# Globally optimal stretching foliations of dynamical systems * 

Sanjeeva Balasuriya ${ }^{\dagger}$ and Erik Bollt ${ }^{\ddagger}$

Abstract. There is a well-established understanding of maximizing and minimizing local stretching associated with a map; this can be couched in terms of an eigenproblem associated with the Cauchy-Green tensor. The local eigenvalues correspond to stretching rates, while the eigenvector directions identify maximal and minimal stretching directions. Generating hyperstreamlines of these eigenvector fields has been popular; however various claims that these collection of curves correspond to global optimizers of stretching remain unsubstantiated. In this paper, a formulation in terms of restricted foliations is presented, for which these curves form a well-defined solution as the most and least unstable objects, with respect to a global definition. The theory works for any map in $n$-dimensions, derived for example from repeated applications of a discrete dynamical system, or from a finite-time flow of a continuous one. Further insight is obtained into the two-dimensional situation via an elementary approach, including a simple singularity classification criterion and an understanding of numerical artefacts in foliation computations (e.g., special curves along which hyperstreamlines veer and/or do not cross). Illustrations and validations of the results to the Hénon map, the double-gyre flow, and the standard (Chirikov) map are provided.

Key words. Foliations, maximum stretching, tensor field lines.
MSC codes.

1. Introduction. A central topic of dynamical systems theory involves analysis of instabilities, since this is the central ideas behind the possibility of forecast time horizon, or even of ease of control of future outcomes. The preponderance of work has involved analysis of local instability, whether by the Hartman-Grobman theorem and center manifold theorem [18] for periodic orbits and similarly for invariant sets [36]. For general orbits, local instability is characterized by Oseledec spaces [31] which are identified via Lyapunov exponents [37] and Lyapunov vectors [41, 34], or their finite-time counterparts. These are associated with local stretching: given a sufficiently differentiable map $\boldsymbol{F}$ on $\mathbb{R}^{n}$, the local stretching of an infinitesimal sphere placed at a general location leads to the concept of the Cauchy-Green tensor, whose largest eigenvalue represents the stretching rate in the corresponding eigenvector direction. Similarly, the smallest eigenvalue encodes the direction of least stretching, and therefore a complete understanding of locally optimizing stretching is available. The concepts of Lyapunov vectors and fields $[37,34,41,3]$ as well as (finite-time versions of) Oseledets spaces [31], are all connected to these ideas.

In this paper, we take the view that global stretching is related to a global view of instabilities. The related organizing skeleton of orbits must therefore be premised on analysis

[^0]of globally optimal stretching. Here, orbits will be in relation to $n$-dimensional maps which can be derived from various sources: a finite sequence of discrete maps, or a flow occurring over a finite time period. The latter situation is particularly relevant when seeking regions in unsteady flows which remain 'coherent' over a given time period [3]. In all these cases, we emphasize that we are not seeking to understand stretching in the infinite-time limit-which is the focus in many classical approaches [31, 36]-but rather stretching associated with a one-step map derived from any of these approaches. From the applications perspective, the one-step map would be parametrized by the discrete or continuous time over which the map operates, and this would of necessity be finite in any computational implementation. Neither are we seeking the propagation of stretching due to model uncertainty, a topic which is receiving attention recently [2].

Local stretching information in deterministic systems has been used by many to apparently infer global stretching properties. By drawing hyperstreamlines [11] which are locally tangential to the leading eigenvector direction, one presumably obtains a curve which is associated with 'maximal stretching' in some way. However, this statement is hard to justify: in what way is such a curve the solution to a global maximization problem? There have been attempts by several researchers to pose such questions, but each is problematic (see Section 2.2). In this paper, therefore, we seek to pose a global optimization problem for which these curves form a solution. It turns out that in order to do this, the concept of a restricted foliation (closely related to a punctured foliation [29]) needs to be introduced. This theory is presented in Section 2, in which we introduce the concept of Stretching Optimizing Restricted Foliations (SORFs).

Our theory works in $\mathbb{R}^{n}$ in general, but the two-dimensional problem is particularly attractive since explicit formulas can be derived using elementary arguments. In Section 3, we focus on several aspects of the global optimizing foliations: (i) their singularity classification, (ii) inevitability of a 'branch-cut' type phenomenon when attempting to construct global vector fields from the eigenfields, (iii) numerically computed foliation curves stopping abruptly when attempting to cross certain specialized curves (which we can characterize) horizontally or vertically, and (iv) veering along spurious curves. The singularities associated with the Cauchy-Green tensor strongly influence these issues. The first two aspects above are wellknown in the literature based on local stretching, and the (non-degenerate in our terminology) singularities are referred to as either "1-prong" and "3-prong," or "wedge" and "trisector," depending on the context. However, the last two do not seem to have been realized by practitioners, and consequently some existing numerical computations may indeed be incorrect.

While our formulation here is motivated by stretching in dynamical systems, we note that there are related structures in other contexts in the mathematical literature. The so-called tensor fields in topology are closely related to the SORFs that we define here. Specifically, Teichmüller theory [24] concerning so called isomorphism class of "marked" Riemann surfaces has been shown to be closely related to the geometric trajectory structure in the special case of quadratic differentials [23]. Our SORFs are moreover related to an equivalence relation between measured foliations $[32,16]$ called Whitehead-equivalence as discussed in braid theory [1]. These include in [32] elements closely related to the topological classification of nondegenerate singularities we define here from our interest in dynamical systems. In braid theory $[16,27,4]$, the so called " 1 -prong" and "3-prong" singularities appear significantly comparable
to the nondegenerate singularities of our SORFs, and moreover feature in the estimation of the entropy of a braid [4]. In a more applied direction, tensor fields and their singularities (usually called "wedge" and "trisector" in this literature) arise in two-dimensional graphical visualization $[8,42,12,40]$. The applications range from solid mechanics [8] to computer graphics $[42,12,40]$. These studies variously focus on extracting topological skeletons of vector fields [12], designing tensor fields for capturing graphical features [42], visualization for asymmetric tensors [8] and in understanding the local topology and bifurcations [40].

This paper is organized as follows. In Section 2, we present our theory for restricted foliations of globally optimal stretching. In doing so, we first discuss local stretching in Section 2.1, as well as attempts to extend these ideas globally in Section 2.2, before presenting our theory in Section 2.3. Section 3 is devoted to the two-dimensional interpretation, with a strong effort to provide alternative but simple expressions by using elementary ideas while taking advantage of two-dimensional intuition. In the subsections, we find equivalent expressions to those in $\mathbb{R}^{n}$ in Section 3.1, categorize singularities in Section 3.2, show the impossibility of forming a vector field from the eigenfields in the presence of singularities in Section 3.3, and finally describe numerical artefacts arising from specialized curves and develop an integral curve solution in Section 3.4. In Section 4, we demonstrate computations of globally optimal restricted foliations for several well-known examples: the Hénon map [22], the Chirikov (standard) map [9], and the double-gyre flow [37], each implemented over a finite time. The aforementioned numerical issues are highlighted in these examples. We conclude with some short remarks in Section 5.
2. Optimizing stretching. Let $\Omega$ be an open bounded subset of $\mathbb{R}^{n}$ consisting of a finite union of connected sets, each of whose closure has at most a finite number of boundary components. So $\Omega$ may, for example, consist of disconnected open sets, or (if in $\mathbb{R}^{2}$ ) entities which are topologically equivalent to the interior of an annulus. We consider a map $\boldsymbol{F}: \Omega \rightarrow$ $\mathbb{R}^{n}$ such that $\boldsymbol{F} \in \mathrm{C}^{2}(\Omega)$, and are interested in the stretching resulting from the application of $\boldsymbol{F}$.

Our approach embraces the fact that $\boldsymbol{F}$ might be generated in various ways. It can be simply one iteration of a given map, multiple (finitely-many) iterations of a map, or even the application of a finite sequence of maps. It can also be the flow-map generated from a nonautonomous flow over a finite time. In this sense, $\boldsymbol{F}$ encapsulates the fact that finiteness is inevitable in any numerical, experimental or observational situation, while allowing for both discrete and continuous time, as well as nonautonomy. The time over which the system operates can be thought of as a parameter which is encoded within $\boldsymbol{F}$, and its effect can be investigated if needed by varying this parameter. In general, we ignore how $\boldsymbol{F}$ were generated, conscious though that it can arise in various ways. Once specified, it will only make sense to consider $\boldsymbol{F}$ applied one time to $\Omega$. Thus standard approaches to dynamical systems which may, for example, examine the sign of the real part of eigenvalues associated with a fixed point in infer stability, no longer make sense. This is because the map $\boldsymbol{F}$ is not applied repetitively, and hence connecting stability with the rate of growth of powers of the eigenvalues is not justifiable. The issue in this instance is to understand stretching optimization in relation to the one iteration of $\boldsymbol{F}$.
2.1. Local stretching. In this subsection, we briefly outline the well-known understanding for local stretching, for subsequent extension to a notion of global stretching. Consider a small line segment of length $\delta$ (where $0<\delta \ll 1$ ) emanating from a location $\boldsymbol{x} \in \Omega$. Choosing this deviation in the direction of a unit vector $\hat{\boldsymbol{n}}$, we see that the deviation between the endpoints of the line once it has been mapped by $\boldsymbol{F}$ is therefore $\boldsymbol{F}(\boldsymbol{x}+\delta \hat{\boldsymbol{n}})-\boldsymbol{F}(\boldsymbol{x})$. The local stretching engendered by $\boldsymbol{F}$ with respect a vanishingly small initial line in the direction $\hat{\boldsymbol{n}}$ at a location $\boldsymbol{x}$ is therefore

$$
\begin{equation*}
\Lambda(\boldsymbol{x}, \hat{\boldsymbol{n}}):=\lim _{\delta \rightarrow 0} \frac{\|\boldsymbol{F}(\boldsymbol{x}+\delta \hat{\boldsymbol{n}})-\boldsymbol{F}(\boldsymbol{x})\|}{\delta}=\|\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x}) \hat{\boldsymbol{n}}\| . \tag{2.1}
\end{equation*}
$$

The standard approach to determine the directions $\hat{\boldsymbol{n}}$ to optimize the local stretching at each point $\boldsymbol{x}$ uses the observation

$$
\begin{align*}
\Lambda^{2}(\boldsymbol{x}, \hat{\boldsymbol{n}}) & =[\boldsymbol{\nabla} F(\boldsymbol{x}) \hat{\boldsymbol{n}}]^{\top}[\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x}) \hat{\boldsymbol{n}}]=\hat{\boldsymbol{n}}^{\top}[\nabla \boldsymbol{F}(\boldsymbol{x})]^{\top}[\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x})] \hat{\boldsymbol{n}} \\
& =\hat{\boldsymbol{n}}^{\top} C(\boldsymbol{x}) \hat{\boldsymbol{n}} ; \quad\|\hat{\boldsymbol{n}}\|=1, \tag{2.2}
\end{align*}
$$

in which the Cauchy-Green (strain) tensor field on $\Omega$ is defined by

$$
\begin{equation*}
C(\boldsymbol{x}):=[\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x})]^{\top}[\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x})] \quad ; \quad \boldsymbol{x} \in \Omega . \tag{2.3}
\end{equation*}
$$

Since $C$ is symmetric, its eigenvalues $\lambda_{j}(j=1,2, \cdots, n)$ are real and nonnegative, and it enjoys a full set of eigenvectors $\boldsymbol{w}_{j}$. This is true even if the eigenvalues have multiplicity greater than 1 [5, c.f.], but in this case each eigenvalue with multiplicity $m$ has $m$ linearly independent eigenvectors. All eigenfields depend on $\boldsymbol{x}$, a fact we will sometimes suppress in the notation for brevity.

If the orthogonal transformation matrix $T(\boldsymbol{x})$ has the normalized eigenvectors $\boldsymbol{w}_{j}$ as columns, then $T^{\top}(\boldsymbol{x})=T^{-1}(\boldsymbol{x})$ has the $\boldsymbol{w}_{j}$ as rows, and $C(\boldsymbol{x})$ is diagonalizable in the form $C(\boldsymbol{x})=T(\boldsymbol{x}) \operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \cdots, \lambda_{n}\right) T^{\top}(\boldsymbol{x})$. We will agree to order the eigenvalues in an increasing fashion, i.e., $0 \leq \lambda^{-}=\lambda_{1} \leq \lambda_{2} \leq \lambda_{3} \leq \cdots \leq \lambda_{n-1} \leq \lambda_{n}=\lambda^{+}$, where we are using the notation $\lambda^{ \pm}$for the largest and smallest eigenvalues. Consequently

$$
\Lambda^{2}(\boldsymbol{x}, \hat{\boldsymbol{n}})=\left[T(\boldsymbol{x})^{\top} \hat{\boldsymbol{n}}\right]^{\top} \operatorname{diag}\left(\lambda_{1}, \lambda_{2}, \cdots, \lambda_{n}\right)\left[T(\boldsymbol{x})^{\top} \hat{\boldsymbol{n}}\right],
$$

in which the $j$ th component of $T(\boldsymbol{x})^{\top} \hat{\boldsymbol{n}}$ is the projection of $\boldsymbol{w}_{j}$ the direction of the unit vector $\hat{\boldsymbol{n}}$. Thus if $\hat{\boldsymbol{n}}$ were chosen to be exactly $\boldsymbol{w}_{j}$, we get $\Lambda^{2}\left(\boldsymbol{x}, \boldsymbol{w}_{j}\right)=\lambda_{j}$, thereby picking out the corresponding eigenvalue. So if $\boldsymbol{x}$ is fixed, and the problem of determining the direction $\hat{\boldsymbol{n}}$ such that the local stretching $\Lambda(\boldsymbol{x}, \hat{\boldsymbol{n}})$ is optimized (either maximized or minimized) is considered, the following standard results are clear:

- Choosing $\hat{\boldsymbol{n}}$ in the direction of the eigenvector $\boldsymbol{w}^{+}(\boldsymbol{x})=\boldsymbol{w}_{n}(\boldsymbol{x})$ of $C(\boldsymbol{x})$ corresponding to its largest eigenvalue $\lambda^{+}(\boldsymbol{x})$, represents the direction of the infinitesimal line segment such that $\Lambda^{2}$ (and consequently $\Lambda \geq 0$ ) is maximized;
- Choosing $\hat{\boldsymbol{n}}$ in the direction of the eigenvector $\boldsymbol{w}^{-}(\boldsymbol{x})=\boldsymbol{w}_{1}(\boldsymbol{x})$ of $C(\boldsymbol{x})$ corresponding to its smallest eigenvalue $\lambda^{-}(\boldsymbol{x})$, represents the direction of the infinitesimal line segment such that $\Lambda$ is minimized.

