IOP FTC

IOP Publishing — Journal of Physics A: Mathematical and Theoretical

J. Phys. A: Math. Theor. **42** (2009) 412004 (18pp) doi:10.1088/1751-8113/42/41/412004

FAST TRACK COMMUNICATION

On the origin and nature of finite-amplitude instabilities in physical systems

R Krechetnikov[1] **and J E Marsden**[2]

[1] University of California at Santa Barbara, Santa Barbara, USA
[2] California Institute of Technology, Pasadena, USA

Received 26 May 2009, in final form 10 September 2009
Published 29 September 2009
Online at stacks.iop.org/JPhysA/42/412004

Abstract

Finite-amplitude instabilities are ubiquitous, but their theory and precise definitions require clarification. In this work, we discuss the interrelation of various notions connected with finite-amplitude instabilities and offer a precise context for these phenomena. Then we establish a connection between non-normality of linear operators, energy conservation by nonlinear operators and the existence of finite-amplitude instabilities in finite- and infinite-dimensional dynamical systems, both in the conservative and dissipative cases. Such a connection may at first appear counter-intuitive since it relates intrinsically linear and nonlinear phenomena, but it follows naturally from the properties of linear and nonlinear operators when they appear together in a dynamical system. In particular, the main theorem of this communication proves that non-normality is a necessary condition for a finite-amplitude instability. It is demonstrated that this phenomenon is relevant to a wide class of physical systems with energy-conserving nonlinearities.

PACS numbers: 47.15.Fe, 47.27.Cn

(Some figures in this article are in colour only in the electronic version)

1. Introduction

1.1. Motivation

This work originated from the question: is it possible to have finite-amplitude instabilities in evolution problems having the standard form

$$\frac{\mathrm{d}\mathbf{u}}{\mathrm{d}t} = A\mathbf{u} + \mathbf{N}(\mathbf{u}), \tag{1}$$

with purely dissipative linear operators A (i.e., when the spectrum of A is located in the left half-plane) and energy-conserving nonlinear terms $\mathbf{N}(\mathbf{u})$? In recent years, there has been a growing interest in physical systems, models for which have energy-conserving nonlinear terms, such as the classical Navier–Stokes equations (NSEs). One can write these equations for the perturbation velocity field $\mathbf{u}$ superimposed on a base flow $\mathbf{U}$ in the form (1) with the linear and nonlinear terms given by

$$A\mathbf{u} = \mathbb{P}[-\mathbf{U}\cdot\nabla\mathbf{u} - \mathbf{u}\cdot\nabla\mathbf{U} + Re^{-1}\Delta\mathbf{u}], \tag{2a}$$

$$\mathbf{N}(\mathbf{u}) = -\mathbb{P}[\nabla\cdot(\mathbf{u}\otimes\mathbf{u})], \tag{2b}$$

where $\mathbb{P}$ is the Leray projection onto the space of divergence-free vector fields and Re is the Reynolds number. The NSEs are used, in particular, in the analysis of the transition to turbulence. However, it is known that, for example, in plane Couette flow, the experimentally observed transition to turbulence is not explained by a linearized spectral stability analysis of the NSEs. The latter is usually done using a Fourier transform in the stream-wise direction (see, for instance, [1]), and it predicts spectral and nonlinear stability for all Reynolds numbers (nonlinear stability being due to [2]). This stimulated the following reasoning [3–6]: assume that the spectrum of the linearized operator (2a) is in the left half-plane [1], and the nonlinear terms of the NSEs are energy conserving[3] [5], i.e.,

$$\int_{\Omega}(\mathbf{u},\mathbf{N}(\mathbf{u}))\,\mathrm{d}x = 0; \tag{3}$$

then it is usually argued that the only way to get energy growth of a perturbation $\mathbf{u}$ under these conditions is through the transient growth mechanism [3], i.e. transient growth of a solution of the linearized problem (1) before its eventual decay to zero as $t\to+\infty$. The latter mechanism is due to generic non-normality of the linearized NSE operator[4] $A\,A^{\dagger}\neq A^{\dagger}\,A$, where $A^{\dagger}$ is the operator adjoint to A, and the existence of sensitive directions in the geometry of this non-normal operator A. This has led to the assertion that the transition is 'essentially linear' [3]. Analogous to [3], Reddy and Henningson [7] argued that subcritical bifurcations of Couette and Poiseuille flows are related to the presence of transient growth. While the linear operator (2a) in the NSEs is generically non-normal due to non-trivial base states and/or boundary conditions, there are situations when it is normal[5]. In such cases, the linear stability analysis is significantly simplified as the operator is diagonalizable[6].

It should be pointed out that the notion of Lyapunov stability, cf definition 1, used both in classical linearized stability studies [1] and in the nonlinear stability analysis of Couette flow [2], does not exclude the possibility of finite-amplitude instabilities, which were in fact found numerically in 3D Couette flow [9]—there are also experimental indications of them [10]. Since finite-amplitude instabilities are an exhibition of nonlinear properties of a system (in fact, they are intrinsically nonlinear phenomena), Manneville [6] advocated that the non-normality of the linear operator is not itself responsible for the transition. One of the purposes of this communication is to contribute to the clarification of the interplay of the linear and nonlinear effects in the occurrence of finite-amplitude instabilities in this and many other physical systems. As we show in section 3, the presence of non-normal operators is a necessary condition for finite-amplitude instabilities.

The energy-conservation property of the nonlinear terms in the NSEs is also important in wave turbulence [11], where the main role of nonlinear terms (advection) is to transfer energy

[3] In fact this is true only for very special boundary conditions, e.g., no-slip on finite domains, periodicity and decay at infinity. While this argumentation is often applied to the flows such as pipe or channel flows [5], these flows do not obey such boundary conditions!

[4] Though, not all non-normal operators lead to transient growth, as will be shown in section 3.1.

[5] A linear operator A on $\mathcal{H}$ is normal if A is closed, densely defined, and $A^{\dagger}A = A\,A^{\dagger}$ [21].

[6] Moreover, in finite-dimensional and certain infinite-dimensional spaces a normal linear operator is unitarily diagonalizable. Also, if its eigenvalues are real, then it is self-adjoint (the converse is true as well), i.e. $A = A^{\dagger}$, as in the classical Rayleigh–Benard convection [8].

from the large eddies at the integral scale where the system is forced to smaller eddies and eventually to the viscous sink; there is also an inverse energy cascade in 2D turbulence [12]. In addition, there are various physical systems from geophysical fluid dynamics to plasma physics, where nonlinearities are energy conserving, e.g., low-order models of large-scale atmospheric variability [13] derived by truncation of the spectral Galerkin projection. It has been realized that using a metric in a Galerkin projection that guarantees energy-conserving nonlinear terms in the resulting dynamical system improves the performance of the low-order model considerably [13]. Analogously, simplified models of the resonant nonlinear interaction of equatorial baroclinic and barotropic Rossby waves [14] have energy-conserving nonlinear terms. These equations were shown to have a Hamiltonian structure and to admit analytic solitary wave solutions. Finally, a two-field model for collisionless trapped electron mode turbulence [15] is also known to have energy-conserving nonlinearity and to exhibit a finite-amplitude instability as opposed to single-field models, which are therefore inadequate for the description of collisionless plasmas as argued in [15].

Subcritical bifurcations are also ubiquitous in various convection problems, e.g., double-diffusive convection [16], non-Boussinesq convection [8], magneto-convection [17], etc. In most of these cases, the nonlinear terms are energy conserving, e.g., classical double-diffusive convection, but the linear operator fails to be normal, which allows for subcritical bifurcations.

1.2. Paper outline

This communication is organized as follows. First, in section 2.1 we introduce and discuss the notion of finite-amplitude instabilities along with a few auxiliary definitions, used in the formulation of the main results in section 2.2 and construction of the theory in the subsequent text. In section 3, we provide a proof of theorem 1 in finite and infinite dimensions. Section 4 gives an illustration, both analytical and graphical, of the theorem and corollaries, as well as how the interplay of non-normality of the linear operator and energy conservation of the nonlinear operator gives rise to a finite-amplitude instability. Also, in section 4.3 we discuss the conditions for nonlinear terms to be energy conserving, and in section 4.2 the implication of theorem 1 for normal form analysis and Hamiltonian systems. Finally, in section 4.4 we formulate a key open question requiring further study.

