

On the Observable Shape of Black Hole Photon Rings

Samuel E. Gralla^{1,*} and Alexandru Lupsasca^{2,3,†}

¹*Department of Physics, University of Arizona, Tucson, AZ 85721, USA*

²*Princeton Gravity Initiative, Princeton University, Princeton, NJ 08544, USA*

³*Society of Fellows, Harvard University, Cambridge, MA 02138, USA*

Motivated by the prospect of measuring a black hole photon ring, in previous work we explored the interferometric signature produced by a bright, narrow curve in the sky. Interferometric observations of such a curve measure its “projected position function” $\mathbf{r} \cdot \hat{\mathbf{n}}$, where \mathbf{r} parameterizes the curve and $\hat{\mathbf{n}}$ denotes its unit normal vector. In this paper, we show by explicit construction that a curve can be fully reconstructed from its projected position, completing the argument that space interferometry can in principle determine the detailed photon ring shape. In practice, near-term observations may be limited to the visibility *amplitude* alone, which contains incomplete shape information: for convex curves, the amplitude only encodes the set of projected diameters (or “widths”) of the shape. We explore the freedom in reconstructing a convex curve from its widths, giving insight into the shape information probed by technically plausible future astronomical measurements. Finally, we consider the Kerr “critical curve” in this framework and present some new results on its shape. We analytically show that the critical curve is an ellipse at small spin or inclination, while at extremal spin it becomes the convex hull of a Cartesian oval. We find a simple oval shape, the “phoval”, which reproduces the critical curve with high fidelity over the whole parameter range.

I. INTRODUCTION

General relativity predicts that gravitational lensing will generically produce narrow “photon rings” on images of sources near a black hole [1–6]. Recently, horizon-scale emission from the black hole in the galaxy M87 was resolved using ground-based interferometry [7–12], and future space missions may be able to measure the much narrower photon ring via its universal signature on long baselines [5, 13]. A precise measurement of the photon ring shape could be used to infer black hole parameters and/or as a test of general relativity.

The observable signature of this shape in interferometry is the “projected position function” of the underlying curve [13]. By counting of (functional) degrees of freedom, one expects to be able to reconstruct the full curve from this information, but no proof was given in Ref. [13]. This paper provides a simple constructive proof of this fact, completing the argument that the universal signature discussed in Refs. [5, 13] encodes the photon ring shape and giving a straightforward method to recover it.

Measuring the *full* projected position function of a photon ring requires absolute phase tracking, which may be experimentally challenging with present technology. A more plausible near-term goal is the measurement of the visibility amplitude alone, which provides only partial information: for a closed, convex curve, the amplitude encodes only the set of projected diameters, or *widths*, of the shape [13]. The second purpose of this paper is to explore the information contained solely in the widths of a closed, convex shape. We show that the freedom in reconstructing a closed, convex shape from its widths

consists of one choice of an odd function on the circle, and we give a constructive method for producing all shapes that share a given set of widths.

The shape of astrophysical photon rings is expected to closely approximate that of the theoretical “critical curve” predicted by general relativity for Kerr black holes. The third purpose of this paper is to explore the detailed shape of this critical curve as encoded in its interferometric observables. We analytically show that the critical curve is an ellipse to quadratic order in (either) the black hole spin or observer inclination, and derive the associated ellipse parameters. At extremal spin (and any inclination), we show that the critical curve is the convex hull of a Cartesian oval, a classical shape first studied by Descartes in 1637. We also compute analytically the projected position function of the equatorial critical curve at any spin. Finally, we find approximate expressions for the critical curve that provide an excellent fit over the entire parameter range, including the extremal limit. These expressions are obtained by using a certain notion of “shape addition” that naturally arises in our study. By adding together a circle, an ellipse, and a certain cuspy triangle, we obtain a closed curve that we dub the *phoval*. This four-parameter family of shapes encompasses the three-parameter critical curve of the Kerr black hole to very high accuracy.

This paper is organized as follows. In Sec. II, we show how to reconstruct a general plane curve from its universal signature on long interferometric baselines. We then specialize in Sec. III to closed, convex curves (including a certain cuspy generalization). Next, we present a series of examples in Sec. IV, before finally turning to the study of the Kerr critical curve in Sec. V. In the appendices, we also connect with the Cauchy surface area theorem, review the method of implicitization of plane curves, and study the algebraic geometry of the Kerr critical curve.

* sgralla@email.arizona.edu

† lupsasca@princeton.edu

II. GENERAL CURVES

Consider a smooth (possibly disjoint) plane curve, together with a set of Cartesian axes. Fix a polar angle φ and let $\hat{\mathbf{u}}$ be the outward unit vector in that direction,

$$\hat{\mathbf{u}} = (\cos \varphi, \sin \varphi). \quad (1)$$

This vector is not to be viewed as a vector field, but rather as a fixed vector (for each choice of φ) that can be placed at any point in the plane.

For each $\varphi \in [0, \pi]$, consider the set of points \mathbf{r}_I on the curve such that $\hat{\mathbf{u}}$ is normal to the curve at these points. Consider the generic case in which the curvature is non-vanishing at these points.¹ Each point has a sign S_I defined by whether $\hat{\mathbf{u}}$ points toward (+) or away (-) from the center of curvature. Letting $\hat{\mathbf{n}}_I$ denote the unit normal pointing toward the center of curvature, we have

$$S_I(\varphi) = \hat{\mathbf{n}}_I(\varphi) \cdot \hat{\mathbf{u}}(\varphi) \in \{\pm 1\}. \quad (2)$$

For each point, we further define the *projected position*,

$$z_I(\varphi) = \mathbf{r}_I(\varphi) \cdot \hat{\mathbf{u}}(\varphi). \quad (3)$$

The information directly available in the universal interferometric signature of a narrow curve is the set $\{S_I(\varphi), z_I(\varphi)\}$ for all $0 \leq \varphi < \pi$ (see Eq. (36) and surrounding discussion in Ref. [13]).

Above, we considered fixing φ and finding a list of points \mathbf{r}_I . We may instead imagine fixing I and varying φ , such that the functions $\mathbf{r}_I(\varphi)$ parameterize the curve. Each segment I may have a different range of φ . The direction of increasing φ defines an orientation for each segment, but this does not necessarily provide a consistent orientation across segments that join. In particular, the orientation flips when curves join at inflection points. The normal $\hat{\mathbf{n}}_I$ is continuous over each segment I , and points in the direction of the center of curvature, which is given by a 90° counter-clockwise rotation of the tangent vector $\mathbf{r}'_I(\varphi)$. A Frenet-Serret formula relates the two directly by

$$\mathbf{r}'_I(\varphi) = -\mathcal{R}_I(\varphi)\hat{\mathbf{n}}'_I(\varphi), \quad (4)$$

where $\mathcal{R}_I(\varphi) > 0$ is the radius of curvature at $\mathbf{r}_I(\varphi)$.

From Eqs. (1), (2) and (4), we obtain expressions for the tangent and normal to the curve at $\mathbf{r}_I(\varphi)$,

$$\mathbf{r}'_I(\varphi) = S_I \mathcal{R}_I(\varphi) (\sin \varphi, -\cos \varphi), \quad (5)$$

$$\hat{\mathbf{n}}_I(\varphi) = S_I (\cos \varphi, \sin \varphi). \quad (6)$$

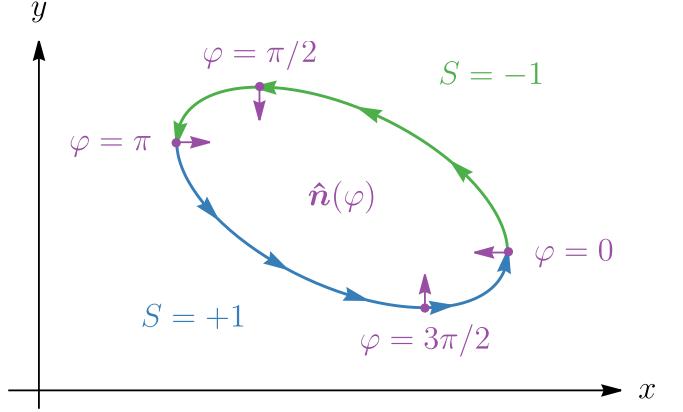


FIG. 1. In the general framework capable of handling arbitrary plane curves (Sec. II), a closed, convex curve is described by two segments: its “top” ($S = -1$, green), and its “bottom” ($S = +1$, blue). Each of these segments is separately parameterized by $\varphi \in [0, \pi]$ (not shown). In the framework of Sec. III, we instead parameterize the whole curve with a single angle φ (purple) ranging over $[0, 2\pi]$. This provides a counter-clockwise orientation around the curve (depicted with arrows). We use a single inward-pointing normal vector $\hat{\mathbf{n}}$.

Here, we dropped the φ -dependence of S_I , imagining that we restrict to a single segment I , over which it is constant.

Together, Eqs. (3), (5), and (6) imply that the projected position z_I and its first derivative are given by

$$z_I(\varphi) = x_I(\varphi) \cos \varphi + y_I(\varphi) \sin \varphi, \quad (7)$$

$$z'_I(\varphi) = -x_I(\varphi) \sin \varphi + y_I(\varphi) \cos \varphi, \quad (8)$$

where x_I and y_I are the Cartesian positions $\mathbf{r}_I = (x_I, y_I)$. Inverting gives the parameterization as

$$x_I(\varphi) = z_I(\varphi) \cos \varphi - z'_I(\varphi) \sin \varphi, \quad (9a)$$

$$y_I(\varphi) = z_I(\varphi) \sin \varphi + z'_I(\varphi) \cos \varphi. \quad (9b)$$

Taking a derivative and comparing with Eq. (5) gives the curvature radius as

$$\mathcal{R}_I(\varphi) = -S_I [z_I(\varphi) + z''_I(\varphi)]. \quad (10)$$

Since Eqs. (9) parameterize the curve over some collection of segments I , this completes the proof that the entire curve can be uniquely reconstructed from the observational data. Note that the sign S_I contains redundant information; the full curve follows entirely from $z_I(\varphi)$.

III. CLOSED, CONVEX CURVES

A convex curve is one for which every tangent line intersects the curve precisely once. In the formalism developed above, a closed, convex curve consists of a “top” segment with $S_I = -1$ joined with a “bottom” segment with $S_I = +1$. The angle φ takes the full range $[0, \pi]$ on each segment, jumping discontinuously between 0 and π

¹Isolated points where the curvature vanishes (and $\hat{\mathbf{n}}_I$ is ill-defined) are to be treated as a limit, where $\hat{\mathbf{n}}_I$ can change sign and the number of different points I can change. A finite segment of the curve with vanishing curvature (straight line) can be approximated by a very small curvature over the region.

at the joining points. We may eliminate these jumps as well as the signs by instead parameterizing the bottom segment with $\varphi \in [\pi, 2\pi)$, so that the whole curve is described by $\varphi \in [0, 2\pi)$ —see Fig. 1. That is, if an interferometric observation reveals two segments with opposite signs that smoothly join (i.e., a closed, convex curve), then it is convenient to promote φ to the full range $[0, 2\pi)$ and thereby deal with a single parameterized curve $\mathbf{r}(\varphi)$.

