J. Fluid Mech. (1990), *vol.* 211, *pp.* 393–416
Printed in Great Britain

Stability of stationary barotropic modons by Lyapunov's direct method

By H. SAKUMA AND M. GHIL
Climate Dynamics Center, Department of Atmospheric Sciences, and Institute of Geophysics and Planetary Physics, University of California, Los Angeles, CA 90024-1565, USA

(Received 22 September 1988 and in revised form 1 August 1989)

A new Lyapunov stability condition is formulated for the shallow-water equations, using a gauge-variable formalism. This sufficient condition is derived for the class of perturbations that conserve the total mass. It is weaker than existing stability criteria, i.e. it applies to a wider class of flows. Formal stability to infinitesimally small perturbations of arbitrary shape is obtained for two classes of large-scale geophysical flows: pseudo-eastward flow with constant shear, and localized coherent structures of modon type.

1. Introduction and motivation

Since the pioneering study of the Fermi–Pasta–Ulam (FPU) problem by Zabusky & Kruskal (1965), interest in the dynamics of solitary waves has grown in many branches of the physical sciences. Fluid dynamicists have investigated solitary waves in plasmas and ideal flows with various horizontal scales, extending from laboratory to planetary dimensions. In the field of geophysical fluid dynamics (GFD), work on solitary waves of Rossby type has been particularly intensive in the last decade. Consequently, a new approach toward the understanding of localized, coherent, synoptic-scale motions is emerging: long-lived vortical structures which often obtain in the geophysical fluid environment may be explained in terms of the dynamics of solitary waves.

Quasi-geostrophic solitary-wave solutions in GFD are usually classified into two types (Flierl *et al.* 1980): the first is known as Rossby solitons (Redekopp 1977) and the second is represented by modons (Stern 1975). The former type is a class of approximate solutions to the full equations of motion in the following sense: the derived Korteweg–de Vries (KdV) equations, for which they are exact solutions, result from truncated small-parameter expansion of the full equations. The latter type is a class of exact solutions to the quasi-geostrophic or barotropic vorticity equations, though they fail to be continuous either in their vorticity field or in their derivatives of some order.

These two types of solutions have been used as models of such long-lived vortices as the Great Red Spot of Jupiter (Maxworthy & Redekopp 1976; Ingersoll & Cuong 1981), atmospheric blocks (McWilliams 1980) and Gulf Stream rings (Flierl *et al.* 1980; Malanotte-Rizzoli 1982). Direct observational evidence to confirm the validity of these models is suggestive, but not conclusive. Stability of solitary-wave solutions is therefore an important indicator of their realism and plausibility as the explanation of isolated vortex structures.

Arnol'd's (1965) stability criterion follows from a fairly general application of Lyapunov's direct method to fluid mechanics. Blumen (1968) adapted Arnol'd's criterion to rotation-dominated geophysical flows. The stability criteria of Arnol'd and Blumen were derived by checking the definiteness of the second variation of a certain integral invariant. For positive definiteness of such a quantity, Blumen's criterion assumes the form

$$\nabla\psi/\nabla(f+\Delta\psi) > 0, \tag{1.1}$$

where ψ and f are the stream function and the Coriolis parameter, respectively. However, quasi-geostrophic flows in mid-latitudes with large spatial scales and low frequencies are pseudo-eastward in the sense that the absolute vorticity $f+\Delta\psi$ increases to the left of geostrophic velocity (Andrews 1984). Pseudo-eastward shear flows violate Blumen's criterion (1.1). So do modon solutions, for which $f+\Delta\psi = -\lambda\psi$, with λ a positive constant, in the interior of a circular disk.

Benzi *et al.* (1982) derived a modified criterion that corresponds to a negative-definite norm for an equivalent-barotropic model on an infinite β-plane. Andrews (1984) pointed out, however, that steady, inviscid, quasi-geostrophic flows that satisfy the Arnol'd–Blumen and Benzi *et al.* stability criteria, subject to given zonally symmetric boundary conditions, must themselves be zonally symmetric parallel flows in the entire domain. In Appendix A, we extend Andrews' argument to show that for any two-dimensional, incompressible flow problem with a given (boundary) symmetry, Arnol'd-stable solutions must have the same symmetry. Hence Arnol'd stability of these solutions is essentially a nonlinear extension of classic results on linear stability for separable problems.

Linear stability analyses by Pierini (1985), Laedke & Spatschek (1986) and Swaters (1986) provide partial stability results for propagating modons. The numerical study of McWilliams *et al.* (1981) indicated that eastward-propagating modons are quite robust to some finite-amplitude perturbations and that scale-dependent thresholds exist for the amplitude of perturbations beyond which modons lose their stability. Carnevale *et al.* (1988) showed recently that the linear stability analyses of Pierini (1985) and Swaters (1986) were invalid, since their (formal) stability criteria are satisfied not only by stable but also by a rather unrestricted class of unstable modes. By using a simple but informative model of triadic interactions, Carnevale and colleagues provided a counterexample for the stability criteria derived and discussed by Swaters (1986) and reviewed by Flierl (1987); they concluded that the stability of eastward-moving modons is still an open question, while the derivation of Laedke & Spatschek (1986) for westward-drifting modons stands.

The purpose of this paper is to give an analytic proof of the stability of a stationary modon (Stern's solution) as a preliminary step toward the stability analysis of drifting modons (Sakuma 1989). In order to achieve this end we introduce a slightly more general setting, which is free from the symmetry restriction found by Andrews (1984), and seek a weaker Lyapunov-type stability criterion than the ones derived so far. To start with, we consider the stability of steady flows governed by the shallow-water equations, in which the velocity field is not necessarily non-divergent. In studying Lyapunov stability of various fluids and plasmas, Holm *et al.* (1983) have already obtained stability criteria for numerous systems, including the shallow-water equations. Our criterion turns out to be weaker than that of Holm and colleagues, and the stability of a stationary modon in a quiescent background is easily covered by our, but not by Holm's, criterion.

To be more precise, recall that two different notions of Lyapunov stability were introduced by Arnol'd and by Holm and colleagues, i.e. nonlinear stability and formal stability. *Nonlinear stability* means essentially the stability of a fixed point of a given system to finite-amplitude perturbations. Arnol'd (1969) provided a nonlinear stability proof for two-dimensional non-divergent flows by giving (global) convexity estimates of a certain conservative quantity; Holm and colleagues further developed these estimates to derive nonlinear stability criteria for other flows.

Formal stability is defined by Holm *et al.* (1985) as follows: An equilibrium solution U_e of a system $\dot{\boldsymbol{u}} = \boldsymbol{x}(\boldsymbol{u})$ is formally stable if a conservative quantity exists whose second variation at this solution is positive (or negative) definite; i.e. only local convexity of the quantity is required. These authors showed that formal stability is sufficient for linear stability of Hamiltonian systems. In their stability analyses, Blumen (1968) and Benzi *et al.* (1982) checked the formal stability of inviscid, rotating two-dimensional flows.

The stability discussed in this paper is formal stability. Our mathematical formulation of the problem recovers the formal stability criterion for the shallow-water equations of Holm and colleagues and, from it, one can easily get the Arnol'd–Blumen criterion (1.1) for non-divergent flows. In addition, our new formulation reduces the number of 'formally independent variational variables', in a precise sense to become apparent forthwith, and hence it yields a weaker criterion than the classical one.

In order to clarify the above statement, let us consider the heuristic model of triadic interaction used by Carnevale *et al.* (1988). The model is

$$\dot{x} = ayz; \quad \dot{y} = bzx; \quad \dot{z} = cxy, \tag{1.2a–c}$$

with x, y and z the amplitudes of three plane waves having wavenumbers l, m and n, respectively. The coefficients a, b and c are related to these wavenumbers by $a = n^2-m^2$, $b = l^2-n^2$, $c = m^2-l^2$. The dynamics represented by (1.2) can be regarded as a simplification of two-dimensional non-divergent, inviscid flow. The flow conserves energy E and enstrophy Z given respectively by

$$E = x^2+y^2+z^2; \quad Z = l^2x^2+m^2y^2+n^2z^2. \tag{1.3a, b}$$

The quadratic form L

$$L = E-Z/m^2 \tag{1.4}$$

is useful in studying the stability of a stationary solution $(0, Y, 0)$ with wavenumber m. Clearly, L is analogous to the Arnol'd stability norm δ^2H for two-dimensional non-divergent flows,

$$\delta^2H = \int\!\!\int\left[\left(\frac{\partial\delta\psi}{\partial x}\right)^2+\left(\frac{\partial\delta\psi}{\partial y}\right)^2+\frac{\nabla\psi}{\nabla\Delta\psi}(\Delta\delta\psi)^2\right]\mathrm{d}x\,\mathrm{d}y. \tag{1.5}$$

The analogy assumes a more concrete form if we formally consider L as the function of independent variables Z and x, y, z in E, i.e.

$$L = x^2+y^2+z^2-\frac{1}{m^2}Z. \tag{1.6}$$

With these formally independent variables, however, L cannot be positive definite, while for δ^2H positive definiteness is attained by $\nabla\psi/\nabla\Delta\psi > 0$. Yet if we take advantage of the simple dependence of Z on x, y and z (equation (1.3b)), then L becomes

$$L = (cx^2-az^2)/m^2. \tag{1.7}$$

So L becomes definite and hence a given stationary solution $(0, Y, 0)$ is stable if $ac < 0$. For the stability norm δ^2H of divergent flows, to the best of our knowledge, no simplified expression has been derived so far which corresponds to the form (1.7). In this paper it is shown that such a modification is actually possible for the stability norm $\delta^2 H$ of the shallow-water equations.

