

Envelopes of lines, unfoldings and breaking symmetry

Peter Giblin

Alexander Wettig

1 Introduction

It is very striking that when certain families of straight lines are drawn in the plane it is not only the lines which are apparent to the eye but also a *curve* which has the lines for its tangents. A simple example is shown in Figure 1, left where the curve, called the *envelope* of the lines, has two lobes which join at two “sharp points” called *cusps*. The center diagram of this figure is more complicated: the curve drawn heavily is in fact tangent to all the lines and it has one cusp (and also the curve crosses itself in three places, all on the horizontal axis of symmetry). To make this more obvious the envelope alone is drawn in the right-hand diagram. These diagrams of lines and their envelope are often called “embroidery diagrams” since the lines are all formed by joining two points of a unit circle, in rather the way that cotton might be stretched across a circular embroidery hoop. For each envelope we choose a rational number $m = a/b$, where a and b are coprime, then join the point $(\cos t, \sin t)$ to the point $(\cos mt, \sin mt)$ for many values of t . The respective values of m are given in the caption. We show how to count the cusps and tangencies with the circle in §2. The value of m will be restricted to *rational* values to ensure a closed envelope curve: there is no continuously varying parameter in this construction which allows us to observe continuous changes or evolutions of the envelope curve.

In this article we briefly explore these envelopes in §2, but then move on to some generalizations, also involving envelopes of straight lines, but with continuous parameters which can be varied. Envelopes constructed from straight lines are a natural and simple way in which to construct curves with *singularities* or *singular points*, that is points at which the speed of the curve becomes zero; usually this manifests itself as a cusp (sharp point) such as those in Figure 1. There are other circumstances in which “higher cusps” occur and it is the way in which these break up and evolve—the technical and rather descriptive word is “unfold”—as the parameters change which will be our main topic. Such unfoldings are a major concern of “Singularity Theory” and this article is intended to introduce some of the ideas in a concrete context.

In §4 we introduce one continuous parameter $r > 1$: the straight lines join corresponding points $(\cos t, \sin t)$ and $(r \cos mt, r \sin mt)$ of two circles centered at the origin, of radii 1 and r . (Here m is a rational number which, in any one family of envelopes, will be held fixed.) In §5 we allow the center of the second circle to move from $(0, 0)$ to $(0, d)$, introducing a second continuous parameter d : we now have a two-parameter family of lines with parameters r, d . The point of doing this is to show how singularities evolve as the two parameters r, d change. As above, this process is called “unfolding the singularity”, revealing its inner structure. Some unfoldings may not be sufficient to reveal the full structure; those that do are called “versal unfoldings”. There is an illustration in Figure 9, which shows how the “butterfly singularity”, which is the “sharp point” to the left

in the center diagram, unfolds completely (versally) by varying the parameters r, d with m fixed at $3/4$. Note that until we see the unfolding there is no particular visual difference between the two sharp points in the center diagram, but the one further to the right appears unaffected by the unfolding—as indeed it is. This cusp is stable, there is nothing further to reveal by unfolding.

Unfolding is one of the central concepts of Singularity Theory and Catastrophe Theory, originally introduced by René Thom and Hassler Whitney in the 1950s and 1960s and then hugely developed by many others including John Mather. There is some information on this in [2, 13, 14] but no comprehensive history of the subject as yet exists.

Using readily available online graphics programs it is possible to explore these envelopes experimentally and to witness the changes in the number and nature of the singularities as parameters vary. Writing the instructions for such programs requires calculation of the envelope curves starting from the family of lines and we include some standard results and calculations in this direction in §3. The two-parameter case studied in §5 is available in the Desmos interactive program [6]; in §3 and elsewhere we shall explain where the formulas in this program come from. The program allows the rational number $m = a/b$ to be chosen, the lines and the envelope curve to be drawn or withheld and the parameters $r =$ the radius of the second circle and $d =$ the displacement of the center in the y -direction to be changed. In fact the center of the second circle can be moved more generally to (c, d) .

In Remark 4.5 we explain briefly how some of the envelopes in §4 can be generated in a different way.

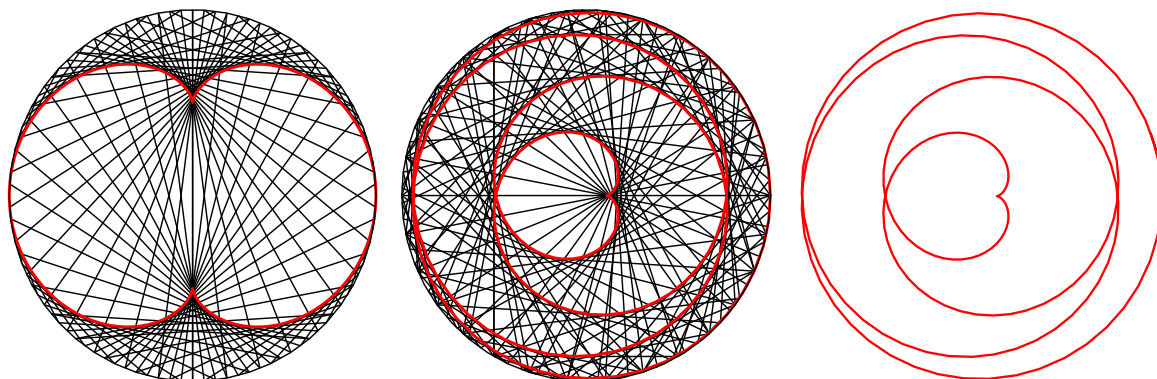


Figure 1: All the figures show the “embroidery” envelope and also the circle. Left: $m = 3$, with 2 cusps and no self-crossings. Centre: $m = \frac{5}{4}$, which has 1 cusp and 3 self-crossings. Right: just the envelope of the center diagram without the lines being drawn.

The envelopes considered in this article are examples of a wide class where corresponding points of two plane curves C_1 and C_2 are joined by straight lines. The two curves can also coincide, and the “correspondence” can be any relationship, differentiable when expressed in local parameters; it does not have to be one-to-one. For example the envelope of lines joining points of C_1 and C_2 at which the tangents are parallel gives the *center symmetry set*, initially studied by Janeczko [11] but later developed into a substantial theory extending to higher dimensions by many authors; see [10] for example. In [9] the authors study the related “Wigner caustic” which is the envelope of lines *halfway between* parallel tangents and which first arose in physics. In all cases a central question is the nature of the singular points (cusps etc.) and the extent to which they are fully unfolded, with their inner structure revealed, by whatever smoothly varying parameters are available.

2 An envelope from one circle

A well-known way of forming an envelope from chords of a single circle is as follows (see for example [4, §5.7(2), §7.14(6)]). For each point $(\cos t, \sin t)$, $0 \leq t \leq 2\pi$, on the unit circle construct the chord from this point to $(\cos(mt), \sin(mt))$ where m is a positive integer > 1 . (For $t = 0$ we use the tangent at the origin instead of a chord.) These lines have an envelope which has $m - 1$ cusps and no self-crossings. See Figure 1, left, for the case $m = 3$. To generalise, we let $m = \frac{a}{b}$ be any rational number, not equal to 0 or ± 1 and in its lowest terms. We consider three problems, in increasing order of difficulty: (i) find the range of values of t which allows the envelope to close; (ii) find the number of cusps on the envelope; and (iii) find the number of self-crossings on the envelope. Two examples are given also in Figure 1.