2. Key notions and the main result

2.1. Notion of finite-amplitude instability

First, we discuss a general notion of finite-amplitude instabilities not restricted to physical systems with energy-conserving nonlinear terms (and hence the examples in this section need not comply with this restriction). Informally, the basic meaning of a finite-amplitude instability of a Lyapunov stable equilibrium point of (1), e.g. the origin $\mathbf{u} = \mathbf{0}$, is the existence of initial conditions, which constitute an open set or, at least, have positive measure[7], outside of a ball of finite radius centered at the equilibrium point, such that the solution either tends to infinity,

[7] This requirement on the initial conditions is crucial since only generic initial conditions are experimentally feasible. The importance of this condition is illustrated with the following example. Consider a particular saddle-node bifurcation in $\mathbb{R}^3$, which should occur at a finite distance from the stable equilibrium at the origin, cf figure 1. Let this bifurcation create an unstable node and a saddle with two-dimensional unstable manifold, and the origin be the unique sink on the latter manifold. Then the invariant set consists of the unstable node and the one-dimensional stable manifold of the saddle. Apparently, there is no finite-amplitude instability of the origin, since there are neither other attractors nor trajectories that escape to infinity.

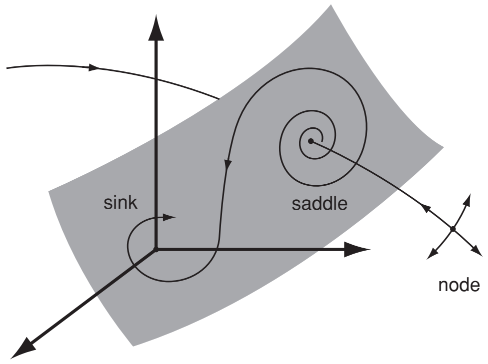

Figure 1. A saddle-node bifurcation in $\mathbb{R}^3$.

$\|\mathbf{u}\| \to \infty$ as $t \to \infty$, or to some invariant set other than the origin. A simple mathematical example of finite-dimensional instability is the first-order ordinary differential equation,

$$\frac{\mathrm{d}x}{\mathrm{d}t} = -\epsilon x + x^2, \qquad 0 \leqslant \epsilon < 1. \tag{4}$$

Obviously, $x = 0$ is a linearly stable equilibrium state. If $\epsilon > 0$, then $x = 0$ is asymptotically stable. But the system has a finite-amplitude instability of critical amplitude $\sigma_c = \epsilon$. In general, in dynamical systems with dimension higher than 1, the critical perturbation usually has a directional dependence, and the hyper surface defined by $\boldsymbol{\sigma}_c$ is a boundary of the basin of attraction of the stable equilibrium at the origin.

First, let us state the classical Lyapunov's notion of stability to arbitrarily small perturbations.

Definition 1. *The origin of (1) is said to be Lyapunov stable, if, for every* $\epsilon > 0$, *there exists a* $\delta(\epsilon) \geqslant 0$ *such that, if* $\|\mathbf{u}(0)\| < \delta$, *then the solution* $\mathbf{u}(t)$ *with the initial condition* $\mathbf{u}(0)$ *exists and* $\|\mathbf{u}(t)\| < \epsilon$ *for all* $t \geqslant 0$. *If, in addition,* $\delta(\epsilon)$ *can be chosen such that* $\lim_{t\to+\infty} \|\mathbf{u}(t)\| = 0$, *then the origin of (1) is called asymptotically stable.*

If the initial condition $\mathbf{u}(0)$ in the above definition can be chosen with arbitrarily large norm $\|\mathbf{u}(0)\|$, then the origin is called *globally Lyapunov stable* (*globally asymptotically stable*, respectively).

A formal notion of *finite-amplitude* instability can also be naturally based upon the notion of a neighborhood of the origin.

Definition 2. *The origin of (1) is said to be finite-amplitude unstable if*

(i) it is Lyapunov stable, and

(ii) there exist a finite ball and a set of initial conditions, which constitute an open set or, at least, have positive measure, outside this ball such that the corresponding ω*-limit set*[8]*, which may include infinity, does not contain the origin.*

Definition 2 can be illustrated with the following two examples in figure 2. The first one is taken from the Takens–Bogdanov bifurcation, which also illustrates the existence of

[8] A point $\mathbf{u}_0$ is an ω-limit point of the forward orbit of $\mathbf{u}(0)$ if there is a sequence $\{t_i\}$, $t_i \to +\infty$, such that $\mathbf{u}(t_i) \to \mathbf{u}_0$; the set of all ω-limit points is called the ω-limit set [18]. Note that the trajectory $\mathbf{u}(t)$ may go to infinity, $\|\mathbf{u}(t)\| \to \infty$, if the state space is unbounded, e.g., $\mathbb{R}^n$ in finite dimensions (or $\mathbb{R}^{2n}$ for canonical Hamiltonian systems).

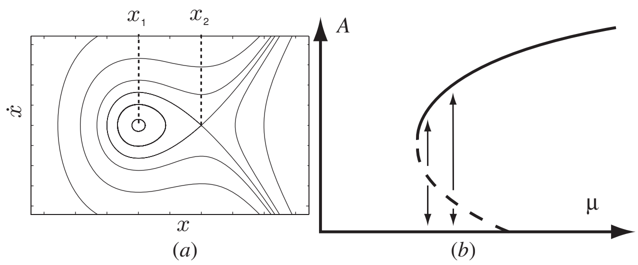

Figure 2. On the notion of finite-amplitude instabilities. (*a*) Takens–Bogdanov bifurcation: $\ddot{x} = -\frac{1}{4} + x^2$. (*b*) Subcritical pitch-fork bifurcation: $\dot{A} = \mu A + A^3$. The dashed curve corresponds to unstable, and solid curves to stable equilibria.

initial conditions at a finite distance from the origin (outside the region defined by the solid curve), such that the solution tends to infinity as $t \to \infty$. This example can be considered as an illustration of a *hard bifurcation* in the terminology of Arnold [19]. The second example—subcritical pitchfork bifurcation—illustrates the fact that while the origin $A = 0$ is Lyapunov stable, there may exist stable equilibria $A(\mu) \neq 0$ at a finite distance from the origin with non-zero basin of attraction. As can be seen from the above two examples, finite-amplitude instabilities occurring due to subcritical bifurcations are a subclass of general finite-amplitude instabilities even in the category of real smooth dynamical systems: for a finite-amplitude instability to happen, (1) need not experience a subcritical bifurcation at the origin. This is important to stress, as the notion of subcritical bifurcations is often imprecisely used to substitute the notion of finite-amplitude instabilities, leaving aside the fact that the former are 'events' and the latter are 'properties'. Namely, (local) bifurcation phenomena occur in dynamical systems, which depend upon a parameter such that a change in the latter causes the stability of an equilibrium (or fixed point) to change, while instabilities are the phenomena occurring at fixed values of the parameter upon which a given system may depend. Finally, it should be noted that global asymptotic stability precludes a finite-amplitude instability, but global Lyapunov stability does not (the proofs of the corresponding Lyapunov stability theorem are usually silent about this fact as, for example, in theorem 4.1 in [20]); the latter is illustrated in figure 3. This distinction will be useful for the subsequent discussion. One of the popular tests for global Lyapunov stability is based on Lyapunov's direct method and thus on the notion of a Lyapunov function [20]. Namely, if there exists a continuously differentiable, radially unbounded (i.e. $V(\mathbf{u}) \to \infty$ as $\|\mathbf{u}\| \to \infty$), positive-definite Lyapunov function $V(\mathbf{u})$ such that $\dot{V}(\mathbf{u})$ is negative semi-definite for all $\mathbf{u} \in \mathbb{R}^n$, then the equilibrium point is globally Lyapunov stable. If, in addition, $\dot{V}(\mathbf{u})$ is negative definite, then the origin is globally asymptotically stable. More general conditions for global asymptotic stability based on LaSalle's invariance principle can be found in [20].

