every direction in the plane there are n real tangents to the curve Γ. For the dual curve the corresponding statement is that every real (projective) line through the origin meets the curve in n real points; this is known as the *RZ condition.*[13] When n is even the curve consists of $\frac{n}{2}$ simple closed curves (*ovals*) $W_1, \ldots W_{\frac{n}{2}}$, with $W_i \subset W_{i+1}$ around the origin $[0, 0, 1]$. When n is odd there is an additional non-separating curve.

The proof of Lax conjecture provides the converse of this result. Specifically, Helton and Vinnikov prove that if a polynomial $q(u, v)$ satisfies $q(0, 0) = 1$ and for all $(u, v) \neq (0, 0) \in \mathbb{R}^2, p(t) = q(tu, tv) = 0$ has all roots real, then there are real symmetric matrices H and K with

$$q(u, v) = \det(uH + vK + I).$$

The convex region bounded by the dual curve to $q(u, v)$ is the boundary of the numerical range of $A = H + iK$.

For example, the quartic curve

$$q(u, v) = (u^2 + 6u + v^2 - 52)^2 + 144(4u - 17) = 0$$

satisfies the RZ condition, so there is a 4×4 matrix whose numerical range is bounded by the convex curve dual to the inner curve of q. The outer curve has a double tangent and two inflection points, so the dual curve has an inner real branch which has two cusps and a double point.

The equation of the dual curve is

$$\begin{aligned} &11025u^6 + u^4(838 + 37444v + 79393v^2) + u^2(-2087 - 13628v - 7326v^2 \\ &+ 89804v^3 + 139112v^4) + (4 + 17v)^2(1 - 12v - 68v^2 + 48v^3 + 256v^4) = 0. \end{aligned}$$

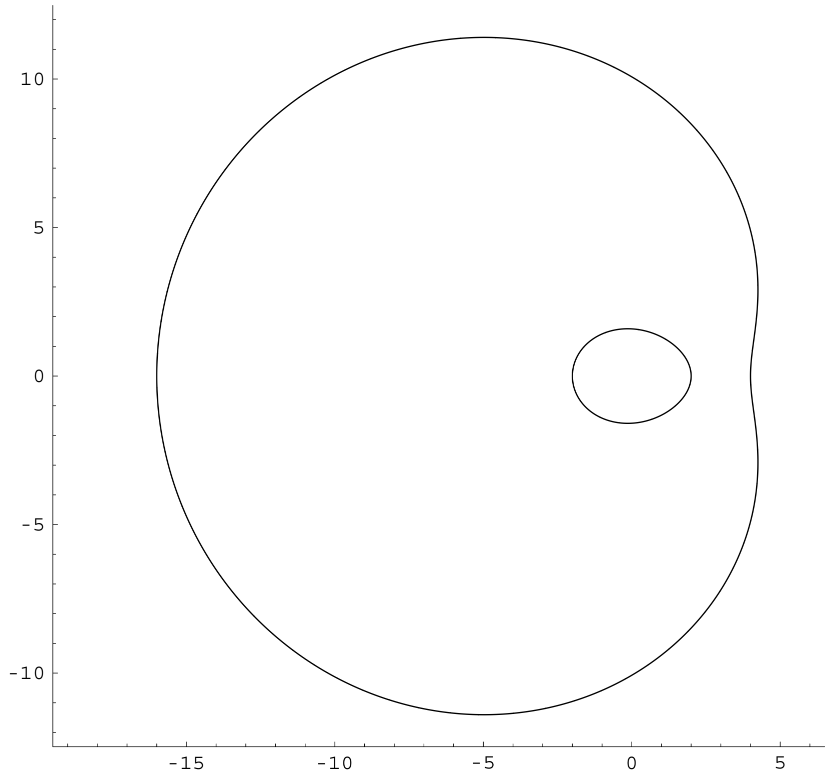

FIGURE 7.1. The curve $q(u, v) = 0$ is an RZ curve.

According to the solution of the Lax conjecture, there is a 4×4 matrix for which the outer curve of Figure 7.2 bounds the numerical range.

8. Inversion

The map from $\mathbb{C} - 0$ to itself taking z to $1/z$ can be extended to almost all of $\mathbb{C}P^2$ by the formula:

$$\mathcal{I}[X, Y, Z] = [XZ, -YZ, X^2 + Y^2].$$

To be precise, let us first exclude the three lines $X + iY = 0$, $X - iY = 0$, and $Z = 0$. On the complement of these three lines, that is, the degenerate cubic $Z(X^2 + Y^2) = 0$, the map is an involution:

$$\mathcal{I}^2[X, Y, Z] = [XZ(X^2 + Y^2), YZ(X^2 + Y^2), X^2Z^2 + Y^2Z^2] = [X, Y, Z].$$

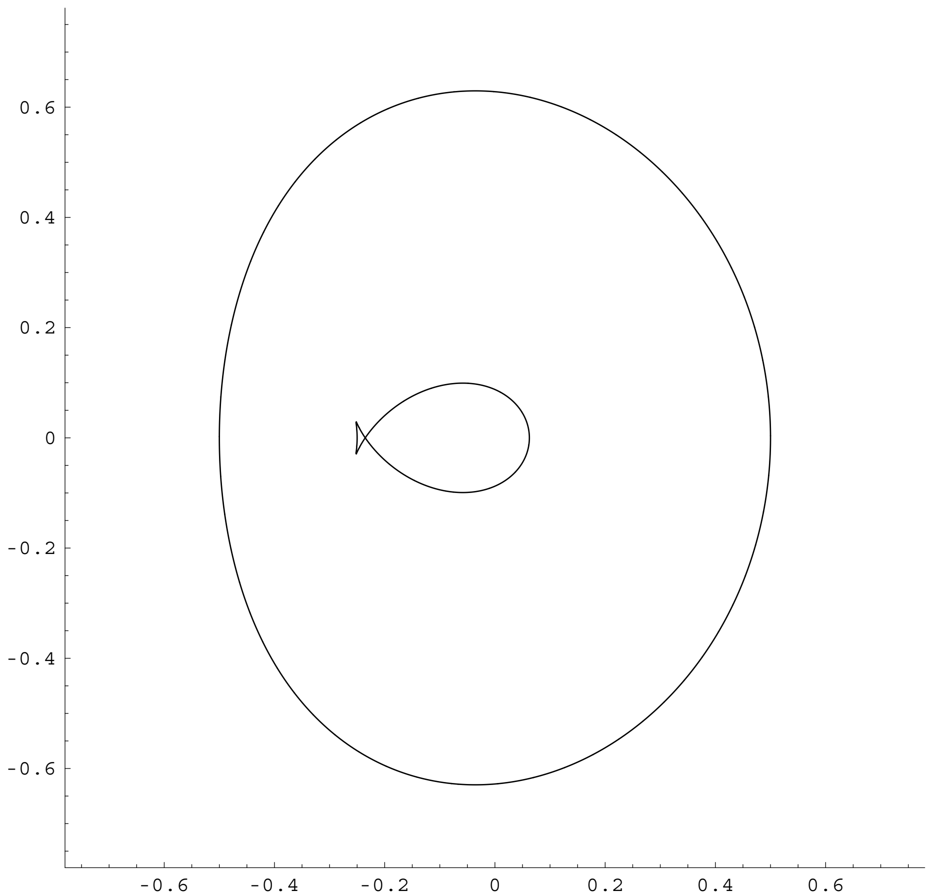

FIGURE 7.2. The dual ("fishbowl") curve.

If an algebraic curve Γ is irreducible, it will meet the degenerate cubic in finitely many points, at each of which the map $\mathcal{I}$ can be extended in a natural way, unless the curve passes through the origin or one of the circular points. On the complement of these three points the map is well-defined, but it fails to be one-to-one. Thus the isotropic line $\{[U, iU, 1] | U \in \mathbb{C}\}$ (with the circular point $[1, i, 0]$ excised), is collapsed to the circular point $[1, -i, 0]$, and a similar collapse occurs for the other lines. Schematically, the degenerate cubic can be thought of as a triangle, and the map takes each edge to the opposite vertex. The transformation Φ takes curves missing the three vertices to curves passing through these vertices. For example, a real ellipse not passing through the origin is taken to a circular, quartic curve which has the origin as an isolated real point. (The inverted ellipse will be discussed in more detail below. It is actually an example of a *bicircular curve*, as it has multiplicity *two* at the circular points.)

Now suppose ℓ is an isotropic line given by the equation $X+iY-wZ = 0$, where $w = u+iv \neq 0$ is a complex number, and u and v are real. A point P on this line which is not an ideal point has (homogeneous) coordinates $[X, Y, \frac{X+iY}{w}]$. We compute

$$\mathcal{I}(P) = [\frac{X(X+iY)}{w}, -\frac{Y(X+iY)}{w}, X^2+Y^2] = [\frac{X}{w}, -\frac{Y}{w}, X-iY]$$

or letting $z = \frac{1}{w}$,

$$\mathcal{I}(P) = [X, -Y, \frac{X-iY}{z}].$$

That is, the finite points on ℓ go to the finite points on the line $\ell' : X + iY - zZ = 0$. Thus the isotropic line through the real point (u, v) is taken to the isotropic line through the real point $(\frac{u}{u^2+v^2}, -\frac{v}{u^2+v^2})$.

Since inversion as defined above takes such an isotropic tangent line to an isotropic line tangent to the image of Γ, we conclude

Theorem 8.1. *Inversion takes the ordinary foci of a curve* Γ *to the ordinary foci of the inverse curve.*

Remark 8.2. The result just stated is actually a special case of *conformal invariance of foci.* Roughly speaking, foci are invariant as they are the singularities of Schwarzian reflection, which is itself invariant. Indeed, any conformal map $H(z)$ yields a local mapping $TH(R, B)$, as defined earlier, which takes isotropic lines to isotropic lines, preserves tangency, the real plane, etc. The invariance property is awkward to formulate precisely, however, because the argument holds only as long as the entire construction lies in the domain of definition of the extended mapping.

9. Schwarz functions and their Singularities: the generic case

Consider a real algebraic curve $\gamma \subset \mathbb{R}^2$ defined by an n^{th}-degree polynomial equation $0 = f(x, y) = g(x+iy, x-iy)$. The corresponding complex curve $\Gamma \subset \mathbb{C}^2$ may be expressed using either standard or isotropic coordinates: $0 = F(X, Y) = G(R, B)$. The n^{th}-degree polynomials F, G have the same coefficients as f, g, respectively; our notation emphasizes the distinction, e.g., between $g(z, \bar{z})$, a real valued function on the real plane, and the function $G : \mathbb{C}^2 \to \mathbb{C}$, which depends on independent variables $R = X + iY, B = X - iY$. In this section, for simplicity, we make the blanket assumption that Γ is *regular* (though singular curves are ultimately important): the partial derivatives G_R and G_B never simultaneously vanish on Γ. Further, the present section will use certain *general position* assumptions (some of which will be relaxed in subsequent sections).

