Frontiers of chaotic advection

Hassan Aref
Engineering Science & Mechanics, Virginia Tech, Blacksburg, Virginia 24061, USA

John R. Blake
School of Mathematics, University of Birmingham, Edgbaston, Birmingham, United Kingdom

Marko Budišić
Department of Mathematics and Clarkson Center for Complex Systems Science (C3S2), Clarkson University, Potsdam, New York 13699, USA

Silvana S. S. Cardoso
Department of Chemical Engineering and Biotechnology, University of Cambridge, Cambridge, United Kingdom

Julyan H. E. Cartwright
Instituto Andaluz de Ciencias de la Tierra, CSIC-Universidad de Granada and Instituto Carlos I de Fisica Teórica y Computacional, Universidad de Granada, Granada, Spain

Herman J. H. Clercx
Fluid Dynamics Laboratory, Department of Applied Physics and JM Burgerscentrum, Eindhoven University of Technology, Eindhoven, The Netherlands

Kamal El Omari
Université Pau & Pays Adour, Laboratoire des Sciences de l’Ingénieur Appliquées à la Mécanique et au Génie Électrique—Fédération IPRA, EA4581, Pau, France

Ulrike Feudel
Institut für Chemie und Biologie des Meeres, Carl von Ossietzky Universität Oldenburg, Oldenburg, Germany

Ramin Golestanian
Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, United Kingdom

Emmanuelle Gouillart
Surface du Verre et Interfaces, UMR 125 CNRS/Saint-Gobain, Aubervilliers, France

GertJan F. van Heijst
Fluid Dynamics Laboratory, Department of Applied Physics and JM Burgerscentrum, Eindhoven University of Technology, 5600 MB Eindhoven, The Netherlands

Tatyana S. Krasnopolskaya
Institute of Hydromechanics, National Academy of Sciences of Ukraine, Kiev, Ukraine

Yves Le Guer
Université Pau & Pays Adour, Laboratoire des Sciences de l’Ingénieur Appliquées à la Mécanique et au Génie Électrique—Fédération IPRA, EA4581, Pau, France
This work reviews the present position of and surveys future perspectives in the physics of chaotic advection: the field that emerged three decades ago at the intersection of fluid mechanics and nonlinear dynamics, which encompasses a range of applications with length scales ranging from micrometers to hundreds of kilometers, including systems as diverse as mixing and thermal processing of viscous fluids, microfluidics, biological flows, and oceanographic and atmospheric flows.

DOI: 10.1103/RevModPhys.89.025007

CONTENTS

I. Introduction 3
   A. Synopsis of key concepts 4
   B. A definition of chaotic advection 4
   C. Stirring and mixing 5

II. Unbounded Flows 5
   A. The chaotic saddle and its invariant sets 5
   B. Partial mixing, fractals, and fractal dimensions 7

III. The Role of Walls 11
   A. Things are not always exponential 11
   B. Passive scalar near a wall 11
   C. Exponential approach to a free-slip wall 12
   D. Algebraic approach to a no-slip wall 13
   E. Transport barriers and Kolmogorov-Arnol’d-Moser islands: The effective dimension 8
I. INTRODUCTION

A dynamical process like stirring a fluid holds a great deal more interest than a static system for physicists just as it does for everyone else. Stirring and mixing of scalars (additives, nutrients, heat, etc.) by fluid flows is the common denominator in a wide variety of natural and industrial fluid systems of sizes extending from micrometers to hundreds of kilometers. Industrial examples range from mixing in the rapidly expanding field of microfluidics, encompassing applications as diverse as microelectronics cooling, microreactors, “labs on a chip” for molecular analysis and biotechnology and “smart pills” for targeted drug delivery, up to the mixing and thermal processing of viscous fluids with compact processing equipment. Examples in nature include magma transport in the Earth’s mantle and dispersion of hydrocarbons within fractured rock, as well as gas exchange in lung alveoli and the distribution of blood-borne pathogens, and large-scale dispersion of pollutants in the Earth’s atmosphere and oceans.

Given its ubiquity in industry and nature, insight into the mechanisms underlying mixing and ways purposefully to employ and control them have great scientific, technological, and social relevance and are imperative for further development of fluid-processing technologies in, especially, microfluidics applications and process engineering. Although this insight remains incomplete to date, important physical and mathematical approaches for the analysis and understanding of mixing in laminar flows have become available during the last three decades. Moreover, both the advancement of measurement technologies to investigate mixing in laboratory setups (such as laser-induced fluorescence, particle-tracking velocimetry, microparticle image velocimetry, and so on) and the rapid development of passive and active mixing elements for microfluidic devices open the perspective for quantitative mixing studies for industrial and microfluidic applications. Key challenges for advancing this field may occur to the reader during the course of this article; we defer our thoughts on these to the discussion in Sec. IX.

The context of our review is the dynamical-systems and mathematical-physics perspective on fluid transport and mixing. Early works on the subject from the 1960s are owed to Arnol’d (1965) and Hénon (1966); the 1980s brought Arter (1983) and Aref (1984). Three decades have passed since Aref (1982, 1984, 2002) introduced the term chaotic advection, and over 25 years since a textbook on the field appeared (Ottino, 1989). During that time a great deal of research has been done. It is time to summarize 30 years of chaotic advection. Our review focuses on theoretical and mathematical-physics concepts and numerical approaches of stirring and mixing in viscous fluids. Its scope should be seen within the broader context of mixing in fluid flows and of earlier reviews. Thus we complement the textbook and review of Ottino (1989, 1990) with developments from the last two decades. We have a more physical approach than the strongly mathematical perspective of Arnol’d and Khesin (1992) [see also Arnol’d and Khesin (1998)], which reviewed mathematical concepts of transport for hydrodynamics. Mezić (2013) discussed the Koopman operator approach and Haller (2015) examined Lagrangian coherent structures; for
this reason these two topics are reviewed herein only to a limited extent. And, as this is a review paper and not a book, we do not document the ever growing field of applications in exhaustive detail, but we provide some illustrative examples. Phenomena such as chemical reactions in flows, transport of active matter, aggregation processes, droplet dynamics, and granular flow [see, for example, Ottino and Khakhar (2000)] are outside the scope of the present review and are touched upon only when they affect our core concern. The 2000 Laporte Prize lecture (Aref, 2002) links chaotic advection with dynamical-systems theory, which is the central theme of this review; its basic ideas are our starting point.

The number of papers in chaotic advection is now so numerous that anything approaching a complete coverage of the field is impossible. Here we present a view arising from a group of researchers predominantly working in the field of stirring and mixing in viscous fluids from a theoretical, numerical, and/or mathematical-physics point of view, although experimental excursions are not excluded. Some of the choices to illustrate concepts are unavoidably colored by the interests of the researchers involved in this review and should not be considered as exhaustive.

We begin with an informal definition of chaotic advection. We then look at the differences between open and bounded flows; in Sec. II we look at unbounded flows, and in Sec. III we discuss ideas on the role of walls. The new frontier is, undoubtedly, three-dimensional (3D) unsteady flow, treated in Sec. IV. There is ongoing interest in elucidating chaotic advection from numerical and experimental data that we discuss in Sec. V. In Sec. VI we discuss what is meant by a laminar or a turbulent flow; these terms are frequently used in this discipline and yet are open to a range of interpretations. In Sec. VII we examine the mixed state itself; there are interesting aspects of the mixed state, in particular, the overarching problem of the quality of mixing and mixing measures that are well worth looking at in more geometrical and topological detail. “Chaotic advection plus,” treated in Sec. VIII, is a more and more active area; the increasing range of application of these ideas is most encouraging. Last, in Sec. IX we conclude with perspectives on these frontiers of chaotic advection.

A. Synopsis of key concepts

At its simplest we may consider a flowing fluid to consist of only fluid particles, aggregations of material elements small enough to satisfy the requirements to treat the fluid as a continuum. Each fluid particle, if a particle is conceptually or actually marked, often called a tracer, is denoted by its position \( x \) and moves passively with the fluid velocity \( u \) according to the kinematic equation

\[
\frac{dx}{dt} = u, \quad x(0) = x_0.
\]  

(1)

Deceptively simple, the kinematic equation can be taken as the elementary definition of velocity or as defining a dynamical system in which a given velocity field generates so-called Lagrangian trajectories for the fluid particles. Indeed, Eq. (1) has the formal solution \( x(t) = \Phi_t(x_0) \) describing the Lagrangian trajectory of a tracer released at \( x_0 \) with a corresponding Poincaré map defined by \( x_{k+1} = \Phi(x_k) \), where \( x_k = x(kT) \) is the tracer position after \( k \) periods of a time-periodic flow. The velocity \( u \) is often derived from the (steady) Navier-Stokes and continuity equations

\[
Reu \cdot \nabla u = -\nabla p + \nabla^2 u, \quad \nabla \cdot u = 0,
\]  

(2)

here given in nondimensional form for incompressible fluids, with \( Re = UL/\nu \) the Reynolds number.

A special, but important case occurs for two-dimensional (2D), divergence-free flows where the velocity \( u \) can be derived from a so-called stream function \( \Psi(x,y) \) and the kinematic equation written in terms of the stream function as

\[
\frac{dx}{dt} = \frac{\partial \Psi}{\partial y}, \quad \frac{dy}{dt} = -\frac{\partial \Psi}{\partial x}.
\]  

Equation (3) is identical to Hamilton’s equations of motion for a 1-degree-of-freedom dynamical system with the identification of the position coordinates \( x \) and \( y \), respectively, as canonical position and momentum coordinates along with the identification of \( \Psi \) with the Hamiltonian function. This crucial insight has allowed over a century of theoretical developments from Hamiltonian mechanics to be brought to bear on 2D flow problems and links 2D fluid flow to many other areas of physics.

Under quite general conditions some trajectories or sets of trajectories advected by Eq. (1) form barriers to material transport, manifolds such as material lines or sheets that are invariant under the flow. These barriers are persistent and even in transient flows are long lived. In informal terms, these material curves or sheets are the Lagrangian coherent structures (LCSs) (see Sec. V.A) whose material lines attract or repel the neighboring material. LCSs can both facilitate and retard transport fluxes; they organize and mediate all transport and interaction of matter and energy in the flow. Finding, classifying, and manipulating these structures plays the central role in both analysis of and design with flows. These organizing structures can meander wildly throughout the space of interest, and the wild meanders associated with chaotic dynamics give the title chaotic advection to this entire field of study.

B. A definition of chaotic advection

In many applications one wants to maximize the rate of mixing of a fluid. In the simplest setting, this means that we want to reduce as much as possible the time it takes for molecular diffusion to homogenize an initially inhomogeneous distribution of a scalar tracer. If there is no advection, molecular diffusion to homogenize an initially inhomogeneous distribution of a scalar tracer. If there is no advection, molecular diffusion by itself takes a very long time to achieve homogeneity, even in quite small containers. So we use advection to accelerate this process. The classical and more well-known way to do so is through turbulence: by imposing a high Reynolds number in a 3D flow, we trigger the formation of a Kolmogorov energy cascade (Kolmogorov, 1941a, 1941b; Tritton, 1988; Frisch, 1996; Kundu and Cohen, 2008) whereby energy flows from large to small scales. This energy cascade is mirrored by a corresponding cascade in any scalar field advected along with the flow, whose distribution develops in this process small-scale structures, which are then
rapidly homogenized by molecular diffusion. From the point of view of mixing, such turbulence is therefore a way to create, quickly, small-scale structures in the spatial distribution of advected fields, resulting in their being smoothed by diffusion.

Chaotic advection (Aref, 1984) is a different way to generate small-scale structures in the spatial distribution of advected fields, by using the stretching and folding property of chaotic flows. This chaotic dynamics quickly evolves any smooth initial distribution into a complex pattern of filaments or sheets, depending on the dimensionality of the system, which tends exponentially fast to a geometric pattern with a fractal structure. Owing to the stretching, the length scales of the structures in the contracting directions decrease exponentially fast, and when they become small enough, they are smoothed out by diffusion. This is a purely kinematic effect, which does not require high Reynolds numbers and exists even in time-dependent 2D Stokes flows. Chaotic advection can thus be defined as the creation of small scales in a flow by its chaotic dynamics. Mixing by chaotic advection has the advantages over turbulence that it does not require the larger input of energy needed to maintain the Kolmogorov cascade that turbulent mixing does, and it can be set up in situations, such as microfluidics, in which a high Reynolds number is not an option.

C. Stirring and mixing

The terminology regarding mixing and stirring is not always consistent in the literature. We suggest the following (although it is impossible to be taxative, as both words are so embedded in common usage). Stirring is advective redistribution, i.e., purely kinematic transport, and mixing is stirring together with diffusive effects. To add molecular diffusion to the mathematical conception of advection laid out in Sec. I.A, one has the advection-diffusion equation

$$\frac{\partial \phi}{\partial t} + \mathbf{v} \cdot \nabla \phi = \frac{1}{\text{Pe}} \nabla^2 \phi$$

with appropriate boundary and initial conditions, where $\phi$ is the scalar concentration and the Péclet number $\text{Pe}$ balances diffusive and advective time scales. The natural scale for mixing is the Batchelor scale (Batchelor, 1959)

$$\ell_B = \frac{\kappa}{\sqrt{\lambda}},$$

where $\kappa$ is the molecular diffusivity and $\lambda$ is the Lyapunov exponent of the flow. At scales smaller than $\ell_B$, diffusion smooths out concentration gradients and mixing is achieved at the molecular scale.

The sections of this review concentrate upon stirring (Secs. II, III, IV, V, and VI) and mixing (Secs. VIII and IX). Section VII is the bridge between the two parts.

II. UNBOUNDED FLOWS

Chaos in open flows manifests itself through the appearance of fractal structures in the advection of an initially smooth distribution of passive tracers (Tél et al., 2005). These fractal patterns arise as a direct result of the existence of a nonattracting chaotic set in the advection dynamics (Lai and Tél, 2011).

Let us begin by considering a 2D channel flow past an obstacle (Fig. 1). Fluid particles come from an inflow region, may stay in the wake of the obstacle for some time, and then leave through the outflow. If we consider a limited region around the obstacle as our observation region $R$, most fluid particles stay only a finite time in $R$ before escaping to the outflow region (Jung, Tél, and Ziemniak, 1993). The dynamics of advection in an open flow is therefore transient. The transient nature of the dynamics makes mixing in open flows qualitatively different from the closed flow case discussed in Sec. III: the very definition of mixing and its mathematical formulation are different.

A. The chaotic saddle and its invariant sets

In open chaotic flows, the perpetual alternation of stretching and folding results in filaments that grow ever thinner because of the fluid’s escape, as depicted in Fig. 2. As they become thinner, they grow longer and more convoluted. This happens because of the stretching and folding properties of the
dynamics. To understand the consequences of this stretching and folding, let us take an initial region of the flow, located in a place where mixing occurs, for example, immediately behind the wake of an obstacle. If the flow is chaotic, this initial region will be repeatedly stretched and bent back on itself. After some time, some of the fluid in our original region has escaped, but some has come back and intersects the original region. This intersection is the fundamental characteristic of a Smale horseshoe (Smale, 1967, 1998; Shub, 2005), and it immediately follows from Smale’s results that there are infinitely many unstable periodic orbits in the intersection region, as well as an uncountable infinity of aperiodic orbits. These orbits are collectively referred to as the chaotic saddle. The set of initial conditions giving rise to trajectories that approach one of these orbits asymptotically consists of orbits that never escape (they are trapped) since they converge to orbits bound to a compact region of space: the intersection region previously discussed.

Open flows are therefore characterized by a transient dynamics of most fluid particles. More precisely, if one randomly picks an initial condition in the inflow region, the corresponding trajectory will escape to the outflow region with probability 1. In other words, the set of initial conditions which stay trapped forever in an open flow has Lebesgue measure zero.

Although they have zero measure in phase space, these trapped orbits are very important for open flows, because they govern the long-time advection dynamics: those orbits that take a long time to escape correspond to initial conditions lying close to the trapped trajectories. If the advective dynamics of the open flow is chaotic, each trapped trajectory converges asymptotically as $t \to \infty$ to one of the orbits in the chaotic saddle (Lai and Tél, 2011); orbits in the chaotic saddle do not go to the outflow region for $t \to \infty$, and they do not go to the inflow region for $t \to -\infty$. So these orbits lie on a confined portion of space, where mixing takes place in open flows; we refer to this region as the mixing region from now on. In the example of a flow past an obstacle, the mixing region and the chaotic saddle are located in the wake of the obstacle, and the mixing region typically extends for no more than a few times the length of the obstacle (Jung, Tél, and Ziemniak, 1993).

The set of trapped trajectories corresponds to the stable manifold of the chaotic saddle (Lai and Tél, 2011). Conversely, the set of orbits that converge to the chaotic saddle in backward time (i.e., for $t \to -\infty$) is the saddle’s unstable manifold. Both the stable and unstable manifolds have important physical interpretations for the advection dynamics. Long-lived orbits are close to the stable manifold. The physical meaning of the unstable manifold comes from the fact that those trajectories that stay a long time in the mixing region, that is, those lying close to the stable manifold in the inflow region, will trace out the unstable manifold on their way out toward the outflow region. An initial blob of dye, or anything else that passively follows the flow, is repeatedly stretched and folded by the flow in the mixing region, generating a convoluted filamentary structure that converges to the unstable manifold; a sketch of this process is shown in Fig. 2. As a consequence, the unstable manifold can be directly observed in experiments by following a dye as it is advected in the fluid (Sommerer, Ku, and Gilreath, 1996). Once the bulk of the dye has escaped, what still remains in the observation region shadows the unstable manifold.

An experimental example is shown in the top line of Fig. 3, where a potential flow dipole in a Hele-Shaw cell is periodically reoriented (Metcalfe et al., 2010a, 2010b); the bottom line shows a computed version of the same flow. The initial condition of pink fluid is removed by inflowing blue fluid. As the flow proceeds, the persistent pink lines thin but never leave the cell; these are examples of filamentary manifolds. The large pink blob is an island that incoming fluid also never displaces.

FIG. 3 Filamentary manifolds evolving and thinning in a (top) Hele-Shaw flow experiment and (bottom) computer simulation with an opposite inflow and outflow creating a dipole that is on for a scaled time $\tau$, after which the dipole is reoriented by an angle $\Theta$. A number of iterations are shown at the bottom for $\tau = 0.1$ and $\Theta = \pi/4$. From Metcalfe et al., 2010a, 2010b.
B. Partial mixing, fractals, and fractal dimensions

The repeated stretching and folding of fluid elements caused by the presence of the chaotic saddle results in mixing. In open flows this kinematic mechanism has only limited time to act because of the transient nature of the advection dynamics. One can therefore say that open chaotic flows induce partial mixing: an initial blob of dye is deformed into a dynamics. One can therefore say that open chaotic flows act because of the transient nature of the advection.

In open flows this kinematic mechanism has only limited time caused by the presence of the chaotic saddle results in mixing. The repeated stretching and folding of fluid elements described above results in a chaotic mixing process. This is a defining feature of the Sinai-Ruelle-Bowen (SRB) measure, which describes the natural probability distributions of transient chaotic systems. Although we have been focusing on the unstable manifold, the same properties are shared by the stable manifold as well.

From the point of view of mixing, it is clear from Fig. 4, and from the previous discussion on the structure of the unstable manifolds, that there is efficient mixing in the directions locally transversal to the unstable manifold, but no mixing happens in the direction of the unstable manifold. Contrast this to the case of closed flows (Sec. III), where the unstable manifold is space filling, and mixing will eventually take place along all directions, given enough time for the system to evolve.

How does one quantify the amount of mixing in an open flow? Looking at Fig. 4, one would intuitively want to measure mixing by how much area in the picture is occupied by points lying close to both the black and white regions. But what does “close to” mean? We have stated that the natural scale for mixing is the Batchelor scale, Eq. (5). The unstable manifold separates the black and white regions in the limit $t \to \infty$. Let us thus define $A(\varepsilon)$ to be the area of the set $S$ of

$$r = -C/r, \quad \dot{\phi} = K/r^2. \quad (6)$$

Solving these with initial conditions $r_0$ and $\phi_0$, we get

$$r(t) = (r_0^2 - 2Cr)^{1/2}, \quad \phi(t) = \phi_0 - \frac{K}{C} \ln \frac{r(t)}{r_0}. \quad (7)$$

Without loss of generality, we choose the positions of the vortices at $x = \pm a, y = 0$, where $a$ is a parameter of the system. Since we have an analytical expression for the motion of fluid particles for each of the half periods, we may find an expression for the new position $r_{n+1}$ after one period as a function of the position $r_n$ at the beginning of the period. This is best done using a complex representation for the position of a fluid particle $z = x + iy$. The mapping from the initial position $z_n$ to the new one $z_{n+1}$ is then given by

$$z_{n+1/2} = (z_n + a) \left( 1 - \frac{CT}{|z_n + a|^2} \right)^{1/2 - iK/(2C)} - a,$$

$$z_{n+1} = (z_{n+1/2} - a) \left( 1 - \frac{CT}{|z_{n+1/2} - a|^2} \right)^{1/2 - iK/(2C)} + a. \quad (8)$$

Here $z_{n+1/2}$ is an intermediate variable representing the particle’s position after the first half period. One can see in Fig. 4 that the filaments are arranged in intricate layers, such that if one zooms in around a given filament, the nearby filaments are oriented along roughly the same direction. In other words, in the case of 2D flows, due to stretching and folding one finds, in a small volume, an infinite number of sections of the manifold, lined up in parallel, and densely packed in the perpendicular direction. More precisely, the unstable manifold is locally the direct product of a Cantor set and a 1D smooth curve (Lai and Tél, 2011). This means that at any given point the asymptotic distribution of any tracer advected by the flow varies smoothly in the direction along the unstable manifold at that point, while it varies wildly in directions transversal to the unstable manifold (Tél et al., 2005). This is a defining feature of the Sinai-Ruelle-Bowen (SRB) measures (Ott, 1993), which describe the natural probability distributions of transient chaotic systems. Although we have been focusing on the unstable manifold, the same properties are shared by the stable manifold as well.

![FIG. 4. Blinking vortex-sink flow: Evolution of a set of particles. From Károlyi and Tél, 1997, and Tél et al., 2005.](image-url)
points such that their distance from the unstable manifold is smaller than \(e\), restricted to a finite observation region \(\mathcal{R}\) containing the chaotic saddle. \(A\) can be estimated by covering \(\mathcal{R}\) with a grid of size \(\varepsilon\) and counting the number \(N(\varepsilon)\) of grid elements that intersect with the unstable manifold. \(A\) is then given by \(A(\varepsilon) = N(\varepsilon)\varepsilon^2\). The way \(N(\varepsilon)\) scales with \(\varepsilon\) is governed by the fractal dimension \(D\) of the unstable manifold (Ott, 1993; Falconer, 2003):

\[
N(\varepsilon) \sim \varepsilon^{-D}.
\]

(9)

The area \(A(\varepsilon)\) therefore scales with \(\varepsilon\) as

\[
A(\varepsilon) \sim \varepsilon^{2-D}.
\]

(10)

For a regular, nonchaotic 2D flow, the unstable set is a simple 1D curve, and thus \(D = 1\). In this case \(A\) is proportional to \(e\). If the flow displays chaotic advection, \(D\) satisfies \(1 < D \leq 2\), and \(A\) decreases sublinearly with \(e\). This means that in a chaotic flow, the mixing area \(A(\varepsilon)\) decays very slowly with \(e\). Using \(e = \varepsilon^\frac{1}{D}\), from Eq. (5) we see that this results in a slow decay of \(A\) with the diffusivity. This slow decrease of \(A(\varepsilon)\) with \(\varepsilon\) has many consequences for the dynamics of processes taking place in the flow. These include a singular increase in the rate of chemical reactions in open chaotic flows (Tel et al., 2005) and an anomalous scaling in the collision rate of particles (De Moura, 2011).

We have concentrated on the meaning of the fractal dimension of the unstable manifold. The fractal dimension of the stable manifold also has a physical meaning: it is a measure of the sensitivity of the dynamics of fluid particles in open flows to the initial conditions (Grebogi et al., 1983). Let \(p(\varepsilon)\) be the probability that the trajectories corresponding to two initial conditions separated by a small distance \(\varepsilon\) eventually separate before escaping, so that they escape following completely different paths. \(p(\varepsilon)\) can be numerically calculated by choosing a large number of pairs of points located randomly in space and following their trajectories until they escape. If, for example, they escape in different cycles (assuming that the flow is time periodic), we consider them to have separated. The initial conditions within a distance \(\varepsilon\) of the stable manifold are at risk of separating, and thus \(p(\varepsilon)\) is proportional to \(A(\varepsilon)\) and scales as

\[
p(\varepsilon) \sim \varepsilon^{2-D}.
\]

(11)

\(p(\varepsilon)\) can be considered as a measure of the uncertainty in the prediction of the ultimate fate of the trajectory of a given fluid particle, when its initial condition is given with an experimental error of size \(\varepsilon\). Decreasing \(\varepsilon\) means an increase in accuracy in the determination of the initial condition. For nonchaotic flows, \(D = 1\), and therefore \(p(\varepsilon) \sim \varepsilon\); decreasing \(\varepsilon\) by a factor of 10 would decrease the uncertainty by the same factor, as one might expect. If the flow displays chaotic advection, however, \(p(\varepsilon)\) does not scale linearly with \(\varepsilon\), and the uncertainty decreases more slowly. For the case of \(D = 1.9\), for example, it would take a decrease of 10 orders of magnitude in \(\varepsilon\) to reduce \(p(\varepsilon)\) by a factor of 10.

C. Hyperbolicity and the Grassberger-Kantz relation

Open hyperbolic systems have exponential decay: if we keep track of the time evolution of a typical area of flow, the amount \(Q(t)\) of this initial area still remaining in the mixing region at time \(t\) decays exponentially with \(t\) for large \(t\): \(Q(t) \sim \exp(-\kappa_e t)\). \(\kappa_e\) is the escape rate of the flow. It satisfies \(\kappa_e < \lambda\), where \(\lambda\) is the chaotic saddle’s Lyapunov exponent. The fractal dimension \(D\) of the unstable manifold, the Lyapunov exponent \(\lambda\), and the escape rate \(\kappa_e\) are related by the Grassberger-Kantz formula (Kantz and Grassberger, 1985):

\[
D = 2 - \frac{\kappa_e}{\lambda}.
\]

(12)

More rigorously, we should have \(D_1\), the information dimension (Falconer, 2003), instead of the box-counting dimension \(D\) in Eq. (12), but since \(D\) and \(D_1\) are almost always very close for open flows, this approximation is valid in most cases.

D. Robustness of the chaotic saddle

In the previous discussion, and in most of what follows in this section, we concentrate on the case of 2D flows. Furthermore, we have concentrated on the motion of fluid particles, that is, of passive tracers that assume exactly the velocity of the surrounding fluid. The fractal structure of the chaotic saddle and its associated invariant manifolds persist, however, in the case of actual, finite-sized particles, which have inertia and whose velocities do not coincide with that of the fluid’s velocity field (Vilela, de Moura, and Grebogi, 2006; Vilela et al., 2007; Cartwright et al., 2010). There are some considerable differences between the dynamics of fluid particles and that of inertial particles, in particular, the possibility of the appearance of attractors in the latter case (Benczik, Toroczkai, and Tel, 2002; Cartwright et al., 2002, 2010; Motter, Lai, and Grebogi, 2003). But even when the global dynamics has attractors, chaotic saddles are still present, and the system is still governed by fractal structures in phase space connected to a chaotic saddle, as in the simpler case of passive advection.

The same overall picture remains valid for 3D systems (Cartwright, Feingold, and Piro, 1996; de Moura and Grebogi, 2004a; Tuval et al., 2004); in this case, the stable and unstable manifolds are a fractal set of sheets, instead of segments. Periodicity is also not required for the existence of the chaotic saddle: aperiodic and random flows can also result in well-defined fractal structures in phase space (Károlyi et al., 2004; Rodrigues, de Moura, and Grebogi, 2010).

The conclusion is that the concepts of chaotic saddle and its stable and unstable manifolds are remarkably robust and are not consequences of oversimplified models of flows.

E. Transport barriers and Kolmogorov-Arnol’d-Moser islands: The effective dimension

In discussions about chaotic open flows and the chaotic saddle it is often assumed, sometimes tacitly, that the dynamics is hyperbolic. The reason is partly that the hyperbolic case is more tractable, and there are more rigorous results available. However, nonhyperbolicity occurs in many
important cases and is to be expected in many very general scenarios in fluid dynamics. For example, it can be shown that the dynamics of 2D advection of a flow past an obstacle becomes chaotic immediately after the transition of the flow from stationary to time dependent, as the Reynolds number is increased beyond a critical value; furthermore, the dynamics is nonhyperbolic for a range of Reynolds numbers past the transition point, independently of the shape of the obstacle or the particular features of the flow (Biemond et al., 2008). Many other systems of interest are nonhyperbolic, so it is imperative to understand the mixing dynamics in the nonhyperbolic case.

Nonhyperbolicity is manifested through the appearance of stable orbits in space. These orbits are surrounded by stable Kolmogorov-Arnold-Moser (KAM) islands (MacKay and Meiss, 1987). KAM vortices are well known in closed flow (Sec. III), and they have been extensively studied in that context. What is perhaps less well known is that they can also appear in open flows, for example, in the flow of Fig. 3, and when they do, they play a crucial role in the mixing dynamics. They have been observed in geophysical 2D flows, such as the stratospheric polar vortex, which plays a crucial role in the process of ozone depletion (Koh and Legras, 2002) and also in ocean circulation patterns (Abraham, 1998; Abraham et al., 2000; Boyd et al., 2000). The islands form a fractal hierarchical structure, with large islands being surrounded by smaller islands, and these in turn are surrounded by even smaller islands, and so on (Fig. 5). The presence of KAM islands means that there is a finite volume of initial conditions in the mixing region whose orbits do not escape, corresponding to those initial conditions lying in the islands. Moreover, fluid particles with initial conditions outside the interaction region cannot enter the islands. As a result, the set of initial conditions outside the mixing region whose trajectories end up trapped there still has zero measure, as in the hyperbolic case. However, the islands have deep consequences for the transient dynamics, resulting in important differences between the hyperbolic and nonhyperbolic cases.