2.2. Main result

While we consider system (1) and its solutions real, defined on a complete inner product (i.e. Hilbert) space $\mathcal{H}$ (this could be a Sobolev space, as dictated by the existence and smoothness of solutions of (1), but we are not concerned with these issues here), we will need an inner product to be defined for complex elements of $\mathcal{H}$ in the standard way [21], with the convention $\langle a\,\mathbf{u}, b\,\mathbf{v}\rangle = a^* b\langle \mathbf{u}, \mathbf{v}\rangle$ for scalars $a, b \in \mathbb{C}$, a^* being the complex conjugate of a.

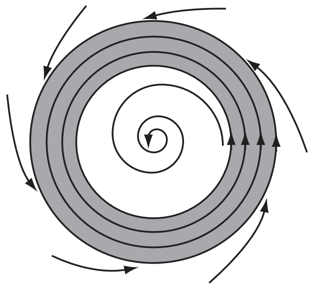

Figure 3. Example of a system with globally Lyapunov stable origin (which is also locally asymptotically stable), but subject to a finite-amplitude instability—the shaded region is filled with periodic orbits.

We define the energy to be the squared norm of the solution of (1):

$$E = \frac{1}{2}\|\mathbf{u}\|^2, \tag{5}$$

where the norm $\|\cdot\|$ is induced by the inner product; for example, in finite dimensions $\mathbf{u} \in \mathbb{R}^n$ and the norm is the standard Euclidean norm. Note that we use the term 'energy' even though E, defined by (5), has such a meaning only in some special situations, such as the NSEs (where it is the kinetic energy). In many other cases, it should be understood as a Lyapunov function, or generalized energy, or Euclidean norm in the state space of (1). As discussed in section 4.3, the form (5) is a natural choice for the NSEs in view of its meaning as a Hamiltonian of the underlying Hamiltonian structure of the conservative part of the NSEs, and also because it allows one to exploit the following property of the nonlinear terms. We define the nonlinear terms to be *energy conserving* if

$$\langle \mathbf{u}, \mathbf{N}(\mathbf{u})\rangle = 0, \tag{6}$$

e.g. in finite dimensions it would be $\mathbf{u}^T\mathbf{N}(\mathbf{u}) = 0$ for all $\mathbf{u} \in \mathbb{R}^n$.

For simplicity, we consider the case when the domain of the definition of the linear operator A is dense in some separable Hilbert space $\mathcal{H}$, which in the context of the NSEs is usually the case if the flow domain is compact (i.e. it is bounded by solid boundaries and/or periodic boundary conditions) or (semi-)infinite with the boundary condition of sufficiently fast decay at infinity, e.g. $L^2(\mathbb{R}^n)$; however, the case of semi- and unbounded domains may lead to non-separable Hilbert spaces, which will be commented in section 3.2.

Then the main result, which applies to the classes of physical systems with energy-conserving nonlinear terms discussed in section 1.1, can be formulated as follows.

Theorem 1. *Assume that for the evolution system (1), posed on a separable Hilbert space $\mathcal{H}$, the linear (possibly unbounded) operator A is normal, and the following conditions hold:*

(a) the initial value problem for (1) is well-posed;
(b) there exists an eigenbasis of A that is an orthonormal basis of $\mathcal{H}$[9]*;*
(c) the spectrum of A is discrete with all eigenvalues located in the left half-plane, as well as there exists $\epsilon < 0$ such that $\mathrm{Re}\,\lambda_i(A) < \epsilon < 0$ *(spectral gap condition)*[10]*;*

[9] This is automatic in finite dimensions, but requires the assumption of a discrete spectrum in infinite dimensions: cf [22] for the case of a symmetric operator A. While every separable Hilbert space $\mathcal{H}$ has an orthonormal basis, the operator A may have a continuous spectrum and thus a basis different from that of $\mathcal{H}$.

[10] Note that in finite dimensions, the spectral gap condition follows from the fact that all eigenvalues are located in the left half-plane, while in infinite dimensions the spectral gap is a separate technical requirement.

(d) the nonlinear terms $\mathbf{N}(\mathbf{u})$ *are energy conserving with respect to the same inner product in which normality of* A *is established.*

Then the origin of (1) is globally asymptotically stable and thus no finite-amplitude instabilities occur in the dynamics of (1).

Of particular relevance to our discussion are the following straightforward corollary and proposition.

Corollary 1. *Under the assumptions of theorem 1 a necessary condition for a finite-amplitude instability to occur in the dynamics of the evolution system (1) with the linear operator* A*, all eigenvalues of which have* $\operatorname{Re}\lambda_i(A) < 0$*, and with energy-conserving nonlinear terms* $\mathbf{N}(\mathbf{u})$ *is non-normality of the linear operator* A*.*

Proposition 1. *If the linearization of the evolution system (1) is Hamiltonian*[11]*, finite-amplitude instabilities may occur even when the linear operator is normal and the nonlinear terms are energy conserving.*

It is worth noting that corollary 1 implies that finite-amplitude instabilities can also occur in evolution systems with energy-non-conserving nonlinear terms and a normal linear operator. This, of course, is a trivial consequence of the fact that finite-amplitude instability phenomena are covariant, i.e. independent of the coordinate system, while non-normality is not.

As we establish here, the behavior stated in the theorem, corollary and proposition is universal for various physical problems. The rest of the communication is devoted to a proof and interpretation of the above results.

3. Proof and Discussion

The arguments in the proof of theorem 1 are somewhat analogous to the energy stability method, but the objectives of the latter is to determine the bifurcation parameter value below which the energy decays, while our goal here is to analyze the bounds on the energy decay when the linear operator is dissipative (see the definition in the Introduction and in [23]).

Proof. The energy evolution equation is readily obtained from (1) by taking an inner product with the solution $\mathbf{u}$

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2 = \langle \mathbf{u}, A\mathbf{u}\rangle, \tag{7}$$

since the nonlinear terms are energy conserving.

Since the operator A, defined on a separable Hilbert space $\mathcal{H}$, is normal, then the spectral theorem [21, 24, 25] implies that it is diagonalizable and, by the assumption of the theorem, its eigensolutions $\{\phi_i\}_{i=1,+\infty}$ can be chosen as an orthonormal basis on $\mathcal{H}$ (a more transparent analysis will be given in finite dimensions in section 3.1); recall that every separable Hilbert space has an orthonormal basis. Note that A can be unbounded, as is relevant to differential operators. Then, using the spectral representation of the solution

$$\mathbf{u}(t,\mathbf{x}) = \sum_i a_i(t)\phi_i(\mathbf{x}),$$

it follows that

$$\|\mathbf{u}\|^2 = \langle \mathbf{u}, \mathbf{u}\rangle = \sum_{i,j}\langle a_i\phi_i, a_j\phi_j\rangle = \sum_{i,j} a_i^* a_j\langle\phi_i, \phi_j\rangle = \sum_i |a_i|^2$$

[11] In this case, of course, for the origin of (1) to be stable, the eigenvalues $\lambda_i(A)$ must lie on the imaginary axis, $\operatorname{Re}\lambda_i(A) = 0$.