The standard *real structure* (antiholomorphic involution) on $\mathbb{C}^2$, given by complex conjugation $(X, Y) \xrightarrow{\sigma} (\bar{X}, \bar{Y})$, takes Γ to itself. In fact, F is *self-conjugate*: $F^*(X, Y) = \overline{F \circ \sigma(X, Y)} = F(X, Y)$ (since F has real coefficients). The induced real structure $\sigma : \Gamma \to \Gamma$ has fixed point set γ. In isotropic coordinates, the real structure is expressed $(R, B) \xrightarrow{\sigma} (\bar{B}, \bar{R})$, and the self-conjugacy condition $G^* = G$ yields the identity $G(R, B) = \bar{G}(B, R)$; here the bar over G denotes complex conjugation of coefficients.

The *Schwarz function of* γ according to [7] is the analytic function $S(z) = S_\gamma(z)$ locally defined (near a given point $(x_0, y_0) \in \gamma$) by $0 = g(x + iy, S(x + iy))$; i.e., under the identification $(x, y) \leftrightarrow z = x + iy$, γ is the locus of the equation $\bar{z} = S(z)$. Thus, e.g., $S(z) = z$ for the real line $y = 0$ and $S(z) = r^2/z$ for the circle $x^2 + y^2 = r^2$. Schwarzian reflection in γ—locally characterized as the unique antiholomorphic map $\mathcal{R} = \mathcal{R}_\gamma$ fixing points of γ—is consequently given by the formula $\mathcal{R}_\gamma(z) = \overline{S(z)}$ for z near $z_0 = x_0 + iy_0$.

A more global treatment of Schwarz functions and Schwarzian reflection cannot avoid issues of multivaluedness and singularities. Such matters may be handled cleanly in case γ is the real part of a real algebraic curve $\Gamma \subset \mathbb{C}P^2$. We begin by describing the Schwarz function of such Γ as a multivalued function $S : \mathbb{C} \to \mathbb{C}^*$ and subsequently extend to $S : \mathbb{C}^* \to \mathbb{C}^*$ as required for global results, below. Let the n^{th}-order curve Γ be defined by $0 = G(R, B) = \sum_{j=0}^{n} a_j(R) B^j$. Here, $a_j(R)$ is a polynomial of degree at most $n - j$ and, in the present section, we assume $a_n \neq 0$. For a given value of $R_0 = x_0 + iy_0 \in \mathbb{C}$, the isotropic line $X + iY = R_0$ intersects Γ in n points (counting multiplicities)—the solutions to the equation $0 = G(R_0, B)$. Replacing R_0 by a variable R, the equation $0 = G(R, S(R))$ defines S as an algebraic function whose role is to give the n "blue" coordinates corresponding to a given "red" coordinate along Γ. Thus, Γ is the "graph" of its Schwarz function $B = S(R)$. (This is the point of view adopted by Study; see page 236.)

Remark 9.1. The exceptional but important case $a_n = 0$ will be considered in the next section. This is precisely the *circular* case already mentioned. For $a_n = 0 \Leftrightarrow$ the equation $0 = G(R_0, B)$ has fewer than n roots $\Leftrightarrow$ the isotropic line $R = R_0$ meets Γ at an ideal point p_0. But then $p_0 = c_r$ since only the ideal line in $\mathbb{C}P^2$ contains two distinct ideal points. Hence, c_r must lie on Γ and, by symmetry, $c_b \in \Gamma$.

Remark 9.2. Note that complex conjugating the identity

$$0 = G(z, S(z)) = \bar{G}(S(z), z)$$

and substituting B for $\bar{z}$ gives $0 = G(\bar{S}(B), B)$. Thus, $R = \bar{S}(B)$ holds along Γ, and S, $\bar{S}$ may be regarded as formal (or local) inverses of each other. (In examples γ often has the additional reflection symmetry in the x-axis; in that case $f(x, y) = f(x, -y)$ implies $G(R, B) = G(B, R)$ and we may write $S = \bar{S} = S^{-1}$. Lines and circles satisfy this involutivity condition globally.) The formal identity $S^{-1} = \bar{S}$ is discussed in [7] as a necessary condition for an analytic function $S(z)$ to be a *Schwarz function*—meaning, the Schwarz function of *some* analytic curve $\Gamma \subset \mathbb{C}$. On the other hand, the condition holds as well, say, for the function $S(z) = -1/z$, which is not regarded in [7] as a Schwarz function, since the equation $\bar{z}z = -1$ has empty solution set in $\mathbb{C}$. However, it suits our purposes to allow *any* algebraic function satisfying $S^{-1} = \bar{S}$ as a Schwarz function; thus, e.g., $S(z) = -1/z$ is the Schwarz function of a *circle of radius i*—a real algebraic curve which happens to have no real points. (However, the circle of radius i lies outside the present discussion, which is limited to non-circular curves!) [Presumably, it can be verified that an algebraic function satisfying $S^{-1} = \bar{S}$ formally also satisfies $0 = G(R, S(R))$ for some polynomial $G = G^*$.]

To express additional formal properties of Schwarz functions, we introduce notation for *red* and *blue isotropic projections*. First, for $p = (R, B) \in \Gamma$, let $\rho(p) = R$ denote projection from the "red" circular point $c_r = [1, i, 0]$ onto the real plane $\mathbb{R}^2 \simeq \mathbb{C}$; here we use the identification $(a, b) \leftrightarrow z = a+ib$. Then one directly verifies that ρ *intertwines Schwarzian reflection in* γ *and the real structure on* Γ:

$$\rho \circ \sigma = \mathcal{R} \circ \rho.$$

Unlike its well-behaved twin σ, $\mathcal{R}$ is multivalued due to the non-injectivity of ρ (as seen in Study's description, which we have merely reworded here).

Similar remarks hold for $\beta(p) = \rho^*(p) = B$—projection from the "blue" circular point $c_b = [1, -i, 0]$ (but to interpret β as a mapping onto the real plane one must be careful to use the identification $(a, b) \leftrightarrow z = a-ib$

instead). Closely related to the above intertwining formula, we have the following expression for the Schwarz function itself in terms of the multivalued function ρ^{-1}:

$$S = \beta \circ \rho^{-1} = \rho^* \circ \rho^{-1}.$$

We mention that a similar-looking representation of the Schwarz function (as a *symmetric element*) was used in [3], [4] to discuss *symmetric space multiplication* on the space of analytic curves (Schwarzian reflection of one curve in a nearby curve).

Next we consider the local behavior of S in the vicinity of a point $R_0 \in \mathbb{C}$. All but finitely many values R_0 are *regular* points of S: the isotropic line $X + iY = R_0$ meets Γ in n distinct points and $S(R)$ has n analytic branches $B_1 = S_1(R), \ldots, B_n = S_n(R)$, for R near R_0, with corresponding derivatives $S_j'(R) = -\frac{\partial G}{\partial R}(R, B_j)/\frac{\partial G}{\partial B}(R, B_j)$. On the other hand, when the line $X + iY = R_0$ meets Γ in fewer than n distinct points, at least one of these must be a point of tangency $r = (R_0, B_{j0}) \in \Gamma$, according to the usual accounting of Bezout's Theorem. Such a *red isotropic point* r is a zero of the partial derivative $\frac{\partial G}{\partial B}(r)$, and $S_j'(R_0)$ fails to be defined. In this case, R_0 is a branch point for $B = S(R)$ (with order of branching $\mu - 1$ at R_0 agreeing with order of tangency at r). Equivalently, $R = \bar{S}(B)$ has corresponding point of multiplicity $\mu \geq 2$: $\bar{S}_j'(B_{j0}) = -\frac{\partial G}{\partial B}(r)/\frac{\partial G}{\partial R}(r) = 0$.

Likewise, a *blue isotropic point* $b = (R_{j0}, B_0)$ results in a branch point B_0 of $\bar{S}_j$ and a zero of S_j'. The involution σ interchanges red and blue points: $b_1 = \sigma(r_1), \ldots, b_N = \sigma(r_N)$. It follows that Schwarzian reflection in γ interchanges the corresponding branch points and critical points of S: $B_k = \rho(b_k) = \rho(\sigma(r_k)) = \mathcal{R}(\rho(r_k)) = \mathcal{R}(R_k)$. As the points $R_k \in \mathbb{R}^2 \simeq \mathbb{C}$ are the (real) *foci* of γ, according to classical terminology, we will refer to the reflected images $B_k = \mathcal{R}(R_k)$ as *defoci* of γ!

As elsewhere in this section, let us assume genericity: $\frac{\partial G}{\partial B}(r)$ has a simple zero and $S(R)$ has a branch point of order *one* at R_0. Applying Bezout's Theorem to the intersection of the two algebraic curves $G = 0$ and $\frac{\partial G}{\partial B} = 0$, we conclude that there are $m = n(n-1)$ (distinct) red isotropic points. (The specific genericity assumption ruling out ideal points of intersection will be discussed below). We conclude that a generic curve of degree n has $m = n(n-1)$ defoci and as many foci. The number m is the class of the curve (see page 238). Note that the value derived here is a special case of the Plucker formula (5.7).

Finally, to complete the above discussion, we need to consider the ideal points $p_1, \ldots, p_n$ of Γ. If $G_n(R,B) = \sum_{j=0}^{n} A_j R^{n-j} B^j$ is the n^{th}-degree part of $G(R,B)$, then in homogeneous coordinates we may write $p_k = [X,Y,Z] = [R_k + B_k, -i(R_k - B_k), 0]$ where $0 = G_n(R_k, B_k)$ is assumed to give n distinct ratios R_k/B_k. Viewed more intrinsically, the ideal points of Γ may be identified as the poles of isotropic projection. That is, the isolated singularities of $\rho : \Gamma \setminus \{p_1, \ldots, p_n\} \to \mathbb{C}$, initially defined as a holomorphic function of degree n, are necessarily poles for the meromorphic extension to Γ. For if the singularity of ρ at p_k were removable, say, $\rho(p_k) = R_0 \in \mathbb{C}$, then it would follow by continuity that the isotropic line $R = R_0 \in \mathbb{C}$ would contain p_k; but such a line $R = R_0 \in \mathbb{C}$ cannot contain p_k under the present non-circularity assumption on Γ ($a_n \neq 0$) since only the ideal line contains two ideal points. We mention that the red embedding $\mathbb{C}^* \subset \mathbb{C}P^{2*}$ provides a good context for the above argument; ideal points look just like finite points for the continuous mapping $\rho : \Gamma \to \mathbb{C}^* \subset \mathbb{C}P^{2*}$. Similarly, using the blue embedding $\mathbb{C}^* \subset \mathbb{C}P^{2*}$, the corresponding remarks apply to β.

