HEWLETT PACKARD

Affine Invariant Distances, Envelopes and Symmetry Sets

Peter J. Giblin*, Guillermo Sapiro
Computer Peripherals Laboratory
HPL-96-93
June, 1996

affine invariance, distance, envelopes, symmetry sets, shape representation

Affine invariant symmetry sets of planar curves are introduced and studied in this paper. Two different approaches are investigated. The first one is based on affine invariant distances, and defines the symmetry set as the closure of the locus of points on (at least) two affine normals and affine-equidistant from the corresponding points on the curve. The second approach is based on affine bitangent conics. In this case the symmetry set is defined as the closure of the locus of centers of conics with (at least) three-point contact with two or more distinct points on the curve. This is equivalent to conic and curve having, at those points, the same affine tangent, or the same Euclidean tangent and curvature. Although the two analogous definitions for the classical Euclidean symmetry set are equivalent, this is not the case for the affine group. We present a number of properties of both affine symmetry sets, showing their similarities with and differences from the Euclidean case. We conclude the paper with a discussion of possible extensions to higher dimensions and other transformation groups, as well as to invariant Voronoi diagrams.

Internal Accession Date Only
*Department of Mathematics, *University of Liverpool*, Liverpool, UK.

1 Introduction

Symmetry sets for planar shapes have received a great deal of attention from the mathematical, biological, and computer vision communities since original work by Blum [7]. The symmetry set of a planar curve is defined as the closure of the locus of points equidistant from at least two different points on the given curve, providing the distances are local extrema [9]. The fact that the distances are local extrema means that the symmetry set point lies on the intersection of the (Euclidean) normals at the corresponding curve points. This leads to an equivalent definition of the symmetry set, as the closure of centers of bitangent circles. Blum actually defined the symmetry set in a different way: If fire is turned on at the boundary of the shape, and it travels with uniform speed, the symmetry set is the points were two or more fire fronts collapse. These points are called *shocks* in wave equations [31].

A sub-set of the symmetry set, usually denoted as *medial axis* or *skeleton*, is defined as those points corresponding to centers of bitangent disks completely inside the shape. Figure 1 shows an example of this structure in the discrete Euclidean plane. As we see from Figure 1, the medial axis is a very attractive shape representation. If we observe the axis itself, it represents the shape 'without width.' This turns out to be very important for scale invariant shape representation for example. Moreover, the medial axis, together with the corresponding radii of bitangent disks, can be used to reconstruct the original shape [11, 29, 30]. The shape is just the union of those disks, and its boundary is the envelope of the disk boundaries [10]. These are the main reasons why the computer vision and shape analysis communities became very interested in studying symmetry sets. Reviewing the incredible large computer vision literature in the subject is out of the scope of this paper; see for example [21, 25, 29, 30] and references therein.

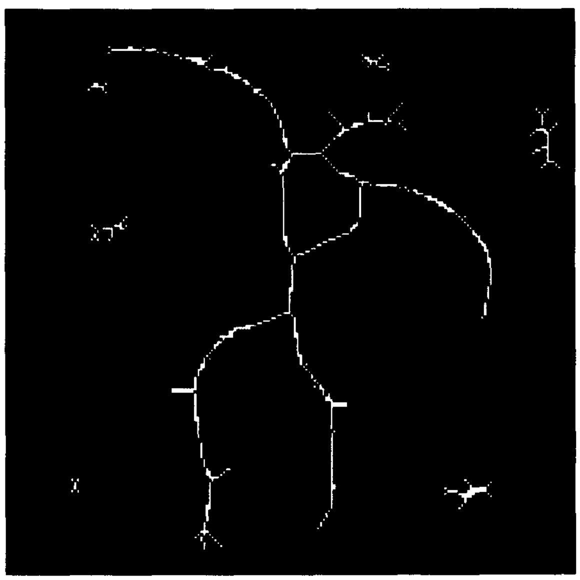

Figure 1: *Example of the Euclidean medial axis on a Matisse painting. This figure was computed using discrete mathematical morphology on a discrete image, resulting in a discrete medial axis. Note how the medial axis captures the general shape of the figure. (This is a color figure.)*

The computation of symmetry sets is not straightforward, specially for noisy shapes obtained from real images, and this has been the subject of extensive research as well [15, 25, 26, 29].

Approaches include the derivation of the set via the computation of Voronoi diagrams, [1] redefining the medial axis for discrete shapes, and simulations of Blum's prairie-fire transform. Since for noisy shapes, the symmetry set is very noisy, 'pruning' algorithms were proposed in the literature as well.

Symmetry sets and medial axes found their way beyond computer (or "artificial") vision. Symmetry is in general an attractive feature for human vision [22]. [2] Recent results show that symmetry sets may play a role in biological vision as well [18, 19, 20]. Both psychophysical and physiological studies show that biological visual systems might have better performance when observing points in certain positions on the medial axis. The experiments in [18] for example are based on measuring the human detection sensibility to a Gabor-type patch placed at different positions inside a given closed curve, defined itself also by oriented Gabor-type patches. Increased sensitivity was found at certain points on the medial axis of the shape defined by the given closed curve. The interested reader is referred to the mentioned papers for more details on these very exciting results.

In addition to its properties for shape analysis, the symmetry set turned out to be a very interesting mathematical structure as well. Following Blum's work, who already presented a number of mathematical properties of the medial axis, researchers studied the formal representation properties of the symmetry set [11, 29], its differential structure [8, 9, 10, 15], its topological structure (see Matheron's chapter in [30]), its relation with curvature [21], and many other mathematical properties. Relations between singularity and bifurcation theory [1, 2, 3, 4] and symmetry sets are described for example in [9], where an extensive list of references on the subject can be found as well.

Symmetry sets, medial axes, and Voronoi diagrams, are based on the classical definition of the l_2 Euclidean distance between points, being therefore only Euclidean invariant (l_1 and l_∞ metrics were used to define Voronoi diagrams as well). In this work we present and study symmetry sets which are affine invariant. Two alternatives to affine invariant symmetry sets are presented. The first one is based on a definition of affine invariant distances, obtaining the affine symmetry set as the closure of locus of points affine-equidistant from at least two points on the curve, providing that the distances are local extrema. In the Euclidean case, the analogous requirement is that the Euclidean distances from a point of the symmetry set to two points on the curve are equal and local extrema.

The second approach is based on affine bitangent conics. In this case, the symmetry set is defined as the closure of the locus of centers of conics having at least a three-point contact with two or more different points on the curve. This is equivalent to saying that the conic and the curve have the same affine tangent at those points. This also is analogous to the Euclidean case where 'three-point contact conic' is replaced by 'two-point contact circle.' As pointed out above, the analogous Euclidean definitions are equivalent. This is not true for the affine group.

Following the formal definitions of the affine symmetry sets, we present a number of their properties, studying mainly their local structure, but also investigating the possibility of deducing global affine symmetry of a curve from its affine symmetry set. We conclude the paper with a discussion of the extension of this work to higher dimensions, other transformation groups, and invariant Voronoi diagrams.

[1] This diagram is defined by the plane partition obtained dividing the planar points according to their proximity to points in a given set: Points closer to a given set point than to any other one in the set, belong to the same group, and groups boundaries give the Voronoi diagram.

[2] Note that the symmetry set is a collection of straight segments for symmetric shapes.

2 Planar affine differential geometry

We now present basic concepts on affine differential geometry of planar smooth curves (boundaries of planar shapes). For more details on classical results, see [6].

Let $\gamma(t) : [0,1] \rightarrow I\!R^2$ be a planar curve parametrized by t. The basic idea behind planar affine differential geometry is to define a new parametrization, s, which is affine-invariant (we restrict our analysis to special, area-preserving, affine transformations). It can be proved that the simplest affine-invariant parametrization s is given by requiring at every curve point $\gamma(s)$ the relation

$$[\gamma_s, \gamma_{ss}] = 1. \tag{1}$$

In (1), subscripts denote derivatives [3] and $[\cdot,\cdot]$ stands for the determinant of the 2×2 matrix defined by the $I\!R^2$ vectors. The vector γ_s is the *affine tangent* and γ_{ss} the *affine normal.* The affine normal is the locus of centers of conics with at least four-point contact with the curve. It is clear that (1) can not hold at inflection points, and affine differential geometry is not defined at inflection points. Therefore, we assume from now on that γ is a strictly convex curve (this curve is also called an *oval*). [4]

From (1) we observe that given an arbitrary parametrization,

$$ds = [\gamma_t, \gamma_{tt}]^{1/3} dt.$$

For example, if v is the Euclidean arclength ($\langle \gamma_v, \gamma_v \rangle = 1$), then

$$ds = [\gamma_v, \gamma_{vv}]^{1/3} dv = [\vec{\mathbf{t}}, \kappa \vec{\mathbf{n}}]^{1/3} dv = \kappa^{1/3} dv,$$

where $\vec{\mathbf{t}}$, $\vec{\mathbf{n}}$, and κ are the Euclidean unit tangent, unit normal, and curvature respectively. This relation was introduced in [27, 28] for the study of affine invariant curve deformations. From the above, the relation between the affine and Euclidean tangents is obtained,

$$\gamma_s = \kappa^{-1/3} \vec{\mathbf{t}}. \tag{2}$$

Differentiating (1) we obtain

$$[\gamma_s, \gamma_{sss}] = 0.$$

Therefore,

$$\gamma_{sss} + \mu \gamma_s = 0 \tag{3}$$

for some $\mu(s) \in I\!R$. This function μ is the *affine-invariant curvature*, and is the simplest non-trivial affine differential invariant, defining γ up to an affine transformation. It is easy to see that

$$\mu = [\gamma_{ss}, \gamma_{sss}]. \tag{4}$$

Curves have constant affine curvature if and only if they are conic sections. There are at least six points such that $\mu_s = 0$ in a closed convex curve [6, §19]. These points are called *sextactic* and are the affine vertices of a curve. Osculating conics passing through $\gamma(s)$ have five-point contact with the curve if $\gamma(s)$ is not sextactic, and at least six-point contact (generically exactly six-point) if it is.

[3] We will also denote derivatives by the prime $'$.

[4] Since inflection points are affine-invariant, segmenting the curve into convex portions does not bring a major problem for most shape understanding problems.

3 Affine symmetry sets: Invariant distances

In this Section we present and study the first of our affine invariant symmetry sets. Since it is based on distance functions, we begin with the presentation of an affine invariant distance [6, 17, 24] and its main properties.

3.1 Affine invariant distance

Definition 1 *Let* $\mathbf{X}$ *be a generic point on the plane,* [5] *and* $\gamma(s)$ *a strictly convex planar curve parametrized by affine arclength. The affine distance between* $\mathbf{X}$ *and a point* $\gamma(s)$ *on the curve is given by*

$$d(\mathbf{X}, s) := [\mathbf{X} - \gamma(s), \gamma_s(s)]. \tag{5}$$

Note: In order to be consistent with the Euclidean case and the geometric interpretation of the affine arclength, $d(\mathbf{X}, s)$ should be defined as the 1/3 power of the determinant above. Since this does not imply any conceptual difference, we keep the definition above to avoid introducing further notation. In [17] the above function $d(\mathbf{X}, s)$ is called the *affine distance-cubed function.*

In contrast with the Euclidean case, the affine distance is defined between a generic $I\!R^2$ point and a curve point. Since the basic geometric affine invariant is area, we need at least three points or a point and a line segment to define affine invariant distances. This is the reason there is no affine distance between two points on Euclidean space.

The above definition of affine distance was used in [17] to study the affine evolute and the curve γ_{ss} (a number of properties of the curve γ_{ss} are given in [27] as well). We present now some of the basic properties of $d(\mathbf{X}, s)$ (see also [17, 24]). Replacing the affine concepts by Euclidean ones, the same properties hold for the classical (squared) Euclidean invariant distance $\langle \mathbf{X} - \gamma, \mathbf{X} - \gamma \rangle$.

Proposition 1 *The affine distance satisfies:*

1. $d(\mathbf{X}, s)$ *is an extremum (i.e.,* $d_s = 0$*) if and only if* $\mathbf{X} - \gamma(s)$ *is parallel to* γ_{ss}*, i.e.* $\mathbf{X}$ *lies on the affine normal to the curve at* $\gamma(s)$*.*
2. *For non-parabolic points (where the affine curvature is non-zero),* $d_s(\mathbf{X}, s) = d_{ss}(\mathbf{X}, s) = 0$ *iff* $\mathbf{X}$ *is on the affine evolute.*
3. γ *is a conic and* $\mathbf{X}$ *its center if and only if* $d(\mathbf{X}, s)$ *is constant.*
4. $\mathbf{X} - \gamma(s) = \alpha\gamma_s(s) - d(\mathbf{X}, s)\gamma_{ss}(s)$*, for some real real number* α*.*

Proof:

1.

$$d_s(\mathbf{X}, s) = [\mathbf{X} - \gamma(s), \gamma_s(s)]_s = [\mathbf{X} - \gamma(s), \gamma_{ss}(s)].$$

Then, $d_s = 0$, that is, d has a local extremum, iff $\mathbf{X} - \gamma$ is parallel to the affine normal γ_{ss}, which means that $\mathbf{X}$ is on the affine normal. □

[5] We use capital bold letters to indicate points in $I\!R^2$.