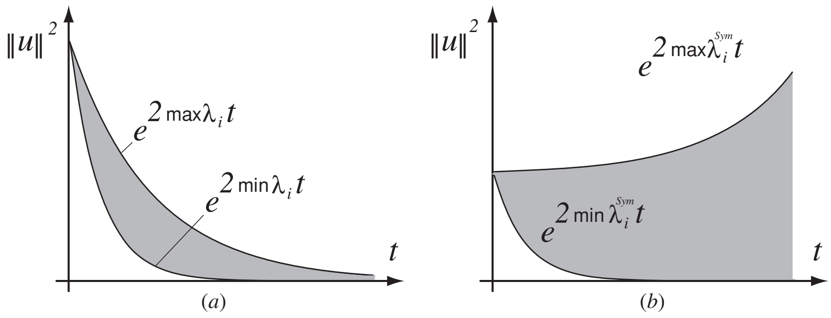

Figure 4. On the bounds on the energy norm $\|\mathbf{u}\|^2$ evolution: the admissible dynamics is limited to shaded regions. (*a*) The case of a normal linear operator A: the dynamics is bounded by its maximal and minimal eigenvalues (9). The case of real eigenvalues is shown. (*b*) The case of a non-normal linear operator A: the dynamics is bounded by minimal and maximal eigenvalues (15) of the symmetric part of A.

and

$$\langle \mathbf{u}, A\mathbf{u}\rangle = \sum_{i,j} \langle a_i\phi_i, a_i A\phi_i\rangle = \sum_{i,j} a_i^* a_j \lambda_j \langle \phi_i, \phi_i\rangle = \sum_i \lambda_i |a_i|^2,$$

where the obvious relation $A\phi_i = \lambda_i\phi_i$ was utilized. As a result, (7) takes the form

$$\sum_i \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}|a_i|^2 = \sum_i \lambda_i |a_i|^2. \tag{8}$$

Because all eigenvalues are located in the left half of the complex plane, i.e. $\mathrm{Re}(\lambda_i) < 0$, then from (8) it follows that the bounds on the energy norm (5) are given by the maximal and minimal real parts of its eigenvalues

$$\min_i \mathrm{Re}(\lambda_i)\|\mathbf{u}\|^2 \leqslant \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2 \leqslant \max_i \mathrm{Re}(\lambda_i)\|\mathbf{u}\|^2, \tag{9}$$

i.e. the evolution of $\|\mathbf{u}\|^2$ is bounded to a region shown in figure 4(*a*). Note that $\mathrm{Re}(\lambda_i) < \epsilon < 0$ in view of the assumed spectral gap condition. Therefore, regardless of the initial amplitude and the geometry of a perturbation, its energy decays to zero, which proves global asymptotic stability and hence prohibits the existence of any finite-amplitude instabilities and thus proves theorem 1. □

Corollary 1 and proposition 1 are a direct consequence of theorem 1. In particular, proposition 1 follows from the fact that if the origin of a Hamiltonian system is Lyapunov stable, then the eigenvalues of the linear operator are located on the imaginary axis and thus the bounds (9) do not apply. A specific example is given below.

3.1. Finite dimensions

To simplify the discussion, we first consider the case when the operators in (1) are defined on a finite-dimensional inner product space, which is also relevant to discretized versions of the NSEs used in various direct numerical simulations. In this case, the linear operator is just a real $n \times n$ matrix: $A \in M_{n\times n}(\mathbb{R})$.

For the purpose of the subsequent discussion, it is instructive to sketch the proof of theorem 1 from the standpoint of direct matrix computation. The energy evolution equation (7) is readily obtained from (1) by multiplying it by $\mathbf{u}^T$ on the left, i.e. by taking an inner product with $\mathbf{u}$.

Since the matrix A is normal, then it is diagonalizable [26] (see also [27, 28]), i.e. there exists a matrix P such that $A = PDP^{-1}$, where D is the diagonal matrix containing the eigenvalues of A. Then, equation (7) can be rewritten as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{u}^T P \cdot P^{-1}\mathbf{u} = \mathbf{u}^T P \cdot D \cdot P^{-1}\mathbf{u}. \tag{10}$$

Since A is normal, the matrix P can be chosen to be unitary (more generally, the matrix P is unitary if and only if A is normal[12]). Because all eigenvalues $\mathrm{Re}(\lambda_i) < 0$, then the bounds on the energy norm (5) are given by the maximal and minimal real parts of its eigenvalues (9), i.e. the evolution of $\|\mathbf{u}\|^2$ is bounded to a region shown in figure 4(*a*). In the derivation of (9) from (10) in finite dimensions, the facts that $(P^{-1}\mathbf{u})^H = \mathbf{u}^T P$, with H denoting the Hermitian transpose, and thus $\|P^{-1}\mathbf{u}\| = \|\mathbf{u}\|$ are utilized. This proves theorem 1 in the finite-dimensional case.

Next, to illustrate proposition 1 in finite dimensions, the Takens–Bogdanov system in figure 2(*a*) can be written using $x = -\frac{1}{2} + \xi$ and $\dot{\xi} = \eta$, as

$$\begin{pmatrix}\dot{\xi}\\ \dot{\eta}\end{pmatrix} = \begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix}\begin{pmatrix}\xi\\ \eta\end{pmatrix} + \begin{pmatrix}0\\ \xi^2\end{pmatrix}, \tag{11}$$

where the linear operator is evidently skew-symmetric and thus normal, but its eigenvalues are on the imaginary axis; hence the bounds (9) are not applicable. The nonlinear terms can easily be modified to make them energy conserving, while still retaining the fact of the presence of finite-amplitude instabilities. The latter can be achieved, if, for example,

(a) the nonlinear term is replaced with the vector $[a(\xi,\eta), b(\xi,\eta)]^T$ such that the energy-conservation condition $\xi\, a + \eta\, b = 0$ holds,
(b) there exists an equilibrium point different from the origin $(\xi_0, \eta_0) \neq \mathbf{0}$, which is the solution of the system $\eta = -a$, $\xi = -b$, and
(c) the linearized operator $\begin{pmatrix} a_\xi & a_\eta + 1\\ b_\xi - 1 & b_\eta\end{pmatrix}$, evaluated at $(\xi_0, \eta_0) \neq \mathbf{0}$, has both eigenvalues $\lambda_{1,2}$ with $\mathrm{Re}\,\lambda_{1,2} \leqslant 0$.

Then there exists a stable equilibrium other than the origin, and hence one gets a finite-amplitude instability. As a particular choice of the functions $a(\xi,\eta)$ and $b(\xi,\eta)$, one can have $a = \eta\, k(\xi,\eta)$ and $b = -\xi\, k(\xi,\eta)$ with the function $k(\xi,\eta)$ satisfying two conditions: $k(\xi,\eta) = -1$ and $k_\eta(\xi,\eta) \geqslant 0$, at the equilibrium point $(1, 0)$. Clearly, there are an infinite number of possible choices for $k(\xi,\eta)$. Note that the system resulting from such a modification of the Takens–Bogdanov system is not Hamiltonian, in general, but its linearization around the origin is, as reflected in proposition 1.

To develop an intuition about corollary 1, re-consider equation (10) in the case when P is non-unitary. Since $(P^{-1}\mathbf{u})^H \neq \mathbf{u}^T P$ for a non-unitary matrix P, then one cannot expect that the energy decays monotonically if the matrix A has eigenvalues with $\mathrm{Re}\,\lambda_i(A) < 0$. In fact, the energy evolution is dictated not only by the eigenvalues of A but also by the eigenvalues of its symmetric part, since

$$\mathbf{u}^T A\mathbf{u} = \mathbf{u}^T A_{\mathrm{Sym}}\mathbf{u}. \tag{12}$$

[12] Indeed, let A be normal, i.e. $AA^T = A^T A$. Since AA^T is symmetric and thus orthogonally diagonalizable, and it commutes with A and A^T, then the eigenvectors of AA^T are also eigenvectors of A. Because AA^T has a complete orthonormal set, then $A = PDP^{-1}$ is also diagonalizable. The necessity follows from the assumption that P is orthonormal, which leads to the conclusion that A must be normal.

For example, if $a_1 \neq a_2$, the matrix A given by

$$A = \begin{pmatrix} -a_1 & 1 \\ 0 & -a_2 \end{pmatrix}, \tag{13}$$

as used later in the example in section 4, is diagonalizable with eigenvalues $\lambda_i = -a_i, i = 1, 2$, and the matrix P is given by $\left(\begin{smallmatrix} 1 & 1 \\ 0 & -a_1 \end{smallmatrix}\right)$. Then, for A from (13) we get $A_{\mathrm{Sym}} = \left(\begin{smallmatrix} -a_1 & 1/2 \\ 1/2 & -a_2 \end{smallmatrix}\right)$ and its eigenvalues

$$\lambda_{1,2} \simeq \pm\tfrac{1}{2} - \tfrac{1}{2}(a_1 + a_2) \pm \tfrac{1}{4}(a_1 + a_2)^2 + O\left(a_1^4, a_1^2 a_2^2, a_2^4\right), \tag{14}$$

i.e. the matrix A_{Sym} is not definite.