The transport of fluid in the vicinity of the islands is dominated by cantori, which are remnants of broken up KAM tori. Cantori are invariant sets of the dynamics, as are KAM islands; but in contrast to those, fluid particles can cross from one side of a cantorus to the other (MacKay, Meiss, and Percival, 1984; MacKay and Meiss, 1987). However, it typically takes very long times to do so, and as a consequence the cantori act as partial transport barriers. The overall picture of nonhyperbolic transport is sketched in Fig. 5.

Figure 6 shows Poincaré sections for a flow simulation with nonhyperbolic advection dynamics. The magnification shows the striking self-similar organization of the islands. The effect of cantori on the advection dynamics can be seen in the cloud of points surrounding the subislands on the upper right and to the left of the main island in the magnified figure. These points are snapshots taken at the start of every period of a single orbit that meanders inside the cantorus surrounding these islands. This orbit eventually escapes after thousands of cycles. Another cantorus can just be seen surrounding the main island. These cantori are in turn surrounded by a larger cantorus encircling the whole structure, which is apparent from the higher density of points in the region around the complex of islands in the bottom figure of Fig. 6. An experimental example showing KAM islands in an open flow motivated by mixers in the food industry was investigated by Gouillart et al. (2009).

---

**FIG. 5.** KAM tori and cantori: hierarchical structure. Solid lines indicate KAM tori and cantori are represented by the dotted lines. From Tél et al., 2005.

**FIG. 6.** Blinking vortex-sink flow [Eq. (8)]: Poincaré map with $C = T = a = 1, \ K = 18$, parameters for which the flow is nonhyperbolic. The map shows the orbits of a few fluid particles, with positions taken at discrete times, at the beginning of every period of the flow. The picture on the bottom is a magnification of a small region of the top image and shows the self-similar structure of the KAM islands.
The partition of space by the KAM islands and cantori into distinct domains separated by transport barriers has no counterpart in hyperbolic systems and is the cause of the profound differences in the dynamics of hyperbolic and nonhyperbolic flows. A direct consequence of the self-similar structure of the transport barriers depicted in Fig. 5 is the phenomenon known as stickiness: in nonhyperbolic flows, many trajectories spend extremely long times inside cantori, leading to very long typical escape times compared to hyperbolic dynamics. Once inside, an orbit may enter an inner cantorus located within another cantorus, and so on to arbitrarily high levels in the cantorus hierarchy. So once a fluid particle is inside a cantorus, it will wander within a fractal labyrinth from which escape is likely to take a very long time. Note that the preceding discussion holds for 2D systems only. However, with more degrees of freedom there is also the possibility of Arnol’d diffusion (Arnol’d, 1964).

Even in nonhyperbolic flows it is still true that fluid particles with initial conditions outside of KAM islands will eventually escape with 100% probability: the component of the chaotic saddle outside the islands has zero measure. But stickiness makes escape subexponential, in marked contrast with hyperbolic flows. In nonhyperbolic flows, the number \( N(t) \) of particles with initial conditions randomly chosen in a region with no intersection with KAM islands that have not escaped up to time \( t \) follows a power law (Meiss and Ott, 1985):

\[
N(t) \sim t^{-\gamma},
\]

with \( \gamma > 0 \).

It has been shown that a direct consequence of the slower escape dynamics described by Eq. (13) is that the fractal dimension \( D \) of the stable (and unstable) manifold is equal to the dimension of the embedding space \( D = 2 \) (Lau, Finn, and Ott, 1991). From the interpretation of the fractal dimension as a measure of uncertainty of transient systems, expressed mathematically by Eq. (11), the fact that \( D \) assumes the maximum possible value in nonhyperbolic systems suggests that these systems have an extreme sensitivity to initial conditions. Indeed, the exponent in Eq. (11) vanishes for \( D = 2 \), which means that the “uncertainty probability” \( p(\varepsilon) \) decreases more slowly than a power law for small \( \varepsilon \). The fact that \( D = 2 \) in nonhyperbolic open flows suggests that predicting asymptotic properties of trajectories in these systems is an almost impossible task. The reason for this unpredictability is the very long time it takes initial conditions inside cantori to escape: two initially close trajectories will have much more time to spend in the mixing region to separate and follow independent paths before they escape. Figures 5 and 6 also suggest that the unpredictability is greater for initial conditions located in deeper levels of the cantorus hierarchy, as they have longer escape times.

To investigate the sensitivity to initial conditions in different areas of space in Fig. 6 one may calculate numerically the fraction \( f(\varepsilon) \) of \( \varepsilon \)-separated pairs of points whose escape times differ by one period or more. For a sufficiently large sample, we expect \( f(\varepsilon) \propto p(\varepsilon) \). The result, for initial conditions taken in two different cantori, is plotted in Fig. 7. Figure 7 seems to go against the claim that \( D = 2 \) for nonhyperbolic systems, since this would predict that the plot of \( f(\varepsilon) \) vs \( \varepsilon \) should be a line with zero slope. But in nonhyperbolic systems, the \( \varepsilon \to 0 \) limit in Eq. (11) converges sublogarithmically with \( \varepsilon \) (Lau, Finn, and Ott, 1991). This extremely slow convergence means that reaching this limit usually requires values of \( \varepsilon \) so small they are not physically meaningful. Any model of a physical system has a lower scale below which the model is no longer valid: for example, the size of advected particles or the finite resolution of our measurements. This implies that the dimension that is physically relevant for realistic systems is not the mathematical definition Eq. (11) with its unreachable limit, but is given instead by an effective dimension \( D_{\text{eff}} \) (de Moura and Grebogi, 2004b; Motter et al., 2005), defined as an approximation of the fractal dimension for a finite range of \( \varepsilon \):

\[
D_{\text{eff}}(\varepsilon) = 2 - \frac{d \ln f(\varepsilon)}{d \ln \varepsilon} \approx \text{const} \quad \text{for } \varepsilon_1 < \varepsilon < \varepsilon_2.
\]

valid in a range \( (\varepsilon_1, \varepsilon_2) \), with \( \varepsilon_1 < \varepsilon_2 \), \( D_{\text{eff}} \) satisfies \( D_{\text{eff}}(\varepsilon) \to 2 \) as \( \varepsilon \to 0 \), in accordance with Eq. (11). From Eq. (14), the results in Fig. 7 can be interpreted as yielding the effective fractal dimensions of the stable and unstable manifolds for two different locations in space: \( D_{\text{eff}} = 1.86 \) inside the outermost cantorus, and \( D_{\text{eff}} = 1.98 \) inside one of the inner cantori. The effective dimension therefore depends on the position in nonhyperbolic systems, in contrast to the actual fractal dimension, which is 2 anywhere. The greater escape time in the inner cantori means that the invariant manifolds of the chaotic saddle have more time to be stretched and folded and distorted by advection, hence the greater effective dimension.

Because of time-reversal symmetry, the stable and unstable manifolds have the same fractal dimensions and also the same

FIG. 7. Blinking vortex-sink flow: Fraction of uncertain pairs as a function of separation \( \varepsilon \) between points in a pair, with \( \eta = 1 \), \( \xi = 18 \); see Fig. 4. Initial conditions for the bottom points are taken in the outermost cantorus, on the segment \( x_0 = 1.75 \), \( y_0 \in (-1.18, -1.175) \); for the top points, initial conditions are in an inner cantorus, on \( x_0 = 1.75 \), \( y_0 \in (-1.193, -1.192) \). The numbers beside each line are the slope coefficients obtained from fitting \( f(\varepsilon) \sim \varepsilon^s \).

Rev. Mod. Phys., Vol. 89, No. 2, April–June 2017 025007-10
effective fractal dimensions. We argued above that the fractal dimension of the unstable manifold is a measure of lower-scale mixing efficiency for open flows. This means that the fluid in regions of space surrounded by cantori will be extremely well mixed, and the efficiency of mixing increases as we go deeper into the cantorus structure and reaches the maximum limit of $D_{\text{eff}} \to 2$ for regions buried deep within the cantori. This picture is somewhat at odds with an idea prevalent in this field that KAM islands are obstacles to mixing. That view is justified in closed flows (Sec. III), where one wants to mix the fluid homogeneously throughout the container; this is not possible in the presence of KAM islands. In open flows, however, the fluid to be mixed usually comes from the inflow region, and thus from outside the KAM islands, and so this is not an issue if the material to be mixed is injected into the flow outside the islands. For open flows, one wants the flow to be well mixed by the time it reaches the outflow region. The cantori surrounding KAM islands greatly enhance this kind of mixing, by causing fluid to spend very long times within themselves. This comes at a cost: the time it takes for any given piece of fluid to escape a cantorus to the outflow is very much increased by the stickiness. If one has a protocol directly related to the concept of geometric phases (Arrieta et al., 2007, 2008, 2009; Gouillart, Thiffeault, and Dauchot, 2010b) have confirmed this hypothesis, and also showed that for a significant period of time the rate of decay of variance is dramatically reduced, even away from the walls, due to the entrainment of unmixed material into the central mixing region.

In this section we describe the limiting effect of boundaries on chaotic mixing, following Gouillart et al. (2007, 2008). We then explain how creating closed orbits near the wall alleviates the problem somewhat, by “shielding” the central mixing region from the detrimental effect of walls (Gouillart, Dauchot, and Thiffeault, 2011; Thiffeault, Gouillart, and Dauchot, 2011). We end by exploring mixing by nonreciprocal contractible loops of the wall’s positions, a class of protocols directly related to the concept of geometric phases (Arrieta et al., 2015).

III. THE ROLE OF WALLS

Many studies of mixing over the years have used maps of flows in periodic domains (cat map, standard map, etc.) to great effect to generate insight into the evolution of chaotic dynamics. However, in actual containers the solid boundaries throw up several new effects whose consequences for mixing are not confined to thin boundary layers but penetrate into the bulk of the flow. Mixing follows a different dynamics when the flow is confined to a closed space. As discussed, in general stirring induces chaotic advection in the flow, causing fluid elements to be repeatedly stretched and folded, generating over time a very fine pattern of thin filaments with a complex structure. In a closed container, these filaments eventually spread throughout the available space, as long as the flow has no prominent regular islands, and total mixing is achieved in the asymptotic limit of $t \to \infty$. The physically relevant questions are then related to the time scale over which mixing is achieved, and how the system approaches the limit of becoming perfectly mixed.

A. Things are not always exponential

The prototype of chaotic mixing is as follows: stirring a fluid promotes chaotic advection, which leads to an exponential stretching of fluid elements. These fluid elements carry some concentration of a substance to be mixed, and as they are stretched gradients of concentration increase exponentially. This allows molecular diffusion to act efficiently, and the uniformization of the concentration proceeds at a much faster rate than it would have in the absence of stirring (Eckart, 1948; Welander, 1954; Batchelor, 1959). Typically, this decay is exponential in time; the decay constant is not, however, simply the average rate of stretching (infinite-time Lyapunov exponent), but is obtained from the distribution of finite-time Lyapunov exponents in a nontrivial manner (Antonsen, Fan, and Ott, 1995; Antonsen et al., 1996; Balkovsky and Fouxon, 1999; Falkovich, Gawędzki, and Vergassola, 2001; Thiffeault, 2008). This is the “local” picture of chaotic mixing; in some cases it must be supplemented by a more global approach, where one directly analyzes the advection-diffusion operator (Pierrehumbert, 1994; Fereday et al., 2002; Wonhas and Vassilicos, 2002; Pikovsky and Popovych, 2003; Thiffeault and Childress, 2003; Fereday and Haynes, 2004; Haynes and Vanneste, 2005). However, whether the decay of concentration is locally or globally controlled, the rate is still exponential.

This exponential decay framework is helpful, but it is complicated by the presence of walls. In this case, several (Jones, Thomas, and Aref, 1989; Jones and Young, 1994; Chertkov and Lebedev, 2003; Lebedev and Turitsyn, 2004; Schekochihin, Haynes, and Cowley, 2004; Popovych, Pikovsky, and Eckhardt, 2007; Salman and Haynes, 2007; Chernykh and Lebedev, 2008; MacKay, 2008; Boffetta, De Lillo, and Mazzino, 2009; Zaggout and Gilbert, 2012) have suggested that the no-slip boundary condition and the presence of separatrices on the walls slow down mixing: the decay is power law rather than exponential. (This is connected to a breakdown of hyperbolicity.) Recent experiments (Gouillart et al., 2007, 2008, 2009; Gouillart, Thiffeault, and Dauchot, 2010b) have confirmed this hypothesis, and also showed that for a significant period of time the rate of decay of variance is dramatically reduced, even away from the walls, due to the entrainment of unmixed material into the central mixing region.

B. Passive scalar near a wall

Consider the experiment shown in Fig. 8(a): a dark blob of ink in a light fluid has been stretched and folded repeatedly by the periodic movement of a rod. The movement of the rod defines a figure-eight stirring protocol with period $T$, shown as a dashed line in Fig. 8(b). 3D effects are negligible, and the fluid flow can be treated as a Stokes flow. The mixing pattern has a kidney shape, and it slowly grows and approaches the wall. The distance of closest approach at the top is $d(t)$, where $t$ is time.

Inside the central mixing region, we assume the action of the flow is that of a simple chaotic mixer. By this we mean that fluid elements are stretched, on average, at a given rate $\lambda$ (the Lyapunov exponent). Hence, after a time $t$ a blob of initial size $\delta$ will have length $\delta e^{\lambda t}$. However, because of
Persistence before it is wiped out by diffusion? The answer is that the mixing pattern grows uniformly along the periphery of the wall, we can write the width of the mixing region by a factor \( \exp(\lambda t) \) (Batchelor, 1959; Balkovsky and Fouxon, 1999; Thiffeault, 2008).

We emphasize that the flow in Fig. 8(a) is globally chaotic in the sense that it does not possess visible islands, as evidenced by the numerical Poincaré section in Fig. 8(b). The chaotic region extends all the way to the wall, but there are clearly two special points at the wall, at the top and bottom (shown as dots in Fig. 8(b)) corresponding to separatrices. They are associated with the stable (top) and unstable (bottom) manifolds of two distinguished nonhyperbolic (parabolic) fixed points at the wall.

Each period the pattern gets progressively closer to the wall. Assuming molecular diffusion can be neglected, because of area preservation some white fluid must have entered the central mixing region. It does so in the form of white strips, visible as layers inside the pattern of Fig. 8(a). If we assume that the mixing pattern grows uniformly along the periphery of the wall, we can write the width \( \Delta(t) \) of a strip injected at period \( t/T \) as

\[
\Delta(t) = d(t) - d(t + T) = -T \dot{d}(t) \geq 0, \quad (15)
\]

where we also assumed that \( d(t) \) changes slowly in time.

Now if a white strip is injected at time \( \tau \), how long does it persist before it is wiped out by diffusion? The answer is the solution \( t \) to the equation

\[
\Delta(t)e^{-\lambda(t-\tau)} = \varepsilon_B. \quad (16)
\]

This means that the strip initially had width \( \Delta(\tau) \) when it was injected, it gets compressed by the flow in the central mixing region by a factor \( \exp(-\lambda(t-\tau)) \) depending on its age \( t - \tau \), and once it is compressed to the Batchelor length \( \varepsilon_B \) it quickly diffuses away. Thus, we can solve Eq. (16) to find the age the strip was when it gets wiped out by diffusion,

\[
t - \tau = \lambda^{-1} \log[\Delta(\tau)/\varepsilon_B]. \quad (17)
\]

Eventually, at time \( t_B \), any newly injected filament will have width equal to the Batchelor length. This occurs when

\[
\Delta(t_B) = \varepsilon_B. \quad (18)
\]

which can be solved for \( t_B \) given a form for \( \Delta(t) \). After this time it makes no sense to talk of newly injected filaments as “white,” since they are already dominated by diffusion at their birth. Hence, the description we present here is valid only for times earlier than \( t_B \), but late enough that the edge of the mixing pattern has reached the vicinity of the wall.

In their experiments, Gouillart et al. (2007, 2008) measured the intensity of pixels in the central mixing region. They observed for \( T \ll t \lesssim t_B \) that the concentration variance is dominated by the proportion of strips in the central region that are still white at that time. Because of area conservation, the total area of injected white material that is still visible at time \( t \) is proportional to

\[
A_w(t) = d(\tau(t)) - d(t), \quad (19)
\]

where we use Eq. (17) to solve for \( \tau(t) \), the injection time of the oldest strip that is still white at time \( t \). Hence, the goal is to estimate \( A_w(t) \) for times \( T \ll t \lesssim t_B \), since \( A_w \) is directly proportional to the concentration variance. To do this we need \( \tau(t) \), which requires specifying \( \Delta(t) \). We now look at three possible forms, corresponding to a free-slip wall, a no-slip wall, and a moving no-slip wall.

C. Exponential approach to a free-slip wall

Consider first the case where \( d(t) = d(0)e^{-\mu t} \) for some positive constant \( \mu \). We have \( \Delta(t) = -T \dot{d} = \mu T d(0)e^{-\mu t} = \Delta(0)e^{-\mu t} \). From Eq. (18), we have \( t_B = \mu^{-1} \log[\Delta(0)/\varepsilon_B] \), and from Eq. (17),

\[
t - \tau = \frac{\mu}{\lambda - \mu} (t_B - t). \quad (20)
\]

By assumption, \( \tau < t < t_B \), so for consistency we require \( \mu < \lambda \), i.e., the rate of approach toward the wall is slower than the natural decay rate of the chaotic mixer. The area of white material in the mixing region is then obtained from Eq. (19):

\[
A_w(t) = d(0)e^{-\mu t} \left[ \exp \left( \frac{\mu^2}{\lambda - \mu} (t_B - t) \right) - 1 \right], \quad (21)
\]

which in the wall-dominated regime \( (\lambda/\mu \gg 1) \) can be approximated by

\[
A_w(t) \sim d(0)\lambda^{-1} \mu^2 (t_B - t)e^{-\mu t}, \quad t \lesssim t_B. \quad (22)
\]

The decay rate of the white area is completely dominated by the walls. The central mixing process is potentially more efficient \( (\lambda > \mu) \), but it is starved by the boundaries.

If \( \mu > \lambda \), we have \( t > t_B \) in Eq. (20), since newly injected strips reach the Batchelor length before strips that were injected previously. This violates our assumptions, and we
conclude that in that case the white strips can be neglected; the decay rate of the concentration variance is then given by the natural decay rate \( \lambda \).

As an example of an exponential approach to the wall, consider the velocity field near a free-slip boundary,

\[
\begin{align*}
    u(\theta, y) &= u_0(\theta) + O(y), \\
    v(\theta, y) &= -u_0'(\theta)y + O(y^2),
\end{align*}
\]

(23a, 23b)

which satisfies the incompressibility constraint. Here \( u \) is the direction parallel to the wall, and \( v \) is perpendicular to the wall. The perpendicular distance from the wall is \( y \), and \( \theta \) is an angle around the circular boundary. (Since the dynamics near the wall are slow, we can use a steady flow here to model the time-\( T \) Poincaré map.) A separatrix is a distinguished streamline that ends at the boundary at some position \( \theta = \theta_s \). Along a separatrix at \( \theta = \theta_s \), we have \( u_0(\theta_s) = 0 \) since the velocity field changes sign. The rate of approach along the separatrix is thus given by \( \dot{y} = v(\theta, d) = -u_0'(\theta_s)d \), so that \( \mu = u_0'(\theta_s) \). Hence, if \( u_0'(\theta_s) > \lambda \) the rate of decay of concentration variance will not be limited by wall effects.

D. Algebraic approach to a no-slip wall

If the fluid at the wall is subject to no-slip boundary conditions, the Taylor expansion Eq. (23) is modified to become

\[
\begin{align*}
    u(\theta, y) &= u_1(\theta)y + O(y^2), \\
    v(\theta, y) &= -\frac{1}{2}u_1'(\theta) y^2 + O(y^3).
\end{align*}
\]

(24a, 24b)

The rate of approach along the separatrix at \( \theta = \theta_s \) is given by \( \dot{d} = v(\theta_s, d) = -\frac{1}{2}u_1'(\theta_s)d^2 \), with asymptotic solution \( d(t) \sim 2/u_1'(\theta_s)t \), for \( d(0)u_1'(\theta_s)t \gg 1 \). This is independent of the initial condition \( d(0) \): asymptotically, a fluid particle forgets its initial position; this explains why material lines bunch up against each other faster than they approach the wall, as reflected by the front in the upper part of Fig. 8(a). The total area of remaining white strips at time \( t \) as given by Eq. (19) is proportional to

\[
A_w(t) = \frac{2}{u_1'(\theta_s)} - \frac{2}{u_1'(\theta_s)} \frac{t - \tau}{\tau t}.
\]

(25)

The width of injected strips is \( \Delta(t) = -2Td/u_1'(\theta_s)t^2 \). Equation (17) cannot be solved exactly, but since \( \tau(t) \) is algebraic its right-hand side is not large, implying that \( t/\tau \approx 1 \) for large \( t \). We can thus replace \( \tau \) by \( t \) in Eq. (17) and the denominator of Eq. (25), and find

\[
A_w(t) \approx \frac{2}{u_1'(\theta_s)} \frac{\log[\Delta(t)/L_B]}{\lambda t^2}, \quad \frac{1}{d(0)u_1'(\theta_s)} \ll t \ll t_B.
\]

(26)

Compare this to the exponential case of Eq. (22): the decay of concentration variance is now algebraic \((1/t^2)\), with a logarithmic correction. The form of Eq. (26) has been verified in experiments using a simple map model (Gouillart et al., 2007, 2008).

E. Dynamics near a moving no-slip wall

Now consider the case of a rotating wall, where we add a constant speed \( \Omega > 0 \) to the velocity \( u \) in Eq. (24). Again we look for fixed points: all the parabolic fixed points on the wall have disappeared, as well as the two separatrices. Since \( u_1(\theta) \) is continuous, has two zeros, and \( u_1'(0) > 0 \), \( u_1(\theta) \) must have a minimum at some angle \( \theta_s \), where \( u_1'(\theta_s) = 0 \) and hence \( v(\theta_s, y) = 0 \) for all \( y \). Enforcing that the along-wall velocity also vanishes, there will be a fixed point at \( y_s = -\Omega/u_1(\theta_s) \).

Now we look at the linearized dynamics near the fixed point. Let \( (\theta, y) = (\theta_s + \Theta, -\Omega/u_1(\theta_s) + Y) \); then

\[
\begin{align*}
    \dot{\Theta} &= u_1(\theta_s) Y + O(\Theta^2, Y^2, \Theta Y), \\
    \dot{Y} &= -\frac{1}{2}u_1''(\theta_s) y^2 \Theta + O(\Theta^2, Y^2, \Theta Y).
\end{align*}
\]

(27a, 27b)

The linearized motion thus has eigenvalues \( \lambda_{\pm} = \pm \lambda = \pm \sqrt{-u_1''(\theta_s)/2u_1'(\theta_s)U} \), where the argument in the square root is non-negative since \( u_1(\theta_s) < 0 \) and \( u_1''(\theta_s) \geq 0 \). For \( u_1''(\theta_s) > 0 \) and \( \Omega > 0 \), this is a hyperbolic fixed point, and the approach along its stable manifold is given by \( Y(t) \sim Y_0 \exp(-\lambda t) \) for \( (\Theta_0, Y_0) \) initially on the stable manifold. Compare this to the \( 1/t \) approach for a fixed wall: the approach to the fixed point is now exponential, at a rate proportional to the speed of rotation of the wall. One expects that this exponential decay will dominate if it is slower than the mixing rate in the bulk. Otherwise, if \( \lambda \) is large enough, then the rate of mixing in the bulk dominates.

Figure 9(a) shows a numerical simulation of the flow pattern for a wall rotating at a rate \( \Omega T = 0.2 \). The hyperbolic fixed point is indicated by a dot, as is the distance \( d(t) \) between the mixing pattern and the hyperbolic point. Note the unmixed region between the rotating wall and the mixing pattern. Figure 9(b) is a Poincaré section that shows the presence of the unmixed region, which consists of closed orbits. Numerical simulations have confirmed that the decay rate of a passive scalar in the central region is indeed

\[ \text{FIG. 9. Rod mixer: (a) Numerical simulation of dye advection for a wall rotating at velocity } \Omega T = 0.2, \text{ with the rod moving in a figure-eight pattern. (b) Poincaré section, which shows a large chaotic region and closed orbits near the wall, with a separatrix in between. From Thiffeault, Gouillart, and Dauchot, 2011.} \]
exponential, so the rotating wall can help recover exponential mixing (Thiffeault, Gouillart, and Dauchot, 2011). However, the price to pay is that there is now an unmixed region surrounding the region of good mixing. Whether this is a price worth paying depends on the specific application.

Another strategy to mimic a moving wall, and thus recover exponential mixing, is to move the rod in a looping “epitrochoid” motion, shown in Fig. 10. This motion creates closed trajectories near the wall, as is evident in the Poincaré section, Fig. 10(b). Thiffeault, Gouillart, and Dauchot (2011) verified experimentally that the decay of the passive scalar in this case is indeed exponential, for the same reason as for the moving wall. However, the analysis of the near-wall map for this system is more complicated than for a moving wall and 3D effects also remain to be investigated: these could hold some surprises, since the nature of separatrices at the wall is potentially much richer.

F. Geometric mixing

Protocols aimed at efficient mixing of fluids heavily rely upon the generation of a chaotic kinematic template. In most cases, mixing protocols can be designed free of major geometrical constraints. In particular, no limitation is usually imposed upon the relative displacement of boundaries. But, can we achieve fluid mixing if we limit the allowed motion of boundaries to the subset returning to their original position after each iteration of the protocol? The answer to this question is related to the concept of geometric phases (Shapere and Wilczek, 1989a, 1989b): the failure of system variables to return to their original values after a closed circuit in the parameters.

In the zero Reynolds number limit, fluid inertia is negligible, fluid flow is reversible, and an inversion of the movement of the walls leads, up to perturbations owing to particle diffusion, to unmixing, as Heller (1960) and Taylor (1960) demonstrated. This would seem to preclude the use of reciprocating motion to stir fluid at low Reynolds numbers; it would appear to lead to perpetual cycles of mixing and unmixing. But, is that always the case? Can cyclic changes in the shape of the containers lead to efficient mixing?

The well-known 2D mixer based on the journal-bearing flow (Aref and Balachandar, 1986; Chaiken et al., 1986; Ottino, 1989; Tabor, 1989) may be used as an example of how nonreciprocal cycling of the deformable boundaries of a container can be used as a tool for fluid mixing at low Reynolds number (Arrieta et al., 2015); see Fig. 11. Considering as parameters in this device the positions of the outer and inner cylindrical walls of the container specified, respectively, with the angles $\theta_1$ and $\theta_2$ from a given starting point, a geometric phase might arise from driving this system around a loop in the parameter space. In a Heller-Taylor-type unmixing demonstration the parameter loop is very simple: $\theta_1$ first increases a certain amount and then decreases the same amount while $\theta_2$ remains fixed. This loop encloses no area, and reversibility ensures that the phase is zero. To obtain a finite-area nonreciprocal contractible loop we can, for instance, rotate first one cylinder, then the other, then reverse the first, and finally reverse the other. If we now perform a parameter loop by the sequence of rotations, we arrive back at our starting point from the point of view of the positions of the two cylinders, so it is, perhaps, surprising that the fluid inside does not return to its initial state. We illustrate the presence of this geometric phase in Fig. 11(a) in which an example of the trajectory of a fluid particle is shown as the walls are driven through a nonreciprocal contractible loop.

The long-term fluid dynamics elicited by a repeated realization of the same contractible nonreciprocal loop is shown in Fig. 11(b). A single fluid particle has covered most
of the area available to it between the two cylinders. This is fluid mixing induced entirely by a geometric phase; we call it geometric mixing. Geometric mixing therefore creates chaotic advection. The geometric phase scales with the area of the parameter loop and it is independent, at least for small enough loops, of the specifics of the trajectory in parameter space (Fig. 11(c)).

This example illustrates a general class of protocols in which mixing arises as a consequence of a geometric phase induced by a contractible nonreciprocal cycle in the parameters defining the shape of the container. It turns out that the mixing efficiency estimated from the stretching of material lines is roughly proportional to the geometric phase. Mixing in the corresponding flows can also be considered as the result of chaos arising in the mapping describing the motion of fluid elements during one cycle. When the cycle is reciprocal, this map is the identity and a small departure from reciprocity corresponds to a small departure from the identity map. Hence, the problem of mixing by nonreciprocal cycles is closely related to the class of dynamical systems constituted by perturbations of the identity (Arrieta et al., 2015). The structure of chaos in this class of dynamics has been greatly overlooked in the literature, which points to a much needed revisiting of this associated problem.

IV. THE NEW FRONTIER: 3D UNSTEADY FLOW

Most results to date of chaotic advection, indeed of dynamical systems in general, have been found by examination of 2D maps or flows where stable elliptic fixed points and the stable and unstable 1D manifolds of hyperbolic points define just a few Lagrangian coherent structures that control all of the transport behavior. However, in 3D there is an explosion of complexity in the number of possible Lagrangian structures and connections between them. This is due both to the impossibility of the existence of any stable fixed points and to hyperbolic manifolds existing as both sheets and curves. It is still an open question, especially in experiment, how these structures fit together to control mixing rates and the distribution of material and energy in a 3D stirred flow.