We will call this φ the *normal angle*. It is the angle that the *inward* normal vector makes with the *negative* x -axis. (With these conventions, φ reduces to the usual polar angle when the curve is an origin-centered circle.) The formulas for the tangent (5) and normal (6) become

$$\mathbf{r}'(\varphi) = \mathcal{R}(\varphi)(-\sin \varphi, \cos \varphi), \quad (11)$$

$$\hat{\mathbf{n}}(\varphi) = -(\cos \varphi, \sin \varphi). \quad (12)$$

These are the counter-clockwise tangent (positive orientation) and the inward normal. One may regard Eq. (12) as the definition of the normal angle φ of a convex curve (stipulating that $\hat{\mathbf{n}}$ is the inward unit normal).

To represent projected position for a closed, convex curve we use $f = -Sz$ instead of z , or equivalently,

$$f(\varphi) = -\mathbf{r}(\varphi) \cdot \hat{\mathbf{n}}(\varphi). \quad (13)$$

More explicitly, we have

$$f(\varphi) = x(\varphi) \cos \varphi + y(\varphi) \sin \varphi. \quad (14)$$

The curvature radius (10) is now given by

$$\mathcal{R}(\varphi) = f(\varphi) + f''(\varphi), \quad (15)$$

while the parameterization (9) takes an identical form,

$$x(\varphi) = f(\varphi) \cos \varphi - f'(\varphi) \sin \varphi, \quad (16a)$$

$$y(\varphi) = f(\varphi) \sin \varphi + f'(\varphi) \cos \varphi. \quad (16b)$$

This is just the vector $(f(\varphi), f'(\varphi))$ rotated by φ .

The function $f(\varphi)$ contains information both about the “curve itself” and about the choice of Cartesian axes. Rotations of the plane induce rotations on the circle, i.e., translations in φ . By contrast, translations of the plane $\mathbf{r} \rightarrow \mathbf{r} + (X, Y)$ change $f(\varphi)$ by [see Eq. (14)]

$$f \rightarrow f + X \cos \varphi + Y \sin \varphi. \quad (17)$$

In practice, one will typically want to compute the projected position function $f(\varphi)$ from a plane curve given in parametric form $(x(\sigma), y(\sigma))$. The angle φ that we have defined satisfies

$$\tan \varphi(\sigma) = -\frac{x'(\sigma)}{y'(\sigma)}, \quad (18)$$

which allows one to directly compute $\varphi(\sigma)$. Plugging into Eq. (14) then provides the projected position in terms of σ , but comparison with observations (and/or use of the framework of this section) requires knowledge of f in terms of φ . The needed inversion to find $\sigma(\varphi)$ is in general highly nontrivial. Note, however, that one can easily plot $f(\varphi)$ as a parametric curve $(\varphi(\sigma), f(\varphi(\sigma)))$.

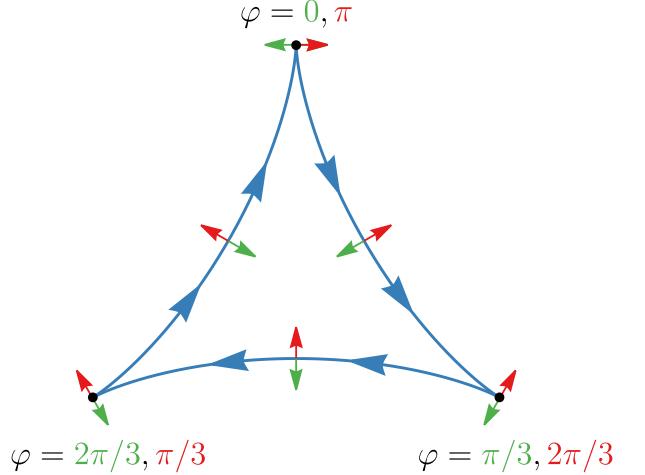


FIG. 2. The framework developed for closed, convex curves is naturally extended to a class of cuspy curves by allowing $f(\varphi)$ to be an arbitrary C^1 function on the circle. The parameterization (16) provides a consistent orientation around the curve (depicted with arrows), and the normal vector (12) varies continuously. In this example of $f(\varphi) = \sin 3\varphi$, the full range $\varphi \in [0, 2\pi)$ traces over the closed curve twice. After one loop around the figure (green), the normal vector has switched sign (top point); after the second loop (red), it has returned to its original leftward orientation.

A. Extension to cuspy curves

In the previous section, we showed that for a smooth convex curve, the projected position function f provides a parameterization of the curve by the normal angle φ . Conversely, any C^1 function $f(\varphi)$ on the circle defines a closed curve via Eqs. (16). However, this curve is in general not convex or even simple or smooth. To guarantee a closed, convex curve one must choose $f(\varphi)$ to be a C^2 function such that $f(\varphi) + f''(\varphi)$ is strictly positive.² This ensures that the curvature never vanishes, proving that the closed curve is convex.

However, it is interesting (and important for the decomposition discussed below) to ask what sort of non-convex curves may be represented within this framework. That is, suppose one chooses some $f(\varphi)$ that is C^2 on the circle (so as to keep the curvature radius finite), but otherwise arbitrary. What kind of curve is defined by the parameterization (16)? To examine this question, we will take Eq. (15) to be the *definition* of \mathcal{R} (which can now be negative), and we will take Eq. (12) to be the *definition* of $\hat{\mathbf{n}}$ (which can now also point away from the center of

²One can also obtain a convex shape by choosing $f(\varphi)$ such that $f(\varphi) + f''(\varphi)$ is strictly negative. However, in this case the normal vector $\hat{\mathbf{n}}$ defined by (12) points outward, such that the shape is not properly parameterized by the normal angle.

curvature). The other equations in Sec. III hold without modification.

Computing the signed curvature k from Eqs. (16), we find that $k = 1/|\mathcal{R}|$. That is, the curvature is always positive, although it diverges wherever \mathcal{R} vanishes, signaling the presence of a sharp cusp. At such a cusp, the unit tangent $\mathbf{r}'(\varphi)/|\mathbf{r}'(\varphi)|$ flips sign [see Eq. (11)]. This means that the parameter φ provides a consistent orientation of the whole closed curve, with the tangent changing direction at cusps. Notice, however, that the normal defined by Eq. (12) does not change sign: it is by definition continuous on the whole curve (but no longer always points toward the center of curvature).

In summary, the curves defined using Eqs. (16) with a generic function $f(\varphi)$ that is C^2 on the circle are either convex, or composed of individually convex segments joined at sharp cusps. An example of the latter is shown in Fig. 2.

As $f(\varphi)$ is defined on the circle, it is tempting to analyze its Fourier series. At first, this approach works beautifully: the zeroth moment is related to the perimeter of the curve (App. A), and the first moment can always be set to zero by a translation of the Cartesian origin. However, the higher moments are not as convenient because the basis functions $\{\cos(m\varphi), \sin(m\varphi)\}$ define pointed stars with sharp cusps for integers $|m| > 1$. For odd m , these are m -pointed stars, with their parameterization tracing their shape twice (e.g., Fig. 5 bottom), while for even m , these are $2m$ -pointed stars, with their parameterization tracing them only once.

B. Antipodal decomposition

A closed, convex curve admits a natural, well-defined notion of antipodal points: those whose tangent lines have the same slope. In our framework, two antipodal points have normal angles φ that differ by π . For each normal angle φ , we introduce the antipodal displacement vector \mathbf{A} and the midpoint position \mathbf{m} along the antipodal chord,

$$\mathbf{A}(\varphi) = \mathbf{r}(\varphi) - \mathbf{r}(\varphi + \pi), \quad (19)$$

$$\mathbf{m}(\varphi) = \frac{\mathbf{r}(\varphi) + \mathbf{r}(\varphi + \pi)}{2}. \quad (20)$$

Each of these vector functions defines a closed curve on the plane. The curve defined by \mathbf{A} is symmetric under reflections through the origin (or equivalently, rotations by π), since \mathbf{A} flips sign under $\varphi \rightarrow \varphi + \pi$. It encodes the overall dimensions of the original curve and for this reason will be called its *hull*. The curve defined by \mathbf{m} is in general displaced from the origin and always features sharp cusps, since \mathbf{m} returns to itself after a lapse of only φ , during which time the continuous normal $\hat{\mathbf{n}}$ flips sign. This *midpoint curve* encodes the original curve's position on the plane as well as its fine features. The parameterization $\mathbf{m}(\varphi)$ traces out the midpoint curve twice over the range $\varphi \in [0, 2\pi]$, as illustrated in Fig. 2.

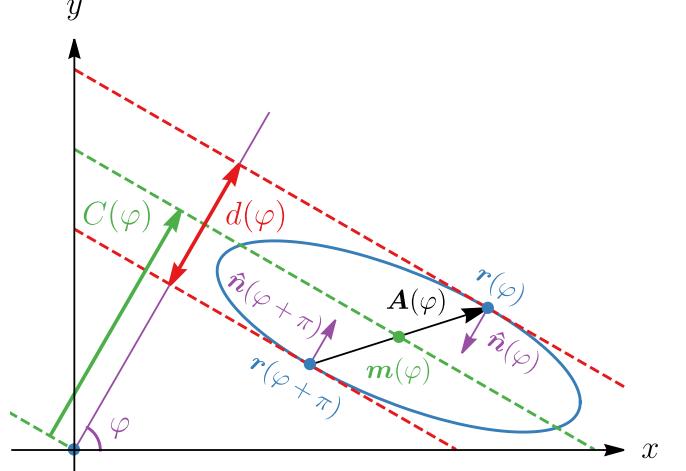


FIG. 3. For a closed, convex curve, each angle φ defines a pair of antipodal points $\mathbf{r}(\varphi)$ and $\mathbf{r}(\varphi + \pi)$ (blue dots). In interferometry, one measures the projection of these points onto a line at angle φ (purple). We denote by $d(\varphi)$ the projected distance (red), which is the angle-dependent projected diameter, or width, of the shape. The midpoint between the antipodes is denoted $\mathbf{m}(\varphi)$, and its projection $C(\varphi)$ is interpreted as an angle-dependent projected centroid (green).

In other words, the notion of antipodal points naturally leads to a decomposition of any closed, convex curve into a symmetric, centered hull and a cuspy, displaced midpoint curve. The full curve is recovered by vector addition of the hull and midpoint curves, identifying points that share a common normal angle. Each point on the midpoint curve has two normal angles that differ by π , so that the full midpoint curve is thus included twice in the sum. An example is illustrated in Fig. 4.

The projected positions of the hull and midpoint curves will be denoted d and C , respectively,

$$d(\varphi) = -\mathbf{A} \cdot \hat{\mathbf{n}} > 0, \quad (21)$$

$$C(\varphi) = -\mathbf{m} \cdot \hat{\mathbf{n}}. \quad (22)$$

These are to be interpreted as the projected diameter and centroid, respectively (Fig. 3). The projected diameter is always positive, while the projected centroid can take either sign. The full projected position f is decomposed into d and C as

$$f(\varphi) = \frac{1}{2}d(\varphi) + C(\varphi), \quad (23)$$

or equivalently,

$$d(\varphi) = f(\varphi) + f(\varphi + \pi), \quad (24)$$

$$C(\varphi) = \frac{1}{2}[f(\varphi) - f(\varphi + \pi)]. \quad (25)$$

That is, the projected diameter and centroid are the parity-even and parity-odd parts of the projected position, respectively. This decomposition is very natural in

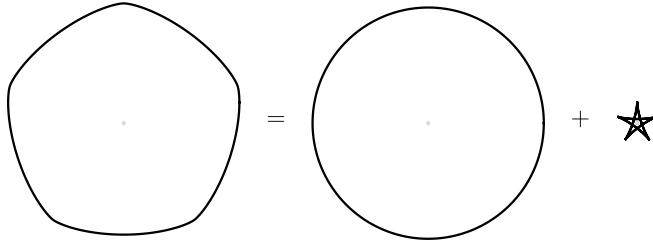


FIG. 4. Illustration of the decomposition of a convex shape into its hull and midpoint curve. The projected position function $f(\varphi) = 1 + \sin(5\varphi)/28$ produces a rounded pentagon of constant width (left). Its parity-even part $d(\varphi) = 2$ defines a circle (middle), while its parity-odd part $C(\varphi) = \sin(5\varphi)/28$ defines a sinestar (right). We may therefore regard the rounded pentagon as the sum of a circle and a sinestar.

the context of interferometry, where d is encoded in the visibility amplitude, while C appears as an overall phase (see Eq. (3) in Ref. [13]).