Thus, if the matrix A is non-normal, then as follows from the above discussion, it is more insightful to deal with eigenvalues of its symmetric part:

$$\min_i \lambda_i^{\mathrm{Sym}} \|\mathbf{u}\|^2 \leqslant \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2 \leqslant \max_i \lambda_i^{\mathrm{Sym}} \|\mathbf{u}\|^2, \tag{15}$$

which, in turn, dictates the dynamics as in figure 4(*b*), i.e. the evolution of $\|\mathbf{u}\|$ is bounded by the maximal and minimal eigenvalues of the symmetric part of a non-normal matrix A, the substantial difference of which from (9) is due to the transient growth. Obviously, the bounds (15) are too loose for long-time evolution, as it is controlled by the maximal and minimal real parts of the eigenvalues of the non-normal matrix A similar to the normal case (9), where such control is valid for all times. Therefore, the energy argument in this case does not prohibit an existence of finite-amplitude instabilities for evolution systems with non-normal linear operators and energy-conserving nonlinear terms, as suggested by corollary 1.

Next, it should be illustrated that non-normal operators, as for example

$$A = \begin{pmatrix} 0 & 1 \\ -2 & 0 \end{pmatrix}, \tag{16}$$

do not necessarily lead to transient growth as defined in section 1.1: the solution of the corresponding linear initial-value problem, $\mathbf{u}_t = A\mathbf{u}$, is just oscillating in time with frequency $\sqrt{2}$. However, for any non-normal operator with all eigenvalues located in the left half-plane, $\mathrm{Re}(\lambda_i) < 0$, there always exist initial conditions which lead to transient growth. The latter follows from non-orthogonality of eigenvectors and will be illustrated in section 4.

It should be stressed, however, that the occurrence of finite-amplitude instabilities is not guaranteed by non-normality of the linear operator itself, but is also controlled by the existence of ω-limit sets not containing the origin (e.g. at least one stable fixed point different from the origin should exist; also, recall that we defined an ω-limit set as possibly containing infinity, cf section 2.1).

As an alternative to studying the eigenvalues of A_{Sym}, one can analyze singular values of A, which are an important tool used in transient growth analysis [29]. Indeed, for a real square matrix $A \in M_{n\times n}(\mathbb{R})$, its singular value decomposition is $A = U\Sigma V^T$, where U and V are orthogonal matrices and Σ is a diagonal matrix containing non-negative singular values σ_i of A. Then, from the properties of all these matrices it follows that

$$\min_i \sigma_i \|\mathbf{u}\|^2 \leqslant \left|\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\mathbf{u}\|^2\right| = |\mathbf{u}^T A\mathbf{u}| \leqslant \max_i \sigma_i \ \|\mathbf{u}\|^2, \tag{17}$$

which illustrates the geometric meaning of $A = U\Sigma V^T$—the matrix A maps the unit sphere into an ellipsoid with semi-axes equal to singular values of A, since U and V^T are orthogonal. Clearly, the lower bounds obtained from the singular value decomposition analysis (17) are less informative than those obtained from the analysis of the symmetric part of A, as in (15), because of non-negativeness of singular values. Also, the upper bounds obtained from the singular value decomposition analysis may be looser than those obtained from the analysis of the symmetric part of A; for example, in the case of the matrix (13) the singular values are given by

$$\sigma_1 \simeq 1 + O\big(a_1^2, a_2^2\big), \qquad \sigma_2 \simeq 0 + O\big(a_1^2, a_2^2\big), \tag{18}$$

which certainly give an 'overshoot' upper bound compared to (14). It should be noted, though, that singular and eigenvalue decompositions coincide if the matrix is symmetric positive semi-definite. Indeed, recalling that singular values are positive square roots of the eigenvalues of the matrix $A\,A^T$, and that $A\,A^T = A_{\mathrm{Sym}}^2 - A_{\mathrm{Skew}}^2$ with A_{Sym} being symmetric and A_{Skew} skew-symmetric parts of A, one can observe that for symmetric and almost symmetric matrices, the main contribution to singular values comes from A_{Sym}[13], while for highly non-normal matrices both symmetric and skew-symmetric components contribute more equally to the singular values of A.[14] However, the use of singular value decomposition is indispensable in identifying the directions of optimal disturbances; namely, if v_1 and u_1 are column vectors of V and U corresponding to σ_1, one obtains

$$A v_1 = \sigma_1 u_1, \tag{19}$$

which describes the amplification of the initial condition v_1 by a factor σ_1, $\|A v_1\| = \sigma_1$.

3.2. *Infinite dimensions*

With regard to the bounds (9) and figure 4(*b*), it should be noted that in the infinite-dimensional case the real part of the minimal eigenvalue may tend to $-\infty$, e.g. in the case of the Dirichlet problem for the Laplace operator $-\Delta$ on a bounded domain.

Next, the fact that the linear operator A can be unbounded, i.e. the condition that there exists $M > 0$ such that for all $\mathbf{u} \in \mathcal{H}$ the inequality $\|A\mathbf{u}\| \leqslant M\|\mathbf{u}\|$ does not hold, does not contradict the fact of the existence of an orthonormal eigenbasis, which spans the space $\mathcal{H}$. Indeed, the fact that the eigensolutions $\{\phi_i\}_{i=1,+\infty}$ of $A\phi_i = \lambda_i\phi_i$ span $\mathcal{H}$ implies that $\|A\mathbf{u}\|^2 = \sum_i |\lambda_i|^2\,|a_i|^2$, which can be infinite, while the solution norm $\|\mathbf{u}\|^2 = \sum_i |a_i|^2$ should be finite. Thus, the boundedness of the linear operator is not implied by the existence of an orthonormal eigenbasis.

The proof of theorem 1 can be extended to the case of infinite-dimensional operators in (1) on non-separable Hilbert spaces, which often occur when the evolution problem (1) is defined on (semi-) infinite spatial domains and thus leads to continuous spectra; while the spectral theorem for normal bounded operators on non-separable Hilbert spaces is known [25], the case of unbounded operators may need some further elaboration [21]. The case of continuous spectra may also occur on separable Hilbert spaces and thus requires special treatment, as, for example, the Laplace operator on $L^2(\mathbb{R})$, which is a separable Hilbert space, has a continuous spectrum [30], namely $\lambda = -k^2$, $k \in \mathbb{R}$ with generalized eigenfunctions $\mathrm{e}^{\mathrm{i}kx}$, $x \in \mathbb{R}$. In summary, it should be stressed that there are a number of technical conditions, including normality of the linear operator which goes beyond the formal equality $A\,A^\dagger = A^\dagger A$, to be

[13] As an illustration, consider the matrix (13), where the symmetric part is given by $A_{\mathrm{Sym}} = \begin{pmatrix} -a_1 & 1/2 \\ 1/2 & -a_2 \end{pmatrix}$ and skew-symmetric one is $A_{\mathrm{Skew}} = \begin{pmatrix} 0 & 1/2 \\ -1/2 & 0 \end{pmatrix}$. Clearly, the symmetric part dominates for $|a_i| \gg 1$.

[14] If A is almost skew-symmetric, then it is almost normal.