Henceforth, let ρ, β denote the extended meromorphic functions on $\Gamma \subset \mathbb{C}P^2$, and let S denote the algebraic function $S = \beta \circ \rho^{-1} : \mathbb{C}^* \to \mathbb{C}^*$ (allowing $\mathbb{C}^*$ to denote both "red" and "blue" copies of the Riemann sphere as in the previous paragraph). Note that $R = \infty$ thus becomes a simple pole for each branch of this extended S, since both ρ and β have simple poles at each of the n ideal points (Γ has n simple intersections with the ideal line).

We summarize our conclusions thus far:

Proposition 9.3. *Let $\gamma \subset \mathbb{R}^2$ be a nonsingular, real algebraic curve $0 = g(z, \bar{z})$, and assume γ is generic (as specified above in this section—in particular, non-circular). Let $S : \mathbb{C}^* \to \mathbb{C}^*$ be the Schwarz function of γ—the algebraic function $B = S(R)$ defined by the n^{th} degree equation $0 = G(R, S(R)$. Then $S(R)$ has the following singularities (where S fails locally to have n analytic, univalent branches):*

i) *there are $m = n^2 - n$ foci R_k of γ, the branch points of S*
ii) *there are m defoci $B_k = \mathcal{R}_\gamma(R_k)$, the critical points of S: $S'(B_k) = 0$*
iii) *there are n simple poles of S, one for each of the branches of S at $R = \infty$.*

Example 9.4. Consider the ellipse γ: $x^2/a^2 + y^2/b^2 = 1,\ 0 < b < a$. We let $c^2 = a^2 - b^2$, and use coordinates $R = X + iY, B = X - iY$ to express the equation and Schwarz function of the ellipse:

$$0 = G(R, B) = R^2 - 2\frac{a^2 + b^2}{c^2}RB + B^2 + 4\frac{a^2b^2}{c^2}$$

$$B = S(R) = \frac{a^2 + b^2}{c^2}R - \frac{2ab}{c^2}\sqrt{R^2 - c^2}.$$

The ellipse has foci $\pm c = \pm\sqrt{a^2 - b^2}$; these are simple branch points for $S = \bar{S}$. The defoci $d = \overline{S(\pm c)} = \pm\frac{a^2+b^2}{\sqrt{a^2-b^2}}$ are critical points for (one branch of) S. Also, both branches of S have poles at infinity corresponding to the two ideal points $[X, Y, Z] = [a, \pm ib, 0]$ on the complex ellipse $\Gamma \subset \mathbb{C}P^2$. As these poles are simple, the ideal points are regular for ρ, and the only points of ramification for ρ are therefore the two red isotropic points. Γ may thus be viewed as a surface of genus zero obtained by gluing together two copies of the slit domain $\mathbb{C}^* \setminus [-c, c]$.

10. Circular curves

Next we explain how the results of the previous section are modified for a circular curve. When $c_r = [1, i, 0]$ lies on Γ, it counts as one (or more) of the n intersections with a given isotropic line and results in a lower degree for the isotropic projection ρ. This is how a Schwarz function may have a pole in the finite plane, as we shall presently show.

We first illustrate the situation with the circle $x^2 + y^2 = a^2$. In homogeneous coordinates, $X^2 + Y^2 - a^2Z^2 = 0$, $c_r = [1, i, 0]$ is one of two intersection points of Γ with a given isotropic line $X + iY - R_0Z = 0$, $R_0 \neq 0$, and ρ has degree *one*. For $R_0 = 0$, the isotropic line $X + iY = 0$ is a tangent—it intersects Γ twice at c_r. In this respect, the *singular focus* $R_0 = 0$ resembles a focus somewhat, but only the "second" of the two intersections counts as $\rho^{-1}(0)$. Further, using $G(R, B) = RB - a^2$ we find that $S(R) = a^2/R$ has a pole rather than a branch point at $R_0 = 0$. It is also important to note that c_r counts as a *double* isotropic point; there being none other, this gives $m = 2(2 - 1) = 2$ red isotropic points as expected. (To "count" c_r, observe that the equation $0 = \frac{\partial G}{\partial B} = R$ happens to describe the tangent line to γ at c_r, which point is therefore a double intersection of $G = 0$ and $\frac{\partial G}{\partial B} = 0$.) Finally, we mention that isotropic projection on $\Gamma \simeq \mathbb{C}^*$ behaves unlike the usual *stereographic projection* on S^2: ρ has a zero at the "north pole" c_r and a pole at the "south pole" c_b! As the roles of c_r, c_b are reversed for β, ideal points no longer all look the same under red and blue light.

More generally, for a circular curve Γ, let $\mu \geq 1$ be the multiplicity of c_r as an intersection point of Γ with a given isotropic line $X + iY = R_0$. If $\mu = 1$, the remaining branches of $S(R)$ will be bounded near R_0. If $\mu = 2$, the line $X + iY = R_0$ is (first order) tangent to Γ at c_r and there results a meromorphic branch of S with simple pole at R_0. This can be seen as follows. In a sufficiently small punctured neighborhood $\dot{U}_{c_r} = U \setminus \{c_r\} \subset \Gamma$, the projections β and ρ are injective, c_r being a regular point for β and a point of multiplicity *two* for ρ. Therefore, the corresponding branch $S_j(R) = \beta \circ \rho_j^{-1}(R)$ is analytic, unimodular, and unbounded in a punctured neighborhood of R_0, and the claim follows by standard results on isolated singularities of analytic functions. If $\mu \geq 3$, the situation resembles that of an ordinary isotropic point; namely, it is not hard to show that R_0 is a branch point of S of order $\mu - 1$ (say, by considering the locally holomorphic mapping $\rho \circ \beta_j^{-1}$ obtained by taking the branch $\beta_j^{-1}(\infty) = c_r$).

For simplicity, we summarize the situation for circular curves which are "otherwise generic" (in the sense of the preceding section); in particular, there are still $m = n(n-1)$ red isotropic points, but now "two" of them are the ideal point c_r, which does not yield a branch point of S.

Proposition 10.1. *Let $\gamma \subset \mathbb{R}^2$ be a nonsingular, real algebraic curve $0 = g(z, \bar{z})$ of order n, and assume γ is a generic circular curve. Then the Schwarz function $S : \mathbb{C}^* \to \mathbb{C}^*$ of γ has the following singularities:*

i) *there are $m - 2 = n^2 - n - 2$ foci R_k of γ, the branch points of S*
ii) *there are $m-2$ defoci $B_k = \mathcal{R}_\gamma(R_k)$, the critical points of S: $S'(B_k) = 0$*
iii) *there is a singular focus $R_0 \in \mathbb{C}$ where S has a simple pole*
iv) *there are $n - 1$ other simple poles of S, one for each of the branches of S at $R = \infty$.*

Though we do not attempt here a systematic discussion of exceptional curves, examples will illustrate some possible behavior not represented in the last two propositions. For a circular curve, for instance, c_r could result in multiple finite poles of S (multiple singular foci, in case c_r is a multiple point) or a branch point of S (if c_r is an *inflection point* of Γ). Like singular curves and curves for which S has higher order poles or branch points, such examples arise naturally and are hardly "pathological".

Example 10.2. **A circular cubic.** The elliptic curve

$$P(x, y) = (x - 1)(x^2 + y^2) - x$$

is a circular cubic whose singular focus is at the origin and lies on the curve. The curve is self-inverse with respect to the unit circle, so its four foci are $a, \overline{a}, -1/a$ and $-1/\overline{a}$, where $a = (-1 + i) + \sqrt{1 - 2i}$. (It is a general fact that a circular cubic is self-inverse with respect to four circles. If the curve has two components, as in this example, then three of the four circles will be real. ([14, page 218]))

The equation of the dual curve is

$$5v^6 + 4(1 - u)^3(-1 + u + u^2) + 6v^4(-6 - 4u + u^2) \\ - v^2(4 + 20 - 24u^2 - 28u^3 + u^4) = 0.$$

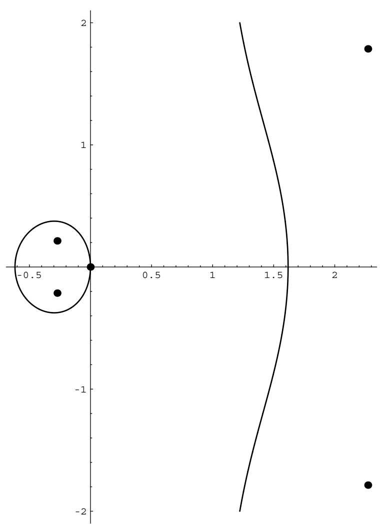

FIGURE 10.1. The circular cubic has four ordinary foci and a singular focus

11. Meromorphic functions and differentials on Γ

Concrete knowledge of even one meromorphic function on a Riemann surface may be very useful. For instance, according to (a special case of) the *Riemann-Hurwitz formula*, if a meromophic function $f : \Gamma \to \mathbb{C}^*$ has degree

d and total branching number $\mathcal{B}$, then Γ has genus: $g = \frac{1}{2}(\mathcal{B}-2d+2)$. To apply this result to the isotropic projection ρ on a generic real algebraic curve Γ (as in the earlier proposition) we insert the values $d = n, \mathcal{B} = n^2 - n = \#$ *red isotropic points* (note the ideal points are simple poles and therefore unbranched). Thus we obtain the *genus formula* (valid for a nonsingular algebraic curve of order n): $g = \frac{1}{2}(n^2 - n - 2n + 2) = \frac{1}{2}(n-1)(n-2)$. For example, in the case of the ellipse, we obtain the expected value $g = 0$. (Note that if we had taken instead a circular curve as in the previous section, we would have obtained $g = \frac{1}{2}((n^2 - n - 2) - 2(n-1) + 2) = \frac{1}{2}(n-1)(n-2)$, the same result.)