2. From (1) and the relation $d_{ss}(\mathbf{X}, s) = [\mathbf{X} - \gamma, \gamma_{sss}] - 1$, it follows that the first and second derivatives of the affine distance vanish iff (non-parabolic points) $\mathbf{X} = \gamma + \mu^{-1}\gamma_{ss}$, which is the definition of the affine evolute. □

3. The proof is left to the reader (or see [24]).

4. We have $\mathbf{X} - \gamma(s) = \alpha\gamma_s + \beta\gamma_{ss}$ for some real numbers α and β. Now take the 'bracket' [,] of both sides with γ_s to show $\beta = -1$. □

3.2 Affine invariant symmetry sets

From the definition above, we are now ready to present the first possible affine invariant symmetry set. As pointed out in the Introduction, the same definition is frequently used for the classical Euclidean symmetry set, replacing all affine concepts by Euclidean ones.

Definition 2 *The affine distance symmetry set ($\mathcal{ADSS}$) of a planar convex curve $\gamma(s)$, also called an oval, is the closure of the the locus of points $\mathbf{X}$ in $\mathbb{R}^2$ on (at least) two affine normals and affine-equidistant from the corresponding points on the curve. In other words, $\mathbf{X} \in \mathbb{R}^2$ is a point in the affine distance symmetry set of $\gamma(s)$ if and only if there exist two different points s_1, s_2 such that $d(\mathbf{X}, s_1) = d(\mathbf{X}, s_2)$, and $d_s(\mathbf{X}, s_1) = d_s(\mathbf{X}, s_2) = 0$, or if $\mathbf{X}$ is a limit of such points.*

We present now a number of basic properties for the $\mathcal{ADSS}$.

Proposition 2 *Let $\mathbf{X}(v) : \mathbb{R} \to \mathbb{R}^2$ be a segment of the affine invariant symmetry set, and $s_1(v), s_2(v)$ the corresponding points on γ, where we use affine arclength s as parameter on γ (see §3.4 for smoothness conditions). Then*

1. *The tangent to $\mathbf{X}(v)$ is parallel to $\gamma_s(s_1(v)) - \gamma_s(s_2(v))$.*

2. *The $\mathcal{ADSS}$ has an inflection point iff*

$$(1 - [\gamma_s(s_2), \gamma_{ss}(s_1)])^2(1 - \mu_1 d(\mathbf{X})) = (1 - [\gamma_s(s_1), \gamma_{ss}(s_2)])^2(1 - \mu_2 d(\mathbf{X})),$$

   where $d(\mathbf{X})$ is the common affine distance and μ_1, μ_2 the affine curvatures.

3. *The tangent to the $\mathcal{ADSS}$ and the two tangents to the oval γ at the corresponding points s_1 and s_2, all intersect at one point.*

4. *$[\gamma(s_1(v)) - \gamma(s_2(v)), \gamma_{ss}(s_1(v)) - \gamma_{ss}(s_2(v))] = 0$. Conversely, if this holds, then the intersection $\mathbf{X}$ of the affine normals at $\gamma(s_1)$ and $\gamma(s_2)$ satisfies the conditions of Definition 2 above.*

Note: The last result provides a method of drawing the affine distance symmetry set: Using arclength parameter on γ, find the solutions of the equation in two variables s_1, s_2

$$[\gamma(s_1) - \gamma(s_2), \gamma_{ss}(s_1) - \gamma_{ss}(s_2)] = 0. \tag{6}$$

This set in (s_1, s_2)-space can be called the *pre-$\mathcal{ADSS}$*. Then for each solution, find the intersection $\mathbf{X}$ of the affine normals at $\gamma(s_1), \gamma(s_2)$. Of course it is an easy matter to rewrite the above condition in terms of any given parametrization of γ. If $\gamma(t) = (X(t), Y(t))$ is an arbitrary regular parametrization of γ then (using $'$ for d/dt) write $k(t) = X'Y'' - X''Y'$. We have

$$\gamma_{ss} = k^{-2/3}\gamma'' - \frac{1}{3}k'k^{-5/3}\gamma'.$$

Proof of Proposition 2:

1. Let s_1 and s_2 be the corresponding curve points for $\mathbf{X}(v) \in \mathcal{ADSS}$, and v the Euclidean arclength. From the equal distance condition we have

$$[\mathbf{X}(v) - \gamma(s_1), \gamma_s(s_1)] = [\mathbf{X}(v) - \gamma(s_2), \gamma_s(s_2)].$$

Tacking derivatives with v, and defining $h(s) := ds/dv$, we get

$$[\mathbf{X}_v(v) - h(s_1)\gamma_s(s_1), \gamma_s(s_1)] + [\mathbf{X}(v) - \gamma(s_1), h(s_1)\gamma_{ss}(s_1)] =$$
$$[\mathbf{X}_v(v) - h(s_2)\gamma_s(s_2), \gamma_s(s_2)] + [\mathbf{X}(v) - \gamma(s_2), h(s_2)\gamma_{ss}(s_2)].$$

Since $\mathbf{X}$ is in the affine symmetry set, the second term on each side of the equality is zero, and we get

$$[\mathbf{X}_v(v), \gamma_s(s_1)] = [\mathbf{X}_v(v), \gamma_s(s_2)].$$

Therefore, $\mathbf{X}_v$ is parallel to $\gamma_s(s_1) - \gamma_s(s_2)$,

$$\mathbf{X}_v = f(v)(\gamma_s(s_1) - \gamma_s(s_2)).$$

□

2. Now, let's compute the Euclidean curvature:

$$\mathbf{X}_{vv} = f_v(v)(\gamma_s(s_1) - \gamma_s(s_2)) + f(v)(h(s_1)\gamma_{ss}(s_1) - h(s_2)\gamma_{ss}(s_2)).$$

The curvature $\kappa_{\mathbf{X}}$ is $[\mathbf{X}_v, \mathbf{X}_{vv}]$. Therefore

$$\begin{aligned}\kappa_{\mathbf{X}} &= [f(v)(\gamma_s(s_1) - \gamma_s(s_2)), f_v(v)(\gamma_s(s_1) - \gamma_s(s_2)) + \\ &\quad f(v)(h(s_1)\gamma_{ss}(s_1) - h(s_2)\gamma_{ss}(s_2))] = \\ &\quad [f(v)(\gamma_s(s_1) - \gamma_s(s_2)), f(v)(h(s_1)\gamma_{ss}(s_1) - h(s_2)\gamma_{ss}(s_2))].\end{aligned}$$

Since $[\gamma_s, \gamma_{ss}] = 1$ for all s, the condition for inflection point on the $\mathcal{ADSS}$, that is $\kappa_{\mathbf{X}} = 0$, is

$$h(s_1) - h(s_2)[\gamma_s(s_1), \gamma_{ss}(s_2)] - h(s_1)[\gamma_s(s_2), \gamma_{ss}(s_1)] + h(s_2) = 0.$$

Let's compute now the function h. Since $\mathbf{X} \in \mathcal{ADSS}$, we have

$$[\gamma((s_1(v)) - \mathbf{X}(v), \gamma_{ss}(s_1(v))] = 0.$$

Differentiating with respect to v we have,

$$(ds_1/dv)(1 - d\mu_1) = [x_v(v), \gamma_{ss}(s_1)],$$

where d is the common value of the affine distance $[\gamma(s_i) - x, \gamma_s(s_i)]$, $i = 1, 2$, and μ_1 the affine curvature at $\gamma(s_1)$. A similar result holds for ds_2/dv. We also know from before that $\mathbf{X}_v(v)$ is a multiple of $\gamma_s(s_1) - \gamma_s(s_2)$. This multiple cancels out from the inflection equation when substituting for ds_1/dv and ds_2/dv, obtaining

$$(1 - [\gamma_s(s_2), \gamma_{ss}(s_1)])^2(1 - \mu_1 d) = (1 - [\gamma_s(s_1), \gamma_{ss}(s_2)])^2(1 - \mu_2 d).$$

Incidentally if $\mu_1 = \mu_2$ (the shape is locally affine symmetric, see next) and $[\gamma_s(s_2), \gamma_{ss}(s_1)] = [\gamma_s(s_1), \gamma_{ss}(s_2)]$, then the inflection condition holds. The second of these conditions is the same as requiring that the tangent to the $\mathcal{ADSS}$ passes through the midpoint of the chord joining $\gamma(s_1)$ and $\gamma(s_2)$. □

3. We assume the tangent at the point $\mathbf{X} \in \mathcal{ADSS}$ corresponding to the pair $(\gamma(s_1), \gamma(s_2)) \in \gamma(s)$ is the x-axis. We also know by now that the tangent to the $\mathcal{ADSS}$ is parallel to $\gamma'(s_1) - \gamma'(s_2)$. Then

$$(1, 0)^T = f(v)(\gamma'(s_1) - \gamma'(s_2)),$$

where $f(v)$ is the normalization factor. We then have that

$$\gamma'(s_2) = \gamma'(s_1) - (1/f, 0)^T,$$

or, writing the explicit coordinates of the points in $\gamma(s) = (X(s), Y(s))$,

$$X'(s_2) = X'(s_1) - 1/f, \;\; Y'(s_2) = Y'(s_1).$$

Now, the intersection of $\gamma'(s_1)$ with the x-axis is given by

$$\frac{[\gamma(s_1), \gamma'(s_1)]}{Y'(s_1)} = \frac{X(s_1)Y'(s_1) - X'(s_1)Y(s_1)}{Y'(s_1)}.$$

In the same way, the intersection of $\gamma'(s_2)$ with the axis x is (using the relation above between $\gamma'(s_1)$ and $\gamma'(s_2)$)

$$\frac{[\gamma(s_2), \gamma'(s_2)]}{Y'(s_2)} = \frac{[\gamma(s_2), \gamma'(s_1) - (1/f, 0)^T]}{Y'(s_1)}.$$

We have to prove that both intersections are the same, that means we have to prove the equality

$$\frac{[\gamma(s_1), \gamma'(s_1)]}{Y'(s_1)} = \frac{[\gamma(s_2), \gamma'(s_1) - (1/f, 0)^T]}{Y'(s_1)},$$

which is equivalent to proving

$$[\gamma(s_1), \gamma'(s_1)] = [\gamma(s_2), \gamma'(s_1)] - [\gamma(s_2), (1/f, 0)^T].$$

We have now to connect $\gamma(s_1)$ with $\gamma(s_2)$. For this, we use the fact that those points define a certain point on the $\mathcal{ADSS}$. This point has coordinates $\mathbf{X} = (p, 0)^T$. Then, from the equal affine distance condition we have

$$[\mathbf{X} - \gamma(s_1), \gamma'(s_1)] = [\mathbf{X} - \gamma(s_2), \gamma'(s_2)].$$

We again substitute $\gamma'(s_2)$ to obtain

$$[\mathbf{X} - \gamma(s_1), \gamma'(s_1)] = [\mathbf{X} - \gamma(s_2), \gamma'(s_1) - (1/f, 0)^T].$$

Then, since $\mathbf{X} = (p, 0)^T$ ($[\mathbf{X}, (1/f, 0) = 0]$), simplifying the equality above we obtain

$$-[\gamma(s_1), \gamma'(s_1)] = -[\gamma(s_2), \gamma'(s_1)] + [\gamma(s_2), (1/f, 0)^T],$$

which proves the relation above, that is, the condition for the $\mathcal{ADSS}$ and the two corresponding oval tangents to all intersect at one point. □

4. Let us drop the variable v from the notation. Since $[\mathbf{X}-\gamma(s_1), \gamma_{ss}(s_1)] = [\mathbf{X}-\gamma(s_2), \gamma_{ss}(s_2)] = 0$, we can write
$$\mathbf{X} - \gamma(s_1) = \lambda\gamma_{ss}(s_1), \quad \mathbf{X} - \gamma(s_2) = \mu\gamma_{ss}(s_2) \tag{7}$$
for some real numbers λ, μ. However bracketing these equations with $\gamma_s(s_1)$ and $\gamma_s(s_2)$ respectively and using the affine arclength condition $[\gamma_s, \gamma_{ss}] = 1$ gives $\lambda = \mu$. Subtracting the two equations in (7) now gives $\gamma(s_1) - \gamma(s_2)$ parallel to $\gamma_{ss}(s_1) - \gamma_{ss}(s_2)$ as required. Conversely it is easy to show that if the latter condition holds and if the first vector is $-\lambda$ times the second, then $\mathbf{X}$ defined by $\mathbf{X} = \gamma(s_1) + \lambda\gamma_{ss}(s_1) = \gamma(s_2) + \lambda\gamma_{ss}(s_2)$ satisfies the conditions to be on the $\mathcal{ADSS}$. □

In passing we note that the last result of the proposition makes it possible to draw also the affine invariant equivalent of the *midpoint locus* or *smoothed local symmetry* (see [15]), where the midpoint of the chord joining $\gamma(s_1)$ to $\gamma(s_2)$ is taken in preference to the point $\mathbf{X}$. Figure 2 shows examples of the $\mathcal{ADSS}$ and other affine invariant characteristics of the curve. These drawings were obtained using the relation in Proposition 2 to compute the pre-$\mathcal{ADSS}$, and were implemented in Maple [13] and the Liverpool Surface Modeling Package [23].