Since $\lambda^{+}(\boldsymbol{x})$ is the largest value that $\Lambda^{2}(\boldsymbol{x}, \hat{\boldsymbol{n}})$ takes, the maximum local stretching is given by the field $\sqrt{\lambda^{+}(\boldsymbol{x})}$, which is in fact the spectral norm $\|\boldsymbol{\nabla} \boldsymbol{F}(\boldsymbol{x})\|$ by (2.1). While $\boldsymbol{F}$ is general here, it is worth remarking that if $\boldsymbol{F}$ were derived from a finite-time flow over a time-length of $T$, then this field is a scaled version of the well-known Finite-Time Lyapunov Exponent (FTLE) field

$$
\begin{equation*}
\operatorname{FTLE}(\boldsymbol{x})=\frac{1}{|T|} \ln \sqrt{\lambda^{+}(\boldsymbol{x})} \tag{2.4}
\end{equation*}
$$

representing the exponential rate-of-stretching along a trajectory beginning at $\boldsymbol{x}$.
At points $\boldsymbol{x} \in \Omega$ at which $\lambda_{j}(\boldsymbol{x})<\lambda_{k}(\boldsymbol{x})$ for some $j \neq k$, the fact that these eigenvalues of the symmetric matrix $C$ are different ensures that the corresponding eigenvalues are orthogonal, i.e., $\boldsymbol{w}_{j}(\boldsymbol{x})^{\top} \boldsymbol{w}_{k}(\boldsymbol{x})=0$. This property is indeed valid even if either of $\lambda_{j}$ or $\lambda_{k}$ had multiplicity greater than 1 : any eigenvector associated with $\lambda_{j}$ is orthogonal to any eigenvector associated with $\lambda_{k}$.

For each fixed $j$, the eigenvectors $\boldsymbol{w}_{j}$ are sometimes referred to as the Lyapunov vectors, which are finite-time analogs of Oseledets vectors for infinite-time flows/maps. The quantity $\boldsymbol{w}_{j}(\boldsymbol{x})$ does not in general represent a $\mathrm{C}^{1}$-vector field on $\Omega$, since eigenvectors are only defined modulo scalar multiples and only define a direction (field). (A 'reversed' direction is also permitted by choosing a negative multiple.) However, locally in regions where the multiplicity of $\lambda_{j}$ is 1 , it is possible to construct a vector field by choosing $\boldsymbol{w}_{j}$ in consistent directions. Moreover, this vector field will be locally $\mathrm{C}^{1}$ because the components of $C$, comprising first spatial derivatives of $\boldsymbol{F}$, are $\mathrm{C}^{1}$ and the eigenproblem for $C$ is constructed using an algebraic combination of these components. From this, one can locally construct a family of curves which are tangential to these directions:

Definition 2.1 (Hyperstreamlines). In regions where the multiplicity of the eigenvalue $\lambda_{j}$ is 1 , then $\boldsymbol{w}_{j}(\boldsymbol{x})$ can be chosen to be a $\mathrm{C}^{1}$-vector field, and the associated hyperstreamlines [11, e.g.] are curves which are everywhere tangential to the vector field $\boldsymbol{w}_{j}(\boldsymbol{x})$.

The $j=1$ streamlines are associated with curves tangential to the local directions of least stretching, since they are identified with the scalar field $\lambda_{1}(\boldsymbol{x})=\lambda^{-}(\boldsymbol{x})$. The hyperstreamlines generated from the local stretching minimization are therefore tangential to the vector field $\boldsymbol{w}^{-}$which is well defined if the multiplicity of $\lambda^{-}$is 1 . Similarly, the tangent vectors to the $j=N$ streamlines, being associated with the scalar field $\lambda_{n}(\boldsymbol{x})=\lambda^{+}(\boldsymbol{x})$, represent the directions which are locally maximally stretching. Notice that by extending the hyperstreamlines over all parts of $\Omega$ for which the multiplicity condition is satisfied, one obtains global objects constructed from local stretching properties. In what sense do these entities satisfy a global optimization problem? In this paper, we formulate a notion of global stretching, and will establish how these hyperstreamlines coincide with the solution to a global stretching optimization problem.
2.2. Global optimization and hyperstreamlines. There have been several attempts at casting the hyperstreamlines obtained from the local stretching analysis, or entities derived from them, in terms of some sort of non-pointwise optimization problem. In this subsection, we briefly describe some of these approaches, and comment on why they do not necessarily fit the bill as global optimizers. In response to these issues we will in the next subsection
establish a formulation which casts the hyperstreamlines as a solution to a specific type of global stretching optimization problem.

We will first reiterate an issue related to determining streamlines: repeated eigenvalues of the Cauchy-Green tensor. If an eigenvalue at any point $\boldsymbol{x}$ has multiplicity greater than 1 , this causes difficulty in unambiguously identifying the 'corresponding' eigenvectors. If for example $\lambda_{3}=\lambda_{4}=\lambda_{5}$ represents an eigenvalue of multiplicity 3 , which vector from the threedimensional eigenspace at each point should be taken as $\boldsymbol{w}_{3}$ ? The lack of clarity of this choice means that it is not then possible to use appropriately scaled versions of any chosen $\boldsymbol{w}_{3}$ as a $\mathrm{C}^{1}$-vector field at that point. Since we will be principally concerned with minimum and maximum stretching, our focus is on the constructed eigenvector fields $\boldsymbol{w}^{-}(\boldsymbol{x})=\boldsymbol{w}_{1}(\boldsymbol{x})$ and $\boldsymbol{w}^{+}(\boldsymbol{x})=\boldsymbol{w}_{n}(\boldsymbol{x})$, which will have potential difficulties if $\lambda^{-}(\boldsymbol{x})=\lambda_{1}(\boldsymbol{x})$ and $\lambda^{+}(\boldsymbol{x})=\lambda_{n}(\boldsymbol{x})$ (respectively) have multiplicity greater than 1.

A two-dimensional situation in which hyperstreamlines appear is when subsets of them are apparently solutions to a global variational problem of length optimization [21]. The map $\boldsymbol{F}$ in this case [21] is derived from a flow over a specified finite time period, and here since $n=2$ the relevant eigenvalues are $\lambda^{-}=\lambda_{1} \leq \lambda_{2}=\lambda^{+}$and with corresponding eigenvectors $\boldsymbol{w}_{1,2}$. Here, the intuitive problem is in determining curves whose mapped length is maximized or minimized in comparison to nearby curves (our emphasis, highlighting the variational problem which is seemingly solved). Solutions to this problem are then defined to be 'hyperbolic transport barriers' [21, Defs. 1,2]. Though these definitions do not place any restrictions on the classes of curves considered, the 'most length maximizing' development quickly limits to seeking curves from among the hyperstreamlines associated with $\lambda_{2}=\lambda^{+}$(as described in the quick argument via geodesics in Appendix B and the algorithm in $\S 7.1$; see also the boundary conditions of Eq. (15) [21]). Intuitively, of course, this makes sense, since locally at any point, the direction $\boldsymbol{w}_{2}$ is the one in which there is maximal stretching. An infinitesimal curve at that point certainly should be in this direction for maximality of length increase.

However, it is not clear how a curve comprising an amalgamation of infinitesimal curves with directionality $\boldsymbol{w}_{2}$ are those whose 'length is maximized' because the there is no specified restriction of the classes of curves considered, and thus other (non-hyperstreamline) curves should be permissible. (The optimization problem posed [21] is not, for example, selecting from among curves with fixed endpoints which have maximal stretching; this problem cannot have a hyperstreamline solution if the endpoints are not on the same hyperstreamline.) From among these hyperstreamlines, the development [21] comes up with conditions-locally at each point-which ensures that a particular hyperstreamline has maximal stretching in comparison to nearby hyperstreamlines [21]. In other words, the hyperstreamlines seem to be the class of curves to which the optimization problem is restricted. Our emphasis here is different: is there a global optimization problem for which the hyperstreamlines appear as a solution?

Hyperstreamlines also play a prominent role in the influential article by Haller [19], where (subsets of) the hyperstreamlines help identify what he calls 'hyperbolic Lagrangian Coherent Structures' (hyperbolic LCSs) in $n$-dimensional flows. This is the basis of what is now a well-established method for seeking flow barriers over a finite-time in fluid flows [20, 15, 30]. A hyperbolic LCS is defined to be "locally most repelling or attracting material surface" [19, Def. 1] where in this case the map is derived explicitly from a flow on an $n$-dimensional open set over a specified finite-time interval. The intuitive idea is to determine a co-dimension-

1 surface at the initial time, such that the attraction/repulsion towards it due to the flow over the time-interval is optimal in comparison to nearby surfaces. The 'local' here implies a comparison to locally-close surfaces, and so this is in reality a global problem because the surfaces are extended in space.

The surfaces are initially confined to "finite-time hyperbolic material surfaces" [19, Def. 3] which are surfaces which pointwise have a normal stretching rate of at least 1 , and with the normal stretching rate being larger than the tangential one. This automatically implies that at each point one has the normal vector being $\boldsymbol{w}_{n}$, and $\lambda_{n}>\lambda_{n-1}$ [19, Theorem 7, Proposition 8]. The first condition is obvious because the direction of optimal local stretching is $\boldsymbol{w}_{n}$, which must then be chosen in the normal direction. The insistence of normal stretching being strictly larger than any tangential stretching moreover precludes the possibility of $\lambda_{n-1}=\lambda_{n}$ and consequently identifying a tangential direction $\boldsymbol{w}_{n-1}$ which has an equal local stretching. We emphasize that the surfaces considered by this process therefore obey these conditions pointwise as opposed to there being any definition of repelling from a surface. The surfaces are therefore, effectively by definition, amalgamated objects constructed from maximal stretching directions (when they exist) at each point. In our language, they are simply surfaces which are everywhere orthogonal to the $\boldsymbol{w}_{n} \mathrm{~s}$ (whenever they are well-defined). Let us call such surfaces (which may or may not be connected) $\mathcal{S}$. It is from surfaces in $\mathcal{S}$ that those which are locally most repelling/attracting in comparison to nearby surfaces in $\mathcal{S}$ are selected as "locally most repelling or attracting material surfaces" [19, Def. 1]. (A measure for "most repelling" for a global, as opposed to an infinitesimal, surface would seem to be relevant here; we relegate a more detailed discussion on this and how the algorithm selects from among the surfaces in $\mathcal{S}$ to Appendix A.) Again, this is different from what we seek here, instead trying to understand to what global optimization problem do the hyperstreamlines which orthogonally pierce $\mathcal{S}$ form a well-defined solution.

Returning to the general framework, we know that at a given location, an infinitesimal curve in the direction of $\boldsymbol{w}_{n}$ (associated with the maximal eigenvalue $\lambda_{n}=\lambda^{+}$, as long as $\boldsymbol{w}_{n}$ is well-defined in the sense that $\lambda_{n}$ has multiplicity 1) provides a curve which stretches most, while one in the direction of $\boldsymbol{w}_{1}$ (again subject to similar caveats) stretches least. Given this local property, to what global problem associated with non-infinitesimal curves do the relevant hyperstreamlines form a solution? Can this problem be posed such that the higher multiplicity issue is dealt with automatically? This problem is different from those discussed in this section; in those cases [21, 19, 20, 15, 30], the optimization problem is more on how one can select from among these hyperstreamlines (or surfaces orthogonal to them).
2.3. Global stretching optimization. We will now establish our theory for how the hyperstreamlines can be posed as the solution to a global stretching optimization problem. We require some definitions.

Definition 2.2 (Singularity sets). The minimal singular set $S_{\min } \subseteq \Omega$ is defined by

$$
\begin{equation*}
S_{\min }:=\left\{\boldsymbol{x} \in \Omega: \text { multiplicity }\left(\lambda_{1}(\boldsymbol{x})\right)>1\right\}, \tag{2.5}
\end{equation*}
$$

while the maximal singular set $S_{\max } \subseteq \Omega$ is

$$
\begin{equation*}
S_{\max }:=\left\{\boldsymbol{x} \in \Omega: \text { multiplicity }\left(\lambda_{n}(\boldsymbol{x})\right)>1\right\} . \tag{2.6}
\end{equation*}
$$

The corresponding restricted sets will be defined by

$$
\begin{equation*}
\Omega_{\min }:=\Omega \backslash S_{\min } \quad \text { and } \quad \Omega_{\max }:=\Omega \backslash S_{\max } . \tag{2.7}
\end{equation*}
$$

on which the eigenvectors $\boldsymbol{w}^{-}$and $\boldsymbol{w}^{+}$respectively form well-defined direction fields.
When the multiplicity of the eigenvalue $\lambda^{+}$is $m>1$, the corresponding eigenvectors span $m$-dimensional space and consequently a unique directionality cannot be inferred for $\boldsymbol{w}^{+}$. This means that the hyperstreamlines associated with $\lambda^{+}$become ill-defined. Such points are collected together in the singularity set $S_{\text {max }}$. Similarly, the singularity set $S_{\text {min }}$ consists of points in $\Omega$ such that a locally minimal stretching direction is not well-defined. Generically, these sets will consist of a finite number of points (as will be apparent in the numerics in Section 4), but we will allow a broader class as given below:

Definition 2.3 (Puncture sets). The set of puncture sets is defined by

$$
\begin{equation*}
\mathcal{P}:=\{P \subseteq \Omega: P \text { is a finite union of compact connected sets }\} . \tag{2.8}
\end{equation*}
$$

Definition 2.4 (Restricted foliation). $A$ restricted foliation $f$ on the open bounded set $\Omega$ consists of a family of curves defined on a restricted set $\Omega_{P}=\Omega \backslash P$, where $P \in \mathcal{P}$, such that (a) The curves of $f$ ('the leaves of the foliation') are disjoint;
(b) The union of all these curves covers $\Omega_{P}$;
(c) The tangent vector varies in a $\mathrm{C}^{1}$-smooth fashion along each curve.