It is often convenient to regard $d(\varphi)$ and $C(\varphi)$ as being defined on the smaller range $\varphi \in [0, \pi]$. That is, we may view a closed, convex curve as being specified by a single function $f(\varphi)$ ranging over the full circle $\varphi \in [0, 2\pi]$, or equivalently, it can be repackaged into two functions $d(\varphi)$ and $C(\varphi)$ ranging over $\varphi \in [0, \pi]$.

Under shifts of the origin $\mathbf{r} \rightarrow \mathbf{r} + (X, Y)$, we have

$$d(\varphi) \rightarrow d(\varphi), \quad (26)$$

$$C(\varphi) \rightarrow C(\varphi) + X \cos \varphi + Y \sin \varphi. \quad (27)$$

That is, all the coordinate origin information is contained in $C(\varphi)$, while $d(\varphi)$ remains invariant under translations, as required by its interpretation as the projected diameter. To study shapes intrinsically, one might fix a canonical choice of C by demanding that the dipole vanish, $\int C(\varphi) e^{\pm i\varphi} d\varphi = 0$. However, for interferometric measurements tracking absolute phase, the coordinate origin is “known” and the dipole of C contains important information about where the curve is located on the image.

We can view a measurement of $d(\varphi)$ as a measurement of the hull of the underlying curve. The remaining freedom is a choice of closed plane curve made of individually convex segments joined together at cusps.

IV. EXAMPLES

We now illustrate the framework of the previous section (Sec. III) with a series of examples.

A. The point

The most trivial example is that of a curve that has degenerated to a single point (X, Y) . Its projected position function is the pure dipole

$$f_{\text{point}} = X \cos \varphi + Y \sin \varphi. \quad (28)$$

This form may be added to any other $f(\varphi)$ to induce a translation. It can also be used to add sharp corners to a shape by demanding that it hold exactly over some finite range of φ . The parameterization then “hovers” at this point for that lapse of φ , during which time the normal vector still advances, creating a discontinuity in the normal (and tangent) vector of the shape. The next example will illustrate this behavior.

B. Curves of constant width

Curves with $d'(\varphi) = 0$ are generally called curves of constant width. These are the curves whose hull is a circle. The simplest example is the circle itself,

$$f_{\text{circle}} = R, \quad (29)$$

where R is the radius. In this framework, one can easily construct other curves of constant width by adding an arbitrary function $C(\varphi)$ that is parity-odd,

$$f_{\text{c.w.}} = R + C(\varphi). \quad (30)$$

For example, choosing $C(\varphi) = A \sin 3\varphi$ with sufficiently small A creates a triangular object with rounded sides and rounded corners. Similarly, $C(\varphi) = A \sin 4\varphi$ creates a rounded square, and so on. The rounded pentagon with $C(\varphi) = A \sin 5\varphi$ is shown in Fig. 4.

The most famous example of a non-circular constant-width curve is perhaps the Reuleaux triangle, which is composed of three circular arcs joined at 120° angles. This is a slightly subtle example in our formalism, as it contains sharp corners. These are handled by using the point form (28) over the lapses of φ corresponding to the jumps in $\hat{\mathbf{n}}$ at the corners. The remainder of the shape consists of three circular arcs, each with curvature center at the opposite vertex of the triangle. These are constructed using appropriate segments of the shifted circles $f = D + X \cos \varphi + Y \sin \varphi$ with appropriately chosen centers (X, Y) . Choosing the geometric center of the Reuleaux triangle to be the origin and letting the top vertex lie on the positive y -axis, the projected position function is

$$f_{\text{Rx}} = D \begin{cases} 1 - \frac{1}{2\sqrt{3}} \sin \varphi - \frac{1}{2} \cos \varphi & 0 \leq \varphi \leq \frac{\pi}{3}, \\ \frac{1}{\sqrt{3}} \sin \varphi & \frac{\pi}{3} \leq \varphi \leq \frac{2\pi}{3}, \\ 1 - \frac{1}{2\sqrt{3}} \sin \varphi + \frac{1}{2} \cos \varphi & \frac{2\pi}{3} \leq \varphi \leq \pi, \\ -\frac{1}{2\sqrt{3}} \sin \varphi - \frac{1}{2} \cos \varphi & \pi \leq \varphi \leq \frac{4\pi}{3}, \\ 1 + \frac{\sin \varphi}{\sqrt{3}} & \frac{4\pi}{3} \leq \varphi \leq \frac{5\pi}{3}, \\ \frac{1}{2} \cos \varphi - \frac{1}{2\sqrt{3}} \sin \varphi & \frac{5\pi}{3} \leq \varphi \leq 2\pi. \end{cases} \quad (31)$$

Although complex to write down, this function has a simple behavior, oscillating continuously (but not smoothly) between maxima and minima in a manner reminiscent of a pure sinusoid (Fig. 5). It can be very closely approximated by a function of the form $A + B \sin 3\varphi$, but it is not exactly of this form.

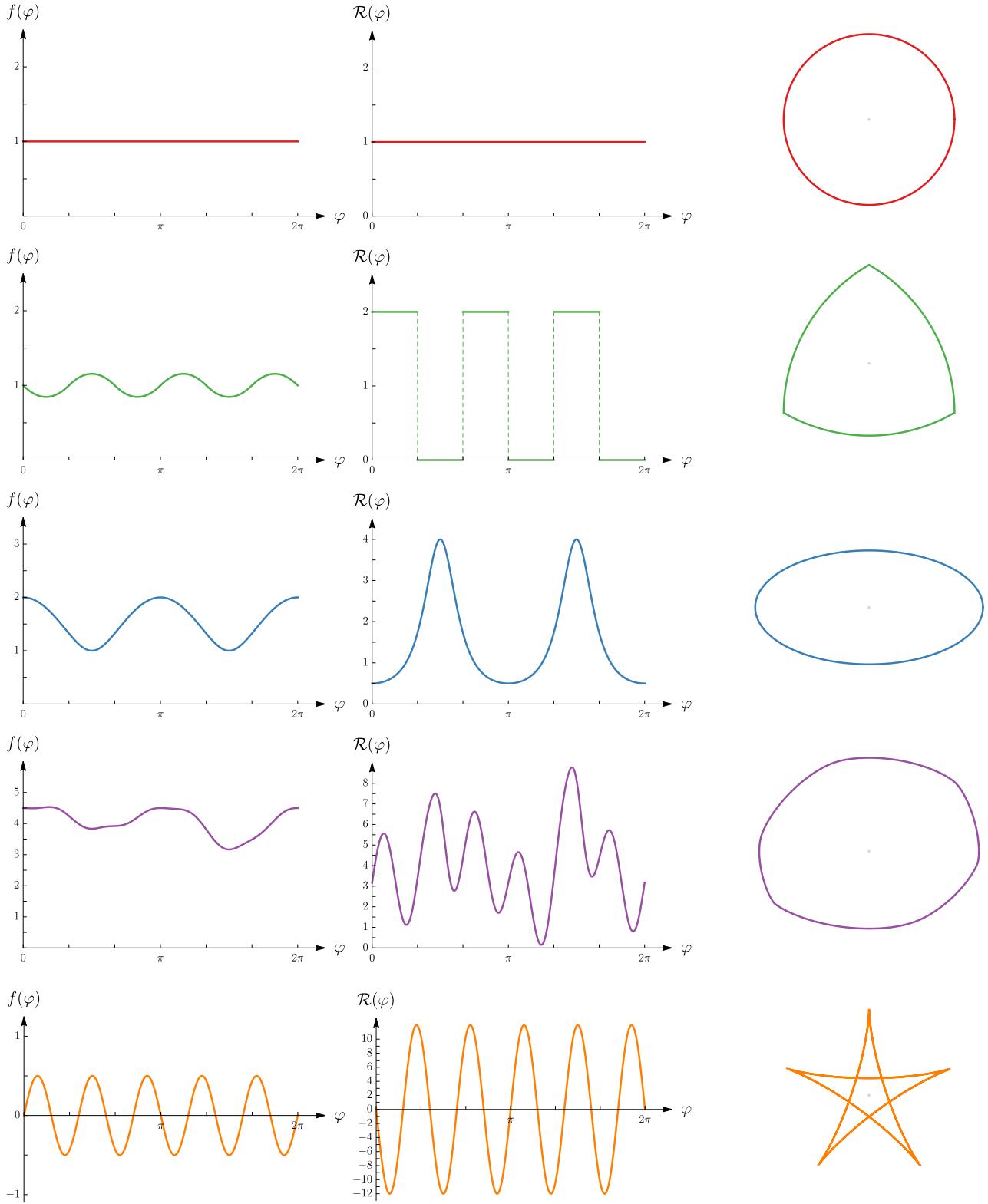


FIG. 5. Illustration of the relationship between the projected position function $f(\varphi)$ (left) and its associated curve (right). We also show the function $\mathcal{R}(\varphi)$ (middle), whose absolute value is the radius of curvature. When this function passes through zero, the curve develops a cusp. (The star shown in the bottom-right is traced twice as the parameter ranges over the full circle $\varphi \in [0, 2\pi]$ and hence has only five cusps, despite the ten zero-crossings of \mathcal{R} .) From top to bottom, we show the circle [Eq. (29) with $R = 1$], the Reuleaux triangle [Eq. (31) with $D = 2$], the ellipse [Eq. (38) with $R_1 = 2$ and $R_2 = 1$], the pebble [Eq. (43)], and the five-pointed sinestral $f(\varphi) = \sin 5\varphi$.

C. Ellipse

The parameterization of the ellipse by polar angle θ is

$$\mathbf{R}(\theta) = (R_1 \cos \theta, R_2 \sin \theta), \quad (32)$$

where R_1 and R_2 are the semi-axis lengths in the x and y directions. The normal angle $\varphi(\theta)$ is found from Eq. (18), which becomes

$$\tan \varphi = \frac{R_1}{R_2} \tan \theta. \quad (33)$$

This relation is simple enough that it can be inverted as

$$\cos \theta = \frac{R_1 \cos \varphi}{\sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}}, \quad (34)$$

$$\sin \theta = \frac{R_2 \sin \varphi}{\sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}}. \quad (35)$$

Plugging into Eq. (32) with $\mathbf{r}(\varphi) = \mathbf{R}(\theta(\varphi))$ then yields the parameterization of the ellipse by normal angle,

$$x(\varphi) = \frac{R_1^2 \cos \varphi}{\sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}}, \quad (36)$$

$$y(\varphi) = \frac{R_2^2 \sin \varphi}{\sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}}. \quad (37)$$

The projected position is calculated from Eq. (14) as

$$f_{\text{ellipse}} = \sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}, \quad (38)$$

which is invariant under antipodal exchange $\varphi \rightarrow \varphi + \pi$. Hence, we find

$$d_{\text{ellipse}} = 2\sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}, \quad C_{\text{ellipse}} = 0. \quad (39)$$

That is, the ellipse has antipodal symmetry.