Theorem 1 *Let $\gamma(s)$ be a strictly convex planar closed curve. For every non-sextactic point $\gamma(s_0)$ there exists a point $\mathbf{X} \in \mathbb{R}^2$ and at least a second point $\gamma(s_1)$ such that $\mathbf{X}$ is in the affine distance symmetry set ($\mathbf{X} \in \gamma_{ss}(s_0)$, $\mathbf{X} \in \gamma_{ss}(s_1)$, and $d(\mathbf{X}, s_0) = d(\mathbf{X}, s_1)$). Exceptionally, the point $\mathbf{X}$ will be 'at infinity,' when the affine normals through $\gamma(s_0)$ and $\gamma(s_1)$ are parallel.*

Proof: Given a Cartesian coordinate system (x, y), we can assume that $\gamma(s_0) = (0,0)$, $\gamma_s(s,0) = (\alpha, 0)$, and $\gamma_{ss}(s_0) = (0, \beta)$, where $\alpha, \beta \in \mathbb{R}$. The orientation of the affine tangent and normal vectors at s_0 can be selected since any angle can be transformed into 90 degrees by an affine transformation. It is also easy to prove, see Equation (2), that $\alpha = \kappa^{-1/3}$, where κ is the Euclidean curvature at $\gamma(s_0)$. Since $d(\mathbf{X}, s_0)$ is minimum, $\mathbf{X} \in \gamma_{ss}(s_0)$. Therefore, $\mathbf{X} = (0, \lambda)$, $\lambda \in \mathbb{R}$. In order to find a second point $\gamma(s_1)$ having the same affine distance to $\mathbf{X}$ as $\gamma(s_0)$, we have to show that there exists s_1 such that
$$[\mathbf{X} - \gamma(s_0), \gamma_s(s_0)] = [\mathbf{X} - \gamma(s_1), \gamma_s(s_1)]$$
and
$$[\mathbf{X} - \gamma(s_0), \gamma_{ss}(s_0)] = [\mathbf{X} - \gamma(s_1), \gamma_{ss}(s_1)] = 0.$$
The first condition comes from the equal affine distance while the second one comes from the fact that $\mathbf{X}$ must be on the affine normals corresponding to s_0 and s_1. Writing $\gamma(s) = (x(s), y(s))$, combining the two conditions above, and substituting $\mathbf{X} = (0, \lambda)$ and $\gamma_s(s_0) = (\kappa^{-1/3}, 0)$, we obtain that a necessary and sufficient condition for the existence of a second point s_1 is given by the function
$$h(s) = xy_{ss} - x_{ss}y - x\kappa^{1/3},$$
having a zero for $s_1 \neq s_0$. This gives a finite point for $\mathbf{X}$ provided the affine normals at $\gamma(s_0)$ and $\gamma(s_1)$ are not parallel ($x_{ss}(s_1) \neq 0$). The affine distance symmetry set 'goes to infinity' if they are parallel.

If $\gamma(s_0)$ is not a sextactic point, then it is easy to prove that the function h has a 3-fold zero at $(0,0)$, and since it is smooth and periodic, it must have another zero. For sextactic points, a 4-fold zero is obtained, and nothing can be deduced for the existence of another zero. □

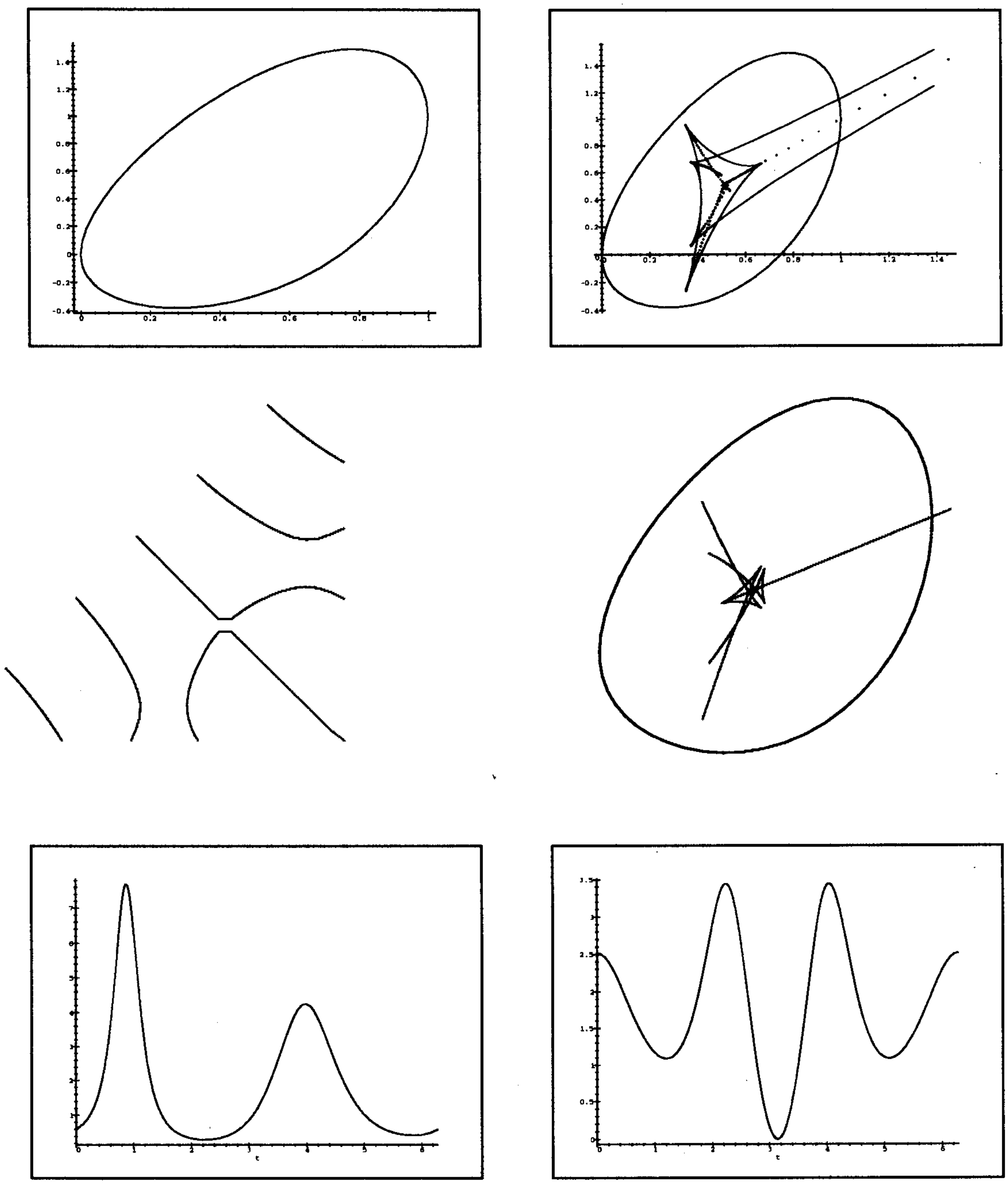

Figure 2: *Example of the $\mathcal{ADSS}$. The first row shows the oval (left), together with the affine evolute (green) and points on the $\mathcal{ADSS}$ (red) on the right. Note how the $\mathcal{ADSS}$ ends at evolute points, as proved in this paper. The second row shows the pre-$\mathcal{ADSS}$ (left) and the complete $\mathcal{ADSS}$ (right). On the third row we see the Euclidean (left) and affine (right) curvatures of the oval.*

This result means that (almost) all the points on the curve have an associated second point such that the pair defines a point on the affine distance symmetry set. Sextactic points will define the ends of the $\mathcal{ADSS}$, where s_0 becomes equal to s_1 in the limit, as we see in the following Section. In fact with a little more trouble one can strengthen the result of Theorem 1 to show that, if $\gamma(s_0)$ is not a sextactic point, then there are at least *two* other points $\gamma(s_1), \gamma(s_2)$ giving points of the $\mathcal{ADSS}$. We are grateful to S. Tabachnikov for pointing this out to us. We do not, in this paper, enter into a detailed description of the $\mathcal{ADSS}$, and the result just presented is enough for our goal. For relations to general questions see [32].

As mentioned before, fronts motion can be used to obtain the Euclidean medial axis (a subset of the symmetry set such that the corresponding disks are included in the shape). More specifically, being τ the evolution time, this is obtained via the shocks (front collisions) of the flow $\frac{\partial \gamma}{\partial \tau} = \vec{\mathbf{n}}$, where $\vec{\mathbf{n}}$ is the Euclidean unit normal. The analogous flow in the affine case will be to (locally) move the front with constant affine distance in the direction of the affine normal, that is, $\frac{\partial \gamma}{\partial \tau} = \frac{\partial^2 \gamma}{\partial s^2}$, which is exactly the *affine heat flow* introduced and studied in [27]. Note than in contrast with the Euclidean constant motion above, this affine motion does not create singularities (the normal is not preserved), and smoothly deforms the shape to an elliptic point [27]. More hints on why the affine heat flow might be considered the equivalent of the Euclidean constant motion are given in [27].

3.3 Straight affine distance symmetry sets

If the Euclidean symmetry set is a straight line, then the curve γ consists of two portions which are symmetric by a reflexion in that line. It seems that there is no completely general analogy to this in the case of the $\mathcal{ADSS}$. We give a brief account of this matter below. As we shall see in §4, there is an alternative construction for affine invariant symmetry sets which does have a property completely analogous to the Euclidean case.

First of all, let's define what we mean by symmetry in the affine case.

Definition 3 *A planar shape is affine symmetric if it is the affine transformation of a shape which is symmetric by a reflexion in a straight line.*

Note that an affine symmetric curve γ is composed of two arcs related by an affine transformation, that is, $|\mu_1| = |\mu_2|$ for all corresponding points in the two arcs γ_1 and γ_2. (Here, μ is affine curvature.)

It is straightforward to show that the $\mathcal{ADSS}$ of an affine symmetric curve contains a straight line. We are now concerned with the question if a straight $\mathcal{ADSS}$ implies an affine symmetric shape. Suppose then that for a given curve divided in two segments, the two pieces of curve define an $\mathcal{ADSS}$ which is a straight line. Let the straight line be the x-axis, and let one of the pieces of curve be $\gamma = (X(s), Y(s))$, where s is affine arclength. Let the affine normal to this curve meet the x-axis at $(a(s), 0)$ and let the tangent line meet the x-axis at $(b(s), 0)$. Let $(U(s), V(s))$ be another piece of curve which, together with γ, gives the $\mathcal{ADSS}$ along the x-axis. Note that we cannot assume that s is arclength on the second piece, corresponding points of the two pieces of curve have the same parameter s so we cannot guarantee that affine arclength is 'transferred' from one to the other. We do know the following:

(i) The tangents to the two pieces meet on the x-axis, since this is tangent to the $\mathcal{ADSS}$ at every point. Let the point of the x-axis where they meet be $(b(s), 0)$, so that

$$b = \frac{XY' - X'Y}{Y'} = \frac{UV' - U'V}{U'}.$$

(ii) The tangent to the ASS is in the direction of

$$(X',Y') - k^{-1/3}(U',V'), \quad \text{where} \quad k = U'V'' - U''V'.$$

This uses the formula '$\gamma_1' - \gamma_2'$' (see part 1 of Proposition 2) for the tangent direction, and allows for the fact that s is arclength on (X,Y) but not necessarily on (U,V). Since this tangent is along the x-axis we know that

$$Y' = k^{-1/3}V' \text{ for all } s.$$

(We expect that if increasing s induced anticlockwise orientation on (X,Y) then it will induce clockwise orientation on (U,V), i.e. that k will be < 0.)

There are also two other conditions which we know to hold:
(iii) The affine normals meet on the x-axis, at say $(a(s),0)$;
(iv) The affine distances from $(a(s),0)$ to the two pieces of curve are equal.

It is of course possible to write down these two last conditions too, but it is not hard to show that they actually *follow* from (i) and (ii) above.