The restricted foliation therefore depends on the choice of the set $P$ from the puncture set, with the foliation curves of $f$ only being defined within the restricted set $\Omega_{P}=\Omega \backslash P$.

Our definition is consistent with the local properties expected from a formal definition of foliations on manifolds [28], but bears in mind that $\Omega_{P}$ is not generally a manifold because of the omission of the closed set $P$ from $\Omega$. We remark that if $P$ consists of a finite number of points, our restricted foliation definition is equivalent to that of a 'punctured foliation' [29] or 'measured foliation [39] on $\Omega$, where the punctures are at the points in $P$. Given that we allow $P$ to be more general, we use the term 'restricted foliation.' Definition 2.4 also allows for the possibility of $P$ being the empty set, in which case the chosen restricted foliation is a genuine foliation on $\Omega$.

The properties of Definition 2.4 ensure that every restricted foliation $f$ is associated with a unique $\mathrm{C}^{1}$-smooth direction field on its chosen restricted set $\Omega_{P}$ in the following sense. Given a point $\boldsymbol{x} \in \Omega_{P}$, there exists a unique curve from $f$ which passes through it. The unit tangent line drawn at this point is unique up to multiplication by -1 , and thereby defines a direction $\hat{\boldsymbol{n}}$. Conversely, suppose a $\mathrm{C}^{1}$-smooth direction field $\hat{\boldsymbol{n}}$ is defined on $\Omega_{P}$. Given an arbitrary point $\boldsymbol{x}_{\alpha} \in \Omega_{P}$, the existence of solutions to the differential equation

$$
\frac{\mathrm{d} \boldsymbol{r}(s)}{\mathrm{d} s}=\hat{\boldsymbol{n}}(\boldsymbol{r}(s)) \quad ; \quad \boldsymbol{r}(0)=\boldsymbol{x}_{\alpha}
$$

ensures that there is an integral curve given in parametric form by $\boldsymbol{r}(s)$, which is $\mathrm{C}^{1}$-smooth in the parameter $s$. The curve can be evolved in both directions from $\boldsymbol{x}_{\alpha}$ by either going forward or backward in $s$. The fact that $\hat{\boldsymbol{n}}$ is only defined up to a multiple $\pm 1$ therefore does not interfere with this; multiplying $\hat{\boldsymbol{n}}$ by -1 is equivalent to simply reversing $s$ in the
equation. Finding an integral curve in this way is possible for each and every $\boldsymbol{x}_{\alpha} \in \Omega_{P}$, and the uniqueness property ensures that separate curves do not intersect one another. Moreover, $\Omega_{P}$ is spanned by these curves because $\Omega_{P}=\bigcup_{\alpha} \boldsymbol{x}_{\alpha}$, ensuring that there is a curve passing through every point $\boldsymbol{x}_{\alpha}$. Hence, this process generates a unique restricted foliation $f$ on $\Omega_{P}$.

We are now in a position to define the global stretching which we seek to optimize.
Definition 2.5 (Global stretching). The global stretching, dependent on a puncture set $P \in$ $\mathcal{P}$ and a restricted foliation $f$ defined on $\Omega_{P}=\Omega \backslash P$, is

$$
\begin{equation*}
\Sigma(P, f):=\int_{\Omega_{P}} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x} \tag{2.9}
\end{equation*}
$$

in which $\hat{\boldsymbol{n}}_{f}(\boldsymbol{x})$ is the (unit) direction field induced on $\Omega_{P}$ by a choice of restricted foliation $f$, and $\mathrm{d} \boldsymbol{x}=\mathrm{d} x_{1} \mathrm{~d} x_{2} \cdots \mathrm{~d} x_{n}$ is a shorthand notation for the volume element in $\Omega$.

As central premise of this work, we seek sets $P \in \mathcal{P}$ and restricted foliations $f$ which optimize (maximize, as well as minimize) $\Sigma$. Partitions of $\Omega$ which are extremal in this way represent the greatest instability or most stability associated with the dynamical system, and so orbits associated with these are distinguished for their corresponding difficulties in forecasting, or alternatively, relative coherence.

Theorem 2.6 (Stretching Optimizing Restricted Foliations (SORF)). The global optimizing restricted foliations have the following properties:
(a) If $S_{\max } \in \mathcal{P}$, any restricted foliation $f^{+}$which maximizes the global stretching (2.9) comprises the hyperstreamlines associated with the direction field $\boldsymbol{w}^{+}$on the restricted set $\Omega_{\text {max }}$.
(b) If $S_{\min } \in \mathcal{P}$, any restricted foliation $f^{-}$which minimizes the global stretching (2.9) comprises the hyperstreamlines associated with the direction field $\boldsymbol{w}^{-}$on the restricted set $\Omega_{\min }$.

Proof. We first tackle the maximization problem. We initially consider $P \in \mathcal{P}$ to be fixed, and suppose that $f$ is any restricted foliation associated with this choice (i.e., the $\mathrm{C}^{1}$-curves associated with $f$ are defined on $\left.\Omega_{P}=\Omega \backslash P\right)$. Recall that $S_{\max }$ is the singularity set such that the largest eigenvalue of $C, \lambda^{+}=\lambda_{n}$, has multiplicity greater than 1 . Now, since $\Omega$ is the union of the disjoint sets $\Omega_{\text {max }}$ and $S_{\text {max }}$,

$$
\begin{aligned}
\Sigma(P, f) & =\int_{\Omega \backslash P} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x} \\
& =\int_{\Omega_{\max } \backslash P} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x}+\int_{S_{\max } \backslash P} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x} \\
& =: I_{\Omega}(P)+I_{S}(P) .
\end{aligned}
$$

Consider first the integral $I_{\Omega}(P)$, the first term on the right-hand side of (2.10). For any $\boldsymbol{x} \in \Omega_{\max } \backslash P$, both $\hat{\boldsymbol{n}}_{f}$ and $\boldsymbol{w}^{+}$are well-defined. Moreoever, the local stretching obeys

$$
\Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \leq \Lambda\left(\boldsymbol{x}, \boldsymbol{w}^{+}(\boldsymbol{x})\right)=\sqrt{\lambda^{+}(\boldsymbol{x})}
$$

because the maximal local stretching is associated with the eigenvalue $\lambda^{+}$, and is associated with the direction $\boldsymbol{w}^{+}$. If the unit vector $\hat{\boldsymbol{n}}_{f}$ is not in the same direction (modulo a plus/minus
$\operatorname{sign})$ as $\boldsymbol{w}^{+}$, then strict inequality occurs because $\boldsymbol{w}^{+}$is well-defined as a direction within $\Omega_{\max }$ in which the multiplicity of $\lambda^{+}$is 1 . By the $\mathrm{C}^{1}$-nature of the restricted foliation curves of $f$, and the local differentiability of $\boldsymbol{w}^{+}$, this implies the presence of a neighborhood $N$ of $x$ in $\Omega_{\max } \backslash P$ such that

$$
\int_{N} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x} \leq \int_{N} \Lambda\left(\boldsymbol{x}, \boldsymbol{w}^{+}(\boldsymbol{x})\right) \mathrm{d} \boldsymbol{x}
$$

Exact equality can be obtained above if and only if the $\hat{\boldsymbol{n}}_{f}$ is chosen to be exactly in the same direction as $\boldsymbol{w}^{+}$throughout $N$. Thus, within $N$, one can make the contribution to the integral larger by adjusting $f$ to be identical to the hyperstreamlines associated with $\boldsymbol{w}^{+}$. Since this argument works for every point $\boldsymbol{x} \in \Omega_{\max } \backslash P, I_{\Omega}(P)$ takes on its maximum value exactly if $\hat{\boldsymbol{n}}_{f}$ is identical to $\boldsymbol{w}^{+}$. Such a restricted foliation will be called $f^{+}$, and this argument tells us the nature of the foliation curves of $f^{+}$on $\Omega_{\max } \backslash P$.

It now remains to optimize (2.10) over all $P \in \mathcal{P}$. By changing $P$ such that $P \cap \Omega_{\max }$ is larger, but $P \cap S_{\max }$ is the same, we will get a smaller contribution to $I_{\Omega}(P)$ (since it will be over a smaller domain). Therefore, we should choose $P \cap \Omega_{\max }$ to have zero measure. Given that $P$ is a finite union of compact connected sets, this implies that $P \cap \Omega_{\max }$ should be at most a finite collection of points. Any such point $p$ can be a puncture associated with the restricted foliation $f$, i.e., a point at which the tangent direction to the foliation curve passing through it is undefined. However, we have argued that we need to choose the foliation to be exactly the hyperstreamlines associated with $\boldsymbol{w}^{+}$in a neighborhood around $p$ because this punctured neighborhood is purely within $\Omega_{\max }$. Since $\boldsymbol{w}^{+}$is locally $\mathrm{C}^{1}$, this means that the puncture point $p$ can be "filled-in" with its hyperstreamlines, and the generated restricted foliation therefore does not have a singularity at $p$. Consequently, we can maximize $I_{\Omega}(P)$ by choosing $P \cap \Omega_{\max }=\emptyset$; that is, $P \subseteq S_{\text {max }}$.

Within this condition for $P$, the question now is the impact on $I_{S}(P)$. Note that

$$
I_{S}(P)=\int_{S_{\max } \backslash P} \Lambda\left(\boldsymbol{x}, \hat{\boldsymbol{n}}_{f}(\boldsymbol{x}) \mathrm{d} \boldsymbol{x} \leq \int_{S_{\max }} \sqrt{\lambda^{+}(\boldsymbol{x})} \mathrm{d} \boldsymbol{x}\right.
$$

While $\boldsymbol{w}^{+}$is not well-defined on $S_{\max }$ (it is any unit vector in the span of the eigenvectors associated with $\lambda^{+}$, which is at least two-dimensional), it is conceivable that the foliation curves of $f$ can be chosen in $S_{\max } \backslash P$ to be consistent with the span of the eigenvectors of $C$ associated with $\lambda^{+}$, and that these curves can connect up with the hyperstreamlines of $\boldsymbol{w}^{+}$at the boundary between $S_{\max }$ and $\Omega_{\max }$ to form the maximal foliation $f^{+}$. Moreover, choosing $P$ to have measure zero will enable equality above. In any case, choosing $P \subseteq S_{\max }$ in maximizing $I_{S}(P)$ cannot affect the conclusion of the theorem: that any restricted foliation $f^{+}$associated with maximizing $\Sigma$ coincides with the hyperstreamlines on $\Omega_{\max }$.

The proof of minimizing $\Sigma$ is analogous, and will be skipped.
The restricted foliation $f^{+}$defines a $\mathrm{C}^{1}$ collection of curves on $\Omega_{\text {max }}$, and we call these the Stretching Optimizing Restricted Foliation - Maximum, or SORF $_{\text {max }}$. Correspondingly, the curves $f^{-}$on $\Omega_{\min }$ will be called SORF $\min ^{\text {curves for the system. We will present many }}$ examples of such curves in Section 4.
3. Two-dimensional global stretching. Here, we specialize to the case where $\Omega$ is twodimensional. We will show here we can provide explicit formulas for the preceding general theory by using elementary methods and two-dimensional intuition. These lead to a straightforward classification of the singularities of the relevant restricted foliation, and new insights into the numerical artefacts that can arise when attempting to compute foliations via streamlines.

In this case where $n=2$, we use the standard Cartesian coordinates $\boldsymbol{x}=(x, y)$ for $\Omega \subset \mathbb{R}^{2}$, and write the map $\boldsymbol{F}$ in component form as

$$
\boldsymbol{F}\left(\binom{x}{y}\right)=\binom{u(x, y)}{v(x, y)}
$$

where the functions $u$ and $v$ are in $\mathrm{C}^{2}(\Omega)$. Thus, the functions

$$
\begin{equation*}
\phi(x, y):=\frac{u_{x}(x, y)^{2}+v_{x}(x, y)^{2}-u_{y}(x, y)^{2}-v_{y}(x, y)^{2}}{2} \tag{3.1}
\end{equation*}
$$

and

$$
\begin{equation*}
\psi(x, y):=u_{x}(x, y) u_{y}(x, y)+v_{x}(x, y) v_{y}(x, y) \tag{3.2}
\end{equation*}
$$

defined in terms of the partial derivatives $u_{x}, u_{y}, v_{x}$ and $v_{y}$ of the mapping $\boldsymbol{F}$, are $\mathrm{C}^{1}$-smooth. These are related to the elements of the Cauchy-Green tensor (2.3)

$$
\begin{gathered}
C(x, y)=\left(\begin{array}{ll}
c_{11} & c_{12} \\
c_{21} & c_{22}
\end{array}\right) \\
\phi=\frac{c_{11}-c_{22}}{2} \text { and } \psi=c_{12}=c_{21}
\end{gathered}
$$

3.1. Equivalent characterizations. In this section, we establish alternative, simpler, characterizations of the objects developed in Section 2 under the condition of two-dimensionality. We are able in many instances to develop these using elementary methods-without appealing to the Cauchy-Green tensor, for example - even though the results are equivalent.

In two dimensions, a unit vector direction $\hat{\boldsymbol{n}}$ can be easily parametrized by an angle $\theta \in[-\pi / 2, \pi / 2)$ such that

$$
\begin{equation*}
\hat{\boldsymbol{n}}=\binom{\cos \theta}{\sin \theta} . \tag{3.3}
\end{equation*}
$$

Since any multiple (including negative) of $\hat{\boldsymbol{n}}$ is permitted as a direction, it suffices to restrict $\theta$ to $[-\pi / 2, \pi / 2)$ to be able to represent all possible directions.