D. Circlipse

Adding the projected position functions of two shapes produces a third shape corresponding to the vector sum of points that share a common normal angle. A simple nontrivial example is the sum of a circle and an ellipse, or a “circlipse”:

$$f_{\text{circlipse}} = R_0 + \sqrt{R_1^2 \cos^2 \varphi + R_2^2 \sin^2 \varphi}. \quad (40)$$

This is an antipodally symmetric shape with an oval form. The semi-axes in the x and y directions are $R_0 + |R_1|$ and $R_0 + |R_2|$, respectively. As discussed in Sec. V below, the hull of the Kerr critical curve is closely approximated by a circlipse over the entire parameter

space (see also Fig. 6). We are unaware of previous discussion of this shape, which arose in our studies of the Kerr critical curve. Using the method of implicitization reviewed in App. B, we find that the circlipse is part of the vanishing locus of an order 8 polynomial, which does not appear to correspond to previously studied octic curves.

When $R_0 = 0$, the circlipse reduces to an ellipse, and when $R_1 = R_2 = 0$, it becomes a circle. If only one of R_1 or R_2 vanishes, then we obtain a singular limit that produces a “racetrack” shape,

$$f_{\text{racetrack}} = R_0 + R_2 |\cos \varphi|. \quad (41)$$

This function produces two half-circles of radius R_0 separated by a distance of $2R_2$, constructed by cutting a circle vertically through its center and separating the pieces horizontally. Joining the two with straight lines produces a racetrack shape. The circlipse (40) is closed at any non-zero R_1 , meaning that its $R_1 \rightarrow 0$ limit is indeed the full closed racetrack. We may also take Eq. (41) to represent the full closed curve if we adopt the natural convention that kinks in $f(\varphi)$ are to represent straight lines.

E. A cuspy triangle

Whereas studying $d(\varphi)$ for the Kerr critical curve led us to consider the circlipse, studying $C(\varphi)$ led us to

$$f_{\text{cuspy triangle}} = \arcsin(\chi \cos \varphi), \quad (42)$$

where $\chi \in [-1, 1]$. This shape is a cuspy triangle, ranging from small and equilateral as $\chi \rightarrow 0$ to finite and isosceles as $|\chi| \rightarrow 1$. The limiting case $|\chi| = 1$ contains a straight line encoded by kinks in f at $\varphi \bmod \pi = 0$. This parameterization also moves the triangle rightward with increasing χ . This cuspy triangle closely approximates the midpoint curve of the Kerr critical curve (Fig. 6 bottom), with the parameter χ an increasing function of dimensionless spin a/M .

F. Pebble

As a final example we consider the function

$$f_{\text{pebble}} = 3 + \sqrt{(3/2)^2 \cos^2 \varphi + (1/2)^2 \sin^2 \varphi} + \frac{1}{4} \sin^3 2\varphi + \frac{1}{3} \sin^3 \varphi, \quad (43)$$

which produces an irregularly shaped “pebble” (Fig. 5).

V. THE KERR CRITICAL CURVE

The Kerr critical curve [1] is a theoretical closed curve on the image plane of a camera aimed at a black hole. It

is defined by the asymptotic arrival positions of photons that orbit the black hole arbitrarily many times before escaping to the camera. The shape depends on the black hole spin parameter a and the observer inclination θ_o relative to the spin axis, with the overall size set by the black hole mass M . Successive images of bulk matter emission asymptote to the critical curve (e.g., Ref. [6]). In this section, we study the critical curve shape as encoded by its interferometric observables. We discuss general features of the curve, derive analytic results in various limits, and give analytic formulae that faithfully approximate the shape over the entire parameter space.

A. General Properties

The original work of Bardeen [1] provides a parametric formula for the critical curve. In the notation of Refs. [6, 14], Bardeen's formula is

$$\alpha(\tilde{r}) = -\frac{\lambda(\tilde{r})}{\sin \theta_o}, \quad (44a)$$

$$\beta_{\pm}(\tilde{r}) = \pm \sqrt{\eta(\tilde{r}) + a^2 \cos^2 \theta_o - \lambda(\tilde{r})^2 \cot^2 \theta_o}, \quad (44b)$$

with

$$\lambda(\tilde{r}) = a + \frac{\tilde{r}}{a} \left[\tilde{r} - \frac{2(\tilde{r}^2 - 2M\tilde{r} + a^2)}{\tilde{r} - M} \right], \quad (45a)$$

$$\eta(\tilde{r}) = \frac{\tilde{r}^3}{a^2} \left[\frac{4M(\tilde{r}^2 - 2M\tilde{r} + a^2)}{(\tilde{r} - M)^2} - \tilde{r} \right]. \quad (45b)$$

Here, α and β are Cartesian “screen coordinates” (with units of M) describing the image. We set $G = c = 1$ and use the range $0 < a < M$, treating the edges as limits. The curve is parameterized by \tilde{r} in two separate segments \pm , corresponding to its upper-half (+) and lower-half (−). The orientation is clockwise in the upper-half plane and counter-clockwise in the lower-half plane (see, e.g., Fig. 3 in Ref. [6]).

The parameter \tilde{r} represents the radius at which photons orbit before reaching the detector. The allowed range is the subset of the region $\tilde{r} \in [M, 4M]$ for which β remains real. (Allowing \tilde{r} to take any values defines a larger object, the *critical locus*, that we study in App. C.) This range shrinks to zero at vanishing spin and/or inclination, where all photons orbit at the same radius. Thus, this parameterization can only be used at non-zero spin and inclination.

The normal angle φ is determined from Eq. (18) using Eqs. (44). After some simplification, we find that

$$\tan \varphi(\tilde{r}) = \frac{\beta(\tilde{r})}{\alpha(\tilde{r}) - a \sin \theta_o \left(\frac{\tilde{r}+M}{\tilde{r}-M} \right)}. \quad (46)$$

Inverting for $\tilde{r}(\varphi)$ would provide $f(\varphi)$ analytically, but unfortunately, this requires solving a sextic polynomial.

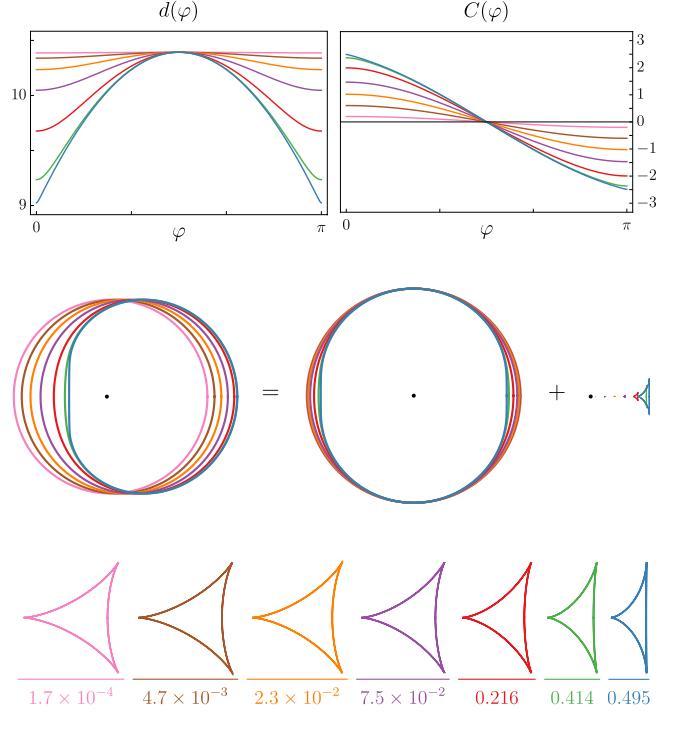


FIG. 6. The shape decomposition of the critical curve. The hull is closely approximated by a circlipse [Eq. (40)], while the midpoint curve is closely approximated by the cuspy triangle (42). The bottom panel shows a blow-up of these triangles with scale indicated in units of M . We have set $M = 1$ and chosen an equatorial observer $\theta_o = \pi/2$ with black hole spins 10%, 30%, 50%, 70%, 90%, 99%, 99.99%.

Nonetheless, it is straightforward to use $\varphi(\tilde{r})$ to make parametric plots of $f(\varphi)$. Defining $\arctan(x, y) \in (-\pi, \pi]$ to be the usual principal argument of the complex number $x + iy$, we have

$$\varphi_{\pm}(\tilde{r}) = \arctan \left(\alpha(\tilde{r}) - a \sin \theta_o \left(\frac{\tilde{r}+M}{\tilde{r}-M} \right), \beta_{\pm}(\tilde{r}) \right), \quad (47)$$

where φ_{\pm} is the normal angle in the upper-half (+) and lower-half (−) planes. This formula holds modulo 2π ; to obtain our canonical range $\varphi \in [0, 2\pi]$, one must add 2π to φ_- .

Together, Eqs. (14), (44), and (47) imply that $f(\varphi)$ can be plotted as the two-segment parametric curve $(\varphi_{\pm}(\tilde{r}), f(\varphi_{\pm}(\tilde{r})))$. However, particular values of $f(\varphi)$ must still be extracted numerically (say, by a graphical method using the plot data). Similarly, the decomposition into $d(\varphi)$ and $C(\varphi)$ requires a numerical method.

In Fig. 6, we illustrate the properties of the Kerr critical curve in the language of this paper. We pick the equatorial observer $\theta_o = \pi/2$ for which the shape variation with spin is greatest. Similar patterns are seen at smaller inclination angles, albeit with less overall variation.

In studying the projected position function of the critical curve, we have found some simple functional forms that approximate it to remarkably high accuracy. In particular, the even part $d(\varphi)$ is well-described by the circlipse form shown above in Eq. (40), while the odd part $C(\varphi)$ is well-described by the cuspy triangle (42) together with a horizontal translation (28). The precise critical curve is therefore well-described by the projected position function

$$f(\varphi) = R_0 + \sqrt{R_1^2 \sin^2 \varphi + R_2^2 \cos^2 \varphi} + (X - \chi) \cos \varphi + \arcsin(\chi \cos \varphi). \quad (48)$$

We will call this a *phoval*, for “photon ring oval”.

We have chosen the translation parameter X such that $X \cos \varphi$ is the dipole term in the Fourier cosine series of this function. The intrinsic shape of the phoval has four parameters, namely: three nonnegative radii R_0, R_1, R_2 , together with an asymmetry parameter $-1 \leq \chi \leq 1$. Provided the parameters are chosen such that the shape is convex, the phoval has two preferred axes that intersect it orthogonally. In the form (48) with $X = 0$, these are the x and y axes. The horizontal radius is $R_0 + R_1$ and the vertical radius is $R_0 + R_2$. The relative size of R_1 and R_2 determines how flattened or rounded the associated edges are, and the parameter χ introduces a horizontal asymmetry in this roundedness.

The phoval provides an excellent fit to the Kerr critical curve. For each choice of black hole spin and observer inclination, we plot the critical curve projected position function $f(\varphi)$ as described above. We then fit the functional form (48) to the plot data, finding best-fit parameters and normalized root-mean-square (RMS) residuals, defined as the average of the squared-deviation divided by the span, i.e., $\langle (\delta f)^2 \rangle / (f_{\max} - f_{\min})$. Repeating this procedure over the whole parameter space, sampled uniformly in spin $a \in (0, M)$ and inclination $\theta_o \in (0, \pi/2]$ (in practice, we start at values very close to the edges of these ranges), we find a median normalized RMS deviation of 10^{-5} . The largest residuals are a few times 10^{-3} , occurring in the extremal limit $a \rightarrow M$. (We have computed for a as large as $a = .9999M$.) That is, the fit works to a part in 10^5 over the vast majority of the parameter space, and to a part in 10^3 near extremality.