We want to use (i) and (ii) to deduce the relation between (X,Y) and (U,V), that is, we assume (X,Y) are given, with $X'Y'' - X''Y' = 1$ for all s, and try to find U and V. Now (i) and (ii) can be written as

$$V'^3 = (U'V'' - V'U'')f, \quad UV' - U'V = V'g, \tag{8}$$

where we have temporarily written f for the 'known' function Y'^3 and g for the 'known' function $(XY' - X'Y)/Y'$. We deduce from the first equation of (8) that

$$\frac{d}{ds}\left(\frac{U'}{V'}\right) = -\frac{V'}{f},$$

so that

$$U' = -V'\int \frac{V'}{f}ds + c_1V' \quad (c_1 = \text{ constant}).$$

On the other hand from the second equation of (8) we deduce

$$\frac{d}{ds}\left(\frac{U}{V}\right) = -\frac{V'g}{V^2}, \text{ so } U = -V\int \frac{V'g}{V^2}ds + c_2V,$$

for a constant c_2.

Differentiating the last equation and equating with the other expression for U' we obtain

$$-\int \frac{V'}{f}ds + (c_1 - c_2) = -\int \frac{V'g}{V^2}ds - \frac{g}{V},$$

which on differentiation gives

$$VV' = fg' = YY',$$

the last "=" here being by direct calculation from the expressions for f and g, using $X'Y'' - X''Y' = 1$. We can therefore deduce that

$$Y^2 = V^2 + c$$

for a constant c.

Suppose first that $c = 0$. We then have $V' = YY'/V$, and substituting this and $V^2 = Y^2$ in the equation of (i) above we get

$$(X - U)Y' = (X' - U')Y, \text{ which gives } \frac{d}{ds}\left(\frac{X - U}{Y}\right) = 0.$$

This gives $U = X +$ const.Y which together with $V = \pm Y$ shows that *the two pieces of curve are related by an affine transformation* (and also shows that $k = \pm 1$ so that in fact the parameter s *is* affine arclength on the second piece of curve). Recall that, as above, we expect $k < 0$ when the two pieces of curve lie on opposite sides of the x-axis.

It is, however, possible to have $c \neq 0$ and in that case we can construct examples where the two pieces of curve are *not* related by an affine transformation. For example, if $\gamma(s) = (-s/2, 1 - s^2)$, we can follow the construction just presented to find the corresponding segment (U, V) that defines together with γ a straight $\mathcal{ADSS}$, although it is not related to γ by an affine transform. This example is shown in Figure 3, for a specific selection of the constants in the construction above. The $\mathcal{ADSS}$ is part of the x-axis, and the two arcs are not related by an affine transform.

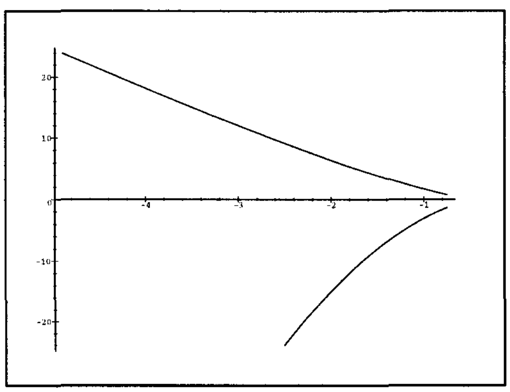

Figure 3: *Example of a straight $\mathcal{ADSS}$ defined by two arcs which are not related by an affine transformation. One of the arcs is given by $\gamma(s) = (-s/2, 1 - s^2)$, and the second one is obtained from the construction developed in the paper. The $\mathcal{ADSS}$ is on the x-axis.*

Note that from the condition of inflection points in Proposition 2, if c is not zero, then the tangent to the $\mathcal{ADSS}$ (the x-axis) does not pass trough the midpoint of the chord joining $\gamma(s_1)$ and $\gamma(s_2)$. Then, $(1 - [\gamma_s(s_2), \gamma_{ss}(s_1)])^2 \neq (1 - [\gamma_s(s_1), \gamma_{ss}(s_2)])^2$, and since the equality holds (the $\mathcal{ADSS}$ is a straight line), $\mu_1 \neq \mu_2$ and the shape is not affine symmetric. We then obtain that for the shape to be affine symmetric, in addition to the $\mathcal{ADSS}$ to be a straight line, the midpoint of the chord joining two corresponding points in the curve γ has to lie on the $\mathcal{ADSS}$ as well. This second property will automatically hold in the alternative affine symmetry set that we present in §4, and the fact that the symmetry set is a straight line will be enough to conclude that the shape is affine symmetric.

3.4 Singularity theory and the affine distance symmetry set

It is proved in Proposition 1 and in [17] that the function d of Definition 1, regarded as a family of functions on γ parametrized by the points of the plane, has bifurcation set which is precisely the classical affine evolute [6] of the curve γ. The bifurcation set is by definition the set of points $\mathbf{X}$ in the plane (or, exceptionally, at infinity) where there exists a value of s for which $d_s(\mathbf{X}, s) = d_{ss}(\mathbf{X}, s) = 0$, s being, as usual, affine arclength. This makes the function d completely analogous to the well-known *distance-squared function* of Euclidean differential geometry, whose bifurcation set is the Euclidean evolute of γ (for an exposition of this see for example [9, §§2.27, §§7.1]). For instance, the fact that the affine evolute has ordinary cusps corresponding to sextactic points of γ is exactly analogous to the fact that the Euclidean symmetry set has cusps corresponding to vertices of γ ([17]). (This was also deduced in a similar but less straightforward way in [14].) Likewise the affine evolute is the envelope of the affine normals, just as the Euclidean evolute is the envelope of the Euclidean normals.

We can include the affine distance symmetry set in the bifurcation set of the family of functions d: it then becomes the *full bifurcation set* or *levels bifurcation set* (see e.g. [8]) which includes all non-Morse functions in the family. A function is non-Morse if it has a degenerate singularity (first two derivatives zero, as above) *or* if it has two singularities (where $d_s = 0$) at which the function has the same value. The latter case is precisely the affine distance symmetry set as defined in Definition 2. The limit points will be points which belong to the affine evolute, in exactly the same way as the Euclidean symmetry set [8, 10, 15] has end-points in the cusps of the Euclidean evolute.

Geometrically, Definition 2 amounts to saying that $\mathbf{X}$ is the center of two conics, having 4-point contact with γ at $\gamma(s_1), \gamma(s_2)$, and having the same affine distance from γ (see Proposition 1 and Theorem 1). In terms of these, and using the full bifurcation set structure as in [8], we find the local structure of the affine distance symmetry set.

Theorem 2 *Locally, the affine distance symmetry set is:*

- *smooth when both conics have exactly 4-point contact with γ,*
- *an ordinary cusp when one of the conics has 5-point contact with γ ($\mathbf{X}$ is then on the affine evolute too but at a smooth point of it),*
- *an endpoint when $\mathbf{X}$ is the center of a 6-point contact conic (tangent to γ at a sextactic point): the endpoint is then in a cusp of the affine evolute,*
- *a triple crossing when there are three conics centered at $\mathbf{X}$ having equal affine distance and 4-point contact with γ; this is a stable phenomenon.*

Note that the pre-$\mathcal{ADSS}$ in the parameter space of pairs of points of γ defined by the equation (6) can be used, as in [15], to analyze the simpler structures of the $\mathcal{ADSS}$. For example, the curve (6) will have a tangent parallel to the s_1-axis when the $\mathcal{ADSS}$ has a cusp on account of the conic tangent at $\gamma(s_1)$ having 4-point contact there. The pre-$\mathcal{ADSS}$ is smooth unless *both* conics have 5-point contact.

One can also use the techniques of [8] to analyze the events in the evolution of affine distance symmetry sets when γ undergoes a generic change in shape. For example, we can analyze the behavior of the $\mathcal{ADSS}$ for a curve deforming according to the affine invariant heat flow studied in [27]. This is the subject of further work and will be reported elsewhere; see §5.

3.5 Remarks on affine-invariant Voronoi diagrams

Assume we have a finite set of points $\{\mathcal{P}_i\}_{i=0}^{i=N-1} \in \mathbb{R}^2$, defining a polygon $\gamma(\mathcal{P}_i)$, and we want to extend the classical Euclidean Voronoi diagram to the affine group. Following the results on previous sections, the basic idea is to define an affine distance between a given point $\mathbf{X} \in \mathbb{R}^2$ and the points $\mathcal{P}_i$. For this, we have to define the affine tangent to the polygon at $\gamma(\mathcal{P}_i)$. One possible way of doing this is to consider the vector $\mathcal{P}_{i+1} - \mathcal{P}_i$ (points are counted modulo N). (Note that all points on a parallel to a line segment have the same affine distance with the segment according to this definition.) In order to avoid bias towards one direction on the polygon, both vertices around $\mathcal{P}_i$ can be used to define the discrete tangent. Another possibility is to follow the approach approach in [12], that is, construct the conic defined by $(\mathcal{P}_{i-2}, \mathcal{P}_{i-1}, \mathcal{P}_i, \mathcal{P}_{i+1}, \mathcal{P}_{i+2})$, and compute the affine tangent to this conic at $\mathcal{P}_i$.

After the discrete affine tangent is defined, the affine distance follows as before, and the affine Voronoi diagram is defined as in the Euclidean case. Note that the affine Voronoi diagram can be defined for non-convex polygons as well.

4 Affine symmetry sets: Envelopes of conics or lines

We will now consider an alternative definition of affine invariant symmetry set. Let γ be a simple closed smooth curve in the plane. Strictly we do not any longer need γ to be an oval, though we shall explore the non-oval case in more detail elsewhere. The alternative definition is based on conics having 3-point contact with the oval in at least two places. In the Euclidean case there is only one reasonable definition of symmetry set, based on bitangent circles, and indeed it can be argued that in the Euclidean case the distance based and envelope based definitions reduce to the same thing. But for the affine group there do seem to be two alternative constructions.

We shall see that the new definition has natural relations with the theory of envelopes and for that reason we call it the *affine envelope symmetry set* ($\mathcal{AESS}$).

Definition 4 *Given a simple closed smooth curve γ (which we generally take to be an oval below), the affine envelope symmetry set ($\mathcal{AESS}$) is the closure of locus of centers of conics with at least 3-point contact with the curve in two or more different points. This is equivalent to saying that the curve and the conic have the same affine tangent in at least two points.*

Note: From Proposition 1, the affine distance from the conic center, as defined in Definition 1, will be equal for both curve points involved in the $\mathcal{AESS}$. In contrast with the $\mathcal{ADSS}$ as defined in Definition 2, this distance has not to be a local extremum, and the center of the conic has not to be on the corresponding affine normals (in Section 4.2 we show what happens with the $\mathcal{AESS}$ when this center is on at least one of the affine normals). Also, in the $\mathcal{ADSS}$, two distinct 4-point contact conics are involved in the construction. On the other hand, in the $\mathcal{AESS}$, only one conic, with at least a double 3-point contact, is needed.

The properties of the $\mathcal{AESS}$ can be obtained by considering either envelopes of *conics* or envelopes of *lines*. We shall give details of each approach here since they give different insights into the $\mathcal{AESS}$. The 'envelope of lines' approach is analogous to that adopted in [16] for the Euclidean case, where the lines are perpendicular bisectors of segments joining points on the curve γ and these lines are shown to envelope the symmetry set. There does not seem to be an analogue of the 'envelope of conics' approach in the Euclidean case.

4.1 Envelopes of conics

Consider an oval γ and a fixed point of the oval, which we take to be the origin. Suppose that γ is tangent to the x-axis at the origin. We seek a conic which has 3-point contact with γ at the origin and at another point. Initially we consider conics which have 3-point contact with γ at the origin and pass through a second point. We shall in particular be interested in the *centers* of such conics. Note again that from (2), 3-point contact is equivalent to having the same affine tangent (γ_s). We shall see that the $\mathcal{AESS}$ can be described as the set of *points of regression* on a certain envelope of conics.

Consider a general conic $\mathcal{C}$ with equation

$$ax^2 + by^2 + 2hxy + 2gx + 2fy + c = 0.$$

Using the definition of the center as the pole of the line at infinity with respect to the conic, the coordinates (p, q) of the center satisfy

$$ap + hq + g = 0, \quad hp + bq + f = 0.$$

If $\mathcal{C}$ is to touch γ at the origin then $g = c = 0$. Suppose γ has local parametrization of the form

$$y = \alpha x^2 + \beta x^3 + \dots \tag{9}$$

near the origin. We may suppose $\alpha > 0$. The (Euclidean) curvature κ at the origin is 2α, and substituting the parametrization into the equation of the conic we find that the condition for 3-point contact between $\mathcal{C}$ and γ is $a + 2f\alpha = 0$. If $a = 0$ this means that $f = 0$, and $\mathcal{C}$ is the degenerate conic consisting of the tangent to γ at the origin together with the chord joining the origin to $\gamma(t)$. If $a \neq 0$ then we can take $a = 1$ and f is determined by the 3-point contact condition: $f = -1/2\alpha = -1/\kappa(0)$. Note that this makes $f < 0$, assuming $\alpha > 0$.