Lemma 3.1 (Local stretching in terms of angle). The local stretching (2.1) can be expressed in terms of the angle by

$$
\begin{equation*}
\Lambda^{2}=\phi \cos 2 \theta+\psi \sin 2 \theta+\frac{|\nabla u|^{2}+|\nabla v|^{2}}{2} \tag{3.4}
\end{equation*}
$$

Proof. See Appendix B.
Next, we note that since $\Omega$ is two-dimensional, there are only two eigenvalues related to stretching. Thus, if multiplicity is greater than 1 at any point $\boldsymbol{x}=(x, y)$, we get $\lambda^{-}=\lambda_{1}=$ $\lambda_{2}=\lambda^{+}$. The eigenspace at any such point spans two dimensions. Hence, $S_{\max }=S_{\min }$, and we will refer to this singularity set as the isotropic set $I$ since points in this set exhibit equal local stretching irrespective of the directionality $\hat{\boldsymbol{n}}$ of the line chosen in (2.1); an infinitesimal circular disk maps to a circular disk with no ellipticity. Now, (3.4) gives us the easy characterization:

Lemma 3.2 (Isotropic set). The isotropic set $I=S_{\min }=S_{\max }$ can be equivalently characterized by

$$
\begin{equation*}
I:=\{(x, y) \in \Omega: \phi(x, y)=0 \text { and } \psi(x, y)=0\} \tag{3.5}
\end{equation*}
$$

Proof. See Appendix B.
We define $\Omega_{I}:=\Omega \backslash I$ to be the restricted set such that $I$ is excluded from $\Omega$; all relevant eigenfields are $\mathrm{C}^{1}$-smooth in $\Omega_{I}$.

Lemma 3.3 (Equivalent characterizations of directions $w^{ \pm}$). On $\Omega_{I}$, the maximal and minimal local stretching directions $\boldsymbol{w}^{ \pm}$can be associated with directions encoded in the angles $\theta^{ \pm} \in[-\pi / 2, \pi / 2)$, for which two equivalent representations are

$$
\begin{align*}
\theta^{+}(x, y) & =\frac{1}{2} \tan ^{-1}(\psi(x, y), \phi(x, y)) \quad(\bmod \pi)  \tag{3.6a}\\
& =\tan ^{-1} \frac{-\phi(x, y)+\sqrt{\phi(x, y)^{2}+\psi(x, y)^{2}}}{\psi(x, y)} \quad(\bmod \pi) \tag{3.6b}
\end{align*}
$$

and

$$
\begin{align*}
\theta^{-}(x, y) & =\frac{\pi}{2}+\frac{1}{2} \tan ^{-1}(\psi(x, y), \phi(x, y)) \quad(\bmod \pi)  \tag{3.7a}\\
& =\tan ^{-1} \frac{-\phi(x, y)-\sqrt{\phi(x, y)^{2}+\psi(x, y)^{2}}}{\psi(x, y)} \quad(\bmod \pi) \tag{3.7b}
\end{align*}
$$

Proof. See Appendix B.
The notation $\tan ^{-1}$ in (3.6a) and (3.7a) represents the four-quadrant inverse tangent, in which $\tan ^{-1}(y, x)$ picks the angle in the appropriate quadrant dependent on the $(x, y)$ coordinates. This is useful since it is built into many computational packages (e.g., atan2 in Matlab). For the representation in (3.6b) and (3.7b), we note that while it appears that points where $\psi=0$ but $\phi \neq 0$ are not in the domain, these are removable singularities in the sense of keeping $\phi$ constant and letting $\psi \rightarrow 0$. Specifically,

$$
\left.\theta^{+}(x, y)\right|_{\psi=0}=\left\{\begin{array}{ll}
-\pi / 2 & \text { if } \phi<0  \tag{3.8}\\
0 & \text { if } \phi>0
\end{array} \quad \text { and }\left.\quad \theta^{-}(x, y)\right|_{\psi=0}= \begin{cases}0 & \text { if } \phi<0 \\
-\pi / 2 & \text { if } \phi>0\end{cases}\right.
$$

From the numerical perspective, there are advantages and disadvantages of each of the representations in Lemma 3.3, as will be demonstrated in subsequent computations. The orthogonality of the vectors $\boldsymbol{w}^{+}$and $\boldsymbol{w}^{-}$is easily observable in the fact that $\theta^{+}-\theta^{-}=\pi / 2(\bmod \pi)$.


Figure 3.1: Topological classification of nondegenerate singularities with respect to $\mathrm{SORF}_{\max }$ or -min (a) a 1-pronged (wedge) point, and (b) a 3-pronged (trisector) point. See Property 1, and compare with Fig. B.1.
3.2. Behavior near singularities. The fact that generic singularities in two-dimensional symmetric tensor fields typically are 'wedges' and 'trisectors' is well known in the computational community [11, 42, 14, e.g.]. Here, we show how our characterization in the previous sections can help categorize these using elementary arguments. By Lemma 3.2, we know that singularities are points where both $\phi$ and $\psi$ are zero. Since both $\phi$ and $\psi$ are $\mathrm{C}^{1}$-smooth in $\Omega$, their gradients are well-defined on $\Omega$. For a singularity to be nondegenerate, we must preclude either $\phi$ or $\psi$ possessing critical points at $\boldsymbol{p}$. Thus, we cannot get self-intersections of either $\phi=0$ or $\psi=0$ contours at $\boldsymbol{p}$, have local extrema of $\phi$ or $\psi$ at $\boldsymbol{p}$, or have a situation where $\phi$ or $\psi$ is constant in an open neighborhood around $\boldsymbol{p}$. Nondegeneracy also precludes $\phi=0$ and $\psi=0$ contours intersecting tangentially at $\boldsymbol{p}$ (although we will make some remarks about this situation later). These considerations allow us to say that a singularity $\boldsymbol{p}$ is nondegenerate if the orientation function

$$
\begin{equation*}
g(x, y):=\left.(\boldsymbol{\nabla} \phi \times \boldsymbol{\nabla} \psi) \cdot \hat{\boldsymbol{k}}\right|_{(x, y)}=\left.\operatorname{det} \frac{\partial(\phi, \psi)}{\partial(x, y)}\right|_{(x, y)} \tag{3.9}
\end{equation*}
$$

is nonzero. We explain in Appendix C how we obtain the following complete classification for nondegenerate singularities, as illustrated in Fig. 3.1:

Property 1 (1- and 3 -pronged singularities). Let $\boldsymbol{p} \in I$ be a nondegenerate singularity, and let $\hat{\boldsymbol{k}}$ be the unit-vector in the $+z$-direction (i.e., 'pointing out of the page' for a standard right-handed Cartesian system). Then,
(a) If $g(\boldsymbol{p})>0$, then $\boldsymbol{p}$ is a 1-pronged singularity ('wedge'), with nearby foliation of both $f^{+}$ and $f^{-}$topologically equivalent to Fig. 3.1(a); and
(b) If $g(\boldsymbol{p})<0$, then $\boldsymbol{p}$ is a 3-pronged singularity (a 'trisector'), with nearby foliation of both $f^{+}$and $f^{-}$topologically equivalent to Fig. 3.1(b).
The singularities occur in opposite directions for the two orthogonal foliations $f^{ \pm}$.
The ' 1 -pronged' and ' 3 -pronged' terminology is from the theory of measured foliations [39, 23], while the terms 'wedge' and 'trisector' are common in the computational literature [11, 42, 14, 38]. We also note that in the case of all singularities being nondegenerate, the curves


Figure 3.2: Some possible topologies for $\mathrm{SORF}_{\text {max }}$ near $\boldsymbol{p}$ when transversality is relaxed (see Appendix C for explanations of these structures).
on $\Omega_{I}$ may be thought of as a punctured foliation [29, e.g.] on $\Omega$. These two singularities also correspond to the index of the foliation being $+1 / 2$ and $-1 / 2$ respectively [35, 11]. The alternative form of classification as given in Property 1 does not seem to appear in the literature. These two topologically distinct singularities serve as the organizing skeleton around which the rest of the SORF smoothly vary.

To see why the topology of $f^{-}$is similar to that of $f^{+}$, imagine reflecting these curves in Fig. 3.1 about the vertical line going through $\boldsymbol{p}$. This generates an orthogonal set of curves, which are the complementary (orthogonal) foliation. Thus, $f^{+}$and $f^{-}$have the same topology near $\boldsymbol{p}$.

At the next-order of degeneracy, we will have $\phi=0$ and $\psi=0$ contours continuing to be curves, but now intersecting at $\boldsymbol{p}$ tangentially. In that case, it turns out that Fig. 3.2 gives the possible topologies for $\mathrm{SORF}_{\text {max }}$, which are explained in detail in Appendix C. If $\boldsymbol{p}$ is not an isolated point in $I$, then many other possibilities exist. The SORF $_{\min }$ in the mildly degenerate situations in Fig. 3.2 represent curves which are orthogonal to the pictured ones; their topology will be identical.
3.3. Discontinuity in Lyapunov vectors. We have determined slope fields $\theta^{+}$and $\theta^{-}$ which while optimizing local stretching, also maximize and minimize global stretching according to our formulation. Here we argue that it is generically not possible to express a $\mathrm{C}^{0}$-vector field on the closure of $\Omega_{I}$ from the $\theta^{ \pm}$angle fields. This specifically impacts numerical computations, and we give insight into numerical artefacts which can arise.

To determine a curve from the $\operatorname{SORF}_{\text {max }}$, we need to pick an initial point in $\Omega_{I}$, and evolve it according to 'the' vector field generated from $\theta^{+}$from (3.6a). A simple possibility would be to take the (unit) vector field

$$
\begin{equation*}
\boldsymbol{w}^{+}(x, y):=\binom{\cos \left[\theta^{+}(x, y)\right]}{\sin \left[\theta^{+}(x, y)\right]}, \tag{3.10}
\end{equation*}
$$

in which $\theta^{+}$is computed from (3.6a). Similarly, a natural vector field for $S O R F_{\min }$ would be

$$
\begin{equation*}
\boldsymbol{w}^{-}(x, y):=\binom{\cos \left[\theta^{-}(x, y)\right]}{\sin \left[\theta^{-}(x, y)\right]}, \tag{3.11}
\end{equation*}
$$



Figure 3.3: The map from $\Omega$ to $(\phi, \psi)$-space, illustrating the sets $I^{\prime}$ and $B^{\prime}$ to which the sets $I$ and $B$ map. In red, we have stated the value of the field $\theta^{+}$in (3.6a) in each quadrant.
where $\theta^{-}$is defined from (3.7a).
Property 2 (Generating foliation curves using vector fields). If generating a $\mathrm{SORF}_{\max }$ or $\mathrm{SORF}_{\min }$ curve in $\Omega_{I}$, we can in general find solutions to

$$
\begin{equation*}
\frac{d}{d s}\binom{x}{y}=\boldsymbol{w}(x(s), y(s)) \quad ; \quad\binom{x(0)}{y(0)}=\binom{x_{0}}{y_{0}} \tag{3.12}
\end{equation*}
$$

where $s$ is the parameter along the curve and $\left(x_{0}, y_{0}\right) \in \Omega_{I}$, and we can choose a Lyapunov vector field in the form

$$
\begin{equation*}
\boldsymbol{w}(x, y)=m(x, y) \boldsymbol{w}^{ \pm}(x, y) \tag{3.13}
\end{equation*}
$$

for a suitable scalar function $m$.
If we use $m \equiv 1$ on $\Omega_{I}$, the parametrization $s$ along the trajectory is exactly the arclength. However, more general scalar functions $m$ can be used in (3.12), reflecting the fact that the vector fields which generate the foliations are actually direction fields, and thus can be multiplied at each point by a scalar. The only restrictions are (i) $m$ can never be zero, because if it is, we introduce a spurious fixed point in the system (3.12) which 'stops' the curve, and (ii) $m$ is sufficiently smooth to ensure that the equation (3.12) has unique $\mathrm{C}^{1}$-smooth solutions. The choice of $m$ simply adjusts the parametrization along the curve. Notice that if we flip the sign of $m$ we would be going along the curve in the opposite direction.

To understand the generation of curves from (3.13), it helps to think of the mapping from $\Omega$ to $(\phi, \psi)$-space, illustrated in Fig. 3.3. We have already characterized an important subset of $\Omega$ in relation to this mapping: the isotropic set $I$ is the kernel of this mapping (by Lemma 3.2). Its image is denoted by $I^{\prime}$, the origin in $(\phi, \psi)$-space.

Definition 3.4 (Branch cut). The branch cut $B$ is the set of points $(x, y) \in \Omega$ such that

$$
\begin{equation*}
B:=\{(x, y) \in \Omega: \phi(x, y)<0 \text { and } \psi(x, y)=0\} \tag{3.14}
\end{equation*}
$$


(a)

(b)

Figure 3.4: Vector field of (3.12) using $\boldsymbol{w}^{+}$, near a nondegenerate singularity $\boldsymbol{p}$, with the branch cut $B$ shown in green: (a) if $g(\boldsymbol{p})>0$ and (b) if $g(\boldsymbol{p})<0$.

The image $B^{\prime}$ of the branch cut is the negative $\phi$-axis in Fig. 3.3. In each of the four quadrants we have stated the value of the $\theta^{+}$field in terms of the standard inverse tangent function. We focus here near a nondegenerate singularity $\boldsymbol{p}$, where the $\phi=0$ and $\psi=0$ contours must cross $\boldsymbol{p}$ transversely, given that $g(\boldsymbol{p}) \neq 0$. The axis-crossings in Fig. 3.3 will have the same topology as these contours if $g(\boldsymbol{p})>0$. The relevant set $B$ in $\Omega_{I}$, near $\boldsymbol{p}$, must therefore have the structure as seen in Fig. 3.4(a). Consider a small circle around $\boldsymbol{p}$ as drawn in Fig. 3.4(a), and indicated via arrows the directions of the vector field $\boldsymbol{w}^{+}$along it. The reasons for these directions stems directly from Fig. 3.3; we need to take the cosine (for the $x$-component) and the sine (for the $y$-component) of the angle field defined therein. While $\boldsymbol{w}^{+}$must vary smoothly along the circle, it exhibits a discontinuity across the branch cut $B$, because the angle has rotated around from $-\pi / 2$ to $+\pi / 2$. Clearly, the same qualitative behavior occurs if $g(\boldsymbol{p})<0$ : in this case we need to consider Fig. 3.3 with the $\psi$-axis flipped (this orientation-reversing case is indeed pictured in Fig. B.1(b)). Once again, it is the $\phi_{-}$axis to which the branch cut $B \in \Omega_{I}$ gets mapped. The intuition of Fig. 3.4 gives us a theoretical issue related to using a vector field to find curves:

Theorem 3.5 (Impossibility of continuous Lyapunov vector field). If there exists at least one nondegenerate singularity $\boldsymbol{p} \in \Omega$, then no nontrivial scalar function $m$ in (3.12) exists such that the right-hand side (i.e., vector field associated with the angle field $\theta^{+}$) is a $\mathrm{C}^{0}$-smooth nonzero vector field in $\Omega_{P}$. The same conclusion holds for vector fields generated from $\theta^{-}$.