Our formulation assumes that the variation of the free surface is constrained in such a way that total mass be conserved. This assumption is satisfied automatically by two-dimensional non-divergent motions and is physically natural in the more general case of locally divergent two-dimensional motions, without sources or sinks of fluid.

In §2 we introduce a gauge formalism associated with a vector potential $\vec{\boldsymbol{\phi}}$ for the 3-vector of height h and horizontal momentum components hv_1 and hv_2. A Lyapunov functional related to the Hamiltonian is obtained from the action integral of the shallow-water equations; it is quadratic in gauge-independent variations. Formal stability of stationary solutions is derived at the end of §2 for an inertial frame of reference and in §3 for a rotating frame. The new stability criterion is applied to zonal shear flows in §4, and to Stern's modon in §5. Concluding remarks follow in §6. The connection between symmetry and Arnol'd stability is discussed in Appendix A, and certain technical details concerning the gauge formalism are given in Appendix B. Appendix C contains a simple example which illustrates geometrically some of the arguments in §2.

2. Derivation of the new stability criterion

2.1. *The energy–Casimir convexity (E–CC) method in gauge variables*

The shallow-water equations in an inertial reference frame are

$$\frac{\partial \boldsymbol{v}}{\partial t}+(\boldsymbol{v}\cdot\boldsymbol{\nabla})\boldsymbol{v}=-\boldsymbol{\nabla}gh, \quad \frac{\partial h}{\partial t}+\boldsymbol{\nabla}\cdot(h\boldsymbol{v})=0; \tag{2.1a, b}$$

here h is the height of the free surface and $\boldsymbol{v}$ the Cartesian velocity of a column of fluid. For the subsequent analysis, it is mathematically convenient to use a three-dimensional coordinate space with

$$x_0=ct, \quad x_1=x, \quad x_2=y. \tag{2.2a–c}$$

No physical significance is implied for this space, and c is taken arbitrarily to be unity. The mass conservation equation (2.1b) implies that the '3-vector' $\vec{\boldsymbol{M}}=(hc, hv_1, hv_2)$ is non-divergent in this three-dimensional space and that it admits a '3-vector' potential $\vec{\boldsymbol{\phi}}=(\phi_0,\phi_1,\phi_2)$:

$$hc=\frac{\partial\phi_2}{\partial x_1}-\frac{\partial\phi_1}{\partial x_2}, \quad hv_1=\frac{\partial\phi_0}{\partial x_2}-\frac{\partial\phi_2}{\partial x_0}, \quad hv_2=\frac{\partial\phi_1}{\partial x_0}-\frac{\partial\phi_0}{\partial x_1}. \tag{2.3a–c}$$

We shall use bold type with an arrow over to denote 3-vectors, and bold type for 2-vectors.

For given hc, hv_1, hv_2, the components ϕ_i, $i=0,1,2$, of the 'vector' potential $\vec{\boldsymbol{\phi}}$ are undetermined to within the 'three-dimensional gradient' of an arbitrary function $G(x_0,x_1,x_2)$. An additional relation, a *gauge condition*, must be added to determine G (Landau & Lifshitz 1962). In particular, any one of ϕ_i, say ϕ_0, can equal an arbitrary smooth function by adjusting G. This arbitrariness of G in our gauge formulation is used to best advantage in determining a Casimir function.

The first step in the energy–Casimir convexity (E–CC) method (e.g. Salmon 1988) is to introduce a functional K of the form

$$K=K_A+K_B, \tag{2.4}$$

where K_A and K_B represent the total energy and a Casimir functional respectively. K_B is some function of advectively invariant quantities, i.e. of quantities for which the material derivative vanishes identically. Both K_A and K_B, separately, are integral invariants of the system. The reason for the existence of an invariant K_B, in addition to the total energy K_A, is the degeneracy of the Poisson bracket of the underlying Hamiltonian formulation for Eulerian variables. The functional dependence of K_B on advectively invariant quantities is determined so that the invariant K be extremal at an equilibrium. The E–CC method was first applied by Arnol'd (1965), using an arbitrary function of vorticity as the second, Casimir invariant. It was generalized by Abarbanel *et al.* (1986, especially Appendix A).

In the present formulation, dimensional considerations suggest a Casimir

$$K_B \equiv -\iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\phi_0\left(\frac{\partial v_2}{\partial x_1}-\frac{\partial v_1}{\partial x_2}\right), \tag{2.5}$$

where Ω is the two-dimensional domain in which the flow occurs. This K_B is non-unique since ϕ_0 is still arbitrary. Hence an appropriate choice of G corresponds to the right determination of the Casimir K_B. In other words, in our formulation, the choice of a Casimir is equivalent to the reduction of the number of independent gauge variables ϕ_i. The freedom of the gauge formalism allows us to choose K_B so that the first variation of K vanish at an equilibrium, and that $\delta^2 K$ be positive definite, without requiring the term-wise definiteness of the second variation in total energy and in the Casimir, as in the Arnol'd, Blumen and Holm approaches.

In order to find an appropriate form for the gauge condition, we notice first that K_B becomes an invariant if the advective derivative of ϕ_0 vanishes,

$$\frac{\mathrm{D}\phi_0}{\mathrm{D}t} \equiv \frac{\partial \phi_0}{\partial t}+v_1\frac{\partial \phi_0}{\partial x_1}+v_2\frac{\partial \phi_0}{\partial x_2} = 0. \tag{2.6}$$

Actually, if ϕ_0 is an advective quantity, so is $\phi_0 Q$, where Q is the potential vorticity given by

$$Q \equiv \frac{1}{h}\left(\frac{\partial v_2}{\partial x_1}-\frac{\partial v_1}{\partial x_2}\right). \tag{2.7}$$

Using (2.6), (2.7), the integrand in

$$\frac{\partial K_B}{\partial t} = -\iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\frac{\partial h\phi_0 Q}{\partial t}$$

can be rewritten as

$$\begin{aligned}-\frac{\partial h\phi_0 Q}{\partial t} &= -\frac{\partial h}{\partial t}\phi_0 Q - h\frac{\partial \phi_0 Q}{\partial t} = \phi_0 Q\boldsymbol{\nabla}\cdot(h\boldsymbol{v})+h\boldsymbol{v}\cdot\boldsymbol{\nabla}(\phi_0 Q)\\ &= \boldsymbol{\nabla}\cdot(\phi_0 Q h\boldsymbol{v}).\end{aligned}$$

Thus
$$\frac{\partial K_B}{\partial t} = \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\boldsymbol{\nabla}\cdot(\phi_0 Q h\boldsymbol{v}) = \oint_{\partial\Omega}\mathrm{d}l\,\phi_0 Q h v_\mathrm{n} = 0,$$

since the normal component v_n of $\boldsymbol{v}$ vanishes along the boundary $\partial\Omega$ of the flow domain Ω. Using (2.3), the gauge condition (2.6) becomes

$$\begin{aligned}0 &= \frac{\partial \phi_0}{\partial x_0}+\frac{v_1}{c}\left(\frac{\partial \phi_1}{\partial x_0}-hv_2\right)+\frac{v_2}{c}\left(\frac{\partial \phi_2}{\partial x_0}+hv_1\right)\\ &= \frac{\partial \phi_0}{\partial x_0}+\frac{v_1}{c}\frac{\partial \phi_1}{\partial x_0}+\frac{v_2}{c}\frac{\partial \phi_2}{\partial x_0}.\end{aligned} \tag{2.8}$$

Remark 2.1. The integrand of (2.5), $h\phi_0 Q$, has the same form as that of the perfect-fluid Casimir, $\rho f(q)$ (Salmon 1988, equation (6.8)), where ρ is density, q is potential vorticity and $f(q)$ an arbitrary function. Thus a choice of $\phi_0 = \phi_0(Q)$ here corresponds to the choice of $f(q)$ in the perfect-fluid equations (Salmon 1988).

Equations (2.3) indicate that an arbitrary variation $\delta\vec{\boldsymbol{\phi}}$ defines a 3-momentum variation $\delta\vec{\boldsymbol{M}}$ that conserves mass in two-dimensional physical space,

$$\vec{\nabla}\cdot\delta\vec{\boldsymbol{M}} = 0. \tag{2.9}$$

Any gauge condition imposed on $\delta\vec{\boldsymbol{\phi}}$ does not interfere with (2.9) being satisfied. A suitable gauge condition is suggested by the following observation: an infinitesimally small change of $\vec{\boldsymbol{\phi}}$ induced by the fluid motion with time at a fixed location, i.e. $\delta_t\phi_i \equiv (\partial\phi_i/\partial x_0)\,\mathrm{d}x_0$, can be looked upon as a particular variation $\delta\phi_i$; the gauge condition (2.8) is equivalent for such a variation to $\vec{\boldsymbol{M}}\cdot\delta_t\vec{\phi} = 0$. Hence a natural extension of the condition (2.8) to arbitrary $\delta\vec{\boldsymbol{\phi}}$ appears to be

$$\vec{\boldsymbol{M}}\cdot\delta\vec{\boldsymbol{\phi}} = 0. \tag{2.10}$$

It is easy to check that (2.10) is indeed a gauge condition on $\delta\vec{\boldsymbol{\phi}}$. To see that let $\delta\vec{\boldsymbol{\phi}}^*$ be a permissible choice of $\delta\vec{\boldsymbol{\phi}}$ for a given $\delta\vec{\boldsymbol{M}}$. For every $\delta\phi^*$, δG may be determined in such a way that

$$c\left(\delta\phi_0^* + \frac{\partial\delta G}{\partial x_0}\right) + v_1\left(\delta\phi_1^* + \frac{\partial\delta G}{\partial x_1}\right) + v_2\left(\delta\phi_2^* + \frac{\partial\delta G}{\partial x_2}\right) = 0, \tag{2.11a}$$

namely,
$$c\frac{\partial\delta G}{\partial x_0} + v_1\frac{\partial\delta G}{\partial x_1} + v_2\frac{\partial\delta G}{\partial x_2} = -(c\delta\phi_0^* + v_1\delta\phi_1^* + v_2\delta\phi_2^*). \tag{2.11b}$$

Here v_1, v_2 and $\delta\vec{\boldsymbol{\phi}}^*$ are all known quantities in determining δG so that (2.11*b*) is a linear, first-order partial differential equation with the unknown δG. Therefore δG can be determined subject to a suitable boundary condition. Then (2.11*a*) says that, in particular, a gauge condition (2.10) can be imposed on $\delta\vec{\boldsymbol{\phi}}$.

