SIAM J. Applied Dynamical Systems
Vol. 13, No. 3, pp. 1106–1136

Differential Geometry Perspective of Shape Coherence and Curvature Evolution by Finite-Time Nonhyperbolic Splitting*

Tian Ma† and Erik M. Bollt‡

Abstract. Mixing and coherence are essential topics for understanding and describing transport in fluid dynamics and other nonautonomous dynamical systems. Only recently has the idea of coherence gained more serious footing, particularly with the recent advances of finite-time studies of nonautonomous dynamical systems. Here we define *shape coherent sets* as a means to emphasize the intuitive notion of ensembles which "hold together" for some period of time, and we contrast this notion to other recent perspectives of coherence, notably "coherent pairs," and likewise also to the geodesic theory of material lines. We will relate shape coherence to the differential geometry concept of curve congruence through matching curvatures. We show that points in phase space where there is a zero-splitting between stable and unstable foliations locally correspond to points where curvature will evolve only slowly in time. Then we develop curves of points with zero-angle, meaning nonhyperbolic splitting, by continuation methods in terms of the implicit function theorem. From this follows a simple ODE description of the boundaries of shape coherent sets. We will illustrate our method with popular benchmark examples and further investigate the intricate structure of foliation geometry.

Key words. shape coherent set, curvature evolution, finite-time stable and unstable foliations, implicit function theorem, continuation, mixing, transport

AMS subject classifications. 37C60, 57R30

DOI. 10.1137/130940633

1. Introduction.

Finite-time mixing and transport mechanisms in two-dimensional fluid flows are a classic and long-standing problem area in dynamical systems. We are interested here especially in coherent structures in the flow. For the autonomous case, there are several studies on almost invariant sets; see [6, 16, 4], for example. For the nonautonomous case, Lagrangian coherent structures (LCSs) based on finite-time Lyapunov exponents (FTLEs) that focus on maximal local stretching have become a popular computational way to study transport (see [21, 22, 37]), but caveats regarding the possibility of false positives have been revealed [23, 12]; similarly finite sized Lyapunov exponents (FSLEs) show differences and likely false positives [33]. On the other hand, transfer operator methods based on Galerkin–Ulam matrices which evolve distributions of ensembles of initial conditions have also been successfully developed for finding so-defined coherent pairs [11, 16, 18]. There are also some recent works on improving, comparing, and generalizing these methods [39, 30]. An exciting

*Received by the editors October 21, 2013; accepted for publication (in revised form) by B. Sandstede April 29, 2014; published electronically July 24, 2014.
http://www.siam.org/journals/siads/13-3/94063.html
†Department of Mathematics and Computer Science, Clarkson University, Potsdam, NY, 13699 (mat@clarkson.edu).
‡Corresponding author. Department of Mathematics and Computer Science, Clarkson University, Potsdam, NY, 13699 (bolltem@clarkson.edu).

recent development involves minimally stretching material lines by studying the Cauchy–Green strain tensor [1, 24] in a so-called theory of geodesic curves. In [11], the authors developed a theory of geodesic curves whereby they define a principle that coherent sets should correspond to boundaries that are minimally stretching. Our definition here of shape coherent sets is agreeable with this geodesic curve principle by Farazmand–Haller, but from a different perspective, since we will discuss that evolution of curvature is also related to evolution of arc length, whereas they discuss evolution of length of boundary curves.

However, popularly general notions of coherence broadly interpreted may include allowing for sets that may stretch and fold when considering material curves of set boundaries and ensembles of set of points therein that advect with the flow. We note that if they are highlighted by coloring (partition), then general sets automatically catch the eye, especially when remembering the simple notion that dynamical systems, being continuous, will preserve properties such as connectedness. From this perspective, almost any connected set that is highlighted may appear coherent in the sense of continuity. We take a narrower view here.

Here we introduce a different but related mathematical definition of coherence by defining *shape coherent sets* which we so named to distinguish our definition from other possible notions of the phrase coherence. We will show that our intuitively motivated notion of shape coherence directly addresses the idea of sets, evolving in such a way as to "hold together." Our perspective is strictly Lagrangian and is meant to capture the typical idea that a coherent set should "catch your eye," as often described roughly when suggesting coherence. We mathematically define this concept in terms of stating that *shape coherence* should correspond (approximately at least) to a region of phase space where, over the finite-time epoch of consideration, the flow restricted to that region is equivalent to a member of the group of rigid body motions.

Analysis of our definition of shape coherence from a perspective of differential geometry reveals that sets which have boundaries that have relatively slowly changing curvature of their boundary curves correspond to shape coherent sets. From this observation we relate that curves with slowly evolving curvature are related to curves with a zero-splitting property, meaning a shear-like scenario in which, when stable and unstable foliations are parallel, and a curve of such points exists, these curves evolve curvature relatively slowly. This geometric interpretation by curvature evolution is agreeable with the theory of geodesic curves derived through variational principles in [24], from which it follows that such curves follow as solutions of the ODE whose vector field is defined by unstable foliations; the connection to coherence is intuitive in that minimally stretching boundary curves suggest slowly changing shapes.

We highlight [8], in which the authors make an observation in discussing a "Lyapunov diffusivity." They propose a quantity that depends on particle dispersion in the forward time Lyapunov exponent (future) correction in terms of angle α between stable and unstable Lyapunov vectors, $\tilde{\lambda} = \lambda \sin^2 \alpha$; from this they note that when $\alpha = 0$, shear would dominate, and they note that shear is consistent with a transport barrier. There have been studies of evolution of curvature in the literature [29, 10, 9, 34, 25] focused on strongly chaotic regions or turbulent regions. In more recent studies, Thiffeault [41, 40, 42], relating a distribution of curvature growth rate, found long tails and a corresponding scaling law describing the folding points, and a small range describing the stretching regions, but with a relatively flat intermediate scale suggestive of ellipticity. The key difference between these previous works and the appearance of curvature here is that past work was focused on the role of curvature in the

description of the chaotic set, whereas we show that curvature can also be used to describe the complement, which is indicative of an equivalence to shape-coherent sets. See also [32, 43, 27] for studies of curvature of streamlines as descriptive of the topological aspects of experimental turbulent flows.

Since it is well known in differential geometry [7] that curvature directly defines curve congruence, here we link directly to our definition of a shape coherent set through the Frenet–Serret formulae of curvature [7, 14, 36] and the so-called Frenet frame. We present a constructive analysis based on the implicit function theorem; our analysis shows that such zero-splitting curves may be shown to exist and they may be numerically constructed by adaptations of standard continuation methods. We give two examples with the popular benchmarks of transport in the Rossby wave system and the double gyre. We also make some geometric observations that angles between stable and unstable foliations develop zeros with increasing time epoch in a manner that relates to the prominence of the coherent sets, and also related to the accumulation of stable and unstable manifolds progressively revealing Cantor-like structure.

2. Shape coherent sets.

Here we will describe our perspective of coherence in terms of our definition of "shape coherence." We will compare and contrast this definition with the popular coherent pairs definition [18] based on measurable dynamics and transfer operators, and also to the geodesic curve theory [11], from which we will realize that there are coincidences in the outcomes but efficiencies to be gained in a different perspective leading to methods that are computationally very different from but aim toward the same basic goal. From a differential geometry perspective, we will draw strong links between shape coherence to evolution of curvature of the boundaries of shape coherent sets.

The goal in developing a definition of coherence in a nonautonomous dynamical system is to find sets that "hold together"; in some sense that phrase requires a good mathematical definition. Suppose we have a dynamical system,

$$\dot{z} = G(z,t), \tag{2.1}$$

and for a vector field $G : M \times R \to R$, for an open subset $M \subset R^2$, and sufficient regularity such that there is a flow, $\Phi(z,t;\tau) : M \times R \times R$. In an autonomous dynamical system, "hold together" leads to almost invariant sets [16], which describes weak transitivity [15]. That is, A is almost invariant if $\Phi(A,0;0) \approx \Phi(A,0;T)$ describes almost invariance, for a time epoch, $t = 0 : T$. The idea of almost invariance breaks down in the nonautonomous setting; to demand that the image of a set must mostly overlap itself in finite time is too strict since sets generally move in time. If we wish to say that two sets A_t and $A_{t+\tau}$ are *coherent pairs* [18] (see also [4]), then in measure μ we may demand that

$$\rho_\mu(A_t, A_{t+\tau}) := \frac{\mu(A_t \cap \Phi(A_{t+\tau}, t+\tau;\tau))}{\mu(A_t)} \geq \rho_0 \tag{2.2}$$

for some fraction $\rho_0 \approx 1$. Further, in [18] they demand that $\mu(A_t) = \mu(A_{t+\tau})$. But, by this (not yet completely described) definition with conditions as stated so far here, any set is a coherent pair to its own image, *no matter how contorted*. In [18] they further add a third condition to the definition of coherent pairs that $\mu(A_t)$ and $\mu(A_{t+\tau})$ must be "robust" to small perturbations, while not explicitly part of the computation. This third condition is covered

automatically, in some sense, by the computational diffusive effect [17] of the discretized grid used to approximate transfer operators in the Ulam–Galerkin based method [4]. In our notation, we will take A_t to be the evolution of an initial set A to time t under the flow, so $A_{t_0} \equiv A$, and for simplicity, we write $\Phi_t(z) := \Phi(z, t_0, t)$ with $t_0 = 0$.

Alternatively by our perspective we will appeal directly to the geometry of the idea that a coherent set should "hold together" to capture the intuition that such sets should "capture our eyes" when viewing the flow.

Definition 2.1 (finite-time shape coherence). *The shape coherence factor α between two measurable nonempty sets A and B under an area preserving flow Φ_t after a finite time epoch $t \in [0, T]$ is*

$$\alpha(A, B, T) := \sup_{S(B)} \frac{m(S(B) \cap \Phi_T(A))}{m(B)}, \tag{2.3}$$

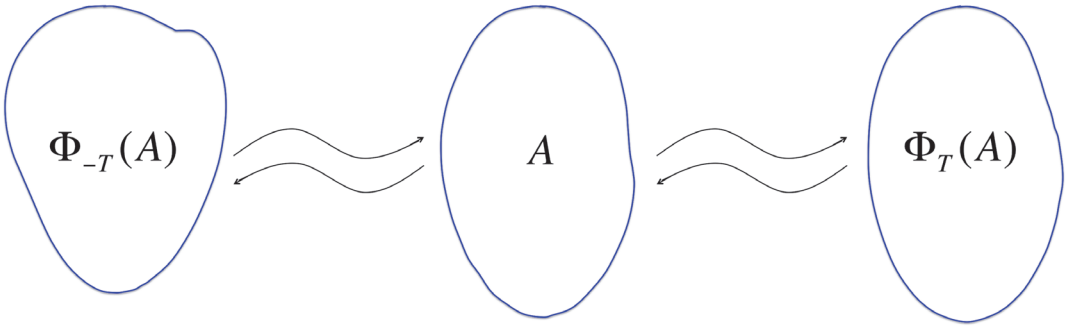
where $S(B)$ is a group of transformations of rigid body motions of B, specifically translations and rotations descriptive of frame invariance. We interpret $m(\cdot)$ to denote Lebesgue measure, but one may substitute other measures as desired. Then we say A is finite-time shape coherent to B with the factor α under the flow Φ_T after the time epoch T, but, stated briefly, A is shape coherent to B. We call B the reference set and A the dynamic set.

Whereas generally a nonlinear flow can make for quite complicated distortions when stretching and folding on general sets, on a shape coherent set this definition could be interpreted by the statement that when the flow Φ_T is restricted to a shape-coherent set A, then $\Phi_T|_A$ is roughly equivalent to a simpler transformation, a member of the group S of rigid body transformations. That is a degree of simplicity on the spatial scale of the set A, and on the time scale T, even if on finer scales within the set A there may be complicated flow.

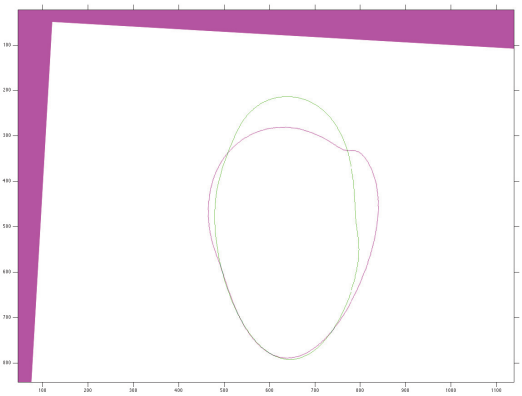
Remark 2.1.

1. While generally we can choose any kind of set B as the reference set, the idea was designed so that a set corresponding to the intuitive idea of "holding together" is more relevant. So, for example, one may choose a circle of the same area as A as a reference set.
2. It is interesting to choose the reference set B to be the dynamic set A itself. Then we are considering if a set is *finite-time shape coherent relative to itself* over time. That is, *does A maintain its shape?*
   In such a case, we may choose $B = A$ at $t = t_0 = 0$ which is then compared to $\Phi_T(A)$ where $T = t_0 + T$. We are interested here in time epochs centered on t_0, and as such, we may choose time intervals $t_1 < t_0 < t_2$ perhaps to balance stretching. Generally, in this paper we will discuss simply centered time, $t_1 = -T < t_0 = 0 < t_2 = T$.
3. Direct computation of α for specific shapes is called the "registration" problem within the computational geometry and image processing community, and there are fast algorithms, including those based on the FFT, to be used for these [5, 2, 4].
4. The definition yields that $0 \le \alpha(A, B, T) \le 1$.