Knowing *two* meromorphic functions may be even nicer. For instance, as meromorphic functions on the algebraic curve $\Gamma \subset \mathbb{C}P^2$, the red and blue projections ρ, β in fact define a *primitive pair* in the sense that they separate points on Γ. Namely, consider the n points of intersection $P_1, \ldots, P_n$ of Γ with a red isotropic line $R = R_0$ (assuming the generic case for the moment). For $j = 1, \ldots, n$, we have $\rho(P_j) = R_0$, whereas β assumes n distinct values $\beta(P_1), \ldots, \beta(P_n)$; for if two of these values were the same they would determine a blue isotropic line coinciding with the red isotropic line $R = R_0$, which is impossible (the line at infinity being the only "purple" line). It follows by a standard result that all meromorphic functions on Γ are given by rational expressions in ρ and β:

Proposition 11.1. *Let $f : \Gamma \to \mathbb{C}^*$ be meromorphic. Then there exist rational functions $a_0(z), \ldots, a_{n-1}(z)$ such that $f(P) = \sum_{j=0}^{n-1} a_j(\rho(P))\beta(P)^j$ for $P \in \Gamma$.*

Proof. We reproduce the proof on pp. 249–250 of [9] and observe that the argument applies without modification in the case of a circular curve; note that in the latter case the isotropic projections are meromorphic functions of degree $n = deg(\rho) = deg(\beta) < ord(\Gamma)$ (rather than $n = ord(\Gamma)$). Consider the linear system $f(P_k) = \sum_{j=0}^{n-1} \beta(P_k)^j a_j(R_0),\ k = 1, \ldots, n$, for a generic value R_0. The unique solution $a_0(R_0), \ldots, a_{n-1}(R_0)$ is obtained by applying Cramer's Rule: $a_j(R_0) = \det M_j / \det M, j = 0, \ldots, n-1$, where M has i^{th} row $(1\ \beta(P_i)\ \ldots\ \beta(P_i)^{n-1})$ and M_j is obtained from M by replacing the j^{th} column by $(f(P_1), \ldots, f(P_n))^t$. As the values $\beta(P_1), \ldots, \beta(P_n)$ are distinct, the Vandermonde determinant $detM = \Pi_{k<j}(\beta(P_j) - \beta(P_k)) \neq 0$.

Now replace R_0, above, by a variable R near R_0. Even though the coefficients $f(P_k), \beta(P_k)^j$ in the resulting linear system depend on the multivalued (algebraic) function $\rho^{-1}(R)$, observe that the solution $a_0(R), \ldots,$ $a_{n-1}(R)$ is well-defined and analytic—except at the finitely many values of

R where the points $P_1, \ldots, P_n$ are not all distinct—since the expression obtained from Cramer's Rule depends symmetrically on the n branches. Thus, the given formula defines f as a meromorphic function $f : \Gamma \to \mathbb{C}^*$. □

It follows as a simple corollary (see [9, p. 250]) that the field of meromorphic functions on Γ is isomorphic to an algebraic extension of the field of rational functions of R: $\mathcal{K}(\Gamma) \simeq \mathbb{C}(R)[B]/P(R, B)$.

Next we consider the *meromorphic differentials* on Γ, beginning with the 1-*differentials* (*abelian differentials*). These are defined by expressions $\omega = g(\zeta)d\zeta$, a meromorphic function $g(\zeta)$ for each local coordinate ζ on Γ such that $\omega = g(\zeta)d\zeta$ is suitably invariant; namely, if $\zeta = \varphi(\eta)$ defines an overlapping local coordinate η, then $\omega = g(\varphi(\eta))\varphi'(\eta)d\eta$ is the corresponding local representation of ω. In particular, it follows by the chain rule that each meromorphic function $f \in \mathcal{K}(\Gamma)$ determines an *exact* differential $df = f'(\zeta)d\zeta$. Also, if $\omega_1 = g_1(\zeta)d\zeta$ and $\omega_2 = g_2(\zeta)d\zeta$ are two such abelian differentials then their *ratio* $\omega_1/\omega_2 = g_1(\zeta)/g_2(\zeta)$ gives a well-defined meromorphic function $f(\zeta)$ on Γ. Thus if ω_0 is any given abelian differential on Γ, all others may be represented in the form $\omega = f(\zeta)\omega_0$ for $f \in \mathcal{K}(\Gamma)$.

Returning to the present context, it is a nice fact that the differentials of isotropic projection are (generically) easy to describe. For instance $d\rho$ ($= dR$) gives an alternate route to the genus formula. Namely, the degree $D = \#zeros - \#poles$ of a meromorphic differential on a Riemann surface of genus g satisfies $D = 2g - 2$ (a formula which may be viewed as a very special case of Riemann-Roch). In the generic case, $d\rho$ has first order zeros at each of the $n^2 - n$ red isotropic points and second order poles at each of the n ideal points, thus: $g = \frac{1}{2}(D+2) = \frac{1}{2}(n^2-n-2n+2) = \frac{1}{2}(n-1)(n-2)$ as before.

More significantly, combining several of the above observations, we conclude:

Proposition 11.2. *Any meromorphic (abelian) differential on* Γ *may be written*

$$\omega = \sum_{j=0}^{n-1} a_j(R)B^j dR,$$

where each $a_j(R)$ *is a rational expression in* R.

12. Periods of abelian differentials on Γ, the outer function of $S(z)$, moments of planar domains, etc.

We begin by recalling one of the original applications of the Schwarz function considered in [8] and elsewhere. Let $\mathcal{D}$ be a planar domain, bounded by a piecewise smooth curve γ, and let $f(x,y)$ be a smooth, complex-valued function defined on a neighborhood of the closure of $\mathcal{D}$. (For simplicity, let's assume $\mathcal{D}$ is bounded and simply connected to begin with, though much of what follows does not require this.) By abuse of notation, write $f(x,y) = f(x+iy)$, $\frac{\partial}{\partial \bar{z}} = \frac{1}{2}(\frac{\partial}{\partial x} + i\frac{\partial}{\partial y})$ and express Green's Theorem in complex notation as: $\int_\gamma f(z)dz = i\int\int_{\mathcal{D}} \frac{\partial f}{\partial x} + i\frac{\partial f}{\partial y}dxdy = 2i\int\int_{\mathcal{D}} \frac{\partial}{\partial \bar{z}} f dxdy$. Now consider, say, the special case $f(x,y) = z^m\bar{z}^n$ and let $S(z)$ be the Schwarz function of γ. Then since the identity $\bar{z} = S(z)$ holds along γ, we may compute the $(m,n)^{th}$ *complex moment* of $\mathcal{D}$, for $m,n = 0,1,\dots,$ as a line integral:

$$\begin{aligned} M_{m,n} &= \iint_{\mathcal{D}} z^m\bar{z}^n dxdy \\ &= \frac{1}{2i(n+1)}\int_\gamma z^m\bar{z}^{n+1}dz \\ &= \frac{1}{2i(n+1)}\int_\gamma z^m(S(z))^{n+1}dz. \end{aligned}$$

In particular, $n=0$ gives the sequence of *Richardson moments*:

$$M_m = \iint_{\mathcal{D}} z^m dxdy = \frac{1}{2i}\int_\gamma z^m S(z)dz, \quad m = 0,1,\dots$$

Remark 12.1. Setting now $m=0$ results in a formula for area of the (simply or multiply connected) domain $\mathcal{D}$ bounded by the curve(s) with Schwarz function $S(z)$: $A(\mathcal{D}) = M_0 = \frac{1}{2i}\int_\gamma S(z)dz$. A basic fact about Schwarz functions follows at once: $S(z)$ cannot be holomorphic throughout $\mathcal{D}$—otherwise, Cauchy's Theorem would give $A = 0$.

The Richardson moments are the coefficients in the expansion about $z=\infty$ of the *outer function* of the Schwarz function with respect to γ:

$$S_+(z) = \frac{1}{2\pi i}\int_\gamma \frac{S(w)dw}{z-w} = \frac{1}{2\pi i}\sum_{m=0}^{\infty} M_m z^{-m-1}.$$

The holomorphic function $S_+ : \mathbb{C}^* \setminus \mathcal{D} \to \mathbb{C}$ has the following physical interpretation: $\overline{S_+(z)}$ represents the *gravity field of* $\mathcal{D}$ (that is, the gravitational field of a homogeneous solid occupying the generalized cylinder with cross section $\mathcal{D}$). More precisely, writing $u+iv = \int\int_{\mathcal{D}} \frac{dxdy}{x_0+iy_0-(x+iy)} = \frac{1}{2i}\int_\gamma \frac{S(w)dw}{x_0+iy_0-w}$, the gravity field at a point (x_0, y_0, z_0) outside the cylinder is given by $u\hat{\mathbf{i}} - v\hat{\mathbf{j}}$. The inverse problem, to reconstruct $\mathcal{D}$ from its gravity field (if possible) has been studied quite extensively (see, e.g., [29]). In particular, it is known that a domain $\mathcal{D}$ is *locally* determined by its sequence of Richardson moments $M_0, M_1, \dots$; though one can construct distinct domains having the same moments, there cannot be a continuous family of such domains. This fact is essential, e.g., in certain planar models in which two fluids of very different viscosity occupy complementary regions $\mathcal{D}$ and $\mathcal{D}^c$ of a thin fluid film or layer. For the evolving fluid region $\mathcal{D}$ in the models considered in [29], it is shown that each M_m changes linearly in time, and corresponding evolution of $\mathcal{D}$ is thus determined.

The above formulas apply exceptionally well when $\mathcal{D}$ happens to be a *quadrature domain*. By definition, such a domain satisfies an identity of the form: $\int_{\mathcal{D}} hdxdy = \sum_{k=1}^{N} c_k h(z_k)$. Here, h is any holomorphic (or harmonic) function and the points $z_k \in \mathcal{D}$ and coefficients c_k are fixed. (In general, derivatives of h at these points may also be required in the sum.) In the present context, the existence of such a formula may be understood as follows: The Schwarz function turns out to be meromorphic on such $\mathcal{D}$ and evaluation of the above integral reduces to residue calculus. Namely, letting $f = \bar{z}h$ in Green's Theorem, above, gives $\displaystyle\int_\gamma h(z)S(z)dz = 2i\int\int_{\mathcal{D}} hdxdy$ for h holomorphic in $\mathcal{D}$. When all poles of S in $\mathcal{D}$ are simple, one obtains the quadrature identity $\displaystyle\int\int_{\mathcal{D}} hdxdy = \pi\sum_{k=1}^{N} res(hS; z_k) = \pi\sum_{k=1}^{N} res(S; z_k)h(z_k)$ (and when there are higher order poles, the quadrature identity involving derivatives of h follows easily from the first sum).

As we have seen in an earlier section, $S(z)$ can have poles in the finite plane only at singular foci, so a quadrature domain $\mathcal{D}$ is very special: $\mathcal{D}$ contains only singular foci and no ordinary foci!