The following diagram

Mass continuity in M-variables	Gauge conditions for Casimir in ϕ-variables
$\vec{\nabla}\cdot\vec{\boldsymbol{M}} = 0$ (2.1*b*)	$\mathrm{D}\phi_0/\mathrm{D}t = 0$ (2.6)
$\downarrow$	$\downarrow$
(natural extension to $\delta\vec{\boldsymbol{M}}$)	(natural extension to $\delta\vec{\boldsymbol{\phi}}$)
$\downarrow$	$\downarrow$
$\vec{\nabla}\cdot\delta\vec{\boldsymbol{M}} = 0$ (2.9) $\Rightarrow$	$\vec{\boldsymbol{M}}\cdot\delta\vec{\boldsymbol{\phi}} = 0$ (2.10)

illustrates the relation between mass continuity in $\vec{\boldsymbol{M}}$-variables and the gauge condition for the Casimir in $\vec{\boldsymbol{\phi}}$-variables. Notice that the condition (2.10) combines the two constraints (2.6) and (2.9) in a consistent manner and defines a two-dimensional manifold $\mathscr{M}_\phi^2$ embedded in three-dimensional ϕ-space; the tangent 'plane' of $\mathscr{M}_\phi^2$ is given by (2.10). On this manifold, mass continuity of the flow itself, (2.1*b*), and of its variations, (2.9), is automatically satisfied, and so is (2.6). In §2.2, we develop a Hamiltonian formulation for our dynamical system (2.1) based on the principle of least action defined on the cotangent bundle π_ϕ of $\mathscr{M}_\phi^2$ (Arnol'd 1978). This will yield a Lyapunov functional K, cf. (2.4), for variations restricted by (2.10).

2.2. *A Lyapunov functional*

In §2.1, we showed that the functional K becomes an invariant of the system if condition (2.6) holds. For the invariant K to be a Lyapunov functional, steady states must correspond to its extremal points. Sufficient conditions for this to be so are given in the following proposition.

PROPOSITION 2.1. *Imposing the gauge condition* (2.6) *on* $\vec{\boldsymbol{\phi}}$, *the functional K becomes a generalized Hamiltonian defined on the cotangent bundle π_ϕ of the manifold $\mathscr{M}_\phi^2$. The extremal points of K correspond to the steady states of the system if $\boldsymbol{v}$ satisfies Arnol'd's equivortical condition*

$$\delta\oint_{\partial\Omega} \mathrm{d}l v_l = 0,$$

and the mass-conservation condition

$$\delta\oint_{\partial\Omega} \mathrm{d}l\,\delta\phi_l = 0;$$

here $(\cdot)_l$ denotes the tangential component of the 2-vector along the boundary $\partial\Omega$.

Proof. The action integral A is defined as

$$A \equiv \int \mathrm{d}x_0 \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\mathscr{L}\left[\frac{\partial\phi_i}{\partial x_j}\right], \tag{2.12a}$$

where

$$\mathscr{L} \equiv \tfrac{1}{2}h(v_1^2+v_2^2)-\tfrac{1}{2}gh^2. \tag{2.12b}$$

is the Lagrangian density and (2.3) are used to express $\mathscr{L}$ in terms of the derivatives of $\phi_i, i = 0, 1, 2$. The variation δA with respect to $\delta\phi_i$ yields

$$\delta A = \delta A_V+\delta A_S, \tag{2.13a}$$

where the interior contribution δA_V and boundary contribution δA_S are

$$\delta A_V = \int \mathrm{d}x_0 \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\left[-\frac{1}{c}\left(\frac{\partial v_2}{\partial t}+\frac{\partial B}{\partial x_2}\right)\delta\phi_1 + \frac{1}{c}\left(\frac{\partial v_1}{\partial t}+\frac{\partial B}{\partial x_1}\right)\delta\phi_2+\left(\frac{\partial v_2}{\partial x_1}-\frac{\partial v_1}{\partial x_2}\right)\delta\phi_0\right], \tag{2.13b}$$

$$\delta A_S = \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2[v_2\,\delta\phi_1-v_1\,\delta\phi_2]-\iint \mathrm{d}x_0\,\mathrm{d}x_2\left[v_2\,\delta\phi_0+\frac{B}{c}\delta\phi_2\right] + \iint \mathrm{d}x_0\,\mathrm{d}x_1\left[v_1\,\delta\phi_0+\frac{B}{c}\delta\phi_1\right], \tag{2.13c}$$

and $B \equiv \frac{1}{2}(\boldsymbol{v}\cdot\boldsymbol{v})+gh$ is the Bernoulli function. The natural boundary conditions for the variations are that $\delta\phi_i = 0$, $i = 0, 1, 2$, making δA_S vanish.

The gauge condition (2.10) permits one to eliminate $\delta\phi_0$ in (2.13*b*), yielding

$$\delta A_V = \int \mathrm{d}x_0 \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\frac{1}{c}(-F_2\,\delta\phi_1+F_1\,\delta\phi_2), \tag{2.14a}$$

where F_1 and F_2 are

$$F_1 = \frac{\partial v_1}{\partial t}+\frac{\partial B}{\partial x_1}-\left(\frac{\partial v_2}{\partial x_1}-\frac{\partial v_1}{\partial x_2}\right)v_2, \tag{2.14b}$$

$$F_2 = \frac{\partial v_2}{\partial t}+\frac{\partial B}{\partial x_2}+\left(\frac{\partial v_2}{\partial x_1}-\frac{\partial v_1}{\partial x_2}\right)v_1. \tag{2.14c}$$

Since $\delta\phi_1$ and $\delta\phi_2$ are arbitrary, the necessary and sufficient condition for $\delta A = \delta A_V = 0$ is that $F_1 = 0$ and $F_2 = 0$. We see that these two conditions are the momentum equation (2.1 a) in our shallow-water system. Hence the principle of least action holds on the tangent bundle of $\mathscr{M}^2_\phi$.

A momentum 2-vector $\boldsymbol{\pi} = (\pi_1, \pi_2)$ conjugate to $\boldsymbol{\phi} = (\phi_1, \phi_2)$ is now defined by

$$\pi_i \equiv \frac{\partial \mathscr{L}}{\partial\left(\dfrac{\partial \phi_i}{\partial x_0}\right)} = c\left(\frac{\partial \phi_i}{\partial x_0} - \frac{\partial \phi_0}{\partial x_i}\right) \Big/ \left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}\right), \quad i = 1, 2. \tag{2.15a, b}$$

Using (2.3), we easily see that

$$\pi_1 = v_2, \quad \pi_2 = -v_1. \tag{2.16}$$

The Hamiltonian density $\mathscr{H}$ is defined by $\mathscr{L}$ through the Legendre transformation (e.g. Goldstein 1980, equation (12-55), p. 563):

$$\mathscr{H} = \sum_{i=1}^{2} \pi_i \frac{\partial \phi_i}{\partial x_0} - \mathscr{L}. \tag{2.17}$$

Remark 2.2. The absence of $\pi_0 \partial\phi_0/\partial x_0$ in the Hamiltonian density $\mathscr{H}$ reflects the fact that, according to (2.10), ϕ_0, ϕ_1 and ϕ_2 are no longer independent in our formulation. As far as the relation between ϕ_i and π_i is concerned, we may extend (2.15) to define π_0 as

$$\pi_0 \equiv \frac{\partial \mathscr{L}}{\partial\left(\dfrac{\partial \phi_0}{\partial x_0}\right)}. \tag{2.15c}$$

It is then easily seen that $\pi_0 = 0$, since $\mathscr{L}$ does not depend explicitly on $\partial\phi_0/\partial x_0$.