Very recently, Ref. [15] showed that the critical curve may also be approximated to great accuracy by a limaçon curve, building on an older observation of Ref. [16]. A direct comparison of the fit quality is difficult, since the fit diagnostics differ. In particular, our diagnostic is tied to the interferometric signature $f(\varphi)$, whereas theirs is tied to the radial distance from an origin on the image plane.

B. Expansion in small spin or inclination

At zero spin or inclination, the critical curve is precisely circular, as required by symmetry. It remains cir-

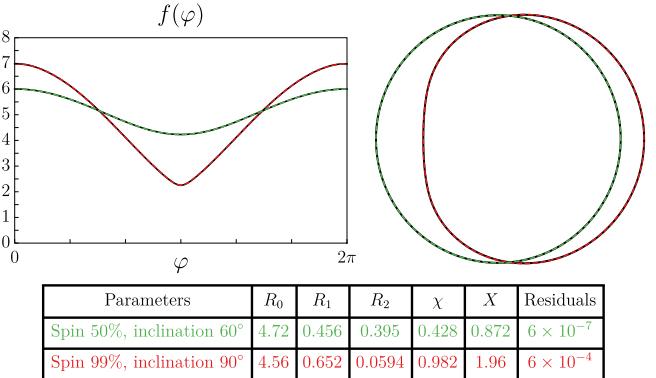


FIG. 7. The Kerr critical curve as a phoval. We show two examples of fitting the phoval shape (dashed curves) to critical curve (solid). The resulting phoval is visually indistinguishable from the critical curve, both in terms of the projected position function (left) and the image plane curve (right). The parameters and normalized RMS residuals are displayed in the table. We have set $M = 1$ in this figure.

cular up to second order in spin, when it takes the shape of an ellipse [17]. Here, we reproduce this small-spin result, correcting some errors in the formulas for the ellipse parameters. We also provide the analogous result at small inclination, proving that the critical curve is also an ellipse at second order in inclination (for any spin), and finding closed-form expressions for its parameters.

As discussed below Eqs. (45), the parameter \tilde{r} breaks down in the limit of small spin or small inclination. Studying these limits requires an alternative expression of the critical curve. We can eliminate the problematic parameter by using the formula

$$\tilde{r} = M + 2M\Delta \cos \left[\frac{1}{3} \arccos \left(\frac{1 - \frac{a^2}{M^2}}{\Delta^3} \right) \right],$$

$$\Delta = \sqrt{1 - \frac{a(a - \alpha \sin \theta_o)}{3M^2}}, \quad (49)$$

which follows from the $k = 0$ case³ of Eq. (122) in Ref. [14] [wherein $\tilde{\lambda} = \lambda(\tilde{r})$], along with Eq. (44a) above. This expresses \tilde{r} in terms of α , which is one of the screen coordinates. Since the curve itself remains finite in the limit, the range of α must similarly remain finite, indicating that it can be used as a parameter in these limits.

That is, we may now discuss the critical curve as a pair of ordinary functions $\beta_{\pm}(\alpha)$. Combining Eqs. (44a) and (44b) gives

$$\beta_{\pm}(\alpha) = \pm \sqrt{\eta(\tilde{r}(\alpha)) + (a^2 - \alpha^2) \cos^2 \theta_o}, \quad (50)$$

³The other inversions $k = 1$ and $k = 2$, reproduced in Eq. (C2) below, are not needed to recover the critical curve at sufficiently small spin or inclination. See Fig. 1 of Ref. [6] and Eqs. (123)–(125) of the revised arXiv version of Ref. [14].

with $\tilde{r}(\alpha)$ provided by Eq. (49). Another way of organizing this information is to write

$$\alpha^2 + \beta^2 = a^2 \cos^2 \theta_o + F(a, \alpha \sin \theta_o), \quad (51)$$

where

$$F = \frac{\tilde{r}^3}{a^2} \left[\frac{4M(\tilde{r}^2 - 2M\tilde{r} + a^2)}{(\tilde{r} - M)^2} - \tilde{r} \right] + \alpha^2 \sin^2 \theta_o, \quad (52)$$

with \tilde{r} given by Eq. (49). In regarding F as a function of a and $\alpha \sin \theta_o$, we are suppressing dependence on the overall scale M .

This function F has a regular expansion at small spin (for any inclination) as well as at small inclination (for any spin). We may imagine expanding to some fixed order in either parameter and determining the critical curve from Eq. (51). The functional form of (51) shows immediately that, at quadratic order in either small parameter, the critical curve is determined by the roots of a quadratic polynomial in α and β . This is on general grounds a conic section, and since the critical curve is closed, it must be an ellipse. That is, Eq. (51) immediately shows that the critical curve is an ellipse at small spin and/or inclination.

It is rather straightforward to extract the parameters of the ellipse by performing the expansion in F . When expanding in spin, we find

$$F = 27M^2 + (4\alpha \sin \theta_o)a - \left[\frac{\alpha^2 \sin^2 \theta_o}{9M^2} + 4 \right] a^2 + \mathcal{O}(a^3). \quad (53)$$

Let us first consider the shape at linear order. Plugging into Eq. (51), we obtain

$$(\alpha - 2a \sin \theta_o)^2 + \beta^2 = 27M^2 + \mathcal{O}(a^2), \quad (54)$$

demonstrating that the curve is a circle of radius $3\sqrt{3}M$ centered at $\alpha = 2a \sin \theta_o$. At quadratic order, we find

$$\left(\frac{\alpha - 2a \sin \theta_o}{1 - \frac{a^2 \sin^2 \theta_o}{18M^2}} \right)^2 + \beta^2 = 27M^2 - 3a^2 \cos^2 \theta_o + \mathcal{O}(a^3). \quad (55)$$

This can be put in the canonical form of an ellipse,

$$\left(\frac{\alpha - \alpha_0}{R_1} \right)^2 + \left(\frac{\beta}{R_2} \right)^2 = 1, \quad (56)$$

with

$$\alpha_0 = 2a \sin \theta_o + \mathcal{O}(a^3), \quad (57)$$

$$R_1 = 3\sqrt{3}M \left(1 - \frac{a^2}{18M^2} \right) + \mathcal{O}(a^3), \quad (58)$$

$$R_2 = 3\sqrt{3}M \left(1 - \frac{a^2 \cos^2 \theta_o}{18M^2} \right) + \mathcal{O}(a^3). \quad (59)$$

The shape of the critical curve of a slowly spinning black hole was previously analyzed in Ref. [17]. Our Eq. (57) is in agreement with their Eq. (22), but our Eqs. (58) and (59) differ from their Eqs. (23) and (24) by some simple factors. We have checked numerically that our results are correct.

The ellipse parameters at small inclination (for any spin) can be found by the same method: expand F to quadratic order in $\alpha \sin \theta_o$ and compare to the canonical form. This produces rather unwieldy analytic formulas that we do not display here. In App. C, we find somewhat less unwieldy formulas using an algebraic geometry approach. In particular, we are able to write the ellipse parameters as rational functions of the spin a and the radius $\tilde{b}(a)$ of the zero-inclination critical curve [see Eqs. (C27)–(C33) below].

C. Extremal limit

Finally, we consider the extremal ($a \rightarrow M$) limit of the critical curve. In this case, the curve picks up a straight line (the “NHEKline”) that can be attributed to the near-horizon geometry [1, 18]. This straight line means that our description in terms of functions $\beta_{\pm}(\alpha)$ breaks down as $a \rightarrow M$. However, it is still instructive to analyze the portion of the curve for which it remains valid. This will lead us to the result that the extremal critical curve is the convex hull of a pair of Cartesian ovals.

As noted in the previous section, the substitution of $\tilde{r}(\alpha)$ [Eq. (49)] into $\beta_{\pm}(\tilde{r})$ [Eq. (44)] provides a valid parameterization $\beta_{\pm}(\alpha)$ of the entire critical curve only for sufficiently small spin and inclination. Above a certain threshold in these parameters, we recover only a part of the critical curve with this substitution (with the other portion requiring another inversion of $\lambda(\tilde{r})$ given in Eq. (C2) below). We will see that in the extremal limit $a \rightarrow M$ (for any inclination), this portion gives precisely the non-straight part, i.e., the curve minus its NHEKline. Thus, to study this portion, we set $a = M$ in Eqs. (49) and (52), finding the comparatively simple expressions

$$\tilde{r} = M \left(1 + \sqrt{2 + \frac{\alpha \sin \theta_o}{M}} \right), \quad (60)$$

and

$$F = \alpha^2 \sin^2 \theta_o - \frac{\tilde{r}^3(\tilde{r} - 4M)}{M^2}, \quad (61)$$

so that Eq. (51) becomes

$$(\alpha - M \sin \theta_o)^2 + \beta^2 - 12M^2 = 8M^2 \sqrt{2 + \frac{\alpha \sin \theta_o}{M}}. \quad (62)$$

This curve fails to close for $\sin \theta_o > \sqrt{3} - 1$. Comparing with Eq. (A6) of Ref. [18], we see that the missing segment is precisely the NHEKline $\alpha = -2M \csc \theta_o$ with

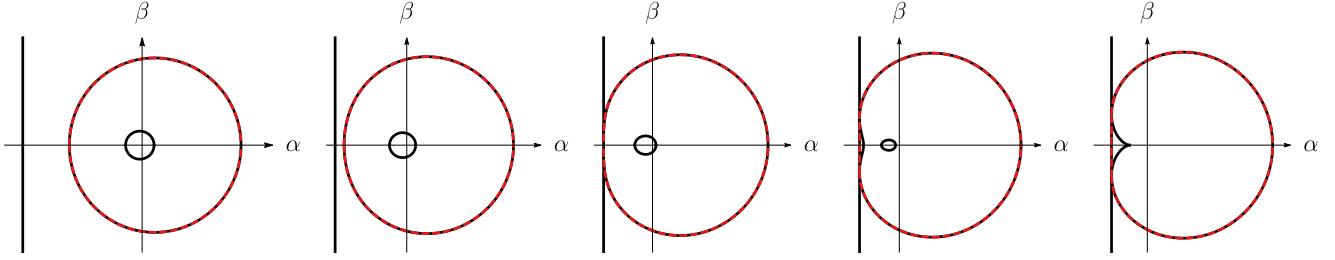


FIG. 8. The extremal critical locus (black) consists of a vertical line together with a pair of Cartesian ovals. The physical critical curve (dashed, red) is a subset of this locus that can be succinctly described as the convex hull of the ovals. At small observer inclinations, the ovals are nearly circular and concentric about the origin, and the line is far away to the left. As the inclination is increased ($\sin \theta_o = 0.3, 0.5, \sqrt{3} - 1 \approx 0.732, 0.9, 1$ from left to right), the line moves towards the larger oval, which flattens to meet it; the smaller oval moves leftward. The line touches the oval at the critical inclination $\sin \theta_o = \sqrt{3} - 1$, where the NHEK spacetime becomes visible [18]. After the kiss, the line remains attached to the oval, which develops a dimple as it smiles in response. The inner oval continues to move left, eventually merging with the dimple at the cusp of a cardioid.

endpoints $\beta_{\pm} = \pm M \sqrt{3 + \cos^2 \theta_o - 4 \cot^2 \theta_o}$, as claimed above. To elucidate the shape of the extremal critical curve, we may square both sides of Eq. (62) to obtain a quartic equation $\mathcal{E} = 0$, with

$$\mathcal{E} = \left[(\alpha - M \sin \theta_o)^2 + \beta^2 - 12M^2 \right]^2 - 64M^3(2M + \alpha \sin \theta_o). \quad (63)$$

After a simple translation in α to Cartesian coordinates $(x, y) = (\alpha - M \sin \theta_o, \beta)$, we recognize \mathcal{E} to be the defining equation for a pair of Cartesian ovals,

$$(x^2 + y^2)^2 + k(x^2 + y^2) + lx + m = 0, \quad (64)$$

with parameters $k = -24M^2$, $l = -64M^3 \sin \theta_o$, and $m = 16M^4(1 - 4 \sin^2 \theta_o)$.⁴ Squaring Eq. (62) introduced an unphysical curve in Eq. (63): solving for β gives

$$\beta^2 = 12M^2 - (\alpha - M \sin \theta)^2 \pm 8M^{3/2} \sqrt{2M + \alpha \sin \theta}, \quad (65)$$

with only the plus sign corresponding to the physical critical curve (62) (more precisely, its non-straight portion).