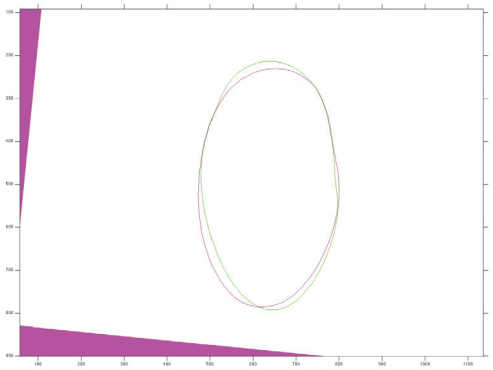
Any group of transformations could be chosen, but we have restricted our attention here to rigid body motions simply to match to the notion of curvature congruence for the boundary curve in the subsequent section. However, a shape coherence could be defined to also



(a) A set A in a flow Φ.



(b) A registration between $\Phi_{-T}(A)$ and A



(c) A registration between $\Phi_T(A)$ and A

Figure 1. (b) *and* (c) *show that* A *almost keeps its shape in the flow for both forward time and backward time. Here shapes were matched by direct application of* (2.3) *but specifically by traditional image processing registration algorithms.*

include transformations such as dilation, shear, and even a nonlinear transformation such as a conformal mapping, whether or not it has a group structure. From the perspective of image warping, registration, and image processing, the group of transformations S serves to describe the local transformation Φ_T *as restricted to a chosen set* A and has the characteristics of the simpler transformation from S. In this sense we are asking if B "warps" to A, $\Phi_T(A) \approx S(B)$ as a transformation. Note that the form of the definition (2.3) allows a set which may have "mixing" in its interior A but nonetheless "holds together" relative to B as a whole. This comes from the specific form of the shape coherence measure, α.

We may wish to contrast a stronger condition that a set remains coherent constantly throughout the time epoch, rather than possibly just at the terminal times as defined by the coefficient α in Definition 2.1. In such a case we may develop a coefficient β.

Definition 2.2 (finite-time shape coherence throughout the time epoch). *The shape coherence factor throughout the time epoch between between two measurable nonempty sets* A *and* B *under a flow* Φ_t *throughout a complete finite time epoch* $[t_1, t_2]$ *is defined as*

$$\beta(A, B, [t_1, t_2]) := \max_{t \in [t_1, t_2]} \alpha(A, B, t). \tag{2.4}$$

While α and β are well defined, and therefore shape coherence is well defined when $\alpha \approx 1$, or throughout the time epoch if $\beta \approx 1$, direct application may resort to image registration methods, and this may seem awkward. See Figure 1, where shapes $\Phi_{-T}(A)$ and A as well as A and $\Phi_T(A)$ were matched by "traditional" image registration algorithms in MATLAB, namely, the *imregister* algorithm. Instead we will now describe how these can be easily related to quantities from differential geometry that are easily computed and analyzed, namely, the evolution of curvature of the boundaries.

3. Curvature and frame invariance of shape coherence. Here we relate our definition of shape coherence to the readily computed quantities of the evolution of curvature of the boundary curves from differential geometry. In brief, by Definition 2.1 above, a set is shape coherent if it "mostly" maintains its shape in the sense of a rigid body motion. Here we relate matching the shape of a set A to its image $\Phi_T(A)$ after the flow to the problem of matching the boundary curves. From differential geometry, shapes are matched if their boundary curves are matched and curves are matched if the curvatures of the curves are matched. Furthermore, we show the regularity statement that a slowly changing boundary curve implies shape coherence. Then in a subsequent section we will explore the dynamical properties that correspond to slowly evolving boundary curves. We begin by recalling the following definition.

Definition 3.1 (see [7]). Congruence. *Two space curves are congruent if they differ only by rigid body motions, meaning translation and rotations.*

Congruence is generally accepted as the equivalence relationship between curves in differential geometry, as it allows the following theorem, which says that matching curvature and torsion lead to congruence.

Theorem 3.2 (see [7]). Fundamental Theorem of Curve Theory. *Two space curves C and $\tilde{C}$ are congruent if and only if their corresponding arcs, $C : \gamma(s) = (x(s), y(s), z(s)), 0 \le s \le 1$, and $\tilde{C} : \tilde{\gamma}(s) = (\tilde{x}(s), \tilde{y}(s), \tilde{z}(s)), 0 \le s \le 1$, both parameterized by unit arc length s, have curvature and torsion that can be matched, in the sense that there exists a "shift" parameter a such that $\kappa(s) = \tilde{\kappa}(s')$, and $\tau(s) = \tilde{\tau}(s')$ for all s, for $s' = mod(s - a, 1)$.*

We refer to a choice of $s' = mod(s - a, 1)$ that matches curvatures and torsion as exactly "lined up" parameterizations, as depicted in Figure 2.

Remark 3.1. In this paper, since we will specialize to planar flows and curves outlining planar sets, torsion is zero, $\tau = \tilde{\tau} = 0$, and $z(s) = \tilde{z}(s) = 0$. So we write simply $\gamma(s) = (x(s), y(s))$ and likewise for $\tilde{\gamma}(s)$.

Now recalling the Frenet–Serret formula [7, 14, 36],

$$\begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}' = \begin{bmatrix} 0 & \kappa & 0 \\ -\kappa & 0 & \tau \\ 0 & -\tau & 0 \end{bmatrix} \begin{bmatrix} \mathbf{T} \\ \mathbf{N} \\ \mathbf{B} \end{bmatrix}, \tag{3.1}$$

but in our planar setting we get the specialized case

$$\mathbf{T}' = \kappa \mathbf{N}, \mathbf{N}' = -\kappa \mathbf{T}. \tag{3.2}$$

The general solution of (3.2) can be readily found,

$$\begin{aligned}\mathbf{T} &= C_1 \sin \int_s \kappa(\sigma)d\sigma - C_2 \cos \int_s \kappa(\sigma)d\sigma,\\ \mathbf{N} &= C_1 \cos \int_s \kappa(\sigma)d\sigma + C_2 \sin \int_s \kappa(\sigma)d\sigma,\end{aligned} \tag{3.3}$$

where C_1 and C_2 are constant vectors.

This specifies the tangent and normal vectors of a curve, $T(s)$ and $N(s)$, in terms of a given curvature function, $\kappa(s)$, as an ODE in terms of the derivative of the (unit) parameterization, $' := \frac{d}{ds}$. Furthermore, since $T(s) = \gamma'(s)$, solving (3.2) for $T(s)$ (and simultaneously $N(s)$) for a given curvature $\kappa(s)$ allows the curve to be found:

$$\gamma(s) = \int_0^s \mathbf{T}(\sigma)d\sigma + \gamma_0. \tag{3.4}$$

So we see that the curvature specifies the curve, up to an initial position and orientation, γ_0. This clarifies part of the interest in the Frenet frame for the definition of congruence in matching curves. We wish to point out that the Frenet frame is closely related to the dynamical notion of frame invariance.

Now in the following two theorems, we show that close curvature functions correspond to close curves. Then subsequently we show that it then follows that the shape coherence must be significant. The proofs of these two theorems are given in Appendix A.

Theorem 3.3. *Given two curvature functions $\kappa_1(s)$ and $\kappa_2(s)$ of two closed curves $\gamma_1(s)$ and $\gamma_2(s)$ with the same arc length, if $\sup_s |\kappa_1(s) - \kappa_2(s)| < \varepsilon$, then there exists*

$$\delta(\varepsilon) = s^2\varepsilon(\|C_1\|_2 + \|C_2\|_2) > 0, \tag{3.5}$$

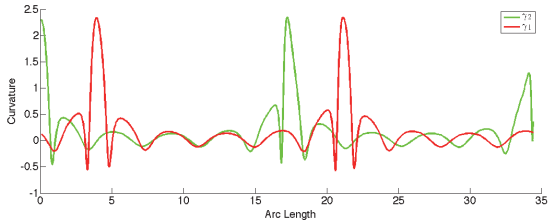
where C_1 and C_2 are the constant vectors in (3.3) and s is the arc length from initial point of the two lined-up curves, such that

$$\|\gamma_1(s) - \gamma_2(s)\|_2 < \delta(\epsilon). \tag{3.6}$$

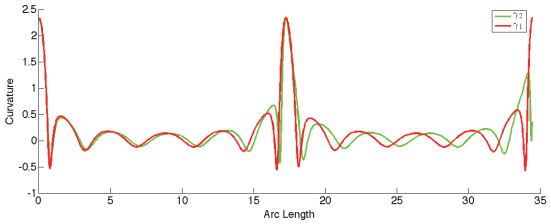
The idea of congruence in Definition 3.1 is that comparison of two curves should be immune to details such as that the parameterization of one curve may be shifted relative to the other despite if they are otherwise (mostly) the same. Even if comparing two very similar curves, but which are not oriented in a convenient way, it is possible that $\kappa_1(s_1)$ and $\kappa_2(s_2)$ may not be closely matched pointwise; nonetheless, they may define curves $\gamma_1(s)$ and $\gamma_2(s)$ that are close or even the same in the sense of congruence. In such a case, the only problem in seeing that curves are so similar by comparing the curvatures is that the parameterizations of each of the two curves are not ideally "aligned." Such a situation is depicted in Figure 2. In other words, there may exist some shift a, such that if we can define

$$s = s_1 \text{ and } s_2 = s + a \text{ such that } \sup_s |\kappa_1(s) - \kappa_2(s)| < \varepsilon, \tag{3.7}$$

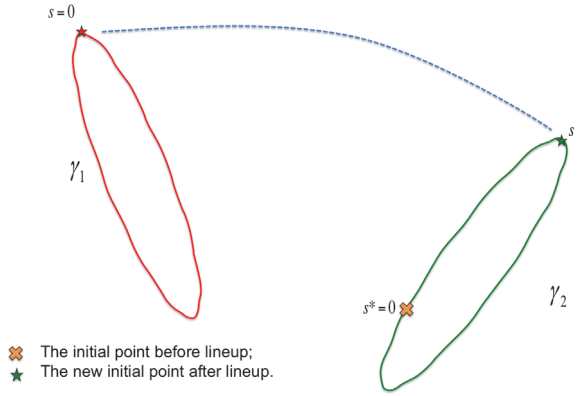
and if it turns out that $\varepsilon \geq 0$ is sufficiently small, then by Theorem 3.3, it can be shown that $\gamma_1(s)$ and $\gamma_2(s)$ are close. In practice, before resorting to Theorem 3.3, and assuming



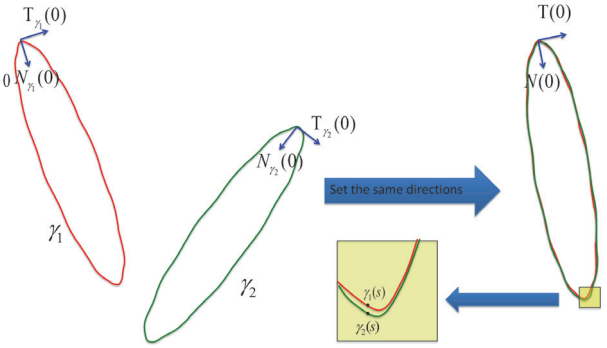
(a) Two curves: Similar curvatures & same arc length.



(b) Two curves: Similar curvatures & same arc length.



(c) Two curves that have been rectified (lined up).



(d) Two curves that have been rectified (lined up).

Figure 2. (c) *Two curves with the same arc length are shown, but the curvatures differ pointwise; note in particular the initial points* $s = 0$ *(red star on* γ_1*) and* $s^* = 0$ *(orange cross on* γ_2*). In* (a) *and* (b)*, we apply a circular convolution on the two curves to find a new initial point* $s = 0$ *(green star) on* γ_2 *so that it has pointwise almost the same curvature function as* γ_1*, as described in* (3.7)–(3.8). (d) *With the new initial point on* γ_2*, the curves are approximately lined up by overlapping and setting the same directions for both the unit normal and unit tangent vectors of the two initial points.*

continuous functions, it is useful to choose

$$a = \operatorname{argmin}_s \max |\kappa_1(s) - \kappa_2(s)|, \text{ where } s = s_1 \text{ and } s_2 = s + a. \tag{3.8}$$

The curves are approximately lined up. One useful way we find to estimate such a shift is by convolution.

In the next theorem we describe how closely matching curves leads to closely matched shapes as measured in terms of shape coherence. This, together with the above statement that closely matching curvature functions lead to closely matched curves, informs us in the following sections that studying the propensity of slowly changing curvature is the key to understanding shape coherence. To most usefully apply the following theorem, we should again assume generally that the parameterizations are designed to associate points along the curves in an efficient way as described in the previous paragraph; otherwise there may exist a parameterization that causes a small ϵ, but we may not realize it without considering (3.8).