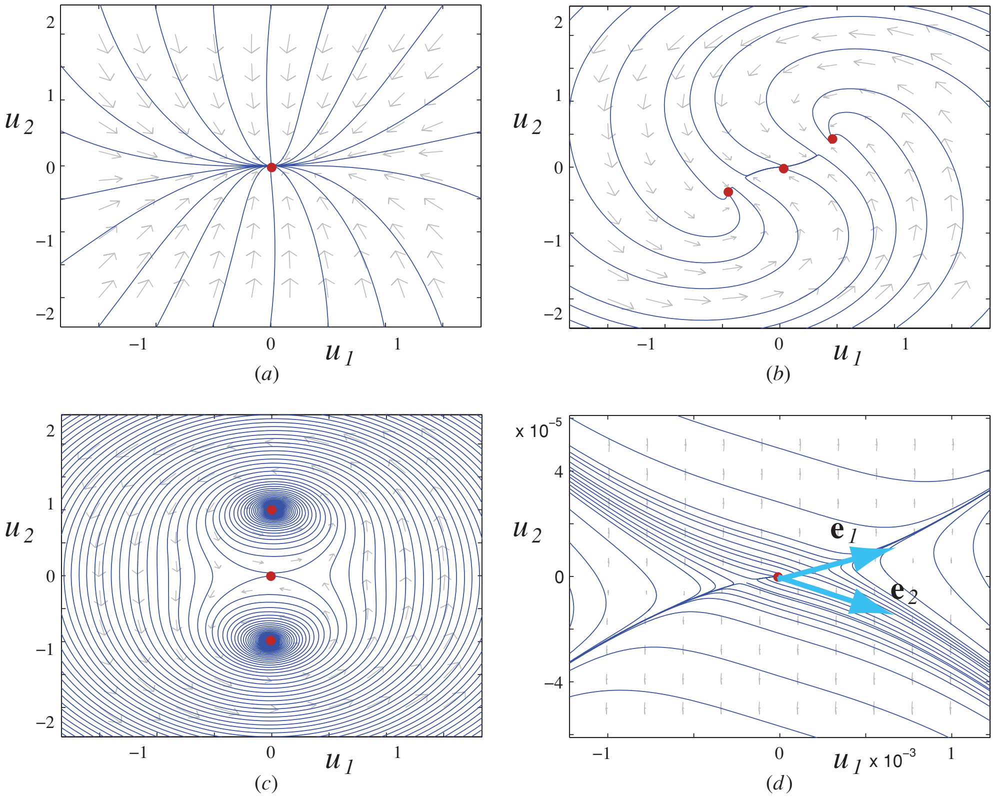

Figure 5. Finite-amplitude instability phenomena from the phase-space geometry viewpoint: phase portraits of the example (20). Red dots are stable equilibria: the origin is a stable node and the rest are stable foci (at $Re = 2\sqrt{2}$ two saddles and two stable nodes appear; further growth of Re turns these nodes into stable foci shown in (*b*) and (*c*)). (*a*) Almost normal operator: $Re = 0.1$. No finite-amplitude instability of the origin (stable node). (*b*) Non-normal operator: $Re = 2.85$. Finite-amplitude instability of the origin due to the presence of two stable foci. (*c*) Highly non-normal operator: $Re = 10^2$. Finite-amplitude instability of the origin due to the presence of two stable foci. (*d*) Highly non-normal operator: $Re = 10^2$. Zoomed out (*c*) in the neighborhood of the origin.

checked in order to exclude finite-amplitude instabilities from the dynamics of (1) based on theorem 1.

4. Interpretations and discussion

4.1. Geometric interpretation of theorem 1

While the proof above highlights some of the salient features of the interrelation of the occurrence of finite-amplitude instabilities and non-normality of linear operators in systems with energy-conserving nonlinearities, one can still gain more insight from particular examples. First, we would like to analyze the example, originally given by Trefethen *et al* [3], from the phase-space geometry viewpoint. Its phase portraits are a good illustration of theorem 1: in the normal case, or almost normal case as in figure 5(*a*), there is no finite-amplitude instability, while starting with some degree of non-normality, controlled by the parameter Re, finite-amplitude instabilities appear as in figures 5(*b*) and (*c*). Namely, the example is of the form (1) with the linear and nonlinear operators defined as

$$A = \begin{pmatrix} -a_1 & 1 \\ 0 & -a_2 \end{pmatrix}, \qquad \mathbf{N}(\mathbf{u}) = \|\mathbf{u}\| B\mathbf{u}, \tag{20}$$

respectively, where $a_1 = Re^{-1}$, $a_2 = 2\,Re^{-1}$, $B = \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right)$ and $\mathbf{u} = (u_1, u_2)^T$. This model problem exhibits non-normality of the linear operator for $Re \gg 1$ and a finite-amplitude instability [3], as illustrated in figure 5. It should be pointed out that the nonlinear terms in (20) are energy conserving, $\mathbf{u}^T\mathbf{N}(\mathbf{u}) = 0$, in view of the skew-symmetry of the matrix B. Since the eigenvectors of A are almost parallel for $Re \gg 1$:

$$\mathbf{e}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad \mathbf{e}_2 = \begin{bmatrix} 1 \\ -Re^{-1} \end{bmatrix}, \tag{21}$$

the linearized system $\dot{\mathbf{u}} = A\mathbf{u}$ experiences a transient growth on the time scale $t = O(Re)$, e.g. for the initial conditions $\mathbf{u}(0) = [0, 1]^T$:

$$\mathbf{u} = \begin{bmatrix} t + O\left(\frac{t^2}{Re}\right) \\ 1 - \frac{2t}{Re} + O\left(\frac{t^2}{Re^2}\right) \end{bmatrix}.$$

This is also reflected on the phase portrait of the system: with an increase in non-normality, the trajectories are deformed from the picture in the almost modal coordinates in figure 5(*a*) to the highly deformed one in figure 5(*c*), the geometry of which near the origin is defined by the eigenvectors, as shown in figure 5(*d*). It is also notable that the domain of attraction near the origin shrinks with Re increasing.

4.2. Relation to normal form analysis

Examples such as (20) considered above are often derived by (invariant manifold) reduction from a more complete model, e.g. NSEs, and thus for the sake of local bifurcation analysis can be simplified further through normal form transformations [31]. A natural question to ask is whether the property of energy conservation by nonlinear terms is invariant under such transformations or not? The answer to this question certainly affects the ability of the transformed model to capture finite-amplitude instability phenomena present in the original model.

As can easily be inferred from the simple case (20) with $\mathbf{N}(\mathbf{u})$ smoothed out (so that application of the normal form analysis is possible), the normal form is

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = J\mathbf{x} + \begin{pmatrix} 0 \\ ax_1^2 \end{pmatrix}, \qquad \text{where } \mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \tag{22}$$

where the nonlinear term becomes non-energy conserving and thus energy conservation by nonlinear terms is not invariant under the normal form transformations. Conversely, one can have the situation when the linear operator is non-resonant and the original nonlinear operator is not energy conserving and thus after the normal form transformation one arrives at the linear system with no nonlinear terms, which is of course a trivial case of energy-conserving nonlinear terms. This non-invariance property is due to the fact that the normal form analysis is local, while the energy conservation by nonlinear terms dictates the global phase (or state) space structure, in particular finite-amplitude instabilities.

There are, however, normal forms which exhibit finite-amplitude instabilities, e.g. the normal form

$$\dot{A} = i(\lambda A - |A^2|A - \epsilon A^*) - \delta A,\ \lambda, \epsilon, \delta \in \mathbb{R},\ A \in \mathbb{C}, \tag{23}$$

studied in [32, 33]. The latter works illustrate instabilities in Hamiltonian systems occurring due to symmetry-breaking deformations, such as in the problem of a vortex filament in a cylindrical vessel the cross-section of which is perturbed from a circular to an elliptical shape of small eccentricity. The normal form (23) has both the non-normal linear operator and non-energy-conserving nonlinear terms, which allows for the existence of finite-amplitude instabilities according to theorem 1. It would be interesting, though, to identify a class of normal forms with the property of energy conservation by nonlinear terms as such normal forms would have a better reflection of some of the global features of the NSEs.

4.3. The conditions for and meaning of energy conservation by nonlinear terms

Since in this work we consider physical systems with energy-conserving nonlinear terms, it is worth commenting on the mathematical and physical conditions under which nonlinear terms possess the property (6). Apparently, the condition (6) holds if $\mathbf{N}(\mathbf{u})$ is orthogonal to $\mathbf{u}$ for all $\mathbf{u}$, or, as suggested by example (20), if $\mathbf{N}$ has the form $\mathbf{N}(\mathbf{u}) = N'(\mathbf{u})\mathbf{u}$, with the matrix $N'(\mathbf{u})$ being skew symmetric for all $\mathbf{u}$. The same applies to the infinite-dimensional case, but with an appropriate definition of the inner product, e.g. the nonlinearity in the NSEs for incompressible fluid (2*b*) gives,

$$\langle \mathbf{u}, \mathbf{N}(\mathbf{u})\rangle = \int_\Omega (\mathbf{u}, \nabla \cdot (\mathbf{u} \otimes \mathbf{u}))\, \mathrm{d}x = \int_\Omega \nabla \cdot (\mathbf{u}|\mathbf{u}|^2)\, \mathrm{d}x = 0, \tag{24}$$

which is due to the divergence-free velocity field and thus due to the ability to put the product of the solution vector with nonlinear one in the divergence form. In other words, the nonlinear operator in the NSEs is skew in the case of no-slip, periodic or decay at infinity boundary conditions and because $\nabla \to -\nabla$ when $\mathbf{x} \to -\mathbf{x}$. From a physical point of view, this conservation property is clearly related to the second (spatial) component in the material derivative, $D/Dt = \partial_t + (\mathbf{u} \cdot \nabla)$. Namely, if the transport of some scalar C is considered, the convective part $(\mathbf{u} \cdot \nabla)C$ represents the time rate of change of C due to the fact that the fluid particle, which at time t is at the position $\mathbf{x}$, moves to a new position in space. Thus, locally C can change, but globally it is conserved if there is no net flux of C through the boundaries, i.e. all local changes cancel each other when summed up over the entire domain.