We shall prove the following.

Theorem 2.1 *For the envelope constructed as above, with $m \neq -1, 0, 1$ a rational number $\frac{a}{b}$ in its lowest terms, with $b > 0$,*

- (i) *the envelope, starting at $t = 0$, closes for $t = 2b\pi$,*
- (ii) *the number of cusps is $|a - b|$, occurring when $(m - 1)t$ is an odd multiple of π ; there are also $|a - b|$ points of tangency between the envelope and the circle, occurring when $(m - 1)t$ is an even multiple of π .*

In addition it can be shown that

- (iii) *the number of self-crossings is*

$$\begin{aligned} &(|a| - 1)|a - b| \quad \text{if } |m| < 1 \\ &(b - 1)|a - b| \quad \text{if } |m| > 1. \end{aligned}$$

As examples, see those in Figure 1. See §3 below for a more general discussion of envelopes of lines.

First, we shall need a formula for the envelope itself. The line joining the points $(\cos t, \sin t)$ and $(\cos mt, \sin mt)$ ($m \neq 1$) has equation

$$x(\sin mt - \sin t) - y(\cos mt - \cos t) = \sin(m - 1)t. \quad (1)$$

The standard way to obtain the envelope of such a family of lines—that is, a curve which is tangent to all the lines—is to differentiate with respect to t and then to solve for x and y as functions of t from the two equations. The solution will be written as $x = X(t), y = Y(t)$ in what follows. After some trigonometric work we find

$$X(t) = \frac{1}{m+1}(m \cos t + \cos mt), \quad Y(t) = \frac{1}{m+1}(m \sin t + \sin mt). \quad (2)$$

2.1 Tangencies with the circle

It is important to note that in the process of deriving this parametrization there is a cancellation from numerator and denominator of $1 - \cos(m - 1)t$ which vanishes when $(m - 1)t$ is an even multiple of π . This corresponds to those values of t for which the two ends of the chord joining $(\cos t, \sin t)$ and $(\cos mt, \sin mt)$ actually *coincide*, but because cancellation takes place in both numerator and denominator it is a “removable singularity” and the parametrization (2) remains

valid at these points. Indeed these are exactly the points at which the envelope is *tangent* to the circle. To count them $(m-1)t = 2n\pi$ (n an integer) means $t = \frac{2nb\pi}{|a-b|}$ (recall $t \geq 0$) and the values of n which make this angle between 0 and $2b\pi$ are $n = 0, 1, \dots, |a-b|$. Thus there are $|a-b|$ points of tangency between the envelope and the circle, proving the second part of (ii) This is also illustrated in Figure 1.

2.2 Cusps

It is clear that with $m = \frac{a}{b}$ in its lowest terms, and $b > 0$, the functions giving x and y will begin to repeat when $t = 2b\pi$. (This is not the same as saying that the smallest t with $X(t) = X(0)$ and $Y(t) = Y(0)$ is $t = 2b\pi$: the envelope may, as in (iii) of the above theorem, intersect itself for smaller values of t .) For m an integer this is of course just a period of 2π . Thus we need to consider in general the interval $0 \leq t < 2b\pi$ when discussing cusps or self-crossings of the envelope. In practice we will regard parameter values t as defined modulo $2b\pi$.

A cusp forms on the curve $(X(t), Y(t))$ when its velocity vector $(X'(t), Y'(t))$ is zero, the prime ' here representing differentiation with respect to t . These require both $\sin t + \sin mt = 0$ and $\cos t + \cos mt = 0$, which in particular implies that $\sin t \cos mt = \cos t \sin mt$, that is $\sin(m-1)t = 0$ or $(m-1)t = n\pi$ for an integer n . Substituting $mt = n\pi + t$ into the equations $X'(t) = Y'(t) = 0$ however requires $\cos n\pi = -1$ in both cases, so in fact n has to be *odd*. The even values correspond with the two points on the unit circle coinciding, a case already mentioned above.

Thus for the formation of a cusp $|m-1|t$ takes values $(2k-1)\pi$ for $k = 1, 2, \dots$. When m is an integer, we are looking for values of $0 \leq t < 2\pi$ and the largest value of k is then $|m-1|$. In general when $m = \frac{a}{b}$ in lowest terms we have $0 \leq t < 2b\pi$ and the largest value of k is $|a-b|$. Hence there are $|a-b|$ cusps in general, proving (ii) of the theorem.

Remark 2.2 These cusps are always what is called “simple cusps”, which means that by a change of coordinates in the plane close to any one cusp it can be reparametrised as the “normal form” (t^2, t^3) . The standard test for a simple cusp on a curve parametrised as $(X(t), Y(t))$ is that $X' = Y' = 0$ but $X''Y''' - X'''Y'' \neq 0$ where the derivatives are evaluated at the cusp point. We have more to say about this condition in Lemma 5.1 below. (There is some information about reduction to normal form in [4, pp.154–159]; more formal treatments are found in more advanced books such as [13, Ch.8]. Note that a simple cusp is referred to as an A_2 *singularity* of a plane curve.)

In the present case $X''Y''' - X'''Y''(m \neq 0, -1)$ is zero precisely when $m \cos(m-1)t + m^2 - m + 1 = 0$. However when $\cos(m-1)t = -1$ this requires $m = 1$, a case already eliminated. Thus all cusps are simple.

3 General envelopes of lines

Let

$$F(t, x, y) = A(t)x + B(t)y + C(t), \quad (3)$$

so that $F = 0$ is a family of straight lines, parametrized by t , with envelope given by $F = 0$, $F_t = A't + B'y + C'' = 0$, using here and in what follows subscripts to denote partial derivatives. Solving

for $x = X(t)$, $y = Y(t)$ gives (omitting the variable t)

$$X = \frac{BC' - B'C}{AB' - A'B}, \quad Y = \frac{A'C - AC'}{AB' - A'B},$$

provided of course that $AB' - A'B \neq 0$.

For the envelope considered above in §2 (see (1)), $A = \sin mt - \sin t$, $B = -\cos mt + \cos t$, $C = -\sin(m-1)t$. Calculation then shows that $AB' - A'B = (m+1)(1 - \cos(m-1)t)$ which is zero when $(m-1)t$ is an even multiple of π . These are exactly the points for which the two ends of the chord joining $(\cos t, \sin t)$ to $(\cos mt, \sin mt)$ coincide, corresponding to lines in the family which (as limits) are tangent to the circle. Fortunately this does not invalidate the parametrization (2) and indeed the envelope is actually smooth at these points, the cusps occurring at intermediate values of t where $(m-1)t$ is an *odd* multiple of π . It is worth recording the following well-known lemma.