Theorem 3.4. *For two closed curves* $\gamma_1(s) = (x_1(s), y_1(s))$ *and* $\gamma_2(s) = (x_2(s), y_2(s))(0 \leq s < 2\pi$*, which are boundaries of sets* A_1 *and* A_2 *(see Figure 3), let the boundaries of* $A =$

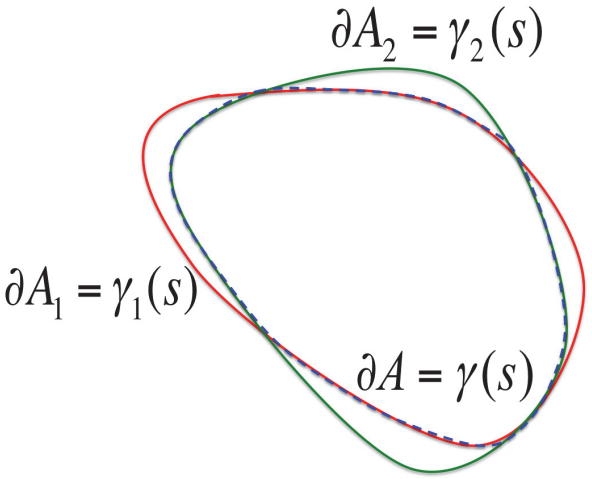


Figure 3. *The red, green, and blue closed curves are the boundaries of sets A_1, A_2, and A.*

$A_1 \cap A_2$ $A = A_1 \cap A_2 \neq \emptyset$ *such that* $area(A) > 0$ *is* $\gamma(s) = (x(s), y(s))$ *and*

$$\begin{aligned} \varepsilon &= \max\{|x - x_2|, |y - y_2|, |x' - x_2'|, |y' - y_2'|\}, \\ M &= 2\max\{|x|, |x_2|, |y|, |y_2|, |x'|, |x_2'|, |y'|, |y_2'|\}. \end{aligned} \tag{3.9}$$

Then there exists a $\Delta(\epsilon)$*, which is defined as*

$$\Delta(\epsilon) = \frac{2\pi M \varepsilon}{Area(A_2)}, \tag{3.10}$$

such that

$$1 \geq \alpha(A_1, A_2, 0) \geq 1 - \Delta(\epsilon). \tag{3.11}$$

4. Finite-time stable and unstable foliations. Stated simply, the stable foliation at a point describes the dominant direction of local contraction in forward time, and the unstable foliation describes the dominant direction of contraction in "backward" time. See Figures 4–5. Generally, the Jacobian matrix $D\Phi_t(z)$ of the flow $\Phi_t(\cdot)$ evaluated at the point z has the same action as any matrix in that a circle maps onto an ellipse. In Figures 4–5 we illustrate the general infinitesimal geometry of a small disc of variations ϵw from a base point $\Phi_t(z)$. At z, we observe that a circle of such vectors, $w = \langle \cos(\theta), \sin(\theta) \rangle, 0 \leq \theta \leq 2\pi$, centered at the point $\Phi_t(z)$ pulls back under $D\Phi_{-t}(\Phi_t(z))$ to an ellipsoid centered on z. The major axis of that infinitesimal ellipsoid defines $f_s^t(z)$, the stable foliation at z. Likewise, from $\Phi_{-t}(z)$, a small disc of variations pushes forward under $D\Phi_t(\Phi_{-t}(z))$ to an ellipsoid, again centered on z. The major axis of this ellipsoid defines the unstable foliations, $f_u^t(z)$.

To compute the major axis of ellipsoids corresponding to how discs evolve under the action of matrices, we may resort to the singular value decomposition [19]. Let

$$D\Phi_t(z) = U \Sigma V^*, \tag{4.1}$$

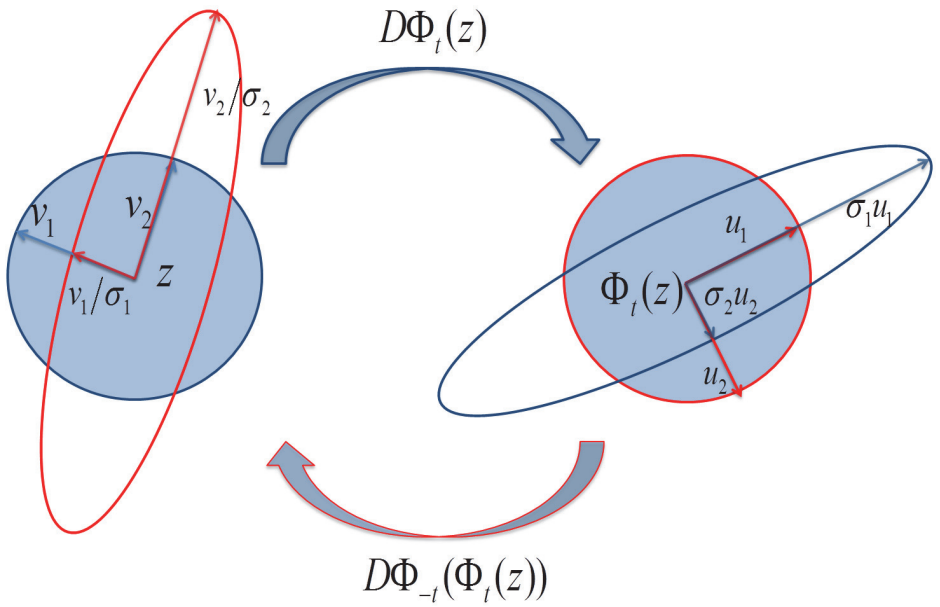


Figure 4. *The SVD (4.1) of the flow $\Phi_t(z)$ can be used to infer the finite time stable foliation $f_s^t(z)$ (and likewise the finite time unstable foliation $f_u^t(z)$) at z in terms of the major and minor axes as shown and described in (4.5)–(4.6).*

where * denotes the transpose of a matrix, U and V are orthogonal matrices, and $\Sigma = diag(\sigma_1, \sigma_2)$ is a diagonal matrix. By convention we choose the index to order, $\sigma_1 \geq \sigma_2 \geq 0$. As part of the standard singular value decomposition theory, principal component analysis provides that the first unit column vector of $V = [v_1|v_2]$ corresponding to the largest singular value, σ_1, is the major axis of the image of a circle under the matrix $D\Phi_t(z)$ around z. That is,

$$D\Phi_t(z)v_1 = \sigma_1 u_1, \tag{4.2}$$

as seen in Figure 4, describes the vector v_1 at z that maps onto the major axis, σu_1 at $\Phi_t(z)$. Since $\Phi_{-t} \circ \Phi_t(z) = z$ and $D\Phi_{-t}(\Phi_t(z))D\Phi_t(z) = I$, then, recalling the orthogonality of U and V, it can be shown that

$$D\Phi_{-t}(\Phi_t(z)) = V\Sigma^{-1}U^*, \tag{4.3}$$

and $\Sigma^{-1} = diag(\frac{1}{\sigma_1}, \frac{1}{\sigma_2})$. Therefore, $\frac{1}{\sigma_2} \geq \frac{1}{\sigma_1}$, and the dominant axis of the image of an infinitesimal circle from $\Phi_t(z)$ comes from $D\Phi_t(z)u_2 = \frac{1}{\sigma_2}v_2$.

Summarizing, the stable foliation at z is

$$f_s^t(z) = v_2, \tag{4.4}$$

where v_2 is the second right singular vector of $D\Phi_t(z)$, according to (4.1). Likewise, by the description above, the unstable foliation is

$$f_u^t(z) = \overline{u}_1, \tag{4.5}$$

where $\overline{u}_1$ is the first left singular vector of the matrix decomposition,

$$D\Phi_t(\Phi_{-t}(z)) = \overline{U}\,\overline{\Sigma}\,\overline{V}^*. \tag{4.6}$$

An important concept here is the included angles between the stable and unstable foliations as follows.

Definition 4.1. *The included angle of the finite-time stable and unstable foliations is defined as* $\theta(z,t) : \Omega \times \mathbb{R}^+ \to [-\pi/2, \pi/2]$:

$$\theta(z,t) := \arccos \frac{\langle f_s^t(z), f_u^t(z)\rangle}{\|f_s^t(z)\|_2 \|f_u^t(z)\|_2}. \tag{4.7}$$

See the included angle indicated in Figure 5; it plays a role in evolution of curvature and shape coherence, as discussed in the following section.

5. Curvature evolution near local hyperbolicity and nonhyperbolicty. In section 3 we presented the theory that shapes whose boundary curves' curvature is slowly varying in time correspond to shape coherent sets. In this section we will argue on geometric grounds that curves whose points correspond to tangency between stable and unstable foliations, as presented in section 4, tend to have slowly evolving curvature. Thus stable and unstable foliations are related to shape coherence. For this reason, in the next section we will present theory and later a constructive method to find curves of such tangencies as a means to construct shape coherent sets.

Before tackling the general problem of nonlinear flows as evolved in finite time, we discuss the linear flow, representing the evolution of curvature in the neighborhood of a base point $z = (x_1, x_2)$ with each of several hyperbolic and nonhyperbolic scenarios.

Here we cover a hyperbolic saddle which we contrast to three (precursor of) nonhyperbolic transformation types—scaling, rotation, and shear [3, 28]. We recover the curvature formula under each of these linear transformation maps, representing the local behaviors of the flow near a point of zero-splitting foliation.

1. Hyperbolic saddle: Let

$$\dot{x}_1 = \lambda_1 x_1, x_1(0) = x_{1,0}, x_2(0) = x_{2,0}, \dot{x}_2 = \lambda_2 x_2, \lambda_2 > 0 > \lambda_1, \tag{5.1}$$

which by design is a hyperbolic saddle and decoupled, so the stable and unstable manifolds are orthogonal. Now consider the evolution of a curve of initial conditions,

$$x_{2,0} = f(x_{1,0}). \tag{5.2}$$

As an example of how this curve evolves in such a manner as to increase curvature at the origin, see Figure 5(a). To verify this statement mathematically, let the evolution of points on the curve starting at initial conditions $(x_{1,0}, x_{2,0}) = (x_{1,0}, f(x_{1,0}))$ evolve according to the linear flow,

$$(x_1(s,t), x_2(s,t)) = (se^{\lambda_1 t}, f(s)e^{\lambda_2 t}), \tag{5.3}$$

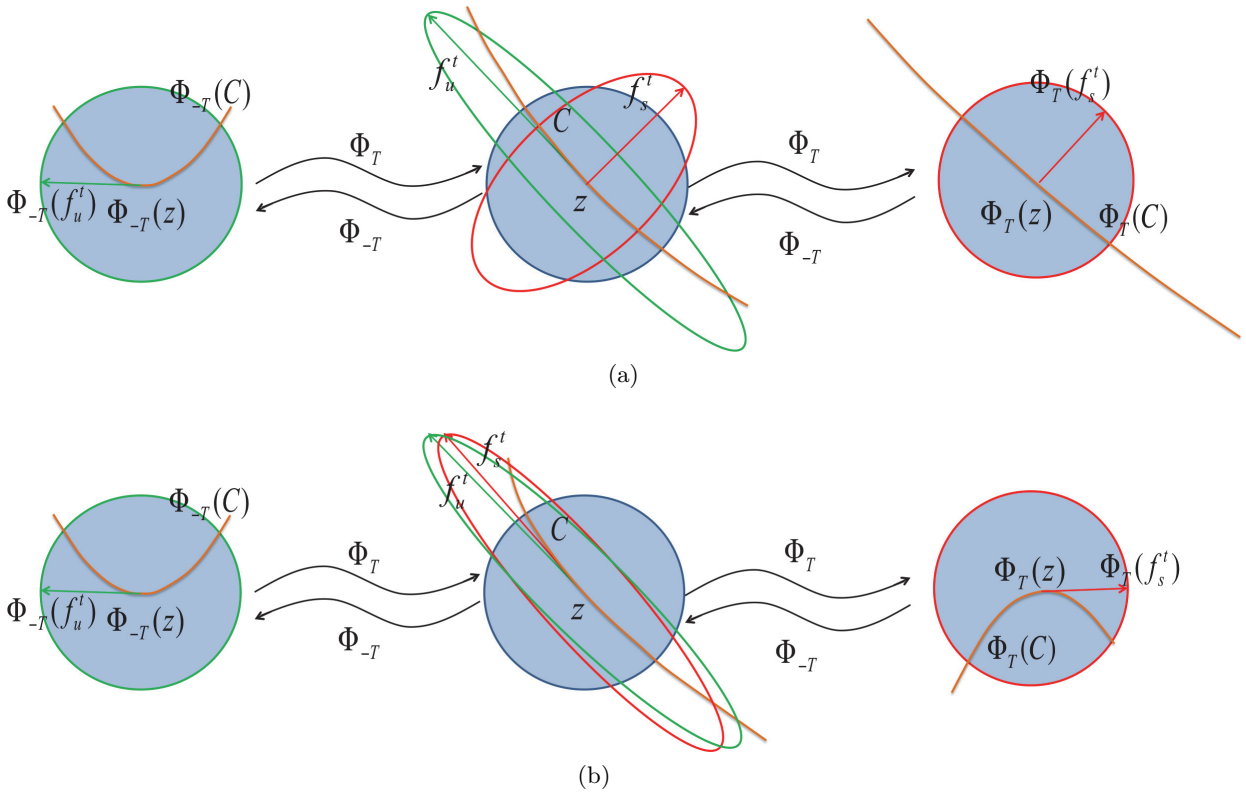


Figure 5. (a) *A curve C goes through a small neighborhood of a point z with 90 degree foliations; the angle changes its shape from time $-T$ to T. Notice that the curve changes its curvature significantly in time, and it can increase or decrease curvature depending on the details of how the curve is oriented relative to $f_s^t(z)$ and $f_u^t(z)$;* (b) *The same curve C but with almost zero-splitting foliations roughly keeps its shape as noted by inspecting the curvature at z through time.*

taking $s = x_{1,0}$ to be the chosen parameterization of the curve in terms of the initial x_1-position of a point on the curve. See Figure 5. Then using the standard curvature computation of a two-dimensional parameterized curve yields

$$k(s,t) = \frac{|x_1' x_2'' - x_2' x_1''|}{(x_1'^2 + x_2'^2)^{3/2}} = \frac{e^{(\lambda_2 - 2\lambda_1)t}|f''(s)|}{[1 + f'(s)^2 e^{2(\lambda_2 - \lambda_1)t}]^{3/2}}. \tag{5.4}$$

Thus we may estimate asymptotically in long time,

$$k(s,t) \approx e^{(\lambda_1 - 2\lambda_2)t}|f''(s)f'(s)^{-3}| \quad \text{when } t >> 1, \tag{5.5}$$

where long time is interpreted as $t > 0$ when $e^{2(\lambda_1 - \lambda_2)t} < f'(s)^2$. Conversely, asymptotically in short time,

$$k(s,t) \approx e^{(\lambda_2 - 2\lambda_1)t}|f''(s)| \quad \text{when } t << 1, \tag{5.6}$$

where short time is interpreted as $e^{2(\lambda_2 - \lambda_1)t} < f'(s)^{-2}$. We can interpret that since we have assumed that $\lambda_2 > 0 > \lambda_1$, then in short time curvature grows, but after

