BEYOND THE EPSILON NEIGHBORHOOD

JIM KELLIHER

15 March 2002 (with small update 21 Feb 2006)

1. Acknowledgements

Most of the information presented here is from *Local Symmetry of Plane Curves*, P. J. Giblin and S. A. Brassett, American Mathematical Monthly, Dec. 1985, p. 689-707. (Except, that is, for the mistakes.)

2. Calculating the Symmetry Set

Parameterize a smooth, rectifiable curve $\gamma : \mathbb{R} \to \mathbb{R}^2$ by arc length, where γ is periodic, or consider γ as mapping S^1 into $\mathbb{R}^2$. Then two points on the curve can be viewed as a single point (t_1, t_2) in $\mathbb{R}^2$ or on the torus—the *phase space*.

If $\gamma(t_1)$ and $\gamma(t_2)$ are bi-tangent points—or, we might say, (t_1, t_2) is a bi-tangent point in phase space—then there is some nonzero radius, r, such that

$$\begin{aligned}
&\gamma(t_1) + rN_1 = \gamma(t_2) + rN_2 \\
&\Leftrightarrow \gamma(t_1) - \gamma(t_2) = -r(N_1 - N_2) \\
&\Leftrightarrow \gamma(t_1) - \gamma(t_2) \text{ is parallel to } N_1 - N_2 \\
&\Leftrightarrow \gamma(t_1) - \gamma(t_2) \text{ is perpendicular to } T_1 - T_2 \\
&\Leftrightarrow g(t_1, t_2) = (\gamma(t_1) - \gamma(t_2)) \cdot (T_1(t_1) - T_2(t_2)) = 0.
\end{aligned}$$

where T_i and N_i are unit tangent and normal vectors to the curve and where $g : \mathbb{R}^2 \to \mathbb{R}$. There is an issue concerning the orientation of the curve and of its associated normal and tangent vectors that we are going to shove under the rug until near the end.

Hence, $g^{-1}(0)$ gives the points in phase space corresponding to pairs of points on the curve that share a tangent circle. We call the locus of the centers of such pairs of points the *symmetry set*.

Actually, there is a problem with this in that the points on the curve needn't be distinct. In fact, when $t_1 = t_2$, $g(t_1, t_2) = 0$, and the points (being the same point) share a bi-tangent circle of any radius. Applying the epsilon-neighborhood theorem, and using the fact that the curve is a compact one-manifold (crossing itself is only a minor issue), and so has bounded curvature (being rectifiable), we can see that the set

$$A = \{(t, t) : t \in \mathbb{R}\},$$

is a component of $g^{-1}(0)$ that is bounded away from all points in phase space corresponding to the symmetry set, so we don't really have a problem.

What I say here isn't true; take a circle for instance. It is generically true, though, but for the wrong reason. In any case, how closely a point on the curve makes contact with it osculating circle enters in.

So in what follows, we should properly use $g^{-1}(0) \setminus A$ in place of $g^{-1}(0)$.

Suppose we know that the point $(t_1, t_2) \in g^{-1}(0)$. Then at (t_1, t_2)

$$\gamma(t_1) - \gamma(t_2) = -r(N_1 - N_2),$$

and

$$\begin{aligned}\partial g/\partial t_1 &= \gamma'(t_1) \cdot (T_1 - T_2) + (\gamma(t_1) - \gamma(t_2)) \cdot T_1'(t_1) \\ &= T_1 \cdot T_1 - T_1 \cdot T_2 - r(N_1 - N_2)k_1 N_1, \text{ since } \gamma'(t_1) = T_1 \\ &= (1 - T_1 \cdot T_2) - rk_1(N_1 \cdot N_1 - N_2 \cdot N_1) \\ &= (1 - T_1 \cdot T_2)(1 - rk_1),\end{aligned}$$

where k_1 is the curvature of γ at t_1 and where we used the fact that $T_1 \cdot T_2 = N_1 \cdot N_2$. A similar result for $\partial g/\partial t_2$ shows that