Example 12.2. **Neumann's quadrature domain.** Aside from a disk, the first example of a quadrature domain to be discovered was the domain $\mathcal{D}$ whose boundary is obtained by inversion of an ellipse in the unit circle. Specifically, pulling back $\tilde{F}(X,Y,Z) = X^2/a^2 + Y^2/b^2 - Z^2$ by inversion $\Phi[X,Y,Z] = [XZ, -YZ, X^2+Y^2]$ leads to a *bicircular* quartic curve Γ defined by the equation $0 = F(X,Y,Z) = X^2Z^2/a^2 + Y^2Z^2/b^2 - (X^2+Y^2)^2$. Dehomogenizing and transforming to isotropic coordinates, one may solve the resulting equation for the Schwarz function $B = S(R)$ of the inverted ellipse. A shortcut is to use the transformation rule for Schwarz functions under conformal mappings (see [3], [4]): If $\tilde{S}$ is the Schwarz function of $\tilde{\Gamma}$, then the conformal image $\Gamma = \varphi(\tilde{\Gamma})$ has Schwarz function $S = \bar{\varphi} \circ \tilde{S} \circ \varphi^{-1}$. In this case, the Schwarz function $\tilde{S}(R) = \frac{a^2+b^2}{c^2}R - \frac{2ab}{c^2}\sqrt{R^2-c^2}$ of the ellipse defined by $\tilde{F}$ transforms under $R \mapsto 1/R$ to

$$S(R) = 1/\tilde{S}(1/R) = \frac{R(a^2+b^2+2ab\sqrt{1-c^2R^2})}{c^2+4a^2b^2R^2}.$$

We note that S has a meromorphic branch in $\mathcal{D}$ with simple poles at $R_\pm = \dfrac{\pm ci}{2ab}$; the residues are given by $res(S;R_\pm) = \dfrac{a^2+b^2+\sqrt{4a^2b^2+c^4}}{4a^2b^2} = \dfrac{a^2+b^2}{2a^2b^2}$.

(The meromorphic branch of S corresponds to the positive value of $\sqrt{4a^2b^2+c^4}$.) From the earlier computation, we obtain the quadrature formula for $\mathcal{D}$:

$$\int_{\mathcal{D}} h dx dy = \pi \frac{a^2+b^2}{2a^2b^2}\left(h\left(\frac{ci}{2ab}\right) + h\left(\frac{-ci}{2ab}\right)\right).$$

Figure 1.2 on page 229 shows the canonical foliation of the algebraic curve bounding $\mathcal{D}$ in the case $a = 5, b = 3, c = 4$. Thus, the highlighted (heavier) curve is obtained by inversion of the ellipse of Figure 1.1.

Proof of Theorem 1.1. We are assuming that the curve γ that bounds the numerical range is nonsingular and all the eigenvalues, and therefore all of the foci of γ, are in the interior. If we perform an inversion in a circle centered at some point in the interior of the numerical range, the resulting curve γ^{-1} will have its foci in the exterior (unbounded) region of the plane. Then the Schwarz function S of γ^{-1} is meromorphic on the interior $\mathcal{D}$ and the residue theorem can be applied. Thus the interior will be a simply connected bounded quadrature domain in the plane. □

Example 12.3. An elliptical region $\mathcal{D}$ itself is *not* a quadrature domain, but one may still usefully apply residue calculus. To illustrate, we compute the outer function of the Schwarz function $S(z) = \frac{a^2+b^2}{c^2}z - \frac{2ab}{c^2}\sqrt{z^2-c^2}$ of the ellipse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. If $z \in \mathcal{D}^c$, then $w/(z-w)$ is holomorphic for $w \in \mathcal{D}$, so $S_+(z) = \frac{abi}{c^2\pi}\int_\gamma \frac{\sqrt{w^2-c^2}dw}{z-w}$. This integral may be evaluated by summing the residues of (the relevant branch of) $g(w)dw = \frac{\sqrt{w^2-c^2}}{z-w}dw$ at poles in $\mathcal{D}^c$. The simple pole at $w = z$ has residue $-\sqrt{z^2-c^2}$, and the double pole at $w = \infty$ has residue z, so $S_+(z) = \frac{2ab}{c^2}(z - \sqrt{z^2-c^2})$. The correct branch satisfies the usual properties of an outer function: $S_+(z)$ vanishes at $z = \infty$, and $S(z) - S_+(z) = \frac{(a-b)^2}{c^2}z$ is holomorphically continuable to $\mathcal{D}$. Further, writing $z = 1/\zeta$, $S_+(1/\zeta) = \frac{2ab}{c^2\zeta}(1 - \sqrt{1-c^2\zeta^2}) = \frac{ab}{c^2}\zeta + \frac{abc^2}{4}\zeta^3 + \ldots$, one may read off the Richardson moments of $\mathcal{D}$ from the Taylor expansion for $\sqrt{1-c^2\zeta^2}$ about $\zeta = 0$. ($M_0 = \pi ab/c^2, M_1 = 0, M_2 = \pi abc^2/4$, etc.) Below, we discuss a slightly different computation which illustrates a general approach to moments of algebraic domains.

Next, we reconsider such integrals from a different point of view. Namely, suppose γ happens to be the real part of an algebraic curve Γ, and suppose, for simplicity, γ divides Γ into two *halves*: $\Gamma \setminus \gamma = \Gamma_+ \cup \Gamma_-$ is a disjoint union of two connected open subsets of Γ. Then instead of working with the planar domain $\mathcal{D}$ itself, we apply the residue calculus to Γ_+. In this more "intrinsic" setting, we no longer express the required integrals in terms of the Schwarz function but in terms of the abelian forms $\omega = \sum_{j=0}^{n-1} a_j(R)B^j dR$ considered earlier; for instance, we compute area of the planar domain $\mathcal{D}$ as $\frac{1}{2i}\int_\gamma S(z)dz = \frac{1}{2i}\int_\gamma BdR$, and evaluate the latter by residues of the differential $BdR = \beta d\rho$ on Γ_+ (or on Γ_-).

Example 12.4. To illustrate the strategy, we recompute the Richardson moments of the ellipse by evaluating the integrals $M_m = \frac{1}{2i}\int_\gamma R^m BdR$. We have already made use of the fact that ρ projects one of the two halves—say, Γ_+—injectively onto the exterior $\mathcal{E}_\gamma$ of γ, and that the Schwarz function consequently has a single-valued meromorphic branch on $\mathcal{E}_\gamma$. Here we show that Γ_- may be used just as well, even though $\rho(\Gamma_-)$ includes branch points of S.

To begin the computation recall that $B = \beta$ is holomorphic except for simple poles at both ideal points, where also $R = 0$. To compute the residues of $\omega_m = R^m BdR$ at these points we write $R = 1/w, B = wQ, \omega_m = -w^{-m-3}Qdw$ and expand Q in powers of w. In the variables w, Q the previous equation for the ellipse becomes $0 = c^2 + 4a^2b^2w^2 - 2d^2q + c^2q^2$, with $c^2 = a^2 - b^2$ and $d^2 = a^2 + b^2$. Substituting the series expansions

$Q = \sum_{n=0}^{\infty} A_n w^n$ and $Q^2 = \sum_{n=0}^{\infty} w^n \sum_{k=0}^{n} A_k A_{n-k} = A_0^2 + 2A_0A_1w + (A_1^2 + 2A_0A_2)w^2 + \sum_{n=3}^{\infty} w^n(2A_0A_n + \sum_{k=1}^{n-1} A_kA_{n-k})$ into the latter equation, one reads off the following equations for the coefficients of $w^0, w^1, w^2, \ldots$:

$$w^0 : 0 = c^2 - 2d^2A_0 + c^2A_0^2$$

$$w^1 : 0 = 2A_1(-d^2 + c^2A_0)$$

$$w^2 : 0 = 4a^2b^2 - 2d^2A_2 + c^2(A_1^2 + 2A_0A_2)$$

$$w^n : 0 = -2d^2A_n + 2c^2A_0A_n + c^2 \sum_{k=1}^{n-1} A_kA_{n-k}, n \geq 3.$$

One then solves for $A_0 = \frac{(a\pm b)^2}{c^2}, A_1 = 0, A_2 = \mp ab$ and computes $A_3, A_4, \ldots$ recursively:

$$A_n = \frac{c^2}{2(d^2 - c^2A_0)} \sum_{k=1}^{n-1} A_kA_{n-k} = \frac{c^2}{4a^2b^2}A_2 \sum_{k=1}^{n-1} A_kA_{n-k}, n \geq 3.$$

Finally, one obtains $M_m = \mp\pi Res[-w^{-m-3}Qdw; w = 0] = \pm\pi A_{m+2}$, where the $+$ $(-)$ sign results in the final expression when the residue theorem is applied on Γ_+ (Γ_-).

For the ideal point $[a, ib, 0] \in \Gamma_+$ the correct sign choices are: $A_0 = \frac{(a-b)^2}{c^2}, A_2 = ab$, hence, $M_0 = \pi ab$. For $[a, -ib, 0] \in \Gamma_-$, one has $A_0 = \frac{(a+b)^2}{c^2}, A_2 = -ab$, with the same result $M_0 = \pi ab$. Likewise, a sign cancellation explains the fact that identical values for $M_1, M_2, \ldots,$ are obtained whether the residue theorem is applied on Γ_+ or on Γ_-. Ramification points of isotropic projection do not interfere with this computation of moments; i.e., the strategy may be applied (in principle) even when the Schwarz function does not have a meromorphic branch on *either side* of the curve.

13. The fundamental quadratic differential on Γ

The singularities of the Schwarz function of a real algebraic curve $\gamma \subset \mathbb{R}^2$ provide a "skeletal" picture of the embedding of the corresponding complex curve $\Gamma \subset \mathbb{C}^2$. The picture may be fleshed out and placed in a geometric context by considering the trajectories of a certain quadratic differential q canonically associated with $\Gamma \subset \mathbb{C}^2$, as follows. As the equations $dR = 0$ and $dB = 0$ describe the respective foliations of $\mathbb{C}^2$ by red and blue isotropic lines, it seems reasonable to take as a starting point the *fundamental form* on $\mathbb{C}^2$ (see [5]):

$$Q = dX^2 + dY^2 = dRdB.$$

The pull-back of Q to Γ by the inclusion $\iota : \Gamma \to \mathbb{C}^2$ may be expressed using isotropic coordinates:

$$q = \iota^* Q = S'(R)dR^2 = -\frac{G_R}{G_B}dR^2 = -\frac{G_B}{G_R}dB^2.$$

Assuming Γ is regular, either R or B will serve as a holomorphic, local coordinate near $p \in \Gamma$, depending on whether $G_B(p) \neq 0$ or $G_R(p) \neq 0$. Thus, it follows from the last two expressions that q does indeed define a holomorphic quadratic differential on $\Gamma \cap \mathbb{C}^2$. Further, the expression $q = d\rho d\beta$ shows that q is actually a meromorphic quadratic differential on all of $\Gamma \subset \mathbb{C}P^2$, as required for the application of global theorems in the examples below.

One may now apply to q the standard geometric constructions (as in [27] or [20]) which associate to such a quadratic differential on a Riemann surface both a singular flat geometry and a pair of orthogonal, singular foliations. Accordingly, we define the *canonical flat geometry on* $\Gamma \subset \mathbb{C}P^2$ by the metric

$$g = |q| = |S'(R)|dRd\bar{R}.$$

To define the *canonical foliation* we use the *horizontal trajectories of* q—these are given by parametrized curves $p(u) = ((R(u), B(u)))$ satisfying the condition:

$$0 < q(\frac{dp}{du}) = S'(R)(\frac{dR}{du})^2 = \bar{S}'(B)(\frac{dB}{du})^2.$$

The *vertical trajectories* satisfy instead $0 > q(\frac{dp}{du})$.