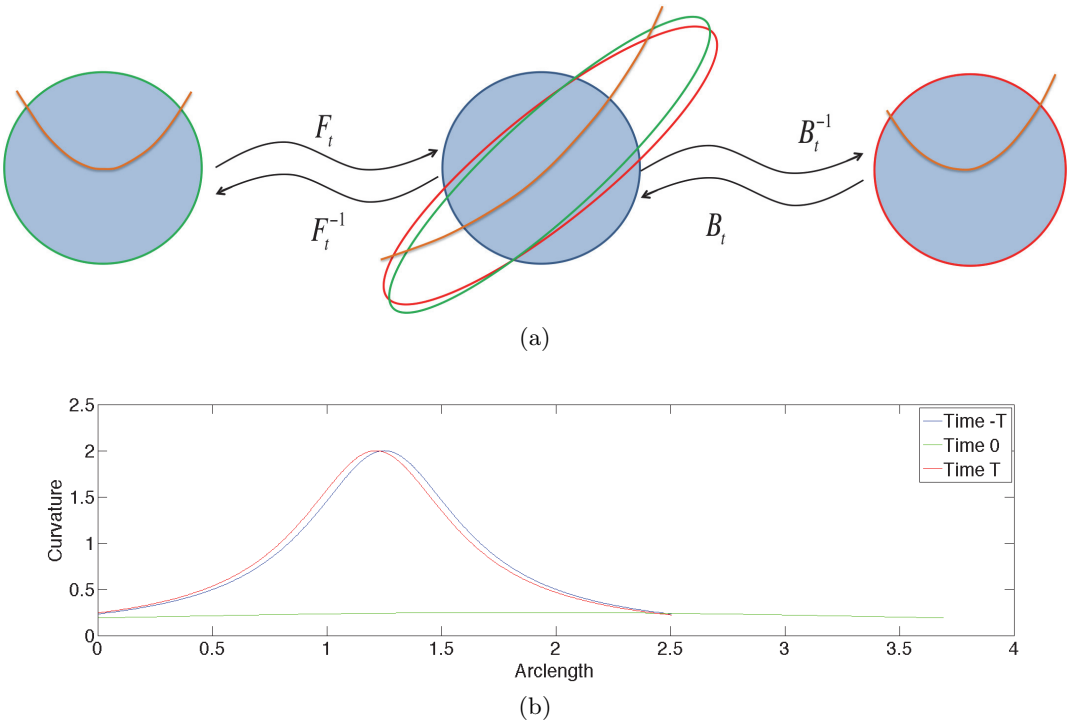


Figure 6. *The rotation and scaling case. We see that curvature changes only slightly over the time epoch.*

reaching a maximum, in long time curvature shrinks. Notably, we add to this story that points of large initial curvature due to large $f''(s)$ and small $f'(s)$ persist longer in the growing stage of the curvature, as noted by $e^{2(\lambda_2-\lambda_1)t} < f'(s)^{-2}$ before transition to $e^{2(\lambda_1-\lambda_2)t} < f'(s)^2$. Thus a highly hyperbolic saddle structure suggests significant change of curvature, which corresponds to significant changes in shape according to the Frenet–Serret theory.

2. Scaling and rotation: For a given time epoch $0 < t < +\infty$ and a point $z = (x, y)$, we suppose the forward flow F_t is

$$F_t(z) = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix} z$$

and the backward transport matrix B_t is

$$B_t(z) = \begin{pmatrix} \cos\beta & -\sin\beta \\ \sin\beta & \cos\beta \end{pmatrix} \begin{pmatrix} b & 0 \\ 0 & \frac{1}{b} \end{pmatrix} z,$$

where $\beta, \alpha > 0$ and $a, b > 1$, without loss of generality. See Figure 6(a). Thus, the flow from $-t$ to t can be considered as $B_t^{-1}F_t(z)$. And the deformation matrix of $B_t^{-1}F_t(z)$ clearly is the same as its coefficient matrix, which is

$$Q = B_t^{-1}F_t = \begin{pmatrix} \frac{1}{b} & 0 \\ 0 & b \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & \frac{1}{a} \end{pmatrix} = \begin{pmatrix} \frac{a}{b}\cos\theta & -\frac{1}{ab}\sin\theta \\ ab\sin\theta & \frac{b}{a}\cos\theta \end{pmatrix},$$

where $\theta = mod(\alpha - \beta, \pi/2)$. And it is easy to show that θ is the included angle of the stable and unstable foliations. We next show how the angle θ is related to the change of curvature of a given curve under the flow from $-t$ to t.

Consider a curve $\gamma_{-t} = (x_{-t}, f(x_{-t}))$ at time $-t$, under the flow through a time interval $[-t, t]$. The curvature of γ_{-t} at $-t$ is known as

$$k(\gamma_{-t}) = \frac{|f''(x_{-t})|}{(1 + f'^2(x_{-t}))^{\frac{3}{2}}}. \tag{5.7}$$

However, the curvature is changed by the flow $B_t^{-1}F_t(z)$ to

$$k(\gamma_t) = \frac{|f''(x_{-t})|}{((\frac{a}{b}\cos\theta - \frac{1}{ab}f'^2(x_{-t})\sin\theta)^2 + (ab\sin\theta + \frac{b}{a}f'^2(x_{-t})\cos\theta)^2)^{\frac{3}{2}}}. \tag{5.8}$$

The ratio between the two curvatures can be written as

$$\frac{k(\gamma_{-t})}{k(\gamma_t)} = \frac{(\frac{a}{b}\cos\theta - \frac{1}{ab}f'^2(x_{-t})\sin\theta)^2 + (ab\sin\theta + \frac{b}{a}f'^2(x_{-t})\cos\theta)^2}{1 + f'^2(x_{-t})}. \tag{5.9}$$

The Taylor expansion of the ratio $\frac{k(\gamma_{-t})}{k(\gamma_t)}$ with respect to small angle θ is

$$\frac{k(\gamma_{-t})}{k(\gamma_t)} = \frac{\frac{a^2}{b^2} + \frac{b^2 f'^2(x_{-t})}{a^2}}{1 + f'^2(x_{-t})} + \frac{2f'(x_{-t})(b^2 - \frac{1}{b^2})}{1 + f'^2(x_{-t})}\theta + O(\theta^2). \tag{5.10}$$

Thus, for $\theta << 1$, if the flow has $a \approx b$, we have $\frac{k(\gamma_{-t})}{k(\gamma_t)} \approx 1$ for all x_{-t} of the curve. See Figure 6(b). The curvatures hardly change through time interval $[-t, t]$. Note that there may be some special x_{-t} such that this ratio may be close to 1 even without a small θ.

3. Shear:
   (a) If we have

$$F_t = \begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}$$

and B_t is

$$B_t = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix},$$

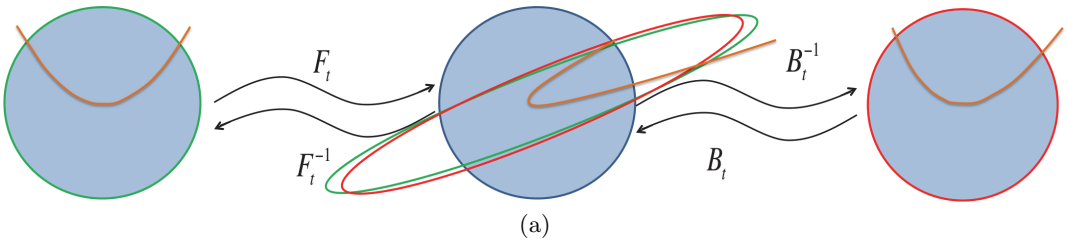
then it follows that

$$Q = B_t^{-1}F_t = \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & a \\ 0 & 1 \end{pmatrix}.$$

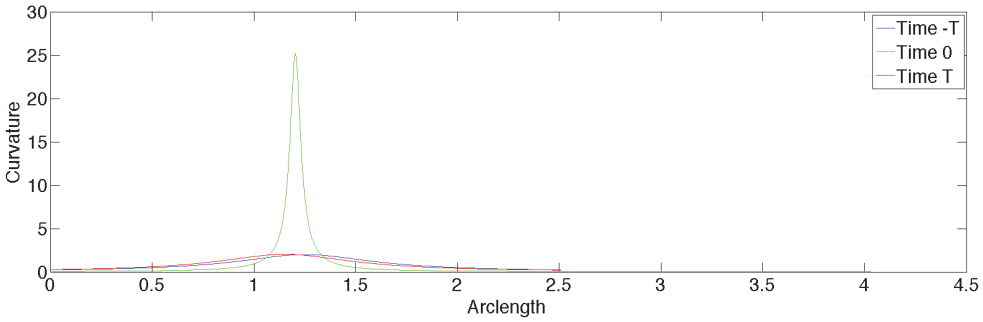
By the same process of case 1, we have the curvatures ratio as

$$\frac{k(\gamma_{-t})}{k(\gamma_t)} = \frac{(1 + (a - b)f'(x_{-t}))^2 + f'^2(x_{-t})}{1 + f'^2(x_{-t})}. \tag{5.11}$$

Thus, $a \approx b$ is necessary and sufficient for the angle between foliations to be small, and from this follows that the ratio $\frac{k(\gamma_{-t})}{k(\gamma_t)} \approx 1$. See Figure 7.



(a)



(b)

Figure 7. *The shear case. Again the curvature changes only slightly over the time epoch.*

(b) If we have

$$F_t = \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}$$

and B_t is

$$B_t = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix},$$

then

$$Q = B_t^{-1} F_t = \begin{pmatrix} 1 & -b \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ a & 1 \end{pmatrix}.$$

In this case, we can see that the angle between foliations never reaches 0 unless $a = b = 0$. However, a small angle can still can keep the curvature relatively constant. Consider the ratio between curvatures,

$$\frac{k(\gamma_{-t})}{k(\gamma_t)} = \frac{(1 - ab - bf'(x_{-t}))^2 + (a + f'(x_{-t}))^2}{1 + f'^2(x_{-t})}. \tag{5.12}$$

Hence, if we have $a \approx 0$ and $b \approx 0$, the foliations' angle is small, and the ratio $\frac{k(\gamma_{-t})}{k(\gamma_t)} \approx 1$.

6. Curves of stable and unstable foliation tangencies. Motivated by these linear hyperbolic and nonhyperbolic dynamics, we consider a nonlinear flow and how it evolves a complete curve of points z, such that each point on the curve locally has a tangency scenario as above. Therefore, we should expect that at each point on such a curve, curvature will change slowly.

We will highlight the difference between hyperbolicity and nonhyperbolicity in terms of evolution of curves by showing a nonhyperbolic curve and contrasting it with a nearby hyperbolic curve. A question which follows this discussion is if there is a curve that consists only of points with a zero-splitting angle; i.e., a zero-splitting curve is a set of points C such that

$$\text{If } z \in C, \text{ then } \theta(z,t) = 0. \tag{6.1}$$

The answer is yes. In this section we will not discuss construction as that will be covered in section 7.

Figure 8 shows the comparison between a zero-splitting curve and a nearby general curve $\tilde{C}$. We show time evolution of C from time $-T$ until time T and, correspondingly, that its curvature changes only slightly. However, even a nearby curve $\tilde{C}$ is shown, and significant changes in curvature develop in the same time epoch. Correspondingly, we see that C evidently encloses a shape coherent set, but $\tilde{C}$ does not. It may seem surprising that a small displacement of C produces such a large change of evolution of curvature, but the explanation, which relates to the geometry, is that finite-time stable and unstable foliations can change direction quite rapidly, even in small neighborhoods, as suggested by Figures 12–13. This is agreeable with the traditional well-known infinite-time concept of stable manifolds accumulating on unstable manifolds, known as the lambda lemma [13]. Thus motivated, in the next section we will discuss a continuation algorithm based on the implicit function theorem to construct curves C of zero-splitting.

7. On continuation curves of zero-splitting.

We enlist the implicit function theorem for proof of existence of curves of zero-splitting and construction. We cite the planar implicit function theorem as follows, as the planar version is sufficient for the purposes of this paper.

Theorem 7.1 (see [31]). Implicit Function Theorem. *If $F : E \to \mathbb{R}$ is continuously differentiable in a domain $E \subset \mathbb{R}$, an equation $F(x,y) = c$ has some point $(x_0, y_0) \in E$ such that $F(x_0, y_0) = c$, and $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, then there exist a neighborhood $U \subset E$ of x_0 and a function $y = y(x)$ in this neighborhood such that $y(x_0) = y_0$ and $F(x, y(x)) = c$ for all $x \in U$.*

If the vector field has enough regularity to ensure that the function $\theta(z,t)$ is at least C^1, then by the implicit function theorem, we have the following.