Proof. See Appendix D.
3.4. Computational issues of finding foliations. In the previous section, we have outlined a theoretical concern in defining a vector field for computing optimal foliations in two dimensions. We show here related numerical issues which emerge when attempting to compute foliating curves.

First, we remark that generating hyperstreamlines for eigenvectors of tensor-which as seen here are equivalent to $\operatorname{SORF}_{\text {max }}$ and $\operatorname{SORF}_{\text {min }}$ curves-is standard practice [11, 42, 38, 14]. Numerical issues in doing so have been observed previously, and ad hoc remedies proposed:

- In generating trajectories following 'smooth' fields from grid-based data, one suggested approach is to keep checking the direction of the vector field within each cell a trajec-
tory ventures into, and then flip the vector field at the bounding gridpoints to all be in the same direction before interpolating [14, 21].
- In dealing with points at which the eigenvector field is not defined, an approach is to mollify the field by multiplying with a sufficiently smooth field which is zero at such points (e.g., the square of the difference in the two eigenvalues [38, 21]).
Our Theorem 3.5 gives explicit insights into the nature of both these issues. Both ad hoc numerical methods relate to choosing the function $m$ (respectively as $\pm 1$, or a smooth scalar field which is zero at singularities). In either case, actual behavior near the singularities gets blurred by this process

The branch cut near singularities also leads to more subtle - and apparently hitherto unidentified in the literature of following streamlines of tensor fields-issues when performing numerical computations. In Appendix E, we explain why the following occur.

Property 3 (Numerical computation of optimal foliations using vector fields). Suppose we numerically compute a $\mathrm{SORF}_{\max }$ (resp. $\mathrm{SORF}_{\min }$ ) curve by using (3.12) with $m=1$ and the vector field $\boldsymbol{w}^{+}$(resp. $\boldsymbol{w}^{-}$), by allowing the parameter $s$ to evolve in both directions. Then
(a) $\mathrm{SORF}_{\max }$ curves will not cross a one-dimensional part of $B$ vertically, and may also veer along $B$ even though $B$ may not be a genuine $\mathrm{SORF}_{\max }$ curve;
(b) SORF $_{\min }$ curves will not cross a one-dimensional part of $B$ horizontally, and may also veer along $B$ even though $B$ may not be a genuine SORF $_{\min }$ curve .

These problems are akin to branch splitting issues arising when applying curve continuation methods in instances such as bifurcations [13]. Is it possible to choose a function $m$ which is not identically 1 to remove these difficulties? The proof of Theorem 3.5 tells us that the answer is no. Either the branch cut gets moved to a different curve connected to $\boldsymbol{p}$ across which there is a similar discontinuity, or it gets converted to a curve which has spurious fixed points (i.e., a center manifold curve) because $m=0$ on it. In either case, the numerical evaluation will give problems.

Thus, there are several numerical issues in computing foliations using the vector fields $\boldsymbol{w}^{ \pm}$as given in (3.10) and (3.11). Lemma 3.3 suggests a straightfoward alternative method for numerically computing such curves in generic situations, while systematically avoiding all these issues. For the maximizing foliation, let

$$
\begin{aligned}
& \Phi_{-}:=\{(x, y): \phi(x, y)<0 \text { and } \psi(x, y)=0\} \quad \text { and } \\
& \Phi_{+}:=\{(x, y): \phi(x, y)>0 \text { and } \psi(x, y)=0\}
\end{aligned}
$$

these are points mapping to the 'negative $\phi$-axis' and the 'positive $\phi$-axis' (see Figs. 3.3 and B.1), and we also note that $\Phi_{-}=B$. We define on $\Omega_{P} \backslash \Phi_{-}$,

$$
h^{+}(x, y)= \begin{cases}\frac{-\phi(x, y)+\sqrt{\phi^{2}(x, y)+\psi^{2}(x, y)}}{\psi(x, y)} & \text { if } \psi(x, y) \neq 0  \tag{3.15}\\ 0 & \text { if } \psi(x, y)=0 \text { and } \phi(x, y)>0\end{cases}
$$

This is essentially the function $\tan \theta^{+}$as defined in (3.6b), and is $\mathrm{C}^{1}$ in $\Omega_{I} \backslash \Phi_{-}$. The reason for not defining $h^{+}$on $\Phi_{-}$is because the relevant tangent line becomes vertical. Hence we
define its reciprocal, $\mathrm{C}^{1}$ on $\Omega_{I} \backslash \Phi_{+}$, by

$$
y^{+}(x, y):= \begin{cases}\frac{\phi(x, y)+\sqrt{\phi^{2}(x, y)+\psi^{2}(x, y)}}{\psi(x, y)} & \text { if } \psi(x, y) \neq 0  \tag{3.16}\\ 0 & \text { if } \psi(x, y)=0 \text { and } \phi(x, y)<0\end{cases}
$$

The minimizing foliation is associated with the angle field $\theta^{-}$. Thus we define on $\Omega_{I} \backslash \Phi_{+}$,

$$
h^{-}(x, y):= \begin{cases}\frac{-\phi(x, y)-\sqrt{\phi^{2}(x, y)+\psi^{2}(x, y)}}{\psi(x, y)} & \text { if } \psi(x, y) \neq 0  \tag{3.17}\\ 0 & \text { if } \psi(x, y)=0 \text { and } \phi(x, y)<0\end{cases}
$$

which gives the slope field associated with $\theta^{-}$, and on $\Omega_{I} \backslash \Phi_{-}$its reciprocal

$$
y^{-}(x, y):=\left\{\begin{array}{ll}
\frac{\phi(x, y)-\sqrt{\phi^{2}(x, y)+\psi^{2}(x, y)}}{\psi(x, y)} & \text { if } \psi(x, y) \neq 0  \tag{3.18}\\
0 & \text { if } \psi(x, y)=0 \text { and } \phi(x, y)>0
\end{array} .\right.
$$

Property 4 (Foliations as integral curves). Within $\Omega_{I}$, a SORF $_{\max }$ curve can be determined by taking an initial point $\left(x_{0}, y_{0}\right)$ and then numerically following

$$
\begin{equation*}
\frac{d y}{d x}=h^{+}(x, y) \quad \text { if } \quad\left|h^{+}(x, y)\right| \leq 1 \quad \text { and } \quad \frac{d x}{d y}=y^{+}(x, y) \quad \text { if else } \tag{3.19}
\end{equation*}
$$

where we keep switching between the equations depending on the size of $\left|h^{+}\right|$. This generates a sequence $\left(x_{i}, y_{i}\right)$ to numerically approximate an integral curve. Similarly, a $\operatorname{SORF}_{\min }$ curve can be determined in $\Omega_{P}$ as integral curves of

$$
\begin{equation*}
\frac{d y}{d x}=h^{-}(x, y) \quad \text { if } \quad\left|h^{-}(x, y)\right| \leq 1 \quad \text { and } \quad \frac{d x}{d y}=y^{-}(x, y) \quad \text { if else } \tag{3.20}
\end{equation*}
$$

Property 4 is an attractive alternative which avoids issues related to the branch cut and vector field discontinuities. Moreover, it is directly expressed in terms of the functions $\phi$ and $\psi$ via the straightforward definitions of $h^{ \pm}$and $Y^{ \pm}$. The switching between the $d y / d x$ and $d x / d y$ forms avoids the infinite slopes which may result if only one of these forms is used. Thus, we can follow a particular curve as it meanders around $\Omega_{I}$, having vertical and horizontal tangents, and also crossing branch cuts, with no problem.
4. Numerical examples of optimal foliations. We will demonstrate the global optimal foliations for several maps $\boldsymbol{F}$, generated from several applications of discrete maps, and from sampling flows driven by unsteady velocities. The examples include situations which are highly disordered (e.g., maps known to be chaotic under repeated iterations, flows known to possess chaos over infinite times). Moreover, the maps $\boldsymbol{F}$ need not be area-preserving. We will in particular highlight the numerical issues outlined in the previous section by example, and demonstrate how the integral curve approach sidesteps many of these issues. Hence, we focus only on two-dimensional maps, for which we use $(x, y)$ as the standard Cartesian coordinates.

Given its relationship to the FTLE field (2.4), we will in our examples also plot the field

$$
\begin{equation*}
\text { FTLE } *(x, y)=\ln \sqrt{\lambda^{+}(x, y)} \tag{4.1}
\end{equation*}
$$

i.e., the logarithm of the largest local stretching value associated with each point (we do not time-scale ibecause our $F$ may be generated in various ways from maps or flows). In order to retain sufficient resolution to view relevant features in the many subfigures that we present in this Section, we will dispense with axes labels since these are self-evident: $x$ will be the horizontal axis and $y$ the vertical as per standard convention.
4.1. Hénon map. As our first example, consider the Hénon map, which is defined by [22]

$$
\mathfrak{H}(x, y)=\binom{1-a x^{2}+y}{b x}
$$

on $\Omega=\mathbb{R}^{2}$, and where we make the classical parameter choices $a=1.4$ and $b=0.3$. We choose $\boldsymbol{F}$ to be four iterations of the Hénon map, i.e., $\boldsymbol{F}=\mathfrak{H}^{4}$. Fig. 4.1 demonstrates the computed foliations and related graphs. The stretching field $\Lambda^{+}$is first displayed in Fig. 4.1(a). In Fig. 4.1(b), we show the zero contours of $\phi$ and $\psi$. In this case, there are no nice transversalities. Indeed, there are several regions of almost tangencies, and the fact that several of the zero contours almost coincide in the two outer streaks in the figure, indicate that degenerate foliations are to be expected in their vicinity. The 'squashing together' that is occurring here is because we are at an intermediate stage in which initial conditions are gradually collapsing to the Hénon attractor.

The vector fields $\boldsymbol{w}^{ \pm}$, shown in Figs. 4.1(c,d) were computed using (3.10) and (3.11). The discontinuities impact the computation of the SORF curves in (e) and (f). These are obtained by seeding 300 initial locations randomly in the domain, and then computing streamlines generated from (3.12) with $m=1$ in forward, as well as backward, $s$. Since the $\phi$ and $\psi$ fields have large variations at small spatial scales because of the chaotic nature of the map, finding the branch cut $B$ (where where $\psi=0$ and $\phi<0$ ) as obtained from (3.14) requires care. We assess each gridpoint, and color it in (in green) if it has a different sign of $\psi$ in comparison to any of its four nearest neighbors, and the $\phi$ value at this point is negative. The lowermost panel overlays the (green) set $B$ on the SORF curves, indicating why some of the apparent behavior in (e) and (f) is not representative of the true foliation; the center vertical line in (f), for example, occurs because of Property $3(\mathrm{~b})$, while the $\mathrm{SORF}_{\max }$ (resp. SORF min ) curves stop abruptly on $B$ if crossing vertically (resp. horizontally).

On the other hand, Fig. 4.1(b) indicates that the zero contours of $\phi$ and $\psi$ almost coincide on two curves: 'outer' and 'inner' parabolic shapes. These are also identified as part of the branch cut set $B$ because $\psi \approx 0$ and $\phi$ is slightly negative here. These curves are 'almost' a curve of $I$, and we see accumulation of $S^{2} F_{\text {max }}$ curves towards these, indicating-at this level of resolution-potential degeneracy of the foliation. We zoom in to this in Fig. 4.2. In conjunction with the explanations in Fig. B.1, what occurs here is that the inner green line in Fig. 4.2(a) must have a slope field which is $-\pi / 2$ (it is in $\Phi_{-}=B$ with respect to Fig. 4.2), while on the inner pink line it should be $-\pi / 4$ (corresponding to $\Psi_{-}$in Fig. B.1(a)). The extreme closeness of the contours means that a very sharp change in direction must be achieved in a tiny region, which then visually appears as a form of degeneracy.

This example highlights an important computational issue which is very general: even though relevant foliations will exist, in order to resolve them, one needs a spatial resolution which can resolve the spatial changes in the $\phi$ and $\psi$ fields.
(a)


(c)

(e)




(b)

(d)
(b)
(f)
(h)

Figure 4.1: Optimal foliation computations for $\boldsymbol{F}=\mathfrak{H}^{4}:$ (a) the FTLE* field, (b) zero contours of $\phi$ and $\psi,(c)$ vector field $\boldsymbol{w}^{+}$generated from (3.10), (d) vector field $\boldsymbol{w}^{-}$generated from (3.11), (e) SORF $_{\text {max }}$ by implementing vector field in (c), (f) SORF $_{\text {min }}$ by implementing vector field in $(\mathrm{d}),(\mathrm{g}) \mathrm{SORF}_{\text {max }}$ with branch cut (green), (h) SORF min with branch cut (green).


Figure 4.2: Zooming in to an area associated with the map $\boldsymbol{F}=\mathfrak{H}^{4}$ (a) the zero contours of $\phi$ and $\psi,(b)$ the $\operatorname{SORF}_{\text {max }}$, and (c) the $\operatorname{SORF}_{\text {min }}$.
4.2. Double-gyre flow. As an example of when $\boldsymbol{F}$ is generated from a finite-time flow, let us consider the flow map from time $t=0$ to 2 generated from the differential equation

$$
\begin{equation*}
\frac{d}{d t}\binom{x}{y}=\binom{-\pi A \sin [\pi g(x, t)] \cos [\pi y]}{\pi A \cos [\pi g(x, t)] \sin [\pi y] \frac{\partial g}{\partial x}(x, t)} \tag{4.2}
\end{equation*}
$$

in which $g(x, t):=\varepsilon \sin (\omega t) x^{2}+[1-2 \varepsilon \sin (\omega t)] x$ and $\Omega=(0,2) \times(0,1)$. This is the wellstudied double-gyre model [37], but we exclude the boundary of the domain. We use the parameter values $A=1, \omega=2 \pi$ and $\varepsilon=0.1$, and the optimal reduced foliations are demonstrate in Fig. 4.3.