Solving the equations for the center of the conic γ we have

$$p = -\frac{fh}{h^2 - ab}, \quad q = \frac{af}{h^2 - ab}.$$

In fact we are more interested in the reverse process, expressing the coefficients in $\mathcal{C}$ in terms of p and q. We suppose from now on that $a \neq 0$, and divide through to make $a = 1$. Then

$$h = -\frac{p}{q}, \quad b = \frac{qh^2 - f}{q} = \frac{p^2 - fq}{q^2}.$$

The conic $\mathcal{C}$ therefore has the form $q^2x^2 + (p^2 - fq)y^2 - 2pqxy + 2fyq^2 = 0$. Now suppose that $\mathcal{C}$ passes through a second point $(X(t), Y(t))$ of γ, which we take to be other than the origin. This amounts to replacing x, y by $X(t), Y(t)$ in the above equation. Omitting t we get

$$q^2X^2 + (p^2 - fq)Y^2 - 2pqXY + 2fYq^2 = 0. \tag{10}$$

Important things to notice for the moment are:

1. For a fixed t, (10) now gives an equation of degree 2 in the variables p, q. This shows that *the locus of centers (p, q) of conics $\mathcal{C}$ having 3-point contact with γ at the origin and passing through $\gamma(t)$ is itself a conic.* We denote it by $\mathcal{D}(t)$, or later by $\mathcal{D}(t_0, t)$ where t_0 is the parameter value of the origin on γ. Note that the conic (10) passes through the origin $p = q = 0$. This corresponds to the degenerate conic $\mathcal{C}$ consisting of the tangent to γ at the origin together with the chord joining the origin to $\gamma(t)$: the 'center' of this conic is at the origin (the case $a = 0$ above).

2. The discriminant of the conic (10) is $-2fY^3$. With $f < 0$ and the oval γ in the region $y \geq 0$ this discriminant is positive, so the conic $\mathcal{D}(t)$ is a *hyperbola.*

3. Equation (10) can be written in the form
$$(pY - qX)^2 = fqY(Y - 2q).$$
Note that, in p, q coordinates, $pY - qX = 0$ is the equation of the line $\mathcal{L}$ joining the origin to (X, Y), $q = 0$ is the equation of the tangent $\mathcal{T}$ at the origin, and $2q = Y$ is the equation of the line $\mathcal{M}$ parallel to $\mathcal{T}$ passing through the midpoint of the segment from the origin to (X, Y). Thus with a suitable choice of equations for these lines (multiplying by constants), the conic $\mathcal{D}(t)$ can be written $\mathcal{L}^2 = \mathcal{M}\mathcal{T}$.

4. One consequence of the last fact is that $\mathcal{D}(t)$ passes through the midpoint of the segment from the origin to (X, Y), and is tangent to the line $\mathcal{M}$ there. Thus it is a hyperbola with one branch having 3-point contact with the oval γ at the origin, and the other branch tangent to $\mathcal{M}$ at the point $\mathcal{L} \cap \mathcal{M}$.

Equation (10) is a 1-parameter family of conics, in (p, q) coordinates, parametrized by t. The envelope of this family is obtained by differentiating (10) with respect to t:
$$q^2XX' + (p^2 - fq)YY' - pq(XY' + X'Y) + fq^2Y' = 0. \tag{11}$$

Equation (11) is the same as the additional condition required for the conic $\mathcal{C}$ to be *tangent* to γ at $(X(t), Y(t))$. Thus *the envelope of the conics (10) is the locus of centers of conics $\mathcal{C}$ having 3-point contact with γ at the origin and tangent to γ somewhere else.* For a given t we expect (10) and (11) to have a unique solution (p, q), since there should be a unique conic having 3-point contact with γ at the origin $\gamma(t_0)$ and 2-point contact with γ elsewhere – at $\gamma(t)$. (If $t = t_0$ then we might still expect a unique solution, namely the osculating conic for γ at $\gamma(t_0)$.)

Taking $Y' \times (10) - Y \times (11)$ gives
$$-pY(XY' - X'Y) + q(fYY' + X(XY' - X'Y)) = 0. \tag{12}$$

Note that assuming $t \neq t_0$, so that $Y \neq 0$, we can recover (11) from (10) and (12). It follows that the envelope point of the family of conics (10) is the unique intersection of (10) for a particular t with the line (12), disregarding the intersection at the origin $\gamma(t_0)$.

A short calculation shows that the limiting line (12) as $t \to t_0$ is the affine normal $2p\alpha^2 + q\beta = 0$, using the local parametrization of γ is (9). The limiting conic (10) is the affine normal together with the tangent to γ at the origin.

4.1.1 Double 3-point contact conics

We now derive the condition for the conic $\mathcal{C}$ to have 3-point contact with γ at $\gamma(t)$ as well as at the origin. This amounts to differentiating (11) again, so corresponds to *points of regression* on the envelope of the conics (10): *The point (p,q) is on the $\mathcal{AESS}$, i.e. is the center of a conic $\mathcal{C}$ having 3-point contact with γ at $\gamma(t_0)$ and $\gamma(t)$, precisely when $\gamma(t)$ is a point of regression ([9, §5.26] – generically a cusp – on the envelope of the conics (10).*

There are various ways of calculating the condition on t (and t_0) for this to hold. One way is to differentiate (10) again, and to derive a second line through the origin which meets $\mathcal{D}(t)$ in the intersection of $\mathcal{D}(t)$ with this second derivative. This comes to

$$\begin{aligned} & p \quad (-2XYY'^2 - XY^2Y'' + 2X'Y^2Y' + X''Y^3) \\ + & q \quad (X^2Y'^2 + X^2YY'' + 2fYY'^2 + fY^2Y'' - XX''Y^2 - X'^2Y^2) = 0. \end{aligned}$$

We can then calculate the condition for this line to coincide with (12), which comes to

$$fY^3(X''Y' - X'Y'') = (XY' - X'Y)^3. \tag{13}$$

Note that the coefficient of f will not be zero so long as $t \neq t_0$: $Y \neq 0$ as the oval is on one side of the tangent at the origin, and $X''Y' - X'Y'' \neq 0$ as the curvature is nonzero.

There are a number of interesting consequences of (13), which we present now.

4.1.2 Euclidean condition for the double 3-point contact conics

We can turn (13) into a simple geometric criterion depending on Euclidean properties of γ. Thus we may assume Euclidean unit speed. Then $XY' - X'Y$ is, up to sign, the perpendicular distance from $\gamma(t_0)$ to the tangent to γ at $\gamma(t)$. Similarly Y is the perpendicular distance from $\gamma(t)$ to the tangent at $\gamma(t_0)$, which is the x-axis. Finally the other terms in (13) can be interpreted as curvatures (κ), and we obtain

Theorem 3 *The (Euclidean) condition for a conic to have 3-point contact with γ at both $\gamma(t_0)$ and $\gamma(t)$ is (up to sign)*

$$\left(\frac{\kappa(t)}{\kappa(t_0)}\right)^{1/3} = \frac{\text{distance from } \gamma(t_0) \text{ to tangent at } \gamma(t)}{\text{distance from } \gamma(t) \text{ to tangent at } \gamma(t_0)}.$$

The reader should not be surprised by the appearance of the 1/3 power in the expression above. From the section on basic affine differential geometry, and the literature on the subject, it is probably clear already that 1/3 is the 'magic' number of planar affine differential geometry.

4.1.3 Affine condition for the double 3-point contact conic

We can turn the condition (13) into an affine invariant condition by using affine arclength s with $X'Y'' - X''Y' = 1$ instead of the arbitrary parametrization t. Then $f = -1/\kappa(s_0) = -X'(s_0)^3$ and the condition reduces to a perfect cube. Taking the cube root gives $X'(s_0)Y(s) = X(s)Y'(s) - X'(s)Y(s)$, that is

$$\begin{vmatrix} X(s) & X'(s) + X'(s_0) \\ Y(s) & Y'(s) + Y'(s_0) \end{vmatrix} = 0,$$

since $Y'(s_0) = 0$. This condition can be written for the general case ($\gamma(s_0) \neq (0,0)$) as well.

Theorem 4 *The (affine) condition for a conic to have 3-point contact with γ at both $\gamma(s_0)$ and $\gamma(s)$ is*

$$[\gamma(s) - \gamma(s_0), \gamma'(s) + \gamma'(s_0)] = 0. \tag{14}$$

This relation defines the *pre-$\mathcal{AESS}$* in parameter space (s_0, s). Note that this is exactly analogous to the Euclidean case, where the condition for a bitangent circle can be written as the above when Euclidean arclength t is used, so that $\gamma'(t) + \gamma'(t_0)$ is along the bisector of the angle between the two oriented tangents to the curve γ at $\gamma(t)$ and $\gamma(t_0)$. In spite of this, Equation (14) might still be a bit of a surprise. The following Theorem gives the geometric interpretation of this relation.

Theorem 5 *If γ is a conic, then (14) holds for every two points on γ.*

Proof: Assume γ is parametrized by affine arclength s. We prove the Theorem for ellipses and parabolas. The proof is left to the reader for hyperbolas.

- *Ellipses:* Since parallelism is preserved under affine transformations, it is enough to present the proof for a circle. We can follow several approaches to prove that (14) holds for any two points on a circle. The first approach is just geometric: Drawing the tangent vectors at two arbitrary points, it is easy to show that $\gamma(s_0) - \gamma(s_1)$ is parallel to $\gamma'(s_0) + \gamma'(s_1)$. Another alternative is just to write $\gamma = (\cos\theta, \sin\theta)$, $\theta \in [0, 2\pi]$, and compute the matrix and determinant involved, showing that it vanishes.

- *Parabolas:* From Equation (3), and since the affine curvature μ is zero for parabolas, we have that $\gamma''' = 0$, or $\gamma = \mathbf{A}s^2 + \mathbf{B}s + \mathbf{C}$, where $\mathbf{A}, \mathbf{B}, \mathbf{C} \in \mathbb{R}^2$. Substituting this on (14), we obtain the desired result. □

4.1.4 Existence of double 3-point contact conics

An important consequence of (13) is that we shall always find a conic having 3-point contact at $\gamma(t_0)$ and *somewhere* else. In fact, writing γ locally as in (9) we find that the limiting value of the right hand side of (13) is $-1/2\alpha$. Note that we have actually chosen f so that this holds: this was in order to make the conic $\mathcal{C}$ have 3-point contact with γ at the origin. So (13), regarded as an equation for t, has a solution at $t = t_0$. In fact a direct calculation shows that there is at least a *triple* zero at t_0. To see this, take $t_0 = 0$ and use $X = t, Y = \alpha t^2 + \beta t^3 + \delta t^4 + \ldots$. Cross-multiplying (13) we find that the coefficients of t on both sides are $6\alpha^2\beta$, and the coefficients of t^2 on both sides are $12\alpha\beta^2 + 9\alpha^2\delta$. The coefficients of t^3 agree on the two sides if and only if

$$\alpha^2\delta - 3\alpha\beta\delta + 2\beta^3 = 0.$$

It is not hard to check that this is the condition for γ to have a sextactic point at $\gamma(t_0)$.

If we assume that γ does not have a sextactic point at $\gamma(t_0)$ then (13) provides us with a function of t on the oval which has exactly a *triple* zero at $t = t_0$. Consequently it must have another zero somewhere else. Thus

Theorem 6 *Provided γ does not have a sextactic point at $\gamma(t_0)$, there is always a conic with 3-point contact at $\gamma(t_0)$ and at some* other *point $\gamma(t), t \neq t_0$. If there is a sextactic point at $\gamma(t_0)$ then of course there is a six-point contact conic at $\gamma(t_0)$, which can be regarded as a limiting case of the double-3-point contact conic.*

4.1.5 Center of the double 3-point contact conic

Suppose that we know two points of the curve γ which satisfy the equation (14). How can we find the center of the conic $\mathcal{C}$ which has 3-point contact with γ at $\gamma(t_0)$ and $\gamma(t)$? First we know that it lies along the line (12). A vector along this line is

$$(fYY' + X(X'Y - XY'),\ \ Y(X'Y - XY')).$$

Now using affine arclength and assuming (14) we have as before $f = -X'(t_0)^3$ and $X'(t_0)Y(t) = X(t)Y'(t) - X'(t)Y(t)$, so that the above vector is parallel to

$$-[\gamma'(t_0), \gamma'(t)]\gamma'(t_0) + (\gamma(t) - \gamma(t_0)).$$

(Note that $[\gamma'(t_0), \gamma'(t)] = X'(t_0)Y'(t)$ since $Y'(t_0) = 0$.)