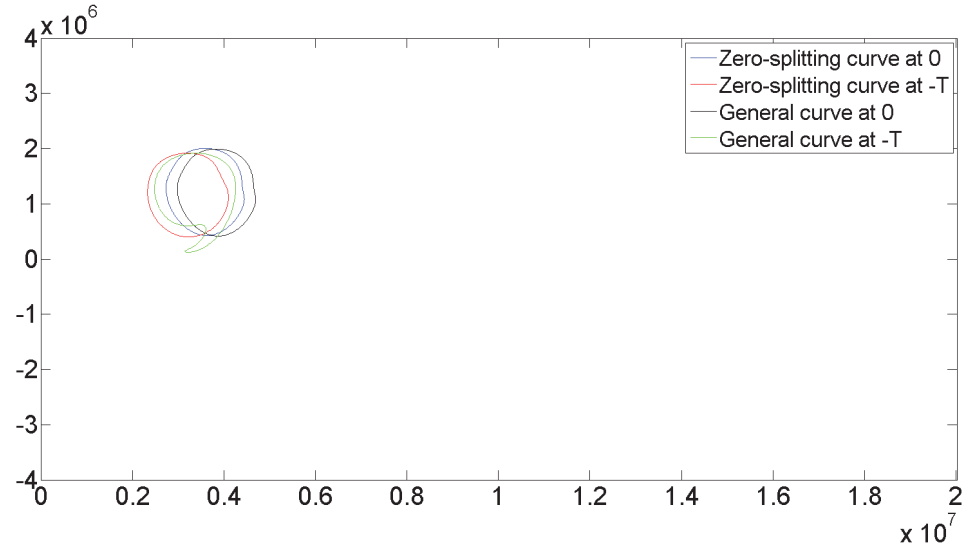
Theorem 7.2 (continuation theorem). *The set of $z = (x,y)$ with $\theta(z,t) = 0$ is a set of C^1 curves, which can be written as C^1 functions such as $y = g_1(x)$ or $x = g_2(y)$ of a finite t, which depends on $\theta_y \neq 0$ or $\theta_x \neq 0$. Furthermore, $dy/dx = -\theta_x/\theta_y$ for a given initial condition z_0 has a solution g_1 or likewise $dx/dy = -\theta_y/\theta_x$, respectively.*

By the discussion in section 6, these curves relate to shape coherence. The following algorithm offers a numerical method to obtain these boundaries.

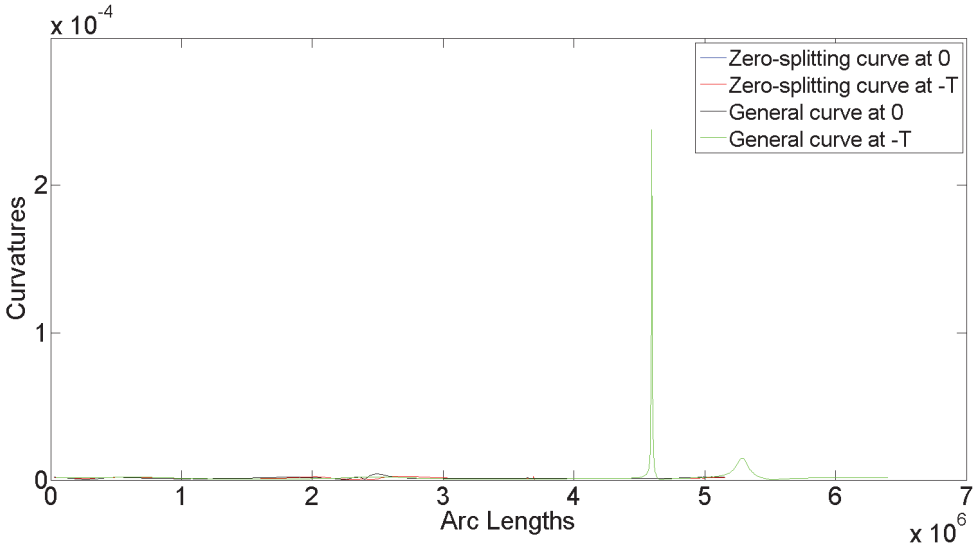
7.1. Numerical continuation.

Theorems 7.1–7.2 lead readily to a numerical continuation method to find zero-splitting curves by adaptation of the idea of continuation and by using the differential equation from the implicit function theorem. The implicit function theorem gives that, solving the initial value problem,

$$\frac{dy}{dx} = -\frac{\theta_x}{\theta_y} \quad \left(\text{or possibly } \frac{dx}{dy} = -\frac{\theta_y}{\theta_x} \text{ if tracking with respect to the other coordinate}\right), \tag{7.1}$$
$$\theta(z_0, t) = 0,$$



(a) A zero-splitting curve versus a general curve.



(b)

Figure 8. (a) *The blue curve is a zero-splitting curve C at time* 0*, and the black curve is a general curve $\tilde{C}$, which is a slightly shifted version of the blue curve. We can see that at time $-T$, compared to the general curve, the zero-splitting curve roughly keeps its curvatures, and the arc length is better.* (b) *In a curvature as a function of the arc length, $k(s)$, notice that the general curve's (black curve above) curvature changes dramatically, as seen by the green bump. Likewise, the green curve extends on the x-axis far from the other three, meaning that the arc length of the general curve grows much more than the zero-splitting curve.*

from a seed point $z_0(x_0, y_0)$ where $\theta(z_0, t) = 0$. A solution of this initial value problem represents a subset segment of the curve of zero-splitting of the stable and unstable foliations. Numerical continuation is a common theme in applied mathematics, and specifically it is used

in dynamical systems in numerical bifurcation theory [20]. There are now many sophisticated algorithms for this purpose, and we will describe only the simplest rudimentary variation of such an algorithm here.

In practice seeding the initial value problem requires finding at least one good point $(\tilde{z_0})$ which is near a zero point, which by continuity of $\theta(z,t)$, if $\|(z_0)-(\tilde{z_0})\| < \delta$, then $|\theta(\tilde{z_0},t)| < \epsilon$. Such a point $(\tilde{z_0})$ for an initial rough threshold ϵ_1 can be found by essentially a random search in the domain of interest, and then an improved seed representing a smaller threshold $0 < \epsilon_2 < \epsilon_1$ can be found by a numerical optimization (root-finding-like) algorithm applied to $\theta(z,t) = 0$ with that rough seed $\tilde{z_0}$, to the useful tiny threshold ϵ_2. For example, we choose $\epsilon_1 = 1 \times 10^{-2}$ and $\epsilon_2 = 1 \times 10^{-10}$. Then a numerical ODE solver continues along solutions of (7.1), but with the caveat that at each step of the numerical ODE solver, a corrective step must be taken by repetition of the root-finding algorithm to the threshold ϵ_2 as a search purely along the y variable when solving along $\frac{dy}{dx} = -\frac{\theta_x}{\theta_y}$, together with the caveat that the roles of x and y can reverse when reaching singularities $\theta_x = 0$ or $\theta_y = 0$ representing points where the curve can become multiply valued as a function over x (or y). So we state the following algorithm as a prediction-then-correction type algorithm, which we describe as steps rather than a complete algorithm.

1. Find a rough seed in the domain $\tilde{z}_0 \in \Omega$, for a given rough threshold, $\epsilon_1 > 0$, and let $n = 1$.
2. An improved seed is found, $z_n = (x_n, y_n)$, to the required precision $0 < \epsilon_2 < \epsilon_1$, such that $|\theta(z_n,t)| < \epsilon_2$ by the trust-region dogleg method [38].
3. Make a predictive step by Euler's method on (7.1), $\tilde{y}_{n+1} = y_n + h - \frac{\theta_x}{\theta_y}(z_n,t)$, and $x_{n+1} = x_n + h$, for a chosen small $h > 0$. (Or $\tilde{x}_{n+1} = x_n + h - \frac{\theta_y}{\theta_x}(z_n,t)$ and $y_{n+1} = y_n + h$ if the roles are reversed.)
4. Make a corrective step by the root-finder, $\tilde{z_0} \to z_0$, again to precision ϵ_2 by the trust-region dogleg method but this time while holding the dependent variable of the ODE constant. (If solving $\frac{dy}{dx} = -\frac{\theta_x}{\theta_y}$, then x is the active variable, and vice-versa if solving $\frac{dx}{dy} = -\frac{\theta_y}{\theta_x}$.)
5. If the right-hand side of the differential equation develops a singularity, $\theta_x(z_n) \approx 0$, or approximately so ($|\theta_x(z_n)| < \epsilon_3$ for some small $\epsilon_3 > 0$) (or $\theta_y(z_n) \approx 0$ depending on which of the two ODEs is currently being used in (7.1), then reverse the roles of x and y; that is, begin tracking the other version of the continuation equation in (7.1).
6. Repeat step 3 until a stopping criterion is reached.
7. Stopping criteria. Keep repeating step 4 until it cannot find zero points in the trust region or find a zero point already exists.
8. Connect the zero-splitting points from step 4; after step 5, we get a curve for one seed.
9. Connecting gap criteria. If the gap is smaller than a given distance l_{max}, we just connect it and claim that the internal region of these curves is a coherent structure candidate.

The stopping criterion generally is the result of the inability of the root-finder to find a root to the required precision. Variations of this could proceed by adaptively reducing step size h and retrying the test step as possibly the loss of the curve is the reason for the inability to find a root, and in such case the method could have stepped past a point of singularity, requiring

a role reversal. Note that such a phenomenon is sensitive to step size, and generally we choose small step sizes. Because of the role of the corrector step, we have not been motivated to choose a higher order numerical ODE integrator, but again we emphasize that there are many more sophisticated continuation algorithms available [20]. We have found that even this simple algorithm gives excellent smooth curves to high precision, and quickly, as evidenced by the examples in Figures 9–11 and 14–15. Finally, note that since there are generally many zero-splitting curves, repeated initial seeding for step 1 can proceed by randomly choosing many initial conditions $\tilde{z}_0 \in \Omega$ in an attempt to satisfy the chosen rough threshold, $\epsilon_1 > 0$.

8. Examples.

In this section, we apply our methods to the Rossby wave system and the double gyre system, both of which have become benchmark examples for studying almost invariance, coherence, and transport [18, 16, 30, 4].

8.1. An idealized stratospheric flow.

Consider the Hamiltonian system

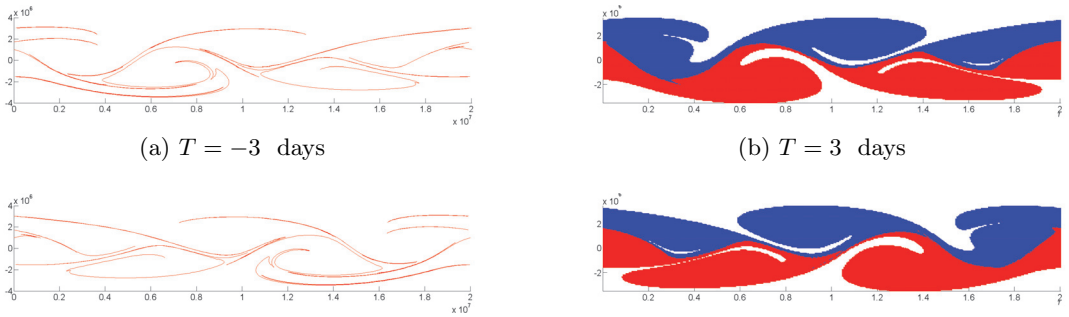
$$\begin{aligned} dx/dt &= -\partial\Phi/\partial y, \\ dy/dt &= \partial\Phi/\partial x, \end{aligned} \tag{8.1}$$

where

$$\begin{aligned} \Phi(x,y,t) = c_3 y &- U_0 L \tanh(y/L) \\ &+ A_3 U_0 L \operatorname{sech}^2(y/L)\cos(k_1 x) \\ &+ A_2 U_0 L \operatorname{sech}^2(y/L)\cos(k_2 x - \sigma_2 t) \\ &+ A_1 U_0 L \operatorname{sech}^2(y/L)\cos(k_1 x - \sigma_1 t). \end{aligned} \tag{8.2}$$

This quasiperiodic system represents an idealized zonal stratospheric flow [35, 18]. There are two known Rossby wave regimes in the system. We will show two cases with different parameters.

1. Let $U_0 = 63.66$, $c_2 = 0.205U_0$, $c_3 = 0.7U_0$, $A_3 = 0.2$, $A_2 = 0.4$, $A_1 = 0.075$, and the other parameters in (8.3) be the same as stated in [35]. We set the time epoch $T = 3$ *days*. By the parameters, we emphasize the main partition and small gaps between zero-splitting curves in the Rossby wave.
   At first, we generate a uniform 2000×200 grid of the domain $[0, 6.371\pi \times 10^6] \times [-2.5 \times 10^6, 2.5 \times 10^6]$. Figure 12(a) shows the finite-time stable and unstable foliations for each of these 4×10^5 points in the domain. Then we obtain, amongst these, 3292 points with foliation angles smaller than $\epsilon_1 = 10^{-2}$.
   Figures 9(a) and (c) are the zero-splitting curves that result from the continuation algorithm in section 7.1 but are then shown as evolved at different times, $T = -3$ *days* and $T = 3$ *days*. Notice that the small gaps in the middle region can be connected, so we can get the middle partition of the zonal flow. However, the small gaps indicate more details of the mixing behaviors of the flow. For comparison we have included in Figures 9(b) and 9(d) the results using the Frobenius–Perron operator based coherent pairs method [18].
2. Let $U_0 = 44.31$, $c_2 = 0.205U_0$, $c_3 = 0.461U_0$, $A_3 = 0.3$, $A_2 = 0.4$, $A_1 = 0.075$, and the other parameters are the same as the above. For this example, we change T



(a) $T = -3$ days (b) $T = 3$ days

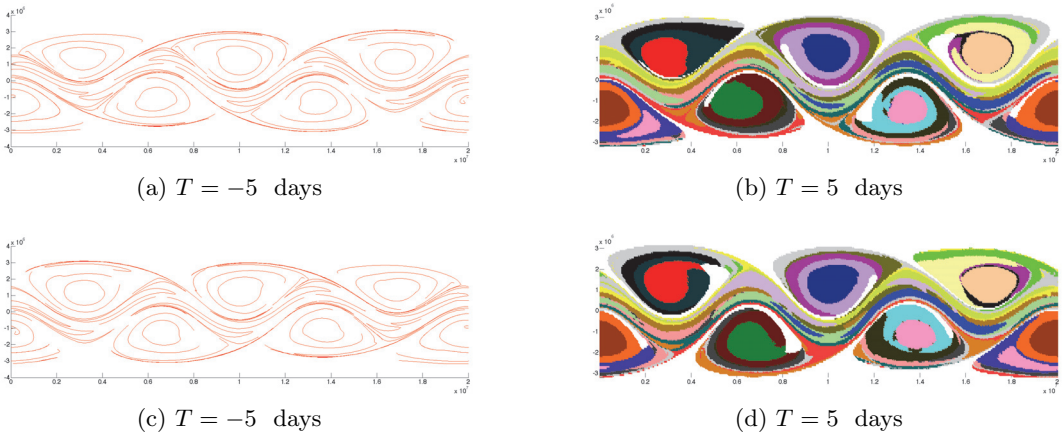
(c) $T = -3$ days (d) $T = 3$ days

Figure 9. (a) *and* (c) *are curves with zero-splitting angles at different times; the small gaps between curves indicate the leaking points of the middle partition from the coherent pair method, which is* (b) *and* (d).

from 3 *days* to 5 *days*. We describe the whole partition and elliptic islands by the parameters.