The flatness of g may be seen by taking the square root of q to get a linear differential ω and its dual vectorfield W:

$$\omega = \sqrt{q} = \sqrt{S'(R)}dR \;\longleftrightarrow\; W = \frac{1}{\sqrt{S'(R)}}\frac{\partial}{\partial R}.$$

The differential ω and complex vectorfield $W = (u + iv)\frac{\partial}{\partial z}$ are locally defined, and globally defined when q happens to be *orientable*—all singularities have even order. The real vectorfields $Re[W] = u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y}$ and $Im[W] = -v\frac{\partial}{\partial x} + u\frac{\partial}{\partial y}$ are orthonormal with respect to g, and commute as a consequence of the Cauchy-Riemann equations (equivalently, ω is closed). The integral curves of $Re[W]$ and $Im[W]$ are, respectively, the horizontal and vertical trajectories, which are geodesics in the metric g.

The previous paragraph simply recalls some of the basics of the general theory. In the case of $q = d\rho d\beta$, the construction has the following additional features. The real part γ of Γ is a horizontal trajectory (or union of finitely many). Here we assume γ to be nonsingular, and recall that the unit tangent vector along $\gamma \subset \mathbb{C}$ is given by $T = 1/\sqrt{S'(R)}$ (see, [7] or [3]); so $q(\frac{dp}{du}) = S'(R)(\frac{dR}{du})^2 = S'(R)(\frac{dR}{ds}\frac{ds}{du})^2 = (\frac{ds}{du})^2 > 0$. Thus, γ itself belongs to the canonical foliation as claimed. Further, the zeros of q are precisely the (red and blue) isotropic points on Γ, so the (finite) singularities of this foliation indeed project to the foci and defoci of γ. Finally, the poles of $q = d\rho d\beta$ are the ideal points of Γ which, in the case of circular points, project to singular foci. The geometry of the flat surface in the vicinity of such a singularity of q is discussed in detail in [20]; we will mention only a few simple examples here.

Example 13.1. Unit speed parametrization of the real circle $x^2 + y^2 = r^2$ extends analytically to a global parametrization of the complex circle $X^2 + Y^2 = RB = r^2$, missing the circular points c_r, c_b:

$$p(s + it) = (r\cos\theta, r\sin\theta),\ \theta = (s + it)/r,\ 0 < s \leq 2\pi r,\ -\infty < t < \infty.$$

Since $q(\frac{dp}{ds}) = d\rho(\frac{dp}{ds})d\beta(\frac{dp}{ds}) = (iR/r)(-iB/r) = 1$, the curves $s \mapsto p(t, s)$ are indeed horizontal trajectories of q. Isotropic projection gives the exponentially shrinking circles $\rho(p) = R = re^{i\theta}$ filling up $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$. Using the Schwarz function of the circle $S(R) = r^2/R$, one can also write $q(\frac{dp}{ds}) = S'(R)(\frac{dR}{ds})^2 = -\frac{r^2}{R^2}(\frac{iR}{r})^2 = 1$ to verify that the above concentric circles are the (projected) horizontal trajectories of q. According to a general result, simple poles of $\omega = \sqrt{q} = irdR/R$ imply $(\mathbb{C}^\times, g)$ has cylindrical ends of circumference $2\pi|res(\omega)| = 2\pi r$ (see [20]); presently, $(\mathbb{C}^\times, g)$ is globally an infinite cylinder of radius r.

The circle is very special in that its isotropic projection is injective, and the canonical foliation can be viewed as a foliation of $\mathbb{C}$. In general, the condition $S'(z)dz^2 > 0$ defines a "multi-line field" on $\mathbb{C}$, and the resulting trajectories form multiple "sheets." In such isotropically projected foliations, the branch points occur at foci, thus distinguishing foci from defoci—as will be illustrated presently. We consider such multi-foliations on $\mathbb{C}$ in examples (as well as the cleaner description on Γ) not just for the sake of graphics, but also because the real plane is the setting of the main classical results on foci of algebraic curves.

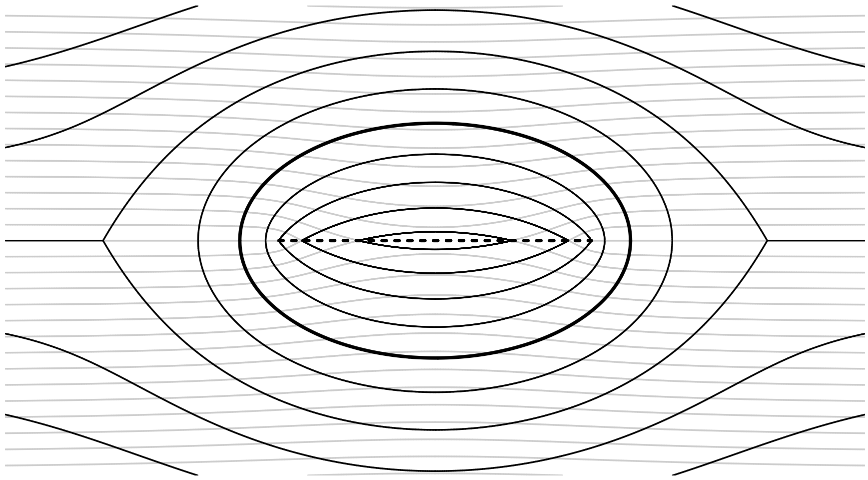

FIGURE 13.1. Canonical geodesic departing from the ellipse $\frac{x^2}{25}+\frac{y^2}{9}=1$.

Example 13.2. Using the same notation as before for the ellipse $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$ the fundamental quadratic differential may be expressed:

$$\begin{aligned} q &= \left(\frac{a^2+b^2}{c^2}-\frac{2abR}{c^2\sqrt{R^2-c^2}}\right)dR^2 \\ &= \frac{c^2R-(a^2+b^2)B}{(a^2+b^2)R-c^2B}dR^2 \\ &= \frac{c^2B-(a^2+b^2)R}{(a^2+b^2)B-c^2R}dB^2. \end{aligned}$$

The third expression for q vanishes at red points $r_\pm=((\pm c, S(\pm c)))$, while the second vanishes at blue points $b_\pm=\sigma(r_\pm)=((S(\pm c),\pm c))$. These isotropic points are first order zeros for q. The only other singularities of q occur at the two ideal points ($[X,Y,Z]=[a,\pm ib,0]$) of the complex ellipse $\Gamma\subset\mathbb{C}P^2$. These are seen to be fourth order poles corresponding to $w=0$ with $R=1/w$ in the expression for q. The order sum $\#zeroes-\#poles=4(1)-2(4)=-4$ fits the general formula $\#zeroes-\#poles=4g-4$ for a quadratic differential on a surface of genus g. Equivalently, the four zeroes of q correspond to singularities of the q-horizontal foliation of index minus one-half, while the poles give singularities of index *two*; thus, we have $\chi=4(-1/2)+2(2)=2$.

We visualize q via the projection $\rho : \Gamma \to \mathbb{C}^*$, a covering of degree $d = 2$ branched at the two red points $r_\pm$. In Figure 13.1, the ρ-projected horizontal trajectories are shown in two sheets. Here we have used the values $a = 5, b = 3$, resulting in singularities at foci $c_\pm = \pm 4$ and defoci $d_\pm = \pm\frac{17}{2}$. Note that foci and defoci look quite different—despite the red-blue symmetry $b_\pm \stackrel{\sigma}{\longleftrightarrow} r_\pm$—because ρ is regular at $b_\pm$ and singular at $r_\pm$.

See the characteristic pattern of the singularity of index minus one-half at d_+, where three trajectories meet at angles $0, 2\pi/3$ and $4\pi/3$. At the same location, the other sheet is foliated by nearly straight lines. Here it should be noted that r_+ is only one of the two values of $((d_+, S(d_+)))$; the other is a regular point of q, $((\frac{17}{2}, \frac{257}{8}))$, accounting for the "lined sheet". Meanwhile, r_+ is doubly-projected by ρ onto c_+, and the horizontal trajectries in Γ forming angles of $2\pi/3$ at r_+ are projected to curves making angles of $4\pi/3$ with each other in the doubly covered neighborhood of c_+.

Example 13.3. Consider $\Gamma \subset \mathbb{C}P^2$ *a hyperelliptic curve of genus* $g \geq 1$:

$$0 = F(X, Y) = Y^2 - P(X) = Y^2 - \sum_{j=0}^{n} A_j X^j.$$

Here we assume, for simplicity, that P has an even number $n = 2g + 2$ of distinct roots $x_1, \dots, x_n$. We may use X as coordinate at all other values $X \in \hat{\mathbb{C}} \setminus \{x_1, \dots, x_n\}$, and there correspond two sheets $Y = \pm\sqrt{P}$ for the coordinate projection $\Gamma \ni (X, Y) \mapsto X$. Now consider the quadratic differential:

$$q = \iota^*(dX^2 + dY^2) = (1 + \frac{(P')^2}{4P})dX^2.$$

The pair of equations $P'(X)^2 + 4P(X) = 0$, $Y^2 = P(X)$ give $4n - 4$ zeroes $(X_k, Y_\pm) \in \Gamma$ of q (which ρ projects onto the $2n - 2$ real foci and as many defoci of Γ.) There are no other finite singularities of q, for at each exceptional point $(x_k, 0)$ we may use Y as local coordinate ($F_X = P' \neq 0$) and write $q = (1 + \frac{4Y^2}{(P')^2})dY^2$.

It remains to determine the behavior of q at the two points $X = \infty_\pm$ completing the two branches of Γ. (From the extrinsic point of view, there is only *one* ideal point $(X, Y, Z) = (0, 1, 0)$, a degenerate double point obtained by homogenizing the above equation; however, we are presently considering the q-foliation of Γ as an intrinsic Riemann surface.) Thus, using $X = 1/w$ at $w = 0$, we find that $q = \frac{P'(1/w)^2 + 4P(1/w)}{4P(1/w)w^4} dw^2$ has a pole of order $n + 2$ at each of these two points. Altogether, we obtain

$\#zeroes - \#poles = (4n-4) - 2(n+2) = 4g - 4$, the correct sum for a quadratic differential on a Riemann surface of genus g.