In the case when $\mathcal{C}$ has 3-point contact at *both* $\gamma(t_0)$ and $\gamma(t)$ there will be a symmetrical formula with the roles of t_0 and t reversed. We then have that the center of the conic $\mathcal{C}$ must be expressible as

$$\begin{aligned} \text{center} &= \lambda_1(-[\gamma'(t_0), \gamma'(t)]\gamma'(t_0) + (\gamma(t) - \gamma(t_0)) + \gamma(t_0) \\ &= \lambda_2(-[\gamma'(t), \gamma'(t_0)]\gamma'(t) + (\gamma(t_0) - \gamma(t)) + \gamma(t) \end{aligned} \tag{15}$$

for real numbers λ_1, λ_2. Using the equality of these two expressions we get

$$(\lambda_1\gamma'(t_0) + \lambda_2\gamma'(t))[\gamma'(t_0), \gamma'(t)] = (\lambda_1 + \lambda_2 - 1)(\gamma(t) - \gamma(t_0)).$$

But we already know that $\gamma(t) - \gamma(t_0)$ is parallel to $\gamma'(t_0) + \gamma'(t)$, so, provided $\gamma'(t)$ and $\gamma'(t_0)$ are not parallel, we can deduce that $\lambda_1 = \lambda_2$. This means that the center of $\mathcal{C}$ is defined by (15) where λ_1 is the number uniquely determined by the condition (14). Explicitly, if $\gamma'(t_0) + \gamma'(t) = \nu(\gamma(t) - \gamma(t_0))$ then

$$\lambda_1 = \frac{1}{2 - \nu[\gamma'(t_0), \gamma'(t)]}.$$

If the denominator here is zero, then $\mathcal{C}$ must be a parabola, with center at infinity. The two lines through $\gamma(t_0)$ and $\gamma(t)$ must also be parallel.

Examples of the $\mathcal{AESS}$ are given in Figure 4, generated as well with the help of the Maple software [13] and the Liverpool Surface Modeling Package [23]. In Figure 5, a second example is presented. Note that the curve is now non-convex. Basically, the inflection points are 'ignored' where searching for the pre-$\mathcal{AESS}$. The analysis of the $\mathcal{AESS}$ for non-convex curves will be reported elsewhere.

4.2 Envelopes of lines

There is an attractive way to describe the $\mathcal{AESS}$ in terms of envelopes of lines instead of envelopes of conics and since this gives certain information which the other approach does not (at any rate not so easily) we give some details here. In fact this approach makes it clear that the $\mathcal{AESS}$ is from the point of view of singularity theory *the dual of a discriminant* as opposed to the $\mathcal{ADSS}$ which is a *full bifurcation set.* The Euclidean symmetry set is *both* of these things; for details see [16].

Consider an oval γ and two points on it, $\gamma(t_1), \gamma(t_2)$. We consider initially conics $\mathcal{C}$ which are tangent to γ at these two points. It is known from the general theory of conics that these conics form a pencil; in fact if $\mathcal{T}_1, \mathcal{T}_2$ are the tangent lines to the oval at the two points and $\mathcal{L}$

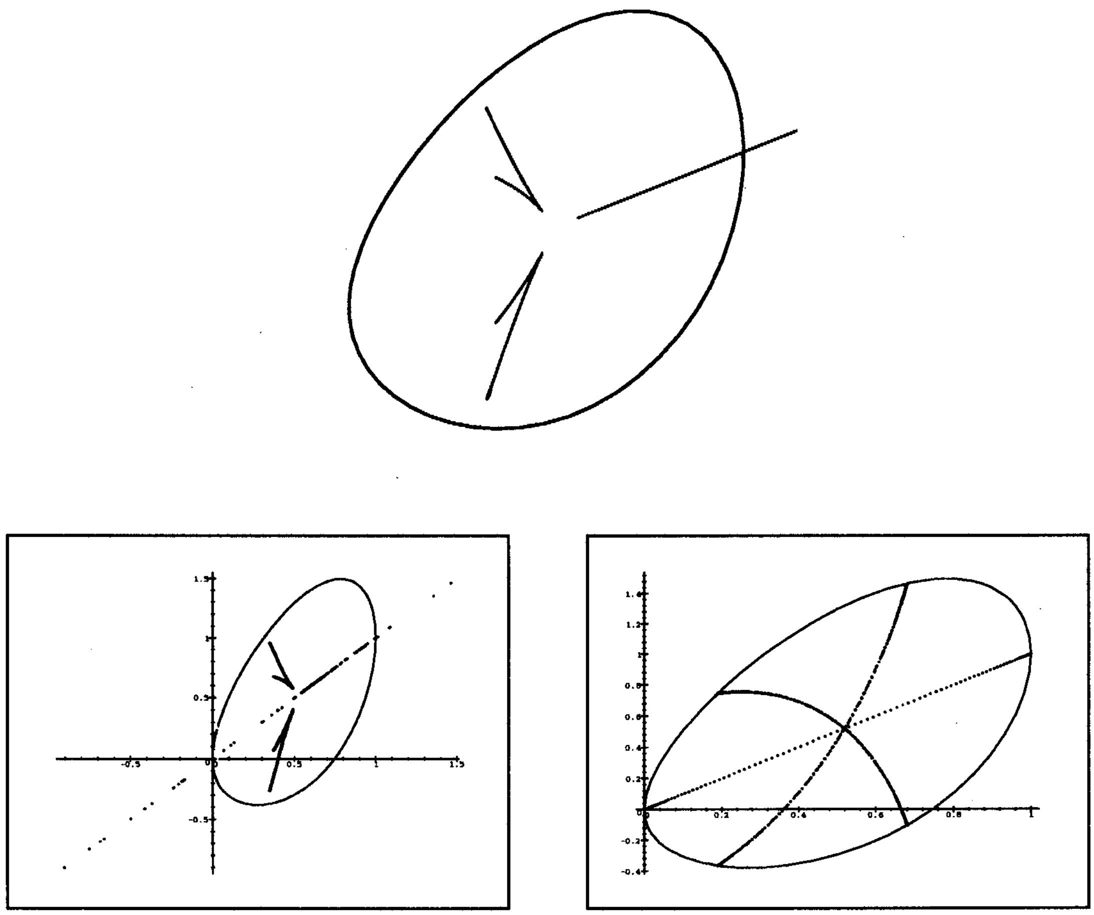

Figure 4: *Example of the $\mathcal{AESS}$. The top figure shows the $\mathcal{AESS}$ for the oval in Figure 2, and the bottom-left shows again points in the $\mathcal{AESS}$, together (right) with the midlocus for the same oval. (This is a color figure.)*

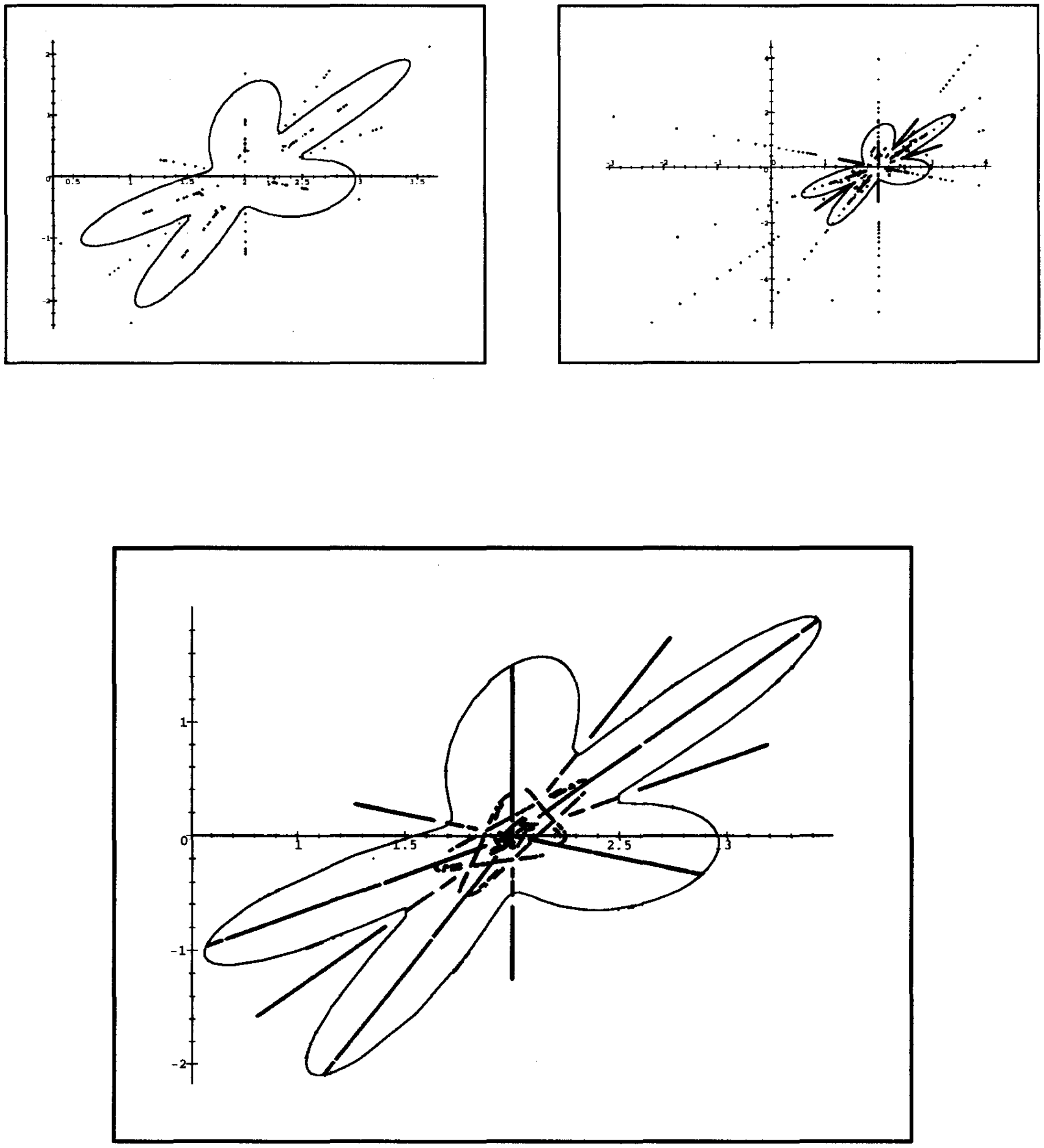

Figure 5: *Example of the $\mathcal{AESS}$ for a non-convex curve. The top-left figure shows a few points on the $\mathcal{AESS}$, while the top-right shows a more detailed part (more points) of it. On the bottom, the midlocus for the same curve is shown. Note how the $\mathcal{AESS}$ and midlocus capture the affine symmetry of the different parts of the shape. (This is a color figure.)*

is the line joining the two points, then the general conic of the pencil is $\mathcal{T}_1\mathcal{T}_2 + \lambda\mathcal{L}^2 = 0$. Also the locus of centers of these conics is the line $\mathcal{M}$ passing through the intersection of $\mathcal{T}_1$ and $\mathcal{T}_2$, and the midpoint of the chord from $\gamma(t_1)$ to $\gamma(t_2)$. (There is another degenerate component of the locus of centers, namely the line $\mathcal{L}$, which corresponds to the double line $\mathcal{L}^2$ of the pencil, whose center is indeterminate on $\mathcal{L}$.) These lines are shown in Figure 6.

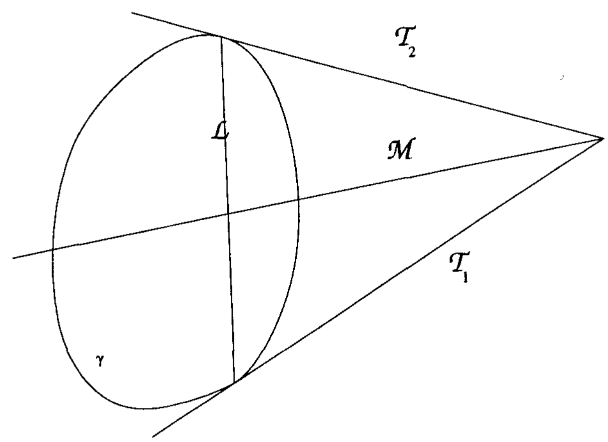

Figure 6: *Example of the lines used for the construction of the $\mathcal{AESS}$ as an envelope of lines (see text).*

We examine this situation in more detail as follows. The conic $\mathcal{C}$ will have six homogeneous coefficients, being the general equation of the form

$$ax^2 + by^2 + 2hxy + 2gx + 2fy + c = 0. \tag{16}$$

Writing $\gamma(t) = (X(t), Y(t))$, we can substitute $x = X(t)$, $y = Y(t)$ into (16) to measure the contact of $\mathcal{C}$ with the oval at $\gamma(t)$. Calling the resulting function $\mathcal{C}(t)$ we say that there is k-point contact when $\mathcal{C}$ and its first $k-1$ derivatives vanish at t while the k-th derivative is non-zero there. Furthermore it is well known that the coordinates of the center (p, q) of $\mathcal{C}$ satisfy the two linear equations $\mathcal{C}_x(p,q) = 0, \mathcal{C}_y(p,q) = 0$, where here and later subscripts x, y, t denote again partial derivatives. Assuming $\mathcal{C}$ is tangent to γ at both points $\gamma(t_1), \gamma(t_2)$ we can eliminate a, b, h, g, f, c from six linear equations, two coming from each point of contact and two from the center of the conic. In order to write down the resulting determinant condition, we introduce some abbreviations. Thus $\mathcal{C}(t_1)$ stands for

$$(X^2,\ Y^2,\ 2XY,\ 2X,\ 2Y,\ 1),$$

evaluated at $t = t_1$. This is the vector in $\mathbb{R}^6$ such that $\mathcal{C}(t_1)(a, b, h, g, f, c)^\top = 0$ gives the condition for the conic $\mathcal{C}$ to pass through $(X(t_1), Y(t_1))$. The vector $\mathcal{C}(t_2)$ is similarly defined.