Eliminating $\partial\phi_i/\partial x_0, i = 1, 2$, in (2.17) by using (2.3), we get

$$\mathscr{H} = \frac{1}{2c}\left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}\right)(\pi_1^2 + \pi_2^2) + \frac{g}{2c^2}\left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}\right)^2 + \pi_1 \frac{\partial \phi_0}{\partial x_1} + \pi_2 \frac{\partial \phi_0}{\partial x_2}. \tag{2.18}$$

Since we can add the divergence of any two-dimensional vector to the Lagrangian density $\mathscr{L}$ without affecting the principle of least action, $\mathscr{H}$ may be redefined as

$$\mathscr{H} = \mathscr{H}_A + \mathscr{H}_B, \tag{2.19a}$$

where

$$\mathscr{H}_A = \frac{1}{2c}\left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}\right)(\pi_1^2 + \pi_2^2) + \frac{g}{2c^2}\left(\frac{\partial \phi_2}{\partial x_1} - \frac{\partial \phi_1}{\partial x_2}\right)^2, \tag{2.19b}$$

$$\mathscr{H}_B = -\phi_0\left(\frac{\partial \pi_1}{\partial x_1} + \frac{\partial \pi_2}{\partial x_2}\right). \tag{2.19c}$$

In the Hamiltonian density (2.19), the first term $\mathscr{H}_A$ is the total energy density, while the second term $\mathscr{H}_B$ equals the Casimir density which we tentatively introduced in (2.5).

With this $\mathscr{H}$, it can be shown that the momentum equation (2.1 a) and (2.3 b, c) are recovered by allowing the first variation of the action integral A

$$A = \int \mathrm{d}x_0 \iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2 \left(\sum_{i=1}^{2} \pi_i \frac{\partial \phi_i}{\partial x_0} - \mathscr{H}\right)$$

to vanish with respect to independent variations of ϕ_1, ϕ_2, π_1, and π_2. Hence we have proved the first half of the proposition.

Now let us consider the first variations of K_A and K_B with respect to four independent variables ϕ_1, ϕ_2, π_1 and π_2. A little manipulation yields

$$\delta K_A = \iint_\Omega \mathrm{d}x_1 \mathrm{d}x_2 \left[\frac{\partial \pi_1}{\partial x_0}\delta\phi_1 + \frac{\partial \pi_2}{\partial x_0}\delta\phi_2 + h\{\pi_1(\delta\pi_1 - Q\delta\phi_2) + \pi_2(\delta\pi_2 + Q\delta\phi_1)\}\right] + \frac{1}{c}\oint_{\partial\Omega} \mathrm{d}l\, B\delta\phi_l, \quad (2.20a)$$

$$\delta K_B = -\iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2 \left[-\frac{\partial \phi_1}{\partial x_0}\delta\pi_1 - \frac{\partial \phi_2}{\partial x_0}\delta\pi_2 + h\{\pi_1(\delta\pi_1 - Q\delta\phi_2) + \pi_2(\delta\pi_2 + Q\delta\phi_1)\}\right] - \oint_{\partial\Omega} \mathrm{d}l\phi_0\, \delta v_l, \quad (2.20b)$$

where Q is the potential vorticity defined in (2.7). In deriving δK_B, the gauge condition (2.10) was used again. For steady states both $\partial\pi_i/\partial x_0$ and $\partial\phi_i/\partial x_0$ vanish, so (2.20a, b) reduce to

$$\delta K_A = \iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2\, h[\pi_1(\delta\pi_1 - Q\delta\phi_2) + \pi_2(\delta\pi_2 + Q\delta\phi_1)] + \frac{B}{c}\oint_{\partial\Omega} \mathrm{d}l\, \delta\phi_l, \quad (2.21a)$$

$$\delta K_B = -\iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2\, h[\pi_1(\delta\pi_1 - Q\delta\phi_2) + \pi_2(\delta\pi_2 + Q\delta\phi_1)] - \frac{\phi_0}{c}\oint_{\partial\Omega} \mathrm{d}l\, \delta v_l, \quad (2.21b)$$

since B and ϕ_0 are constant along the boundaries. The justification of $\partial\phi_i/\partial x_0 = 0$ is not entirely obvious and is given in Appendix B.

The two key assumptions from the statement of Proposition 2.1 are required at this point:

$$\oint_{\partial\Omega} \mathrm{d}l\, \delta v_l = 0; \quad (2.22)$$

$$\oint_{\partial\Omega} \mathrm{d}l\, \delta\phi_l = 0. \quad (2.23)$$

Assumption (2.22) states that the integrated vorticity over the entire domain remains the same, and was already used in Arnol'd's original stability analysis for two-dimensional non-divergent flows. The constraint (2.23) is a new assumption, for the shallow-water equations; it also has a very simple physical meaning, namely the conservation of total mass. Indeed, from (2.3a), we easily see that

$$\iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2\, \delta h = \frac{1}{c}\iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2 \left(\frac{\partial\delta\phi_2}{\partial x_1} - \frac{\partial\delta\phi_1}{\partial x_2}\right) = \frac{1}{c}\oint_{\partial\Omega} \mathrm{d}l\, \delta\phi_l.$$

So (2.23) is equivalent to

$$\delta\iint_\Omega \mathrm{d}x_1\, \mathrm{d}x_2\, h = 0. \quad (2.24)$$

The total mass conservation condition (2.24) makes good physical sense and is consistent with the non-divergence of $\delta\vec{M}$ mentioned earlier. Under the two assumptions (2.22) and (2.23), it follows from (2.21) that

$$\delta K \equiv \delta K_A + \delta K_B = 0 \quad (2.25)$$

for steady states. Thus we have proved Proposition 2.1.

2.3. *Formal stability*

Since K is an integral invariant, a given steady solution of (2.1) is formally stable if the second variation of K is either positive or negative definite near that solution. The second variation of K_A with respect to four independent variations $\delta\phi_1$, $\delta\phi_2$, $\delta\pi_1$ and $\delta\pi_2$ is

$$\delta^2 K_A = \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2[\tfrac{1}{2}h(\delta\pi_1^2+\delta\pi_2^2)+\delta h(\pi_1\delta\pi_1+\pi_2\,\delta\pi_2)+\tfrac{1}{2}g\,\delta h^2], \tag{2.26}$$

where

$$\delta h = \frac{1}{c}\left(\frac{\partial\delta\phi_2}{\partial x_1}-\frac{\partial\delta\phi_1}{\partial x_2}\right).$$

In calculating the second variation of K_B, attention must be paid to the variation of ϕ_0, since by (2.10) it is no longer independent of $\delta\phi_1$ and $\delta\phi_2$. So we temporarily denote it by $\Delta\phi_0$. From (2.19c), one easily gets

$$\begin{aligned}&\mathscr{K}_B[\phi_0+\Delta\phi_0,\pi_1+\delta\pi_1,\pi_2+\delta\pi_2]-\mathscr{K}_B[\phi_0,\pi_1,\pi_2]\\&\quad=-\phi_0\left(\frac{\partial\delta\pi_1}{\partial x_1}+\frac{\partial\delta\pi_2}{\partial x_2}\right)-\Delta\phi_0\left(\frac{\partial\pi_1}{\partial x_1}+\frac{\partial\pi_2}{\partial x_2}\right)-\Delta\phi_0\left(\frac{\partial\delta\pi_1}{\partial x_1}+\frac{\partial\delta\pi_2}{\partial x_2}\right).\end{aligned} \tag{2.27}$$

If ϕ_0 were simply a function of ϕ_1 and ϕ_2, $\Delta\phi_0$ would assume the form

$$\Delta\phi_0 = \sum_{i=1}^{2}\frac{\partial\phi_0}{\partial\phi_i}\delta\phi_i+O(\delta\phi_1^2+\delta\phi_2^2).$$

Thus, in (2.27), not only the third term but also the second term appears to contribute to the second variation of K_B. However, (2.10) should not be interpreted as expressing the first variation of ϕ_0 with respect to independent variations ϕ_1 and ϕ_2. Rather $\delta\phi_0$ must be specified by two arbitrary variations $\delta\phi_1$ and $\delta\phi_2$ subject to (2.10), which does not necessarily mean that $c\partial\phi_0/\partial\phi_i = -v_i$, $i = 1,2$. As mentioned already, in the three-dimensional ϕ-space, $\delta\vec{\boldsymbol{\phi}}$ must be on the tangent plane of $\mathscr{M}^2_\phi$ and (2.10) is the exact condition for that, so it should be treated as a supplementary condition which is imposed after taking formally independent variations of ϕ_i, π_i, $i = 0,1,2$. In actuality, we have shown that, for a given steady state, the generalized Hamiltonian K becomes extremal for the variations $\{\delta\pi_1,\ \delta\pi_2,\ \delta\vec{\boldsymbol{\phi}}\colon\ \vec{\boldsymbol{M}}\cdot\delta\vec{\boldsymbol{\phi}} = 0\}$. The gauge condition (2.10) keeps the variation $\delta\vec{\boldsymbol{\phi}}$ exactly on the tangent plane of the manifold $\mathscr{M}^2_\phi$ on which our principle of least action is defined. Thus setting $\Delta\phi_0 = \delta\phi_0$, one gets

$$\begin{aligned}\delta^2 K_B &= -\iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\,\delta\phi_0\left(\frac{\partial\delta\pi_1}{\partial x_1}+\frac{\partial\delta\pi_2}{\partial x_2}\right).\\&= -\iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\frac{1}{c}(\pi_2\,\delta\phi_1-\pi_1\,\delta\phi_2)\left(\frac{\partial\delta\pi_1}{\partial x_1}+\frac{\partial\delta\pi_2}{\partial x_2}\right).\end{aligned} \tag{2.28}$$

For reasons to become clear later on, we define $\delta\chi$ by

$$\delta\chi_1 \equiv \delta\pi_1-Q\delta\phi_2,\quad \delta\chi_2 \equiv \delta\pi_2+Q\delta\phi_1. \tag{2.29a,b}$$