Lemma 3.1 *In the above notation, suppose that $AB' - A'B \neq 0$ and let $(X(t), Y(t))$ be the resulting parametrization of the envelope, At an envelope point $(X(t), Y(t))$ with corresponding parameter t , given by solving $F(t, X(t), Y(t)) = F_t(t, X(t), Y(t)) = 0$, the second derivative $F_{tt}(t, X(t), Y(t))$ vanishes if and only if the envelope is singular, that is $X'(t) = Y'(t) = 0$.*

Note that $F_{tt}(t, x, y) = A''x + B''y + C''$ so $F_{tt}(t, X(t), Y(t)) = A''X(t) + B''Y(t) + C''$, a function of t only, and similarly for higher derivatives of F .

Proof We have $F(t, X(t), Y(t)) = 0$ and $F_t(t, X(t), Y(t)) = 0$ identically as functions of t . Differentiating these with respect to t we get $F_t + F_x X' + F_y Y' = 0 = F_{tt} + F_{tx} X' + F_{ty} Y' = 0$. It is clear that if $X' = Y' = 0$ then $F_{tt} = 0$. For the converse, assume $F_{tt} = 0$ at an envelope point, so that $F_t = 0$ too. We have $F_x = A, F_y = B, F_{tx} = A', F_{ty} = B'$ so that the two equations for X', Y' are $AX' + BY' = 0, A'X' + B'Y' = 0$ and these imply $X' = Y' = 0$ since $AB' - A'B \neq 0$. \square

Thus away from points where $AB' - A'B = 0$ we can detect singular points of the envelope by the equation $F_{tt} = 0$. When $AB' - A'B = 0$ then in general we might expect the envelope to “go to infinity” since the denominator of X and Y vanishes. For the envelope studied in §2 numerator and denominator always vanish together and these “singularities” are removable. The envelope *never* goes to infinity and its genuine singular points are cusps as determined in Theorem 2.1. In the next section we shall meet an example where the equal denominators of X and Y can vanish with neither numerator, or one numerator, vanishing, though never both at once. When one numerator vanishes this means that the envelope “goes to infinity” in the direction of the x - or the y -axis, and when neither vanishes it means that the envelope goes to infinity in some other direction.

Naturally singularities of envelopes $(X(t), Y(t))$, detected by the conditions $X' = Y' = 0$, are not always simple cusps, in the sense that they do not always “look like” the cusp $y^2 = x^3$. In the case of the singularities present in one-circle envelopes they are in fact simple cusps: a suitable smooth change of coordinates in the plane near to the cusp will transform it into the standard form (t^2, t^3) . The test for this is that, writing $(X(t), Y(t))$ for a parametrization of the envelope, such as (2), we have $X'(t_0) = Y'(t_0) = 0$, but $X''(t_0)Y'''(t_0) \neq X'''(t_0)Y''(t_0)$, where t_0 gives the position of the cusp. The latter condition is equivalent in fact to $F_{ttt}(t, X(t), Y(t)) \neq 0$ at $t = t_0$. Calculation shows that either of these conditions reduces to $2\cos((m-1)t_0) + 1 = 0$, which is incompatible with the condition $\sin((m-1)t_0) = 0$ for $F_{tt} = 0$, so all cusps are simple.

In the next section we shall meet higher singularities. Simple cusps are *stable* in the sense that small changes of continuously varying parameters may move their position but will not destroy the cusp itself. For the one-circle envelopes there are no such parameters to vary (m is restricted to being rational) but we shall now introduce one, and later two, parameters of this kind to vary.

4 An envelope from two circles

There is an interesting extension of the above construction to produce a *one-parameter family of envelopes* in which changing the parameter which increases or decreases the number of cusps and self-intersections. However, as we shall see, the construction is too symmetric to exhibit the most general behavior, and we need to break the symmetry by means of a 2-parameter family to reveal the full structure of the singular points which arise.

We take two circles centered at the origin, of radii 1 and $r > 1$, a fixed rational number m , and join the point $(\cos t, \sin t)$ of the first circle to the point $(r \cos mt, r \sin mt)$ of the second with a straight line. This line has equation $F(t, x, y, r) = 0$, where

$$F(t, x, y, r) = x(r \sin mt - \sin t) - y(r \cos mt - \cos t) - r \sin(m-1)t. \quad (4)$$

Here r is a continuous parameter, given a definite value for each example, and t tells us which line we are considering.

We shall later describe F as a *three-parameter family of functions of the variable t* . Note that m is restricted to rational values, so for example cannot be differentiated, and is regarded as fixed.

The envelope of the lines $F = 0$, obtained from $F = \partial F / \partial t = 0$, is (suppressing r from the notation for X, Y now)

$$\begin{aligned} X(t) &= \frac{r [\cos mt(r \cos(m-1)t - 1) + m \cos t(\cos(m-1)t - r)]}{r(m+1) \cos(m-1)t - (mr^2 + 1)}, \\ Y(t) &= \frac{r [\sin mt(r \cos(m-1)t - 1) + m \sin t(\cos(m-1)t - r)]}{r(m+1) \cos(m-1)t - (mr^2 + 1)}. \end{aligned} \quad (5)$$

Note that in this case the factor $\cos((m-1)t) - 1$, which was cancelled from numerator and denominator when deducing the parametrization (2) in the one-circle case, appears here explicitly: when $r = 1$ the denominator in (5) becomes exactly $(m+1)(\cos(m-1)t - 1)$.

We shall assume, as before, that $m = \frac{a}{b}$, a rational number in its lowest terms, with $b > 0$. We shall exclude certain values of m , namely $-1, 0, 1$, and sometimes 2 and $\frac{1}{2}$.

Remark 4.1 The two-circle envelope above can also be constructed from the family of straight lines joining the point $(\cos bT, \sin bT)$ on the unit circle to the point $(r \cos aT, r \sin aT)$ on the concentric circle of radius r : just write $t = bT$ so that $mt = (a/b)bT = aT$. The range of values of T is $0 \leq T < 2\pi$. The formula corresponding to (4) is

$$\tilde{F}(x, y) = x(r \sin aT - \sin bT) - y(r \cos aT - \cos bT) - r \sin(a-b)T.$$

Note that the denominator in (5) can be zero, indicating that the envelope has “gone to infinity”. This occurs when

$$\cos(m-1)t = \frac{mr^2 + 1}{r(m+1)}. \quad (6)$$

The denominator in *this* expression is not zero! Values of t exist satisfying (6) if and only if the right-hand side lies in the closed interval $[-1, 1]$. Some work with inequalities, using $r > 1$ and $m \neq -1, 0, 1$, shows that this is equivalent to $1 < r \leq \frac{1}{|m|}$ and in particular is only possible if $|m| < 1$. Summing this up:

Proposition 4.2 *The 2-circle envelope parametrized by (5), where $r > 1$, $m = \frac{a}{b}$, $b > 0$ is in lowest terms, goes to infinity if and only if $|m| < 1$ and $1 < r \leq \frac{1}{|m|}$. The values of t for which this happens are given by (6) which has $2|b - a|$ solutions for $0 \leq t < 2b\pi$, except for $r = \frac{1}{|m|}$ when the number is $|b - a|$. \square*

There is an illustration of this in Figure 2.