We choose the same grid to seed the initial conditions as above to apply our method. Figures 10(a) and (c) are the initial status and final status of the resulting zero-splitting curves. Notice the strong similarity here to the results shown in Figures 10(b) and (d) of the relatively coherent pairs method results from [30] that specialized [18] to a hierarchical partition. It is apparent that, in addition to the larger scaled north-south barriers, the interior elliptic island-like structures are also found by both methods. Figure 11 shows a movie of the zero-splitting curves which illustrates directly that as time proceeds the shapes hold together, as was the original motivation. This is a visual presentation of the shape coherence.

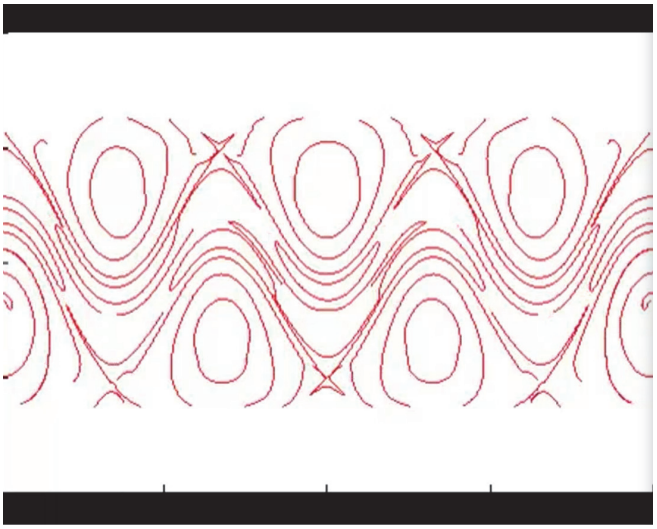
3. It is interesting to inspect the foliation geometry in more detail. Figure 12 (b) describes how the angle changes with x-coordinates for a fixed y-coordinate. We focus on a small region of the domain, Figure 12(c), and the angle function in this restricted domain, Figure 12(d). We believe that the fast switching behaviors between 0 and 90 degrees indicate an efficient mixing system. Note that as the time epoch T is increased, the angle function develops increasing variation, as the foliation switches direction more and more quickly. The eventual development in the limit of long time windows would be that the foliations would change direction infinitely many times in a finite sized small domain representing the behavior of the stable and unstable manifolds that are known to accumulate in horseshoe-like trellis structures in many common chaotic dynamical systems. On the other hand, for small time epochs, we can study how zero-splitting of the foliation first develops. Figure 13 shows how the angle changes by time T for a fixed y-coordinate. From a practical standpoint, we note how those zeros of $\theta(z,t)$ that develop first seem to generally correspond to primary partitioning of the phase space, while those zeros that develop only at larger times seem to correspond to smaller-scaled shape coherent structures.



(a) $T = -5$ days (b) $T = 5$ days

(c) $T = -5$ days (d) $T = 5$ days

Figure 10. (a) *and* (c) *are curves with zero-splitting angles at different times; the elliptic islands are consistent with the results from coherent pair method, which is* (b) *and* (d).



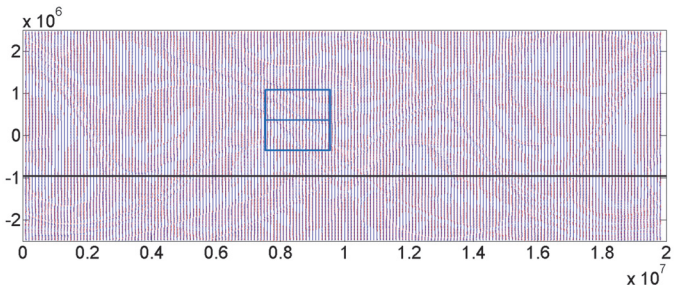
(a) A snapshot of the movie.

Figure 11. *The movie of the stable and unstable foliations with nonhyperbolic splitting for zonal flow from* $T = -5$ *days to* $T = 5$ *days.*

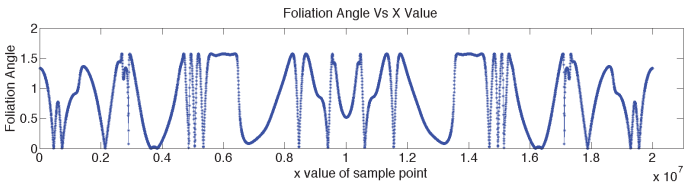
8.2. The nonautonomous double gyre. Consider the nonautonomous double gyre system,

$$\begin{aligned} \dot{x} &= -\pi A \sin(\pi f(x,t)) \cos(\pi y), \\ \dot{y} &= \pi A \cos(\pi f(x,t)) \sin(\pi y) \frac{df}{dx}, \end{aligned} \tag{8.3}$$

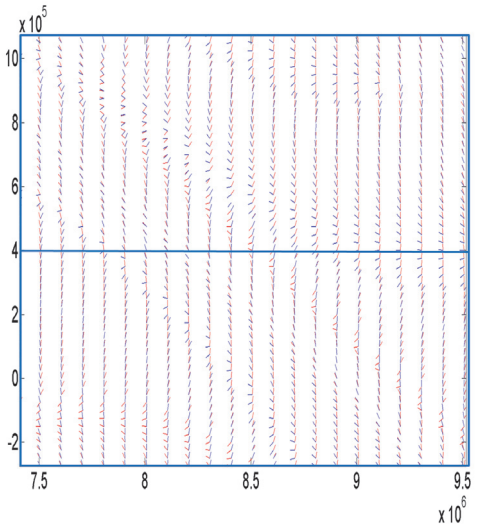
where $f(x,t) = \epsilon \sin(\omega t)x^2 + (1 - 2\epsilon \sin(\omega t))x$, $\epsilon = 0.1$, $\omega = 2\pi/10$, and $A = 0.1$. See [37, 16]. Let the initial time be $t_0 = 0$, and the time epoch to build the foliations is $T = 10$. For double gyre, we focus on the main partition and the closed curves.



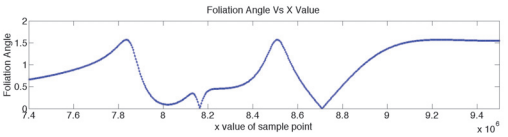
(a) Stable and unstable foliations for the whole domain.



(b) Angles versus x for a fixed y.



(c) Foliations in the blue rectangular.



(d) Angles versus x in the small blue region.

Figure 12. (a) *is the stable and unstable foliations of the Rossby wave at time* $t = 0$*; the wavelike structures result from the differences among angles.* (b) *is the plot of foliation angles versus the* x *value for a fixed* y*.* (c) *is how foliations look like in the small blue area of the whole domain.* (d) *is the angle changes with* x *for a fixed* y *for the small blue region.*

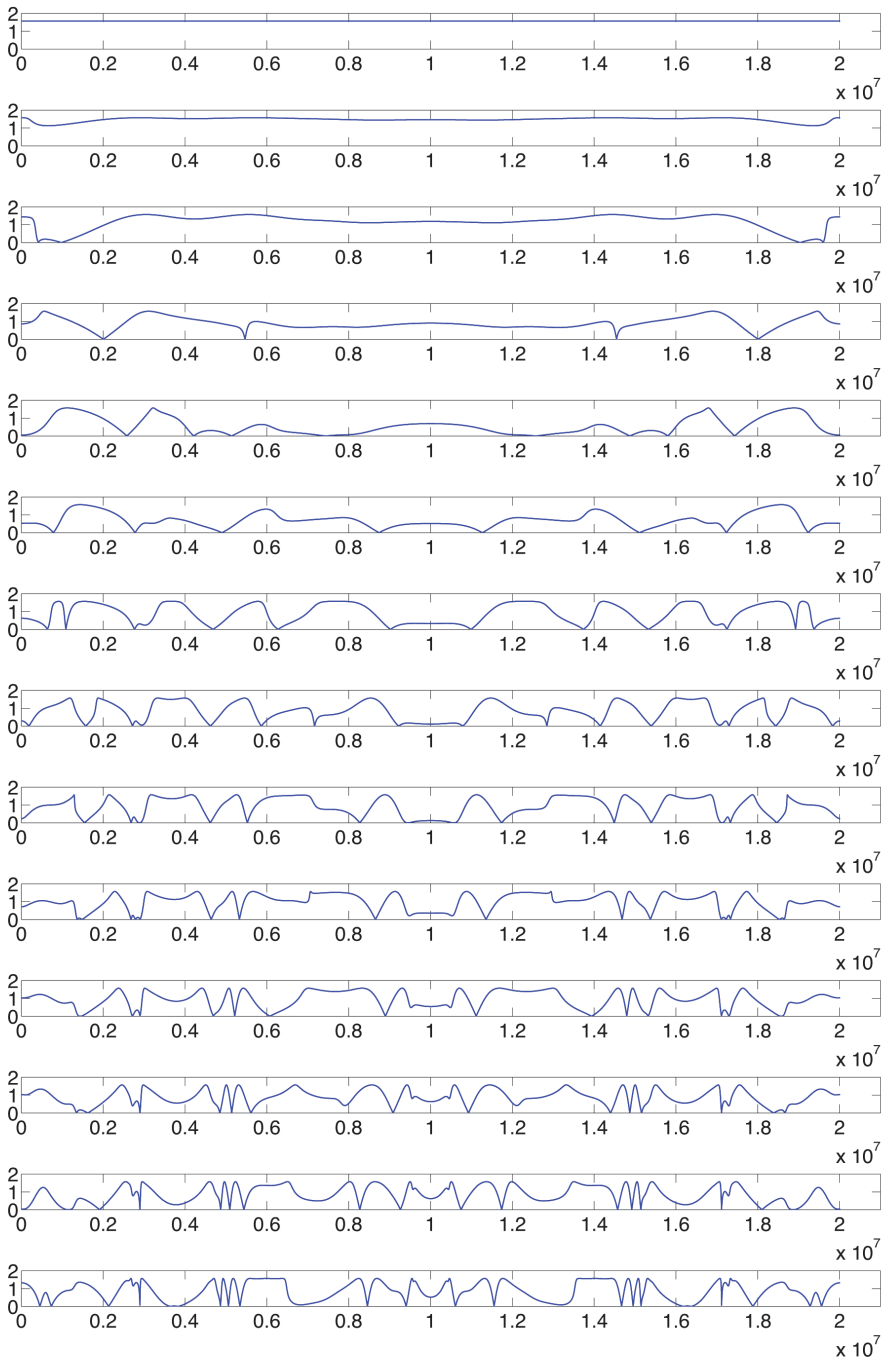
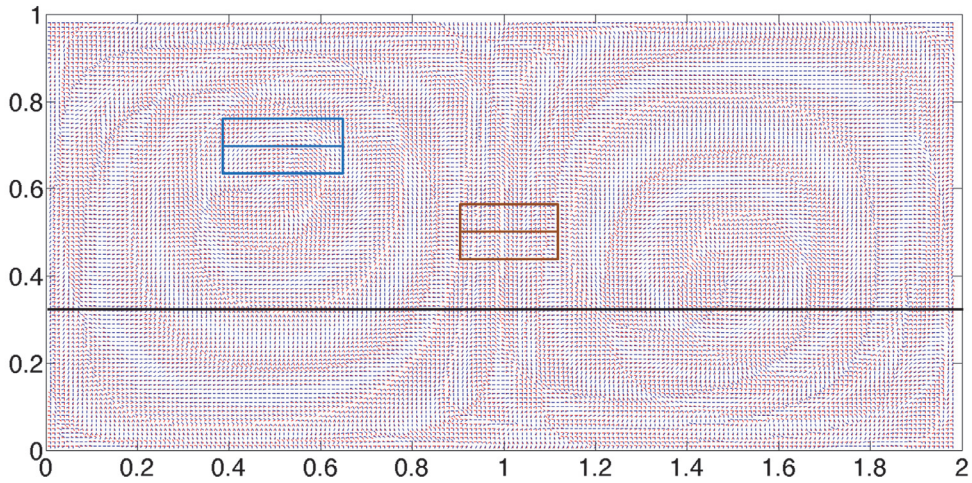


Figure 13. *For $y = -1 \times 10^6$, the figure shows the angles between foliations for different uniformly spaced times from $T = 0$ to $T = \pm 3$ days versus the x-coordinates of the Rossby wave. At the beginning, all the angles are equal to 90 degrees. Then the angles on the two sides become smaller, and more small angles keep emerging for longer times. At last, we get the figure which is in Figure 12*(b).