A. Motivation and background

Coherence and, intimately related to that, invariance are key notions in the investigation and classification of transport phenomena in (laminar) fluid flows. These notions can in general be defined in several ways (see, e.g., Sec. V). The discussion within this section adopts the Lagrangian perspective of the organization of fluid trajectories into coherent structures collectively defining the flow topology that geometrically determines the advective transport of material. It is important to note in this context that basically any group or union of fluid trajectories constitutes a material entity (Sec. V.A), suggesting a certain arbitrariness in the definition of coherent structures. An example of an (in general) nonunique entity is a stream surface in 3D steady flows: any material line advected unobstructed by the flow describes a stream surface. Consider to this end Poiseuille flows, where advection of any family of closed material curves released at the inlet yields a valid foliation into stream surfaces (Sec. VI.A). Connection with other properties or entities of the system renders coherent structures in the web of Lagrangian fluid trajectories unique. Structures tied directly to properties of the kinematic equation governing Lagrangian motion are, arguably, the most fundamental kind and include entities such as separatrices due to discrete symmetries, families of invariant surfaces due to continuous symmetries, 2D manifolds and tubes associated with closed streamlines and periodic lines and 1D and 2D manifolds associated with isolated stagnation and periodic points. Such structures constitute elements of the flow topology or, equivalently, the “ergodic partition” (Sec. VII.F). Coherent structures may also be defined indirectly as, for instance, material entities distinguished by Lyapunov exponents (Sec. V.A), topological deformation of enclosing material curves (Sec. V.D), or leakage from Eulerian regions. However, the following discussion concerns the coherent structures directly formed by the Lagrangian fluid trajectories.

Well-known examples of coherent structures of the “direct” kind are the KAM islands and unstable and stable manifolds of hyperbolic periodic points that constitute the flow topologies of 2D time-periodic flows in bounded domains (Sec. III). Lagrangian transport in other flow configurations has received considerably less attention and remains the subject of ongoing investigations. This section concerns one such class of configurations: 3D unsteady flows.

Current insight into the fundamentals of Lagrangian transport in 3D (un)steady flows is to a great extent based on kinematic properties of divergence-free vector fields and volume-preserving maps. This encompasses any incompressible unsteady flow $\mathbf{\nabla} \cdot \mathbf{u} = 0$ as well as any compressible steady flow $\mathbf{\nabla} \cdot (\rho \mathbf{u}) = \mathbf{\nabla} \cdot \mathbf{u}' = 0$ where $\mathbf{u}' = \rho \mathbf{u}$. Groundbreaking progress has come from the 3D extensions of the KAM theorem (Cheng and Sun, 1990a; Mezić and Wiggins, 1994; Broer, Huijtema, and Sevryuk, 1996) and the Poincaré-Birkhoff theorem (Cheng and Sun, 1990b) that, respectively, describe the fate of nonresonant invariant tori and resonant trajectories in 3D volume-preserving maps. The existence of 3D counterparts to these theorems was first hypothesized on the basis of a classification of volume-preserving maps by the number of action variables (Feingold, Kadanoff, and Vivo, 1987, 1988a, 1988b, 1989). Further important results include generic reductions in flow complexity by symmetries (Mezić and Wiggins, 1994; Haller and Mezić, 1998), the formation of invariant manifolds of various topologies due to constants of motion (Mezić and Wiggins, 1994; Haller and Mezić, 1998; Gómez and Meiss, 2002; Mullowney, Julien, and Meiss, 2005), and the local and global breakdown of invariant manifolds by resonances (Feingold, Kadanoff, and Vivo, 1988b; Cartwright, Feingold, and Vivo, 1994; Mezić, 2001a; Vainchtein, Neishtadt, and Mezić, 2006; Vainchtein, Widloski, and Grigoriev, 2007; Meiss, 2012), and universal properties of the Lagrangian transport between flow regions (MacKay, 1994; Lomeli and Meiss, 2009). These phenomena require in principle only satisfaction of continuity and compliance with certain kinematic conditions. However, whether a real fluid flow indeed admits the latter conditions (and the associated Lagrangian dynamics) depends essentially on momentum conservation.
Consider the 3D steady momentum equation once again
\[ \rho u \cdot \nabla u = -\nabla p + \mu \nabla^2 u, \quad (28) \]
where \( u \), as before, is the velocity, \( p \) is the pressure, \( \rho \) is the density, and \( \mu \) is the dynamic viscosity. We recast this for the present discussion into the alternative form
\[ \rho w \times u = -\nabla \zeta + \mu f, \quad f = \nabla \times w, \quad (29) \]
with \( w \equiv \nabla \times u \) the vorticity, \( w \times u \) the Lamb vector, and \( f \) the flexion field, representing inertia and viscous forces, respectively, and
\[ \zeta = p + \rho g z + e_k \quad (30) \]
the Bernoulli function (Yannacopoulos et al., 1998). Here \( e_k = \rho u \cdot u / 2 \) is the kinetic energy, and gravity is defined as \( g = -ge_z \). This form of the momentum equation directly reveals that it reduces from the 3D steady Navier-Stokes equation to the 3D steady Euler equation for both inviscid (\( \mu = 0 \)) and flexion-free (\( f = 0 \)) flows. 3D steady Euler flows are special in that universal conditions for (the absence of) chaos can be formulated on the basis of momentum conservation. It is well known that they admit chaos upon satisfying the Beltrami condition \( w \times u = 0 \); in all other cases they yield \( u \cdot \nabla \zeta = 0 \) and possess invariant manifolds defined by level sets of \( \zeta \) (Sec. VI.A). Moreover, these invariant manifolds are diffeomorphic to either cylinders or tori (Mezić and Wiggins, 1994; Arnol’d and Khesin, 1998). Hence, the Lagrangian dynamics can happen only in Euler flows (locally) meeting the Beltrami condition. Thus the latter facilitates (yet not per se causes) certain kinematic events. [See Arnol’d and Khesin (1998) for further properties of 3D steady Euler flows.] The Arnol’d-Beltrami-Childress (ABC) flow, for example, always satisfies the Beltrami condition and is the archetypal flow for many studies on 3D chaotic advection (Dombre et al., 1986; Feingold, Kadanoff, and Piro, 1988b; Cartwright, Feingold, and Piro, 1994; Haller, 2001a).

Irrotational flexion fields (\( \nabla \times f = 0 \)) imply \( f = -\nabla \sigma \), with \( \sigma \) the flexion potential, reducing Eq. (29) essentially to an Euler form \( \rho w \times u = -\nabla \zeta' \), with \( \zeta' = \zeta + \mu \sigma \) (Yannacopoulos et al., 1998). Here the Beltrami condition again determines the Lagrangian dynamics. Thus flexion-free flows, although strictly incorporating viscous effects, behave effectively as inviscid flows. It must be stressed that Stokes flows, although governed by \( \mu f = \nabla p \) and thus also meeting \( \nabla \times f = 0 \), are excluded from this behavior. Here the flexion potential equals \( \sigma = -p/\mu + c \), with \( c \) an arbitrary constant, implying \( \zeta' = c \) and, in consequence, the conservation of \( \zeta' \) tells us nothing. Hence, akin to a Beltrami flow, \( \zeta' \) is not a useful constant of motion. This point exposes an intriguing contrast: Euler flows satisfy \( \nabla \zeta' = 0 \) only in the exceptional Beltrami case; Stokes flows, on the other hand, invariably satisfy \( \nabla \zeta' = 0 \). Thus Euler flows generically are nonchaotic, while 3D Stokes flows have no obstacle to chaos. This observation has the fundamental implication that, lacking a universal dynamical restriction, Stokes flows can be integrable only due to symmetries.

Realistic 3D steady flows typically have significant inertia and viscosity, implying a rotational flexion field \([\mu \nabla \times f = \nabla \times (\rho w \times u) \neq 0]\), meaning that in general they are devoid of constants of motion \( u \cdot \nabla \zeta = \mu u \cdot f \neq 0 \) [see Kozlov (1993) for a rigorous discussion]. Here, similarly to Stokes flows, absence of a universal dynamical mechanism means integrability can ensue only from symmetries. These must yield a flexion field perpendicular to \( u \). It is important to note in relation to Euler flows that realistic fluid flows admit 3D chaos without satisfying the Beltrami condition. Thus Beltrami flows (e.g., the widely used ABC flow) may be too restrictive for general studies on 3D chaotic advection in realistic fluid flows (Mezić, 2002).

Laminar flows are for increasing Re progressively better described by the Euler limit of the momentum equation and, given that the Beltrami condition is exceptional, therefore typically tend to become integrable. (It must be stressed that laminar flow is assumed at all times here.) Significant viscous effects, for example, due to (local) breakdown of symmetries or boundary layers (locally), disrupt the integrability of the Euler approximation and thus promote, or at least facilitate, chaotic advection in high-Re (yet laminar) 3D steady flows (Yannacopoulos et al., 1998; Mezić, 2001b). Conversely, increasing Re tends to augment the Euler-flow region and, in consequence, to suppress 3D chaos. Realistic high-Re laminar flows thus generically lean toward a nonchaotic bulk flow; chaos, if occurring, tends to be confined to boundary layers and certain localized areas with symmetry breakdown.

Flows with low Re have significant viscous effects throughout the entire flow domain and, contrary to Euler flows and high-Re flows, in principle admit global chaos. Here chaos (or absence thereof) is intimately related to symmetries. The linearity of the momentum equation in the Stokes limit (\( \text{Re} = 0 \)) causes symmetries in geometry and boundary conditions to be imparted on the flow. Hence, Stokes flows, akin to Euler flows, often are integrable yet due to different mechanisms. Consider, for example, 3D lid-driven cavity flow inside cubic and cylindrical domains; here symmetries result in closed streamlines in the Stokes limit (Shankar, 1997; Shankar and Deshpande, 2000). Nonlinearity due to fluid inertia (\( \text{Re} > 0 \)) or asymmetry in geometry and/or flow forcing are necessary ingredients for chaos to occur in 3D steady viscous flows (Bajer and Moffatt, 1990, 1992; Shankar, 1998; Shankar and Deshpande, 2000). This discloses a remarkable difference with the high-Re regime in that here increasing Re promotes rather than suppresses chaos. This implies an essentially nonlinear dependence of the chaotic Lagrangian dynamics on Re. However, the complete story of the routes between the generically integrable states in the Stokes and Euler limits remains unexplored to date.

The terrain of 3D unsteady flows is even less charted than that of their steady counterparts. Here the left-hand side of the momentum equation (29) becomes augmented by an unsteady term \( \rho \partial u / \partial t \) and in principle any flow, including non-Beltrami Euler flow, is nonintegrable and admits 3D chaos. Integrability thus, reminiscent of 3D steady viscous flows, seems to hinge entirely on symmetries and linearity and for unsteady flows can in all likelihood be expected only in the Stokes limit. Studies on Lagrangian dynamics in realistic 3D unsteady fluid flows have to date been few and far between and restricted to time-periodic flows constructed by systematic reorientation of piecewise steady flows: the biaxial
unsteady spherical Couette flow (Cartwright, Feingold, and Piro, 1995, 1996), the cubic lid-driven cavity (Anderson et al., 1999, 2006), and the cylindrical lid-driven cavity (Malyuga et al., 2002; Speetjens, Clercx, and van Heijst, 2004; Pouransari, Speetjens, and Clercx, 2010).

The discussion hereafter concentrates on 3D unsteady flows and exemplifies typical behavior by way of the cylindrical lid-driven cavity. This system possesses a rich dynamics and thus enables a good demonstration of what may happen in this class of flows. Two topics are considered that, as in the simpler class systems previously discussed, are key to Lagrangian dynamics and 3D chaos: first, the role of symmetries in the integrability of the Stokes limit (Sec. IV.B); and second, the breakdown of this integrability by fluid inertia (Sec. IV.C). These phenomena are examined in terms of the formation of coherent structures and the associated freedom of motion for tracers. The discussion later on the cylindrical lid-driven cavity in essence concerns an overview and recapitulation of the main results of the separate studies by Malyuga et al. (2002), Speetjens, Clercx, and van Heijst (2004, 2006a, 2006b), Pouransari, Speetjens, and Clercx (2010), and Speetjens and Clercx (2013).

1. 3D square cylinder flow

Lagrangian features of 3D unsteady flows are exemplified by way of a simple yet realistic fluid flow: the time-periodic flow inside a 3D square cylinder $[r, \theta, z] = [0, 1] \times [0, 2\pi] \times [-1, 1]$ (Malyuga et al., 2002; Speetjens, Clercx, and van Heijst, 2004). The fluid is set in motion via time-periodic repetition of a sequence of piecewise steady translations ("forcing steps") of the end walls with unit velocity $U = 1$ and relative wall displacement $D = L/R$ ($L$ and $R$ are physical wall displacement and cylinder radius, respectively) by prescribed forcing protocols. Figure 12(a) shows a schematic of the flow configuration; forcing protocols are specified and are composed of the forcing steps indicated by the arrows. Highly viscous flow conditions are assumed such that transients during switching between forcing steps are negligible (i.e., $T_s/T_{step} \ll 1$, with $T_s = R^2/\nu$ the viscous time scale and $T_{step}$ the duration of one forcing step). Under this premise the internal flow consists of piecewise steady flows that are each governed by the nondimensional steady Navier-Stokes and continuity equations (2), $\text{Re} \ u \cdot \nabla u = -\nabla p + \nabla^2 u$ and $\nabla \cdot u = 0$. Nondimensionalization follows from substitution of the scaling $x = Rx$, $u = Uu'$, and $p = Pu'$ in Eq. (28), with primes indicating dimensionless variables (omitted for brevity). The characteristic pressure is given by $P = \mu U/R^2$ and ensues from assuming laminar-flow conditions dominated by a force balance between viscous forces and pressure gradient. Thus the Reynolds number appears before the inertial term and parametrizes perturbation of the Stokes limit.

The motion of passive tracers is governed by the kinematic equation (1), with formal solution $x(t) = \Phi(x_0)$ describing the Lagrangian trajectory of a tracer released at $x_0$. The corresponding Poincaré map [which has also been referred to as a Liouvillian map in the case of 3D volume-preserving flows$^2$ (Cartwright, Feingold, and Piro, 1994, 1995, 1996)] is defined by $x_{k+1} = \Phi(x_k)$, where $x_k = x(kT)$ is the tracer position after $k$ periods of the time-periodic forcing protocol.

The following forcing protocols (denoted protocols $A$, $B$, and $C$ hereafter) are considered:

$$\Phi_A = F_{B}^{+y}F_{B}^{+x}, \quad \Phi_B = F_{T}^{+x}F_{B}^{+y}, \quad \Phi_C = F_{B}^{+y}F_{T}^{+x}F_{B}^{+y}. \quad (31)$$

with subscripts in the forcing steps referring to the top ($T$) and bottom ($B$) end walls and superscripts indicating the translation direction [Fig. 12(a)]. All forcing steps are transformations of the base flow $F_{B}^{+y}$ according to

$$F_{B}^{+y} = \mathcal{F}_{\pi/2}(F_{B}^{+x}), \quad F_{T}^{+x} = \mathcal{S}_{z}(F_{B}^{+x}), \quad (32)$$

with $\mathcal{F}_{\alpha} : \theta \rightarrow \theta + \alpha$ and $\mathcal{S}_{z} : (x,y,z) \rightarrow (x,y,-z)$. Furthermore, the relative displacement is fixed at $D = 5$, leaving only $Re$ as a control parameter for each forcing protocol. Maps $\Phi_{A,B,C}$ as well as the underlying base flow $F_{B}^{+y}$ each exhibit particular 3D dynamics and thus serve to demonstrate fundamental aspects of 3D flows. Results have been obtained via numerical simulations (Pouransari, Speetjens, and Clercx, 2010).

2. Coherent structures in 3D systems

Coherent structures in the flow topology are spatial entities in the web of Lagrangian fluid trajectories that exhibit a certain invariance to the mapping $\Phi$. Four kinds [based on classifications in Guckenheimer and Holmes (1983) and Feingold, Kadanoff, and Piro (1988b)] can be distinguished in 3D time-periodic systems, defined by

$$\lambda^k = \Phi^k(\lambda^k), \quad (33)$$

with $\lambda^k$ constituting periodic points ($P^k$), periodic lines ($\mathcal{L}^k$), invariant curves ($\mathcal{C}^k$), and invariant surfaces ($\mathcal{S}^k$) of order $k$ (i.e., invariant with respect to $k$ forcing cycles). Note that periodic lines consist of periodic points, meaning that each constituent point is invariant; invariant curves and surfaces are only invariant as an entire entity. Brouwer’s fixed-point theorem states that any continuous mapping of a

---

$^2$The present class of divergence-free flows ($\nabla \cdot u = 0$) may in the literature alternatively be denoted “volume-preserving flows” or “solenoidal flows.”
convex space\(^3\) onto itself has at least one fixed point (Speetjens, Clercx, and van Heijst, 2004; Zeidler, 2012). This puts forward periodic points and associated coherent structures as the most fundamental building blocks of 3D flow topologies. Periodic points and lines fall within one of the following categories: node-type and focus-type periodic points and elliptic and hyperbolic periodic lines (Malyuga et al., 2002). (Periodic lines admit segmentation into elliptic and hyperbolic parts.) Isolated periodic points and hyperbolic and elliptic lines imply pairs of stable (\(W^s\)) and unstable (\(W^u\)) manifolds, arising as surface-curve pairs (\(W^s,W^u\)) for points and as surface-surface pairs (\(W^s,W^u\)) for lines. Elliptic lines form the center of concentric tubes. The 1D manifolds of isolated periodic points define invariant curves \(C^1\); 2D manifolds and elliptic tubes define invariant surfaces \(S^2\). Period-1 structures are the most important for the flow topology, as they determine the global organization. Higher-order structures are embedded within lower-order ones and thus concern ever smaller features. The following discussion thus is restricted to period-1 structures.

B. Degrees of integrability in 3D unsteady Stokes flows

Flows often accommodate symmetries due to the geometry of the flow domain and the mathematical structure of the governing conservation laws. Such symmetries, if present, play a central role in the formation of coherent structures and, inherently, in the spatial confinement of tracer motion. Symmetries in fact are the only mechanism that may accomplish integrability in 3D viscous flows (Sec. IV.A). In 2D time-periodic flows this typically results in symmetry groups of coherent structures or physical separation of flow regions by symmetry axes (Franjione, Leong, and Ottino, 1989; Ottino, Jana, and Chakravarthy, 1994; Meleshko and Peters, 1996). In 3D time-periodic flows this may furthermore suppress truly 3D dynamics (Feingold, Kadanoff, and Piro, 1988b; Mezić and Wiggins, 1994; Haller and Mezić, 1998; Malyuga et al., 2002; Speetjens, Clercx, and van Heijst, 2004). Such manifestations of symmetries are demonstrated below for the time-periodic cylinder flow in the noninertial limit \(Re=0\). The impact of fluid inertia is examined in Sec. IV.C.

The flow field \(u\) governed by Eq. (28) collapses to

\[
\begin{align*}
    u_r(x) &= u_r(r, z) \cos \theta, \\
    u_\theta(x) &= u_\theta(r, z) \sin \theta, \\
    u_z(x) &= u_z(r, z) \cos \theta,
\end{align*}
\]

in the noninertial limit. (The italic \(u\)’s refer to the actual velocity components of the 3D velocity; the roman \(u\)’s correspond with the part of each component that depends on \(r\) and \(z\).) This implies closed streamlines in the base flow \(F_B\) that are symmetric about the planes \(x=0\) and \(y=0\) (representing reflections) [see Fig. 12(b)] (Shankar, 1997) and, inextricably connected with that, two constants of motion of the generic form

\[
F_1(x) = f_1(r, z), \quad F_2(x) = f_2(r, z) \sin \theta, \quad \text{(35)}
\]

satisfying \(dF/dt = u \cdot \nabla F = 0\). [An analytical expression for \(F_1\) is given by Malyuga et al. (2002).] The properties of Eq. (35) have essential ramifications for the flow topologies of the forcing protocols of Eq. (31). Figure 13 offers some first insight into the dynamics by way of the Poincaré sections of a single passive tracer. Tracers released under protocol \(A\) [Fig. 13(a)] are confined to invariant spheroidal surfaces on which they perform effectively 2D (chaotic) dynamics. This occurrence of chaos on a submanifold of codimension one is an essentially 3D phenomenon; see, e.g., Gómez and Meiss (2002), Meier, Lueptow, and Ottino (2007), Mullowney, Julien, and Meiss (2005, 2008), and Sturman et al. (2008) for dynamically similar systems. Protocol \(B\) [Fig. 13(b)] restricts tracers to a quasi-2D (chaotic) motion within thin shells parallel to the \(y-z\) plane. Truly 3D (chaotic) dynamics covering the entire flow domain occurs only for protocol \(C\) [Fig. 13(c)]. These dramatic differences in dynamics signify the presence of geometric restrictions on the tracer motion, akin to the KAM islands and cantori of 2D flows (Sec. II.E), in protocols \(A\) and \(B\). This is a direct consequence of symmetries as we elaborate on.

These observations furthermore demonstrate that the quality integrability takes on a subtler meaning in 3D flows, in that various degrees of integrability—and, inherently, spatial confinement—can be distinguished, ranging from restriction to closed trajectories [base flow, Fig. 12(b)] to global 3D

\[
\text{FIG. 13. 3D square cylinder flow: Degrees of integrability of 3D unsteady Stokes flows demonstrated by Poincaré sections of single tracers for the time-periodic forcing protocols according to Eq. (31). (a) Protocol A, (b) protocol B, and (c) protocol C. From Speetjens, Clercx, and van Heijst, 2004.}
\]
chaotic advection [protocol C, Fig. 13(c)]. The classification of 3D time-periodic fluid flows introduced by Cartwright, Feingold, and Piro (1996) may be understood in terms of this notion of degrees of integrability. The cylinder flow in its Stokes limit encompasses all degrees of integrability in 3D time-periodic systems.

1. 2D (chaotic) dynamics within invariant manifolds

The restriction of tracers in protocol $\mathcal{A}$ to invariant surfaces arises from a hidden axisymmetry in the base flow that is retained by any forcing protocol involving reorientations of only one end wall. According to Eq. (35), the constant of motion $F_1$ is invariant under the continuous transformation $\mathcal{F}_\alpha: \theta \rightarrow \theta + \alpha$, with $0 \leq \alpha \leq 2\pi$, i.e., $\mathcal{F}_\alpha(F_1) = F_1$. The level sets of $F_1$ are defined by the surfaces of revolution of the trajectories $dr/dz = u_c(r,z)/u_c(r,z) = g(r,z)$ in the $r$-$z$ plane (Speetjens, Clercx, and van Heijst, 2006a). Figure 14(a) shows members of the infinite family of concentric spheroidal surfaces thus formed. Their emergence is entirely consistent with the generic property that a continuous symmetry in a bounded 3D steady flow (here the base flow) partitions the flow topology into a finite number of families of nested invariant tori or spheroids [see Theorem 4.1 in Mezić and Wiggins (1994)]. Furthermore, spheroidal invariant surfaces imply closed streamlines (Mezić and Wiggins, 1994). This explains the flow topology of the base flow in its Stokes limit [Fig. 12(b)].

Further organization of the flow topology of protocol $\mathcal{A}$ results from discrete symmetries arising from the base flow. Transformations Eq. (32) through $\Phi_\mathcal{A}$ following Eq. (31) translate into

$$
\Phi_\mathcal{A} = S_1 \Phi_\mathcal{A}^{-1} S_1, \quad \Phi_\mathcal{A} = \tilde{S} \Phi_\mathcal{A} \tilde{S},
$$

with $S_1: (x,y,z) \rightarrow (-y,-x,z)$, $\tilde{S} = S_2 F_B^{x,y}$, and $S_2: (x,y,z) \rightarrow (y,x,z)$. (Note that these symmetry operators are consistent with the continuous axisymmetry by acting only within a given invariant spheroid.) The time-reversal reflectional symmetry $S_1$ has the fundamental consequence that the flow must possess at least one period-1 line $\mathcal{L}_1$, viz., within the symmetry plane $I_1 = S_1(I_1)$ (plane $y = -x$) (Speetjens, Clercx, and van Heijst, 2004). Coexistence of $S_1$ with $\tilde{S}$ dictates that $\mathcal{L}_1$ be invariant to both discrete symmetries, i.e.,

$$
\mathcal{L}_1 = S_1(\mathcal{L}_1) = \tilde{S}(\mathcal{L}_1) = S_1 \tilde{S}(\mathcal{L}_1) = \tilde{S} S_1(\mathcal{L}_1),
$$

meaning they essentially shape the period-1 line and its associated structures. The curve in Fig. 14(a) outlines the period-1 line for $D = 5$, where heavy and normal parts indicate elliptic and hyperbolic segments, respectively. The hyperbolic segment is invariant to $\tilde{S}$; left and right elliptic segments form symmetry pairs related via $\tilde{S}$. Higher-order periodic lines are subject to a similar organization (Speetjens, Clercx, and van Heijst, 2006a). Experimental validation of periodic lines and their fundamental link with symmetries was discussed by Znaïen et al. (2012).

The intrasurface topologies within the invariant spheroids are organized by the periodic points defined by their intersection with the periodic lines. The segmentation into elliptic and hyperbolic parts results in multiple kinds of intrasurface topologies. Figure 14(b) shows an invariant spheroid that intersects with the hyperbolic segment of the period-1 line and exposes a topology that consists of two pairs of period-2 islands arranged around the two hyperbolic period-1 points (dots) and enclosed by their heteroclinically interacting manifolds (not shown). The stable ($W^s_i$) and unstable ($W^u_i$) manifolds of each period-1 point relate via $W^s_i = S_1(W^u_i)$; manifolds of the period-1 points interrelate via $W^u_1 = \tilde{S}(W^s_2)$ and $W^s_1 = \tilde{S}(W^u_2)$. (Note that the intrasurface manifolds merge into 2D manifolds $W^u_{25}$ in 3D space with the same symmetry properties.) The two pairs of period-2 islands correspond to two pairs of period-2 elliptic points (facing points make one pair) that are on elliptic segments of two period-2 lines intersecting the invariant surface. These entities relate, in a similar manner as their period-1 counterparts, via symmetries $S_1$ and $\tilde{S}$. Figure 14(c) shows an invariant spheroid that intersects with hyperbolic segments of said periodic lines, leading to fully chaotic intrasurface dynamics.

Periodic points within invariant spheroids belonging to periodic lines imply the map $\Phi_\mathcal{A}$ is locally area preserving in their proximity (Gómez and Meiss, 2002). This has the important implication that the tracer motion within each spheroid is essentially similar to that in 2D area-preserving maps. This underlies the 2D Hamiltonian intrasurface topologies in Fig. 14 with generic composition according to
Sec. IIE; compare Fig. 14(b) with Fig. 6. Imperative for the rich intrasurface dynamics here as well as in other 3D systems (Gomez and Meiss, 2002; Mullowney, Julien, and Meiss, 2005, 2008; Meier, Lueptow, and Ottino, 2007; Sturman et al., 2008) is the existence of multiple isolated periodic points of different type. This, through Brouwer’s fixed-point theorem, suggests that convexity of the invariant surfaces is a necessary (if not sufficient) condition. Absence of such complex dynamics in the classical case of invariant tori, which are nonconvex and generically accommodate a dense winding of a single trajectory, supports this conjecture. However, conclusive establishment of the conditions that admit intrasurface chaos remains outstanding.

2. Quasi-2D (chaotic) dynamics within subregions

An immediate consequence of including both top and bottom walls in \( \Phi_B \) is the vanishing of constant of motion \( F_1 \) and the associated continuous axisymmetry. Thus here tracers are, in contrast to protocol \( A \), no longer restricted to invariant surfaces. Protocol \( B \) nonetheless accommodates discrete symmetries, reading

\[
\Phi_B = \bar{S} \Phi_B^{-1} \bar{S}, \quad \Phi_B = \Phi_B^{-1} \Phi_B S, \quad \Phi_B = S_x \Phi_B S_y, \quad (38)
\]

with \( S_x \colon (x, y, z) \rightarrow (x, -y, z), \ \bar{S} = S_x F_B^+ S_x, \ \bar{S}' = (x, y, z) \rightarrow (-x, y, z), \ \bar{S}' = S_x \bar{S} S_z, \) and \( S \) as before. Note that symmetries \( \bar{S} \) and \( S \), and corresponding symmetry planes \( \bar{I} = \bar{S}(I) \) and \( I' = S(I) \), are conjugate in that they relate via \( S_x \); the latter is a time-reversal symmetry hidden in \( \bar{S} \) and \( S' \). Time-reversal symmetry again implies at least one period-1 line \( L^1 \) within the corresponding symmetry plane. However, coexistence of two such symmetries imposes an additional restriction compared to protocol \( A \) in that period-1 lines identify with the intersections \( L^1 \in \bar{I} \cap I' \) of the conjugate symmetry planes so as to belong simultaneously to both (Speetjens, Clercx, and van Heijst, 2004). The symmetry of \( \bar{I} \) and \( I' \) about \( z = 0 \) implies organization of period-1 lines into the group

\[
M_\bar{B} = \{ \{ L_{z,1}^1, L_{z,2}^1, \ldots, L_{z,m}^1 \}, \{ L_{z,1}^2, S_z(L_{z,1}^2) \}, \ldots, \{ L_{z,m}^2, S_z(L_{z,m}^2) \} \}, \quad (39)
\]

with \( L_{z,i}^2 = S_z(L_{z,i}^1) \) (i \( \in \) \( [1, m] \)) period-1 lines within \( z = 0 \) and \( \{ L_{z,1}^2, S_z(L_{z,1}^2) \} \) symmetry pairs about \( z = 0 \). Note that all period-1 lines possess the self-symmetry \( L_{z,i}^1 \equiv S_z(L_{z,i}^1) \) and \( L_{z,1}^1 \equiv S_z(L_{z,1}^1) \) about the plane \( y = 0 \). Thus here a symmetry group of period-1 lines forms; for protocol \( A \) a single period-1 line consisting of symmetric segments forms.