Using the identities

$$\frac{1}{c}(\pi_2\,\delta\phi_1-\pi_1\,\delta\phi_2) = -\frac{1}{cQ}(\pi_1\,\delta\pi_1+\pi_2\,\delta\pi_2)+\frac{1}{cQ}(\pi_1\,\delta\chi_1+\pi_2\,\delta\chi_2)$$

and

$$\frac{\partial\delta\pi_1}{\partial x_1}+\frac{\partial\delta\pi_2}{\partial x_2} = cQ\delta h+\frac{1}{Q}\left(\frac{\partial Q}{\partial x_1}\delta\pi_1+\frac{\partial Q}{\partial x_2}\delta\pi_2\right)-\frac{1}{Q}\left(\frac{\partial Q}{\partial x_1}\delta\chi_1+\frac{\partial Q}{\partial x_2}\delta\chi_2\right)+\frac{\partial\delta\chi_1}{\partial x_1}+\frac{\partial\delta\chi_2}{\partial x_2},$$

$\delta^2 K_B$ can be rewritten as

$$\delta^2 K_B = \delta^2 K_B^{(1)} + \delta^2 K_B^{(2)}, \tag{2.30}$$

where

$$\delta^2 K_B^{(1)} = \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2 \frac{1}{cQ}(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2)\left[cQ\delta h + \frac{1}{Q}\left(\frac{\partial Q}{\partial x_1}\delta\pi_1 + \frac{\partial Q}{\partial x_2}\delta\pi_2\right)\right], \tag{2.31}$$

$$\delta^2 K_B^{(2)} = -\iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2\left[\frac{1}{cQ}\boldsymbol{\pi}\cdot\delta\boldsymbol{\chi}(cQ\delta h + \nabla\ln Q\cdot\delta\boldsymbol{\pi} - \nabla\ln Q\cdot\delta\boldsymbol{\chi} + \nabla\cdot\delta\boldsymbol{\chi}) - \frac{1}{cQ}\boldsymbol{\pi}\cdot\delta\boldsymbol{\pi}(-\nabla\ln Q\cdot\delta\boldsymbol{\chi} + \nabla\cdot\delta\boldsymbol{\chi})\right]. \tag{2.32}$$

From (2.26), (2.31) and (2.32), we see that $\delta^2 K_A + \delta^2 K_B^{(1)}$ is gauge independent, while $\delta^2 K_B^{(2)}$ is not. That is to say $\delta^2 K_A + \delta^2 K_B^{(1)}$ is expressed in terms of δh, $\delta\pi_1(=\delta v_2)$, $\delta\pi_2(=-\delta v_1)$ and observable field variables for a given steady state, while $\delta^2 K_B^{(2)}$ depends explicitly upon the unobservable quantity $\delta\boldsymbol{\chi}$, equation (2.29). Since the stability criterion should not depend on the particular choice of gauge variables, one is led therewith to conjecture that the stability of a given steady state may not depend upon the definiteness of $\delta^2 K_B^{(2)}$.

This conjecture can be formulated precisely and proven rigorously as:

Statement A. *If $\delta^2 K_A + \delta^2 K_B^{(1)}$ is either positive or negative definite, then a given steady state is linearly stable.*

In order to prove this statement, we consider the contrapositive one, which becomes true in case Statement A is actually true and vice versa. Namely,

Statement B. *If a given steady state is linearly unstable, then $\delta^2 K_A + \delta^2 K_B^{(1)}$ is indefinite.*

Remark 2.3. The proof of Statement B is carried out by following a trajectory of the system through the equilibrium point, call it O, whose stability we wish to prove. It is shown that the existence of any unstable direction issuing from O implies that $\delta^2 K_A + \delta^2 K_B^{(1)}$ vanishes along such a direction, in agreement with Statement B. This type of argument might be unfamiliar to some readers, and is certainly less straightforward than the usual connection between dynamic stability of the system and geometric convexity, near the equilibrium, of a Hamiltonian hypersurface. The latter connection is natural and well-established when all the variables are observables, such as positions and velocities.

In our gauge formalism, one half of the variables are not observables, being constrained vector-potential components. Hence the relation between dynamical stability and local convexity of the Hamiltonian is blurred, owing to the presence of the gauge-dependent part $\delta^2 K_B^{(2)}$ of the quadratic Hamiltonian form $\delta^2 K$. This part is not necessarily constant in time (see also Appendix B), and hence the approach we have taken to the stability proof becomes more natural. The approach is illustrated by a simple example from point-mass mechanics, a conservative nonlinear pendulum, in Appendix C. Readers might want to look at this Appendix before or instead of going through the proof below.

Proof. Let O be the equilibrium point in phase space which corresponds to the steady state. We have seen that the Hamiltonian K is an integral invariant and that O has to be a critical point of K (Proposition 2.1). If K is definite, O is an *elliptic point*: trajectories near O are topologically circles, O is neutrally stable (linearly), and no trajectory passes through O. If K is indefinite, O is a *hyperbolic point*: both stable and unstable trajectories pass through O, and O is (linearly) unstable.

The proof proceeds by following a trajectory through the equilibrium point O. Along any trajectory, the momentum equation (2.1a) and (2.3) hold. Combining these two sets of equations, we obtain

$$\frac{\partial \pi_1}{\partial x_0}+\frac{1}{c}\frac{\partial B}{\partial x_2}+Q\left(\frac{\partial \phi_0}{\partial x_2}-\frac{\partial \phi_2}{\partial x_0}\right)=0, \tag{2.33a}$$

$$-\frac{\partial \pi_2}{\partial x_0}+\frac{1}{c}\frac{\partial B}{\partial x_1}-Q\left(\frac{\partial \phi_1}{\partial x_0}-\frac{\partial \phi_0}{\partial x_1}\right)=0. \tag{2.33b}$$

The additional constraint on ϕ_0 we impose at this point is that $\phi_0=\phi_0(Q)$. It is consistent with Remark 2.1, with (2.6), and with the fact that such a relation has to hold in a steady state. With this constraint, the above equations become

$$\frac{\partial \pi_1}{\partial x_0}-Q\frac{\partial \phi_2}{\partial x_0}=-\frac{\partial}{\partial x_2}\left(\frac{B}{c}+F\right),\quad \frac{\partial \pi_2}{\partial x_0}+Q\frac{\partial \phi_1}{\partial x_0}=\frac{\partial}{\partial x_1}\left(\frac{B}{c}+F\right), \tag{2.34a, b}$$

where

$$\mathrm{D}F/\mathrm{D}\phi_0=Q(\phi_0). \tag{2.34c}$$

In particular, for steady states, we may set $B/c+F=0$, which says that the Bernoulli function B is some function of potential vorticity Q. This fact can be derived directly from the steady-state version of (2.1a). For non-steady states, we have

$$\delta_t \pi_1-Q\delta_t \phi_2=-\frac{\partial}{\partial x_2}\left(\frac{B}{c}+F\right)\mathrm{d}x_0, \tag{2.35a}$$

$$\delta_t \pi_2+Q\delta_t \phi_1=\frac{\partial}{\partial x_1}\left(\frac{B}{c}+F\right)\mathrm{d}x_0, \tag{2.35b}$$

where

$$\delta_t(\cdot)=\frac{\partial(\cdot)}{\partial x_0}\mathrm{d}x_0.$$

Equations (2.35a, b) apply to our trajectory through the equilibrium point O. In the limit, at O,

$$\delta_t \pi_1-Q\delta_t\phi_2=0,\quad \delta_t\pi_2+Q\delta_t\phi_1=0. \tag{2.36a, b}$$

But $\delta_t\pi_i$, $\delta_t\phi_i$, $i=1,2$, can be considered as arbitrary variations $\delta\pi_i$, $\delta\phi_i$ constrained by

$$\delta\pi_1-Q\delta\phi_2=0,\quad \delta\pi_2+Q\delta\phi_1=0. \tag{2.37a, b}$$

Equations (2.21) indicate that the first variation of the total energy, δK_A $(=-\delta K_B)$, vanishes for such variations. The directions (2.37) are also the ones along which $\delta\chi$ in (2.29) vanishes, so that $\delta^2 K_B^{(2)}$ is zero in these directions as well. Indeed, all the terms in the integrand of $\delta^2 K_B^{(2)}$, (2.32), contain either the factor $\delta\chi$ or its divergence, $\nabla\cdot\delta\chi$. But $\delta\chi=\delta\chi\,(x_1,x_2;t)$ is a field quantity, and hence $\nabla\cdot\delta\chi\equiv 0$ whenever $\delta\chi=0$, identity being understood with respect to the coordinates x_1 and x_2.

Remark 2.4. Equations (2.36) hold trivially at the equilibrium point, since $\delta_t\pi_i=0=\delta_t\phi_i$, at steady state, for $i=1$, 2 separately. But away from O, they do not hold in any full neighbourhood of the steady state. They define, rather, the unstable or (asymptotically) stable directions in which we are interested.