The full set of solutions (65) describes a pair of Cartesian ovals. When $\sin \theta_o \leq \sqrt{3} - 1$, the ovals are both convex and the outer one is the full critical curve. When $\sin \theta_o > \sqrt{3} - 1$, the outer oval becomes non-convex, and the critical curve is formed by adjoining the NHEKline to the portion of the oval described by Eq. (62). That is, at any observer inclination, the extremal critical curve is the convex hull of a Cartesian oval.

It is entertaining to ponder how a shape first studied by Descartes in 1637 is thereby embedded into the Kerr metric of general relativity.

In App. C, we recover the extremal critical curve in a different way, by viewing it as a subset of a larger “critical locus”. The critical locus includes additional curves corresponding to unphysical values of the parameter \tilde{r} , but as a result can be described as an algebraic variety (in this case, the vanishing locus of a single polynomial). The extremal critical locus is composed of the full pair of Cartesian ovals (63) together with a straight line, as illustrated in Fig. 8.⁵

D. Extremal, equatorial shape

Recently, Ref. [15] noted that the extremal critical curve of an equatorial observer is the convex hull of a cardioid. This fact is apparent in the rightmost panel of Fig. 8, given that the fully degenerate case of a Cartesian oval is known to be a cardioid. We may also see it algebraically by noting that for $\theta_o = \pi/2$, the polynomial (63) defining the Cartesian ovals can be recast in the canonical form of a cardioid,

$$(X^2 + Y^2)^2 - 4AX(X^2 + Y^2) = 4A^2Y^2, \quad (66)$$

with $X = \alpha + M$, $Y = \beta$, and $A = 2M$.

The normal-angle parameterization for the convex hull of the cardioid (i.e., the extremal, equatorial critical curve) can be computed in closed form. We begin with the parameterization of the full cardioid (66) by polar

⁴Under a different translation, it is also possible to put \mathcal{E} in the form $[(1 - m^2)(X^2 + Y^2) + 2m^2cX + A^2 - m^2c^2]^2 = 4A^2(X^2 + Y^2)$ that is more traditionally given for the Cartesian oval, at the cost of complicated expressions for the parameters m , c , and A .

⁵The critical locus at generic spin does not contain a Cartesian oval. This follows from the fact demonstrated in App. C that its defining polynomial \mathcal{C}_o is of the form $\prod_{i=1}^3 [\beta^2 - f_i(\alpha)]$, with none of the f_i a polynomial [see Eqs. (C8), (C13), (C16)]. At extremality, one of these factors becomes a quadratic describing a simple line (with multiplicity two), implying that the other two factors necessarily multiply to form a quartic polynomial—the Cartesian ovals (63).

angle $\theta \in [0, 2\pi]$,

$$\alpha(\theta) = M(1 - 4 \cos \theta + 2 \cos 2\theta), \quad (67a)$$

$$\beta(\theta) = M(-4 \sin \theta + 2 \sin 2\theta). \quad (67b)$$

The angle $\varphi(\theta)$ is found from Eq. (18), which becomes

$$\tan \varphi = -\cot \frac{3\theta}{2}. \quad (68)$$

This relation is simple enough that it can be inverted as

$$\theta_I(\varphi) = \frac{2\varphi + (2I+1)\pi}{3}, \quad I \in \mathbb{Z}. \quad (69)$$

Since the cardioid is not a closed, convex curve, we must resort to the general framework of Sec. II to describe its shape. Plugging into Eqs. (67) with $\mathbf{r}(\varphi) = \mathbf{R}(\theta_I(\varphi))$ then yields the parameterization by $\varphi \in [0, \pi]$ of the edge-on extremal cardioid in three segments with $I \in \{-1, 0, 1\}$,

$$\alpha_I(\varphi) = \alpha(\theta_I(\varphi)), \quad \beta_I(\varphi) = \beta(\theta_I(\varphi)). \quad (70)$$

Equivalently, these segments can be stitched into a single curve with $\varphi \in [-\pi, 2\pi]$,

$$\alpha(\varphi) = M \left(1 + 4 \cos \frac{2\varphi}{3} + 2 \cos \frac{4\varphi}{3} \right), \quad (71a)$$

$$\beta(\varphi) = M \left(4 \sin \frac{2\varphi}{3} + 2 \sin \frac{4\varphi}{3} \right). \quad (71b)$$

Discarding the $I = -1$ piece results in an open, convex curve consisting of two segments $I = 1$ (“top”) and $I = 0$ (“bottom”), whose convex hull is the critical curve. This is equivalent to restricting the parameterization (71) to the circle $\varphi \in [-\pi, \pi]$.

The projected position on the range $\varphi \in [-\pi, \pi]$ is then calculated from Eq. (14) as

$$f(\varphi) = M \left(\cos \varphi + 6 \cos \frac{\varphi}{3} \right). \quad (72)$$

By making this function 2π -periodic, we can extend it to our canonical range $\varphi \in [0, 2\pi]$, on which it becomes

$$f(\varphi) = M \cos \varphi + \begin{cases} 6M \cos \frac{\varphi}{3}, & 0 \leq \varphi \leq \pi, \\ -6M \cos \frac{\varphi+\pi}{3} & \pi \leq \varphi \leq 2\pi. \end{cases} \quad (73)$$

This formula gives the exact projected position function of the extremal, equatorial critical curve. The kink at $\varphi = \pi$ represents the NHEKline.

E. Projected position for equatorial critical curves

We can in fact analytically compute the projected position of the critical curve for any equatorial observer, not just at extremal spin. This is possible because the sextic polynomial discussed below Eq. (46) degenerates to a cubic when $\theta_o = \pi/2$.

More explicitly, when $\theta_o = \pi/2$, Eq. (46) reduces to

$$\cos^2 \varphi(\tilde{r}) = \frac{\tilde{r}(\tilde{r} - 3M)^2}{4Ma^2}, \quad (74)$$

which is a cubic equation in \tilde{r} with positive discriminant

$$\Delta_3 = \frac{3^3 \cos^2 \varphi}{2^4 a^6} \left(1 - \frac{a^2}{M^2} \cos^2 \varphi \right) > 0. \quad (75)$$

Hence, it has three real roots given for $k \in \{0, 1, 2\}$ by

$$\tilde{r}_{(k)}(\varphi) = 3M + \frac{a \cos \varphi}{\cos \left[\frac{1}{3} \arccos \left(\frac{a}{M} \cos \varphi \right) - \frac{2\pi k}{3} \right]}. \quad (76)$$

Plugging the root $\tilde{r}_{(0)}(\varphi)$ into Eqs. (44) yields the parameterization $(\alpha(\varphi), \beta_{\pm}(\varphi))$ by normal angle φ , where one is to choose $+$ for $\varphi \in [0, \pi)$ and $-$ for $\varphi \in [\pi, 2\pi)$.⁶ The projected position is then calculated from Eq. (14) as

$$f(\varphi) = \alpha(\tilde{r}_{(0)}(\varphi)) \cos \varphi + \beta(\tilde{r}_{(0)}(\varphi)) \sin \varphi, \quad (77)$$

with the choice of sign \pm in $\beta(\tilde{r})$ given by $\text{sign}(\pi - \varphi)$.

ACKNOWLEDGEMENTS

This work was supported in part by NSF grant PHY-1752809 to the University of Arizona. AL acknowledges the Jacob Goldfield Foundation.

Appendix A: Perimeter

The Cauchy surface area theorem states (in two dimensions) that the perimeter of a closed, convex curve is the average of its projected diameters. The framework of Sec. III provides an elegant proof of this fact. Using Eq. (16), we have

$$P = \int_0^{2\pi} \sqrt{x'(\varphi)^2 + y'(\varphi)^2} \, d\varphi \quad (A1)$$

$$= \int_0^{2\pi} |f(\varphi) + f''(\varphi)| \, d\varphi. \quad (A2)$$

For a closed, convex curve, the quantity $f(\varphi) + f''(\varphi)$ is the radius of curvature [Eq. (15)], which is nonnegative. We therefore have

$$P = \int_0^{2\pi} [f(\varphi) + f''(\varphi)] \, d\varphi \quad (A3)$$

$$= \int_0^{2\pi} f(\varphi) \, d\varphi. \quad (A4)$$

⁶On the range $\varphi \in [0, \pi)$, the root $\tilde{r}_{(2)}(\varphi)$ yields a parameterization with the opposite (clockwise) orientation, while the root $\tilde{r}_{(1)}(\varphi)$ traces twice over the upper half of the curve defined by mapping the (λ, η) -space curve \mathcal{C}_- [14] onto the image plane via Eqs. (44).

Under the decomposition (23), the parity-odd piece $C(\varphi)$ does not contribute to the integral, while the parity-even piece contributes

$$P = \int_0^\pi d(\varphi) d\varphi. \quad (\text{A5})$$

Thus, the perimeter of a closed, convex curve is the integral of the projected diameter. When $d(\varphi)$ is constant, we have Barbier's theorem: the perimeter of a constant-width shape is πd .

Appendix B: Review of plane curve implicitization

A classical result in algebraic geometry states that every plane curve with rational parameterization can be recast in implicit form as a polynomial equation. The original curve is composed of one or more of the roots of this polynomial, with additional real roots corresponding to other parameter ranges in the formula defining the original curve. The set of all such curves forms an algebraic variety. The conceptually simplest approach to finding this associated algebraic variety is via the method of resultants. Since many readers may be unfamiliar with this approach, we present a brief sketch here.

Consider a rational curve parameterized by σ ,

$$(x(\sigma), y(\sigma)) = \left(\frac{p_1(\sigma)}{p_0(\sigma)}, \frac{p_2(\sigma)}{p_0(\sigma)} \right). \quad (\text{B1})$$

We may regard this curve as the intersection of two surfaces in \mathbb{R}^3 given by the two polynomials

$$f(\sigma) := x(\sigma)p_0(\sigma) - p_1(\sigma) = 0, \quad (\text{B2})$$

$$g(\sigma) := y(\sigma)p_0(\sigma) - p_2(\sigma) = 0. \quad (\text{B3})$$

Two polynomials $f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_0$ and $g(x) = b_m x^m + b_{m-1} x^{m-1} + \dots + b_0$ admit a common root if and only if the determinant of the $(m+n) \times (m+n)$ Sylvester matrix

$$\mathcal{M}(f, g) = \begin{bmatrix} a_n & a_{n-1} & \cdots & a_0 & 0 & \cdots & 0 \\ 0 & a_n & \cdots & a_1 & a_0 & \cdots & 0 \\ \vdots & & \ddots & & & \ddots & \\ 0 & \cdots & 0 & a_n & a_{n-1} & \cdots & a_0 \\ b_m & b_{m-1} & \cdots & b_0 & 0 & \cdots & 0 \\ 0 & b_m & \cdots & b_1 & b_0 & \cdots & 0 \\ \vdots & & \ddots & & & \ddots & \\ 0 & \cdots & 0 & b_m & b_{m-1} & \cdots & b_0 \end{bmatrix} \quad (\text{B4})$$

vanishes. The determinant $\det \mathcal{M}(f, g)$ of this Sylvester matrix is known as the resultant of the two polynomials f and g . Since the condition $\det \mathcal{M} = 0$ is satisfied at the intersection of the surfaces (i.e., on the rational curve), but does not involve the parameter σ , it provides the desired implicit form.