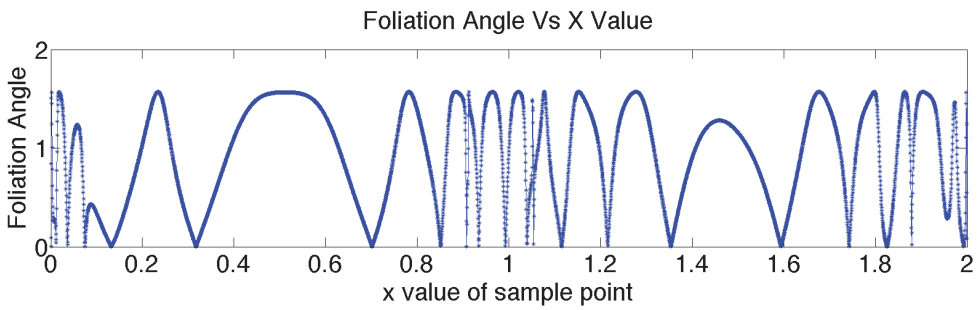
We calculate the finite-time stable and unstable foliations on a uniform 2000×200 grid of the domain $[0, 2] \times [0, 1]$. Figure 14(a) shows the zero-splitting curves at time $T = -10$, and Figure 14(b) shows the same curves evolved $2T$ to time $T = 10$. Investigating finer scaled structures, Figures 14(c) and (d) describe two small different regions in the double gyre system. In particular, those zero-splitting points z with angle functions $\theta(z, T)$ with large variation (with respect to z), which change quickly between 0 and $\pi/2$ in (d) show the well-known fast mixing middle region of the double gyre. These points are similar to those seen in the Rossby system in Figure 13. Figures 15(a) and (b) show the zero-splitting curve in the middle of the double gyre, and (c) and (d) show extra island-like closed curves.

9. Conclusions. Here we have defined a mathematically precise concept of coherence, called "shape coherent sets," to emphasize the intuitive idea of sets that hold together in a flow and roughly catch our eyes. With respect to this definition, it follows that a flow that is locally approximately the same as a simpler rigid body motion is shape coherent. Then we show that the theory of differential geometry which holds the concept of congruence between curves allows us to prove that a set whose boundary curvature changes relatively little during the finite epoch of the flow corresponds to a shape coherent set. Thus, through this equivalence, we then investigate what properties of the flow tend to cause curvature to change relatively slowly. We show that points with tangencies between the finite-time stable and unstable foliations tend to experience curvature change relatively slowly. Therefore, we search for curves of such tangency points, which are usually typical of shear. We prove existence of such curves through the implicit function theorem, which also suggests a constructive algorithm based on solving the ODE from the implicit function theorem, or many other of the typical methods for numerical continuation. Thus we develop shape coherent sets in two benchmark examples, the double gyre and the Rossby wave system. Further, we include investigations of the intricate stable and unstable foliations topology in each of these systems, which is typical of finite versions of chaotic systems, where the complex folding of stable and unstable manifolds suggests that finite-time stable and unstable manifolds will change rapidly in space. Therefore, the angle's function (see Figures 12–13) can have highly intricate structure, and we discuss the role of primary roots for primary barriers.

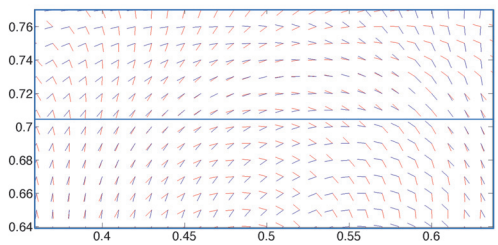
In a broad sense, it might be said that the phrase "coherence" has many different meanings to many different people—some spectral and some set oriented. Some interpretations of the phrase allow for sets that may significantly stretch and fold. Clearly, if a set is specifically highlighted visually by coloring the set as a partition, and then that set is advected, then the set will appear to hold together. However, this could be interpreted as a property of a continuous flow, that connected sets remain connected, and so a coloring alone will guide our eye for almost any coloring. Our definition of shape coherence is not meant to capture all possible perspectives of the popular phrase "coherence." It does, however, describe a phenomenon that we believe exists in a wide variety of chaotic and turbulent systems as a type of simplicity within the otherwise complicated motion. It is evident that it is not a property present in all possible dynamical systems (taking the Arnold Cat map, for example), but it does exist, and we find it to be an interesting phenomenon. Note that shape coherence as we defined it has both spatial and time scales associated with it. Time scales obviously describe sets that may hold together for a while before deforming, but also on a spatial scale, a rather simple set may



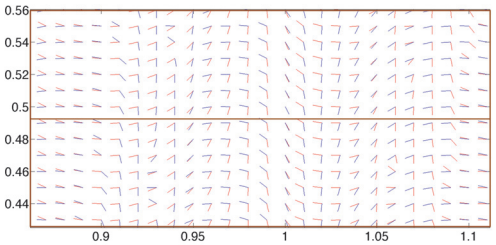
(a) Foliations of the whole domain of the double gyre.



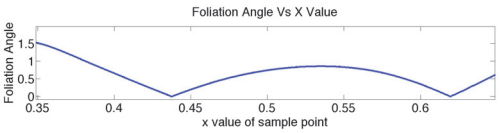
(b) Foliations' angle on the black line.



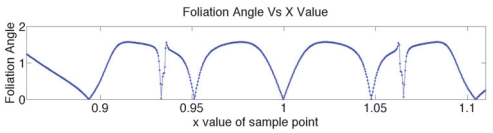
(c) Foliations in the blue rectangular.



(d) Foliations in the brown rectangular.

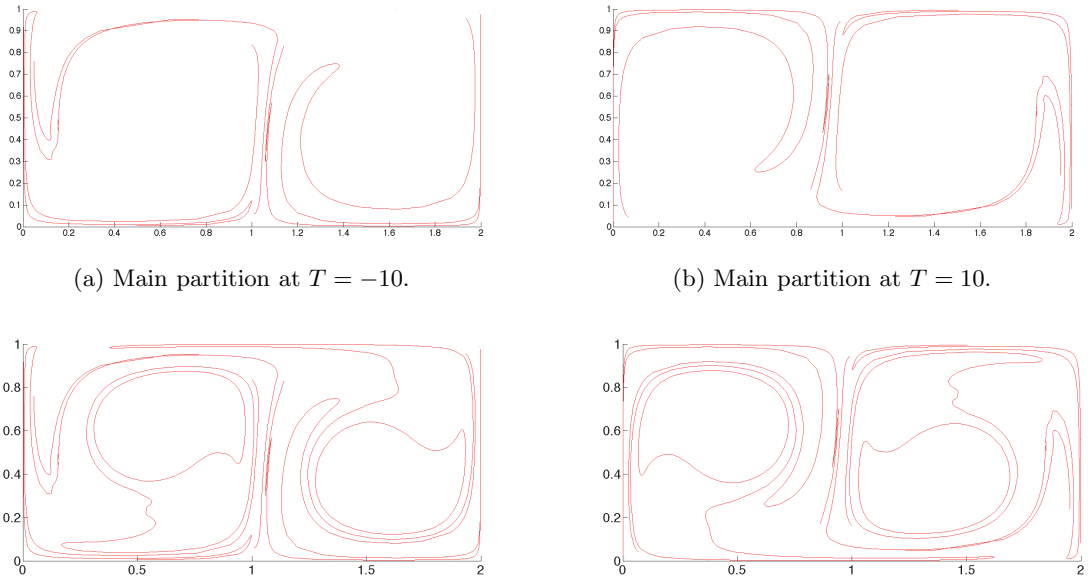


(e) Angles in the blue line.



(f) Angles in the brown line.

Figure 14.



(a) Main partition at $T = -10$.

(b) Main partition at $T = 10$.

(c) Island partition at $T = -10$.

(d) Island partition at $T = 10$.

Figure 15.

maintain its shape, whereas subsets of that set may contain highly turbulent behavior. To that point, consider as an example for the sake of discussion the Great Red Spot of Jupiter, which is known to be quite complex within the great storm which maintains its approximate shape for long times. We have shown in detail two popular examples, and we have found the shape coherent sets to exist in many more examples, including periodic and aperiodic nonautonomous flows. Finally, we remark that fundamentally the study of coherence and specifically shape coherent sets is in some sense the complement to the study of transport and the mixing which occurs in the hyperbolic-like mixing sets. These are fundamentally complementary questions on complement sets.

Appendix A. Proof of Theorems 3.3–3.4. In this section, we give proofs of Theorems 3.3 and 3.4, which we restate here for convenience.

Theorem A.1 (see (3.3)). *Given two curvature functions $\kappa_1(s)$ and $\kappa_2(s)$ of two closed curves $\gamma_1(s)$ and $\gamma_2(s)$ with the same arc length, if $\sup_s |\kappa_1(s) - \kappa_2(s)| < \varepsilon$, then there exists*

$$\delta(\varepsilon) = s^2\varepsilon(\|C_1\|_2 + \|C_2\|_2) > 0, \tag{A.1}$$

where C_1 and C_2 are the constant vectors in (3.3) and s is the arc length from the initial point of the two lined-up curves, such that

$$\|\gamma_1(s) - \gamma_2(s)\|_2 < \delta(\epsilon). \tag{A.2}$$

Proof. Consider comparing the two curves pointwise. While here we assume that the arc lengths of the two curves are the same, we will discuss the generalized case of differing arc

lengths in what follows. We have assumed that the curvatures are as suggested in Figure 2. We will argue that distance between the two points can be controlled by their curvatures.

Without loss of generality and for convenience of the estimates of comparison, assume that the two curves intersect at $s = 0$, $\gamma_1(0) = \gamma_2(0)$, and furthermore at $s = 0$ that the tangent and normal vectors are each parallel, as depicted in Figure 2. The reason this statement can be made without loss of generality is that the Frenet–Seret formula states that a curve can be reconstructed up to initial conditions, which is the reasoning behind the definition of curve congruence that curves are called the same if they are indistinguishable up to a rigid body motion. Therefore, a rigid body motion of $\gamma_2(s)$ can be used to orient the tangent and normal at $s = 0$ with those of $\gamma_1(s)$, and also at $s = 0$, as depicted in Figure 2. This defines constant vectors C_1 and C_2. Note this is not a unique orientation for matching, but sufficient and convenient.

We denote the constants as C_1 and C_2, rather than two different pairs. So we have

$$\gamma_1(s) = \int_0^s \left(C_1 \sin \int_0^s \kappa_1(\sigma)d\sigma - C_2 \cos \int_0^s \kappa_1(\sigma)d\sigma \right) d\sigma, \tag{A.3}$$

$$\gamma_2(s) = \int_0^s \left(C_1 \sin \int_0^s \kappa_2(\sigma)d\sigma - C_2 \cos \int_0^s \kappa_2(\sigma)d\sigma \right) d\sigma. \tag{A.4}$$

Then the distance between the two points $\gamma_1(s_a)$ and $\gamma_2(s_b)$ on the different curves is

$$\begin{aligned}
\|\gamma_1(s) - \gamma_2(s)\|_2 &= \left\| \int_0^s \left(C_1 \sin \int_0^s \kappa_1(\sigma)d\sigma - C_2 \cos \int_0^s \kappa_1(\sigma)d\sigma \right) d\sigma \right. \\
&\quad \left. - \int_0^s \left(C_1 \sin \int_0^s \kappa_2(\sigma)d\sigma - C_2 \cos \int_0^s \kappa_2(\sigma)d\sigma \right) d\sigma \right\|_2 \\
&= \left\| \int_0^s \left(C_1 \left(\sin \int_0^s \kappa_1(\sigma)d\sigma - \sin \int_0^s \kappa_2(\sigma)d\sigma \right) \right) d\sigma \right. \\
&\quad \left. + \int_0^s \left(C_2 \left(\cos \int_0^s \kappa_2(\sigma)d\sigma - \cos \int_0^s \kappa_1(\sigma)d\sigma \right) \right) d\sigma \right\|_2 \\
&= \left\| \int_0^s \left(2C_1 \cos \left(\int_0^s \frac{\kappa_1(\sigma) + \kappa_2(\sigma)}{2} d\sigma \right) \sin \left(\int_0^s \frac{\kappa_1(\sigma) - \kappa_2(\sigma)}{2} d\sigma \right) \right) d\sigma \right. \\
&\quad \left. + \int_0^s \left(2C_2 \sin \left(\int_0^s \frac{\kappa_2(\sigma) + \kappa_1(\sigma)}{2} d\sigma \right) \sin \left(\int_0^s \frac{\kappa_1(\sigma) - \kappa_2(\sigma)}{2} d\sigma \right) \right) d\sigma \right\|_2 \\
&= \left\| \int_0^s 2 \left(C_1 \cos \left(\int_0^s \frac{\kappa_1(\sigma) + \kappa_2(\sigma)}{2} d\sigma \right) + C_2 \sin \left(\int_0^s \frac{\kappa_2(\sigma) + \kappa_1(\sigma)}{2} d\sigma \right) \right) \right. \\
&\quad \left. \times \sin \left(\int_0^s \frac{\kappa_1(\sigma) - \kappa_2(\sigma)}{2} d\sigma \right) d\sigma \right\|_2 \\
&\leq \int_0^s 2 \left((\|C_1\|_2 + \|C_2\|_2) \left| \sin \left(\int_0^s \frac{\kappa_1(\sigma) - \kappa_2(\sigma)}{2} d\sigma \right) \right| \right) d\sigma. \qquad \text{(A.5)}
\end{aligned}$$

By the assumed condition $\sup_s |\kappa_1(s) - \kappa_2(s)| < \varepsilon$, we have at least, for $\epsilon < \frac{\pi}{2}$,

$$\begin{aligned}\|\gamma_1(s) - \gamma_2(s)\|_2 &< \int_0^s 2\left((\|C_1\|_2 + \|C_2\|_2)\left|\sin\left(\int_0^s \frac{\varepsilon}{2} d\sigma\right)\right|\right) d\sigma \\ &= \int_0^s 2\left(\|C_1\|_2 + \|C_2\|_2\right)\left|\sin\left(\frac{s\varepsilon}{2}\right)\right| d\sigma \\ &= 2s\left(\|C_1\|_2 + \|C_2\|_2\right)\left|\sin\left(\frac{s\varepsilon}{2}\right)\right| \le s^2\varepsilon(\|C_1\|_2 + \|C_2\|_2). \end{aligned} \tag{A.6}$$

The last inequality follows from the fact that $\sin(p) \le p$ for all $0 \le p$. Hence we may choose $\delta(\varepsilon) = s^2\varepsilon(\|C_1\|_2 + \|C_2\|_2)$. Note that s is the arc length from the beginning points to the compared points, and recall that by assumption $\epsilon > 0$. ∎

Theorem A.2 (see (3.4)). *For two closed curves $\gamma_1(s) = (x_1(s), y_1(s))$ and $\gamma_2(s) = (x_2(s), y_2(s))(0 \le s < 2\pi)$ which are boundaries of sets A_1 and A_2 (see Figure 3), let the boundaries of $A = A_1 \cap A_2$, $A = A_1 \cap A_2 \neq \emptyset$ such that area$(A) > 0$ be $\gamma(s) = (x(s), y(s))$, and,*

$$\begin{aligned}\varepsilon &= \max\{|x - x_2|, |y - y_2|, |x' - x_2'|, |y' - y_2'|\}, \\ M &= 2\max\{|x|, |x_2|, |y|, |y_2|, |x'|, |x_2'|, |y'|, |y_2'|\}. \end{aligned} \tag{A.7}$$

Then there exists a $\Delta(\epsilon)$, which is defined as

$$\Delta(\epsilon) = \frac{2\pi M\varepsilon}{Area(A_2)} \tag{A.8}$$

such that

$$1 \ge \alpha(A_1, A_2, 0) \ge 1 - \Delta(\epsilon). \tag{A.9}$$

Proof. Suppose that the closed curves $\gamma_1(s)$ and $\gamma_2(s)$ are boundaries of sets A_1 and A_2, respectively. Let $A = A_1 \cap A_2$ and $\gamma(s)$ be the boundary of A. See Figure 3. We annotate these closed curves as $\gamma(s) = (x(s), y(s))$, $\gamma_1(s) = (x_1(s), y_1(s))$, and $\gamma_2(s) = (x_2(s), y_2(s))$.