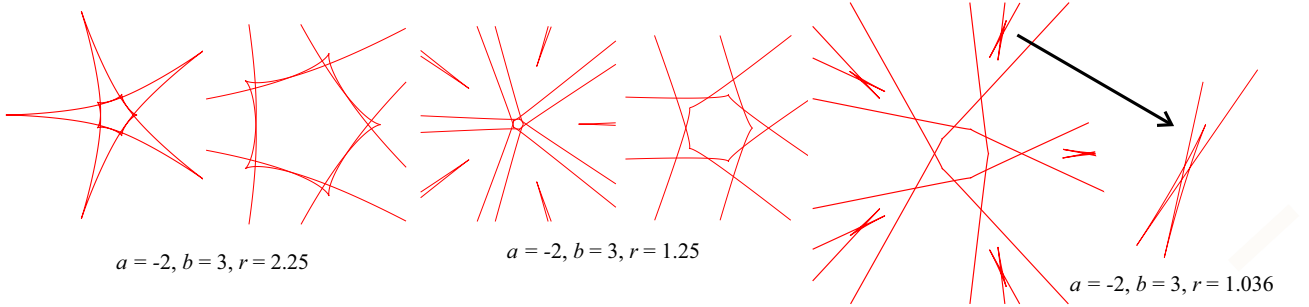


Figure 2: Left pair: $r > \frac{1}{|m|} = 1.5$ and the envelope is finite; the central part is enlarged. Centre pair: the envelope goes to infinity 10 ($= 2(b - a)$) times, because $r < \frac{1}{|m|}$ (see Proposition 4.2); the central part of the envelope is again enlarged for clarity. Right pair: $r < \left| \frac{m-2}{2m-1} \right| = \frac{8}{7}$ which results in 10 additional cusps, as in Proposition 4.4. One of the configurations of three cusps is enlarged at the far right.

4.1 Cusps on the 2-circles envelope

When counting the cusps on the 2-circles envelope it is helpful to use Lemma 3.1. At points given by (5), we have

$$F_{tt} = rm \sin(m-1)t (m-2 + r(m+1) \cos(m-1)t + r^2(1-2m)) \quad (7)$$

and cusps (or possibly higher singularities: see §5) occur when this is zero.

The factor $\sin(m-1)t$ is equal to zero if $|m-1|t$ takes values $k\pi$ for $k = 0, 1, \dots$. Thus, taking $m = \frac{a}{b}$, $b > 0$ in lowest terms and looking at the range $0 \leq t < 2b\pi$, there are $2|a - b|$ solutions for this term, corresponding to $2|a - b|$ “general cusps” for any $r > 1$.

The additional factor of (7) equals zero when

$$\cos(m-1)t = \frac{(2m-1)r^2 + (2-m))}{r(m+1)} = M \text{ say.} \quad (8)$$

Values of t satisfying this expression exist if and only if M lies within the interval $[-1, 1]$. Evaluating the resulting inequalities gives the result that extra cusps occur in the range $1 < r \leq \left| \frac{m-2}{2m-1} \right|$ if $|m| < 1$. Conversely, for $|m| > 1$, there are no extra cusps besides the $2|a - b|$ general cusps for all $r > 1$. When r takes the extreme value $\left| \frac{m-2}{2m-1} \right|$ then it is easy to check that $\cos(m-1)t = \pm 1$, so that $\sin(m-1)t = 0$ and the values of t have coincided with those for the “general cusps” above, possibly changing the nature of these cusps.

Remarks 4.3 (1) The values of t in (8) cannot coincide with those in (6) (given as usual $m \neq 1, r > 1$), but the cusps given by $\sin(m-1)t = 0$ can be at infinity, in fact when $|m| = 1/r$. An example occurs at the intermediate value $r = 1.5$, between the illustrations in Figure 2, left and center, when all these cusps (the outer five in the left-hand diagram) have “gone to infinity”.

(2) Clearly $m = \frac{1}{2}$ is slightly special in the above discussion: with $a = 1, b = 2$ there are $2|a-b| = 2$ additional cusps for all values of $r > 1$. The envelope goes to infinity in this case just for $r \geq 1/|m| = 2$, according to Proposition 4.2, and when $r = 2$ one of the cusps goes to infinity.

From the above discussion we have the following.

Proposition 4.4 *The two-circle envelope parametrized by (5), where as usual $m = \frac{a}{b}$ is in its lowest terms, $r > 1$ and the range of t is $[0, 2b\pi)$, has cusps as follows.*

- (i) $2|a-b|$ cusps always (we call these the general cusps);
- (iia) when $|m| < 1$ and $1 < r < \left\lfloor \frac{2-m}{2m-1} \right\rfloor$, $2|a-b|$ cusps in addition to (i) [two additional cusps for all $r > 1$ when $m = \frac{1}{2}$] See figure 3, right;
- (iib) when $|m| < 1$ and $r = \left\lfloor \frac{2-m}{2m-1} \right\rfloor$, $2|a-b|$ cusps altogether. Not all these cusps may be simple; this is explored in §5. \square

Some further examples are given in Figures 3 and 4.

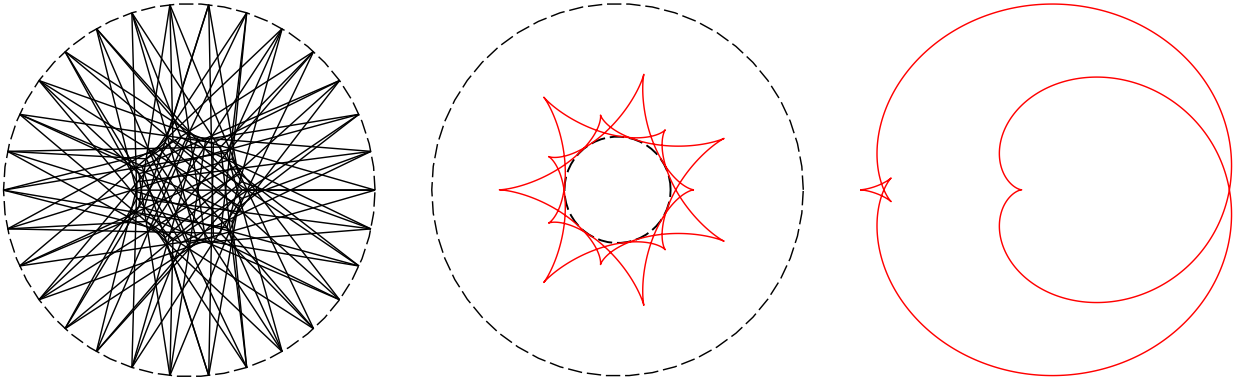


Figure 3: Left and center: the case $a = -4, b = 3, r = 3.5$, with $2|a-b| = 14$ cusps. On the left the lines forming the envelope are drawn and also the circle radius r is drawn dashed. In the center the envelope is the solid curve and the circles of radii 1 and r are drawn dashed. Right: the case $a = 3, b = 4, r = 2$ with 4 cusps, which is $4|a-b|$, as in (iia) of Proposition 4.4. The collection of three nearby cusps at the left of this envelope is called a *butterfly configuration*. As $r \rightarrow |2-m|/|2m-1| = 2.5$ this configuration collapses to a single point which is a *butterfly singularity*. This is not a simple cusp and is discussed further in §5. At $r = 2.5$ there are now $2|a-b| = 2$ cusps as in (iib) of Proposition 4.4. See also Figure 9.