Appendix C: The Kerr critical locus as an algebraic variety

The basic object of study in algebraic geometry is an algebraic variety (the set of solutions to a system of polynomials). As reviewed in App. B, the implicitization of a rational plane curve always results in a polynomial (the resultant) whose vanishing locus (an algebraic variety) contains the plane curve. In general, this vanishing locus may also include other curves that are traced by the original parameterization with different ranges of the parameter. In such cases, rather than examine the original curve individually, it may be more profitable to view it as a subset of its containing algebraic variety. Here, we adopt this algebro-geometric perspective and study the Kerr critical curve as a subset of its associated algebraic variety, which we dub the *critical locus* to distinguish it from the physical critical curve. A similar approach was adopted for the study of timelike orbits in Ref. [19].

Following Ref. [14], we will first study the critical curve in the space of photon conserved quantities (λ, η) , rather than on the image plane (α, β) . This is because there is not just one image plane but rather infinitely many, one for each observer inclination θ_o . By working in (λ, η) -space, we can derive general results that map back to the image plane(s) via Eqs. (44). In this appendix, we will use the inverse form of these equations,

$$\lambda = -\alpha \sin \theta_o, \quad \eta = (\alpha^2 - a^2) \cos \theta_o + \beta^2. \quad (\text{C1})$$

For any choice of observer θ_o , each point (λ, η) defines either two or zero points on the image plane, according to whether there exist real solutions of Eqs. (C1) for (α, β) .

We define the (λ, η) -space *critical locus* to be the set of all points $(\lambda(\tilde{r}), \eta(\tilde{r}))$ in the real (λ, η) -plane obtainable from the parameterization (45) for some real value of \tilde{r} . One may invert $\lambda(\tilde{r})$ to obtain $\tilde{r}(\lambda)$ by solving a cubic equation, resulting in three roots (Eq. (122) of Ref. [14])

$$\tilde{r}^{(k)}(\lambda) = M + 2M\Delta_\lambda \cos \left[\frac{2\pi k}{3} + \frac{1}{3} \arccos \left(\frac{1 - \frac{a^2}{M^2}}{\Delta_\lambda^3} \right) \right],$$

$$\Delta_\lambda = \sqrt{1 - \frac{a(a + \lambda)}{3M^2}}, \quad (\text{C2})$$

with $k \in \{0, 1, 2\}$. In particular, note that Eq. (49) is recovered as $\tilde{r}^{(0)}(-\alpha \sin \theta_o)$. Substituting Eq. (C2) into $\eta(\tilde{r})$ [Eq. (45)] defines three functions

$$\eta_i(\lambda) := \eta(\tilde{r}^{(i+1)}(\lambda)), \quad i \in \{1, 2, 3\}, \quad (\text{C3})$$

which together parameterize the entire critical locus.

While this description allows us to write each part of the critical locus as an ordinary function, the price to be paid is that this involves rather complicated irrational functions. However, returning to the original parameterization (45) of the critical locus, we see that the latter is a rational curve. Therefore, as reviewed in App. B, it is possible to recast the critical locus in implicit form as the

vanishing locus of a polynomial in λ and η . Hence, the critical locus is an algebraic variety (though the physical critical curve by itself is not).

From the point of view of the complicated functions $\eta_i(\lambda)$, this is rather remarkable: Although none has a terminating series expansion in λ , the *product*

$$\mathcal{P}(\lambda, \eta) := \prod_{i=1}^3 [\eta - \eta_i(\lambda)] \quad (\text{C4})$$

must in fact be a terminating polynomial. It is by definition cubic in η , and we shall see that it has degree 6 in λ . Furthermore, since the original parameterization (45) is also rational in a , the function \mathcal{P} is a polynomial in a as well. This fact will be useful for small-spin expansions.

1. Polynomial defining the Kerr critical locus

We now derive the polynomial defining the Kerr critical locus as an algebraic variety. Comparing the rational parameterization (45) with Eqs. (B1), (B2), and (B3), we infer that the associated polynomials f and g may be taken to be

$$f(\tilde{r}) = \tilde{r}^2(\tilde{r} - 3M) + a^2(\tilde{r} + M) + a(\tilde{r} - M)\lambda, \quad (\text{C5})$$

$$g(\tilde{r}) = \tilde{r}^3 \left[\tilde{r}(\tilde{r} - 3M)^2 - 4a^2M \right] + a^2(\tilde{r} - M)^2\eta. \quad (\text{C6})$$

Computing their resolvent yields

$$\det \mathcal{M}(f, g) = 4(M^2 - a^2)M^2a^6\mathcal{C}(\lambda, \eta), \quad (\text{C7})$$

with

$$\begin{aligned} \mathcal{C}(\lambda, \eta) &= (M^2 - a^2)\eta^3 + c_2(\lambda)\eta^2 + c_1(\lambda)\eta \\ &\quad + M^2(\lambda - a)^3c_0(\lambda), \end{aligned} \quad (\text{C8})$$

$$\begin{aligned} c_2(\lambda) &= (3M^2 - 2a^2)\lambda^2 - 4a^3\lambda \\ &\quad - (27M^4 - 33M^2a^2 + 2a^4), \end{aligned} \quad (\text{C9})$$

$$\begin{aligned} c_1(\lambda) &= (3M^2 - a^2)\lambda^4 - 4a^3\lambda^3 \\ &\quad - 6(9M^4 - 5M^2a^2 + a^4)\lambda^2 + 4(27M^4 - a^4)a\lambda \\ &\quad - (54M^4 + 33M^2a^2 + a^4)a^2, \end{aligned} \quad (\text{C10})$$

$$\begin{aligned} c_0(\lambda) &= \lambda^3 + 3a\lambda^2 - 3(9M^2 - a^2)\lambda \\ &\quad + (27M^2 + a^2)a. \end{aligned} \quad (\text{C11})$$

By construction, the resolvent $\det \mathcal{M}(f, g)$ is a polynomial whose vanishing locus in the real (λ, η) -plane is precisely the Kerr critical locus. Therefore, it must be proportional to the polynomial \mathcal{P} defined in Eq. (C4). Comparing the coefficients of their $\mathcal{O}(\eta^3)$ terms then shows that (for $|a| \neq M$)⁷

$$\mathcal{C}(\lambda, \eta) = (M^2 - a^2)\mathcal{P}(\lambda, \eta). \quad (\text{C12})$$

We refer to \mathcal{C} as the *critical polynomial*.

Note that it would have been very challenging (if at all possible) to derive the polynomial expression (C8) by explicit multiplication of the product (C4).

2. Polynomial description on the image plane

Now consider the image plane (α, β) of an observer with inclination θ_o . We define the *image plane* critical locus as the θ_o -dependent set of all (real) points (α, β) obtained by solving Eqs. (C1), when λ and η lie on the (λ, η) -space critical locus. This is equivalent to the set of all real curves parameterized by Eqs. (44) for any real value of the parameter \tilde{r} .

As in the previous subsection, we may view this locus either as a set of ordinary functions or as an algebraic variety. For the former, we use the local description $\eta_i(\lambda)$ of the (λ, η) -space critical locus (C3). Plugging into Eq. (C1) gives a description of the corresponding curves $\beta_i(\alpha)$ (if they exist) on the observer screen. By the same procedure as in main text [see Eqs. (51)–(52)], they are obtained as the roots of the quadratic equations for β_i given by

$$\mathcal{C}_i := \alpha^2 + \beta_i^2 - a^2 \cos^2 \theta_o - F_i(a, \alpha \sin \theta_o) = 0, \quad (\text{C13})$$

$$F_i(a, \alpha \sin \theta_o) = \eta_i(-\alpha \sin \theta_o) + a^2 \sin^2 \theta_o, \quad (\text{C14})$$

with $i \in \{1, 2, 3\}$. Note that only a portion of the full critical locus in (λ, η) -space is mapped to the observer screen in this way, namely the portion for which $\beta^2 \geq 0$. The actual critical curve is a further subset illustrated in Fig. 1 of [6]. Note that F_1 recovers Eq. (52) in the main body, which defines the physical critical curve on the observer screen for sufficiently small spin and inclination.

The situation here is analogous to that described for (λ, η) -space above Eq. (C4): While explicit, the functions $\mathcal{C}_i(a, \alpha \sin \theta_o)$ are rather complicated due to their dependence on the cubic roots (C2). However, the curves $\mathcal{C}_i = 0$ together form the vanishing locus of the critical polynomial under the substitution (C1),

$$\mathcal{C}_o(\alpha, \beta) := \mathcal{C}(-\alpha \sin \theta_o, (\alpha^2 - a^2) \cos^2 \theta_o + \beta^2). \quad (\text{C15})$$

Since $\mathcal{C}(\lambda, \eta)$ is a polynomial in λ , η , and a , we see from the form of Eq. (C15) that \mathcal{C}_o is polynomial in all of α , β , $\sin \theta_o$ (degree 6), and a (degree 8). More explicitly, the product $\mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$ must be a polynomial proportional to \mathcal{C}_o , and comparison of the coefficients of their $\mathcal{O}(\beta^6)$ terms shows that (for $|a| \neq M$),

$$\mathcal{C}_o(\alpha, \beta) = (M^2 - a^2)\mathcal{C}_1\mathcal{C}_2\mathcal{C}_3. \quad (\text{C16})$$

Though each of the expressions \mathcal{C}_i for $i \in \{1, 2, 3\}$ is not individually a polynomial (with their expansion in small parameters a or $\sin \theta_o$ producing infinite power series), their product $\mathcal{C}_1\mathcal{C}_2\mathcal{C}_3$ is a polynomial of finite degree.

⁷When $a = M$, the polynomial \mathcal{C} remains well-defined but degenerates to a simpler quadratic in η . Hence it defines the extremal critical locus as a quartic in β (see discussion in Sec. C3 below).

3. Extremal spin factorization

As discussed in Sec. V C, the extremal limit $a \rightarrow M$ of the critical curve is rather subtle, requiring two separate $a \rightarrow M$ limits [18]. Here we have passed to the critical *locus* and described it as an algebraic variety, which depends only on the physical parameters a and M , and hence has only a single limit $a \rightarrow M$. Setting $a = M$ in Eq. (C8), the critical polynomial $\mathcal{C}(\lambda, \eta)$ degenerates to a quadratic polynomial in η , which factorizes into

$$\mathcal{C}(\lambda, \eta) = M^2(\lambda - 2M)^2 \mathcal{E}(\lambda, \eta), \quad (\text{C17})$$

$$\begin{aligned} \mathcal{E}(\lambda, \eta) &= \eta^2 + 2(\lambda^2 + 2M\lambda - 11M^2)\eta \\ &+ (\lambda - M)^3(\lambda + 7M). \end{aligned} \quad (\text{C18})$$

Upon mapping to the image plane (α, β) via Eq. (C1), the (multiplicity 2) factor $\lambda - 2M = 0$ defines a double line $\alpha = -2M \csc \theta_o$, a portion of which forms the NHEKline that becomes visible when $\sin \theta_o > \sqrt{3} - 1$. On the other hand, the remaining factor becomes a quartic polynomial in both α and β , already given as Eq. (63) in the main text, which defines a pair of Cartesian ovals. The full extremal critical locus is illustrated in Fig. 8. Note that the infinite vertical line is associated with roots of \mathcal{C}_o that become complex for any $a < M$; no vestige of this feature remains for non-extremal black holes.