Finally, we note that isotropic projection $\rho : \Gamma \to \mathbb{C}^*$ has degree $d = n$, with branch points of order *one* at the $2n-2$ foci, and branch points of order $\frac{n}{2} - 1$ at the two ideal points. (Note $\rho = X \pm i\sqrt{P(X)}$ has poles of order $\frac{n}{2}$ at $X = \infty$.) The total order of branching is thus $\mathcal{B} = (2n-2) + 2(\frac{n}{2} - 1)$, in agreement with Riemann-Hurwitz: $\mathcal{B} = 2g + 2d - 2 = (n-2) + 2n - 2 = 3n - 4$.

We return now to the representation of the canonical foliation of Γ in terms of arclength parametrization of γ, as discussed in the introduction. Example 13.1 will serve as the prototypical example for the more local version of Theorem 1.2 which we wish to describe. To give the expanding/shrinking circle $c_t = R(s + it) = re^{i\theta}$ a more dynamical flavor we will, alternatively, regard c_t as a geodesic in the symmetric space of regular analytic curves, as mentioned in the introduction. This point of view emphasizes the fact that the families of curves in our figures should not be viewed merely as leaves in a foliation—they are separated by a fixed "time step" (which amounts to the same thing as fixed *distance* in the flat metric $g = |q|$).

One checks that c_t evolves in time by "continuous Schwarzian reflection" by noting that the corresponding time-dependent Schwarz function $S_t(z) = r^2 e^{-2t/r}/z$ satisfies the geodesic equation $\frac{\partial^2 S}{\partial t^2} - (\frac{\partial S}{\partial t}/\frac{\partial S}{\partial z})\frac{\partial^2 S}{\partial t \partial z} = 0$. The "initial point" c_0 of this geodesic is the circle $x^2 + y^2 = r^2$ and the "initial velocity" is the (inward) unit normal along c_0; in other words, c_t is the *canonical geodesic departing from* c_0, as defined in the introduction. With respect to the flat metric $g = |q| = r^2 dz\overline{dz}/z^2$, the circle c_t sweeps out the cylinder $(\mathbb{C} \setminus \{0\}, g)$ with speed *one*, asymptotically approaching the *ends* $z = 0, \infty$ at $t = \pm\infty$.

More generally, if the initial curve γ is non-circular, the canonical geodesic will be defined only on a finite time interval $-\tau < t < \tau$; here we note that Schwarzian reflection in the initial curve defines *time-reversal symmetry* in the geodesic equation. The geodesic is defined on a finite maximal time interval $|t| < \tau$ because at $t = \pm\tau$, the curve becomes singular and no longer belongs to the space of regular analytic curves.

A representation of such a canonical geodesic may be obtained, in principle, as follows. Let γ have arclength $2\pi r$. Parametrize γ by arclength using the notation $\gamma(s) = f(re^{is/r})$, $0 \le s < 2\pi r$, where $f(z)$ maps $c_0 = \{z : |z| = r\}$ onto γ with unit speed $|f'(re^{i\theta})| = 1$. Here, we may assume that $f(z)$, defined via analytic continuation of $\gamma(s)$, is analytic on an annular domain $A = \{z : re^{-a/r} < |z| < re^{a/r}\}$, which is symmetric with respect to inversion in c_0. Now let $\Gamma(s+it) = f(re^{i\theta}),\ \theta = (s+it)/r,\ 0 \leq s < 2\pi r,\ -a < t < a$. By conformal invariance of the geodesic equation (see [3], Calini-Langer2), Γ defines a geodesic. In fact, since Γ restricts to an arclength parametrization of $\gamma = f(c_0)$, it follows that, for $|t| < a$, $\Gamma_t = \Gamma(s+it)$ is the canonical geodesic departing from $\Gamma_0 = \gamma$, i.e., Γ locally parametrizes the canonical foliation defined by q.

In fact, such a representation is valid over the maximal time interval $|t| < \tau$, and τ may be described as the time it takes to reach the nearest foci (or defoci) of γ:

Theorem 13.4. *Let $\gamma = x + iy$ be a nonsingular, algebraic plane curve of length $2\pi r$, and assume γ is not a circle. Let γ be given an arclength parametrization, expressed in the form $\gamma(s) = f(re^{is/r})$, as above. Then there is a finite "maximal time" $\tau > 0$ such that: $\gamma(s)$ may be continued analytically to*

$$\Gamma(s+it) = f(re^{i\theta}),\ \theta = (s+it)/r,\ 0 \leq s < 2\pi r,\ -\tau < t < \tau,$$

where the closure of the image of Γ contains at least one focus and one defocus of γ. $\Gamma_t = \Gamma(s+it)$ represents the maximal, canonical geodesic departing from γ.

Proof. We regard γ as the real part of an algebraic curve $\Gamma \subset \mathbb{C}P^2$. We know that γ is a closed trajectory of $q = d\rho d\beta$, which induces the Euclidean arclength parametrization on γ. It is known (see [27, pp. 38–42]) that such a closed trajectory may be embedded in a *maximal ring domain* $\mathcal{U}$ in the Riemann surface $M = \Gamma$. In the abstract setting, in case M is a torus and the quadratic differential has no singularites, $\mathcal{U}$ might be a non-singular ring domain. In the present case, however, we know that q must have singularities on Γ. In case $M = S^2$, it is also possible for the only singularities to be a pair of double poles or a single pole of order *four*. However, both of these cases may be ruled out for q, when γ is not a circle or straight line. In all other cases, $\mathcal{U}$ (which is uniquely determined) must be a *singular ring domain* of finite *modulus*, bounded by a union of singular trajectories and finite singularities of q—i.e. finite isotropic points in Γ. (The modulus of $\mathcal{U}$ corresponds to τ.) Further, the natural parameter (here "complex arclength") in this case is known to define a global conformal map from $\mathcal{U}$ onto the annulus of the same modulus (after transformation of the corresponding rectangle by the exponential map). Inversion of the natural parameter then gives the parametrization of $\mathcal{U}$ by horizontal trajectories (corresponding to the concentric circles foliating the annulus).

By uniqueness of analytic continuation, it follows that the latter parametrization, followed by isotropic projection ρ onto $\mathbb{C}$ agrees with the function $\Gamma(s+it)$. Since isotropic points are mapped by ρ to foci and defoci (which are symmetrically paired), the image of $\Gamma(s+it)$ must therefore have both types of points as limit points. □

14. The hyperbolic ellipse

The ellipse in the hyperbolic plane can be represented in the Klein model of the hyperbolic plane $x^2+y^2<1$ by a compact quadratic curve given in orthogonal coordinates as $\frac{x^2}{a^2}+\frac{y^2}{b^2}=1, 0<b<a<1$ (see [29]). Schilling ([23]) proved that the ellipse defined in this manner is the locus of a point the sum of whose hyperbolic distances from the two geometric foci are constant. As noted in [26], the coordinates $\pm\delta$ of these geometric foci are farther from the origin than the Euclidean foci $\pm c$. The precise relationship is

$$c^2=a^2-b^2=\frac{\delta^2(1-a^2)}{1-\delta^2} \tag{14.1}$$

or equivalently

$$\delta^2=\frac{c^2}{1-b^2}. \tag{14.2}$$

The *Klein-to-Poincaré map* $z=KP(w)=\frac{w}{1-\sqrt{1-|w|^2}}$ is the unique isometry from the disc D with the Klein metric to the disc with the Poincaré metric which keeps the ideal boundary pointwise fixed. It takes the straight line segment between points on the boundary to the NE geodesic between the same two points. (See [26]). For our purpose we need the inverse map

$$KP^{-1}(z)=\frac{2z}{1+z\overline{z}}$$

or

$$KP^{-1}(x,y)=(\frac{2x}{1+x^2+y^2},\frac{2y}{1+x^2+y^2}).$$

Pulling back the equation of the ellipse, the hyperbolic ellipse (in the Poincaré disc) is given by

$$\Gamma:\frac{4x^2}{a^2}+\frac{4y^2}{b^2}-(x^2+y^2+1)^2=0. \tag{14.3}$$

The geometric foci are located at $\pm\epsilon$, where $\delta = \frac{2\epsilon}{1+\epsilon^2}$ or

$$\epsilon = \frac{\delta}{1+\sqrt{1-\delta^2}} = \frac{\sqrt{1-b^2}-\sqrt{1-a^2}}{c}. \tag{14.4}$$

The algebraic curve (14.3) is an example of a *bicircular quartic* (see, e.g., [14, p. 304]) This curve has two real components, one interior to and the other exterior to the unit disc; the two components are related by reflection through the unit circle. As an algebraic curve it is of class 8 and deficiency (genus) 1, having four real (algebraic) foci and two singular foci.

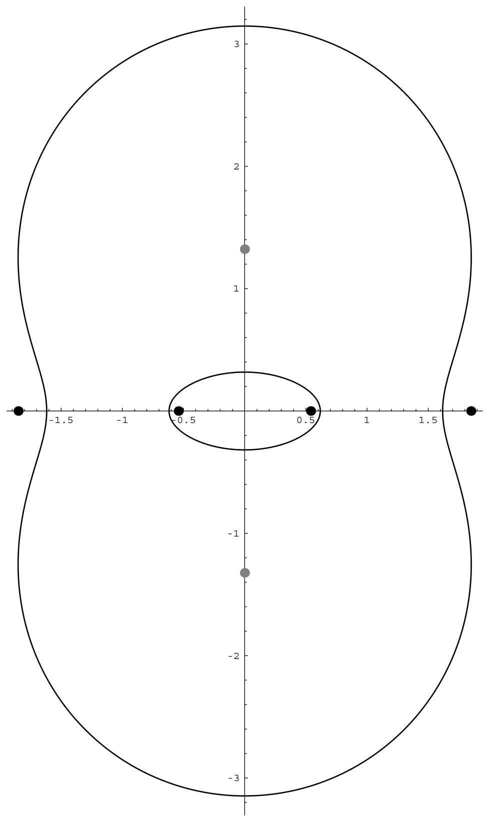

FIGURE 14.1. The hyperbolic ellipse $5x^2+12y^2 = (1+x^2+y^2)^2$ and its foci.