Mass conservation implies at least one period-1 line that, for given symmetries, must sit in the plane \( z = 0 \), meaning that \( L_{z,1}^1 \) always exists. Figure 15(a) shows a typical symmetry group \( M_\bar{B} = \{ L_{z,1}^1, L_{z,2}^1, S_z(L_{z,1}^1) \} \) of period-1 points, with \( L_{z,1}^1 \) fully hyperbolic and the conjugate pair \( L_{z,1}^2, S_z(L_{z,1}^2) \) fully elliptic. The stable and unstable manifolds \( W^u/W^s \) of line segment \( p_1 - p_2 \) of \( L_{z,1}^1 \), relating via \( W_{2D}^u = S_z(W_{2D}^u) \), are shown in Fig. 15(b) and envelop the elliptic region comprising concentric tubes (the 3D counterparts to KAM islands) centered on the elliptic lines (not shown). The manifolds extend primarily in the direction normal to \( L_{z,1}^1 \) (i.e., parallel to the \( y-z \) plane) and exhibit only marginal \( y \)-wise variation. Moreover, they exhibit transversal interaction, which is a fingerprint of chaotic dynamics in 2D systems. This manifold behavior causes the quasi-2D chaotic tracer motion within a thin layer normal to the period-1 lines, as illustrated in Fig. 13(b). Tracers released near the elliptic segments exhibit similar behavior by remaining confined to thin slices of elliptic tubes (not shown). Hence, tracer dynamics within each layer is of a basically Hamiltonian nature and is thus intimately related to the intrasurface dynamics of protocol \( A \). A primary difference with the latter is that tracers are not strictly confined to an invariant surface. Whether these less restrictive conditions may be of any consequence is an open question. Protocols \( A \) and \( B \) thus reveal that time-reversal symmetries, through their link with periodic lines, imply effectively (quasi-)2D dynamics. This suppression of truly 3D dynamics is an essentially 3D manifestation of this kind of symmetries; their role in 2D systems primarily concerns formation of symmetry groups.

3. Unrestricted 3D (chaotic) dynamics

Inclusion of a third forcing step results in a flow that is devoid of global symmetries. This in principle paves the way...
to 3D chaotic advection for viscous flows. In particular the absence of time-reversal symmetries is of fundamental consequence in that periodic lines must thus no longer be present. However, the current flow must, according to Brouwer’s fixed-point theorem, accommodate at least one isolated period-1 point (Sec. IV.A.2). Two node-type period-1 points indeed exist and have associated manifold pairs ($W_{2D}^s-W_{1D}^u$) with essentially 3D foliations that densely fill the entire flow domain [Fig. 15(c)]. (The 2D manifolds have a dominant stretching direction and thus assume a particular shape.) Here the stable and unstable manifolds, in contrast with those associated with periodic lines, are not related via a time-reversal symmetry. This asymmetry in time results in essentially 3D transport and is a key element in the truly 3D chaotic dynamics demonstrated in Fig. 13(c). Moreover, the role of manifolds is fundamentally different here compared with protocols $A$ and $B$. The stability properties of the manifolds admit only homoclinic or heteroclinic $W_{2D}^s-W_{1D}^u$ interactions; any interactions between the 2D manifolds are impossible on grounds of identical stability. [Figure 17(a), for example, shows a homoclinic transversal $W_{2D}^s-W_{1D}^u$ interaction for protocol $A$ at Re = 100.] Furthermore, interactions between 2D manifolds of isolated periodic points can generically happen only through either merger into heteroclinic surfaces or formation of a heteroclinic orbit connecting both points by transversal intersection. 3D counterparts to transversal manifolds is fundamentally different here compared with 2D manifolds of isolated periodic points.4 This suggests that, contrary to the (quasi-)2D dynamics associated with periodic lines, the intrinsic hyperbolicity of isolated period-1 points generically suffices for 3D chaotic advection to occur (Speetjens, Clercx, and van Heijst, 2004). Formation of transport barriers by heteroclinic merger of 2D manifolds is the sole nonchaotic case. However, that is atypical case in 3D unsteady flows.

C. Breakdown of invariant manifolds by fluid inertia

Fluid inertia $Re > 0$ has a strong impact upon transport properties. It eliminates any symmetry that hinges upon linearity of the Stokes limit of the momentum equation and may thus facilitate (if not directly cause) 3D chaotic advection (Bajer and Moffatt, 1992; Cartwright, Feingold, and Piro, 1996). In our example, the cylinder flow, inertia introduces a secondary circulation to the base flow $F_B^{x,y}$ transverse to its primary circulation in the noninertial limit [Fig. 16(a)]. This breaks the symmetry about $x = 0$ and causes the closed streamlines to become nonclosed and wrapped around invariant tori defined by the level sets of a new constant of motion $F$ that emerges in favor of constants of motion $F_{1,2}$ (Cartwright, Feingold, and Piro, 1996). (The symmetry of the base flow about $y = 0$ is preserved for $Re > 0$.) These tori, in turn, undergo a progressive disintegration into tori with winding numbers\(^5\) $W > 1$ [island chains in Fig. 16(b)] and chaotic seas [Fig. 16(c)] with increasing Re, according to the Hamiltonian response scenario for 2D time-periodic systems to perturbations and described by the KAM and Poincaré-Birkhoff theorems (Ottino, 1989; Cartwright, Feingold, and Piro, 1996). The cross sections in Fig. 16 are dynamically equivalent to said 2D systems (Bajer, 1994; Cartwright, Feingold, and Piro, 1996); compare their structure with Fig. 5.

1. Response of invariant tori

Toroidal invariant surfaces in the Poincaré maps of 3D time-periodic flows exhibit similar behavior in that they also partly survive (inertia-induced) perturbations; this results in Poincaré sections of composition similar to the 3D streamline portrait in Fig. 16 yet typically comprising various arrangements of multiple families of tori (Dombre et al., 1986; Feingold, Kadanoff, and Piro, 1987, 1988b; Cartwright, Feingold, and Piro, 1994). This survival of invariant tori is described by a 3D counterpart to the KAM theorem (Cheng and Sun, 1990a; Mezić and Wiggins, 1994; Broer, Huitema, and Sevryuk, 1996). Invariant tori accommodating resonant (i.e., closed)...

---

\(^4\)The periodic points by definition are asymptotic limits of any transversal intersections. Topological consistency then admits two situations: (i) periodic points are part of the intersection (i.e., the heteroclinic orbit), and (ii) intersections are isolated points ($W_{2D}^s-W_{1D}^u$ interactions).

\(^5\)The winding number or rotation number $W$ represents the number of revolutions around the axis of rotation required for completing a full loop on a closed trajectory. The closed streamlines in Fig. 12(b), for example, have $W = 1$. 
trajectories with $W > 1$ disintegrate into tori with $W > 1$ in a way similar to the 3D steady case [Fig. 16(b)] (Cartwright, Feingold, and Piro, 1996). The fate of such resonant trajectories is described by the 3D counterpart to the Poincaré-Birkhoff theorem (Cheng and Sun, 1990b).

A special case exists in perturbed time-periodic systems near the limiting case $W = 1$. Trajectories in the Poincaré map in general wrap themselves densely around invariant tori in a manner akin to the perturbed base flow (Fig. 16). However, so-called resonant sheets, material surfaces in the unperturbed flow composed of period-$p$ points with $p \geq 1$, interrupt this process and, if occurring, cause local defects in the tori by which tracers can randomly jump between their intact sections. This behavior is termed resonance-induced dispersion (or resonance-induced diffusion) and facilitates global tracer distribution (Feingold, Kadanoff, and Piro, 1988b; Piro and Feingold, 1988; Cartwright, Feingold, and Piro, 1994, 1995, 1996). Resonance-induced dispersion is extremely slow compared to truly 3D chaotic advection, since tracers are confined to segments of tori in between jumps. Tracers nonetheless describe space-filling trajectories that visit the entire domain in the course of time. The dynamics within the jump zones is governed by isolated periodic points that remain from the (originally entirely periodic) resonance sheets upon perturbation. The associated manifold pairs $(W_{2D}^u, W_{1D}^u)$ intersect these zones transversally and thus enable the jumping between segments of tori that underlies resonance-induced dispersion (Mezić, 2001a).

2. Response of invariant spheroids

Studies on the responses of invariant surfaces to perturbations are almost exclusively restricted to tori. However, the classification theorem for closed surfaces states that any orientable closed surface, which includes level sets of a constant of motion in bounded flows, is topologically equivalent to a sphere or a connected sum of tori (Alexandroff, 1961). This puts forward invariant spheroids, besides invariant tori, as a second fundamental form of invariant surfaces relevant in the present context (Speetjens, 2001); recall in this regard that tori and spheroids are also the key invariant surfaces in 3D steady flows (Sec. IV.B.1). Essentially different dynamics upon perturbation are likely on grounds of fundamental topological properties. Tori are doubly connected nonconvex manifolds, and spheroids are singly connected convex manifolds. Tori accommodate trajectories that are either closed or densely wrapped, and spheroids typically accommodate chaotic seas and elliptic islands. Model flows offer a way to investigate the response of invariant spheroids to inertial perturbations under realistic conditions. This may contribute to a more complete picture of the fate of invariant surfaces subject to inertial perturbations.

The effect of fluid inertia upon invariant spheroids can be demonstrated for protocol $\mathcal{A}$; this is representative of generic forcing protocols with such a topology (Pouransari, Speetjens, and Clercx, 2010). Inertia breaks both the time-reversal reflectional symmetry $S_1$ in Eq. (36) and the continuous axisymmetry due to constant of motion $F_1$; only symmetry $\tilde{S}$ is preserved for $Re > 0$ (Speetjens, Clercx, and van Heijst, 2006a). This causes the period-1 line $L^1$ [Fig. 14(a)] to give way to a focus-type isolated period-1 point with a $(W_{2D}^u, W_{1D}^u)$ manifold pair that, for sufficiently high $Re$, completely destroys the invariant spheroids and, in consequence, yields 3D chaotic tracer motion. Figure 17 demonstrates this process for $Re = 100$, where, consistent with permissible manifold interactions of isolated periodic points (Sec. IV.B.3), homoclinic transversal $W_{2D}^u-W_{1D}^u$ interaction occurs. However, even minute departures from the noninertial limit may change the flow topology drastically. This is discussed next.

Secondary circulation causes progressive drifting of tracers transverse to the invariant spheroids that grows stronger with increasing $Re$. This is demonstrated in Fig. 18 by means of the $r$-$z$ projection of a Poincaré section of a single tracer for $10^4$ forcing periods. The drifting tracers remain confined within thin shells centered upon the invariant spheroids for $Re \lesssim 0.1$ [Figs. 18(a) and 18(b)] for time spans of $O(2 \times 10^4)$ periods due to the averaging out of transverse excursions (“averaging principle”) (Arnol’d, 1978). Thus invariant spheroids survive in an approximate way as so-called adiabatic shells (Speetjens, Clercx, and van Heijst, 2006a). However, this survival occurs only for regions with chaotic intrasurface dynamics. This puts forth the rather intriguing notion that intrasurface chaos promotes the persistence of partial transport barriers, or equivalently that 2D chaos suppresses the onset of 3D chaos. The formation of a complete adiabatic shell in Fig. 18(b) thus signifies an underlying invariant spheroid with fully chaotic tracer motion. Chaotic and nonchaotic regions of invariant spheroids [e.g., shown in Fig. 14(b)] transform into incomplete adiabatic shells and elliptic tubes, respectively, that merge into intricate adiabatic structures by a mechanism termed resonance-induced merger (Speetjens, Clercx, and van Heijst, 2006a, 2006b). Figure 19 shows an adiabatic structure formed by a resonance-induced merger, comprising an inner and outer adiabatic shell, connected via an elliptic tube on each elliptic segment of the period-1 line for $Re = 0.1$. Both tubes, similarly to the underlying elliptic segments of the period-1 line of the noninertial limit, form a symmetry pair related via $\tilde{S}$. A resonance-induced merger thus results in a
family of nested closed adiabatic structures that are topologically equivalent to tori. A similar adiabatic structure due to a resonance-induced merger is shown in Fig. 20, but with the following two differences: (i) connecting tubes undergo a bifurcation from period-2 \((P_2)\) to period-4 \((P_4)\) structures; (ii) inner \((A_i)\) and outer \((A_o)\) shells are “leaky.” This leakiness causes tracer exchange with a chaotic environment (not shown) and thus sets up a net tracer circulation through the structure from outer to inner shell. This temporary tracer entrapment within a coherent structure \([\text{residence time } O(10^3) \text{ periods}]\) bears a certain resemblance to prolonged confinement in the chaotic saddle region (Sec. II.A) and stickiness near cantori (Sec. II.E) in 2D (open) flows.

This exposes a fundamental difference in response scenarios of invariant tori and spheroids. The case of tori seems fairly well understood, although a complete rigorous description of the breakdown and survival mechanisms remains outstanding. The dynamics of perturbed invariant spheroids, on the other hand, remains largely an open problem. Fully chaotic spheroids survive weak inertia as complete adiabatic shells and constitute transport barriers akin to those of the noninertial limit. Nonchaotic regions on subfamilies of invariant spheroids have fundamental ramifications by causing the formation of intricate adiabatic structures through a resonance-induced merger. Its occurrence for a wide range of forcing protocols \((\text{Speetjens, Clercx, and van Heijst, 2006a, 2006b; Pouransari, Speetjens, and Clercx, 2010})\) and different systems \((\text{Moharana et al., 2013})\) suggests that resonance-induced merger is a universal phenomenon and is part of an essentially 3D route to chaos. The study by \textit{Wu et al. (2014)} supports this assertion by demonstrating that a resonance-induced merger can in fact be triggered by weak perturbations of arbitrary nature and provides the first experimental evidence of its physical existence.

Resonances as a key mechanism in response scenarios are the common denominator for tori and spheroids. However, their emergence and manifestation are very different. Known resonances on tori in time-periodic systems include isolated closed trajectories and, in the case of the very dense windings underlying resonance-induced dispersion, resonant sheets: material surfaces of period-\(p\) points. The former cause disintegration of the entire torus in question; the latter break up into disconnected segments. Whether further types may exist is open. Topological considerations suggest strings of isolated period-\(p\) points as a special case of isolated closed
trajectories (i.e., a localized counterpart to resonant sheets). Resonances on spheroids emerge as periodic points belonging to periodic lines in the 3D domain. Resonance-induced merger seems to rely, in particular, on isolated degenerate periodic points connecting elliptic and hyperbolic segments of periodic lines. Moreover, instead of causing disintegration and chaos, resonance-induced merger reorganizes and reorders the flow topology through the formation of distinct coherent structures. Resonance-induced dispersion and resonance-induced merger nonetheless exhibit a remarkable similarity that hints at a comparable mechanism. Tracer switching effectuated by the isolated periodic points inside the perturbed resonant sheets during resonance-induced dispersion, demonstrated by Mezić (2001a) and supported by experiments in Solomon and Mezić (2003), bears a great resemblance to the migration of tracers from an adiabatic shell into an elliptic tube in the case of resonance-induced merger. Further explorations are needed and it may be of interest to compare other studies on resonance phenomena (Litvak-Hinenzon and Rom-Kedar, 2002; Vainchtein, Neishtadt, and Mezić, 2006).

V. CHAOTIC ADVECTION FROM DATA

An accurate prediction of material transport in geophysical flows is growing in importance in the aftermath of large-scale catastrophic events such as the volcanic eruption of Eyjafjallajökull (Iceland, 2010), the Deepwater Horizon oil spill (Gulf of Mexico, 2010) [see, e.g., Domingos and Cardoso (2013)], and the nuclear disaster in Fukushima (Japan, 2011) [see, e.g., Domingos and Cardoso (2015)]. The main avenues for the material transport have been connected to the skeleton of chaotic advection, and detecting these routes from measured data in (near) real time has become one of the major goals at the intersection of dynamical systems and fluid dynamics. Most approaches use data in the form of numerical vector fields, interpolated and extrapolated from telemetry, e.g., from high-frequency radar and satellite altimetry, although Lagrangian sensor data are also available and studied. The approaches reviewed here are aimed at defining and computing skeletons of chaotic advection directly from available data, without intermediary equation-based models.

A. Lagrangian coherent structures

As we have seen, invariant manifolds play a central role in the analysis of time-invariant velocity fields as they act as barriers for material transport. If in addition an invariant manifold attracts or repels nearby material, it organizes behavior of the flow in its vicinity. Since invariant manifolds align locally with eigenvectors of the linearized velocity field, methods based on flow linearization are able to accurately compute routes for material transport in time-invariant flows. For time-varying flows, however, manifolds that align with eigenvectors of linearized flow do not act as barriers to material transport. Therefore, a different strategy is needed to find lower-dimensional structures that organize the flow.

All material lines advected by the flow are barriers to material transport. In informal terms, Lagrangian coherent structures are those material lines that strongly attract or repel the neighboring material. The initial operational definition specified attracting (respectively, repelling) LCSs as material lines that are linearly stable in forward (respectively, backward) time for a longer period than any of their neighbors (Haller, 2000, 2001a, 2001b, 2002; Haller and Yuan, 2000). LCSs were initially approximated by ridges of finite-time Lyapunov exponent (FTLE) and finite-size Lyapunov exponent (FSLE) fields, which were assumed to exist near true LCSs (Joseph and Legras, 2002; Shadden, Lekien, and Marsden, 2005; Haller and Sapsis, 2011; Tallapragada and Ross, 2013). FTLEs and FSLEs measure the rate of separation between nearby trajectories, which is associated with the linear stability of those trajectories. Since FTLE or FSLE fields are relatively easy to compute and understand, they were quickly adopted as proxies to the routes of material transport in geophysical studies, e.g., by Rypina et al. (2007), Shadden et al. (2009), and Bettencourt, Lopez, and Hernandez-Garcia (2012), but also in analyses of mixing, e.g., by Lukens, Yang, and Fauci (2010). Postprocessing and the use of FTLE or FSLE fields remains a challenge as they may indicate the presence of LCSs where there are none (Karrasch and Haller, 2013). In certain cases, FTLE or FSLE analysis is justified only in a qualitative, but not quantitative sense (BozorgMagham, Ross, and Schmale, 2013). Nevertheless, the ease of interpretation of FTLE fields means that they remain an active area of research in experimental studies (Raben, Ross, and Vlachos, 2013).

Lagrangian coherent structures have recently been redefined with the requirements that the defined structures be objective with respect to the reference frame, observable over finite-time periods, coherent under Lagrangian transport, and spatially continuous, as detailed by Haller (2015). Under these requirements, a variational principle for LCSs has been formulated, which is then used as an operational procedure for calculating the LCS (Haller, 2011; Haller and Beron-Vera, 2012; Blazevski and Haller, 2014). According to the variational definition, LCSs are shadowed by minimal geodesics induced by the Riemannian metric of the Cauchy-Green strain tensor $C(x) = \nabla \Phi(x)^\top \nabla \Phi(x)$, i.e., the “absolute-value squared” of the Jacobian matrix $\nabla \Phi(x)$, of the flow map. Figure 21 shows an example of the invariant manifolds found in this manner in a jet flow. The theory presently covers a broad array of LCSs in 2D and 3D velocity fields, classified according to the type of deformation they exert on the surrounding material (Haller, 2015).

Compared to FTLE or FSLE computations, the variational definition results in a computationally more involved procedure. The algorithm has been made more accessible through a freely available LCS toolbox for MATLAB.\(^6\) As an example of the use of this approach, we highlight detection of an elliptic LCS that matches a physically observed phenomenon known as the Agulhas ring, associated with the transport of cold water from the southern tip of Africa (Beron-Vera et al., 2013) (see Fig. 22). A recent detailed review of the Lagrangian coherent structure theory (Haller, 2015) provides further references to studies that have applied variational LCSs to the analysis of geophysical flows, now casting and forecasting, as well as discussions of numerical considerations associated with the technique.

\(^6\)http://georgehaller.com/software/software.html.
B. Coherent sets

Instead of looking at barriers to transport, as is the case with Lagrangian coherent structures, a skeleton of chaotic advection can be proposed to be a collection of regions from which tracer material does not leak out. Such collections are routes through which the material travels, therefore serving as a road map for transport.

For a tracer distributed in the set \( A \) at initial time \( t_0 \), \( \Phi_t(A, t_0) \) represents the set containing the tracer after it is advected by the flow over the time interval \( t \). The sets \( C_0 \) (source) and \( C_t \) (target) are a coherent pair if the tracer placed in the source fills the target entirely, without leaking, that is

\[
\mu(\Phi_t(C_0, t_0) \cap C_t) \approx \mu(C_t), \tag{40}
\]

where the size of sets is measured by an application-relevant measure \( \mu \), e.g., fluid volume. Finding pairs of coherent sets can be formulated as a variational problem based on the linear Perron-Frobenius transfer operator. The transfer operator acts on tracer distributions \( \lambda \) by composing them with the flow map \([P_{t_0}, \lambda]|(A) = \lambda(\Phi_{-t}(A, t_0))\). A solution to the search for coherent pairs is conveniently given by the eigenfunctions of the transfer operator. Similar methods were previously used successfully in the reconstruction of invariant measures of dynamical systems in Dellnitz and Junge (1999, 2002),

FIG. 21. Elliptic and forward and backward hyperbolic Lagrangian coherent structures (LCSs) in a time-periodic Bickley jet, detected based on the variational redefinition of LCSs. From Haller and Beron-Vera, 2012.

FIG. 22. A satellite image showing chlorophyll being carried by an Agulhas ring away from the southern tip of Africa (bottom right in the image). Superimposed in black is an elliptic Lagrangian coherent structure corresponding to the ring. From Beron-Vera et al., 2013.

FIG. 23. Coherent set of the Antarctic atmospheric vortex, visualized using the Perron-Frobenius technique. Images within each row are different views of the same set; rows are visualizations for data recorded seven days apart. From Froyland, Santitissadeekorn, and Monahan, 2010.
almost-invariant sets in time-independent dynamical systems in Froyland and Dellnitz (2003) and Froyland and Padberg (2009), and coherent pairs in time-dependent flows in Froyland, Lloyd, and Santitissadeekorn (2010) and Froyland, Santitissadeekorn, and Monahan (2010).

Numerical approaches predominantly rely on Ulam approximations of transfer operators, which discretize the fluid domain into smaller cells. Launching a large number of Lagrangian particles from each set, the transfer operator is approximated by a large, but sparse, Markov chain transition matrix. Discretization cells corresponding to nonzero elements of eigenvectors of the Ulam matrix form sets that approximate coherent pairs, as explained in the review by Froyland and Padberg-Gehle (2013) (see Fig. 23).

In the context of fluid flows, the transfer operator approach was applied to geophysical transport: e.g., atmospheric flows in Santitissadeekorn, Froyland, and Monahan (2010), and oceanic flows in Froyland et al. (2012) and Sebille, England, and Froyland (2012). Case studies comparing Lagrangian coherent structures and coherent sets (Froyland and Padberg, 2009; Tallapragada and Ross, 2013) show that both approaches capture similar features in flows, which is explained by Froyland and Padberg-Gehle (2012). A further connection between coherent sets and braid-theoretic studies of mixing (see Sec. V.D) can be found in Grover et al. (2012). That study points out that while material in coherent sets might not be well mixed with the surrounding fluid, the coherent sets may act as "ghost" rods that stir the fluid around them, promoting efficient mixing in the rest of the fluid.

C. Mesochnotic analysis

Given any observable function, i.e., a field over states \( f(x_0, t_0) \), we can generate its time-averaged, or mesochronic, counterpart \( \tilde{f}(x_0, t_0) \) by evolving a trajectory \( x_t(x_0, t_0) \) from the initial condition \( (x_0, t_0) \) and computing the Lagrangian average of \( f \) along it:

\[
\tilde{f}(x, t_0) = \frac{1}{t} \int_{t_0}^{t_0+t} f(x_t(x, t_0), t) dt. \tag{41}
\]

It is a well-known result in ergodic theory that such functions will be constant over sets invariant under dynamics when \( t \to \infty \), regardless of \( f \) initially chosen. This result was exploited by Mezić (1994) and later by Mezić and Wiggins (1999), Levnajic and Mezić (2010), and Budišić, Mohr, and Mezić (2012) to construct an ergodic partition for closed, autonomous, and periodic dynamical systems (see also Sec. VII.F).

In data-based systems, the averaging period \( t \) is always finite, but dynamically relevant information can still be extracted from \( \tilde{f} \), in particular, if the observable \( f \) is chosen to relate to the velocity vector field \( \mathbf{u} \) of the flow. [These methods are referred to as parcel schemes by Samelson (2013) and as Lagrangian descriptors by Mancho et al. (2013).] The first attempt at doing so in the context of chaotic advection was by Malhotra, Mezić, and Wiggins (1998), who computed so-called patchiness plots by averaging the magnitude of the velocity field \( ||u|| \). The boundaries in patchiness plots were recognized by Poje, Haller, and Mezić (1999) to correspond to certain analogs of invariant manifolds in 2D turbulence (Haller and Poje, 1998). The same quantity was again studied by Madrid and Mancho (2009) and Mancho et al. (2013) and shown to align with ridges of the finite-time Lyapunov exponent field. Instead of the norm of velocity, Haller and Iacono (2003) computed time averages of scalar quantities related to the Jacobian matrix \( \nabla \Phi \) and showed that it is possible to produce a finer characterization of deformation in a flow which, again, agreed with proxies of Lagrangian coherent structures.

Instead of averaging scalars related to the velocity field, Mezić et al. (2010) computed time averages of the velocity vector field itself, producing the so-called mesochronic velocity. These averages \( \tilde{\mathbf{u}} \), computed using Eq. (41), are directly related to the flow map \( \Phi_t \) as the average velocity is just the displacement caused by dynamics, divided by the duration of the motion:

\[
\tilde{\mathbf{u}}(x, t_0) = (\Phi_t(x, t_0) - x)/t. \tag{42}
\]

The mesochronic velocity Jacobian matrix \( \nabla \tilde{\mathbf{u}}_t \) is therefore directly related to the Jacobian matrix of the flow map \( \nabla \Phi_t \). Since the spatial derivative \( \nabla \) and the averaging integral \( t^{-1} \int_{t_0}^{t_0+t} \) do not commute, this analysis is different from the analysis by Haller and Iacono (2003) mentioned earlier. Unlike maximal finite-time Lyapunov exponents \( r^{-1} \log ||\nabla \Phi_t|| \) that measure the magnitude of deformation that the material experiences, the scalar field \( \det \nabla \tilde{\mathbf{u}} \) uncovers the character of deformation for planar incompressible flows, e.g., strain (mesohyperbolicity) or rotation (mesoellipticity). Partitioning the state space based on values of \( \det \nabla \tilde{\mathbf{u}} \) is a direct generalization of the Okubo-Weiss deformation.

FIG. 24. Satellite images of bacterial abundance in the (a), (c) Gulf of Mexico and (b), (d) computed mesohyperbolic regions show that finite-time deformation computed from time-averaged velocity indicates chaotic advection activity. From Valentine et al., 2012.
criterion (Okubo, 1970; Weiss, 1991) from infinitesimal to finite advection times. Note that Greene (1968, 1979) previously used a similar quantity to predict the order of destruction of KAM tori in perturbed Hamiltonian maps, the phenomenon shown to be relevant to mixing in Sec. IV.C. The mesochronic calculations were further generalized to 3D flows by Budišić et al. (2016).

The analysis of the mesochronic Jacobian has been applied to the prediction of the oil-slick transport in the aftermath of the Deepwater Horizon spill. Mezić et al. (2010) showed that regions of hyperbolicity correspond to jets that dispersed the slick, while elliptical zones correspond to eddies in which the slick accumulated. In a follow-up paper, the technique contributed toward resolving the problem of “missing oil,” indicating the correct locations and volumes of the oil post-spill (Valentine et al., 2012) (see Fig. 24).

D. Braids of Lagrangian trajectories

Although flows are most commonly analyzed through their vector fields, in certain situations data available about the flow are too sparse to allow interpolation of the (continuous) vector field. An example is the Argo program7 that measures global oceans using sensors relaying their positions, temperature, and salinity as they are advected by ocean currents. In other cases, recorded data are sparse because detailed data would be too complex to use and act upon. For example, industrial stirring devices often use stirring rods following simple protocols that induce complicated flows surrounding them (Sec. VIII.A). In both of these instances, braid-based approaches are able to estimate properties of the flow from only a few Lagrangian trajectories, whether sensor or stirring-rod paths, significantly reducing the need for detailed measurements.

A seminal paper by Boyland, Aref, and Stremler (2000) [anticipated a decade earlier by MacKay (1990)] takes Lagrangian trajectories of a time-varying planar flow and treats them as entangled strands in the 3D space-time domain, resulting in an approach termed braid dynamics. The tangle of trajectories is represented by a braid, which is a sequence of symbols encoding only relative positions of trajectories and their exchanges. Braids can be interpreted as reduced models of flows, acting on closed material lines tightened around trajectories to form so-called loops, which are deformed by interchanges of trajectories.

Regardless of the number and locations of trajectories forming the braid, topological entropy of the braid is always the lower bound for the topological entropy of the flow, which, in turn, is connected to mixing rates (see Sec. VII). Such lower bounds have been used to optimize stirring protocols used for mixing in viscous fluids (Boyland, Aref, and Stremler, 2000; Thiffeault and Finn, 2006; Finn and Thiffeault, 2011) (see Fig. 25). As an analysis procedure, the entropy of the flow can be bounded by either computing braids of individual trajectories, as in Thiffeault (2005, 2010), or braids of coherent structures that act as “ghost rods,” stirring the surrounding fluid, as in Gouillart, Thiffeault, and Finn (2006), Stremler et al. (2011), Grover et al. (2012), and Tumasz and Thiffeault (2013).

The topological entropy of a braid is the fastest rate at which a loop can be deformed under a braid. The slowest rates of deformation are also significant, as minimally growing loops can enclose trajectories that remain trapped together in a coherent structure (Allshouse and Thiffeault, 2012). Such loops then provide approximations to boundaries of coherent structures, which may be computed from several trajectories alone, instead of from the fully resolved velocity field; see Fig. 26.

The braid approach allows one to process extremely sparse data of planar flows in a computationally efficient manner and

---


---

FIG. 25. A braid constructed from the motion of stirrers in a vat, along with a sample material line deformed by the motion. From Thiffeault and Finn, 2006.