Remark 4.5 *The special case $m = 2$ (or $m = \frac{1}{2}$). The 2-circle envelope of lines, in this special case, has an interpretation as a different envelope of lines, in fact as the envelope of rays reflected from a circular mirror of radius 1 with a point light source at $(r, 0)$ where $r > 0$ is the radius of the second circle. (By symmetry it is enough to consider a light source on the positive x -axis.) This is the *caustic by reflexion* of the unit circle with a finite light source. For further information on caustics see for example [5, 12]; there are also articles and videos on the internet under headings such as ‘caustic’ (in the mathematical, not chemical, sense).*

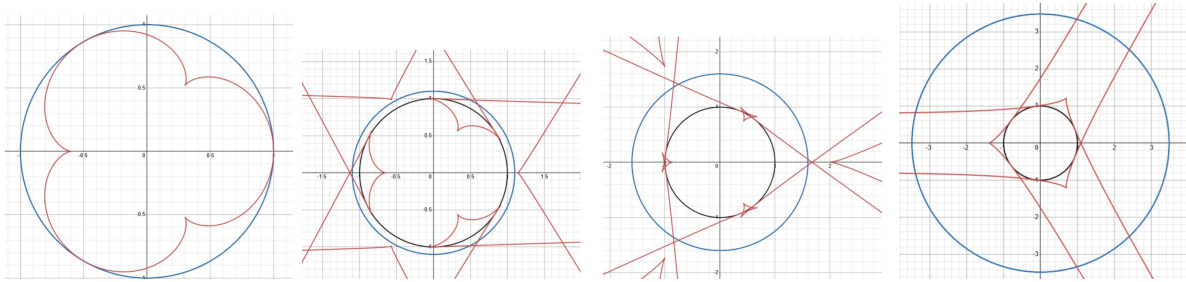


Figure 4: A discontinuous change for $m = 1/4$ from $r = 1$ to $r = 1.1$ and then on to $r = 1.6$ and $r = 3.5$. The last transition shows three butterfly configurations which collapse to three (non-simple) cusps. To reveal the full structure of these cusps we need to introduce a second parameter as in §5 and Figure 9.

Figure 5, left, shows a light ray CP from a point source at $(r, 0)$ reflected at the point P on the unit circle (the “mirror”) and striking the concentric circle of radius r at Q . Note that if the angle t is small in magnitude then the light ray will in practice bounce off the unit circle and not focus at all: we have to produce it “backwards” inside the unit circle for focussing to take place. Thus we are really considering the “mathematical caustic” which is a completion of the “physical caustic”. In practice light rays may be reflected many times inside the mirror; many studies have considered these multiple caustics, for example [3].

The triangles POC and POQ are congruent so that the point Q has coordinates $(r \cos 2t, r \sin 2t)$, that is PQ is the same line as in the construction of the 2-circle envelope with $m = 2$. Any lingering doubts that this always works, even if $r < 1$, can be answered by finding the equation of the reflected ray:

$$\sin t(2r \cos t - 1)x + (\cos t - r \cos 2t)y - r \sin t = 0$$

which always contains the point $(r \cos 2t, r \sin 2t)$ using the standard double angle formulas.

5 Swallowtails and Butterflies

The cusps on an envelope constructed from two concentric circles call for some further analysis. One such envelope is shown in the central diagram of Figure 9 (the circles themselves are not shown), where $m = \frac{3}{4}$ and $r = \frac{5}{2}$. In terms of Proposition 4.4(iib) this is an extreme value of r . The left and right cusps occur for $t = 0$ and $t = 4\pi$ respectively and the initial terms of the Taylor expansions of the envelope curves $(X(t), Y(t))$ close to the two cusps in this example are as follows:

$$\text{Left cusp: } \left(-\frac{5}{7} - \frac{25}{1792}t^4, -\frac{5}{768}t^5 \right); \quad \text{Right cusp: } \left(-\frac{5}{23} - \frac{75}{1058}(t - 4\pi)^2, -\frac{25}{644}(t - 4\pi)^3 \right). \quad (9)$$

The right cusp is therefore a simple cusp, reducible to (t^2, t^3) by a change of coordinates, but the left cusp is not: it is a “(4, 5)” cusp which is called a *butterfly singularity*. Figure 3, right, shows the effect on this singularity of changing r slightly: the butterfly opens its wings into a 3-cusped curve (all these cusps are in fact simple), while the simple cusp remains simple.

In what follows we shall use higher derivatives of the defining equation (4) of the 2-circle envelope to distinguish the various cusps. (These conditions apply to any envelope of lines, not just the 2-circle envelope.) Recall that the envelope itself is determined by $F = F_t = 0$. The

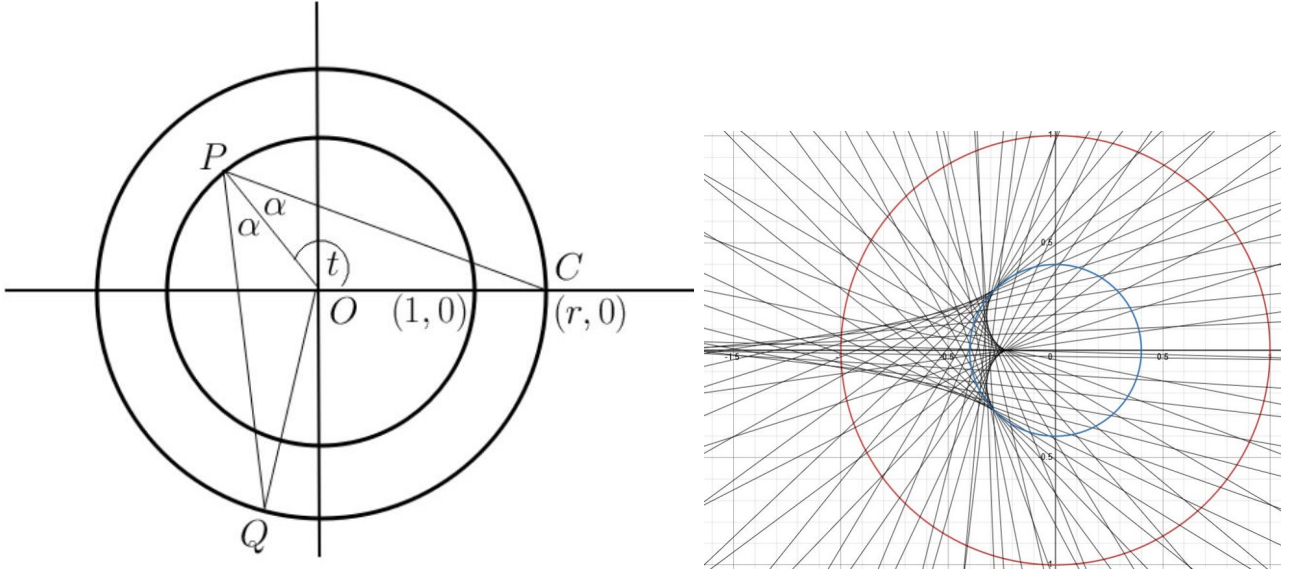


Figure 5: Left: identification of a caustic by reflexion in the unit circle with source at $(r, 0)$ with the 2-circle envelope taking m to be 2. Right: envelope of reflected rays with $r = 0.4$; thus the source of light is at the point $(0.4, 0)$ on the inner circle and the full mathematical caustic is drawn, irrespective of whether the reflected rays focus *inside* the mirror (the outer circle).

condition for a singular envelope is that in addition $F_{tt} = 0$ (evaluated as usual at $(t, X(t), Y(t))$), which from Lemma 3.1 is equivalent to $X' = Y' = 0$. First we introduce a standard example of something simpler (but not simple), unfolding a “(3,4) cusp” or swallowtail point.