4. Small spin factorization

In the case of a non-rotating Schwarzschild black hole ($a = 0$), the critical polynomial $\mathcal{C}_o(\alpha, \beta)$ factorizes as

$$\mathcal{C}_o(\alpha, \beta) = M^2(\alpha^2 + \beta^2)^2(\alpha^2 + \beta^2 - 27M^2). \quad (\text{C19})$$

The vanishing locus of this polynomial is the circular critical curve of radius $3\sqrt{3}M$. Note the “circle of zero radius” factor $\alpha^2 + \beta^2$. At higher order in spin, this becomes a second piece of the critical locus, a small oval inside the larger critical curve. We have not factorized the polynomial at higher order in spin. However, we have checked that the polynomial vanishes to $\mathcal{O}(a^2)$ when evaluated on the ellipse ansatz (56) with the given parameters (57), (58), and (59).

5. Small inclination factorization

The image plane of a polar observer ($\theta_o = 0$) is best described in terms of the impact parameter $b = \sqrt{\alpha^2 + \beta^2}$ (see Eq. (61) of Ref. [6]). On the pole, the critical polynomial $\mathcal{C}_o(\alpha, \beta)$ defined in Eq. (C15) reduces to a cubic polynomial in b^2 ,

$$\begin{aligned} \mathcal{C}_o(b) &= (M^2 - a^2)b^6 - (27M^4 - 30M^2a^2 - a^4)b^4 \\ &- 96M^2a^4b^2 + 64M^2a^6 = 0. \end{aligned} \quad (\text{C20})$$

Its 6 roots (4 are real and 2 complex conjugates) are

$$b = \pm \sqrt{\eta_i(0) + a^2}, \quad i \in \{1, 2, 3\}, \quad (\text{C21})$$

with $\eta_i(\lambda)$ as defined in Eq. (C3). Two of these roots satisfy $b > 0$, and hence are part of the image plane critical locus. The actual critical curve has radius (see Eq. (67) of Ref. [6])

$$\tilde{b} = \sqrt{\eta_1(0) + a^2}. \quad (\text{C22})$$

This formula allows us to pick out the physical part of the critical locus (i.e., the critical curve) at higher order in $\sin \theta_o$. Using the ellipse ansatz (56),

$$\beta^2 = R_2 \left[1 - \left(\frac{\alpha - \alpha_0}{R_1} \right)^2 \right], \quad (\text{C23})$$

together with $R_1 = \tilde{b} + \mathcal{O}(\sin \theta_o)$ and $R_2 = \tilde{b} + \mathcal{O}(\sin \theta_o)$, we plug into the critical polynomial (C15) and find that to leading order $\mathcal{O}(\sin \theta_o)$,

$$\alpha_0 = \tilde{b}^2 X a \sin \theta, \quad R_1 = R_2 = \tilde{b}, \quad (\text{C24})$$

where

$$X = -\frac{2(a^2\tilde{b}^2 - 27M^4 - a^4)}{3(M^2 - a^2)\tilde{b}^4 - 2\mathcal{X}\tilde{b}^2 - 96M^2a^4}, \quad (\text{C25})$$

$$\mathcal{X} = 27M^4 - 30M^2a^2 - a^4. \quad (\text{C26})$$

Plotting numerically shows that $\alpha_0 > 0$ for all positive spins $0 < a \leq M$. That is, the critical curve remains circular at linear order in inclination, with a rightward origin shift given by Eqs. (C24) and (C25).

Repeating the procedure at next order $\mathcal{O}(\sin^2 \theta_o)$, we find the ellipse parameters

$$\alpha_0 = \tilde{b}^2 X a \sin \theta, \quad (\text{C27})$$

$$R_1 = \tilde{b} - \frac{X^3 Y (a \sin \theta)^2}{16\tilde{b} \left(a^2\tilde{b}^2 - 27M^4 - a^4 \right)^3 (M^2 - a^2)^3}, \quad (\text{C28})$$

$$R_2 = \tilde{b} + \frac{\tilde{b}^4 X^2 - 1}{2\tilde{b}} (a \sin \theta)^2, \quad (\text{C29})$$

where

$$\begin{aligned} Y &= 64M^4a^8Z_0 + 32(9M^2 - 7a^2)M^4a^6Z_2\tilde{b}^2 \\ &- M^2a^2Z_4\tilde{b}^4, \end{aligned} \quad (\text{C30})$$

$$\begin{aligned} Z_0 &= 31a^{12} - 1674M^2a^{10} + 7053M^4a^8 - 12636M^6a^6 \\ &+ 12393M^8a^4 - 7290M^{10}a^2 + 2187M^{12}, \end{aligned} \quad (\text{C31})$$

$$\begin{aligned} Z_2 &= 25a^{10} - 721M^2a^8 + 2382M^4a^6 - 3078M^6a^4 \\ &+ 2025M^8a^2 - 729M^{10}, \end{aligned} \quad (\text{C32})$$

$$\begin{aligned} Z_4 &= a^{16} - 3544M^2a^{14} + 61796M^4a^{12} \\ &- 309528M^6a^{10} + 746334M^8a^8 \\ &- 1006344M^{10}a^6 + 784404M^{12}a^4 \\ &- 332424M^{14}a^2 + 59049M^{16}. \end{aligned} \quad (\text{C33})$$

To produce these equations, we have used Eq. (C20),

which is obeyed by \tilde{b} , to reduce the degree of the polynomials in \tilde{b} . The initial result of solving for R_1 yielded a polynomial of degree 14 in the irrational expression \tilde{b} .

[1] J. M. Bardeen, “Timelike and null geodesics in the Kerr metric.,” in *Black Holes (Les Astres Occlus)*, pp. 215–239. Jan., 1973.

[2] J. P. Luminet, “Image of a spherical black hole with thin accretion disk.,” *A&A* **75** (May, 1979) 228–235.

[3] K. Beckwith and C. Done, “Extreme gravitational lensing near rotating black holes,” *MNRAS* **359** no. 4, (June, 2005) 1217–1228, [arXiv:astro-ph/0411339](https://arxiv.org/abs/astro-ph/0411339) [astro-ph].

[4] S. E. Gralla, D. E. Holz, and R. M. Wald, “Black hole shadows, photon rings, and lensing rings,” *Phys. Rev. D* **100** no. 2, (July, 2019) 024018, [arXiv:1906.00873](https://arxiv.org/abs/1906.00873) [astro-ph.HE].

[5] M. D. Johnson, A. Lupsasca, A. Strominger, G. N. Wong, S. Hadar, D. Kapec, R. Narayan, A. Chael, C. F. Gammie, P. Galison, D. C. M. Palumbo, S. S. Dooleman, L. Blackburn, M. Wielgus, D. W. Pesce, J. R. Farah, and J. M. Moran, “Universal interferometric signatures of a black hole’s photon ring,” *Science Advances* **6** no. 12, (Mar., 2020) eaaz1310, [arXiv:1907.04329](https://arxiv.org/abs/1907.04329) [astro-ph.IM].

[6] S. E. Gralla and A. Lupsasca, “Lensing by Kerr black holes,” *Phys. Rev. D* **101** no. 4, (Feb., 2020) 044031, [arXiv:1910.12873](https://arxiv.org/abs/1910.12873) [gr-qc].

[7] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. I. The Shadow of the Supermassive Black Hole,” *ApJ* **875** no. 1, (Apr, 2019) L1, [arXiv:1906.11238](https://arxiv.org/abs/1906.11238) [astro-ph.GA].

[8] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. II. Array and Instrumentation,” *ApJ* **875** no. 1, (Apr, 2019) L2, [arXiv:1906.11239](https://arxiv.org/abs/1906.11239) [astro-ph.IM].

[9] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. III. Data Processing and Calibration,” *ApJ* **875** no. 1, (Apr, 2019) L3, [arXiv:1906.11240](https://arxiv.org/abs/1906.11240) [astro-ph.GA].

[10] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. IV. Imaging the Central Supermassive Black Hole,” *ApJ* **875** no. 1, (Apr, 2019) L4, [arXiv:1906.11241](https://arxiv.org/abs/1906.11241) [astro-ph.GA].

[11] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. V. Physical Origin of the Asymmetric Ring,” *ApJ* **875** no. 1, (Apr, 2019) L5, [arXiv:1906.11242](https://arxiv.org/abs/1906.11242) [astro-ph.GA].

[12] Event Horizon Telescope Collaboration, K. Akiyama, A. Alberdi, W. Alef, K. Asada, R. Azulay, A.-K. Baczko, D. Ball, M. Baloković, and J. Barrett, “First M87 Event Horizon Telescope Results. VI. The Shadow and Mass of the Central Black Hole,” *ApJ* **875** no. 1, (Apr, 2019) L6, [arXiv:1906.11243](https://arxiv.org/abs/1906.11243) [astro-ph.GA].

[13] S. E. Gralla, “Measuring the shape of a black hole photon ring,” *arXiv e-prints* (May, 2020) arXiv:2005.03856, [arXiv:2005.03856](https://arxiv.org/abs/2005.03856) [astro-ph.HE].

[14] S. E. Gralla and A. Lupsasca, “Null geodesics of the Kerr exterior,” *Phys. Rev. D* **101** no. 4, (Feb., 2020) 044032, [arXiv:1910.12881](https://arxiv.org/abs/1910.12881) [gr-qc].

[15] J. R. Farah, D. W. Pesce, M. D. Johnson, and L. L. Blackburn, “On the approximation of the black hole shadow with a simple polar curve,” *arXiv e-prints* (July, 2020) arXiv:2007.06732, [arXiv:2007.06732](https://arxiv.org/abs/2007.06732) [astro-ph.HE].

[16] A. de Vries, “Shadows of rotating black holes approximated by Dürer-Pascal limaçons,” in *Jahresschrift der Bochumer interdisziplinären Gesellschaft e. V. 2003*, pp. 1–20. Jan., 2006. <http://haegar.fh-swf.de/publikationen/pascal.pdf>.

[17] V. Bozza, F. de Luca, and G. Scarpetta, “Kerr black hole lensing for generic observers in the strong deflection limit,” *Phys. Rev. D* **74** no. 6, (Sept., 2006) 063001, [arXiv:gr-qc/0604093](https://arxiv.org/abs/gr-qc/0604093) [gr-qc].

[18] S. E. Gralla, A. Lupsasca, and A. Strominger, “Observational signature of high spin at the Event Horizon Telescope,” *MNRAS* **475** no. 3, (Apr., 2018) 3829–3853, [arXiv:1710.11112](https://arxiv.org/abs/1710.11112) [astro-ph.HE].

[19] L. C. Stein and N. Warburton, “Location of the last stable orbit in Kerr spacetime,” *Phys. Rev. D* **101** no. 6, (Mar., 2020) 064007, [arXiv:1912.07609](https://arxiv.org/abs/1912.07609) [gr-qc].