FIG. 26. An approximation to the boundary of a coherent structure in a stirred vat of fluid estimated from only 40 sampled trajectories. Compare Lagrangian coherent structure computation that would require knowledge of the entire vector field. From Allshouse and Thiffeault, 2012.
VI. LAMINAR VERSUS TURBULENT FLOWS

A fundamental dichotomy between laminar and turbulent flows was mentioned in the Introduction and at points throughout this review and exists in the fluid dynamics literature more generally. We believe, however, that a more refined categorization of fluid flows is required; we provide a road map for rethinking this question, based on what has been learned in the last several decades through studies of chaotic advection.

A. Laminar flows

The basic image of laminar flow is that the flow takes place in layers. Thus a good initial definition of laminar flow is that the velocity field has a codimension-1 foliation whose leaves are invariant. A codimension-1 foliation of an n-dimensional manifold is a decomposition into injectively immersed submanifolds called leaves such that in a neighborhood of any point the decomposition is diffeomorphic to the decomposition of \( \mathbb{R}^n \) into the planes \( x_a = \text{const} \), where \( x_a \) is the last coordinate in \( \mathbb{R}^n \). Our discussion will progress from steady flows to time-dependent flows.

All steady mass-preserving flows in 2D are laminar, or almost so. The leaves can be taken to be the level sets of the stream function. Yet this example already shows that the definition needs generalizing to allow finitely many singular leaves of lower dimension. The problem is that although the generic level set is a smooth 1D submanifold, for critical values of the stream function the level set may be a single point or contain saddle points or worse. There are various ways of prescribing which forms of singular leaf are allowed and how the leaves fit together around it. This is especially well worked out in the theory of pseudo-Anosov maps [8] for references aimed at chaotic advection, see Thiffeault and Budić (2014). But there are various not-quite-equivalent formulations, so we refrain from being prescriptive.

The case of 2D steady mass-preserving flows illustrates another feature of laminar flows that may arise. The 2D manifold is not simply connected then for a vector field \( v \) preserving a nondegenerate 2-form \( \omega \) (formulating mass preservation in a way that is more general than volume preservation) there is not necessarily a single-valued stream function, but the flow still preserves a closed 1-form. For example, the vector field \( \dot{x} = 1, \dot{y} = V \) constant on the 2-torus \( \mathbb{R}^2 / \mathbb{Z}^2 \) preserves area \( \omega = dx \wedge dy \) and the 1-form \( \alpha(x) = \omega(v, \xi) = \xi_y - V \xi_x \) for tangent vectors \( \xi = (\xi_x, \xi_y) \), which is \( d\Psi(\xi) \) for \( \Psi = y - Vx \) on \( \mathbb{R}^2 \), but \( \Psi \) is multivalued [9] for the integral submanifolds of the 1-form, which in general are only injectively immersed submanifolds rather than true submanifolds (e.g., \( V \) irrational in the example). Indeed in this example they densely wind around \( \mathbb{T}^2 \). It is the density of invariant leaves that gives a laminar flow the potential for chaotic mixing, in contrast to those with a single-valued stream function. This simple example does not have chaotic mixing but we will show examples in Sec. VI.B that do.

Moving to steady 3D flows, the prime example of Poiseuille flow in a circular pipe is laminar: the obvious foliation by concentric cylinders works provided we allow the singular 1D leaf down the axis. But this example also allows many other foliations with invariant leaves. Choose any foliation of an initial cross section and let its leaves flow with the vector field to give a foliation of the whole domain. So there is a high degree of nonuniqueness in the foliation. If one adds swirl this still works.

Any steady 3D mass-preserving flow with a continuous symmetry is laminar (Mezić and Wiggins, 1994; Haller and Mezić, 1998), because it has a stream function whose level sets make the leaves. Here is an outline proof: Given a 3-form \( \omega \) representing mass density, a velocity field \( v \) preserving \( \omega \) and a vector field \( u \) representing a symmetry (i.e., also preserving mass and commuting with \( v \)) and independent of \( v \) modulo a set of zero volume, let \( \gamma = i_i, i_0, \omega(u) \). It is a 1-form and calculation shows that it is closed (\( d\gamma = 0 \)). Thus the field of planes defined by \( \gamma = 0 \) is integrable to a foliation. Its leaves are invariant because \( i_i, i_0, \omega(u) \) is zero by antisymmetry of \( \omega \). The constructed foliation is also invariant under the symmetry.

What about dynamically defined classes of steady flow, such as Euler flows or Stokes flows? For steady Euler flows the Bernoulli function [Eq. (30)] is conserved so its level sets provide a foliation with invariant leaves, unless the Bernoulli function is constant. In the latter case curl \( v = \kappa e \) for some function \( \kappa e \) of position, and \( \kappa e \) is conserved so its level sets provide a foliation with invariant leaves, unless \( \kappa e \) is constant. The latter case gives the Beltrami flows, for which a complete understanding is still

---

8MATLAB toolbox braiddlab; see Thiffeault and Budić (2014).
9An injectively immersed submanifold of a manifold \( M \) is the image of a one-to-one differentiable map \( f : N \rightarrow M \) from a reference manifold \( N \) into \( M \) for which the derivative \( Df \) has maximal rank everywhere.
11An n-form is an antisymmetric n-linear functional of n tangent vectors at each point. If n is the dimension of the manifold, the n-form is nondegenerate if at every point there is an n-tuple of tangent vectors on which it is nonzero.
12The operator \( i_i \) is the contraction with the vector field \( v \) as the first argument; thus for all tangent vectors \( \xi \), \( i_i(\alpha(\xi)) = \omega(v, \xi) \). An n-form \( \alpha \) is closed if its exterior derivative \( d\alpha \) is zero; equivalently if the integral of \( \alpha \) over the boundary of any (n + 1)-dimensional ball is zero, which is true here because \( v \) preserves \( \omega \).
13A closed 1-form is called exact if it can be written as \( d\Psi \) for some single-valued function \( \Psi \), which is then a genuine stream function.
lacking (Arnol’d and Khesin, 1998, Chap. II.1). Is every 3D Stokes flow laminar in this sense? Following Boyland, Aref, and Stremler (2000), we feel that this is probably true, because they minimize the dissipation rate, which would make them analogous to Thurston’s simplest representatives of isotopy classes of surface homeomorphisms, each of which has an invariant foliation (two in the case of pseudo-Anosov components), but we are not aware of a definitive answer.

Are there 3D steady mass-preserving flows which are not laminar? One might be tempted to say that any chaotic flow would give an example, because most candidate leaves grow exponentially under the flow, but in fact the better forms of chaotic flow possess invariant foliations. Specifically, each uniformly hyperbolic 3D mass-preserving steady vector field has a forward contracting foliation and a backward contracting foliation (with no singularities, even). An example is Arnol’d’s fast dynamo flow [see, e.g., Chap. II.5.E of Arnol’d and Khesin (1998)], the vector field (0,0,1) on the manifold made by identifying horizontally opposite sides of a cube by translation and the top to the bottom by the matrix

\[
A = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}.
\]

The foliation by the surfaces \(gy - x = C\) is constant, where \(g = (1 + \sqrt{5})/2\) is invariant. To interpret this correctly, one has to identify the values \(C + 1, C - \gamma, \) and \(\gamma^2 C\) with \(C\). Thus each leaf is dense in the manifold. The foliation \(y + \gamma x = \text{const}\) is also invariant. The flow is chaotic in many good senses: in particular, displacements in the direction \((\gamma, 1, 0)\) grow exponentially at rate \(2 \log \gamma\). Although it looks artificial, versions of this flow can be made in 3D containers in Euclidean space (MacKay, 2008).

A genuine obstacle to laminar flow is elliptic periodic orbits. If a leaf passes through an elliptic periodic orbit then after one period the leaf comes back rotated by an amount not equal to 0 or \(\pi\), contradicting the condition for a foliation. One can allow a finite number of elliptic periodic orbits by our decision to allow finitely many singularities in the foliation, but it would probably be a mistake to allow infinitely many, or at least to allow singularities to have accumulation points. Thus any mass-preserving steady 3D flow with a set of elliptic periodic orbits with an accumulation point is not laminar. Generically, within the class of smooth-enough mass-preserving steady 3D flows, if there is one elliptic periodic orbit then it is an accumulation point of others. So, except for the highly exceptional integrable cases (i.e., those with a stream function), and the hyperbolic cases, all the rest are very likely not laminar, unless those arising dynamically are a special subset (as we guess may be the case for Stokes flows).

A generalization of the concept of invariant foliation that may be useful is that of measurable partition (Rokhlin, 1960; Mezić, 1994). In this case, the elements of the partition are not necessarily submanifolds and can even be fractal sets. If a flow possesses an invariant, nontrivial\(^{14}\) measurable partition it could be considered laminar. Nontriviality could be too weak though, because a flow on an annulus that mixes the top half of the annulus and the bottom part of the annulus separately would have only two, thick laminae. Recently, the concept of an ergodic quotient was introduced for flows in bounded domains (Levajčić and Mežić, 2010; Budišić and Mežić, 2012). Roughly speaking, the ergodic quotient space is obtained by identifying each subset of the physical space that the bounded fluid flow “samples” well (precisely speaking, time averages of continuous functions are equal to space averages with respect to a measure defined on that subset; e.g., the invariant subset might be a torus, the flow might induce a dense winding on the torus and the measure is the surface area on the torus, and then the ergodic quotient is a single point). The ergodic quotient of a flow on the annulus already described would consist of two points, representing the top and bottom parts of the annulus. However, a shear flow in such an annulus—produced, e.g., by the outside circle of the annulus rotating and the inside also rotating at smaller angular velocity, as in Couette flow—would have an ergodic quotient that is a closed interval and is thus 1D. We could then say that the flow is \(d\) laminar with (possibly nonuniform) thickness of the laminae, provided the fractal dimension of the nontrivial ergodic quotient is \(1 - d\). In the case of \(d = 0\), we retain the label of “laminar” instead of “0 laminar.”

Let us move now to time-dependent flows. For these flows it is essential to consider the vector field as living in a domain of space-time rather than just space. So an \(n\)-D time-dependent flow corresponds to a vector field on an \((n + 1)\)-D manifold, where the \(t\) component is just 1. Our definition proposes that the flow is laminar if there is an \((n + 1)\)-D foliation of this space-time, with invariant leaves. Let us think about this for \(n = 2\). As for Poiseuille flow, any foliation of the initial domain at \(t = 0\) is transported to a foliation of the whole space-time with invariant leaves. But the leaves may become more and more contorted at large times. So we put a restriction on the types of foliation that we allow, namely, the diffeomorphisms that map the local decompositions onto the planes \(x_n = \text{const}\) should be bounded in \(C^2\). In the case of periodic time dependence, we would furthermore want the foliation to have the same period in time. Thus, because of elliptic islands, the beloved example of blinking vortex flow (Aref, 1984) is almost certainly not laminar.

B. Turbulent flows

A good starting point for considering turbulent flow is the viewpoint of Taylor (1954) that a flow is turbulent if the velocity autocorrelation integral in time converges. To define this, a (possibly time-dependent) velocity field \(v\) in a subset of Euclidean space induces a flow map \(\phi_{ts}\) from time \(s\) to time \(t\). Let

\[
C_{ij}(\tau; t) = \int v_i(\phi_{t+s}, x, t + \tau) v_j(x, t) \mu_i(dx)
\]

where \(\mu_i\) represents the mass density at position \(x\) and time \(t\). Say the flow is turbulent if

\[
\int_0^\infty C_{ij}(\tau; t) d\tau < \infty
\]

\(^{14}\)A trivial measurable partition of a set \(A\) consists of the set itself and the empty set.
for each $t$. Probably if it holds for one value of $t$ then it holds for all, though we have not seen this proved.

Note, however, that Taylor’s condition can hold for some steady flows, contradicting the commonly held notion that turbulent flows must be time dependent. For example, modify the vector field in Arnol’d’s fast dynamo flow to $(0, 0, g(x, y, z))$ for some function $g > 0$ such that $g(x, y, 1) = g(A(x, y), 0)$ and the ratio of period to number of revolutions in $z$ is different for two periodic orbits. Then the flow preserves the modified mass $g^{-1}dx \wedge dy \wedge dz$ and by Anosov’s alternative (Anosov, 1967) it is mixing in the ergodic theorist’s sense, i.e., $C(\tau) = \int \psi(x)\mu(x)dx \rightarrow 0$ as $\tau \rightarrow \infty$ for all pairs of functions $\psi \in L^\infty$, $\chi \in L^1$. A stronger nontriviality condition on $g$ probably makes $\int_0^\infty C(\tau)\, d\tau$ converge. Note that this flow preserves the same foliation as for the case $g = 1$, and so is also laminar.

On the other hand, Taylor’s condition never holds for steady flows with no-slip boundaries. The issue is that the volume remaining in a neighborhood of the boundary for time at least $\tau$ is at least $C/\tau$ for some $C$ (depending on the neighborhood), and thus the velocity autocorrelation integral diverges (Jones and Young, 1994). The effect of walls has been addressed further in Sec. III.

Also, there are flows that are neither turbulent nor laminar. Take, for example, a steady 3D flow with infinitely many elliptic periodic orbits. As discussed in the previous section it is not laminar. Yet if the elliptic periodic orbits are surrounded by a set of positive volume of invariant tori, as given generically by KAM theory for smooth- enough mass-preserving vector fields, then the velocity autocorrelation integral diverges.

Note that Taylor’s notion is in essence a Lagrangian notion, since it requires integration of physical quantities along particle paths. It can be easily reformulated in the context of the Koopman operator (Mezić, 2013), requiring that the Koopman operator spectrum of the solution of the dynamical evolution equation (e.g., the Navier-Stokes equation) with appropriate initial and boundary conditions does not have any eigenvalues and associated eigenfunctions, except for trivial (constant) ones (Arnol’d and Avez, 1968). In other words, the flow generated by the solution has the mixing property. An alternative Eulerian view would be to require integrability over time of the Eulerian velocity autocorrelation function that is given by

$$C_{ij}(\tau; t) = \int v_i(x, t + \tau) v_j(x, t)\, dt$$

for every $x$. This, in fact, can be related to properties of another Koopman operator, the one associated with the velocity phase space. In particular, the property would be implied if the related Navier-Stokes attractor has the mixing property. Nevertheless, even this is not enough to capture the full idea of what most mean by “turbulence”: the definition does not say anything about the spatial distribution of eddies and their size structure, an important notion in the turbulent energy transfer context (Frisch, 1996). Specifically, a flow with a very simple spatial dependence—with all the velocity vectors pointing in one direction—but complex temporal dependence could have the mixing property on the strange attractor in Navier-Stokes phase space, and so be turbulent in the sense set out here, but clearly does not have the energy-cascade property of turbulence. Such a flow would also be laminar in the sense of our previous discussion.

C. Synthesis

Is laminar versus turbulent flow really a dichotomy? The answer is clearly no. We have shown that with common definitions there are flows that are both laminar and turbulent and there are flows that are neither. And we have not even addressed the case of flows that one might want to say are laminar in some regions and turbulent in others. The variety of flows that exist is not captured by the current simplistic laminar versus turbulent classification. More work needs to be done and we hope that this discussion might provide some ideas.

VII. THE STRUCTURE OF THE MIXED STATE: QUALITY AND MEASURES OF MIXING

At the heart of all applications of chaotic advection lies the desire to exploit the inherent complexity in deterministic chaotic dynamics for fluid mixing. There are several different mathematical diagnostics for revealing this complexity, for example, recurrence, transitivity, ergodicity, and entropy; which is the most relevant will depend on the particular application. The most natural one of concern for mixing applications is arguably the measure-theoretic notion of (strong) mixing: a measure preserving (invertible) transformation $f: M \rightarrow M$ is strong mixing if for any two measurable sets $A, B \subset M$,

$$\lim_{n \rightarrow \infty} \mu(A \cap f^n(B)) = \mu(A)\mu(B)/\mu(M).$$

We understand the meaning of Eq. (43) as follows: in a container $M$ of fluid, if $B$, a region originally occupied by a tracer, is mixed to become $S'B$ [in continuous time, or equivalently $f^n(B)$ in discrete time], as depicted in Fig. 27, then for any arbitrary part $A$ of the container, the amount of tracer in $A$ after mixing is $\mu(A \cap f^n(B))/\mu(A)$ and in the limit $n \rightarrow \infty$ this should be $\mu(B)/\mu(M)$. For incompressible fluids the measure $\mu$ can be assumed to be Lebesgue, that is, area or volume and we may normalize the volume of the domain to $\mu(M) = 1$. This definition effectively says that any two sets

15Taylor’s definition of turbulence is not immediately applicable here because Arnol’d’s flow does not live on a subset of Euclidean space, so the meaning of multiplying components of a vector at different points is unclear. However, one can probably turn it into an example in a bounded container with stress-free boundaries for which the velocity autocorrelation function does make sense and is integrable, so this example shows that turbulence might not require time dependence.

16Typically in ergodic theory the short form mixing is used as a synonym for strong mixing, but here we use the full term to avoid confusion both with the everyday sense of the term mixing and with the fluid dynamical definition we put forward in Sec. I.
by In turn weak mixing implies ergodicity, which can be defined
\[ \mu \xi_j \text{ converges to } 0 \text{ in the sense of Cesàro summation, and ergodic} \]

quantity \( \mu \) becomes asymptotically independent of each other for a strong mixing system, as illustrated in Fig. 27. Strong mixing implies the weaker property of weak mixing,

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap f^n(B)) - \mu(A)\mu(B) = 0. \quad (44) \]

In turn weak mixing implies ergodicity, which can be defined by

\[ \lim_{n \to \infty} \frac{1}{n} \sum_{k=0}^{n-1} \mu(A \cap f^n(B)) = \mu(A)\mu(B). \quad (45) \]

Note that for each of these properties we consider the quantity \( \mu(A \cap f^n(B)) \) in relation to \( \mu(A)\mu(B) \). Setting \( \mu(A \cap f^n(B)) - \mu(A)\mu(B) = \xi_{f,n} \), we see that \( f \) is strong mixing if \( \xi_{f,n} \) converges to 0 as \( n \to \infty \), weak mixing if \( \xi_{f,n} \) converges to 0 in the sense of Cesàro summation, and ergodic if \( \xi_{f,n} \) converges to 0 in the Cesàro sense. The hierarchy is strict: strong mixing implies weak mixing, which itself implies ergodicity. The converse is not true; an irrational rotation is an example of an ergodic system that is not weak mixing, while an interval exchange transformation can be weak mixing but not strong mixing.

Mathematical models that underpin the basics of fluid mixing devices can be proven to exhibit this behavior. For example, the cat map \( \mathbf{A} \)—an integral part of the Arnol’d fast dynamo model discussed in Sec. VI—a paradigmatic model of repeated shearing in orthogonal directions, can be shown to be strong mixing and to have positive topological entropy (equal to the logarithm of its largest eigenvalue). Another less simple yet analyzable example is the map of Cerbelli and Giona (2005), which is strong mixing yet shows a multifractal structure [some aspects of this case are further treated by MacKay (2006)]. In practice, however, it may be implausible to expect the strong-mixing property Eq. (43), and impossible to verify, since to do so requires computing over all possible regions \( A \) and \( B \). These issues, coupled with the ubiquity of mixing problems in applications, have led to reformulations of Eq. (43). For example, in Sec. VII.D we discuss weak convergence, and a method to establish this condition, while in Sec. VII.E we describe the decay of correlations of observable functions, which is used in ergodic theory to estimate rates of mixing. These two notions are closely related, but we will describe each in terminology and notation most closely associated with their own bodies of literature.

First we observe that taking a physical measurement such as concentration involves a choice of scale. In Secs. VII.A, VII.B, and VII.C we take a new look at classical ideas of measuring the quality of a mixture, in which the choice of scale provides details of interest about a mixing process. Section VII.D contrasts this with the recent concept of multiscale measures for mixing.

The quality or “goodness” of mixing may be quantified by using statistical quantities like the coarse-grained density and the entropy. These quantities may be employed for this purpose if at any moment in time the distribution of the (marked) fluid to be mixed in the ambient fluid domain is exactly known. We restrict ourselves here to 2D incompressible flow, focusing on the problem of a blob of colored fluid introduced into a fluid in which some specified flow is present. We use the concept of the coarse-grained density of the distribution—introduced by Gibbs (1902)—as a basic measure of the three criteria of the mixed state: the averaged square density (Welander, 1955), the entropy (Gibbs, 1902), and the intensity of segregation (Danckwerts, 1952). By using these criteria we can estimate the time necessary for the mixed state to become uniform within some specified range, for a given volume element size (the “grain”). Note that the three criteria are not independent and that they are statistical measures of the first order. For a more complete description of a mixture, we also use the scale of segregation (Danckwerts, 1952), which is a statistical measure of the second order and represents an average of the size of the clumps of the mixed component. The first-order statistics provide information related to the coarse-grained density distribution calculated in each grain at a time. And the second-order statistics gives information about the correlations of this density in two different grains at the same moment of time.

A. Definitions of mixing quality

We first illustrate the concept of coarse-grained density as a quantification of mixing in an example problem of 2D stirring in a square cavity. The cavity is covered by a grid of \( N_\delta \) square cells of side \( \delta \) and area \( S_\delta = \delta^2 \) and initially contains a square black blob in its center. The area of the cavity is \( S = N_b S_b \), while the blob’s area is \( S_b \), which is conserved under area-preserving stirring. We denote the area of the blob inside cell \( n \) by \( S_b^{(n)} \), and the proportion of black \( D_n = S_b^{(n)}/S_b \), which may be considered a probability density. Calculating the average of \( D_n \) by summing over cells we obtain

\[ \langle D \rangle = \frac{1}{N_b \sum_{n=1}^{N_b} D_n} = \frac{1}{N_b \sum_{n=1}^{N_b} S_b^{(n)}} = \frac{S_b}{S}, \]

where the ratio of the area is \( S_b \) of the colored matter and the total area is \( S \) of the cavity. This value does not change in the course of the stirring and is the mean or uniform density \( \langle D \rangle \).
of the colored blob in the cavity. (The angle brackets here and later denote an average over the cavity.) However, using the square density defined by $D_n^2 = \left( \frac{S_{(n)}^b}{S_b} \right)^2$ and averaging over the area of the cavity, keeping the cell area constant as before, we get the following inequality:

$$\langle D^2 \rangle = \frac{1}{N_{\delta}^2} \sum_{n=1}^{N_{\delta}} D_n^2 = \frac{1}{S} \sum_{n=1}^{N_{\delta}} D_n S_{(n)}^b \leq \frac{S_b}{S}$$  \hspace{1cm} \text{(46)}$$

because $D_n \leq 1$.

Figure 28 shows a square cavity with a central black blob for two different grid sizes. Calculations given in detail by Krasnopolskaya and Meleshko (2009) show how the averages $\langle D \rangle$ and $\langle D^2 \rangle$ depend on the cell size. In Figs. 28(a) and 28(b) grid cells are either completely filled or completely empty, giving $D_n = 1$ or 0 for each $n$. Thus we have $\langle D \rangle = \langle D^2 \rangle = 1/4$. Figures 28(c) and 28(d) show the blob after stirring has taken place. Now in Fig. 28(c) $D_n$ can take values 0, 1, 1/2, or 3/4, and we have $\langle D^2 \rangle = 35/256 < 1/4$, while in Fig. 28(d) all cells are either completely filled or empty, so that $\langle D \rangle = \langle D^2 \rangle = 1/4$. Similarly in Figs. 28(e) and 28(f), after further stirring, we have $\langle D^2 \rangle = 3/40 < \langle D \rangle = 1/4$ when $\delta = 4$, and $\langle D \rangle = \langle D^2 \rangle = 1/4$ when $\delta = 1$. Thus for $\delta = 4$, $\langle D^2 \rangle$ decreases with the thinning of the black striations, while if $\delta = 1$, all cells are filled or empty, and $\langle D^2 \rangle$ does not change.

In terms of statistical mechanics, $D_n$ is the coarse-grained density, which is different from the fine-grained density $f_d$ of the infinitesimal, superdifferential elements $dS_i$, which are always assumed to be small compared to the width of the area of the colored matter. Moreover, $dS_i$ is always so small that either it is located inside the colored matter and $f_d$ or it is outside and $f_d = 0$ (Gibbs, 1902).

It was shown by Gibbs, for the special case of mixing two fluids approaching the perfectly mixed state, that the final state of mixing is characterized by a minimum statistical square density, i.e.,

$$\langle (D_n - \langle D \rangle)^2 \rangle \equiv \langle D^2 \rangle - \langle D \rangle^2 \rightarrow 0. \hspace{1cm} \text{(47)}$$

Thus, going to a uniform mixture in time, the mean square density $\langle D^2 \rangle$ will approach its minimum $\langle D \rangle^2$. The rate of decrease of these values is not only time dependent but also depends on the cell sizes.

It is also possible to use the analogy of entropy, i.e., $-D_n \log D_n$ instead of $D_n^2$ as a statistical measure. If each cell is empty or completely filled, the entropy $-D_n \log D_n$ equals zero. The entropy measure changes only in those boxes where $0 < D_n < 1$. Moreover, for $0 < D_n < 1$, $-\log D_n$ is always positive, so the more boxes (partially) covered by colored material the larger is $-\sum_{n=1}^{N_{\delta}} D_n \log D_n$. As a result, for a good mixing process, the entropy of the mixture $e = -\langle D \log D \rangle$ will grow in time to its maximum. The entropy measure is not independent of the square density measure; both of them have first-order statistics (one element of area at a time).

Danckwerts (1952) defined two properties that are useful in evaluating the quality of mixing with diffusion and chemical reactions: the scale of segregation $L_C$ (the measure of the second-order statistics) and the intensity of segregation $I_C$ (the measure of the first-order statistics). The scale of segregation is a measure of the size of clumps in a mixture, while the intensity of segregation refers to the variance in composition. For the intensity of segregation Danckwerts introduced

$$I_C = \frac{1}{(C - \langle C \rangle)^2} \int (C - \langle C \rangle) dS = \frac{\langle (C - \langle C \rangle)^2 \rangle}{\langle C \rangle (1 - \langle C \rangle)}.$$  \hspace{1cm} \text{(48)}$$

where $C$ is the local concentration, which is in Gibbs’s and Welander’s definitions equal to the fine-grained density $f_d$. It is easy to see that for the fine-grained density $\langle (f_d - \langle f_d \rangle)^2 \rangle = \langle f_d \rangle - \langle f_d \rangle^2$, and in that case $I_C$ always equals to 1. Consequently, for mixing without diffusion and chemical reactions the intensity of segregation $I_C$ is not decreasing, but equals a constant, initial value.
Therefore, one may suggest a modification of the intensity of segregation by making Gibb’s mean square density Eq. (47) dimensionless by dividing by \((D)(1 - (D))\),

\[
I = \frac{\langle (D_n - (D))^2 \rangle}{(D)(1 - (D))}. \tag{49}
\]

For good mixing \(\langle (D_n - (D))^2 \rangle\) tends to zero, which means that \(I\) also tends to zero. [This definition of \(I\) is different from a similar measure \(I_0\) proposed by Ottino (1989), who defined \(I_0\) as the square root of the mean square density divided by \((D)^2\), so \(I_0^2 = \langle (D_n - (D))^2 \rangle/(D)^2\).] For the calculation of \(I\) as proposed in Eq. (49) and what is in fact the coarse-grained modification of \(I_C\), it is necessary to adopt an additional assumption when considering mixing by a set of \(N\) points: it is assumed that each of the \(N\) points (which together represent the colored blob) carries a small undeformed area equal to \(S_0/N\), which cannot be a correct approximation for continuous mixing flow with large stretching and folding.

The scale of segregation \(L_C\) was defined by Danckwerts (1952) by means of the correlation function

\[
K_C(\eta) = \langle (C - \langle C \rangle)(C - \langle C \rangle) \rangle, \tag{50}
\]

which shows how the concentration fluctuations \(C - \langle C \rangle\) at points 1 and 2, separated by the vector \(\eta\), differ from each other. The normalized correlation function is called the correlation coefficient

\[
\rho_C(\eta) = \frac{\langle (C - \langle C \rangle)(C - \langle C \rangle) \rangle}{\langle (C - \langle C \rangle)^2 \rangle}. \tag{51}
\]

It is obvious that \(\rho_C(0) = 1\). When \(|\eta|\) exceeds a certain value, the relationship between the concentrations at points 1 and 2 may become random when \(K_C(\eta)\) is equal to zero. If a mixture consists of clumps \(|\eta|\) at which \(K_C(\eta)\) is equal to zero (say, \(|\eta| = \xi\)) is approximately the average clump size in the direction \(\eta\). More precisely, the average clump radius in the direction of \(\eta\) is \(L_C(\eta) = \int_0^\xi \rho_C(\eta) d|\eta|\).

The mixing patterns which we are discussing do not consist of a random distribution of clumps, but of layered structures. However, the coarse-grained representations of these patterns may look like a collection of clumps. If we indicate cells for which the density \(D_n\) is larger than \(\langle D \rangle\) in black, the cells for which \(D_n\) equals \(\langle D \rangle\) with gray, and cells where \(0 \leq D_n < \langle D \rangle\) with white, then such a representation has the appearance of a collection of white and black clumps with gray clumps in between them as a transition. Moreover, with the coarse-grained correlation function defined as \(K(\eta) = \langle (D_1 - \langle D \rangle)(D_2 - \langle D \rangle) \rangle\) (where \(D_1\) and \(D_2\) correspond to coarse-grained density in boxes 1 and 2 separated by vector \(\eta\)), the short-term regularity (when \(K > 0\)) in the interval \((0, \xi)\) gives important information about the mixture pattern and can be examined. Short-term regularity means that on average in two boxes at any distance \(|\eta| < \xi\) the fluctuations \(D_n - (D)\) have the same sign (i.e., the same color) and thus \(K > 0\). For \(|\eta| = \xi\) the fluctuations become uncorrelated and therefore \(K = 0\). Thus, the distance \(|\eta| = \xi\) in the direction \(\eta\) is related to the average clump size in this direction, and the value of the scale of segregation

\[
L(\eta) = \int_0^\xi \frac{\langle (D_1 - \langle D \rangle)(D_2 - \langle D \rangle) \rangle}{\langle (D - \langle D \rangle)^2 \rangle} d|\eta| \tag{52}
\]
gives the average radius of the clump. Complementary to the intensity of segregation \(I\), the scale of segregation \(L\) can be used as a measure of clump sizes of the coarse-grained description of mixing patterns. The dynamics of such scales should reflect the changes of sizes of unmixed regions, where \(D_n\) is always larger than \(\langle D \rangle\).