5.1 A standard swallowtail

A basic example of a 3-parameter family of functions of t , not arising from 2-circle envelopes, is

$$G(t, x, y, z) = t^4 + x + yt + zt^2 \quad (10)$$

for which the discriminant set $G = G_t = 0$ is parametrised $x(t) = 3t^4 + zt^2$, $y(t) = -4t^3 - 2zt$. (Thus z here plays an analogous role to r in the 2-circle envelope case.) This set is illustrated in (x, y, z) space in Figure 8, left, and the slices $z = \text{constant}$ near to $z = 0$ in Figure 8, right. When $z = 0$ the parametrisation of the curve is $x = 3t^4$, $y = -4t^3$, a “(3,4)” singularity or swallowtail singularity. As z goes through 0, the swallowtail singularity unfolds to reveal two ordinary cusps and a crossing in one direction $z < 0$ and a smooth curve in the other direction $z > 0$.

Note that in this example, $G_{tttt}(0, 0, 0, 0) = 120 \neq 0$. We shall see below that this property distinguishes (3,4) singularities on 2-circle envelopes from the standard example above. It is what prevents a swallowtail point on a 2-circle envelope from being fully unfolded by changing the radius r . We shall see in §5.4 how adding an extra parameter to F , thereby breaking the symmetry of the construction, both swallowtail and higher “butterfly” singularities exhibit full unfoldings.

5.2 Further details of the 2-circle construction

In general terms, the more derivatives of F with respect to t which vanish, the “higher” the cusp. A simple cusp has $F = F_t = F_{tt} = 0$ but $F_{ttt} \neq 0$. Reference**** Here we shall check the next case.

Lemma 5.1 *Suppose that $F = F_t = F_{tt} = 0$ at $(t, X(t), Y(t))$ so that, as above, the envelope of lines $F = 0$ is singular at this point. Then $F_{ttt} = 0$ too if and only if $X''Y''' - X'''Y'' = 0$, which is the condition for a “higher cusp”, typically a “swallowtail point”.*

Proof The equations $F(t, X(t), Y(t)) \equiv 0$, that is $aX + bY + c \equiv 0$ and $F_t(t, X(t), Y(t)) \equiv 0$, that is $a'X + b'Y + c' \equiv 0$, are *identities* and can be differentiated with respect to t any number of times, still giving zero. Differentiating the first and using the second gives $aX' + bY' \equiv 0$, which simply says that the line $F = 0$ is tangent to the envelope curve. Differentiating this and using $X' = Y' = 0$ gives $aX'' + bY'' = 0$ at singular points.

Next, differentiating $F = 0$ three times, and $F_t = 0$ twice, using $X' = Y' = 0$ gives the equations

$$F_{ttt} + 3a'X'' + 3b'Y'' + aX''' + bY''' = 0, \quad F_{ttt} + a'X'' + b'Y'' = 0$$

respectively, from which we deduce $aX''' + bY''' = 2F_{ttt}$ at points where $X' = Y' = 0$. Combining this with $aX'' + bY'' = 0$ and using the fact that a and b cannot vanish together, gives the result of the lemma. \square

At such a swallowtail point, it can be shown that by a local change of coordinates the envelope assumes the “normal form” (t^3, t^4) . Examples of swallowtail points are given below, though in fact the 2-circles envelope does not actually exhibit them: we need to break some of the symmetry to observe such singularities. (See (iib) in the Proposition below.) The fourth derivative of F , that is F_{tttt} , can vanish in addition to the second and third, and in that case the singularity is called a “butterfly”. We have already met this in (9) above. (See also [4, p.129], or, for the more advanced viewpoint of “elementary catastrophes”, [13, §7.5], [1, §13] or [15]. There is also much information, and good illustrations, on the internet, such as [2].)

Recall from Proposition 4.1 that cusps occur on the 2-circles envelope when $F_{tt}(t, X(t), Y(t)) = 0$ where $F = 0$ is the equation of the line as in (4) and F_{tt} is written out in (7). Furthermore $F_{tt} = 0$ if and only if $\sin(m-1)t = 0$ or $\cos(m-1)t = M = ((2m-1)r^2 + 2 - m)/r(m+1)$, which is only possible when $|m| < 1$ and $1 < r \leq |m-2|/|2m-1|$.

Routine but sometimes messy calculations then reveal the following details.

Proposition 5.2 (i) *Suppose that $\sin(m-1)t = 0$, that is $(m-1)t$ is an integer multiple of π . Then the non-simple cusp condition $F_{ttt}(t, X(t), Y(t)) = 0$ (or $X''Y''' = X'''Y''$) requires r to have an extreme value $|m-2|/|2m-1|$. In fact $r = (2-m)/(2m-1)$ when $\cos(m-1)t = +1$, and $r = (m-2)/(2m-1)$ when $\cos(m-1)t = -1$.*

(ii) *When $\sin(m-1)t = 0$ and $F_{ttt} = 0$ as in (i), then also $F_{tttt} = 0$, which raises the singular type of the cusp to the form (t^4, t^5) as in the central figure of Figure 9. Mercifully, the next derivative is not zero here!*

(iii) *When $\cos(m-1)t = M$ then the non-simple cusp condition $F_{ttt} = 0$ requires that r is an extreme value $|m-2|/|2m-1|$, but in that case $M = \pm 1$ so $\cos(m-1)t = \pm 1$ and we are back in case (i), that is $\sin(m-1)t = 0$. \square*

Note that the two cusps in the central figure of Figure 9 are distinguished by (i) and (ii) of the Proposition. For the left hand non-simple cusp, $t = 0$ and $m = 3/4$, $\cos(m-1)t = +1$, $r = 5/2 = (2-m)/(2m-1)$ so $F_{ttt} = 0$, while for the right hand simple cusp, $m = 3/4$, $t = 4\pi$, $\cos(m-1)t = -1$ but $r \neq (m-2)/(2m-1)$ so $F_{ttt} \neq 0$.

Note also that (ii) and (iii) of the Proposition show that swallowtail singularities given by $F_{tt} = 0$, $F_{ttt} = 0$, $F_{tttt} \neq 0$ do not occur on 2-circle envelopes. We shall see below that they occur when we add another parameter besides r to the family of lines defining the envelope.

We can treat the left hand side of equation (4) as $F(t, x, y, r)$, a 3-parameter family of functions of t . For fixed r the *discriminant set* is the envelope in the (x, y) plane as above, given by $F = F_t = 0$.

We say something about the underlying reason why G displays a swallowtail transition and F does not, that is why the fourth derivative is important, in §5.5.