$$\partial g/\partial t_i = (1 - T_1 \cdot T_2)(1 - rk_i), \quad i = 1, 2.$$

So $g^{-1}(0)$ is locally a smooth one-manifold as long as the differential at (t_1, t_2) is nonzero—that is, as long as either $1 - rk_1$ or $1 - rk_2$ is nonzero. It does not follow that the symmetry set is a smooth manifold under the same conditions, an issue we discuss later. However, for a "generic" curve, 0 is a regular value of g, and so $g^{-1}(0)$ is globally a one-manifold.

Since

$$\begin{aligned}1 - rk_i < 0 &\Rightarrow \text{ bi-tangent circle contains osculating circle at } t_i, \\ 1 - rk_i > 0 &\Rightarrow \text{ bi-tangent circle within osculating circle at } t_i, \\ 1 - rk_i = 0 &\Rightarrow \text{ bi-tangent circle is osculating circle at } t_i,\end{aligned}$$

$g^{-1}(0)$ is locally a one-manifold as long as the bi-tangent circle is not simultaneously the osculating circle for both $\gamma(t_1)$ and $\gamma(t_2)$; when this occurs, we will have an isolated point of $g^{-1}(0)$.

These equations are sufficient to compute the symmetry set. For a regularly spaced lattice in phase space we can determine numerically (using bisection, say) any bi-tangent points lying near the lattice points. Starting at each bi-tangent point we walk along a portion of the symmetry set by solving the differential equation

$$\begin{aligned} \partial t_2/\partial t_1 &= -(\partial g/\partial t_1)/(\partial g/\partial t_2) \\ &= -\frac{1 - rk_1}{1 - rk_2} \end{aligned}$$

We walk first forward and then backward from the starting point. Some care is required when either $1 - rk_1$ or $1 - rk_2$ is zero, and it is smart to increment t_1 when $|1 - rk_1| \geq |1 - rk_2|$ and increment t_2 otherwise.

From the sign of $\partial t_2/\partial t_1$ and $\partial^2 t_2/\partial^2 t_1$, we can determine how t_2 changes as a function of t_1 locally.

3. Characterizing the Symmetry Set Locally

To better characterize the local properties of the symmetry set we need to either come up with a new real-valued function whose inverse of a point is the symmetry set, or we need to expand our phase space to include the radius of the bi-tangent circle, and then project from phase space to the symmetry set. We pursue the latter approach.

Let

$$\begin{aligned} &f : \mathbb{R} \times \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2, \\ &f(t_1, t_2, r) = \gamma(t_1) - \gamma(t_2) + r(N_1(t_1) - N_2(t_2)) \end{aligned}$$

and let

$$\begin{aligned} &c : \mathbb{R}^3 \to \mathbb{R}^2, \\ &c(t_1, t_2, r) = \gamma_1(t_1) + rN_1(t_1). \end{aligned}$$

Then $f^{-1}(0)$ is the set of all values (t_1, t_2, r) such that there is a circle bi-tangent to two points on the curve, where $\gamma(t_1)$ and $\gamma(t_2)$ are the points and r is the radius of the bi-tangent circle. Then $c(t_1, t_2, r)$ gives the position of the center of this bi-tangent circle—that is, a point on the symmetry set.

A problem with this is that the points on the curve needn't be distinct. In fact, when $t_1 = t_2$, all values of r give $f(t_1, t_2, r) = 0$, and so the values of $c(t_1, t_2, r)$ cover all points in the plane. So we must eliminate $t_1 = t_2$ from the inverse image of f before applying c. Applying the epsilon-neighborhood theorem, and using the fact that the curve is a compact one-manifold (crossing itself is only a minor issue), and so has bounded curvature (being rectifiable), we can see that the set

What I say here is also wrong; see the margin comment in the previous section regarding the same issue.

$$A = \{(t, t, r) : t, r \in \mathbb{R}\},$$

is a component of $f^{-1}(0)$ such that $c(A)$ contains no points of the symmetry set. Thus, the set

$$c(f^{-1}(0)) \setminus A$$

is the symmetry set.