The existence of the directions given by (2.36) depends on the stability of a given state. If such (asymptotically stable and) unstable directions really exist, then not only $\delta K_A(=-\delta K_B)$ and $\delta^2 K_B^{(2)}$ but also $\delta^2 K_A+\delta^2 K_B^{(1)}$ must vanish in these directions. Indeed, $K(=K_A+K_B)$ is an integral invariant of the motion, and thus $\delta^2 K=0$ along a trajectory through O. Thus we have proved Statement B, and its contraposition, which we restate as:

PROPOSITION 2.2. *If $\delta^2 K_A + \delta^2 K_B^{(1)}$ is either positive or negative definite at a given equilibrium point, then that point is linearly stable.*

From (2.26) and (2.31), we define the gauge-independent part $\delta^2 K^*$ of the stability norm $\delta^2 K$,

$$\delta^2 K^* \equiv \delta^2 K_A + \delta^2 K_B^{(1)} = \iint_\Omega \mathrm{d}x_1\,\mathrm{d}x_2 \Big[\tfrac{1}{2} g\delta h^2 + 2\delta h(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2) + \tfrac{1}{2} h(\delta\pi_1^2 + \delta\pi_2^2) - \frac{h}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\phi_0}(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2)^2\Big]. \quad (2.38)$$

It is easy to see that the integrand $\delta^2\mathscr{K}^*$ in (2.38) reduces to

$$\delta^2\mathscr{K}^* = \tfrac{1}{2} h(\delta\pi_1^2 + \delta\pi_2^2) - \frac{h}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\phi_0}(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2)^2 \quad (2.39)$$

for $\delta h = 0$. For positive definiteness of $\delta^2\mathscr{K}^*$, it is sufficient in this case that

$$\frac{\mathrm{d}Q}{\mathrm{d}(-\phi_0)} > 0. \quad (2.40)$$

From (B 10) in Appendix B it follows that, at steady state, $-\phi_0 = \Psi$, where Ψ is the momentum stream function.

For $\delta h \neq 0$, the integrand can be rewritten as

$$\delta^2\mathscr{K}^* = \tfrac{1}{4} g\left(\delta h + \frac{4}{g}\pi_1\,\delta\pi_1\right)^2 + \tfrac{1}{4} g\left(\delta h + \frac{4}{g}\pi_2\,\delta\pi_2\right)^2 + \tfrac{1}{2} h\left[\left(1 - \frac{8\pi_1^2}{gh}\right)\delta\pi_1^2 + \left(1 - \frac{8\pi_2^2}{gh}\right)\delta\pi_2^2\right] + \frac{h}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\Psi}(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2)^2,$$

and we obtain the sufficient stability criteria

$$1 - \frac{8(\pi_1^2 + \pi_2^2)}{gh} > 0 \quad \text{and} \quad \frac{\mathrm{d}Q}{\mathrm{d}\Psi} > 0. \quad (2.41\,a, b)$$

Except for the value of the constant coefficient of the Froude number $(\pi_1^2 + \pi_2^2)/gh$ in (2.41 a), conditions (2.40) and (2.41) correspond to the (formal) stability criteria of Arnol'd and of Holm and colleagues, respectively. Further use of the identity

$$-(\pi_1\,\delta\pi_1 + \pi_2\,\delta\pi_2)^2 = (\pi_1\,\delta\pi_1 - \pi_2\,\delta\pi_2)^2 - 2(\pi_1^2\,\delta\pi_1^2 + \pi_2^2\,\delta\pi_2^2)$$

leads to

$$\delta^2\mathscr{K}^* = \tfrac{1}{4} g\left(\delta h + \frac{4}{g}\pi_1\,\delta\pi_1\right)^2 + \tfrac{1}{4} g\left(\delta h + \frac{4}{g}\pi_2\,\delta\pi_2\right)^2 - \frac{h}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\Psi}(\pi_1\,\delta\pi_1 - \pi_2\,\delta\pi_2)^2 + \tfrac{1}{2} h\left[\left(1 - \frac{8\pi_1^2}{gh} + \frac{4}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\Psi}\pi_1^2\right)\delta\pi_1^2 + \left(1 - \frac{8\pi_2^2}{gh} + \frac{4}{Q^2}\frac{\mathrm{d}Q}{\mathrm{d}\Psi}\pi_2^2\right)\delta\pi_2^2\right]. \quad (2.42)$$

Equation (2.42) says that $\delta^2\mathscr{K}^*$ still remains positive definite for negative $\mathrm{d}Q/\mathrm{d}\Psi$ provided that

$$0 > \frac{\mathrm{d}Q}{\mathrm{d}\Psi} > \frac{-Q^2}{4(\pi_1^2 + \pi_2^2)}\left\{1 - \frac{8(\pi_1^2 + \pi_2^2)}{gh}\right\}. \quad (2.43)$$

3. Inclusion of rotation

No conceptual difficulty arises when we take the effect of the Earth's solid rotation into account. The equations of motion in a rotating reference frame follow from the principle of least action by formally adding the R-term below to the Lagrangian density $\mathscr{L}$ (Abarbanel 1985). The modified Lagrangian density $\mathscr{L}_R$ is

$$\mathscr{L}_R = \mathscr{L}(\phi_0, \phi_1, \phi_2) + R_1\left(\frac{\partial \phi_0}{\partial x_2} - \frac{\partial \phi_2}{\partial x_0}\right) + R_2\left(\frac{\partial \phi_1}{\partial x_0} - \frac{\partial \phi_0}{\partial x_1}\right). \tag{3.1}$$

It is understood here that $\vec{\phi} = (\phi_0, \phi_1, \phi_2)$ gives the momentum field $\vec{M} = (hc, hv_1, hv_2)$ observed in the corotating reference frame and that the constant two-dimensional vector $R = (R_1, R_2)$ satisfies

$$\frac{\partial R_2}{\partial x_1} - \frac{\partial R_1}{\partial x_2} = f, \tag{3.2}$$

where $f = f(x_2)$ is the Coriolis parameter. Without loss of generality we may assume that $R_1 = R(x_2)$ and $R_2 = 0$. Repeating the procedure used for an inertial reference frame, we obtain a formula analogous to the conditions (2.41 a, b) and (2.43). That is

$$1 - \frac{8(\pi_1^2 + \pi_2^2)}{gh} > 0, \quad \frac{dQ_R}{d\Psi} > 0 \tag{3.3}$$

and

$$0 > \frac{dQ_R}{d\Psi} > \frac{-Q_R^2}{4(\pi_1^2 + \pi_2^2)}\left(1 - \frac{8(\pi_1^2 + \pi_2^2)}{gh}\right), \tag{3.4}$$

where the absolute potential vorticity

$$Q_R = \frac{1}{h}\left(f(x_2) + \frac{\partial \pi_1}{\partial x_1} + \frac{\partial \pi_2}{\partial x_2}\right) \tag{3.5}$$

replaces Q in (2.7). The first criterion, (3.3), is an extended Arnol'd–Blumen criterion for rotating flows in the shallow-water equations. The second criterion (3.4) is new, being obtained by the present formulation.

4. Application to zonal shear flows

As the simplest application of the criteria (3.3), (3.4), let us consider a zonal, i.e. eastward flow in geostrophic balance with constant shear ζ_0:

$$v_1 = U(y) = \zeta_0 y, \quad v_2 = 0, \quad fU = -g\,dh/dy; \tag{4.1 a–c}$$

here $x_1 = x$ is the zonal and $y = x_2$ the meridional coordinate. For this simple flow profile, $dQ_R/d\Psi$ becomes

$$\frac{dQ_R}{d\Psi} = \frac{dQ_R}{dy}\frac{dy}{d\Psi} = -\frac{1}{h^2 U}\left(\frac{df}{dy} + \frac{fU}{gh}\left(f - \frac{dU}{dy}\right)\right). \tag{4.2}$$

For synoptic-scale flows, the magnitude of relative vorticity $|-dU/dy|$ is one order smaller than that of planetary vorticity f, which implies that their Rossby number is of the order of 0.1. So the presence of the planetary vorticity gradient $df/dy > 0$ means that the stability criterion (3.3) of Arnol'd–Blumen type ($dQ_R/d\Psi > 0$) is always violated by westerlies, $U > 0$. This is somewhat counterintuitive since such flows are linearly stable by the Rayleigh–Kuo inflection-point theorem (Lin 1967, Sec. 4.3; Kuo 1973).

For the shallow-water equations with zonal symmetry, Ripa (1983) derived a criterion which does cover the stability of (4.1). His derivation depends, however, crucially on zonal symmetry, while ours does not. Before proceeding to non-zonal, coherent eddies in the next section, it is still of interest to check whether our criterion also works for the simple case of westerlies with constant shear.

Remark 4.1. For an inertial system, the Rayleigh inflection-point theorem can be derived from the Arnol'd criterion by Galilean invariance of the momentum equations. However, in the case at hand, neither the corotating reference frame nor a reference frame moving with a constant velocity relative to it is an inertial system. Galilean invariance means that the equations of motion keep their form, including the value of the coefficients, under translation with arbitrary, but constant, velocity (see also Appendix A). Such a translation in the case (3.1) will change the value of f and of $U(y)$, breaking the invariance (see also Ghil & Childress 1987, p. 41).

Using (4.1c), (4.2) can be rewritten as

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi}=\frac{f}{h}\mathrm{d}\left(\frac{1}{h}\left(f-\frac{\mathrm{d}U}{\mathrm{d}y}\right)\right)\Big/\mathrm{d}(gh). \tag{4.3}$$

It is well known that, for synoptic-scale flows in the Northern Hemisphere, low pressure (small h) is associated with counterclockwise flow (large potential vorticity) and high pressure (large h) accompanies clockwise flow (small potential vorticity) so that in general $\mathrm{d}Q_R/\mathrm{d}\Psi$ is negative for such flows. The situation remains the same for purely zonal geostrophic flows, since planetary vorticity increases to the north while the (scale) height h increases to the south, owing to differential heating by the Sun.