5.3 A standard butterfly

By analogy with the family G in (10) for a swallowtail consider the family

$$H(t, w, x, y, z) = t^5 + w + xt + yt^2 + zt^3 \quad (11)$$

for which the discriminant $H = H_t = 0$ in 4-dimensional (w, x, y, z) space has parametrisation

$$(w, x, y, z) = (4t^5 + 2zt^3 + yt^2, -5t^4 - 2yt - 3zt^2, y, z). \quad (12)$$

This is a 3-dimensional set in 4-space, parametrised by t, y, z . The planar “slices”, given by fixing y and z in the first two coordinates of (12), are curves parametrised by t . As y and z vary near the value 0 we can observe how these curves evolve. When $y = z = 0$ the curve has a $(4, 5)$ singularity (switching the coordinates), which is the same as the singularity in the central diagram of Figure 9. A simple Desmos demonstration enabling you to draw the curves for different (y, z) , appearing in Desmos as (Y, Z) , is at [7]. A different version [8] takes the point (Y, Z) for a tour around the origin $Y = Z = 0$. See also Figure 6 for a “still” from [8].

Two swallowtail transitions for the family (11), observed in a the circular tour of $(y, z) = (0, 0)$, are also displayed in Figure 7.

We now show how to add a parameter to the 2-circles family so that the full structure of a more complicated singularity can be unfolded.

5.4 Adding an extra parameter to the 2-circles construction

We can extend the function $F(t, x, y, r)$ in (4), which gives the family of lines joining corresponding points of two circles both centered at the origin and of radius 1 and r , by allowing the center of the second circle to move to $(0, d)$. That is, the line now joins $(\cos t, \sin t)$ to $(r \cos mt, d + r \sin mt)$, with equation $\mathcal{F} = 0$ where

$$\mathcal{F}(t, x, y, r, d) = x(r \sin mt - \sin t + d) - y(r \cos mt - \cos t) - r \cos t \sin mt + r \sin t \cos mt - d \cos t. \quad (13)$$

There is a correspondingly more complicated formula for the discriminant $\mathcal{F} = \mathcal{F}_t = 0$.

The addition of the parameter d can be described as *breaking the symmetry* of the 2-circles envelope construction and it confers an extra degree of freedom on the envelope curve.

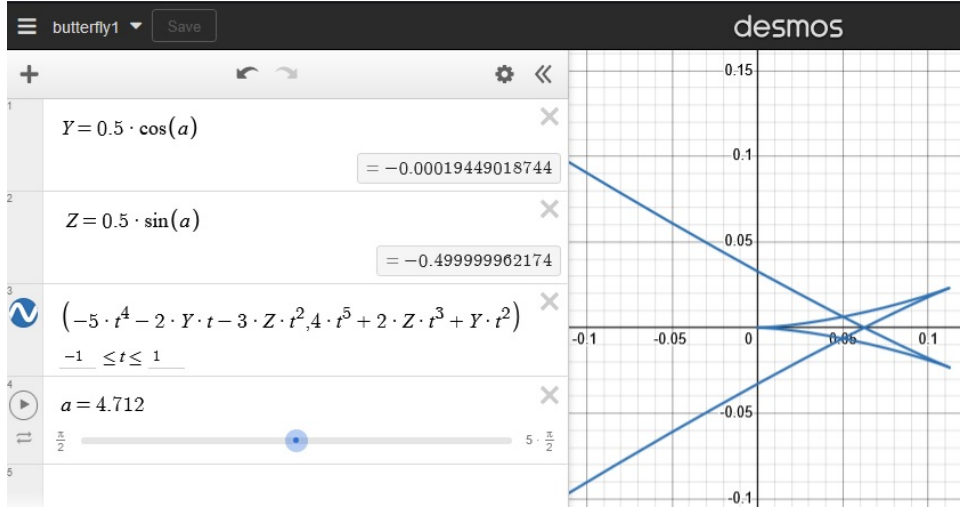


Figure 6: One of the curves given by fixing (y, z) (appearing here as (Y, Z)) close to $(0, 0)$ in (12). This Desmos figure sets $Y = 0.5 \cos a$, $Z = 0.5 \sin a$, $\pi/2 \leq a \leq 5\pi/2$ so that a tour of the origin $(Y, Z) = (0, 0)$ occurs by allowing the angle a to travel through this interval. Clicking on the arrow against a animates this tour. The figure initially displays $a = 3\pi/2 \approx 4.712$.

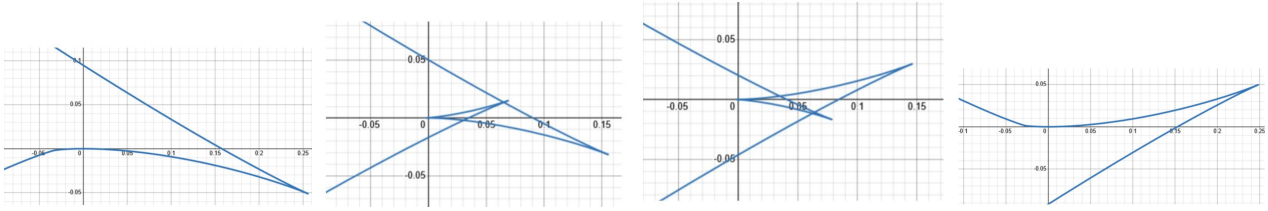


Figure 7: Four diagrams from a tour of the origin $(y, z) = (0.5 \cos a, 0.5 \sin a)$ for, left to right, $a = 4.3, 4.6, 4.8, 5.1$ exhibiting two swallowtail transitions in a neighborhood of a butterfly singularity. The symmetric diagram $a = 4.712$ is in Figure 6.

Figure 9 displays a “clock diagram” of plane curves obtained by taking $m = 3/4$ and taking values of r and d close to $(r, d) = (2.5, 0)$. This base value makes the singularity of the envelope at $(-5.7, 0)$ of type $(4, 5)$: see (9). A two-dimensional array of figures is needed to display the independent variation of the two parameters r, d .

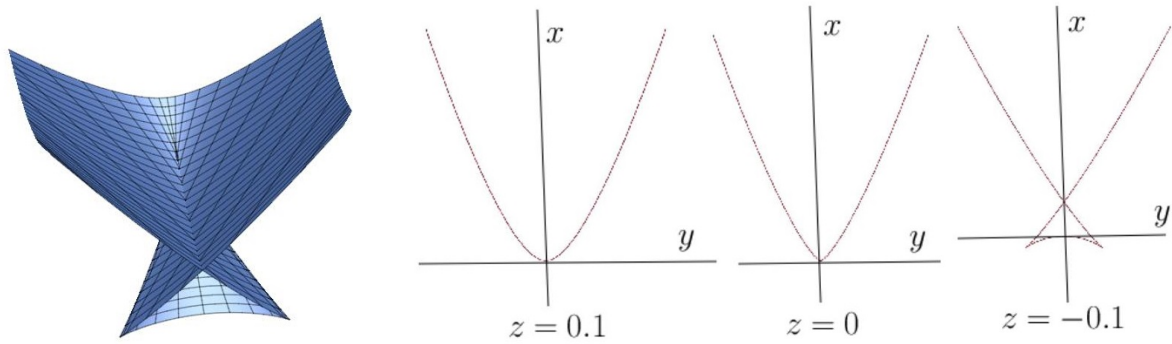


Figure 8: Left: a swallowtail surface in (x, y, z) -space, obtained as the discriminant of the family $G(t, x, y, z) = t^4 + x + yt + zt^2$. Right: slices $z = \text{constant}$ of this surface from “back” to “front”. The reversal of the usual axes arises from using the natural progression x, y, z in the definition of G . The sequence displays a “swallowtail transition” but a similar transition, where two cusps and a crossing merge and evolve into a smooth curve, does not occur on 2-circle envelopes without introducing a second parameter besides r (see Proposition 5.2 and the comments following it). The “swallowtail singularity” itself occurs in the center diagram. It looks like a kink in the curve, a point where the first derivative exists as a limit but the second does not.