Suppose that $(t_1, t_2, r) \in f^{-1}(0)$, $t_1 \neq t_2$. Then $f^{-1}(0)$ will be a smooth curve near (t_1, t_2, r) if f is locally a submersion—that is, if

$$df_p = ((1 - k_1 r)T_1 \quad -(1 - k_2 r)T_2 \quad N_1 - N_2)$$

(a 2×3 matrix) is of full rank (rank 2). (Here we used that fact that $N_i' = -k_i T_i$.) This will occur exactly when the bi-tangent circle is the osculating circle for at most one of the two bi-tangent points.

The orientation of T and N are backwards from what I now generally prefer.

The symmetry set will be smooth when c is locally an immersion. If we let $p \in f^{-1}(0)$, then c is an immersion at p if a nonzero tangent vector in $T_p f^{-1}(0)$ is mapped to a nonzero tangent vector under dc_p. Since $T_p f^{-1}(0) = \ker df_p$, this is equivalent to requiring that

$$\ker df_p \cap \ker dc_p = \{(0, 0, 0)\}.$$

A simple calculation gives

$$dc_p = ((1 - k_1 r)T_1 \quad 0 \quad N_1).$$

Now, suppose the bi-tangent circle is the osculating circle at neither $\gamma(t_1)$ nor $\gamma(t_2)$. Then $\ker df_p \cap \ker dc_p$ is the solution for (a_1, a_2, b) to the simultaneous equations:

$$a_1(1 - k_1 r)T_1 - a_2(1 - k_2 r)T_2 + b(N_1 - N_2) = 0,$$
$$a_1(1 - k_1 r)T_1 + bN_1 = 0.$$

From the second equation, $\ker dc_p = \{(0, a_2, 0) : a_2 \in \mathbb{R}\}$ since $T_1 \perp N_1$ and so from the first equation, $\ker df_p \cap \ker dc_p = \{(0, 0, 0)\}$ and so c is locally an immersion and the symmetry set is locally smooth.

Now let $(a_1, a_2, b) \in \ker df_p$. Then the first of the two equations above still applies, while the LHS of the second equation gives us the tangent vector, v, to the symmetry set at the point $c(p)$. Thus,

$$v = a_1(1 - k_1 r)T_1 + bN_1 = a_2(1 - k_2 r)T_2 + bN_2.$$

Taking the dot product of both sides by N_2,

$$a_1(1-k_1r)T_1 \cdot N_2 + bN_1 \cdot N_2 = b,$$

since T_2 and N_2 are perpendicular and N_2 is a unit vector. Since $N_1 \neq N_2$ (else they could not share a bi-tangent circle), $1 - N_1 \cdot N_2 \neq 0$ and we can solve for b in terms of a_1:

$$b = \frac{a_1(1-k_1r)T_1 \cdot N_2}{1 - N_1 \cdot N_2}.$$

Then

$$\begin{aligned}
v &= a_1(1-k_1r)T_1 + \frac{a_1(1-k_1r)T_1 \cdot N_2}{1 - N_1 \cdot N_2} N_1 \\
&= \frac{a_1(1-k_1r)}{1 - N_1 \cdot N_2}\left[(1 - N_1 \cdot N_2)T_1 + (T_1 \cdot N_2)N_1\right] \\
&= C\left[(1 - N_1 \cdot N_2)T_1 + (T_1 \cdot N_2)N_1\right] \\
&= C\left[T_1 - ((N_1 \cdot N_2)T_1 - (T_1 \cdot N_2)N_1)\right] \\
&= C\left[T_1 - ((T_2 \cdot T_1)T_1 + (T_2 \cdot N_1)N_1)\right] \\
&= C\left(T_1 - T_2\right),
\end{aligned}$$

where we used the following facts:

(i) $N_1 \cdot N_2 = T_1 \cdot T_2$.