Proposition 14.1. *The real foci of the bicircular quartic* Γ *given by equation* (14.3) *coincide with the geometric foci of the hyperbolic ellipse and their reflections in the unit circle.*

Proof. The algebraic foci are the singularities of the Schwarz reflection; they are the points (x, y) in the plane for which the isotropic lines are tangent to the curve Γ. To find these points, we look at the general equation of an isotropic line and see when the equation for its intersection with the curve has a double root. Thus we look at the pair of equations:

$$\frac{4x^2}{a^2} + \frac{4y^2}{b^2} - (x^2 + y^2 + 1)^2 = 0, x + iy = U.$$

Eliminating x leads to the equation

$$(\frac{4}{b^2} - \frac{4}{a^2})y^2 - \frac{8iU}{a^2}y + \frac{4U^2}{a^2} = (1 - 2iUy + U^2)^2. \qquad (14.5)$$

This is quadratic in y, because the curve Γ is bicircular, meaning that it has double points at both circular points. Therefore, a line through a circular point meets the curve twice at the circular point and therefore has only two other intersections. Rewrite this equation as $Ay^2 + By + C = 0$; then

$$A = \frac{4}{b^2} - \frac{4}{a^2} + 4U^2, \ B = 4iU(1 + U^2 - \frac{2}{a^2}), \ C = \frac{4U^2}{a^2} - (1 + U^2)^2. \ (14.6)$$

The discriminant equation $B^2 - 4AC = 0$ would appear to be sixth order in U but there is a cancellation. The equation is actually biquadratic and reduces to

$$U^2 = \frac{1}{c^2}[2 - a^2 - b^2 \pm 2\sqrt{(1 - a^2)(1 - b^2)}]. \qquad (14.7)$$

Taking square roots, we get the four algebraic foci:

$$U = \pm\frac{1}{c}[\sqrt{1 - b^2} \pm \sqrt{1 - a^2}] \qquad (14.8)$$

These four foci are all real, and they correspond to the geometric foci and their reciprocals. □

In addition to the four foci determined above, there are two additional singular foci, which are given by the real points on the lines tangent to the curve at the circular points. By parametrizing the general isotropic line we can find these points:

First rewrite Equation (14.3) in homogeneous form:

$$\Gamma : \frac{4x^2z^2}{a^2} + \frac{4y^2z^2}{b^2} - (x^2 + y^2 + z^2)^2 = 0. \qquad (14.9)$$

Let $x = 1 + tX$, $y = i + tY$, and $z = t$ where X and Y are the (unknown) coordinates of the focus. Plugging these into equation (14.9) gives

$$t^2(2X+2iY+t(X^2+Y^2+1))^2-\frac{4t^2}{a^2}(1+2tX+t^2X^2) \\ -\frac{4t^2}{b^2}(-1+2itY+t^2Y^2)=0. \quad (14.10)$$

The factor of t^2 corresponds to the line passing through the circular point twice, once for each branch. When the line is tangent to the curve at the circular point, there will be an extra factor of t. This occurs when

$$4(x+iY)^2-\frac{4}{a^2}-\frac{4}{b^2}=0.$$

Thus the singular foci occur at

$$X=0,\ Y=\pm\frac{c}{ab}.$$

Remark 14.2. An interesting consequence of the general theory of bicircular quartics ([14, p. 305]) is that there is an ellipse with foci at these singular foci such that the curve is the envelope of the circles centered at points on the ellipse and orthogonal to the unit circle.

It is instructive to compare such computations with the Schwarz function $S(R)$ and fundamental quadratic differential $q=S'(R)dR^2$ for Γ. Writing

$$F(X,Y)=a^2b^2(X^2+Y^2+1)^2-4b^2X^2-4a^2Y^2=0,$$
$$G(R,B)=a^2b^2(RB+1)^2+a^2(R-B)^2-b^2(R+B)^2=0,$$

we solve for $B=S(R)$:

$$S(R)=\frac{(a^2+b^2-a^2b^2)R+ab\sqrt{p(R)}}{a^2-b^2+a^2b^2R^2},$$
$$p(R)=-a^2+b^2+\left(4-2a^2-2b^2\right)R^2+\left(-a^2+b^2\right)R^4.$$

Differentiation yields $q=q1+q2+q3$ where:

$$q1=\frac{\left(-a^2-b^2+a^2b^2\right)\left(-a^2+b^2+a^2b^2R^2\right)}{\left(a^2-b^2+a^2b^2R^2\right)^2}dR^2,$$

$$q2=\frac{2a^3b^3R\sqrt{p(R)}}{\left(a^2-b^2+a^2b^2R^2\right)^2}dR^2,$$

$$q3=\frac{2abR\left(-2+a^2+b^2+\left(a^2-b^2\right)R^2\right)}{\left(a^2-b^2+a^2b^2R^2\right)\sqrt{p(R)}}dR^2.$$

The radicals in $S(R), q2, q3$, represent the same branch of the square root of $P(R)$.

We identify the foci $f_j = \rho(r_j)$ as branch points of $S(R)$, i.e., the four solutions to the biquadratic $p(R) = 0$ given above. The defoci are obtained by Schwarzian reflection in γ:

$$d_j = \beta(b_j) = \overline{S(f_j)} = \frac{\left(a^2b^2 - a^2 - b^2\right)(a' - b')\, c}{2a^2b^2a'b' - a^4b'^2 - b^4a'^2},$$

$$a' = \pm\sqrt{1-a^2}, \quad b' = \pm\sqrt{1-b^2}.$$

(Note that Schwarzian reflection in γ is *two*-valued for regular points—keeping in mind that Γ is a bicircular quartic—and single valued at foci.) Finally, the singular foci are the poles of S; the denominator $a^2 - b^2 + a^2b^2R^2$ vanishes at the two values $R = X + iY = \pm i\frac{c}{ab}$ found above by a different argument.

The quadratic differential q has simple zeros at each of the 8 isotropic points r_j, b_j, and double poles at each circular point. As the latter are double points (and account for all the ideal points of Γ), we have $\#zeroes - \#poles = 8 - 8 = 4g - 4$.

The isotropic projections ρ, β are coverings of degree $d = 2$ with four ramification points apiece; namely, the points r_j and b_j, respectively, have branching order *one* for ρ and β. The other points of interest, the circular points, are regular points for ρ and β! In particular, β has simple poles at (each copy of) the red circular point c_r, as can be seen by considering $R = \infty$ in the above expression for $S(R)$. Likewise, using $\bar{S} = S$, one sees that ρ is unramified at c_b. Meanwhile, ρ maps a neighborhood of (each copy of) c_r regularly onto a neighborhood of the (corresponding) singular focus $\rho(c_r)$ (likewise for β at c_b). This can be seen, e.g., from the earlier computation which shows that a tangent line to Γ at c_r intersects Γ with order *three*—not *four*—hence, nearby isotropic lines contain the double point c_r and exactly one point in a punctured neighborhood of c_r. Note that the conclusions of the present paragraph are consistent with $q = d\rho d\beta$ having double poles at each circular point. Finally, the Riemann-Hurwitz formula is satisfied: $\mathcal{B} = 2g + 2d - 2 = 4$.

References

[1] Hassan Aref and Morten Brons, *On stagnation points and streamline topology in vortex flows.* J. Fluid Mech. **370** (1998), 1–27.

[2] M. Bôcher, *Some propositions concerning the geometric representation of imaginaries.* Ann. of Math. **7** (1892), 70–76.

[3] Annalisa Calini and Joel Langer, *Schwarz reflection geometry I: continuous iteration of reflection.* Math. Z. **244** (2003), 775–804.

[4] ———, *Schwarz reflection geometry II: local and global behavior of the exponential map.* Exp. Math. **16** No. 3 (2007), 321–346.

[5] Élie Cartan, *The Theory of Spinors.* Reprinted by Dover Publications, 1981.

[6] J.L. Coolidge, *A Treatise on Algebraic Curves.* Oxford, Clarendon Press, 1931.

[7] Philip J. Davis, *The Schwarz Function and its Applications.* The Carus Mathematical Monographs, No. 17. The Mathematical Association of America (1974).

[8] Philip Davis and Henry Pollack, On the analytic continuation of the mapping function, TAMS **87** No.1 (1958), 198–225.

[9] H. M. Farkas and I. Kra, *Riemann Surfaces.* GTM 71, Springer Verlag (1980).

[10] Björn Gustafsson, *Quadrature identities and the Schottky double.* Acta Appl. Math. **1** (1983), 209–240.

[11] Björn Gustafsson and Mihai Putinar, *Selected topics on quadrature domains.* Physica D (2007), doi:10.1016/j.physd.2007.04.015.

[12] F. Hausdorff, *Der Wertvorrat einer Bilinearform.* Math. Z. **3** (1919), 314–316.

[13] J.W. Helton and V. Vinnikov, *Linear matrix inequality representation of sets.* Comm. Pure Appl. Math. **59** (2006), 1–21.

[14] H. Hilton, *Plane Algebraic Curves, second edition.* London, Oxford University Press, 1932.

[15] R. A. Horn and C. R. Johnson, *Topics in matrix analysis,* Camb. Univ. Press, New York, 1991.

[16] B. Iverson, *Hyperbolic Geometry.* Cambridge, Cambridge University Press, 1992.

[17] R. Kippenhahn, *Über den Wertevorrat einer Matrix.* Math. Nachr. **6** (1951), 193–228.

[18] A.S. Lewis, P.A. Parillo, and M.V. Ramana, *The Lax Conjecture is True.* Proc. Amer. Math. Soc. **133** (2005), 2495–2499.

[19] M. Marden, *The Geometry of the Zeros of a Polynomial in a Complex Variable.* American Mathematical Society Mathematical Surveys III, New York, 1949.

[20] Jesus Muciño-Raymundo, *Complex structures adapted to smooth vector fields.* Math. Ann. **322** (2002), 229–265.

[21] F. D. Murnaghan, *On the field of values of a square matrix.* Proc. Nat. Acad. Sci. U.S.A. **18** (1932), 246–248.

[22] George Salmon, *A Treatise on the Higher Plane Curves, Third Edition.* G. E. Stechert & Co., New York, 1934.

[23] F. Schilling, *Dei Brennpunktseigenschaften der eigentlichen Ellipse in der ebenen nichteuklidischen hyperbolischen Geometrie.*(German), Math. Ann. **121** (1950), 415–426.

[24] Harold S. Shapiro, *The Schwarz Function and its Generalization to Higher Dimensions.* University of Arkansas Lecture Notes in the Mathematical Sciences 9, John Wiley & Sons, Inc., 1992.

[25] J. Siebeck, *Ueber eine neue analytische Behandlungweise der Brennpunkte.* J. Reine. Angew. Math. **64** (1864), 175.

[26] D. Singer, *The Location of Critical Points of Finite Blaschke Products.* Conform. Geom. Dyn. **10** (2006), 117–124.

[27] Kurt Strebel, *Quadratic Differentials.* Springer-Verlag, 1984.

[28] O. Toeplitz, *Das algebraische Analogon zu einem Satze von Fejer.* Math. Z. **2** (1918), 187–197.

[29] A. N. Varchenko and P. I. Etinghof, *Why the Boundary of a Round Drop Becomes a Curve of Order Four.* University Lecture Series **3**, AMS (1991).

Joel C. Langer and David A. Singer
Dept. of Mathematics, Case Western Reserve University
Cleveland, OH 44106-7058
USA
e-mail: `joel.langer@case.edu`
`david.singer@case.edu`

Lecture held by Joel Langer in the *Seminario Matematico e Fisico* on October 6, 2006

Received: July 2007