Fig. 4.3(a) is a classical figure in this context: the FTLE* field (if divided by the time-of-flow 2, this would be the highly-studied finite-time Lyapunov exponent field for this flow). Fig. 4.3(b) indicates the $\phi=0$ and $\psi=0$ contours, with their intersections defining $I$. We use the 'standard' $\boldsymbol{w}^{ \pm}$unit versions, (3.10) and (3.11), to generate the vector fields in (c) and (d), and the corresponding SORFs are determined in (e) and (f). Figs. 4.3(g) and (h) overlay the branch cuts (green), which are parts of the green curves in Fig. 4.3(b) at which $\phi<0$.

As expected, the $S O R F_{\text {max }}$ curves fail to cross the branch cut vertically, as do the $\operatorname{SORF}_{\text {min }}$ curves horizontally. Moreover, foliation curves which do get pushed in towards the branch cuts tend to meander along them, giving an impact of spurious accumulations. We zoom in towards one of these regions in Fig. 4.4; the $\mathrm{SORF}_{\text {max }}$ curves requirements of having slopes $-\pi / 4$ (resp. $+\pi / 2$ ) on $\Phi_{-}$(resp. $\Phi_{+}$) result in abrupt curving. The accumulation is not exactly to $\Psi_{-}$, but rather to a curve which is very close, as seen in Fig. 4.4(b). Thus, it is not true that there is a one-dimensional part of the isotropic set $I$ along here. The geometric insights of the previous sections allows us to understand and interpret these issues, while appreciating how resolution may give misleading visual cues.

In Fig. 4.5, we zoom in to two difference locations, chosen by zeroeing in to two different intersection points of the zero $\phi$ and $\psi$-contours. The top panels illustrate the $\mathrm{SORF}_{\max }$ (left) and the $\operatorname{SORF}_{\min }$ (right) curves at the same location. The theory related to 1-pronged intruding points is well-demonstrated, with there being two such points adjacent to each other.

(a)
(b)


(c)


(e)
(f)


(g)
(h)

Figure 4.3: Optimal foliation computations for the double-gyre flow: (a) the FTLE* field, (b) zero contours of $\phi$ and $\psi$, (c) vector field $\boldsymbol{w}^{+}$generated from (3.10), (d) vector field $\boldsymbol{w}^{-}$ generated from (3.11), (e) $\operatorname{SORF}_{\text {max }}$ by implementing vector field in (c), (f) $\operatorname{SORF}_{\text {min }}$ by implementing vector field in (d), (g) SORF $_{\text {max }}$ with branch cut (green), (g) SORF $\mathrm{min}_{\text {m }}$ with branch cut.


Figure 4.4: Zooming in to near an 'accumulating' SORF $_{\max }$ from Fig. 4.3: (a) the relevant zero contours of $\phi$ and $\psi$, and (b) the $\operatorname{SORF}_{\text {max }}$.
(a)


(b)

(d)

Figure 4.5: Zooming in to the SORF $_{\text {max }}$ (left) and $\operatorname{SORF}_{\text {min }}$ (right) in the double-gyre. The top and bottom panels correspond to different locations, respectively near two adjacent wedge points, and a trisector point. The branch cut is shown in green. Compare to Fig. 3.1 and Property 1.


Figure 4.6: Optimal foliation computations for the Chirikov map $\boldsymbol{F}=\mathfrak{C}_{2}^{4}$ : (a) the FTLE* field (b) zero contours of $\phi$ and $\psi$, (c) $\operatorname{SORF}_{\text {max }}$ with branch cut (green), (d) $\mathrm{SORF}_{\text {min }}$ with branch cut (green).

The two orthogonal families 'reverse' the locations of the singularities for the maximizing and minimizing foliations, and the branch cut (green) forms vertical/horizontal barriers as appropriate. In contrast, the bottom figures are of a 3 -pronged trisector; again, the numerics validate the theory.
4.3. Chirikov map. The Chirikov (also called 'standard') map is defined on the doublyperiodic domain $\Omega=[0,2 \pi) \times[0,2 \pi)$ by [9]

$$
\mathfrak{C}_{k}(x, y)=\left(\begin{array}{cc}
x+y+k \sin x & (\bmod 2 \pi) \\
y+k \sin x & (\bmod 2 \pi)
\end{array}\right) .
$$

We choose $\boldsymbol{F}=\mathfrak{C}_{k}^{n}$, that is, $n$ iterations of the Chirikov map for a given value of the parameter $k$. Increasing $k$ increases the disorder of the map, as does having $n$ large. (The map is a classical example of chaos, with $\Omega$ consisting of quasiperiodic islands in a chaotic sea, where 'chaos/chaotic' must be understood in the limit $n \rightarrow \infty$.) In more disorderly situations, increasingly fine resolution is required to reveal the structures that we have defined.

Relevant computations for $k=2$ and $n=4$ are shown in Fig. 4.6. There are significant regions where the behavior is quite orderly. There is 'greater disorder' in the region foliated with large values of FTLE* in (a) -indeed, this region is associated with the 'chaotic sea' when the map is iterated many more times - with the outer parts of low FTLE* being associated


Figure 4.7: A degenerate singularity of the map $\boldsymbol{F}=\mathfrak{C}_{1}^{2}$, shown zoomed-in: (a) the zero contours of $\phi$ and $\psi$, (b) $\mathrm{SORF}_{\text {max }}$, and (c) $\mathrm{SORF}_{\text {min }}$.
with quasiperiodic islands and hence order. All features mentioned in previous examples are reiterated in the pictures. Moreover, the $\mathrm{SORF}_{\text {min }}$ foliation somewhat mirrors the structure expected from classical Poincaré section numerics.

If we instead consider $k=1$ and $n=2$, an interesting degenerate singularity (corresponding to the $\psi=0$ contour crossing exactly a saddle point of $\phi$ ) is displayed in Fig. 4.7. The singularity in the SORF $_{\text {max }}$ foliation (b) appears like a degenerate form of a trisector, if thinking in terms of curves coming from above. However, if viewed in terms of curves coming in from below, it appears as a wedge with a sharp (triangular) end. The SORF $_{\text {min }}$ conforms to this, having elements of both a wedge and trisector, as well. (The numerical issue of SORF $\min$ not crossing $B$ horizontally is displayed in Fig. 4.7(c); in reality, the true SORF $_{\min }$ curves should connect smoothly across.)

Next, we demonstrate in Fig. 4.8, using $\boldsymbol{F}=\mathfrak{C}_{2}^{2}$, the efficacy of using the integral-curve forms (3.19) and (3.20) of the foliations, rather than using a vector field. The FTLE* field in Fig. 4.8(a) has several sharp ridges; these are well captured by locations where the $\phi$ and $\psi$ zero-contours in Fig. 4.8(b) coincide. The SORF $_{\max / \min }$ foliations in (b) and (c) are computed respectively using the vector fields $\boldsymbol{w}^{ \pm}$as in previous situations, and exhibit the usual issues when crossing $B$. In contrast, the lower row is generated by using the integralcurve forms (3.19) and (3.20), where we have once again started from 300 random initial conditions. For each initial condition $\left(x_{1}, y_{1}\right)$, we define the next point $\left(x_{2}, y_{2}\right)$ on a $\operatorname{SORF}_{\text {max }}$ curve by $x_{2}=x_{1}+Y^{+}\left(x_{1}, y_{1}\right) \delta y$ where $\delta y>0$ is the spatial resolution in the $y$-direction, and $d x / d y$ is based on (3.19). Similarly, $y_{2}=y_{1}+h^{+}\left(x_{1}, y_{1}\right) \delta x$ using (3.19), and where $\delta x>0$ is the resolution chosen in $x$-direction. This initializes the process. Next, we check the value of $h_{+}\left(x_{2}, y_{2}\right)$, thereby deciding which of the equations in (3.19) to implement. If the $d y / d x$ equation, we take $x_{3}=x_{2}+\operatorname{sign}\left(x_{2}-x_{1}\right) \delta x$, and thus find $y_{3}$ using the ODE solver. Having now obtained $\left(x_{3}, y_{3}\right)$, we again use the last two points to make decisions on which of the two equations to use, and continue in this fashion for a predetermined number of steps. Next, we go back to $\left(x_{1}, y_{1}\right)$ and now set $x_{2}=x_{1}-\psi^{+}\left(x_{1}, y_{1}\right) \delta y$ and $y_{2}=y_{1}-h^{+}\left(x_{1}, y_{1}\right) \delta y$, thereby going in the opposite direction. Having initiated this process, we can then continue this curve using the same continuation scheme.
(a)


(b)

(c)


(d)

Figure 4.8: Comparison between using the integral-curve forms (3.19) and (3.20) and the vector field forms for $\boldsymbol{F}=\mathfrak{C}_{2}^{2}$ : (a) $\ln \Lambda^{+}$field, (b) zero contours of $\phi$ and $\psi$, (c) $\mathrm{SORF}_{\text {max }}$ using the vector field (3.10), (d) $\mathrm{SORF}_{\text {min }}$ using the vector field (3.11), (e) $\mathrm{SORF}_{\text {max }}$ using the integral curve form (3.19), and (f) SORF $_{\text {min }}$ using the form (3.20).

The SORF $_{\text {min }}$ are obtained similarly, using the two equations in (3.20). There is sensitivity in the process to locations where $\phi$ and $\psi$ change rapidly (they are each of the order $10^{5}$ in this situation), and in particular where zeros are near. The resolution scales $\delta x$ and $\delta y$ need to be reduced sufficiently to not capture spurious effects. Notice that there are no branchcut problems in the resulting foliations obtained using the integral-curve approach, since we do not have to worry about a discontinuity in a vector field. Neither are there any abrupt stopping of curves.
5. Concluding remarks. Using the Cauchy-Green tensor for understanding locally optimal stretching is well-established, as is using the hyperstreamlines generated from its associated eigenvector fields. Despite many articles computing these hyperstreamlines and making various claims related to them, genuinely interpreting them in terms of a global optimization problem has remained problematic. In this paper, we formulate a global optimization problem based on restricted foliations, and show how the hyperstreamlines are a solution to this optimization problem. Moreover, we focus on the two-dimensional situation and derive (using elementary methods) conditions on singularities of the foliation, and numerical artefacts that can emerge when attempting to compute foliation curves.

We expect these results to help researchers interpret, and improve, numerical calculations related to optimal stretching paradigms in finite-time situations. In particular, misinterpretations of numerics can be mitigated via the understandings presented here. In two-dimensions, regions of high sensitivity towards spatial resolutions are also identifiable in terms of the near-zero sets of the $\phi$ and $\psi$ functions.

We wish to highlight from our numerical results the role of SORF $_{\text {min }}$ restricted foliations as being effective demarcators of complication flow regimes. These curves-observable for example in blue in Figs. 4.1, 4.3, 4.6 and 4.8 -indicate curves along which there is minimal stretching. Consequently, there is maximal stretching in the orthogonal direction to these curves. This indicates that the $\mathrm{SORF}_{\text {min }}$ curves are barriers in some senses: disks of initial conditions positioned on such a curve experience sharp stretching orthogonal to them. That is, initial conditions on one side of such a curve get separated quickly from those on the other side, with the curve positioned optimally to maximize the separation. Our methodology enables this intuitive idea to be put into a global optimizing foliation framework. Looking at this another way, the dense regions of the $\mathrm{SORF}_{\min }$ (blue) foliations in Figs. 4.1, 4.3, 4.6 and 4.8 are reminiscent of separation curves which attempt to demarcate chaotic from regular regions. We emphasize, though, that 'chaotic' has no proper meaning in the finite-time context since it must be understood in terms of infinite-time limits or repetitive application of maps; in this case, the separation one may try to obtain is between more 'disorderly' and 'orderly' regions. The ambiguity of defining these is reflected in the Figures, in which the $\mathrm{SORF}_{\text {min }}$ foliation nonetheless identifies coherence-related topological structures in $\Omega$ which are strongly influenced by the nature of the singularities in the foliation.

We observe that the interaction of $\phi=0$ and $\psi=0$ level sets as seen in Fig. 4.1(b) bear a striking resemblance to figures regarding zero angle between stable and unstable foliations of Lyapunov vectors such as in Fig. 1 for the Hénon map from [26] that was part of a search for primary heteroclinic tangencies when developing symbolic dynamic generating partitions of the Henon map, $[17,7,6,10]$. Indeed this analysis likely bears a relationship, in that in an infinite time limit, the Lyapunov vectors suggested come to the same point as those much earlier stories underlying the topological dynamics of smooth dynamical systems. What is clear in the finite time discussion here is that when we see a coincidence between the stretching and folding, that in successively longer time windows, these properties repeat in progressively smaller regions. As suggested by Fig. 4.1, e.g. (h), any point of tangency would in turn be infinitely repeated in the long time limit. The perspective of this current work may further understanding of what has always been the intricate topic of why and how hyperbolicity is lost in nonuniformly hyperbolic systems wherein seemingly paradoxically, errors can grow along
the directions related to stable manifolds, such as highlighted by Fig. 5 in [25].
Acknowledgements: Conversations with Thomas Leistner, who pointed out interesting connections to the older pure mathematical literature, is gratefully acknowledged. SB acknowledges with thanks partial support from the Australian Research Council via grant DP200101764. EB acknowledges with thanks the Army Research Office, the Naval Research Office, the Air Force Office of Scientific Research, the National Institutes of Health, and the Defense Advanced Research Projects Agency.

Appendix A. Hyperbolic Lagrangian coherent structures as optimizers.
In Section 2.2, we discuss that in a seminal work [19] which subsequently led to a wellaccepted theory of variational Lagrangian coherent structures [20, 15, 30], the surfaces $\mathcal{S}$ which are everywhere orthogonal to the maximal eigenvector $\boldsymbol{w}_{n}$ (where it is a well-defined direction) are those from which the "surfaces from which there is maximal repulsion " [19] are selected. We pointed out that this is a different approach from ours: we seek a global optimization problem for which the relevant hyperstreamlines are solutions, while in this approach [19] a selection of surfaces from $\mathcal{S}$ is undertaken, such that such a surface is locally maximally repelling in comparison to nearby surfaces in $\mathcal{S}$. For completeness, we discuss here the process of selecting from $\mathcal{S}$, as described in a variety of papers [19, 20, 15, 30].