Alternatively, the property (6) can be understood based on the relation between dissipative and conservative systems. As we saw in the relation between the Euler and NSEs, for example, conservative systems form a basis for dissipative formulations[15]. Therefore, below we discuss the condition (6) just in the context of finite-dimensional conservative formulations (similar arguments apply to continuous systems); linear dissipative terms can be added to make the linear operator as in theorem 1. It is notable that the nonlinearity in the NSEs is independent of the dissipative effects and thus is exactly the same in the Euler equations. The latter have an underlying (non-canonical) Hamiltonian structure with the Hamiltonian $H = \|\mathbf{u}\|^2/2$ [37]; therefore, since Euler's equations are conservative, then H is conserved, which means that the nonlinear terms must conserve H too. Thus, while most of the systems we study in mechanics and physics are dissipative as the NSEs are, this leads to the question: under which conditions do Hamiltonian systems, often underlying dissipative formulations, belong to the class of nonlinearly energy-conserving systems?

[15] If, however, one starts from a dissipative system and would want to decompose it into Hamiltonian and dissipative parts, such decomposition is non-unique in general [34–36].

Let us first look at canonical finite-dimensional Hamiltonian systems. For example, in two dimensions:

$$\frac{\mathrm{d}p}{\mathrm{d}t} = -\frac{\partial H}{\partial q}, \tag{25}$$

$$\frac{\mathrm{d}q}{\mathrm{d}t} = \frac{\partial H}{\partial p}, \tag{26}$$

the nonlinear terms would enter through the cubic or higher order contribution H_N in the Hamiltonian, $H(p,q) = H_L(p,q) + H_N(p,q)$, and would be (phase space) energy conserving if H_N satisfies the first-order linear partial differential equation

$$-p\frac{\partial H_N}{\partial q} + q\frac{\partial H_N}{\partial p} = 0. \tag{27}$$

The latter is feasible only if $H_N(p,q) = H_N(q^2+p^2)$. These results can easily be generalized onto $2n$-dimensional Hamiltonian systems written in a symplectic form

$$\frac{\mathrm{d}z^i}{\mathrm{d}t} = J^{ij}\frac{\partial H}{\partial z^j}, \qquad z^i = (q^i, p_i), \tag{28}$$

i.e. the condition (27) becomes

$$z_i J^{ij}\frac{\partial H_N}{\partial z^j} = 0. \tag{29}$$

Addition of a linear dissipative operator to (28) does not change the condition (29).

Since we are dealing with the energy form (5), which in finite dimensions has the meaning of the Euclidean norm in the phase space, there must be a natural connection to the Liouville theorem on the conservation of the phase-space volume. Since a Hamiltonian system linearized around the origin is still Hamiltonian then the dynamics generated by the nonlinear contribution H_N is still incompressible, i.e. leaves the phase-space volume intact.

The latter observation suggests that there should be a global conservation law associated with the property of energy conservation by nonlinear terms. This fact leads to an interesting insight into the nature of energy-conserving nonlinear terms, as can be seen from the point of view of Lagrangian mechanics. Indeed, let us start with an action integral

$$S = \int L(\mathbf{q},\dot{\mathbf{q}})\,\mathrm{d}t, \qquad \mathbf{q} \in \mathbb{R}^n, \tag{30}$$

with some Lagrangian density $L(\mathbf{q},\dot{\mathbf{q}})$. Application of Hamilton's principle leads to the Euler–Lagrange equations

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\partial L}{\partial \dot{q}_i}\right) - \frac{\partial L}{\partial q_i} = 0. \tag{31}$$

Let us consider simple mechanical systems, i.e. $L(\mathbf{q},\dot{\mathbf{q}}) = T(\dot{\mathbf{q}}) - V(\mathbf{q})$, where $T = \dot{\mathbf{q}}^2/2$. If the potential consists of linear (quadratic) and nonlinear parts, $V(\mathbf{q}) = V_L(\mathbf{q}) + V_N(\mathbf{q})$, then the phase-space energy is conserved by the nonlinear terms iff

$$\dot{q}_i\frac{\partial V_N}{\partial q_i} = 0 \quad\Rightarrow\quad \frac{\mathrm{d}}{\mathrm{d}t}V_N(q_i) = 0 \quad\Rightarrow\quad V_N(q_i) = \mathrm{const}, \tag{32}$$

which implies the existence of a conserved quantity; in this case $\dot{q}_i(\partial V_N/\partial q_i) = 0$ appears to be analogous to the angular momentum conservation and also has a meaning that no work is done by the forces $F_i = -\partial V_N/\partial q_i$, that is $\delta W = \sum F_i\delta x_i = \sum F_i\dot{q}_i\delta t = 0$. Then, if a converse to Noether's theorem applies [38], this fact implies that conservation of energy by nonlinear terms means existence of symmetries in a given physical system.

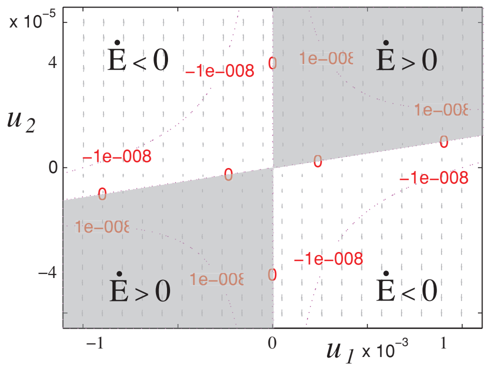

Figure 6. Distribution of the energy gradients, $\dot{E} \equiv \mathrm{d}E/\mathrm{d}t = -u_1^2/Re - 2u_2^2/Re + u_1 u_2$, for figure 5(*d*).

4.4. Open problem: sufficient conditions for finite-amplitude instabilities

In conclusion, it is worth commenting on the general problem of finding finite-amplitude instabilities in a given dynamical system. Since the energy appears to be the cornerstone of the overall discussion, it is natural to appeal to the Lyapunov direct method [39] (based on a consideration of a Lyapunov function). While this method is usually used for studying the stability of the origin of (1), its extension should be suitable for finding finite-amplitude instabilities thanks to its global character. In particular, the level sets of the time gradients of the energy in figure 6 and alternating-sign $\dot{E}$ suggest that some trajectories may leave the neighborhood of the origin if the perturbation amplitude is high enough. However, the fact of stability of the origin in some neighborhood is consistent with LaSalle's invariance principle (theorem 4.4 in [20]), which in turn implies that the sufficient conditions for Lyapunov (not a finite-amplitude) instability in the key instability theorem of Chetayev (theorem 4.3 in [20]) are not met. An intuitive sufficient condition for finding finite-amplitude instabilities can be based on the definition of a finite-amplitude instability formulated in section 2.1. Namely, given a finite-dimensional dynamical system $\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x})$ with the spectrum of the vector field linearized around the origin $\mathbf{x} = \mathbf{0}$ located in the left half-plane, $\mathrm{Re}(\mathbf{D_x f}(\mathbf{0})) < 0$ (so that the origin is Lyapunov stable), a sufficient condition for a finite-amplitude instability to occur is the existence of an (not necessarily compact) invariant set M disjoint from the origin in the state space, $\mathrm{dist}(M, \mathbf{0}) > 0$. Practically, set M is identified by finding a Lyapunov function V such that $\dot{V} \leqslant 0$ in M (note that the space energy $E = \mathbf{x}^2/2$ may grow on some parts of a trajectory in M due to transient growth). A proof of the above sufficient condition is straightforward, and its illustrations are given for two-dimensional state space in figure 7. This sufficient condition is a generalization of the concept of a *trapping region* [18], which is a compact set N such that the flow of (1) gives $\phi_t(N) \subset \mathrm{int}(N)$ with $\mathrm{int}(N)$ being the interior of N. Basically, the vector field points inward everywhere on the boundary of M (or N if it is a compact). It should be noted that the fact of the presence of a transient growth does not invalidate this condition; the size of the subsets defined by $\dot{V} \leqslant 0$ may shrink, through.