**B. Gibbs’s example of mixing**

In his example of fluid stirring, Gibbs (1902) described the case of the advection of colored fluid in the annular domain between two infinite coaxial cylinders rotating at different speeds. The cross-sectional area between the cylinders is shown in Fig. 29, with the dyed (black) fluid occupying a 90° sector between the cylinders. The flow is driven by the rotation of the outer cylinder: the tangential velocity is \(V_2 \neq 0\) at \(r = 2\), while the inner cylinder is kept fixed \((V_1 = 0 \text{ at } r = 1)\). This forcing drives a purely azimuthal flow, with the azimuthal velocity component given by

\[
u_\theta = Ar + \frac{B}{r^2}, \quad A = \frac{V_2 b}{b^2 - a^2}, \quad B = \frac{V_2 a^2 b}{b^2 - a^2}. \tag{53}
\]

For the case under consideration \(a = 1\) and \(b = 2\), therefore, \(A = 2/3V_2\) and \(B = -2/3V_2\). The advection equations are \(dr/dt = u_r\), \(d\theta/dt = u_\theta\), and with initial conditions \(r = r_0, \theta = \theta_0\) at \(t = 0\), these equations describe the motion of a passive Lagrangian particle in a position \((r, \theta)\) at time \(t\) in the known Eulerian velocity field \(v = (u_r, u_\theta)\).

The advection of the colored fluid may be studied by using the contour tracking method (Krasnopolskaya et al., 1996), in which the interface between colored and uncolored fluid is covered by a number of passive tracer particles. After 12 complete revolutions of the outer cylinder the colored fluid...
occupies a spiral-shaped region as shown in Fig. 30(a). For continued rotation of the outer cylinder, the number of windings of the spiral pattern increases steadily: for an infinite number of revolutions the number of windings would be infinite and the mixture uniform.

We now apply the square density and intensity of segregation measures to quantify the mixing quality. Figure 30(b) replicates Fig. 30(a), but now in the form of the square density ($D^n$) distribution for a cell size $\delta = 0.2$. A similar pattern is shown in Fig. 30(c), but for $\delta = 0.1$. Comparison of the latter with Fig. 30(a) reveals the effect of cell size on the mixing quality.

FIG. 30. Gibbs's example: The evolution of the colored fluid (a) after 12 revolutions of the outer cylinder and the corresponding square density distribution ($D^n$) for (b) $\delta = 0.2$ and (c) $\delta = 0.1$. From Krasnopolskaya and Meleshko, 2009.

FIG. 31. Gibbs's example: Calculated evolution of the intensity of segregation $I$ and the square density $\langle D^n \rangle / \langle D \rangle^2 - 1$ as a function of time $t/T$, with $T$ the time of a half revolution of the outer cylinder. From Krasnopolskaya and Meleshko, 2009.

FIG. 32. Wedge-cavity flow: geometry.
two graphics shows that for a coarser grid (larger cells) the mixture looks uniform, while for a finer grid (smaller cells) the mixing is nonuniform, with empty (white) cells occurring in between the black cells.

The evolution of the intensity of segregation $I$ as a function of $t/T$ (with $T$ the time of a half revolution of the outer cylinder) as calculated for $\delta = 0.1$ and the square density $\langle D' \rangle^2 / \langle D \rangle^2 - 1$ is shown in Fig. 31. It is observed that the mixture becomes more uniform with increasing number of revolutions. In this case the intensity of segregation $I$ is more convenient as a criterion because it changes between two fixed values: it decreases from 1 (initial state) to 0 (complete mixing).

**C. A second example: Mixing in a wedge cavity**

**1. Definition of the problem**

As a second example, we consider a 2D creeping flow of an incompressible viscous fluid in an annular wedge cavity $a \leq r \leq b$, $|\theta| \leq \theta_0$ (Fig. 32), driven by periodically time-dependent tangential velocities $V_{\text{bot}}(t)$ and $V_{\text{top}}(t)$ at the curved bottom and top boundaries $r = a$ and $r = b$, respectively. The sidewalls $a \leq r \leq b$ and $|\theta| \leq \theta_0$ are fixed. We consider a discontinuous mixing protocol with the bottom and top walls alternatingly rotating over an angle $\Theta$ in clockwise and counterclockwise directions, respectively. More specifically, we consider the case $V_{\text{bot}}(t) = 2a\Theta/T$, $V_{\text{top}}(t) = 0$, for $kT < t \leq (k + 1/2)T$ and $V_{\text{bot}}(t) = 0$, $V_{\text{top}}(t) = -2b\Theta/T$, for $(k + 1/2)T < t \leq (k + 1)T$, where $k = 0, 1, 2, \ldots$. $\Theta$ is the angle of wall rotation and $T$ is the period of the wall motion.

The radial and azimuthal velocity components $u_r$ and $u_\theta$ can be expressed by means of the stream function $\Psi(r, \theta, t)$ as $u_r = (1/r) \partial \Psi / \partial \theta$, $u_\theta = -\partial \Psi / \partial r$. For a quasi-stationary creeping flow in the Stokes regime the stream function $\Psi$ satisfies the biharmonic equation $\nabla^2 \nabla^2 \Psi = 0$, with $\nabla^2$ the Laplace operator and the boundary conditions $\Psi = 0$, $\partial \Psi / \partial r = -V_{\text{bot}}$ at $r = a$, $|\theta| \leq \theta_0$; $\Psi = 0$, $\partial \Psi / \partial r = -V_{\text{top}}$ at $r = b$, $|\theta| \leq \theta_0$; $\Psi = 0$, $\partial \Psi / \partial \theta = 0$ at $a \leq r \leq b$, $|\theta| = \theta_0$. Therefore, we have the classical biharmonic problem for the stream function $\Psi$ with prescribed values of this function and its outward normal derivative at the boundary. This wedge-cavity flow problem was solved analytically by Krasnopolskaya et al. (1996). Their analytical solution was used for the numerical evolution of the interface line between the marker fluid and the ambient fluid, which was carried out by the “dynamical” contour tracking algorithm.

**2. Wedge-cavity flow**

The results discussed here correspond to a typical wedge cavity with sector angle $\theta_0 = \pi/4$ and radius ratio $b/a = 2$ (see Fig. 32). Using the dimensionless parameter $H = \Theta/\theta_0$ and a fixed value for the period $T_p$, the discontinuous mixing protocol is completely defined. The value of $H$ is proportional to the displacement of the top and bottom walls during one period. Multiplying by the number of periods $M$, the value $HMT_p$ serves as a measure of the energy supplied during the mixing process. In what follows we restrict to the cases $H = 2$ and 4.

![FIG. 33. Wedge-cavity flow: Mixing patterns obtained in the wedge-cavity flow for different initial positions of the tracer blobs [shaded (colored) circles] and different wall displacements $H$. Upper, lower, and right cavities show the results for the same protocol $H = 4$ and $M = 12$ for different initial positions of the blob. Upper: blob initially placed around a single hyperbolic point; right: for a slightly different initial position; lower: blob split up into four smaller blobs, located in different positions. Left cavity: initial blow positioned around an elliptic point of period-1, for $H = 2$. In all cases the energy input was $3HM$. Adapted from Krasnopolskaya et al., 1999.](Image 322x514 to 553x745)
spreading of the blob over the cavity is less efficient. The left cavity shows the worst mixing result: the initial position of the spot was chosen around elliptic points of period 1 (for a single protocol $H = 2$). Under this protocol there is one elliptic period-1 point. An initial blob placed around an elliptical point results in very poor mixing. The areas around the elliptic points are only slightly deformed during the mixing protocol, returning all material in the area around the elliptic point (Meleshko et al., 1996).

3. First-order measures of mixing

For the case shown in the upper cavity of Fig. 33, i.e., with the blob initially located around a hyperbolic point, we consider the statistical measures of the quality of the mixed state. For this purpose the area of the cavity is covered with cells of uniform size. The evolution of the three criteria based on the coarse-grained density $D$ is shown in Fig. 34 for different cell sizes. The mixing process is characterized by a decrease of square density $\langle D^2 \rangle / \langle D \rangle^2$ [Fig. 34(a)], by a decrease of the intensity of segregation $I$ [Fig. 34(b)], and by an increase of the entropy $\delta^2 / \delta a^2$ [Fig. 34(c)]. The three curves in each figure correspond to three different cell sizes, the labels 1, 2, and 3 corresponding to $\delta = 0.1a$, $0.05a$, and $0.025a$, respectively. From Figs. 34(a) and 34(b) it follows that the intensity and the square density show the same evolution but over different scales: the square density has values approximately in the range between $20$ and $2$, and the intensity is in the range between $1$ and $0$. As the intensity always lies in the same range $(0,1)$, using this criterion it is possible to compare mixing processes for different values of the ratio $S_b/S$ and to compare different mixtures with the same ratio $S_b/S$. For instance, it is easy to answer the question after how much time of mixing the intensity of segregation will have some given value for different box sizes (which basically represents the problem of scaling in mixing processes), by drawing a horizontal line: $I = \text{const}$. By stating that a mixture is uniform enough, for a given box size, when $I$ is less than some minimum value $I_{\text{min}}$, we know how long we have to proceed with the mixing for other box sizes, i.e., for differently sized mixers. For example, if $I_{\text{min}} = 0.05$, then, for the box size $\delta = 0.1a$ [curve 1, Fig. 34(b)] the mixture is uniform for $t/T > 25$ according to this requirement.

Alternatively, we can use the entropy [Fig. 34(c)] or the quality $Q$, defined as the reciprocal of the intensity $I$ [see Fig. 34(d)], which both increase during the mixing process. For the same cell size $\delta$, the same value of the mixing quality (or intensity) can be repeatedly reached. The quality of a mixture can decrease for some time during the mixing process, after which it increases again.

Figure 35 shows the results of mixing for the same protocol with $H = 4$ in four equal wedge cavities for different initial locations of the blobs (green circular patches). The left and right cavities have the initial blob centered around a hyperbolic point of period 2 and covering three hyperbolic points of period 6 and one of period 4. For this case the stretching of the contour (length of interface) is the largest, while the distribution of the deformed blob (indicated by the red color) over the cavity is poorest. The graphs show the distribution of the blob material after $18.5$ (right cavity) and $19$ half periods (left). The length of the contour of the initial green circle is extended $1760$ (right) and $2010$ (left) times, respectively. Visual inspection shows, however, that large parts of the cavity, in particular, in the central area, are not covered by the red blob material. The distribution of the colored fluid over

![FIG. 34. Wedge-cavity flow: Evolution of (a) the square density, (b) intensity of segregation, (c) entropy, and (d) quality of mixing. The initial blob is that shown as a green circle in the upper cavity of Fig. 33. From Krasnopolskaya et al., 1999.](image)

![FIG. 35. Wedge-cavity flow: Mixing patterns generated by the same protocol for different initial locations of the circular blobs. In the left and right cavities the stretching is large, the upper cavity shows large stretching and relatively good mixing, and the bottom cavity shows poor stretching combined with best mixing. Adapted from Krasnopolskaya et al., 1999.](image)
the cavity domain can be quantified, for example, by the intensity of segregation $I$; for box size $\delta = 0.025a$ we find $I = 0.21$ (right) and 0.20 (left), respectively. For the calculation shown in the upper cavity, the initial position of the blob was chosen around the other hyperbolic points of period 2. Stretching of the contour line in this case was almost as large as in previous cases: the line stretched with a factor of 1850. These functions are drawn in Fig. 36(b) as dashed lines 4, 5, and 6, respectively. Using these approximations we can roughly estimate after how many periods the averaged scale of segregation $L$ will be smaller than some given value. For example, for box size $\delta = 0.05a$, $L$ becomes less than $\delta/2$ (the unmixed rubbery domain is smaller than the area of the box) after approximately 35 periods. To find the value of periods $(t/T)$ we use the expression for curve 5 (for the chosen $\delta = 0.05a$), which approximates the value of $L$, and subsequently solve the equation $\exp[-c_1(2t/T + c_2)] < \delta/2$. This shows that $2t/T$ should be larger than 70.

Thus, the existence and the evolution of the unmixed rubbery domains in the mixture pattern can be determined from the behavior of the scales of segregation. For one unmixed zone, as in the initial situation with one circular blob, the scale $L$, as defined in Eq. (52), directly gives the size of the blob. For well-mixed patterns of high quality, the nonzero $L$ values indicate the existence of unmixed domains that very slowly decline (or even stabilize when the zone does not diminish in size at all) despite the continuing mixing.

The evolution of the first-order statistics measures (square density $\langle D^2 \rangle$, intensity of segregation $I$, entropy $e$, and quality $Q$) are different from the dynamics of the scale of segregation. The first ones reflect the distribution of filaments over the mixing domain (for the uniform mixture the coarse-grained density should be less than 1 and larger than 0 all over domain, so filaments should be everywhere), while the latter shows the behavior of the unmixed domains of the coherent structure (Danckwerts, 1952). They also change in opposite ways when changing the box size. The smaller the box size $\delta$, the worse is the mixing according to the first measures and the better according to the scale measures, which decrease with decreasing cell size $\delta$. Therefore, it is necessary to use both measures to judge how well or how badly materials are mixed.

### D. Mix norms

The issue with applying these mixing measures has always been the arbitrariness of the choice of scale. But recent proposals for multiscale mixing measures get around this issue, and in this section we discuss a multiscale mixing measure originally introduced by Mathew, Mezić, and Petzold (2005). While the multiscale measure does not require diffusion to represent the amount of homogenization, it is convenient to begin the discussion by recalling the advection-diffusion equation $\partial C/\partial t + u \cdot \nabla C = \kappa \nabla^2 C$, where $C$ is a concentration field in a finite domain $\Omega$, with no-net-flux boundary conditions. We assume without loss of generality that

\[
\int_{\Omega} C d\Omega = 0,
\]

and define the $L^2$ norm, or variance, as

\[
\|C\|_2^2 = \int_{\Omega} C^2 d\Omega.
\]
The variance evolves according to
\[ \frac{d}{dt} \|C\|_2^2 = -2u \| \nabla C \|_2^2, \] (56)
and so decays in time as the system mixes. The variance indicates the extent to which the concentration has homogenized and is thus a good measure of the amount of mixing that has occurred. However, a full computation of variance requires knowledge of small scales in \( C \), which we are not necessarily interested in. It would be better to use a measure that down plays the small scales. This is more in keeping with the definition of mixing in the sense of ergodic theory (Lasota and Mackey, 1994). We thus proceed to consider the pure advection equation
\[ \frac{\partial C}{\partial t} + u \cdot \nabla C = 0. \] (57)
Note that in this case Eq. (56) predicts that the variance satisfies
\[ \frac{d}{dt} \|C\|_2^2 = 0, \] (58)
and cannot therefore be used as a measure of mixing.

The advection equation (57) returns us to the ergodic sense of mixing of Eq. (43). Consider the advection due to the velocity field to be a time-dependent operator \( S^t \cdot \Omega \rightarrow \Omega \) that moves an initial patch of dye according to
\[ C_0(x) \mapsto C(x, t) = S^t C_0(x). \] (59)
If we consider a region \( A \) of uniform concentration defined by
\[ C_0(x) = \begin{cases} 1 & \text{if } x \in A, \\ 0 & \text{otherwise}, \end{cases} \] (60)
then the volume of the region \( A \) remains constant in time by incompressibility and can be associated with the Lebesgue measure \( \mu(A) \). For a flow enjoying the property of strong mixing Eq. (43) is satisfied, with \( f \) representing a stroboscopic map of the operator \( S^t \).

The intersection of the advected patch \( B \) with the reference patch \( A \) as in Fig. 27 is analogous to projection onto \( L^2 \) functions. This motivates the following weak convergence condition:
\[ \lim_{t \to \infty} \langle C(x, t), g \rangle = \bar{g}, \] (61)
for all functions \( g \in L^2(\Omega) \) with spatial mean \( \bar{g} \), where the inner product is defined by
\[ \langle h, g \rangle = \int_{\Omega} h(x)g(x)d\Omega, \] (62)
and \( h \in L^2(\Omega) \) if \( \int_{\Omega} |h|^2d\Omega < \infty \). Weak convergence is equivalent to mixing as a consequence of the Riemann-Lebesgue lemma. The equivalent conditions Eqs. (43) and (61) require computing over all patches \( A \) or functions \( g \), respectively. Thus, neither of these conditions, taken in isolation, is very useful in practice. However, we now describe a theorem that shows that there is a simpler way to determine whether or not weak convergence is satisfied.

Mathew, Mezić, and Petzold (2005) introduced the mix norm, which for mean-zero functions is equivalent to
\[ \|C\|_{H^{-1/2}} = \|\nabla^{-1/2} C\|_2. \] (63)
Doering and Thiffeault (2006), Lin, Thiffeault, and Doering (2011), and Thiffeault (2012) generalized the mix norm to
\[ \|C\|_{H^q} := \|\nabla^q C\|_2, \quad q < 0, \] (64)
which formally is a negative homogeneous Sobolev pseudo-norm. This norm can be interpreted for negative \( q \) via eigenfunctions of the Laplacian operator. For example, in a periodic domain, we have
\[ \|C\|_{H^q}^2 = \sum_k |k|^{2q} |\hat{C}_k|^2, \] (65)
from which we see that, for \( q < 0 \), \( \|C\|_{H^q}^2 \) smooths \( C \) before taking the \( L^2 \) norm. The theorem
\[ \lim_{t \to \infty} \|C\|_{H^q} = 0, \quad q < 0 \Leftrightarrow C \text{ converges weakly to } 0, \] (66)
due to Mathew, Mezić, and Petzold (2005), Lin, Thiffeault, and Doering (2011), and Thiffeault (2012) shows that we can track any mix norm to determine whether a system is mixing. The existence of this quadratic norm facilitates optimization of the velocity field to achieve good mixing. Mathew et al. (2007) used optimal control to optimize the decay of the \( q = -1/2 \) mix norm. Lin, Thiffeault, and Doering (2011) optimized the instantaneous decay rate of the \( q = -1 \) norm using the method of steepest descent, which is easier to compute numerically but yields suboptimal, but nonetheless very effective, stirring velocity fields. A comparison of the methods for optimized mixing is shown in Fig. 37. The solid line decays faster, but this is merely because \( H^{-1/2} \) cannot be compared directly with \( H^{-1/2} \). The corresponding evolution of

![Fig. 37. Mix norms: Comparison for a flow optimized using the separate methods of optimal control and optimal instantaneous decay. From Lin, Thiffeault, and Doering, 2011.](image)
The decay of this correlation function then gives a measure of the rate of mixing of $f$.

In principle, for a mixing chaotic system, the exponential separation of nearby initial conditions should cause $C_n$ to decay to zero at an exponential rate, that is,

$$C_n = O(e^{-\alpha n})$$  \hspace{1cm} (68)

for some constant $\alpha > 0$. One can establish an exponential decay rate in Eq. (67) for the cat map $A$ (recall again the Arnol’d fast dynamo of Sec. VI) and analytic observables relatively easily. Expanding $\varphi$ and $\psi$ as a Fourier series (here also assuming without loss of generality that $\psi$ has zero mean),

$$\varphi(x) = \sum_{k \in \mathbb{Z}^2} a_k e^{ik \cdot x}, \quad \psi(x) = \sum_{j \in \mathbb{Z}^2} b_j e^{ij \cdot x},$$

the analyticity assumption guarantees that the coefficients $a_k$ and $b_j$ decay exponentially quickly. Linearity and orthogonality mean that Eq. (67) can be written as

$$C_n(\varphi, \psi) = \int \sum_{k \in \mathbb{Z}^2} a_k e^{ik \cdot A \cdot x} \sum_{j \in \mathbb{Z}^2} b_j e^{ij \cdot x} dx = \sum_{k \in \mathbb{Z}^2} a_k b_{-k A}.$$

Since $A$ is a hyperbolic matrix, the exponential growth of $|k A^n|$ together with the exponential decay of Fourier coefficients yields superexponential decay of $C_n$ (Baladi, 2000). Note that this calculation is possible only because the map is linear, with constant Jacobian. The same properties also reveal why the cat map is a poor model for real mixing devices. It cannot represent any nonlinear or nonuniform behavior.

In practice it has been observed that the presence of boundaries can slow this rate. This phenomenon can also be demonstrated in the measure-theoretic viewpoint by computation of the decay of correlations for linked twist maps (Sturman, Ottino, and Wiggins, 2006). These form a class of maps that can be thought of as nonuniformly hyperbolic generalizations of the cat map and that include boundary regions at which particular hydrodynamical boundary behavior can be modeled. Using a Young tower (Young, 1998) approach, Springham and Sturman (2013) demonstrated that in linked twist maps $C_n$ decays at a polynomial rate for all choices of scalar observables $\varphi$ and $\psi$ (with Hölder regularity). Moreover, the contribution to $C_n$ from the boundary regions can be computed explicitly, as in Sturman and Springham (2013), suggesting that a linked twist map achieves its upper bound on mixing rate, that is,

$$C_n = O(n^{-b})$$  \hspace{1cm} (69)

for some constant $b > 0$, and hence that the exponent in the underlying power law is determined entirely by the boundary conditions.

Dynamical complexity of polynomial order is also a feature of another class of system. Interval exchange transformations (Keane, 1975) and their generalizations known as piecewise isometries (Goetz, 2000) have no hyperbolic behavior and can
be shown to have zero topological entropy (Buzzi, 2001). Instead, their complexity comes from discontinuities in the system and as such can create mixing by cutting and shuffling, rather than by stretching and folding (Sturman, 2012). A natural place to find isometric dynamics is in materials that exhibit highly localized, discontinuous deformations such as slip surfaces and shear bands, e.g., colloidal suspensions, plastics, polymers, alloys, and granular matter. Moreover, some of these classes of materials, along withvalved fluid systems, e.g., piping or vascular networks, multifunctional microfluidic analysis chips, river networks with locks, or the heart, can undergo a combination of cutting and shuffling along with stretching and folding (Smith et al., 2016).

F. Visualization

To inspect mixing in a fluid, it is reasonable to identify and visualize sets connected by fluid transport over a time period of interest. Any dynamics with mixing property would transport material between any two arbitrary sets, given a sufficiently long time window. Conversely, if there is a barrier separating the fluid into isolated islands, the fluid would not be well mixed globally; however, the fluid inside individual islands may still be mixed. Technically, the existence of material transport connecting all subsets over long time windows is ergodicity: a mixing flow is ergodic a fortiori, but the two concepts are not equivalent, as we discussed earlier. Nevertheless, ergodicity can in practice be used as a first approximation to mixing. Construction of the ergodic partition, i.e., the collection of sets whose interiors contain ergodic dynamics, involves the intersection of level sets of a sequence of static scalar fields averaged over Lagrangian particle paths (Mezić and Wiggins, 1999). This process is the basis for a numerical algorithm described by Budišić and Mezić (2012), which not only approximates the ergodic partition, but also aggregates smaller, dynamically similar ergodic sets into larger invariant structures, allowing inspection of fluid features both on coarser and on finer scales; see Fig. 39.

The averaging algorithm for the ergodic partition can handle relatively sparse and restricted sets of initial points from which particles are launched. However, it requires time periods long enough for the particle paths to explore the mixing region. If only short bursts of dynamics are available, but the initial conditions can be densely sampled, an alternative is to form a Ulam approximation of the flow. The Ulam method approximates the transfer operator of the flow, which describes how densities are advected, by a Markov chain transition matrix. To identify and visualize invariant and almost-invariant sets, eigenvectors of the transition matrix are used to color the state space, as described by Dellnitz and Junge (2002), Froyland and Padberg (2009), and Froyland, Santitissadeekorn, and Monahan (2010).

In the ideal setting, the ergodic partition and the Ulam approximation are two sides of the same coin: they relate to the Koopman and Perron-Frobenius operators (Budišić, Mohr, and Mezić, 2012), respectively, which form a dual pair. Practically, the two approaches have different numerical properties. To discover the true ergodic partition, the ergodic partition algorithm relies on long-time averages of a set of functions. Convergence of such averages can be expected in $O(T^{-1})$ for regular dynamics and $O(T^{-1/2})$ for fully chaotic dynamics. For intermittent chaos, however, convergence can be much slower, although regions of intermittency are small in most problems and do not cause problems as explained by Levačić and Mezić (2010). The Ulam approximation, on the other hand, requires only short bursts of dynamics. To form the Ulam matrix, the state space is discretized into cells, which introduces a numerical stochastic element to the dynamics. This poses a problem in regions with many barriers to transport, since a poor choice of discretization can artificially connect otherwise independent regions. As the ergodic partition algorithm classifies trajectories directly, it does not suffer from the same problem.

G. Summary of results on the structure of the mixed state

We briefly summarize the main conclusions of this section. Based upon the area-preservation property of a closed contour under topological transformation in an incompressible 2D flow, one suggestion for quantifying distributive mixing is to use the modified quantities of intensity of segregation and scale of segregation. All quantitative measures such as square density, intensity and scale of segregation, and entropy reveal non-monotonic behavior in time while approaching their limits for a uniform mixture. In order to properly judge how well or how badly two substances are mixed, it is recommended to use both first-order measures (square density $D^1$, intensity of segregation $I$, entropy $e$, and quality $Q$) and second-order measures (e.g., the scale of segregation $L$) of the mixing quality. Large (exponential) stretching rates do not always correspond to the best quality of mixture. Many of these quantities have a dependence on the scale at which they are computed. The concept of mix norm circumvents this problem.

A related point can be made in the context of decay of correlations in smooth ergodic theory. Equations (68) and (69) represent exponential and polynomial decay of correlations, respectively. Certainly, exponential decay will tend to zero at a faster rate than algebraic decay, but there is nothing in Eqs. (68) and (69) to suggest when this faster rate will occur. Although exponential mixing is a desirable property, it is quite possible that in finite time a chaotic system with an exponential tail may be beaten by a system with a polynomial tail. [An alternative
quantification of the rate of mixing was proposed by MacKay (2008), namely, the decay of the “transportation” distance from an advected measure to the equilibrium measure.]

We comment on the relationship between two different viewpoints on the dynamics. Topological ideas have been fruitful in providing bounds on rates of mixing via rates of stretching. In particular, Boyland, Aref, and Stremler (2000) and Thiffeault and Finn (2006) classified and optimized the efficiency of stirring as measured by the stretching of distinguished material lines in the fluid. The resulting braid dynamics imposes a minimum complexity on the flow, and the Thurston-Nielsen classification theorem reveals particular protocols of moving rods that achieve this complexity. This topological approach leads to a lower bound for the best mixing behavior, while the measure-theoretic approach described leads to an upper bound for the worst behavior. These two differing conclusions are the result of effectively considering the same dynamics in two different ways. An understanding of the gap (or perhaps overlap) between the topological bound and the measure-theoretic bound is likely to be a considerable mathematical challenge, but also promises to offer new insight into the details of mixing processes.

VIII. CHAOTIC ADEVCNITION PLUS: SOME ILLUSTRATIVE APPLICATIONS

The general picture presented in this review shows the chaotic advection of fluid as a consequence of the dynamical system induced by the kinematic equation. Fluid elements can take regular paths, where regular regions are barriers to mixing and transport, or chaotic paths, where chaotic regions have rapid mixing. Aref (1990) termed the geometry created by all the paths that can be taken by passive fluid particles the kinematic template of the flow. But while the kinematic template is at the heart of all applications of chaotic flow, few applications involve only the kinematic template of chaotic fluid advection. In applied chaos this “pure” form of motion is almost never encountered. Applications have multiple physical forces driving multiple modes of fluid or particle motion in addition to the continuum stirred advection of a carrier fluid (Metcalfe, 2010; Metcalfe et al., 2012).

The pathway to creative application lies in understanding the interaction of the kinematic template with these other, multiple application-dependent physical forces. However, little systematic knowledge exists on the question of how nonpassive phenomena interact with the kinematic template to change the organization of their own flow and transport properties. For instance, transport of elements that are not passive scalars (e.g., solid particles, liquid droplets, biological cells, small self-propelling organisms) or of materials that can react and diffuse (e.g., heat, chemical species) plays an essential role in many industrial, geophysical, and biological applications. The interaction of the chaotic flow template and these particular physical forces results in an expanded dynamical system with the particles or scalars that is no longer conservative and is instead dissipative (in the dynamical-systems sense of a contracting phase space, not of fluid energy dissipation). This leads to attractors and repellers in particulate systems (Babiano et al., 2000; Cartwright, Magnasco, and Piro, 2002; Torney and Neufeld, 2007; Metcalfe et al., 2012) and strange eigenmodes with diffusive scalar fields (Pierrehumbert, 1994; Liu and Haller, 2004; Lester et al., 2008, 2009, 2010). The kinematic template can be identical in all these cases, but as the additional physics works through and interactions are mediated by the advection, the transport properties of the expanded systems can be quite different.

Examples occur in every domain in which fluid motion exists; Fig. 40 shows schematically some selected applications where mediation of physical forces by the kinematic template is essential. This section can give only a flavor of the full range of applications possible with chaotic advection; next we have chosen some that we find particularly interesting.