By $\mathcal{C}_t(t_1)$ we mean
$$(2XX',\ 2YY',\ 2(XY'+X'Y),\ 2X',\ 2Y',\ 0),$$
evaluated at $t = t_1$. Thus $\mathcal{C}_t(t_1)(a,b,h,g,f,c)^\top = 0$ is the additional condition for $\mathcal{C}$ to be tangent to the $\gamma = (X,Y)$ curve at $(X(t_1), Y(t_1))$. Likewise $\mathcal{C}_x(p,q)$ denotes the vector obtained by differentiating (16) with respect to x, substituting $x = p, y = q$ and writing down the resulting coefficients of a,b,h,g,f,c. The same goes for $\mathcal{C}_y(p,q)$.

The determinant condition that there should exist a conic as in (16), with not all coefficients zero, tangent to γ at $\gamma(t_1), \gamma(t_2)$ and having center (p,q) is then

$$\mathcal{G}(t_1,t_2,p,q) := \begin{vmatrix} \mathcal{C}(t_1) \\ \mathcal{C}_t(t_1) \\ \mathcal{C}(t_2) \\ \mathcal{C}_t(t_2) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{vmatrix} = 0. \tag{17}$$

Of course, (17) is nothing other than the equation, in (p,q) coordinates, of the locus of centers of conics $\mathcal{C}$. Hence, writing $\mathcal{M}(t_1,t_2,p,q) = 0$ for an equation of the line $\mathcal{M}$ above (using again (p,q) as current coordinates), and $\mathcal{L}(t_1,t_2,p,q) = 0$ for an equation of the chord, it must happen that
$$\mathcal{G} = \mathcal{M}\mathcal{L}.$$

Let $\mathcal{G}(t_1,t_2,p,q) = 0$, that is, (p,q) is the center of a conic tangent to γ at $\gamma(t_1)$ and $\gamma(t_2)$. So long as (p,q) is not the center of the chord, where $\mathcal{M}$ meets $\mathcal{L}$, we can deduce that $\mathcal{L}(t_1,t_2,p,q) \neq 0$. Since tangents to a conic at opposite ends of a diameter are parallel, the only case that gives a problem here is where the tangents to γ at $\gamma(t_1)$ and $\gamma(t_2)$ are parallel. For an oval γ this determines t_2 uniquely from t_1. Apart from that, the zeroes of $\mathcal{G}$ coincide with those of $\mathcal{M}$. The same holds for derivatives; for example
$$\mathcal{G}_t = \mathcal{M}_t\mathcal{L} + \mathcal{M}\mathcal{L}_t,$$
so that $\mathcal{G} = \mathcal{G}_t = 0$ is equivalent, away from $\mathcal{L} = 0$, to $\mathcal{M} = \mathcal{M}_t = 0$.

The function $\mathcal{M}$ gives us a 2-parameter family of lines in the plane, parametrized by the pairs of distinct points of γ. It is natural to ask whether $\mathcal{M}$ can be extended smoothly to all pairs of points. The answer is yes, provided we define $\mathcal{M}(t,t,p,q)$ to be an equation of the affine normal to γ at $\gamma(t)$. (A geometric reason for this is the following: For fixed t_1, t_2, the line $\mathcal{M}$ is the locus of centers of conics having (at least) 2-point contact with γ at $\gamma(t_1)$ and $\gamma(t_2)$. When $t_2 \to t_1$ we shall obtain the locus of centers of conics having at least 4-point contact with γ at $\gamma(t_1)$, which is the affine normal at this point.) In fact $\mathcal{M}$ gives rise to a map (compare with [16])
$$\mathcal{B}:\ \gamma \times \gamma \to \mathcal{DL}, \tag{18}$$
where $\mathcal{DL}$ is the dual plane, or the set of lines in the ordinary plane $\mathbb{R}^2$, and $\mathcal{B}(t_1,t_2)$ is the line whose equation in current coordinates (p,q) is $\mathcal{M}(t_1,t_2,p,q) = 0$.

We can consider $\mathcal{M}$ as defining two envelopes of lines, one by fixing t_1 and the other by fixing t_2. We can also use the function $\mathcal{G}$ to measure contact between the conic $\mathcal{C}$ and γ and use the relation $\mathcal{G} = \mathcal{M}\mathcal{L}$ to compare this contact with the properties of the two envelopes. We collect the main results below, and then sketch the method of proof, which amounts merely to differentiating the functions $\mathcal{G}$ and $\mathcal{M}$ sufficiently many times. In what follows, $\mathcal{C}$ will be the conic tangent to γ at $\gamma(t_1)$ and $\gamma(t_2)$, and having center (p,q).

4.2.1 Results

As before, the $\mathcal{AESS}$ of γ is the locus of centers of conics which have 3-point contact with γ at two different places (or, in the limit, the centers of sextactic conics). We are now ready to derive the $\mathcal{AESS}$ as envelope of lines.

Theorem 7 *For $\mathcal{M}$, $\mathcal{G}$, and $\mathcal{B}$ as defined in §4.2, the following holds:*

1. *Consider the 1-parameter family of lines $\mathcal{M}(t_1, t_2, p, q) = 0$ where t_2 is fixed. A point (p, q) belongs to the envelope of this family, given by $\mathcal{M} = \partial\mathcal{M}/\partial t_1 = 0$, if and only if the conic $\mathcal{C}$ has 3-point contact (at least) with γ at $\gamma(t_1)$.*

2. *Fixing t_2, p, q, the function $\mathcal{G}$ is zero and has a singularity of type A_k at t_1 (i.e. the first k derivatives vanish there but the $(k+1)$-st derivative is nonzero), if and only if the conic $\mathcal{C}$ has $(k+2)$-point contact with γ at $\gamma(t_1)$. The previous result is the case $k = 1$ of this. The functions $\mathcal{G}$ and $\mathcal{M}$ have the same singularity type so long as (p, q) is not on the line $\mathcal{L}$.*

3. *(t_1, t_2) is a critical point of $\mathcal{B}$ if and only if there is a conic with 3-point contact at both $\gamma(t_1)$ and $\gamma(t_2)$. To say that (t_1, t_2) is a critical point of $\mathcal{B}$ amounts to saying that (p, q) is a point of the envelope of both families of lines given by $\mathcal{M}$, keeping t_1 constant and keeping t_2 constant, so this follows from Part 1 above.*

4. *At points where $\mathcal{M} = \partial\mathcal{M}/\partial t_1 = \partial\mathcal{M}/\partial t_2 = 0$, we have $\partial^2\mathcal{M}/\partial t_1\partial t_2 = 0$. (It is not obvious to us what this condition means geometrically.) We note here that taking the Euclidean case where $\mathcal{M}$ is replaced by the perpendicular bisector of the segment joining $\gamma(t_1)$ and $\gamma(t_2)$ (compare with [16]), the formula $\partial^2\mathcal{M}/\partial t_1\partial t_2 = 0$ holds identically since it is easy to see that $\mathcal{M}$ is a sum of functions of t_1 and t_2. But in the present case this is not so: we need $\mathcal{M} = \partial\mathcal{M}/\partial t_1 = \partial\mathcal{M}/\partial t_2 = 0$.*

5. *The line $\mathcal{M}$ is tangent to the $\mathcal{AESS}$ when there is a conic $\mathcal{C}$ having 3-point contact with γ at $\gamma(t_1), \gamma(t_2)$. This says that the tangent to the $\mathcal{AESS}$ passes through the midpoint of the chord joining the two points of γ and through the intersection of the tangents to γ at these two points.*

   Putting this in a more formal way: the critical locus of $\mathcal{B}$ (the image of the critical set) is the set of tangent lines to the $\mathcal{AESS}$ of γ, that is the dual of the $\mathcal{AESS}$ of γ. This is completely analogous to the situation studied in [16].

6. *If the conic $\mathcal{C}$ has k-point contact ($k \geq 4$) with γ at $\gamma(t_1)$, then the dual of the $\mathcal{AESS}$ of γ has an inflexion of order $k-3$. An inflexion of order 1 – an ordinary inflexion – corresponds to an ordinary cusp on the $\mathcal{AESS}$, and arises when the conic $\mathcal{C}$ has 4-point contact.*

4.2.2 Methods

We shall not give the full proofs of the above statements, but indicate the methods, all of which are elementary.

For definiteness, let us fix $\gamma(t_2)$ at the origin and the tangent there as the x-axis. Then the conic $\mathcal{C}$ has the special form

$$ax^2 + by^2 + 2hxy + 2fy = 0,$$

with four (homogeneous) coefficients. We rename t_1 as t for this section. The line $\mathcal{M}$ joins the intersection of the tangent at $\gamma(t)$ and the x-axis to the midpoint of the segment from (0,0) to $\gamma(t)$. The function $\mathcal{G}$ will now be given by eliminating a, b, h, f and will be the 4×4 determinant

$$\mathcal{G}(t,p,q) := \begin{vmatrix} \mathcal{C}(t) \\ \mathcal{C}_t(t) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{vmatrix}. \tag{19}$$

In (p, q) coordinates $\mathcal{G} = 0$ is the equation of the locus of centers of the conics tangent to γ at $\gamma(t)$ and tangent to the x-axis at the origin. Writing $\gamma(t) = (X(t), Y(t))$ we have

$$\mathcal{G} = \mathcal{M}(Xq - Yp)$$

for a suitable equation $\mathcal{M} = 0$ of the line $\mathcal{M}$. We shall assume that the second factor here is nonzero, i.e. that (p, q) is not the midpoint of the chord, which amounts to saying that the tangent to γ at $\gamma(t)$ is not parallel to the x-axis. Then $\mathcal{G}$ and its first k derivatives with respect to t vanish at (t, p, q) if and only if the same applies to $\mathcal{M}$. We need to connect the number of derivatives of $\mathcal{G}$ which vanish with the order of contact of the conic $\mathcal{C}$ with γ at $\gamma(t)$.

We again write $\mathcal{C}(t)$ for the result of substituting $x = X(t), y = Y(t)$ in the equation of $\mathcal{C}$, and, as above, use $\mathcal{C}(t)$ also for the vector $(X(t)^2, Y(t)^2, 2X(t)Y(t), 2Y(t))$. Differentiating $\mathcal{G}$ with respect to t, we obtain

$$\frac{\partial \mathcal{G}}{\partial t}(t,p,q) = \begin{vmatrix} \mathcal{C}_t(t) \\ \mathcal{C}_t(t) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{vmatrix} + \begin{vmatrix} \mathcal{C}(t) \\ \mathcal{C}_{tt}(t) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{vmatrix} = \begin{vmatrix} \mathcal{C}(t) \\ \mathcal{C}_{tt}(t) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{vmatrix}. \tag{20}$$

Suppose that $\mathcal{G} = \partial\mathcal{G}/\partial t = 0$ at (t, p, q). Now the condition for $\mathcal{C}$ to have 3-point contact with γ at $\gamma(t)$ is that the same values of a, b, h, f should allow $\mathcal{C} = \mathcal{C}_t = \mathcal{C}_{tt} = 0$, i.e. that the matrix

$$\begin{pmatrix} \mathcal{C}(t) \\ \mathcal{C}_t(t) \\ \mathcal{C}_{tt}(t) \\ \mathcal{C}_x(p,q) \\ \mathcal{C}_y(p,q) \end{pmatrix} = \begin{pmatrix} X^2 & Y^2 & 2XY & 2Y \\ * & * & * & * \\ * & * & * & * \\ p & 0 & q & 0 \\ 0 & q & p & 1 \end{pmatrix} \tag{21}$$

(the stars representing the second and third rows), should have rank < 4. However, it is easy to check that the three rows given in the right hand matrix are independent, unless one of the following happens:

(i) $p = q = 0$: the center of the conic $\mathcal{C}$ is at the fixed point of γ. This is excluded because we assume the center is on $\mathcal{M}$.

(ii) $Y = q = 0$: this is certainly excluded by γ being an oval, since the only point of γ on the x-axis will be the origin.

(iii) $X = Y = 0$: excluded because $\gamma(t)$ is not at the origin.

(iv) $(p, q) = (\frac{1}{2}X, \frac{1}{2}Y)$: i.e. the center of $\mathcal{C}$ is at the midpoint of the chord, which implies that the tangent to γ at $\gamma(t)$ is parallel to the tangent at the origin. We have excluded this case above.