Extension of the above sufficient condition for finding finite-amplitude instabilities to PDEs is not straightforward in view of infinite dimensionality of the phase space and issues with existence of a solution, etc, which invites future endeavors in view of the practical importance of identifying finite-amplitude instabilities in the behavior of physical systems.

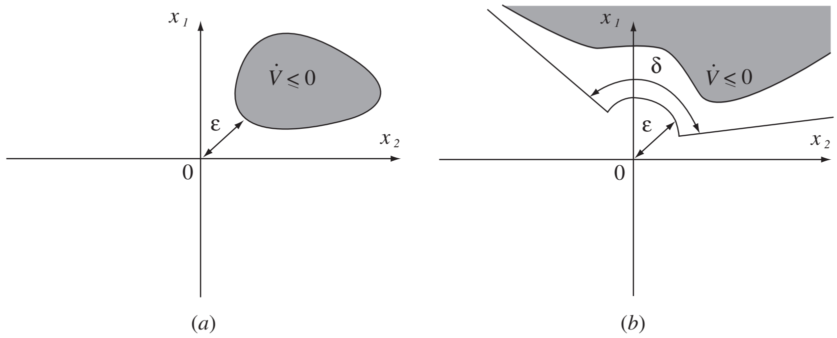

Figure 7. On a sufficient condition for finite-amplitude instability; set M is defined by the points where $\dot{V} \leqslant 0$ with V being a Lyapunov function. (*a*) Compact set M. (*b*) Non-compact set M contained in a sector of angle δ.

Lastly, given the observation that the linear operator in the NSEs is generically non-normal (since the base states, stability of which is studied, are usually non-trivial) one may conjecture that finite-amplitude instabilities are generic phenomena in the context of the NSEs (or systems of similar complexity). Clarification of this conjecture may help to advance our understanding of the transition to turbulence problem. In particular, this may clarify if turbulence in shear flows is a transient [40] or an instability phenomenon.

Acknowledgments

RK is grateful to Edgar Knobloch for a fruitful discussion on the subject. The authors also acknowledge the constructive input of anonymous referees.

References

[1] Drazin P G and Reid W H 1984 *Hydrodynamic Stability* (Cambridge: Cambridge University Press)
[2] Romanov V A 1973 Stability of plane-parallel Couette flow *Funct. Anal. Appl.* **7** 137–46
[3] Trefethen L N, Trefethen A E, Reddy S C and Driscoll T A 1993 Hydrodynamic stability without eigenvalues *Science* **261** 578–84
[4] Waleffe F 1995 Transition in shear flows. Nonlinear normality versus non-normal linearity *Phys. Fluids* **7** 3060–6
[5] Henningson D 1996 Comment on transition in shear flows. Nonlinear normality versus non-normal linearity *Phys. Fluids* **8** 2257–8
[6] Manneville P 2005 Modeling the direct transition to turbulence *Fluid Mech. Appl.* **77** 1–34
[7] Reddy S C and Henningson D S 1993 Energy growth in viscous channel flows *J. Fluid Mech.* **252** 209–38
[8] Straughan B 2004 *The Energy Method, Stability, and Nonlinear Convection* (Berlin: Springer)
[9] Orszag S A and Kellst L C 1980 Transition to turbulence in plane poiseuille and plane couette flow *J. Fluid Mech.* **96** 159–205
[10] Mullin T and Peixinho J 2006 Transition to turbulence in pipe flow *J. Low Temp. Phys.* **145** 75–88
[11] Newell A C, Nazarenko S and Biven L 2001 Wave turbulence and intermittency *Physica* D **152–153** 520–50
[12] Kraichnan R H 1967 Inertial ranges in two-dimensional turbulence *Phys. Fluids* **10** 1417–23
[13] Kwasniok F 2004 Empirical low-order models of barotropic flow *J. Atmos. Sci.* **61** 235–45
[14] Biello J A and Majda A J 2004 The effect of meridional and vertical shear on the interaction of equatorial baroclinic Rossby waves *Stud. Appl. Math.* **112** 341–90
[15] Baver D A, Terry P W, Gatto R and Fernandez E 2002 Nonlinear stability and instability in collisionless trapped electron mode turbulence *Phys. Plasmas* **9** 3318–32
[16] Proctor M R E 1981 Steady subcritical thermohaline convection *J. Fluid Mech.* **105** 507–21

[17] Sunil, Sharma P and Mahajan A 2008 A nonlinear stability analysis for thermoconvective magnetized ferrofluid with magnetic field dependent viscosity *Int. Commun. Heat Mass Transfer* **35** 1281–7
[18] Wiggins S 2003 *Introduction to Applied Nonlinear Dynamical Systems and Chaos* (Berlin: Springer)
[19] Arnold V I 1999 *Bifurcation Theory and Catastrophe Theory* (New York: Springer)
[20] Khalil H K 2002 *Nonlinear Systems* (Englewood Cliffs, NJ: Prentice-Hall)
[21] Conway J B 1985 *A Course in Functional Analysis* (Berlin: Springer)
[22] Robinson J C 2001 *Infinite-dimensional Dynamical Systems* (Cambridge: Cambridge University Press)
[23] Pazy A 1983 *Semigroups of Linear Operators and Applications to Partial Differential Equations* (New York: Springer)
[24] Bernau S J 1966 The spectral theorem for unbounded normal operators *Pacific J. Math.* **19** 391–406
[25] Arveson W 2002 *A Short Course on Spectral Theory* (Berlin: Springer)
[26] Mitchell B E 1953 Normal and diagonalizable matrices *Am. Math. Mon.* **60** 94–6
[27] Horn R A and Johnson C R 1985 *Matrix Analysis* (Cambridge: Cambridge University Press)
[28] Lancaster P and Tismenetsky M 1985 *The Theory of Matrices* (New York: Academic)
[29] Bamieh B and Dahleh M 2001 Energy amplification in channel flows with stochastic excitation *Phys. Fluids* **13** 3258–69
[30] Gelfand I M and Shilov G E 1967 *Generalized Functions, vol. 3: Theory of Differential Equations* (New York: Academic)
[31] Guckenheimer J and Holmes P 1983 *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields* (Berlin: Springer)
[32] Guckenheimer J and Mahalov A 1992 Instability induced by symmetry reduction *Phys. Rev. Lett.* **68** 2257–60
[33] Knobloch E, Mahalov A and Marsden J E 1994 Normal forms for three-dimensional parametric instabilities in ideal hydrodynamics *Physica* D **73** 49–81
[34] Olver P J and Shakiban C 1988 Dissipative decomposition of ordinary differential equations *Proc. Edinb. Math. Soc.* **109** 297–317
[35] Lewis D and Marsden J 1988 A Hamiltonian-dissipative decomposition of normal forms of vector fields *Proc. Conf. on Bifurcation Theory and its Numerical Analysis* (Xi'an: Xi'an Jaitong University Press) pp 51–78
[36] Palacián J F 2005 Dissipative-Hamiltonian decomposition of smooth vector fields based on symmetries *Chaos* **15** 033111
[37] Arnold V I 1997 *Mathematical Methods of Classical Mechanics* (New York: Springer)
[38] Mostepanenko A M and Mostepanenko V M 1975 Converse of the Noether theorem and symmetry in physics *Heuristic Role of Mathematics in Physics and Cosmology* (Leningrad: Nauka) pp 78–95
[39] Merkin D R 1997 *Introduction to the Theory of Stability* (Berlin: Springer)
[40] Hof B, Westerweel J, Schneider T M and Eckhardt B 2006 Finite lifetime of turbulence in shear flows *Nature* **443** 59–62