By Green's theorem, we have

$$\begin{aligned} Area(A) &= \frac{1}{2}\int_{\gamma(s)} xdy - ydx = \frac{1}{2}\int_0^{2\pi} x(s)y'(s)ds - y(s)x'(s)ds, \\ Area(A_2) &= \frac{1}{2}\int_{\gamma_2(s)} x_2dy_2 - y_2dx_2 = \frac{1}{2}\int_0^{2\pi} x_2(t)y_2'(s)ds - y_2(s)x_2'(s)ds \\ &= \frac{1}{2}\int_0^{2\pi} x_2(s)y_2'(s)ds - y_2(s)x_2'(s)ds, \end{aligned} \tag{A.10}$$

where the reason we use 2π is that we set the origin inside A. Then we have

$$\begin{aligned}
|Area(A) - Area(A_2)| &= \left|\frac{1}{2}\int_0^{2\pi} x(s)y'(s)ds - y(s)x'(s)ds - \frac{1}{2}\int_0^{2\pi} x_2(s)y_2'(s)ds - y_2(s)x_2'(s)ds\right| \\
&= \left|\frac{1}{2}\int_0^{2\pi} \Big(x(s)y'(s) - x_2(s)y_2'(s)\Big) + \Big(y_2(s)x_2'(s) - y(s)x'(s)\Big)ds\right| \\
&= \left|\frac{1}{2}\int_0^{2\pi} \Big(x(s)y'(s) - x(s)y_2'(s) + x(s)y_2'(s) - x_2(s)y_2'(s)\Big)\right. \\
&\quad \left. + \Big(y_2(s)x_2'(s) - y(s)x_2'(s) + y(s)x_2'(s) - y(s)x'(s)\Big)ds\right| \\
&= \left|\frac{1}{2}\int_0^{2\pi} \Big(x(s)y'(s) - x(s)y_2'(s)\Big) + \Big(x(s)y_2'(s) - x_2(s)y_2'(s)\Big)\right. \\
&\quad \left. + \Big(y_2(s)x_2'(s) - y(s)x_2'(s)\Big) + \Big(y(s)x_2'(s) - y(s)x'(s)\Big)ds\right| \\
&\le \frac{1}{2}\int_0^{2\pi} \Big|x(s)y'(s) - x(s)y_2'(s)\Big| + \Big|x(s)y_2'(s) - x_2(s)y_2'(s)\Big| \\
&\quad + \Big|y_2(s)x_2'(s) - y(s)x_2'(s)\Big| + \Big|y(s)x_2'(s) - y(s)x'(s)\Big|ds \\
&\le \frac{1}{2}\int_0^{2\pi} |x(s)||y'(s) - y_2'(s)| + |y_2'(s)||x(s) - x_2(s)| \\
&\quad + |x_2'(s)||y(s) - y_2(s)| + |y(s)||x_2'(s) - x'(s)|ds \\
&\le \frac{1}{2}\int_0^{2\pi} \frac{1}{2}M\varepsilon + \frac{1}{2}M\varepsilon + \frac{1}{2}M\varepsilon + \frac{1}{2}M\varepsilon ds \\
&= M\varepsilon \int_0^{2\pi} ds \\
&= 2\pi M\varepsilon. \qquad \text{(A.11)}
\end{aligned}$$

On the other hand, from (2.3),

$$\alpha(A_1, A_2, 0) = \sup_{S(A_2)} \frac{m(S(A_2) \cap \Phi_0(A_1))}{m(A_2)}, \qquad \text{(A.12)}$$

and then we have

$$\alpha(A_1, A_2, 0) \ge \frac{m(S(A_2) \cap A_1)}{m(A_2)} = \frac{m(A)}{m(A_2)} = 1 - \frac{|Area(A) - Area(A_2)|}{Area(A_2)}. \qquad \text{(A.13)}$$

Let $\Delta(\epsilon) = \frac{2\pi M\varepsilon}{Area(A_2)}$; we have

$$\alpha(A_1, A_2, 0) = 1 - \frac{|Area(A) - Area(A_2)|}{Area(A_2)} \ge 1 - \Delta(\epsilon). \qquad \text{(A.14)}$$

■

Note that the simplified details of the above proof assume that there is a region of overlap between A_1 and A_2 that has just one connected component, and we can always assume that the two regions are arranged to have at least one such region. In the more general case in which the two regions overlap in multiple components, the above proof can be easily adjusted to integrate the area across each region, and the statement of the theorem remains the same. Notice also that we have stated the theorem so that each boundary curve has the same arc length, but in the case of different arc lengths a comparable theorem can be developed, proved by defining a "speed" $s' = vs$, to show that regularity between boundary curves leads to regularity of the shape coherence.

REFERENCES

[1] V. I. Arnold, *Mathematical Methods of Classical Mechanics*, Springer, Berlin, 1978.

[2] D. I. Barnea and H. F. Silverman, *A class of algorithms for fast digital registration*, IEEE Trans. Comput., C21 (1972), pp. 179–186.

[3] M. Berger, *Geometry* I, Springer, Berlin, 1987.

[4] E. M. Bollt and N. Santitissadeekorn, *Applied and Computational Measurable Dynamics*, SIAM, Philadelphia, 2014.

[5] E. De Castro and C. Morandi, *Registration of translated and rotated images using finite Fourier transforms*, IEEE Trans. Pattern Anal. Machine Intell., 95 (1987), pp. 700–703.

[6] M. Dellnitz and O. Junge, *Set oriented numerical methods for dynamical systems*, in Handbook of Dynamical Systems II: Towards Applications, World Scientific, Singapore, 2002, pp. 221–264.

[7] M. P. do Carmo, *Differential Geometry of Curves and Surfaces*, Prentice–Hall, Englewood Cliffs, NJ, 1976.

[8] F. D'Ovidio, E. Schuckburgh, and B. Legras, *Local mixing events in the upper troposphere and lower stratosphere. Part* II: *Seasonal and interannual variability*, J. Atmospheric Sci., 66 (2009), pp. 3695–3706.

[9] I. T. Drummond, *Stretching and bending of line elements in random flows*, J. Fluid Mech., 252 (1993), pp. 479–498.

[10] I. T. Drummond and W. Munch, *Distortion of line and surface elements in model turbulent flows*, J. Fluid Mech., 225 (1991), pp. 529–543.

[11] M. Farazmand and G. Haller, *Computing Lagrangian coherent structures from their variational theory*, Chaos, 22 (2012), 013128.

[12] M. Farazmand and G. Haller, *Erratum and addendum to "A variational theory of hyperbolic Lagrangian coherent structures, Physica D,* 240 (2011) 574–598", Phys. D, 241 (2012), pp. 439–441.

[13] E. Fontich and P. Martin, *Differentiable invariant manifolds for partially hyperbolic tori and a lambda lemma*, Nonlinearity, 13 (2000), pp. 1561–1593.

[14] F. Frenet, *Sur les courbes à double courbure*, Abstract of a thesis of the Faculté des Sciences de Toulouse, J. Math. Pures Appl., 17 (1852), pp. 437–447.

[15] G. Froyland, *Statistically optimal almost-invariant sets*, Phys. D, 200 (2005), pp. 205–219.

[16] G. Froyland and K. Padberg, *Almost-invariant sets and invariant manifolds: Connecting probabilistic and geometric descriptions of coherent structures in flows*, Phys. D, 238 (2009), pp. 1507–1523.

[17] G. Froyland and K. Padberg, *Almost-invariant and finite-time coherent sets: Directionality, duration, and diffusion*, in Ergodic Theory, Open Dynamics, and Coherent Structures, Springer Proc. Math. Statist. 70, Springer, New York, 2014, pp. 171–216.

[18] G. Froyland, N. Santitissadeekorn, and A. Monahan, *Transport in time-dependent dynamical systems: Finite-time coherent sets*, Chaos, 20 (2010), 043116.

[19] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 2nd ed., The Johns Hopkins University Press, Baltimore, MD, 1989.

[20] W. J. F. Govaerts, *Numerical Methods for Bifurcations of Dynamical Equilibria*, SIAM, Philadelphia, 2000.

[21] G. Haller, *Finding finite-time invariant manifolds in two-dimensional velocity fields*, Chaos, 10 (2000), pp. 99–108.
[22] G. Haller, *Lagrangian coherent structures from approximate velocity data*, Phys. Fluids, 14 (2002), pp. 1851–1861.
[23] G. Haller, *A variational theory of hyperbolic Lagrangian coherent structures*, Phys. D, 240 (2011), pp. 574–598.
[24] G. Haller and F. J. Beron-Vera, *Geodesic theory of transport barriers in two-dimensional flows*, Phys. D, 241 (2012), pp. 1680–1702.
[25] T. Ishihara and Y. Kaneda, *Stretching and distortion of material line elements in two-dimensional turbulence*, J. Phys. Soc. Japan, 61 (1992), pp. 3547–3558.
[26] D. Karrasch, *Comment on "A variational theory of hyperbolic Lagrangian coherent structures, Physica D,* 240 (2011) 574–598", Phys. D, 241 (2012), pp. 1470–1473.
[27] D. H. Kelley and N. T. Ouellette, *Separating stretching from folding in fluid mixing*, Nature Phys., 7 (2011), pp. 477–480.
[28] S. Lang, *Linear Algebra*, Springer-Verlag, New York, 1987.
[29] M. Liu and F. J. Muzzio, *The curvature of material lines in chaotic cavity flows*, Phys. Fluids, 8 (1996), pp. 75–83.
[30] T. Ma and E. Bollt, *Relatively coherent sets as a hierarchical partition method*, Internat. J. Bifur. Chaos Appl. Sci. Engrg., 23 (2013), 1330026.
[31] J. R. Munkres, *Analysis on Manifolds*, Addison–Wesley, Redwood City, CA, 1991.
[32] N. T. Ouellette and J. P. Gollub, *Curvature fields, topology, and the dynamics of spatiotemporal chaos*, Phys. Rev. Lett., 99 (2007), 194502.
[33] R. Peikert, A. Pobitzer, F. Sadlo, and B. Schindler, *A comparison of finite-time and finite-size Lyapunov exponents*, in Topological Methods in Data Analysis and Visualization III, Springer, New York, to appear.
[34] S. B. Pope, P. K. Yeung, and S. S. Girimaji, *The curvature of material surfaces in isotropic turbulence*, Phys. Fluids A, 1 (1989), 2010.
[35] I. I. Rypina, M. G. Brown, F. J. Beron-Vera, H. Kocak, M. J. Olascoaga, and I. A. Udovydchenkov, *On the Lagrangian dynamics of atmospheric zonal jets and the permeability of the stratospheric polar vortex*, J. Atmospheric Sci., 64 (2007), pp. 3595–3610.
[36] J. A. Serret, *Sur quelques formules relatives à la théorie des courbes à double courbure*, J. Math. Pures Appl., 16 (1851), pp. 193–207.
[37] S. C. Shadden, F. Lekien, and J. E. Marsden, *Definition and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows*, Phys. D, 212 (2005), pp. 271–304.
[38] D. C. Sorensen, *Newton's method with a model trust region modification*, SIAM J. Numer. Anal., 19 (1982), pp. 409–426.
[39] P. Tallapragada and S. D. Ross, *A set oriented definition of finite-time Lyapunov exponents and coherent sets*, Commun. Nonlinear Sci. Numer. Simulat., 18 (2013), pp. 1106–1126.
[40] J.-L. Thiffeault, *Derivatives and constraints in chaotic flows: Asymptotic behaviour and a numerical method*, Phys. D, 172 (2002), pp. 139–161.
[41] J.-L. Thiffeault, *Stretching and curvature of material lines in chaotic flows*, Phys. D, 198 (2004), pp. 169–181.
[42] J.-L. Thiffeault and A. H. Boozer, *Geometrical constraints on finite-time Lyapunov exponents in two and three dimensions*, Chaos, 11 (2001), pp. 16–28.
[43] H. Xu, N. T. Ouellette, and E. Bodenschatz, *Curvature of Lagrangian trajectories in turbulence*, Phys. Rev. Lett., 98 (2007), 050201.