For synoptic-scale flows in the atmosphere, (4.2) can be numerically estimated as

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi}\approx-\frac{1}{h^2U}\left(\frac{\mathrm{d}f}{\mathrm{d}y}+\frac{f^2U}{gh}\right)\approx-10^{-20}\ \mathrm{m^{-5}\,s}<0,$$

where we used $f\approx 10^{-4}\ \mathrm{s^{-1}}$, $\mathrm{d}f/\mathrm{d}y\approx 1.6\times10^{-11}\ \mathrm{m^{-1}\,s^{-1}}$, $U\approx 15\ \mathrm{m\,s^{-1}}$ and an atmospheric scale height approximately equal to 10^4 m for h. On the other hand,

$$\frac{-Q_R^2}{4(\pi_1^2+\pi_2^2)}\left(1-\frac{8(\pi_1^2+\pi_2^2)}{gh}\right)\approx\frac{-Q_R^2}{4(\pi_1^2+\pi_2^2)}\approx-10^{-19}\ \mathrm{m^{-5}\,s}<\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi}.$$

Hence a linearly stable geostrophic flow (4.1) satisfies our newly derived criterion (3.4), though it violates the classical criterion (3.3).

Actually the stability of this constant-shear flow can be verified in a little more general way. Let us assume that Rossby and Froude numbers, R_o and F_r, of a given flow,

$$R_o\equiv U/fL,\quad F_r\equiv U^2/gH,$$

are of the order $R_o<0.1$, $F_r\approx 0.01$; here H, L and U are scale height, horizontal and velocity scales characteristic of the given steady flow. Then we have

$$\begin{aligned}\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi}+\frac{Q_R^2}{4(\pi_1^2+\pi_2^2)}&\approx-\frac{1}{h^2U}\left(\frac{\mathrm{d}f}{\mathrm{d}y}+\frac{f^2U}{gh}\right)+\frac{f^2}{4U^2h^2}\\&=\frac{1}{h^2}\left\{-\frac{1}{U}\frac{\mathrm{d}f}{\mathrm{d}y}+\frac{f^2}{4U^2}\left(1-\frac{4U^2}{gh}\right)\right\}\\&\approx\frac{f^2}{4h^2U^2}\left(1-\frac{4U}{f}\frac{\mathrm{d}\ln f}{\mathrm{d}y}\right),\end{aligned} \tag{4.4}$$

where the last step used the smallness of F_r.

An elementary scale analysis shows that

$$1-\frac{4U}{f}\frac{\mathrm{d}\ln f}{\mathrm{d}y} \approx 1-4R_o L\frac{\mathrm{d}\ln f}{\mathrm{d}y} \approx 1-4R_o > 0. \tag{4.5}$$

So we get

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi} > -\frac{Q_R^2}{4(\pi_1^2+\pi_2^2)},$$

which satisfies the criterion (3.4).

5. Application to Stern's modon

Stern's modon is a non-divergent, piecewise-differentiable solution of the vorticity equation derived from (2.1), in the presence of differential rotation; for this modon

$$f(y)+\Delta\psi = -\lambda\psi \tag{5.1}$$

inside a circle of radius R, while outside of this circle the fluid is at rest, $\psi \equiv 0$. Here ψ is the usual, velocity stream function and R and λ satisfy the relation (Stern 1975):

$$R^2\lambda \approx 26.4. \tag{5.2}$$

For the interior domain of the modon, $\Omega_1 \equiv \{r \leqslant R\}$, where $r^2 = x^2+y^2$,

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi} = \frac{\mathrm{d}Q_R}{\mathrm{d}\psi}\frac{\mathrm{d}\psi}{\mathrm{d}\Psi} = \frac{\mathrm{d}}{\mathrm{d}\psi}\left(\frac{-\lambda\psi}{h}\right)\frac{\mathrm{d}\psi}{\mathrm{d}\Psi}. \tag{5.3}$$

For steady states, both the velocity stream function ψ and the momentum stream function Ψ are invariant along streamlines. Comparison with (2.3b,c) shows that

$$\frac{\mathrm{d}\Psi}{\mathrm{d}\psi} = h. \tag{5.4}$$

Substitution of (5.4) into (5.3) gives

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi} = -\frac{\lambda}{h^2}\left(1-\frac{\psi}{h}\frac{\mathrm{d}h}{\mathrm{d}\psi}\right). \tag{5.5}$$

Use of the geostrophic approximation $\mathrm{d}h/\mathrm{d}\psi \approx f/g$ yields a simpler form of (5.5):

$$\frac{\mathrm{d}Q_R}{\mathrm{d}\Psi} \approx \frac{-\lambda}{h^2}\left(1-\frac{f\psi}{gh}\right). \tag{5.6}$$

Now let L and U be characteristic length and velocity scales of the modon. Since $f\psi = Lf\psi/L \approx U^2/R_o$, we see that $\mathrm{d}Q_R/\mathrm{d}\Psi < 0$ if $F_r < R_o$; this is usually the case for rotation-dominated flows encountered in the geophysical fluid environment. The classical criterion (3.3) is again violated by such flows.

However the new criterion (3.4) permits a negative $\mathrm{d}Q_R/\mathrm{d}\Psi$: it states that

$$0 > \frac{\mathrm{d}Q_R}{\mathrm{d}\Psi} > \frac{-Q_R^2}{4(\pi_1^2+\pi_2^2)} \tag{5.7}$$

provided that $F_r \approx (\pi_1^2+\pi_2^2)/gh$ is negligibly small; $F_r \ll 1$ has to be assumed in order to apply (3.4), derived for (2.1), to non-divergent flows (see Sakuma 1989, Ch. 5, for details). Furthermore, if $F_r \ll R_o$, then

$$0 < \frac{\lambda}{h^2} < \frac{(f+\Delta\psi)^2}{4(\pi_1^2+\pi_2^2)\,h^2} \approx \frac{f^2}{4(\pi_1^2+\pi_2^2)\,h^2}. \tag{5.8}$$

So we obtain with round-off

$$0 < \lambda = \frac{26.4}{R^2} < \left(\frac{f}{2V}\right)^2, \tag{5.9}$$

where $V \equiv \mathrm{Max}\,[(\pi_1^2+\pi_2^2)^{\frac{1}{2}}]$. This can be further simplified as

$$\frac{V}{Rf} < 0.1. \tag{5.10}$$

Hence Stern's modon in Ω_i with a rigid free-slip boundary at $r = R$ is stable to infinitesimal perturbations of arbitrary form, provided its Rossby number R_o is less than 0.1.

Notice that the second variation $\delta^2 K$ in (2.26), (2.28) does not include boundary terms, and that the perturbations need not vanish along the boundary: the natural boundary condition for the variational problem (2.13) under consideration is simply that $\delta v_n = 0$ along $\partial\Omega_i \equiv \{r = R\}$.

For the complete flow field of a stationary modon floating in a quiescent background, an apparent difficulty arises since our stability analysis cannot be applied directly to the entire domain, as the absolute vorticity field $f + \Delta\psi$ has a discontinuity at $r = R$. Indeed our procedure of getting a new stability criterion depends crucially upon the use of (2.36) and (2.37), which cannot be satisfied at the interface $r = R$ since Q_R is discontinuous there. Yet the velocity at the interface is continuous and vanishes there, so it is natural to assume that velocity perturbations $\delta\pi$, as well as mass perturbations δh, are also continuous there.

The proof proceeds by decomposing the stability norm $\delta^2 K$ into two parts, interior and exterior,

$$\delta^2 K \equiv \iint_{\Omega_i} \mathrm{d}x\,\mathrm{d}y\,\delta^2\mathscr{K} + \iint_{\Omega_o} \mathrm{d}x\,\mathrm{d}y\,\delta^2\mathscr{K} \equiv \delta^2 K_i + \delta^2 K_o, \tag{5.11}$$

where $\Omega_o \equiv \{r > R\}$, and showing that the exterior term is positive definite, while the interior term satisfies the proof of Statement B in §2.3. Indeed, the gauge-independent part of $\delta^2 K_i$ is positive definite for arbitrary perturbations, not vanishing along the interface $\partial\Omega_i$, provided (5.10) holds. Contrary to the previous case, where $\delta v_n = 0$ had to be imposed on $\partial\Omega_i$, $\delta\boldsymbol{\pi}$ here can be quite arbitrary for $r = R$, provided it is continuous there, since (2.26), (2.28) for $\delta^2 K_A$ and $\delta^2 K_B$ contain no integrals along $\partial\Omega_i$.

The apparent difficulty with the discontinuity of Q_R at $r = R$ can now be removed by making two observations, which are both specific to the present case of quiescent flow in the exterior domain Ω_o. The first observation is that, since $\delta^2\mathscr{K}_B \equiv 0$ for $\boldsymbol{\pi} \equiv \mathbf{0} \equiv \boldsymbol{v}$, cf. (2.16), (2.28), both $\delta^2\mathscr{K}_A$, the total energy density, and $\delta^2\mathscr{K}_B$ are continuous across $\partial\Omega_i$.