In [19], the most repelling 'hyperbolic (variational) Lagrangian Coherent Structures' are selected from $\mathcal{S}$ using a classical (Hessian) maximality conditions [19, Thereom 7(ii)2]. This is strictly pointwise, rather than using a definition of repelling from a surface. The pointwise tests in the selection process from $\mathcal{S}$ is further amplified in the discussion of the numerics (Algorithm 7.2 [19]).

Indeed, the ambiguity of determining which surfaces to extract from $\mathcal{S}$ based on pointwise conditions is reflected in several follow-up papers. In one paper in two dimensions $(n=2)$ [15], noting now that the 'surfaces' $\mathcal{S}$ are now curves everywhere orthogonal to $\boldsymbol{w}^{+}=\boldsymbol{w}_{2}$, the authors suggest using the average of $\lambda^{+}=\lambda_{2}$ along curves in $\mathcal{S}$ as an effective measure of "repelling from the surface." There is still ambiguity here in thinking of this as a maximal repelling problem: is there any restriction on how long a segment one can choose? In yet another paper, also confined to two dimensions, a computational toolbox LCSTool [30] is introduced; in this, the selection from $\mathcal{S}$ is done quite differently. First, the point $\boldsymbol{p}$ in the domain in which $\lambda^{+}$is maximal is chosen, and the curve from $\mathcal{S}$ which goes across this is selected as a hyperbolic Lagrangian Coherent Structure [30]. It is not at all clear how this corresponds to a curve from which there is maximal repulsion since, once again, the repulsion from the global object (in this case a curve) is not defined, but rather pointwise measures are used. Thus, while $\boldsymbol{p}$ is the point which has the most local repulsion, there is no reason to expect the curve passing through $\boldsymbol{p}$ to be the most repelling curve. Then, curves from a neighborhood of the selected curve are excluded from the domain, and the point in the remaining domain which has the largest $\lambda^{+}$value is chosen, and the process repeated.

Thus, in the computational selection of 'hyperbolic Lagrangian coherent structures' as the "locally most repelling surfaces" [19, 20], the true global optimization problem is somewhat unclear. Rather, a definition which automatically restricts to the set $\mathcal{S}$ is initially used [19, 20]. Then, the selection from among $\mathcal{S}$ is done in various ad hoc ways: using a Hessian argument pointwise [19], taking an average of $\lambda^{+}$along a curve [15], or picking a point with maximal
$\lambda^{+}$and taking curves passing through that [30]. In all these cases, there is ambiguity in how the local (pointwise) stretching behavior and the associated hyperstreamlines are related to a genuine optimization problem for co-dimension-1 surfaces.

## Appendix B. Equivalence of two-dimensional formulation.

Here, we establish proofs of the equivalence results in Section 3.1 using elementary methods.

First, we consider Lemma 3.1. Given a general point $(x, y) \in \Omega_{P}$, let the direction $\hat{\boldsymbol{n}}$ in (3.3) be associated with an angle $\theta \in[-\pi / 2, \pi / 2)$. The local stretching (2.1) is

$$
\Lambda(x, y, \theta)=\sqrt{\left(u_{x} \cos \theta+u_{y} \sin \theta\right)^{2}+\left(v_{x} \cos \theta+v_{y} \sin \theta\right)^{2}} .
$$

where the $(x, y)$-dependence on $u_{x}, u_{y}, v_{x}$ and $v_{y}$ has been omitted from the right-hand side for brevity. Hence,

$$
\Lambda^{2}=\frac{u_{x}^{2}+v_{x}^{2}-u_{y}^{2}-v_{y}^{2}}{2} \cos 2 \theta+\left(u_{x} u_{y}+v_{x} v_{y}\right) \sin 2 \theta+\frac{u_{y}^{2}+v_{y}^{2}+u_{x}^{2}+v_{x}^{2}}{2} .
$$

Using the definitions for the functions $\phi$ and $\psi$ from (3.1) and (3.2) yields (3.4) as desired.
We next address the characterization in Lemma 3.3 using elementary means. We first establish equivalence of the two formulations given, and then later show why these directions are associated with optimal local stretching. Beginning with (3.6a), assuming for now that both $\phi$ and $\psi$ are not zero, we use the double-angle formula to obtain

$$
\frac{2 \tan \theta^{+}}{1-\tan ^{2} \theta^{+}}=\tan 2 \theta^{+}=\frac{\psi}{\phi} .
$$

Solving the quadratic for $\tan \theta^{+}$, we see that

$$
\begin{equation*}
\tan \theta^{+}=\frac{-1 \pm \sqrt{(\psi / \phi)^{2}+1}}{\psi / \phi}=\frac{-\phi \pm \sqrt{\phi^{2}+\psi^{2}}}{\psi} \tag{B.1}
\end{equation*}
$$

We now need to choose the sign in this expression, bearing in mind the usage of the fourquadrant inverse tangent as used in (3.6a). The four quadrants here are in the $(\phi, \psi)$-space, which is indicated in Fig. B.1(a). If $\phi>0$ and $\psi>0$, this implies that $2 \theta^{+}$is in the first quadrant, and thus so is $\theta^{+}$. This means that $\tan \theta^{+}>0$, and consequently the positive sign must be chosen. If $\phi>0$ and $\psi<0,2 \theta^{+}$is in fourth quadrant, or $2 \theta^{+} \in(-\pi / 2,0)$. Thus, $\tan \theta^{+}<0$, and so the positive sign must be chosen in (B.1) to ensure that the division by $\psi<0$ leads to an eventual negative sign. Next, if $\phi<0$ and $\psi>0,2 \theta^{+} \in(\pi / 2, \pi)$, and $\theta^{+} \in(\pi / 4, \pi / 2)$, leading to $\tan \theta^{+}>0$ and the necessity of choosing the positive sign in (B.1). Finally, if $\phi<0$ and $\psi<0,2 \theta^{+} \in(-\pi,-\pi / 2)$ and $\theta^{+} \in(-\pi / 2,-\pi / 4)$, and thus $\tan \theta^{+}<0$ and the positive sign in the numerator of (B.1) must be chosen. Thus, all cases lead to a positive sign, and so

$$
\tan \theta^{+}=\frac{-\phi+\sqrt{\phi^{2}+\psi^{2}}}{\psi}
$$

which is (3.6b). Next, we argue that this expression works even if one or the other of $\phi$ or $\psi$ is zero. The arguments to follow are equivalent to considering the four emanating axes


Figure B.1: SORF $_{\text {max }}$ near a nondegenerate singularity: (a) Value of $\theta^{+} \in[-\pi / 2, \pi / 2)$ in $(\phi, \psi)$-space using (3.6a), (b) as in (a), but shown in a left-hand system, (c) and (d) qualitative slope fields for (a) and (b); (e) 1-pronged 'wedge' associated with the structure (c); (f) 3pronged 'trisector' associated with the structure (d); (g) wedge when axes are tilted; (h) trisector when axes are tilted. Compare to Fig. 3.1 and Property 1.
in Fig. B.1(a). If $\phi=0$ and $\psi \neq 0$, (3.6a) tells us that $2 \theta^{+}=(\pi / 2) \operatorname{sign}(\psi)$ and thus $\tan \theta^{+}=\tan (\pi / 4) \operatorname{sign}(\psi)=\operatorname{sign}(\psi)$. This is consistent with what (3.6b) gives when $\phi=0$ is inserted. If $\psi=0$ and $\phi \neq 0$, (3.6a), which tells us that $2 \theta^{+}=-\pi$ if $\phi<0$, or $2 \theta^{+}=0$ if $\phi>0$. Thus if $\psi=0, \theta^{+}=-\pi / 2$ if $\phi<0$, and $\theta^{+}=0$ if $\phi>0$. This verifies that (3.6b) is equivalent to (3.6a) in $\Omega_{P}$.

Now, $\theta^{-}$in (3.7a) is defined specifically to be orthogonal to $\theta^{+}$. There is only one angle in $[-\pi / 2, \pi / 2)$ which obeys this condition. It is straightforward to verify from (3.6b) and (3.7b) that

$$
\left(\tan \theta^{+}\right)\left(\tan \theta^{-}\right)=-1
$$

in $\Omega_{P}$. Thus, $\theta^{-}$as defined in (3.7b) is at right-angles to $\theta^{+}$as defined in (3.6b), which has been established to be equivalent to (3.6a).

Next, we need to establish that the angles $\theta^{ \pm}$correspond to local optimizers of the local stretching. Using (3.4) associated with a general angle $\theta_{f}$ associated with a restricted foliation $f$, we can write

$$
\begin{align*}
\Lambda^{2} & =\sqrt{\phi^{2}+\psi^{2}}\left[\frac{\phi}{\sqrt{\phi^{2}+\psi^{2}}} \cos 2 \theta_{f}+\frac{\psi}{\sqrt{\phi^{2}+\psi^{2}}} \sin 2 \theta_{f}\right]+\frac{|\boldsymbol{\nabla} u|^{2}+|\boldsymbol{\nabla} v|^{2}}{2} \\
& =\sqrt{\phi^{2}+\psi^{2}}\left[\cos 2 \theta^{+} \cos 2 \theta+\sin 2 \theta^{+} \sin 2 \theta_{f}\right]+\frac{|\nabla u|^{2}+|\nabla v|^{2}}{2} \\
& =\sqrt{\phi^{2}+\psi^{2}} \cos \left[2\left(\theta^{+}-\theta_{f}\right)\right]+\frac{|\boldsymbol{\nabla} u|^{2}+|\boldsymbol{\nabla} v|^{2}}{2} \tag{B.2}
\end{align*}
$$

in which $\theta^{+}=\theta^{+}(x, y)$ satisfies

$$
\begin{equation*}
\cos 2 \theta^{+}=\frac{\phi}{\sqrt{\phi^{2}+\psi^{2}}} \quad \text { and } \quad \sin 2 \theta^{+}=\frac{\psi}{\sqrt{\phi^{2}+\psi^{2}}} \tag{B.3}
\end{equation*}
$$

Thus, $\tan 2 \theta^{+}=\psi / \phi$. If applying the inverse tangent to determine $2 \theta^{+}$from this, we need to take the two equations (B.3) into account in choosing the correct branch. This clearly depends on the signs of $\phi$ and $\psi$, which is automatically dealt with if the four-quadrant inverse tangent is used. Consequently, (B.3) implies that

$$
\theta^{+}(x, y)=\frac{1}{2} \tan ^{-1}(\psi(x, y), \phi(x, y))
$$

which is chosen modulo $\pi$ because of the premultiplier of $1 / 2$ (the four-quandrant inverse tangent is modulo $2 \pi$ ). Thus, $\theta^{+}$as defined here is identical to that given in (3.6a), which by Lemma 3.3 is equivalent to (3.6b). The establishment of $\theta^{-}$as the stretching minimizer is analogous, and will be skipped.

## Appendix C. Singularity classification in two dimensions.

This section provides explanations for the nondegenerate singularity classification of Property 1. Given the transverse intersection of the $\phi=0$ and $\psi=0$ contours at a singularity $\boldsymbol{p}$, we examine nearby contours not in standard $(x, y)$-space, but in $(\phi, \psi)$-space, in which $\boldsymbol{p}$ is at the origin. The angle fields $\theta^{ \pm}$are the defining characteristics of the foliation, and thus we


Figure C.1: SORF $_{\text {max }}$ near $\boldsymbol{p}$ when transversality is relaxed: (a), (b) and (c) show different possibilities for axes to intersect, and the corresponding SORF $_{\text {max }}$ topologies are illustrated in Fig. 3.2.
show in Fig. B.1(a) a schematic of the maximizing angle field $\theta^{+}$. A nonstandard labelling of the $\phi$ and $\psi$ axes is used here because the relative orientations of the positive axes $\phi_{+}$and $\psi_{+}$ (the directions in which $\phi>0$ and $\psi>0$ resp.) and negative axes $\phi_{-}$and $\psi_{-}$is related to the sign of the orientation function-given in (3.9) —at p. Fig. B.1(a) corresponds to a positive orientation. The slope fields and expressions indicated are based on the four-quadrant inverse tangent (3.6a), expressed in terms of the regular inverse tangent in each quadrant. We also express the values of $\theta^{+}$on each of the axes in Figs. B.1(a), along which $\theta^{+}$is seen to be constant.

In Figs. B.1(c), we indicate the angle field $\theta^{+}$by drawing tiny lines which have the relevant slope. What happens when we 'connect these lines' to form a foliation is shown underneath in Figs. B.1(e). The foliation bends around the origin (shown as the blue point p), effectively rotating around it by $\pi$. However, it must be cautioned that while Fig. B.1(e) seems to indicate that the fracture ray lies along $\phi_{+}$, this is in general not the case. The angle fields shown in Figs. B.1(c) and (e) display directions in physical ( $\Omega$ ) space, in which the $\phi=0$ and $\psi=0$ contours intersect in some slanted way. We show one possibility in Fig. B.1(g), in which the fracture ray will be approximately from the northwest. We identify $\boldsymbol{p}$ in this case a wedge or a 1-pronged singularity. The nearby SORF $_{\max }$ curves rotate by $\pi$ around it.

In the right-hand panels of Fig. B. 1 we examine the other possibility of $g(\boldsymbol{p})<0$. This is achieved in Fig. B.1(b) by simply flipping the $\psi_{-}$and $\psi_{+}$axes, and retaining the information that we have already determined in Fig. B.1(a). The corresponding slope field is displayed in Fig. B.1(d). The fracture ray (also along the $\phi_{+}$-axis in this case) now separates out curves coming from the right, rather than causing them to turn around the origin. Fig. B.1(f) demonstrates this behavior, obtained by connecting the angle fields into curves. There are two other fracture rays generated by this process of separation, because curves in the $\phi_{-}$ region are forced to rotate away from the origin without approaching it. Fig. B.1(h) is an orientation-preserving rotation of the axes in Fig. B.1(f), which highlights that the directions
of the three fracture rays are based on the orientations of the axes in physical space. Based on the topology of the foliation, we thus have a trisector or 3-pronged singularity in the case of negative orientation.