(ii) If θ is the angle between T_1 and T_2, then

$$T_1 \cdot N_2 = \cos(\theta + \pi/2) = -\sin\theta = -\cos(\pi/2 - \theta) = -T_2 \cdot N_1.$$

(iii) $\{T_1, T_2\}$ form an orthonormal basis in $\mathbb{R}^2$ and so $(T_2 \cdot T_1)T_1 + (T_2 \cdot N_1)N_1 = T_2$.

Thus, the tangent line to the symmetry set has direction vector $T_1 - T_2$.

More than this can be said about the nature of cusps of the symmetry set. See section 6 of Giblin and Brassett for details.

4. How we shoved an orientation issue under the rug

Notice that in showing that $T_1 \cdot N_2 = -T_2 \cdot N_1$ and also in $T_1 \cdot T_2 = N_1 \cdot N_2$ in the previous section, we have assumed a consistent orientation to the tangent vectors at both point. But the bi-tangent circle can be tangent to the curve on the "inside" at one point and on the "outside" at the other point: this would change the calculation of the direction vector $T_1 - T_2$ from $T_1 + T_2$. This is the issue that we have ignored throughout.

The main issue here is the difficultly of properly determining the orientation in an algorithm. One possibility might be to consider the *two* functions,

$$f_+(t_1, t_2, r) = \gamma(t_1) - \gamma(t_2) + r(N_1(t_1) - N_2(t_2)), \text{ and}$$
$$f_-(t_1, t_2, r) = \gamma(t_1) - \gamma(t_2) + r(N_1(t_1) + N_2(t_2)).$$

Of course, this orientation issue will not occur for points for which the bi-tangent circle is tangent to the closest points on the curve, because the curve must cross the bi-tangent circle (or itself) to reverse the orientation. And these points are of the greatest interest because they lie on the boundary of the largest possible epsilon neighborhood of the curve.

5. Characterizing the Symmetry Set Globally

If 0 is a regular value of g—the generic case—then $g^{-1}(0)$ is a one-manifold, though most likely not connected. Each component of $g^{-1}(0)$ maps to a connected subset of points on the curve via

$$\pi : \mathbb{R}^2 \to \mathbb{R}^2,$$
$$\pi(t_1, t_2) = (\gamma(t_1), \gamma(t_2));$$

that is, to a portion of the curve. This curve then maps continuously to points in the symmetry set via $\gamma(t_1) + rN_1$, where r is also a function of t_1 and t_2. Thus each component in $g^{-1}(0)$ is mapped to a component in the symmetry set.

Let C be a component of $g^{-1}(0)$. By the characterization of one-manifolds, we know that C is (diffeomorphic to) either a circle or an open line segment. Question: Will a circle produce four or more cusps in the symmetry set? I believe this is true, but need to reconcile this view of things with Giblen and Brassett's argument on p. 694 concerning how t_2 can vary with a function of t_1. Observe that the circle cannot intersect the line $t_1 = t_2$ in phase space since (draw a picture) that would require some point (t, t) to be on the circle, and the epsilon-neighborhood theorem excludes such points.

If C is an open line segment, then its image will be an open curve, which could intersect itself. Question: could it have cusps? I believe so, for the same reason as a circle, and I think it must (in a generic sense) have an even number of them. Transversality and intersection theory could be brought in here.

Question: is the symmetry set connected (when 0 is a regular value of g)? I think so. I think the basic idea is to project the t_1 and t_2 coordinates of each component of the one-manifold in phase space onto the curve. This will cover the curve, since every point on the curve has at least one bi-tangent point (the symmetry set having no isolated points when 0 is a regular value). Then argue that overlapping projections lead necessarily to overlapping portions of the symmetry set because of tri-tangent points. Produce a path from any point on the symmetry set to another this way.