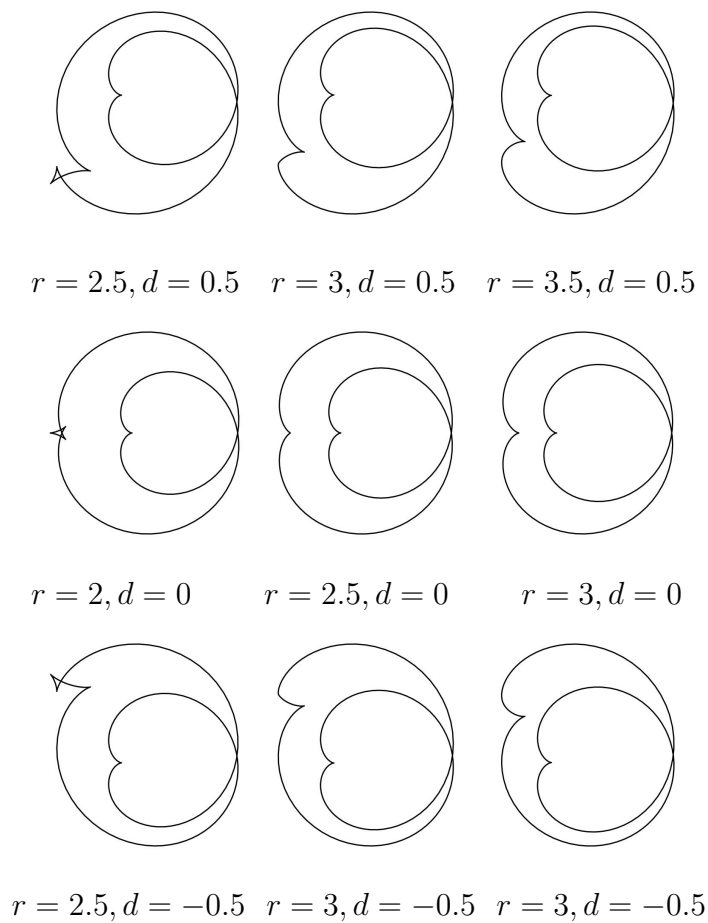


Figure 9: A tour of the origin in the (r, d) -plane, showing the envelope obtained from the unit circle and the circle radius r near to 2.5 and center $(0, d)$ near to $(0, 0)$. Throughout, $m = 3/4$ and all figures are scaled to be approximately the same size. The diagram as a whole is called a “clock diagram” for the butterfly singularity which is to the left in the central diagram. It shows a “universal unfolding” of that singularity, revealing all “nearby singularities” to the butterfly, that is singularities which can occur in any small perturbation. Swallowtail transitions occur top left to top right, and bottom left to bottom right.

5.5 Some more technical comments

How can we distinguish between a family such as H in (11),

Here we shall indicate how it can be decided whether an unfolding will display “all nearby singularities” to a given starting singularity, that is, reveal the full inner structure of that singularity. The methods are simple but the justification is more technical and we must refer to books such as [4, p.148–151],[13] for the details. There are also some details on the webpage [16] though these refer to a slightly different classification.

Using the methods of [4, Ch.7] it can be shown that the two parameters r and d are always sufficient to produce the full unfolding of the butterfly singularity exhibited in this figure. In this diagram there are two instances of swallowtail transitions: top left to top right and bottom left to bottom right. Two cusps disappear via a swallowtail point to a smooth piece of curve.

Besides books such as [1, 13, 15] there is much information and good illustrations of these singularities, or “catastrophes” on the internet, such as [2].

ACKNOWLEDGEMENTS This article is based on a Work Experience placement undertaken by Wettig, supervised by Giblin, at Liverpool University. Subsequently Giblin has added more material, especially in the last sections of the article. Wettig is grateful to his High School, Liverpool College, for suggesting that he undertake a work placement at Liverpool University, and both authors acknowledge the University for the use of computing and other facilities.

References

- [1] Bröcker Th, trans. Lander L. Differentiable Germs and Catastrophes, Cambridge University Press 1975; online 2013.
- [2] Catastrophe Theory, Wikipedia https://en.wikipedia.org/wiki/Catastrophe_theory, accessed August 30, 2024.
- [3] Bor G. and Tabachnikov S. On cusps of caustics by reflection: billiard variations on the four vertex theorem and on Jacobi’s last geometric statement. Amer. Math. Monthly. 2023; 130: 454–467.
- [4] Bruce JW. and Giblin PJ. Curves and Singularities, Second edition 1992, Cambridge University Press.
- [5] Bruce JW, Giblin PJ. and Gibson CG. Caustics through the looking glass. Math. Intelligencer 1984; 6: 47–58.
- [6] Desmos interactive graphics
<https://www.desmos.com/calculator/wq0htznhra>
- [7] Desmos interactive graphics
<https://www.desmos.com/calculator/gb9kxn0khd>
- [8] Desmos interactive graphics
<https://www.desmos.com/calculator/7o5f806hlq>

- [9] Domitrz W. and Zwierzyński M. Singular points of the Wigner caustic and affine equidistants of planar curves. *Bull. Braz. Math. Soc. New Series* 2020; 51: 11–26.
- [10] Giblin P.J. and Zakalyukin V.M. Singularities of centre symmetry sets. *Proc. London Math. Soc.* 2005; 90: 132–166.
- [11] Janeczko S. Bifurcations of the center of symmetry. *Geom. Dedicata* 1996; 60: 9–16.
- [12] Loe B.J. and Beagley N. The coffee cup caustic for calculus students. *College Math. J.* 1997; 28: 277–284.
- [13] Montaldi J. *Singularities, bifurcations and catastrophes*. 2021 Cambridge University Press.
- [14] https://en.wikipedia.org/wiki/Singularity_theory, accessed*****
- [15] Poston T, and Stewart I. *Catastrophe Theory and Applications*. 1998 Dover Publications.
- [16] [https://en.wikipedia.org/wiki/Unfolding_\(functions\)](https://en.wikipedia.org/wiki/Unfolding_(functions))

Peter Giblin, Department of Mathematical Sciences, The University of Liverpool, U.K. pjgiblin@liv.ac.uk

Alexander Wettig, Department of Computer Science, Princeton University, U.S.A. awettig@cs.princeton.edu