A. Heat transfer

Control of heat transfer, its augmentation or suppression, is pivotal for the efficiency of many processes (Gollub et al., 2006). Understanding laminar-flow heat transfer in foods, polymers, energy systems, the Earth’s subsurface, etc., is particularly problematic (Perugini, Poli, and Mazzuoli, 2003; Perugini et al., 2012; El Omari et al., 2015; Petrelli et al., 2016). Without turbulence to continually refresh material near heat transfer boundaries, the temperature gradients at those boundaries lessen and heat transfer slows (Le Guer and El Omari, 2012; Metcalfe et al., 2012). Chaotic advection can provide a means to deal with industrial applications that involve highly viscous complex fluids and highly exothermic reactions that are difficult to control. In such cases, it is important to control the rapid temperature increase resulting from fast reactions in order to obtain the desired products. Heat removal is then a crucial factor in controlling such fast reactions. This problem can be tackled by using a large surface area per unit volume ratio for the reactor (i.e., the case of microreactors), but this solution is sometimes inappropriate when large volumes of fluid are
considered, for example. Another typical industrial solution is to increase surface area by making the fluid transfer conduits very narrow. However, this has a large pumping energy penalty that increases rapidly with the fluid viscosity. These problems are relevant in many conventional industrial domains, in which the potentialities of chaotic advection are not well recognized. In a large majority of industrial reactors, heat transfer for either heating or cooling processes is achieved through the walls of the reactor. Furthermore, many industrial fluids have non-Newtonian rheological behavior (Arratia et al., 2005), and their physical properties can also depend on the temperature. Nowadays, highly energy-efficient heat transfer devices have been designed by considering chaotic flow coupled to the diffusing temperature field (Lester, Metcalfe, and Rudman, 2007; Lester, Rudman, and Metcalfe, 2009; Metcalfe and Lester, 2009).

1. Effect of the Péclet number

As homogenizing the temperature field is equivalent to mixing a diffusing passive scalar, the analysis proceeds via the strange eigenmode solutions to the advection-diffusion equation, Eq. (4). The solution to Eq. (4) is given by a sum of the natural modes \( \phi_k(x,t) \) of the advection-diffusion operator

\[
\phi(x, t) = \sum_{k=0}^{K} \alpha_k \phi_k(x, t) e^{\lambda_k t}, \tag{70}
\]

where the sum is ordered by the magnitude of the real parts of the eigenvalues \( \lambda_k \), with initial weights \( \alpha_k \); and physically the \( \lambda_k \) have negative real parts (Liu and Haller, 2004).

For the problem of material entering a device at one uniform temperature and leaving at a different uniform temperature, only the time to achieve thermal homogenization matters. This time is given by the most slowly decaying term in Eq. (70):

\[
\phi(x, t) \rightarrow \alpha_0 \phi_0(x, t) e^{\lambda_0 t}. \tag{71}
\]

The long-time transport rate is given by \( \text{Re}(\lambda_0) \), and the problem reduces to designing a device that minimizes \( \text{Re}(\lambda_0) \) while keeping relatively open flow conduits, i.e., to simultaneously minimizing the thermal homogenization time and the pumping energy.

Lester, Metcalfe, and Rudman (2007), Lester et al. (2008), and Lester, Rudman, and Metcalfe (2009) used such a procedure to design a heat exchanger based on the rotated arc mixing (RAM) flow (Metcalfe et al., 2006). RAM flow consists of an open tube through which fluid flows. Along the length of the tube portions of the boundary are in motion for a duration \( \tau \), imparting cross-sectional streamlines shown at the top left of Fig. 41. This cross-sectional flow is periodically reoriented by an angle \( \Theta \) as material moves down the tube. The chaotic and symmetry properties of this flow have been extensively studied (Speetjens, Metcalfe, and Rudman, 2006; Speetjens, Rudman, and Metcalfe, 2006; Metcalfe, 2010). The top row in Fig. 41 shows Poincaré sections for passive fluid advection at selected values of \( \tau \) and \( \Theta \) chosen to illustrate the connections between structures of various sizes in the chaotic advection field and the resultant scalar distribution when various levels of diffusion are added. Each column shows the eigenmode associated with the flow at the top for the Péclet number indicated. The initial condition is a Gaussian blob of heat (red color) in the center of the circle. For low enough values of Pe (bottom row) the asymptotic pattern is a deformed version of the initial condition that varies hardly at all even with large structural changes in the fluid advection. With increasing Pe the scalar patterns can either be dominated by large-scale structures in the advection (middle two

FIG. 41. Dominant scalar concentration pattern in a RAM mixer for various reorientation parameter settings of the base flow in the top left, as defined in the text. (Top row) Poincaré sections (essentially infinite Péclet number results). (Columns) Below each Poincaré section are the patterns obtained at the indicated Péclet numbers.

FIG. 42. How a scalar pattern evolves though a cycle of duration \( \tau \) in a RAM flow. At 0 the moving boundary indicated by the arrow moves. Region (a) has been stretched and folded; diffusion heals the pattern along folds. Simultaneously advection stretches and makes new folds around the region (b). The pattern at \( \tau \) resumes its original shape, rotated by \( \Theta \). From Lester, Metcalfe, and Rudman, 2014.
columns) or they can seemingly decouple from the advection field either when island structures in the advection field are small enough (rightmost column) or when there is global chaos (leftmost column). However, other work suggests that the advection and scalar fields are not really decoupled: it is just that Poincaré sections do not show the relation because it occurs in the stretching field.

Figure 42 illustrates the cooperation between advection and diffusion as the pattern $\phi_0$ evolves through one reorientation interval. At 0 the boundary arc indicated by the arrow moves. Region (a) has been stretched and folded during the previous interval and during the subsequent interval diffusion “heals” the pattern along this fold. Simultaneously advection stretches and makes new folds around the region (b). The pattern at $\tau$ resumes its original shape, rotated by $\Theta$. Diffusion heals the pattern wherever folds bring parts of the pattern close together, and folding appears to play a larger role in sustaining the pattern than is often credited.

Experimental data for chaotic heat transfer are more difficult to acquire, depending on what type of data are needed. Overall heat transfer measurements for an optimized RAM heat exchanger show that the heat transfer rates are 4–10 times larger than for an open tube with total pumping energy reductions of 60%–80% compared to an open tube, depending on Pe and fluid rheology. Temperature pattern data are scarce, but experiments are beginning to produce them: Fig. 43 shows the measured temperature field compared with a numerical computation of the same flow (Baskan et al., 2015).

2. Thermal boundary condition

Consider a typical mixer composed of two circular rods inside a cylindrical tank. The tank and the rods are heated or cooled and can rotate around their respective axes. This two-rod mixer is suitable for obtaining global chaotic flow without large KAM regions (El Omari and Le Guer, 2009, 2010a). Consider a highly viscous fluid of $Re \approx 1$ with a high Prandtl number $Pr = 10^4$, so that it is difficult to mix. Because of the combination of mixing with heat transfer, the wall boundary condition plays a particular role that is different from the mixing of a passive scalar considered without heating or cooling. The role of the wall has been discussed in Sec. II where the notion of a chaotic saddle in open flow is presented and in Sec. III where the case of a slip boundary is considered in a 2D bounded flow. El Omari and Le Guer (2010a, 2010b) have shown for both Newtonian and non-Newtonian fluids and for a constant wall temperature boundary condition that the efficiency of thermal exchange is strongly dependent on the choice of the stirring protocol imposed on the walls. The main conclusion from these works is that maximizing heat transfer from the wall boundaries requires that the walls (i.e., the tank or rods here) move alternately. This way, one avoids the development of closed streamlines near the wall, which reduces advection from the wall zones and prevents the persistence of confined hot or cold fluid zones. Continuous modulation of wall movement is not sufficient to produce effective chaotic mixing and the existence of stagnation points on the static walls is necessary to create separatrices from which heteroclinic tangles are created (see Fig. 44). Thus, thermal chaotic mixing is controlled by the topology of the flow near the moving or nonmoving walls. However, the choice of a thermal wall boundary condition, either constant wall boundary condition or constant wall heat flux, gives rise to fundamental differences in the evolution of the fluid temperature and its homogenization.

For a constant wall temperature boundary condition (i.e., Dirichlet condition), the imposed wall temperature represents an asymptotic limit for the evolution of the mean fluid temperature. The dots at the wall indicate stagnation points from which fluid is extracted toward the center of the mixer. Large zones of unheated fluid are observed in the CM case. Case of a constant wall temperature boundary condition. From Le Guer and El Omari, 2012.

![FIG. 43. The two leading eigenmodes (labeled 0 and 1) of the temperature evolution for a RAM flow at $Pe = 10^4$ from (left column) computation and (right column) experiment. For numerical data $Re = 0$; for experimental data $Re = 0.1$. From Baskan et al., 2015.](https://example.com/fig43)

![FIG. 44. Two-rod mixer: Instantaneous temperature fields and streamlines after four periods of stirring. The rods and the tank rotate in the same direction. Continuously modulated stirring protocol (CM—left) vs alternated stirring protocol (ALT—right). The dots at the wall indicate stagnation points from which fluid is extracted toward the center of the mixer. Large zones of unheated fluid are observed in the CM case. Case of a constant wall temperature boundary condition. From Le Guer and El Omari, 2012.](https://example.com/fig44)
temperature, and the rapidity with which this limit is reached is controlled by the efficiency of the mixing. In contrast, for a constant heat flux boundary condition (i.e., Neumann condition), there is no asymptotic limit for the evolution of the mean fluid temperature because its evolution is prescribed by the imposed heat flux density. In this case, the efficiency of mixing (of the stirring protocol) controls the heat homogenization only. As a consequence, the stirring strategy and the choice of the mixing measures must be selected in agreement with the type of the wall heating considered. This is illustrated in Fig. 45 for the two different boundary conditions and for three stirring protocols [nonmodulated (NM), continuously modulated (CM), and alternated (ALT)]. This figure displays the greater efficiency of the alternated stirring protocol for both thermal boundary conditions. Similarly to the nonmodulated stirring protocols, the continuously modulated stirring protocols give rise to closed streamlines in the vicinity of the walls that prevent the radial transport of the temperature scalar inside the mixer. Thus, close to the rotating boundaries, radial heat transfer is achieved only by conduction across the streamlines; these streamlines act as an insulation medium (Gouillart, Thiffeault, and Dauchot, 2010b).

3. Effect of fluid rheology

In order to show the impact of fluid rheological behavior on the pattern of the temperature field, we illustrate in Fig. 46 the temperature fields and the corresponding streamlines for shear-thinning (left) and shear-thickening (right) fluids in the case of the alternated stirring protocol. The Ostwald–de Waele power-law model \( \eta(\dot{\gamma}) = k(\dot{\gamma})^{n-1} \) is chosen for the apparent viscosity. Two values of the behavior index \( n \) are considered, \( n = 0.5 \) and 1.5, corresponding to shear-thinning and shear-thickening fluids, respectively. The values of the consistency index \( k \) are adjusted to keep the generalized Reynolds number the same as the corresponding Newtonian case of Fig. 44. Compared to Newtonian and shear-thinning fluids, the shear-thickening fluid is more driven by the wall due to its higher viscosity, resulting in larger vortices and more extended hot fluid streaks originating from the parietal parabolic points. This greater fluid driving also increases the folding of temperature striations. Thus, the mixing efficiency
for this fluid is higher than that of the two other fluids. The vortices created in the shear-thinning fluid are smaller because of the viscosity reduction in the sheared fluid near the rotating walls. Fluid driving forces and mixing efficiency are then weak. The effects of the temperature dependence of viscosity on chaotic mixing are also important. As discussed, fluid viscosity in the vicinity of the moving walls has a major influence on the fluid displacement and on the creation of large vortices accompanied by separatrices that help to promote mixing. When the temperature dependence is taken into account, the heating and cooling processes have different evolutions. This temperature dependence can be modeled by an exponential model as $\eta = \eta_0 \exp(-BT^*)$ (Le Guer and El Omari, 2012). The results obtained for the Pearson number $B = 5$ (relevant to melted polymers, for example) are shown in Fig. 47 and may be compared to those of the non-temperature-dependent case ($B = 0$) of Fig. 44 (right). In contrast to the non-temperature-dependent case, the velocity field for $B \neq 0$ is not reproducible from period to period owing to the continuous evolution of the viscosity, and therefore the mixing mechanisms are not exactly the same from the beginning of the mixing to its end. The streamlines of Fig. 47 show distinctly different patterns for heating and cooling and both are different from the $B = 0$ case [Fig. 44 (right)]. This difference is salient for the heating process, for which the viscosity is low in a thin fluid layer around the walls and higher elsewhere; thus, there is slippage between the fluid and the wall (i.e., the momentum diffusion is weak), and the fluid is not well driven. On the other hand, the lack of the fluid driving force of the moving walls is balanced by the ease of moving the less-viscous fluid at the static walls. The mixing mechanism of a temperature-dependent fluid is quite complex, with several superimposed effects, even for a Newtonian fluid. Also, flow morphologies evolve in time, and as a remarkable consequence, the strange eigenmodes existing in the non-temperature-dependent case (Fig. 48) are absent in the $B = 5$ case because there are no recurrent patterns in the temperature field.

An important feature of thermal mixing in comparison with the mixing of a scalar substance is that the scalar source is located at the wall boundary. Thus the mode of heating (or cooling) of the boundary directly influences the characteristics of the mixing and heat transfer which are also significantly affected by the rheological behavior of the fluid through the shear rate dependence on viscosity. The thermal strange eigenmodes are always present for a non-Newtonian fluid (with a shear-thinning or shear-thickening behavior) but they disappear when the viscosity is temperature dependent because of the spatial modification of the flow field over time.

B. Microfluidics

In recent years, the transfer of lithographic techniques to the microfluidics industry has led to a renascence of the chaotic advection paradigm for the mixing of fluids at the lowest Reynolds numbers. Although microfluidic mixing is a key process in a host of miniaturized analysis systems, with useful applications for chemical reactions, crystallization, polymerization and organic synthesis, biological screening, polymerase chain reaction amplification, and a great many other fields, it continues to pose challenges owing to constraints associated with operating in an unfavorable laminar-flow regime characterized by a combination of low Reynolds numbers and high Pécelt numbers. A wide variety of micromixing approaches have been explored, most of which can be broadly classified as either active (involving input of external energy) or passive (harnessing the inherent hydrodynamic structure of specific flow fields to mix fluids in the absence of external forces). Here we review some of the most promising approaches.

1. Passive mixing

Passive designs are often desirable in applications involving sensitive species (e.g., biological samples) because they do not impose strong mechanical, electrical, or thermal agitation. The microchannel structures associated with these mixing elements range from relatively simple topological features on one or more channel walls (ridges, grooves, or other protrusions that can, for example, be constructed by means of multiple soft lithography, alignment, and bonding steps) to intricate 3D flow networks. Ultimately, it would be desirable to achieve gentle passive micromixing in the shortest possible downstream distance by using simplified microchannel geometries (ideally, planar 2D smooth-walled) that can be easily constructed, ideally, in a single lithography step.

One of the best-known examples, the staggered-herringbone micromixer (SHM) shown in Fig. 49(a) (Stroock et al., 2002), subjects the fluid to a repeated sequence of rotational and extensional local flow that, as result, produces a chaotic flow. The internal structures endow SHMs with high mixing efficiency, allowing a short mixing length ($\sim 1-1.5$ cm) at high Pécelt number ($Pe \sim 10^5$). A series of improved grooved patterned micromixers has been proposed since then. More convoluted devices use a more faithful realization of the Baker’s map by a “split-and-recombine” strategy where the streams to be mixed are divided or split into multiple channels and redirected along trajectories that allow them to be reassembled subsequently as alternating lamellae

![Fig. 48. Two-rod mixer: Snapshots at different periodic times show the evolution of the $T^*$ field toward a strange eigenmode (top row), and snapshots rescaled between $T^*_\text{min}$ and $T^*_\text{max}$ (bottom row). From Le Guer and El Omari, 2012.](image-url)
Another interesting strategy for passive devices is the use of finite Re effects to force the fluid through an efficient chaotic transformation: examples include the use of Dean vortices in curved channels depicted in Fig. 49(b) (Sudarsan and Ugaz, 2006a, 2006b), or the incorporation of elastic passive elements in the micro-channel to transform a simple steady laminar flow into a forced oscillatory flow through hydroelastic instabilities (Xia et al., 2012).

2. Active mixing

In contrast to the passive types, active micromixers can keep their functions under tighter control in most cases. External perturbation sources, such as pressure, thermal, electrokinetic, magnetic, and acoustic disturbances, can be designed to induce chaotic advection. For instance, externally imposed oscillatory flows driven across a main channel can stretch and fold fluid in the primary flow into itself, significantly enhancing mixing. This secondary oscillatory flow can be generated by means of controlled pressure, as in Fig. 50(a) (Glasgow and Aubry, 2003; Okkels and Tabeling, 2004; Tabeling et al., 2004; Bottausci et al., 2007), via low frequency switching of transverse electro-osmotic flows generated on integrated microelectrodes (Oddy, Santiago, and Mikkelsen, 2001; Song et al., 2010), through ultrafast ion depletion and enrichment by polyelectrolytic gel electrodes (Chun, Kim, and Chung, 2008), or by trapping and exciting acoustic streaming of air bubbles with external piezotransducers, as illustrated in Fig. 50(b) (Liu et al., 2002; Ahmed et al., 2009; Wang, Huang, and Yang, 2011).

3. Mixing in microdroplets

Droplet-based microfluidic systems, sometimes referred to as "digital microfluidics," use two immiscible fluids such as water and oil to produce highly monodisperse emulsions in microchannels with a small size variation. By varying the viscosity of the two phases, the relative flow rates, or the channel dimension, it is possible to tune the dimensions of the microdroplets produced. Such microdroplets can be used to encapsulate, for instance, a number of chemical reagents that need to be mixed quickly for a chemical reaction to take place. This compartmentalization is ideally suited to a large number of applications, including the synthesis of biomolecules, drug delivery, and diagnostic testing (Teh et al., 2008). The discrete nature of microdroplets and the feasibility of individual control of their distinct volumes of fluids contrasts with the continuum nature of other systems and emphasizes their use as a versatile high-throughput platform.

Interaction between the droplet surface and the channel walls can be used to generate recirculating flows within the droplet (Song et al., 2003; Tice et al., 2003). The introduction of turns and bends in the channel induces an asymmetric flow pattern within the drop that can be used to periodically perturb the steady flow and, hence, induce mixing by chaotic advection. The best-known device based on this principle is the "planar serpentine micromixer" (PSM) shown in Fig. 51(a).
(Song et al., 2003; Bringer et al., 2004; Tice, Lyon, and Ismagilov, 2004) in which the extent of mixing can be controlled by altering the number and distribution of turns in the microchannel. Conversely, the microdroplet surface can be perturbed by external forcing. Examples include the generation of surface-tension gradients, for instance, generated by heating from a laser beam (Grigoriev, Schatz, and Sharma, 2006), or the fine-tuning of surface electrowetting induced by microelectrodes (Paik, Pamula, and Fair, 2003; Paik et al., 2003) illustrated in Figs. 51(b) and 51(c), respectively.

These classifications are by no means an exhaustive inventory of the available techniques, as the field is rapidly evolving and new configurations and implementations for effective mixing in microchannels based on the chaotic advection paradigm emerge every day. For further reading on this topic see recent compendia of the field; see, e.g., Squires and Quake (2005), Tabelin and Chen (2005), Lin and Basuray (2011), and Nguyen (2011) and references therein. The ubiquity of the phenomenon in its diverse forms emphasizes the relevance of the physics of mixing for the bulk of the miniaturization industry and, more particularly, for lab-on-a-chip applications.

C. Biology

1. Ciliary and flagellar chaotic advection

Flagella and cilia are microscopic organelles, flexible ropelike driven structures that form part of a single cell, that act as ubiquitous biological stirrers. These organelles are one of the most highly conserved structures in biology (Gibbons, 1981; Gardiner, 2005; Marshall and Nonaka, 2006) and, as such, they have a wide range of biological fluid-mechanical functions: from propulsion in flagellated and ciliated microorganisms (Taylor, 1951; Lighthill, 1976; Brennen and Winet, 1977; Drescher et al., 2010) and the sperm and ovum in reproduction (Fauci and Dillon, 2006; Gaffney et al., 2011) to feeding and filter feeding in sessile organisms (Blake and Sleigh, 1974) and mollusks (Blake and Fulford, 1995), and as a physical guide for diverse developmental processes (Cartwright, Piro, and Tuval, 2009; Freund et al., 2012).

The flow induced by the beating of cilia and flagella, which falls in the regime of very low Reynolds numbers, has been modeled by means of fundamental singularity for Stokes flow: as single point forces (Stokeslets) or torques (rotlets) (Blake and Otto, 1996; Cartwright, Piro, and Tuval, 2004; Niedermayer, Eckhardt, and Lenz, 2008) or as a line distribution of these singular solutions (Smith et al., 2012) with, as expected from Stokes flows, boundaries playing a major role on the extent, topology, and magnitude of the fluid flow (Blake, 1971; Blake and Sleigh, 1974; Montenegro-Johnson et al., 2012). However, the role of cilia and flagella in biological fluid mixing has been much less studied.

The sessile ciliated protozoan Vorticella, which lives in freshwater ponds, has attracted considerable interest. The flow generated by the continual beating of its oral cilia [Fig. 52(a)]

Rev. Mod. Phys., Vol. 89, No. 2, April–June 2017 025007-47
is well captured by a Stokeslet near a flat rigid no-slip boundary, with the force acting normal to the boundary and, hence, generating a toroidal eddy (Blake and Otto, 1996; Nagai et al., 2009; Pepper et al., 2010). Reports on the sudden changes in the length of the stalk linking the cell and the substrate to which it is anchored led to a series of studies on mixing by a “blinking Stokeslet” (Blake, 2001; Orme, Otto, and Blake, 2001a, 2001b; Otto, Yannacopoulos, and Blake, 2001), a model inspired on the now classical “blinking vortex” model (Aref, 1984; Aref and Balachandar, 1986; Meleshko and Aref, 1996) but in which the height of the point force above the no-slip boundary changes periodically. Furthermore, Vorticella’s bell changes its orientation erratically (Pepper et al., 2013) which can be modeled with time-dependent variations in the direction of the applied point force. Both mechanisms might enhance the feeding process of Vorticella by allowing nearby nutrients to be driven through filtering cilia on the cell surface. More recently Vorticella has even inspired a biomimetic solution for mixing in microfluidic chips (Nagai et al., 2014).

Other ciliated organisms (e.g., Opalina, Paramecium, and Volvox) and ciliated tissues (e.g., epithelia, corals) show a remarkable coordination of nearby beating cilia, including complete synchrony (Goldstein, Polin, and Tuval, 2009, 2011) and metachronal waves (i.e., large-scale modulations of the beating phase) (Gueron et al., 1997; Lenz and Ryskin, 2006; Brumley et al., 2012, 2015; Elgeti and Gompper, 2013). The unsteady flow generated by their coordinated beating leads to stretching and folding and, over time, to mixing of the adjacent fluid (Solari et al., 2006; Ding et al., 2014; Hassan Aref et al.: Frontiers of chaotic advection Rev. Mod. Phys., Vol. 89, No. 2, April–June 2017 025007-48...
2. Biological activity in chaotic oceanic flows

The growth of marine organisms is closely related to ocean currents. Since hydrodynamic flows constitute an important physical factor for productivity in the ocean, the interplay between the physical environment and biological growth is an increasing field of research (Mann and Lazier, 1991; Denman and Gargett, 1995; Károlyi et al., 1999; Peters and Marrasé, 2000; Tél et al., 2005). In particular, the growth of plankton, which is the basis of the food web in the ocean, is influenced by currents. Plankton species and their nutrients are only passively transported, while higher organisms (nekton) have the ability to swim actively and are therefore more independent of the flow. Starting with the work of Abrahám (1998) and Bees, Mezić, and McGlade (1998), who first applied the concept of chaotic advection to the study of plankton patterns in the ocean, various others have contributed to this subject (López, Hernández-García et al., 2001; López, Neufeld et al., 2001; Martin et al., 2002; Martin, 2003; Sandulescu et al., 2007).

One of the major requirements for the growth of plankton species is the availability of nutrients, which depends crucially on the hydrodynamic flow patterns in the oceans. Of particular interest in the context of plankton growth are mesoscale flow patterns such as fronts, jets, and vortices, which have a strong impact on plankton patterns. Although the full velocity field in the ocean is 3D, giving rise to horizontal and vertical transport of nutrients and plankton, a 2D field is often a rather good approximation for the study of plankton patterns. On the one hand, the vertical velocity component is frequently much smaller than the horizontal ones; on the other hand, since phytoplankton (being the plants of the ocean) need sunlight for photosynthesis, their abundance is confined to the upper layer of the ocean. These features make the investigation of phytoplankton blooms, i.e., the emergence of high phytoplankton abundances, an ideal application field for chaotic advection. Although ocean circulation models are far more realistic, the study of simple flows leading to typical chaotic advection problems has revealed much insight into the biological-physical interactions. This is due to the much higher spatial resolution of the flow structures which can be achieved by simple, mostly analytically given flows compared to ocean flows. In particular, the emerging fine structure of filamentary plankton patterns is responsible for certain phenomena that cannot be explained by coarse-grained oceanic flows. Concepts of chaotic advection have led to fundamental contributions to the understanding of the mechanisms of plankton-bloom formation as well as of the coexistence of plankton species.

To investigate the impact of mesoscale hydrodynamic structures on plankton growth, a variety of different kinematic models for the flow have been considered. Taking into account the arguments mentioned, 2D horizontal flow patterns have been studied in which the velocity field can be described by a stream function. Three paradigmatic models exhibiting certain important features of an oceanic flow have been used to elucidate the interplay between plankton dynamics and hydrodynamic flows: (i) the blinking vortex flow (Neufeld et al., 2002), (ii) the flow in the wake of an obstacle (Jung, Tél, and Ziemniak, 1993; Sandulescu et al., 2006), and (iii) a jet flow (López, Neufeld et al., 2001). All of these flows are periodically forced, which leads to chaotic advection of passive tracers. Each of them focuses on particular properties of a real flow. While (i) introduces a temporarily changing mixing region, (ii) mimics the dynamics of a von Kármán vortex street in the wake of an island located in an oceanic current, and (iii) is a simple description of a meandering jet. Besides these idealized flow fields, 2D turbulence models are used in more realistic settings (Martin et al., 2002). These velocity fields are coupled to simplified nonlinear plankton growth models that often consist of two or three interacting species, e.g., nutrients, phytoplankton, and zooplankton. In some cases plankton blooms are modeled with excitable dynamics (Neufeld et al., 2002). The combination of these simplified plankton models and the above-mentioned kinematic flows in the corresponding reaction-diffusion-advection equations are the basis of the study of chaotic advection in plankton growth.

Using these approaches it has been shown that the redistribution of nutrients and plankton by horizontal stirring can enhance the growth of plankton (Abrahám, 1998; López, Neufeld et al., 2001; Hernández-García, López, and Neufeld, 2002, 2003; Martin, 2003) and even cause (Károlyi et al., 2000) and sustain (Hernández-García and López, 2004) a plankton bloom. Moreover, mesoscale hydrodynamic vortices can serve as incubators of plankton blooms by keeping the plankton in the interior of the vortex without much exchange with its exterior (Martin, 2003; Sandulescu et al., 2007). In general the time scale of the hydrodynamic processes is much faster than the time scale of biological growth. However, the confinement of plankton within the vortex leads to a much larger residence time of the planktonic organisms in a certain area, so that growth processes can take place; see Fig. 53.

An even simpler approach for growth processes has been studied by considering autocatalytic reactions (Metcalfe and Ottino, 1994; Toroczkai et al., 1998; Károlyi et al., 1999; Tél et al., 2000). Planktonic organisms are considered as point tracers that are advected by the flow and that undergo a reaction $A + B \rightarrow 2B$. Using this modeling technique it has
been shown that the biological activity is enhanced since the fractal filaments formed by chaotic advection lead to an increase in the reaction surface and, hence, of the reaction rate. The fractal skeleton of the reaction is characterized by its capacity dimension $D_0$, which in turn is an essential parameter that can be used to estimate the scaling properties of the reaction rate (Tel et al., 2004). On the other hand, the enhancement of autocatalysis due to chaotic advection has been proposed as an explanation for deracemization during the chiral crystallization of achiral molecules (Metcalf and Ottino, 1994). In this case, an agent-based model constructed to describe the breakdown and subsequent reproduction of growing crystals submitted to shear stresses is immersed in a chaotic flow. As a result an unbalanced population of crystals with both chiralities is obtained (Cartwright et al., 2004). When the agents are endowed also with a dynamics mimicking the Oswald ripening phenomenon, one of the chiralities ends up being completely annihilated (Cartwright, Piro, and Tuval, 2007).

The filamentary structures that develop in chaotic advection are not only advantageous for the growth of species but lead also to an increased separation of different species competing for the same nutrients (Bracco, Provenzale, and Scheuring, 2000; Karolyi et al., 2000; Scheuring et al., 2003; Bastine and Feudel, 2010). As a consequence, competition is diminished, so that more species feeding upon the same resources can coexist. Weaker species with smaller abundance obtain more access to resources along the fractal filaments and are able to survive and compete with stronger competitors. This effect is even stronger when the inertia of the organisms is taken into account (Benczik et al., 2006). In this way, chaotic advection can be considered as a possible mechanism to solve the paradox of plankton (Hutchinson, 1961), an intriguing problem in theoretical ecology.

IX. PERSPECTIVES

For efficient mixing to be achieved, the velocity field must stir together different portions of the fluid to within a scale that is small enough for diffusion to take over and homogenize the concentrations of the advected quantities. One traditional way to do this has been through fully developed turbulence, where the coexisting eddies of many different scales do the job of generating the large concentration gradients at small scales which molecular diffusion smoothes out. The other approach, which we review in this work, is to generate these small-scale structures through the repeated processes of stretching and folding that characterize deterministic chaos. Unlike turbulent advection, chaotic advection does not involve an energy cascade, and it works even in flows that are too viscous or too small scale to have large Reynolds numbers, or are effectively 2D. This generality provokes questions, conjectures, remarks, and reflections from a wide range of applications and theory. In this section we raise some of the current challenges and opportunities in chaotic advection.