Suppose next that the nondegeneracy of $\boldsymbol{p}$ is relaxed mildly by allowing the $\phi=0$ and $\psi=0$ contours (both still considered to be one-dimensional) to intersect tangentially at $\boldsymbol{p}$. To achieve this, imagine bending the $\psi$-axis in Figs. B.1(a) and (c) so that it becomes tangential to the $\phi$-axis, but the axes still cross each other. This degenerate situation is shown in Fig.C.1(a), and we note that the orientation remains positive despite the tangency. Connecting the angle field lines gives the relevant topological structure of Fig. 3.2(a). The topology is very close to the nondegenerate intruding point, but there is an accumulation of curves towards the fracture ray from one side. It is easy to verify (not shown) that there is no change in this topology if the tangentiality shown in Fig. C.1(a) goes in the other direction, with $\psi_{+}$becoming tangential to $\phi_{+}$and $\psi_{-}$to $\phi_{-}$. Fig. C.1(b) examines the impact on the degenerate negative-oriented situation; Fig. 3.2(b) indicates that the fracture ray acquires a similar one-sided accumulation effect, while the remainder of the portrait remains essentially as it was. So this is a degenerate separation point. Finally, in Fig. C.1(c) we consider the case where the tangentiality is such that the $\phi$ - and $\psi$-axes do not cross one another. In this case, drawing connecting curves reveals that the topology is a combination of degenerate intruding and separating points, and is illustrated in Fig. 3.2(c). Testing the other possibilities (interchanging the $\psi_{-}$and $\psi_{+}$axes locations, and doing the same analysis with them below the $\phi$-axis) yields no new topologies. One way to rationalize this is that the relative (degenerate) orientation between the negative axes and that between the positive axes is in this case exactly opposite; one is as if there were a positive orientation, while the other is as if it were negative.

## Appendix D. Proof of Theorem 3.5.

We have established via Fig. 3.4 that if there exists a nondegenerate singularity $\boldsymbol{p}$, then the vector field $\boldsymbol{w}^{+}$as chosen in (3.10) is not continuous across the branch cut $B$. However, attempting to choose a vector field associated with the angle field $\theta^{+}$is ambiguous, as is reflected in the presence of the arbitrary function $m$ in (3.12). The nonuniqueness is equivalent to the potential of scaling Lyapunov vectors in a nonuniform way in $\Omega \backslash I$, by multiplying by a nonzero scalar. The question is: is it possible to remove the discontinuity that (3.10) has across $B$ by choosing a scaling function $m$ ?

From Fig. 3.4, we argue that the answer is no. Imagine going around the black dashed curve, $C$, and attempting to have $\boldsymbol{w}^{+}$be continuous while doing so. Since $\boldsymbol{w}^{+}$has a jump discontinuity across $B$, it will therefore be necessary to choose $m$ to have the opposite jump discontinuity for $m \boldsymbol{w}^{+}$to be smooth. So $m$ must jump from +1 to -1 in a certain direction of crossing. However, since $\boldsymbol{w}^{+}$is continuous on $C \backslash B$, to retain this continuity must also remain continuous along $C \backslash B$. This implies that $m$ must cross zero at some point in $C \backslash B$. Doing so would render the Lyapunov vector $\boldsymbol{w}^{+}$invalid. We have therefore established Theorem 3.5 using elementary geometric means. We remark that this theorem is analogous to the classical "hairy ball" theorem due to Poincaré [33].

## Appendix E. Branch cut effects on computations.

If $\boldsymbol{p}$ is a nondegenerate singularity, then the vector field of (3.12) with $m=1$ and the choice of the positive $\operatorname{sign}\left(\mathrm{SORF}_{\text {max }}\right)$ will locally have the behavior as shown in Fig. 3.4. Now, in
general, in finding a $\operatorname{SORF}_{\text {max }}$ which passes through $\left(x_{0}, y_{0}\right)$, we can implement (3.12) for the choice of $m=1$, in both directions (increasing and decreasing $s$ ), thereby obtaining the curve which crosses the point. An equivalent viewpoint is that we implement (3.12) with $m=1$, and $s>0$, and then implement it with $m=-1$ while retaining $s>0$.

If using (3.12) with $m=+1$ (globally) and $\boldsymbol{w}^{+}$in (3.10) to generate a SORF ${ }_{\text {max }}$ curve, the vector field in Fig. 3.4(a) must be followed. However, it is clear that anything approaching the branch cut $B$ gets pushed away in the vertical direction. Thus, $\operatorname{SORF}_{\max }$ curves near $B$ will in general be difficult to find.

The solution appears to be to set $m=-1$, which reverses the vector field. However, this is essentially the diagram in Fig. 3.4(b), corresponding to a $\boldsymbol{p}$ with negative orientation function. This is of course equivalent to implementing (3.12) with $m=+1$ but in the $s<0$ direction. Curves coming in to $B$ now get stopped abruptly, because the vector field on the other side of $B$ directly opposes the vertical motion. Thus, curves will not cross $B$ vertically. However, since any incoming curve will in general have a vector field component tangential to $B$, this will cause a veering along the curve $B$. The curve will continue along $B$, because the vector field pushes in on to $B$ vertically, preventing departure from it. Thus when numerically finding SORF $_{\text {max }}$ curves, curves which appear to tangentially approach the branch cut $B$ will be seen. These curves are not real SORF $_{\max }$ curves because, as is clear from Fig. 3.4, the actual vector field is not necessarily tangential to $B$. That is, the branch cut is not necessarily a streamline of the direction field $\theta^{+}$.

A similar analysis (not shown) indicates that if using the expression (3.11) for $\boldsymbol{w}^{-}$to generate SORF $_{\text {min }}$ curves, then these curves will not cross $B$ horizontally, and also have the potential for tangentially approaching $B$ in a spurious way. Notice moreover that, while we have discussed the branch cut locally near $\boldsymbol{p}$, these objects extend through $\Omega_{I}$, potentially connecting with several singularities.

Finally, suppose there are parts of $B$ that are two-dimensional regions. In such regions, Fig. B.1(a) indicates that the angle field $\theta^{+}$is vertical. Consequently, $\theta^{-}$is horizontal everywhere. However, numerical issues as above will occur when crossing the one-dimensional boundary $\bar{B} \backslash B$, due to the inevitable issue of the reversal of the vector field along at least one part of this boundary.

## REFERENCES

[1] C. C. Adams, The knot book: An elementary introduction to the mathematical theory of knots, (No Title), (2005).
[2] S. Balasuriya, Stochastic sensitivity: a computable Lagrangian uncertainty measure for unsteady flows, SIAM Review, 62 (2020), pp. 781-816.
[3] S. Balasuriya, n. Ouellette, and I. Rypina, Generalized Lagrangian coherent structures, Physica D, 372 (2018), pp. 31-51.
[4] G. Band and P. Boyland, The Burau estimate for the entropy of a braid, Algebraic \& Geometric Topology, 7 (2007), pp. 1345-1378.
[5] R. Bellman, Introduction to Matrix Analysis, McGraw-Hill, New York, 1960.
[6] E. M. Bollt, T. Stanford, Y.-C. Lai, and K. Życzkowski, Validity of threshold-crossing analysis of symbolic dynamics from chaotic time series, Physical Review Letters, 85 (2000), p. 3524.
[7] E. M. Bollt, T. Stanford, Y.-C. Lai, and K. Życzkowski, What symbolic dynamics do we get with a misplaced partition?: On the validity of threshold crossings analysis of chaotic time-series, Physica

D: Nonlinear Phenomena, 154 (2001), pp. 259-286.
[8] G. Chen, D. Palke, Z. Lin, H. Yeh, P. Vincent, R. Laramee, and E. Zhang, Asymmetric tensor field visualization for surfaces, IEEE Transactions on Visualization and Computer Graphics, 17 (2011), pp. 1979-1988.
[9] B. Chirikov, A universal instability of many-dimensional oscillator systems, Physics Reports, 52 (1979), pp. 263-379.
[10] F. Christiansen and A. Politi, Guidelines for the construction of a generating partition in the standard map, Physica D: Nonlinear Phenomena, 109 (1997), pp. 32-41.
[11] T. Delmarcelle and L. Hesselink, The topology of symmetric, second-order tensor fields, in Proceedings Visualization '94, Washington DC, October 1994, IEEE, pp. 140-147.
[12] T. Delmarcelle and L. Hesselink, The topology of symmetric, second-order tensor fields, in Proceedings Visualization'94, IEEE, 1994, pp. 140-147.
[13] E. J. Doedel, T. F. Fairgrieve, B. Sandstede, A. R. Champneys, Y. A. Kuznetsov, and X. Wang, Auto-07p: Continuation and bifurcation software for ordinary differential equations, tech. report, http://indy.cs.concordia.ca/auto/, 2007.
[14] M. Farazmand, D. Blazevski, and G. Haller, Shearless transport barriers in unsteady twodimensional flows and maps, Physica D, 278-279 (2014), pp. 44-57.
[15] M. Farazmand and G. Haller, Computing Lagrangian coherent structures from their variational theory, Chaos, 22 (2012), p. 013128.
[16] F. T. Farrell and S. K. Roushon, The whitehead groups of braid groups vanish, International Mathematics Research Notices, 2000 (2000), pp. 515-526.
[17] P. Grassberger, H. Kantz, and U. Moenig, On the symbolic dynamics of the Hénon map, Journal of Physics A: Mathematical and General, 22 (1989), p. 5217.
[18] J. GUCKEnheimer and P. Holmes, Nonlinear oscillations, dynamical systems and bifurcation of vector fields, Springer-Verlag, 1983.
[19] G. Haller, A variational theory of hyperbolic Lagrangian Coherent Structures, Physica D, 240 (2011), pp. 574-598.
[20] G. Haller, Lagrangian coherent structures, Annu. Rev. Fluid Mech., 47 (2015), pp. 137-162.
[21] G. Haller and F. Beron-Vera, Geodesic theory of transport barriers in two-dimensional flows, Physica D, 241 (2012), pp. 1680-1702.
[22] M. HÉNON, A two-dimensional mapping with a strange attractor, Commun. Math. Phys., 50 (1976), pp. 69-77.
[23] J. Hubbard and H. Masur, Quadratic differentials and foliations, Acta Mathematica, 142 (1979), pp. 221-274.
[24] Y. Imayoshi and M. Taniguchi, An introduction to Teichmüller spaces, Springer Science \& Business Media, 2012.
[25] L. Jaeger and H. Kantz, Homoclinic tangencies and non-normal Jacobians - effects of noise in nonhyperbolic chaotic systems, Physica D: Nonlinear Phenomena, 105 (1997), pp. 79-96.
[26] L. JaEger and H. Kantz, Structure of generating partitions for two-dimensional maps, Journal of Physics A: Mathematical and General, 30 (1997), p. L567.
[27] Y. KaWashima, A new relationship between the dilatation of pseudo-Anosov braids and fixed point theory, Journal of Knot Theory and Its Ramifications, 28 (2019), p. 1950013.
[28] H. Lawson, Foliations, Bull Amer Math Soc, 80 (1974), pp. 369-418.
[29] L. Mosher, Tiling the projective foliation space of a punctured surface, Trans. Amer. Math. Soc., 306 (1988), pp. 1-70.
[30] K. Onu, F. Huhn, and G. Haller, LCS Tool: A computational platform for Lagrangian coherent structures, J. Comp. Sci., 7 (2015), pp. 26-36.
[31] V. Oseledec, A multiplicative ergodic theorem, Trans. Moscow Math. Soc., 19 (1968), pp. $197-231$.
[32] A. Papadopoulos and G. Théret, On Teichmüller's metric and Thurston's asymmetric metric on Teichmüller space, 2007.
[33] H. Poincaré, Sur les courbes definies par les equations differentielles, J. Math. Pures Appl., 1 (1885), pp. 167-244.
[34] K. Ramasubramanian and M. Sriram, A comparative study of computation of Lyapunov spctra with different algorithms, Physica D, 139 (2000), pp. 72-86.
[35] E. Rykken, Expanding factors for psedo-Anosov homeomorphisms, Michigan Math. J., 46 (1999), pp. 281296.
[36] R. Sacker and G. Sell, Dichotomies and invariant splittings for linear differential equations, J. Differential Equations, 15 (1974), pp. 429-458.
[37] S. Shadden, F. Lekien, and J. Marsden, Definitions and properties of Lagrangian coherent structures from finite-time Lyapunov exponents in two-dimensional aperiodic flows, Physica D, 212 (2005), pp. 271-304.
[38] K.-F. Tchon, J. Dompierre, M.-G. Vallet, F. Guibault, and R. Camarero, Two-dimensional metric tensor visualization using pseudo-meshes, Engineering with Computers, 22 (2006), pp. 121131.
[39] W. Thurston, On the geometry and dynamics of diffeomorphisms of surfaces, Bull. Amer. Math. Soc., 19 (1988), pp. 417-431.
[40] X. Tricoche, X. Zheng, and A. Pang, Visualizing the topology of symmetric, second-order, timevarying two-dimensional tensor fields, Visualization and Processing of Tensor Fields, (2006), pp. 225240.
[41] C. Wolfe and R. Samelson, An efficient method for recovering Lyapunov vectors from singular vectors, Tellus, 59A (2007), pp. 355-366.
[42] E. Zhang, J. Hays, and G. Turk, Interactive tensor field design and visualization on surfaces, IEEE Trans. Visual. Comp. Graphics, 13 (2007), pp. 94-107.


[^0]:    *Submitted to the editors October 4, 2023
    Funding: Partial support from the Australian Research Council (grant DP200101764 to SB) and the Army Research Office (grant N68164-EG to EB) is gratefully acknowledged.
    ${ }^{\dagger}$ School of Computer and Mathematical Sciences, University of Adelaide, Adelaide SA 5005, Australia (sanjeevabalasuriya@yahoo.com).
    ${ }^{\ddagger}$ Department of Electrical and Computer Engineering and the Clarkson Center for Complex Systems Science ( $C^{3} \mathrm{~S}^{2}$ ), Clarkson University, Potsdam NY13699 (bolltem@clarkson.edu).