Now from $\mathcal{G} = 0$ we know that rows 1, 2, 4, 5 of (21) are dependent, i.e. row 2 lies in the space V^3 spanned by rows 1, 4, 5. Similarly from $\partial\mathcal{G}/\partial t = 0$ we know that row 3 lies in V. Hence the matrix in (21) has rank 3 as required, which means that the conic $\mathcal{C}$ has 3-point contact with γ at $\gamma(t)$ when $\mathcal{G} = \partial\mathcal{G}/\partial t = 0$, or equally when $\mathcal{M} = \partial\mathcal{M}/\partial t = 0$. This proves Part 1 of the results in §4.2.1, and repeating the argument with more derivatives of $\mathcal{G}$ proves Part 2.

Part 4 is proved in the same way but using the general form of $\mathcal{G}$ in (17). Part 5 follows from the fact that the critical set of $\mathcal{B}$ is the set of

$$(t_1, t_2) \text{ such that } \mathcal{M} = \frac{\partial\mathcal{M}}{\partial t_1} = \frac{\partial\mathcal{M}}{\partial t_2} = 0 \text{ for some } p, q.$$

Consider in fact the set $\mathcal{S}$ consisting of points

$$(t_1, t_2, p, q) \text{ such that } \mathcal{M} = \frac{\partial\mathcal{M}}{\partial t_1} = \frac{\partial\mathcal{M}}{\partial t_2} = 0.$$

The $\mathcal{AESS}$ is the projection of this set to the (p, q)-plane. It is easy to check that the tangent line to $\mathcal{S}$ projects to the line $\mathcal{M}$ itself, which is therefore tangent to the $\mathcal{AESS}$. (This is analogous to the fact that the tangent to the envelope of a 1-parameter family of lines is always the current line of the family; see for example [9, §5.25].)

Finally Part 6 is just a restatement of Part 2, using the interpretation of Part 5.

4.2.3 Straight affine envelope symmetry sets

We have seen that a straight $\mathcal{ADSS}$ does not necessarily mean that the curve giving rise to it is affine symmetric. For the $\mathcal{AESS}$ the situation is different and we give details below.

Theorem 8 *Suppose that the $\mathcal{AESS}$ for two arcs of an oval γ is a straight line. Then either of these arcs is obtained from the other by an affine transformation of determinant -1.*

Proof: We take the straight line to be along the x-axis. Now, the tangent to the $\mathcal{AESS}$ at a point corresponding to two parameter values (t_1, t_2), always passes through two points (see part 5 of §4.2.1):
(i) the midpoint of the chord joining these two curve points;
(ii) the intersection of the two tangents at these two curve points.
It follows that in the present case, all these midpoints and all these intersections of corresponding tangents lie on the x-axis.

Let one of the arcs be parametrized $(X(t), Y(t))$. The tangent to this arc meets the x-axis at the point with x-coordinate $(XY' - X'Y)/Y'$. The other tangent to the oval through this point meets the oval at say $(U(t), V(t))$, and we therefore have

$$(XY' - X'Y)V' = (UV' - U'V)Y' \tag{22}$$

for all values of t.

Now the midpoint of the chord also lies on the x-axis, so that $Y + V = 0$ for all t. This gives $Y' + V' = 0$ for all t, and substituting in (22) we get, assuming Y and $Y' \neq 0$,

$$XY' - X'Y = UY' - U'Y, \quad \text{i.e.} \quad \frac{d}{dt}\left(\frac{X}{Y}\right) = \frac{d}{dt}\left(\frac{U}{Y}\right),$$

which gives

$$X = U + \lambda Y,$$

for a constant λ. (The assumption $Y \neq 0$ is harmless since the oval will only meet the x-axis in two points, and the condition $Y' \neq 0$ just avoids tangents which are parallel to the x-axis. Clearly these do not affect the result.)

Thus the arc (U,V) is obtained from the arc (X,Y) by the affine transformation with matrix

$$\begin{pmatrix} 1 & -\lambda \\ 0 & -1 \end{pmatrix}$$

which has determinant -1. □

Note of course that the condition $[\gamma(t_1) - \gamma(t_2), \gamma'(t_1) + \gamma'(t_2)] = 0$, which holds at corresponding pairs of the oval which give points of the $\mathcal{AESS}$, automatically holds in the case of the theorem above, because of the affine transformation which takes one piece to the other. What is interesting to note is that if the $\mathcal{AESS}$ has an inflection, then the condition for the $\mathcal{ADSS}$ holds, that is, $[\gamma(t_1) - \gamma(t_2), \gamma''(t_1) - \gamma''(t_2)] = 0$. This means that if the $\mathcal{AESS}$ has an inflection the corresponding pre-$\mathcal{ADSS}$ and pre-$\mathcal{AESS}$ coincide. Then, for affine symmetric shapes, the pre-$\mathcal{ADSS}$ and pre-$\mathcal{AESS}$ have common segments. Furthermore, both affine symmetry sets have common arcs.

In figures 2, 4, and 5, we observe the relation between affine symmetric shapes and straight symmetry sets. Note that both the $\mathcal{ADSS}$ and the $\mathcal{AESS}$ contain straight segments, since the shapes are affine symmetric (observe the affine curvature in Figure 2). Additional segments appear, since the condition for the pre-images, both for the $\mathcal{ADSS}$ and $\mathcal{AESS}$, do not have to hold just for pair of points belonging to the affine related segments of the curve.

5 Discussion and further research

We have presented two different definitions for affine invariant symmetry sets, one based on distance functions and the second one on affine bitangent conics. A number of interesting issues remain open and will be the subject of subsequent reports. We want to conclude this paper by mentioning some of these issues.

The relation between the $\mathcal{ADSS}$ and the $\mathcal{AESS}$ should be further investigated. We already observed that when the shape is affine symmetric, both the pre-images and the symmetry sets have common segments. This raises questions like if the 'differences' between both symmetry sets can give a measurement of the deviation from affine symmetry of the given shape.

Another open question is the reconstruction from the $\mathcal{ADSS}$ and $\mathcal{AESS}$, being this of extreme significance from the point of view of shape representation.

As pointed out before, following [8], the deformation of the $\mathcal{ADSS}$ and the $\mathcal{AESS}$ can be studied. Actually, all the research done for the classical Euclidean symmetry set can be now carried on for the affine ones described in this paper.

The study of affine invariant Voronoi diagrams is an interesting topic as well. We already mentioned how it can be defined from the $\mathcal{ADSS}$. From the $\mathcal{AESS}$, symmetry sets for polygons can be defined extending the definitions in [5] to 3-point contact of conics with polygons.

Probably more interesting is to extend the definitions and analysis of the $\mathcal{ADSS}$ and $\mathcal{AESS}$ to other groups and dimensions. Regarding the extension to higher dimensions, the $\mathcal{ADSS}$ can be defined for any dimension since the the affine distance exists for higher dimensions as well [24]. Some of the properties in Proposition 1 have their analogue in higher dimensions as well. In the same way, the $\mathcal{AESS}$ can be extended to higher dimensions, requesting for the right order of contact with the right high dimensional shapes. Extending the definition of the $\mathcal{AESS}$ to other groups, as the projective one, requires the use of curves of constant projective curvature instead of conics, and of course requesting the right degree of contact. The extension of the $\mathcal{ADSS}$ is not so straightforward, since a projective invariant distance has to be defined first. Of course, following the definitions of the corresponding invariant symmetry sets, their study will come.

To conclude, we believe that beyond the results here presented, this paper opens the door to some new research in the area of symmetry sets.

Acknowledgments

This work was initiated when the authors were visiting the *Isaac Newton Institute for Mathematical Sciences*, Cambridge University, participating at the *Workshop on New Connections between Mathematics and Computer Science*, November 1995. We thank Prof. Jeremy Gunawardena from BRIMS-HP for inviting us to the workshop, the Newton Institute for its hospitality, and the corresponding funding agencies for supporting our visit. GS thanks Prof. Peter Olver and Prof. Allen Tannenbaum for early conversations on symmetry sets.

References

[1] V. I. Arnold, *Singularity Theory - Selected Papers*, London Math. Soc. Lecture Notes 53, Cambridge University Press, Cambridge, 1981.

[2] V. I. Arnold, *Catastrophe Theory*, Springer-Verlag, Berlin, 1986.

[3] V. I. Arnold, *Singularities of Caustics and Wavefronts*, Kluwer, 1990.

[4] V. I. Arnold, S. M. Gussein-Zade, and A. N. Varchenko, *Singularities of Differentiable Maps*, Birkhauser, Basel, 1985.

[5] T. F. Banchoff and P. J. Giblin, "Global theorems for symmetry sets of smooth curves and polygons in the plane," *Proc. Royal Soc. Edinburgh* **106A**, pp. 221-231, 1987.

[6] W. Blashke, *Vorlesungen uber Differentialgeometrie II: Affine Differentialgeometrie*, Springer, 1923.

[7] H. Blum, "Biological shape and visual science", *J. Theor. Biology* **38**, pp. 205-287, 1973.

[8] J. W. Bruce and P. J. Giblin, "Growth, motion and 1-parameter families of symmetry sets," *Proc. Royal Soc. Edinburgh* **104A**, pp. 179-204, 1986.

[9] J. W. Bruce and P. J. Giblin, *Curves and Singularities*, Cambridge University Press, Cambridge, 1984; second edition 1992.

[10] J. W. Bruce, P .J. Giblin and C. G. Gibson, "Symmetry sets," *Proc. Royal Soc. Edinburgh* **101A** pp. 163-186, 1985.

[11] L. Calabi and W. E. Hernett, "Shape recognition, prairie fires, convex deficiencies and skeletons," *Amer. Math. Monthly* **75**, pp. 335-342, 1968.

[12] E. Calabi, P. J. Olver, and A. Tannenbaum, "Affine geometry, curve flows, and invariant numerical approximations," *Advances in Mathematics*, to appear.

[13] B. W. Char, K. O. Geddes, G. H. Gonnet, B. L. Leong, M. B. Monogan, and S. W. Watt, *Maple V*, Springer-Verlag, New York, 1991.

[14] D. L. Fidal, "The existence of sextactic points," *Math. Proc. Cambridge Philos. Soc.* **96**, pp. 433-436, 1984.

[15] P. J. Giblin and S. A. Brassett, "Local symmetry of plane curves," *Amer. Math. Monthly* **92**, pp. 698-706, 1985.

[16] P. J. Giblin and F. Tari, "Perpendicular bisectors, duality and local symmetry of plane curves," *Proc. Royal Soc. Edinburgh* **125A**, pp. 181-194, 1995.

[17] S. Izumiya and T. Sano, "Generic affine differential geometry of plane curves," *Technical Report* **310**, Department of Mathematics, Hokkaido University, Sapporo, Japan, October 1995.

[18] I. Kovács and B. Julesz, "Perceptual sensitivity maps within globally defined visual shapes," *Nature* **370** pp. 644-646, 1994.

[19] T. S. Lee, D. Mumford, K. Zipser, and P. Schiller, "Neural correlates of boundary and medial axis representation in primate striate cortex," *Investigative Opthalmology and Visual Science Supplement (ARVO Abstract)*, 1995.

[20] T. S. Lee and D. Mumford, "Shape representation in long latency V1 responses," pre-print, 1996.

[21] M. Leyton, "Symmetry-curvature duality," *Computer Vision, Graphics, and Image Processing* **38**, pp. 327-341, 1985.

[22] M. Leyton, *Symmetry, Causality, Mind*, MIT-Press, Cambridge, 1992.

[23] Liverpool Surface Modeling Package, written by Richard Morris (rjm@amsta.leeds.ac.uk) for Silicon Graphics and X windows.

[24] K. Nomizu and T. Sasaki, *Affine Differential Geometry*, Cambridge University Press, Cambridge, 1993.

[25] R. L. Ogniewicz, *Discrete Voronoi skeletons,* Hartung-Gorre Verlag Konstanz, Zurich, 1993.

[26] F. P. Preparata and M. I. Shamos, *Computational Geometry*, Texts and Monographs in Computer Science, Springer-Verlag, New York, 1990.

[27] G. Sapiro and A. Tannenbaum, "On affine plane curve evolution," *Journal of Functional Analysis* **119:1**, pp. 79-120, 1994.

[28] G. Sapiro and A. Tannenbaum, "On invariant curve evolution and image analysis," *Indiana University Mathematics Journal* **42:3**, 1993.

[29] J. Serra, *Image Analysis and Mathematical Morphology*, Vol. 1, Academic Press, New York, 1982.

[30] J. Serra, Editor, *Image Analysis and Mathematical Morphology*, Vol. 2, Academic Press, New York, 1988.

[31] J. Smoller, *Shock Waves and Reaction-Diffusion Equations*, Springer-Verlag, New York, 1983.

[32] S. Tabachnikov, "Parametrized plane curves, Minkowski caustics, Minkovski vertices and conservative line fields," *Max Planck Institute Technical Report*, Bonn, Germany, 1996.