We do so under the assumption that an engineer faced with a mixing problem, whether from a traditional industry such as textiles, food, or chemicals, or a more technologically sophisticated field such as microscience or nanoscience, is already well aware of the potential advantages that chaotic advection can bring. The notion of mixing using chaotic dynamics is by now common place. What then are the timely and relevant questions to ask the engineer, and which aspects of theory and practice can now be developed to answer questions posed by the engineer in return? Such questions may include the difficulty of optimizing a mixing process, both in general and in particular circumstances; the effect of specific physical features of the flow, such as fluid inertia, boundaries and walls, and rheological behaviors; and how to describe, analyze, and quantify chaotic advection in the most appropriate way. Answers to such questions require a synthesis between rigorous mathematics, experimental pragmatism, physical understanding, and engineering intuition.

A. How to choose the best stirring protocol for a given mixing problem

To maximize the performance of a mixing device there are typically many parameters, several of which are particular to that device. Some features are common to a range of generic devices, however, and as such suggest general schemes for optimization.

For example, blinking flows and channel mixers have a natural temporal or spatial periodicity to which classical dynamical-systems theory can be applied. Less well understood is the influence of aperiodicity in stirring protocols (Liu, Muzzio, and Peskin, 1994). Between the two extremes of a periodic stirring protocol and a random one, one can introduce a small perturbation in a periodic sequence in order to study the effect of the degree of aperiodicity. The periodic length of a sequence in the stirring protocol is also a parameter to test. It has been shown that the choice of a globally aperiodic sequence is not appropriate for an efficient stirring protocol (Gibout, Le Guer, and Schall, 2006). Moreover, in the case of fixed-time switching between horizontal and vertical shears on a periodic domain, the best mixing (in the sense of Kolmogorov-Sinai entropy) has been rigorously shown to be achieved by the simplest periodic protocol (D’Alessandro, Dahleh, and Mezić, 1999). This was later used to prove optimality of mixing protocols by Boyland, Aref, and Stremler (2000).
The geometrical details of a mixer can also be considered parameters for optimization. For example, is the angle between superimposed crossing streamlines taken at two different times, and their distribution over the fluid domains, a good criterion for the optimization of mixing? One attempt to investigate this can be found in Sturman and Wiggins (2009). Does good mixing always require a fairly uniform distribution of stretching rates within the flow? How does the movement of a mixer wall or its radius of curvature affect the amplitude of the modulation, and in turn the mixing quality and efficiency? What is the relation between stretching fields and curvature fields, or the differences between cases of bounded and unbounded flows for the same stirring protocols? What is the influence of the phase shift between the waveforms applied to different rotating or moving walls? How does the response of the scalar field orientation to a modification of the stirring protocol impact the quality of mixing? Many of these questions do not yet have full answers; see also Balasuriya (2005a, 2005b, 2010) and Cortelezzi, Adrovier, and Giona (2008).

In applications, achieving a sufficiently good mixing performance for the least energy input may be the chief motivation (Alvarez-Hernández et al., 2002). Of the many existing mixing indicators and diagnostics discussed earlier, those which take into account the energy communicated to the fluid, or the energy saved, may be the most practical. Mathew et al. (2007) coupled the two for the first time, using the mix-norm metric supplemented by an energy measure. Lin, Thiffeault, and Doering (2011) advocated the use of an $H^{-1}$ norm, but $H^{-1/2}$ seems to have better physical motivation, as it incorporates the idea of equal distributions in equal volumes for the perfect mixture. As always, the distinction between unsteady 2D flows and 3D steady flows is likely to be challenging, as is that between active and passive modes of generation of the advective flow.

As the question of how to pick a stirring protocol only has a definitive theoretical answer in a specific model system, various computational approaches have been developed to approximate an answer in applications. For diffusive scalar transport the work of Lester et al. (2008, 2010), Lester and Metcalfe (2009), Lester, Rudman, and Metcalfe (2009), and Lester, Metcalfe, and Rudman (2014) has produced a fast method to calculate the most slowly decaying (or the first few) eigenmodes when the stirring flow consists of one or a few basic flows (and symmetries of these base flows) that are governed by a set of control parameters. The general method (Lester et al., 2008) incorporates any symmetry, e.g., rotation, reflection, scaling, or superposition. However, it still remains up to the flow designer to determine the decomposition of fundamental flows and symmetries.

For example, the RAM flow discussed in Sec. VIIA has one fundamental flow, from which all flows in the device are obtained by rotation, and in the simplest case two control parameters $\tau$, the time any one flow operates, and $\Theta$, the fixed rotation angle of the flow. Figure 54 is a contour plot of the asymptotic decay rate $\text{Re}(\lambda_0)$ calculated at $1.2 \times 10^3$ points over the $\tau-\Theta$ plane for $\text{Pe} = 10^3$ and homogeneous Dirichlet boundary conditions; $\lambda_0$ is scaled by the diffusion rate (the most slowly decaying eigenvalue of the diffusion operator in the disk). This figure maps transport enhancement relative to diffusion alone and shows what can be termed a complete parametric solution for scalar transport, in that the optimum stirring protocol—for the goal of thermal homogenization—can be read directly from the graph to the precision of the computation. Lester et al. (2008) reported that for symmetry-derived stirring protocols it is possible to obtain the complete parametric solution about 6000 times faster. The relative performance improves as $\text{Pe}$ increases, but the absolute time to obtain the complete solution also increases.

The structure of the scalar transport solution is worth commenting on. At low values of $\tau$ the enhancement distribution is fractal with many localized maxima, which can be discerned only with highly resolved solutions on the control parameter space. The spiky regions originate from rational values of $\Theta/\pi$ at $\tau = 0$ and grow in width with increasing $\tau$. Inside each region the scalar spatial distribution is locked into a symmetric pattern whose azimuthal wave number $m$ is a rational multiple of the forcing wave number $k = 2\pi/\Theta$; these regions are symmetry-locked “tongues” similar to frequency-locked Arnol’d tongues (Glazier and Libchaber, 1988). This type of behavior was first observed in the context of the micromixer design by Mathew et al. (2004) using the mix norm as the measure of goodness of mixing. As $\tau$ becomes $O(1)$ the spreading tongues interact to produce an order-disorder transition, where the symmetric spatial patterns change into asymmetric patterns. At large values of $\tau > 10^3$ the enhancement ratio everywhere takes on the same value as that on the $\Theta = 0$ line, i.e., no enhancement. For low $\text{Pe}$, such as $\text{Pe} = 10^3$ as in Fig. 54, the maximum transport enhancement occurs in the 1/3 resonance tongue at low values of $\tau$. This seems counterintuitive because the no-diffusion mixing optimum (Metcalfe et al., 2006; Speetjens, Metcalfe, and Hassan Aref et al.: Frontiers of chaotic advection

Rev. Mod. Phys., Vol. 89, No. 2, April–June 2017 025007-51
B. Effect of fluid inertia

In most chaotic advection studies, a Stokes flow regime is envisaged with low Reynolds numbers and a quasi-steady-state approximation of low Strouhal number, Sr. This implies that viscous forces dominate the flow dynamics and leads the Navier-Stokes equations to be linear. The linearity of the problem gives rise to reversibility of the flow and a passive blob of tracer will return to its original place if the effect of diffusion is weak or not amplified by chaotic advection. Indeed, for chaotic advection flows, the trajectories can be so complex in space and time that a particle will lose the memory of its initial starting position if we reverse the flow, due to the accumulation of diffusion responsible for tiny jumps from streamline to streamline, trapping events, or Lévy flights in hyperbolic regions (Solomon, Weeks, and Swinney, 1994; Solomon, Lee, and Fogleman, 2001). If fluid inertia is included (i.e., when Re is increased but the flow is still laminar), the nonlinearity of the Navier-Stokes equations introduces a new source of transient dynamics that leads to the breakdown of invariant manifolds (see Sec. IV.C and the example of 3D unsteady flow in mergera cylinder) and strongly impacts on the transport properties of the flow. How exactly the increase in Re changes mixing properties is an open problem at the intersection between Eulerian and Lagrangian analyses. There has been relatively little work on chaotic advection that considers the effect of fluid inertia (Dutta and Chevray, 1995; Mezić, 2001b, 2002; Horner et al., 2002; Balasuriya, Mezić, and Jones, 2003; Speetjens, Clercx, and van Heijst, 2006a, 2006b; Wang et al., 2009; Pouransari, Speetjens, and Clercx, 2010). Mezić and collaborators focused on possible nonmonotonic mixing created by inertial effects and specifically on the possibility of a decrease of mixing because of these. Follow-up experimental and numerical studies confirm these predictions (Lackey and Sotiropoulos, 2006; Pratt et al., 2014). We emphasize that inertia is not the only mechanism to induce chaos. It is also possible to obtain chaos from a breaking of time reversibility by the time derivative in the time-dependent Stokes equation (Eckhardt and Hascoët, 2005), or simply by minor imperfections in an experimental setup (Wu et al., 2014).

Inertial effects are also beginning to be used in microfluidic systems (Di Carlo, 2009) where chaotic advection is now commonly employed (see Sec. VIII.B). Horner et al. (2002) introduced time dependence through transient acceleration and deceleration of a flat wall driving the flow in a cavity and, despite the fact that the streamline portrait changes very little during the transition, they observed a significant transport enhancement with the increase of the control parameter ReSr. In general, fluid inertia also introduces supplementary secondary flows that bring new fixed parabolic or hyperbolic points from which stable and unstable manifolds can emanate, thereby improving the stirring process. At the same time, inertial forces influence chaotic advection by causing a distortion of the streamlines as well as making them time dependent (Wang et al., 2009).

The effect of inertia is crucial for mixing in chaotic flows through a sequence of alternating curved ducts (Jones, Thomas, and Aref, 1989; Castelain et al., 2001) for which it is necessary to have a large enough Dean number (Dean, 1928) to produce a minimum transverse displacement of a passive scalar along a helical trajectory, otherwise the stirring will be ineffective. Such an arrangement of curved ducts is able to generate very complex flows with the presence of regular, partially, or fully chaotic regions; they are used as mixers or reactors (Boesinger, Le Guer, and Mory, 2005). For a particular geometry, the coexistence of nonmixed islands and chaotic zones in the cross section with the presence of the solid wall boundary highly depends on the flow rate and is a source of multimodality of the finite-time particle distributions along the pipe mixer, while a sole ergodic region in the transverse cross section (globally chaotic flow) gives rise to a narrow distribution of residence times (Mezić, Wiggins, and Betz, 1999).

Fluid inertia is also important in oscillating flows. Hydon (1994) showed, for a pulsed flow in a curved pipe, that particles on certain trajectories are transported by resonant interaction between the secondary orbital motion and the longitudinal oscillation, inducing ballistic transport. For some trajectories trapped behind cantori, particles are transported intermittently by resonance; this mechanism leads to anomalous diffusion (Young and Jones, 1991). An important physiological example is blood flow around a stenosis zone (Schelin et al., 2009; Maiti et al., 2013), for which chaotic advection is triggered by a combination of pulsating flow driven by the heart and constriction of the blood-vessel walls.

In the domain of geophysical flows, Pratt et al. (2014) recently revisited the problem of chaotic advection within a 3D steady flow in a closed cylinder with a rotating bottom. They showed a nonmonotonic variation of the stirring rate with the Reynolds number (more exactly with the Ekman number Ek, which is directly proportional to the inverse of the Reynolds number for a fixed Rossby number). The bulk stirring rate, estimated from the tracer concentration variance, has a maximum at intermediate Ek−1/2 and a complex dependency of the Lagrangian structures is noted, with thin or thick resonant layers sandwiched between KAM barriers; cf. Mezić (2001b).

The issue of fluid inertia also arises when addressing the question of periodicity of the flow which does not necessarily match the geometrical periodicity of the wall boundaries chosen for the generation of chaotic advection (Blancher, Le Guer, and El Omari, 2015). This has a direct implication for the appropriate choice of the periodic velocity field in order to map the flow as in the context of dynamical systems.

When we consider nonpassive scalar transport (i.e., difusive or nonbuoyant scalars) such as finite-mass and finite-size particles, the transient effects due to particle inertia and fluid inertia are intimately linked owing to the presence of new forces. Indeed, a nonpassive particle responds to intrinsic hydrodynamic forces, to externally imposed forces, or to internally generated forces (Metcalfe et al., 2012). The hydrodynamic forces include drag force, lift force, added mass force, buoyancy force, Boussinesq-Basset history force
The dynamics of a finite-size particle in a chaotic flow is thus nontrivial and can lead to very different and rich behaviors because the particle’s trajectory can diverge from the trajectory of the fluid parcel, since a slip velocity exists between the particle and the surrounding fluid whose velocity can eventually fluctuate in a random manner. The fact that the finite-size particle dynamics is dissipative furthermore complicates the story by allowing the existence of attractors in the phase space (Cartwright et al., 2010). Figure 55 shows trajectories of several finite Stokes number particles localizing onto an attracting KAM tube in a 3D laminar flow. Predictions of what level of inertia will cause spontaneous localization of finite inertia particles have been quantified in recent experiments (Wang et al., 2014; Wang, Stewart, and Metcalfe, 2016). A direct consequence is that invariant curves no longer exist for the case of inertial finite-size particles.

C. How to control mixing?

The first question to address is why would we enhance mixing and for what? Very often the answer is to increase its quality (i.e., to obtain a complete homogenization) and its rate (i.e., the speed to achieve the uniform state). For many engineering applications, however, the problem is far from being so simple and the objectives of the mixing problem could be completely different depending on the application field (reactor and polymer engineering, heat exchangers, processes with particles or powders, combustion, etc.). Heat and mass transfer, reactions, multiphase flows, and various coupling mechanisms occur in these systems at a wide range of space and time scales. The geometry is also of primary importance: microfluidics or macrofluidics, closed or open flow, static or moving boundaries, etc. If the goal is carefully chosen for the mixing problem under consideration with all its constraints, the choice of the control parameter(s) can be envisaged. But immediately two other questions arise: what is the optimal manner to reach the goal function and which is the appropriate measure to adopt for the optimization of mixing?

This question of generating effective mixing involves solving an optimal control problem. Pioneering work was done by D’Alessandro, Dahleh, and Mezić (1999) in this domain. They used mathematical tools from control and ergodic theory of dynamical systems to rigorously derive the mixing protocol that maximizes entropy among all the possible periodic sequences composed of two shear flows orthogonal to each other (i.e., the eggbeater flow). Other studies followed in the same spirit (Vikhansky, 2002; Vainchtein and Mezić, 2004; Stremler and Cola, 2006; Mathew et al., 2007; Liu, 2008; Gubanov and Cortelezzi, 2010; Lin, Thiffeault, and Doering, 2011; Gubanov and Cortelezzi, 2012; Lunasin et al., 2012). Aamo, Krstić, and Bewley (2003) addressed the problem of enhancing mixing by means of active boundary feedback control in 2D channel flow. Very efficient mixing was obtained by designing feedback control strategies for the stabilization of the parabolic equilibrium flow, then applying this feedback with the sign of the input reversed. Foures, Caulfield, and Schmid (2014) examined different norms for mixing efficiency and showed the impact of the mix-norm theory in higher-Reynolds-number flows.

Despite the fact that many theoretical ingredients are in place to address the problem of optimizing mixing, most of the studies to date concern idealized model flows. Recently Balasuriya (2010) developed a theoretical tool for the determination of an optimal forcing frequency for a time-periodic mixing strategy in order to increase mixing across a fluid interface between two fluids in a microfluidic device. In engineering processes, the aforementioned strong coupling between different physical phenomena renders the theoretical design of an efficient mixing strategy very challenging. As an example, we can cite the interesting problem of optimization of the formation of an emulsion, a dispersion of droplets in a viscous fluid, in a laminar flow (Caubet et al., 2011) by controlling the droplet size distribution, which influences important properties of the emulsion for cosmetic or food applications such as rheology, texture, shelf life (stability), appearance, or taste. For such an application we have contradictory targets: on the one hand, we need to achieve an efficient mixing in order to avoid KAM regions that could trap large droplets, which implies imposing a modulation of the wall velocity of the stirring device. On the other hand, a continuous movement of the rotating elements is needed in order to achieve a continuous stretching of large droplets to promote their transformation into elongated threads and their rupture by the Rayleigh instability. Indeed, a decrease of the wall velocity will cause the relaxation of the stretched threads toward their initial spherical shape. Thus, in this case, a particular stirring strategy with the consideration of two contradictory objectives needs to be found. Another example of a complex mixing problem is the optimization of chaotic advection in the case of pulsatile switching flows in porous media, with applications to soil remediation or enhanced oil recovery (Trefry et al., 2012; Lester, Metcalfe, and Trefry, 2013).
D. Dynamics of the wall

Since almost all flows of interest have solid boundaries, it is obviously important to establish what effects walls have on the dynamics of mixing. This topic has been surprisingly explored very little, and although there are some important results described later, much remains to be done.

Walls have a dramatic effect on the rate of mixing because the no-slip boundary condition ensures that all points on (stationary) walls are parabolic fixed points. This manifold of nonisolated degenerate fixed points makes the resulting advective dynamical system nonhyperbolic. But this nonhyperbolicity is different from that arising from the presence of KAM islands, because the fractal hierarchy of cantori and inner islands is absent near walls.

In closed flows (Sec. III), the presence of walls slows down mixing, as should be expected from the parabolic fixed points on the wall: as reported by Gouillart et al. (2007), the variance of the concentration distribution of a passively advected scalar field decays in time following an algebraic law, as opposed to the exponential decay observed in the absence of boundaries. This is caused by the slow dynamics of the portions of the fluid in the vicinity of the walls. Interestingly, if the walls of a container are rotated as part of the mixing protocol, it was demonstrated experimentally (Gouillart, Thiffeault, and Dauchot, 2010) that the exponential decay of correlations is recovered. This is achieved because rotation of the walls creates a dynamical barrier surrounding a central chaotic region, isolating it from the wall. This shows that moving boundaries can result in very different mixing dynamics, and there is currently no general theory capable of predicting mixing dynamics in the presence of multiple moving boundaries. In the classic system of fluid in a closed container (fixed walls) being mixed by moving rods (moving boundaries), in some circumstances one can prove that the advective dynamics has a strictly positive topological entropy and is therefore chaotic (Thiffeault et al., 2008). Despite that, the variance of an advected scalar still decays exponentially because of the walls. This highlights the fact that there is no general understanding of the connection between dynamical invariance of the advective system and the mixing dynamics of an advected scalar with finite diffusion.

For open flows in a fluid transported from the inlet to the outlet of a pipe mixing device, it is interesting to relate the form of the distribution of the residence times associated with elementary fluid particles to the chaotic behavior of the flow. It is known that a globally chaotic flow tightens the distribution of residence times, but how the existence of parabolic points along the wall boundaries and the fluid behavior at the vicinity of these hyperbolic points influence the shape of the distribution of residence time has not been studied. This will be particularly interesting for reactive problems in chemical engineering.

If a scalar is being input into the fluid by a continuous source, the variance will not decay to zero in time and can be expected to reach a steady state or a periodic state, depending on the dynamics of the source and the flow. This will almost always be the case for open flows, for example, a chemical reactor. It can also be the case in closed flows, for example, if we expect Ohmic heating of the fluid in its volume, the heating or cooling of the fluid is done through the wall boundaries, and thus temperature acts as a continuous scalar source. In these cases, the asymptotic scalar distribution after the transient depends not only on the chaotic dynamics of the flow, but also on details of the sources. This is again poorly understood at the moment, and it is clearly important in practical applications of chaotic mixing.

Finally, chaos can be directly generated by the walls in more complex duct networks that contain branchings and mergings of ducts, e.g., in pore networks. Many of the branch and merge locations will have hyperbolic node points in the skin friction field. If the branch and merge are symmetric, the unstable and stable manifolds emanating from the branching and merging hyperbolic points, respectively, will join smoothly. However, if the branch and merge pair are twisted, then the hyperbolic manifolds must intersect transversely, generating chaotic fluid trajectories. Lester, Metcalfe, and Trefry (2013) developed this argument and used it to prove that chaos is inherent to porous media flows, where topological considerations guarantee an abundance of hyperbolic points. It turns out that chaos, through walls, has considerable influence on the macroscopic transport properties of natural and engineered porous media (Lester, Metcalfe, and Trefry, 2014). The dispersion of a scalar in an isotropic porous medium with a characteristic length over which neighboring fluid-element paths decorrelate is directly proportional to this length and to the interstitial fluid speed (Phillips, 1991). Such a relation is responsible for the observed fast chemical transport, typically orders of magnitude larger than expected from molecular diffusion, in applications ranging from buoyancy-driven flows in soil (Kalejiaye and Cardoso, 2005) to fluid motion over wavy precipitates in hydrothermal vents (Yang et al., 2016). There is definitely more to be understood about chaotic advection in porous media.

E. Strange eigenmodes

Strange eigenmodes, first discussed by Pierrehumbert (1994), characterize the mixing dynamics of advected fields in flows and have been observed experimentally in both closed and open flows. Because of the non-self-adjoint nature of the advection-diffusion operator, completeness of the eigenfunctions is not guaranteed. Rigorous results on the existence and completeness have been obtained by Liu and Haller (2004). Although strange eigenmodes have been observed experimentally in open flows, a corresponding rigorous result for open flows is missing. More generally, very little is known about strange eigenmodes in open flows. For example, questions about how the eigenmodes are related to the fractal structure of the chaotic saddle and its stable and unstable manifolds are still open.

F. 3D unsteady flows

Great progress has been made on Lagrangian transport phenomena in 3D unsteady flows since the first pioneering studies from the mid-1980s (Dombre et al., 1986; Feingold, Kadanoff, and Piro, 1987, 1988b). However, many challenges remain. First, 3D spaces admit far greater topological complexity of coherent structures compared to their counterparts
in 2D (Alexandroff, 1961). Second, the 3D equations of motion lack the well-defined Hamiltonian structure of 2D configurations. 3D steady systems under certain conditions admit representation as 2D nonautonomous Hamiltonian systems (Bajer, 1994; Mezić, 1994; see Sec. VI). The latter, in particular, continues to be a major obstacle to advances in 3D Lagrangian transport phenomena.

Development of a comprehensive mathematical framework thus is imperative for a systematic description and analysis of 3D Lagrangian transport. Theoretical developments primarily expand on the classical Hamiltonian concept of action-angle variables by (local) representation of coherent structures as invariant surfaces and curves defined by constants of motion, denoted “actions” in Hamiltonian terminology (Feingold, Kadanoff, and Piro, 1987, 1988b; MacKay, 1994; Mezić and Wiggins, 1994; Cartwright, Feingold, and Piro, 1996; Mezić, 2001a). A promising recent concept with Hamiltonian foundations devised specifically for mixing applications is found in the linked twist map (Sturman, Ottino, and Wiggins, 2006; Meier, Lueptow, and Ottino, 2007; Sturman et al., 2008).

Establishment of a Hamiltonian-like formalism for 3D Lagrangian transport, in particular response scenarios to perturbations and routes to chaos, is nonetheless in its infancy. The most important generalization of classical Hamiltonian mechanics to 3D systems is the 3D counterpart to the KAM theorem, describing the fate of invariant tori under weak perturbations (Cheng and Sun, 1990a; Mezić and Wiggins, 1994; Broer, Huitema, and Sevryuk, 1996; Vaidya and Mezić, 2012). However, similar universal response scenarios for coherent structures of different topology, most notably the important case of invariant spheres, remain outstanding. Moreover, the scope must be widened to the effect of strong perturbations on the flow topology. Generically, isolated periodic points and periodic lines will emerge. However, the routes toward such states are largely unknown terrain, although bifurcations similar to those studied by Mullowney, Julien, and Meiss (2005, 2008) are likely to play a pivotal role. Further reconciliation with concepts from mathematical physics (Arnol’d, 1978; Mezić and Wiggins, 1994; Arnol’d and Khesin, 1998; Haller and Mezić, 1998; Bennet, 2006) and magnetohydrodynamics (Moffatt et al., 1992; Biskamp, 1993; Moffatt, 2000) is essential in strengthening the current framework.

Theoretical developments largely concentrate on kinematic properties of divergence-free vector fields and volume-preserving maps and thus in essence account only for the role of continuity in the dynamics. However, the impact of momentum conservation as a facilitator or inhibitor of certain kinematic events in realistic fluid flows must be an integral part of scientific studies. A fundamental question is whether universal dynamic conditions, reminiscent of the Beltrami condition for 3D steady Euler flows, exist for 3D chaos and/or (degrees of) integrability in flows with significant viscous effects. A step toward this is presented in Mezić (2002).

Essentially 3D phenomena similar to those exemplified by the 3D cylinder flow in Sec. IV have been observed in a wide range of 3D unsteady systems. This includes generic 3D volume-preserving maps with non-toroidal invariant surfaces (Gómez and Meiss, 2002; Mullowney, Julien, and Meiss, 2005, 2008), the 3D sphere-driven flow of Moharana et al. (2013), the 3D lid-driven cube of Anderson et al. (1999, 2006), and 3D granular flows inside a spherical cavity (Meier, Lueptow, and Ottino, 2007; Sturman et al., 2008). The emergence of similar dynamics in this great diversity of systems reflects the universality of many of the observed phenomena.

Efforts to date overwhelmingly concern theoretical and computational investigations. Experimental studies remain scarce yet are essential to conclusively establish the physical validity, relevance, and robustness of predicted phenomena. Moreover, experiments enable exposition and exploration of phenomena that are beyond present models and may thus contribute to progress in the field at a far deeper level than validation and verification alone. Available techniques include laser-induced fluorescence for visualization of coherent structures (Fountain, Khakhbar, and Ottino, 1998; Fountain et al., 2000; Alvarez et al., 2002; Mezić and Sotiropoulos, 2002; Sotiropoulos, Webster, and Lackey, 2002) and 3D particle-tracking velocimetry (3DPTV) for measurement of 3D Lagrangian fluid trajectories and 3D velocity fields enabling direct (quantitative) measurements on coherent structures (Luethi, Tsinober, and Kinzelbach, 2005). Experimental studies have demonstrated the capabilities of 3DPTV for such analyses (Speetjen, Clercx, and van Heijst, 2004; Otto, Riegler, and Voth, 2008; Dore et al., 2009; Cheng and Diez, 2011; Znaien et al., 2012; Jilissen, Bloemen, and Speetjens, 2013).

A formidable challenge in 3D transport studies is the visualization and isolation of coherent structures in numerical and experimental data. Poincaré sectioning is the method of choice for time-periodic flows yet has the drawback that results depend critically on the initial tracer positions. The ergodic-partition method following Mezić (1994, 2013) and Budišić, Mohr, and Mezić (2012), reviewed in Sec. VII.F, in essence generalizes this concept and, in principle, enables visualization of all relevant topological features with one generic ansatz. Further promising alternatives exist in the Lagrangian techniques by Haller (2001a) and Branicki and Wiggins (2009) developed specifically for 3D unsteady systems.

G. Synthesis

Why is this subject attractive? First, it is rather interdisciplinary. There are typically at least three areas that are involved: (a) fluid mechanics, (b) dynamical-systems theory, and (c) the application area, which may be from other parts of physics or from engineering, biology, chemistry, etc. Such interdisciplinary problems are typically challenging and thus scientifically exciting. Second, analysis, experiment, and numerics are all possible and productive. Third, the fundamentals are quite close to applications. Fourth, mixing via chaotic fluid motion is a case where chaos is “good.” Usually chaos is associated with something “bad”: undesirable vibrations in machinery, loss of particles in accelerators, difficulty in control problems, yet another type of noise, etc. But in fluid mixing, chaos is good. We want as much of it as we can possibly get. In fact, we go to great lengths to eliminate or design away any vestiges of regularity. Fifth, the concepts are applicable over a substantial range of length scales ($10^{-6}$ to $10^{3}$ m), comparable to the decade range of length scales found in turbulent flows in industrial applications and geophysical flows. Sixth, there are still a number of interesting open problems. Finally, note that
the data from experiments in this field are often aesthetically pleasing, as may be seen from the figures in this paper, which adds an extra attraction to chaotic advection.

The community of researchers working on the fundamentals and applications of chaotic advection has made great strides since the 1980s when the field was thus named. Its impact has been felt in many fields of science and engineering and the area of applications has grown enormously. Yet, the theory is still incomplete and presents great challenges. Moreover, in spite of the hundreds of references cited covering a broad range of areas we are aware that important areas of interest to many possible readers are not covered in this review, such as those on non-Newtonian flows, transport in granular media, chemical reactions in chaotic flows, etc.

The key to advancing in this field is developing a thorough understanding of laminar transport mechanisms in realistic fluid systems as used in industry, including non-Newtonian fluids, chemical reactions, aggregation processes; their rigorous experimental validation; further development of transport formalisms on the basis of principles from mathematical physics; their translation and integration into analysis and design strategies; and further development of numerical and experimental methods for transport studies. As we write, new frontiers are being explored. The physics of mixing is the subject to enter. So say researchers at the famed Lorentz Center.

ACKNOWLEDGMENTS

We acknowledge funding from Netherlands (STW 11054; STW 10458; FOM 05PR2474) and Spain (FIS2013-48444-C2-1-P; FIS2013-48444-C2-2-P; FIS2016-77692-C2-1-P; FIS2016-77692-C2-2-P). This paper has grown out of the workshop “Physics of Mixing” organized by Cartwright, Clercx, van Heijst, and Speetjens, attended by those at the Lorentz Center, Leiden from 24 to 28 January 2011. We thank the Lorentz Center for its financial support and the Lorentz Center staff for making the workshop a great success. We acknowledge the three who are no longer with us, Hassan Aref, John Blake, and Slava (Viatcheslav) Meleshko, with whom we planned and worked on the paper.

REFERENCES


Arnol’d, V. I., and A. Avez, 1968, Ergodic problems of classical mechanics (Benjamin, New York).


Bajer, K., and H. K. Moffatt, 1992, “Chaotic associated with fluid inertia,” in Topological Aspects of the Dynamics of...


Krasnopolskaya, T. S., and V. V. Meleshko, 2009,