To make the second observation, $\delta^2 K_o$ is decomposed further into its two parts, the energy and the Casimir terms,

$$\delta^2 K_o = \delta^2 K_A^{(o)} + \delta^2 K_B^{(o)}. \tag{5.12a}$$

The Casimir vanishes identically in Ω_o, $\delta^2 K_B^{(o)} \equiv 0$, since $\boldsymbol{v} \equiv \mathbf{0}$ there. Therewith, the second observation is that the use of the argument involving (2.37) to prove Statement B and hence Proposition 2.2 becomes superfluous, and no continuity of Q_R across the interface is required. This leaves the energy term,

$$\delta^2 K_A^{(o)} \equiv \iint_{\Omega_o} \mathrm{d}x\,\mathrm{d}y\,(\tfrac{1}{2}g\,\delta h^2 + \tfrac{1}{2}h\,\delta|\boldsymbol{v}|^2), \tag{5.12b}$$

which is also positive definite for arbitrary perturbations.

Hence (5.10) is a sufficient condition for the formal stability of Stern's modon in a quiescent background flow. The unrestricted character of the perturbations along the rim $\{r = R\}$ indicates that the modon is stable to small distortions of the circular boundary of the active part of the flow. But the present analysis does not provide any upper bound on the finite amplitude of such distortions for which the modon is still stable.

6. Concluding remarks

By using a gauge-variable formulation of the shallow-water equations, we derived a weaker formal stability condition than the classical ones (§2), and extended it to rotation dominated, geophysical flows (§3). Our new criterion does not require stable planetary flows with constant shear to be pseudo-westward (§4) and it permits one to show that Stern's stationary modon is stable to perturbations with arbitrary scale but with infinitesimally small amplitude (§5).

The success in deriving a meteorologically and oceanographically useful stability criterion is due to the new representation of the stability norm, rather than to changing the basic idea of Arnol'd's method. The Hamiltonian K, (2.19), consists of two terms: one corresponds to total energy, the other is also a conservative quantity, related to vorticity, called a Casimir function. This two-term structure of the integral invariant K is similar to that of the conservative functional H introduced by Arnol'd (1965). Therefore, in phase space, the new criterion has a geometrical interpretation similar to Arnol'd's. The weakening of the stability criterion is merely due to a sharper numerical evaluation of the relative contributions of the two terms. This also leads to a stability criterion free from the symmetry restriction referred to in §1 and Appendix A.

The approach presented here is extended in Sakuma (1989) to drifting modons by modifying the Lagrangian density $\mathscr{L}$: for drifting modons no problem of matching conditions arises, since their absolute vorticity Q is continuous everywhere (although its derivative is not). The possibility of obtaining a stability criterion for three-dimensional baroclinic flows is also being investigated along similar lines.

It is a pleasure to acknowledge stimulating discussions or correspondence with H. Abarbanel, A. Arakawa, D. Holm, R. Littlejohn, P. Malanotte Rizzoli, J. Marsden, D. Neelin, S. Putterman, P. Ripa, P. Roberts, J. Tribbia, S. Venkateswaran and G. Wolansky. We are also grateful to G. Carnevale for providing us with a preprint of Carnevale *et al.* (1988). Comments from D. G. Andrews and two anonymous referees were very valuable in improving both the content and the style of the manuscript. One of the referees disagrees with our treatment of matching conditions for Stern's modon and with our interpretation of the theorem in Appendix A. This work was supported by NASA grant NAG 5-713, NSF grant ATM 86-15424 and ONR grant N00014-87-K-0331. B. Gola turned the usual mess into a first-rate typescript.

Appendix A. Symmetry and Arnol'd stability

The purpose of this Appendix is to extend a result of Andrews (1984) on Arnol'd-stable solutions sharing the symmetry of the underlying problem. Further generalization is provided by Chern & Marsden (1990).

Symmetry of a problem is mathematically equivalent to its invariance under a group of transformations that maps the domain on which the problem is defined into itself. By problem we mean an initial-boundary-value problem for a system of partial differential equations. Symmetry then requires both the coefficients of the differential operators and the initial and boundary conditions to be invariant under the group of transformations.

The simplest case we use for illustration is invariance under x-translation. Consider the two-dimensional vorticity equation

$$\Delta\psi_t + J(\psi, \Delta\psi) = 0, \tag{A 1a}$$

defined on the entire plane, or in a channel periodic in the x-direction and with solid walls in the y-direction

$$\psi_x = 0 \quad \text{at} \quad y = \pm 1; \tag{A 1b,c}$$

here we use, for brevity, subscript notation for partial derivatives. Clearly the problem (A 1) is symmetric as defined above. Stationary solutions of (A 1) satisfy

$$\Delta\psi = F(\psi), \tag{A 2}$$

and the Arnol'd stability criterion is (compare (1.1) in the main text) that

$$F' \equiv \mathrm{d}F/\mathrm{d}\psi > 0. \tag{A 3}$$

Differentiating (A 2) with respect to x, and multiplying by ψ_x, yields

$$\psi_x \Delta\psi_x = \psi_x^2 F'. \tag{A 4}$$

Integrating (A 4) by parts in both x and y, we get

$$-\iint (\nabla\psi_x)^2 \,\mathrm{d}x\,\mathrm{d}y = \iint \psi_x^2 F' \,\mathrm{d}x\,\mathrm{d}y. \tag{A 5}$$

The right-hand side of (A 5) is non-negative for an Arnol'd-stable solution, and it is positive unless

$$v \equiv \psi_x \equiv 0. \tag{A 6}$$

The left-hand side is non-positive, hence (A 6) follows from (A 3) and the translational invariance.

The argument of Andrews (1984) was brought to the attention of one of us (M.G.) by R. T. Pierrehumbert (personal communication 1983) in the simpler form (A 4)–(A 6). The following theorem is an obvious extension of this argument.

THEOREM A. *For a two-dimensional incompressible flow in an elliptic geometry, governed by a problem with given symmetry, an Arnol'd-stable solution must have the same symmetry as the problem.*

Remark 1. Symmetry permits the use of the eigenfunctions of the associated group of transformations as a basis for solutions. Thus x-translations lead to Fourier expansion, as an integral in the unbounded case and as a series in the periodic case. Likewise cylindrical or spherical symmetry lead to the use of Fourier–Bessel functions or spherical harmonics (Tribbia 1984), respectively. Generally speaking, this implies that the standard Arnol'd criterion, owing to its very simplicity, only permits ascertaining the Lyapunov stability of flows for which linear stability can be studied by the classical methods of separation of variables. An example of this in the cylindrical-symmetry case is the work of Wan & Pulvirenti (1985).

Remark 2. The elliptic geometry of the problem is only restricted by the existence of an arbitrary symmetry group. Elliptic means that the Riemannian metric tensor is positive definite. This assumption is necessary, as will become clear at the end of the proof, (A 19).

Proof. This consists simply in reiterating (A 4)–(A 6) in an abstract setting. Let $\mathscr{M}$ be the underlying two-dimensional manifold, $\boldsymbol{x} \in \mathscr{M} \subset \mathscr{R}^3$, on which the problem is defined. $\mathscr{S}$ is the symmetry group, $S_t \in \mathscr{S}$, such that $S_t \mathscr{M} \subset \mathscr{M}$, or $S_t \boldsymbol{x} \in \mathscr{M}$ for all t, and

$$S_t S_s \boldsymbol{x} = S_{t+s} \boldsymbol{x} \quad \text{for all } t \quad \text{and } s. \tag{A 7}$$

A function (or functional, or operator on functions) on $\mathscr{M}, f: \mathscr{M} \rightarrow \mathscr{M}$, $\boldsymbol{y} = f(\boldsymbol{x}) \in \mathscr{M}$ if $\boldsymbol{x} \in \mathscr{M}$, is invariant under $\mathscr{S}$ if

$$f(S_t \boldsymbol{x}) = f(\boldsymbol{x}) \quad \text{for all } t. \tag{A 8}$$

$\mathscr{S}$ is a continuous group if $t \in \mathscr{R}$ (the unbounded case above, of translation by an arbitrary x-amount) and is discrete if $t \in \mathscr{N}$ (the periodic case, of translation by multiples of a fixed length in the x-direction). One deals accordingly with integrals or sums over t-values.

Functions which satisfy (A 8) are constant along each trajectory of the group. Let for instance $\mathscr{M} = \mathscr{R}^2$ be the (x, y)-plane and

$$S_t : x \rightarrow x + t, \quad t \in \mathscr{R}.$$

Then

$$f(S_t x, y) \equiv f(x+t, y) = f(x, y) \quad \text{for all } t$$

implies simply that $f(x, y) \equiv f_0(y)$. If t were restricted to the integers, $f(x+1, y) = f(x, y)$, f would be merely periodic in x, and could be expanded in a Fourier series in x.

Let $\Sigma_\mu \equiv (\mu^1, \mu^2)$ be a system of curvilinear coordinates on $\mathscr{M}$. In all cases of practical interest one can choose $\mu^1 = t$ without loss of generality. In Σ_μ the covariant form of the steady vorticity and mass-continuity equations is

$$v^j_{,j}(v_{2,1} - v_{1,2}) + v^j(v_{2,1} - v_{1,2})_{,j} = 0, \tag{A 9a}$$

$$(g^{\frac{1}{2}} v^j)_{,j} = 0. \tag{A 9b}$$

Here v^i and v_i are the contravariant and covariant velocity components, respectively, partial differentiation is indicated by commas, e.g. $v_{i,j} \equiv \partial v_i / \partial \mu^j$, and the summation convention is used for same-letter (upper and lower) indices; g is the determinant of the metric tensor g_{ij} (e.g. Landau & Lifshitz 1962). By definition, an elliptic geometry is one for which $g_{ij}\, \mathrm{d}\mu^i\, \mathrm{d}\mu^j$ is positive definite, and hence $g > 0$ under our assumptions. If $\mathscr{M}$ is compact without boundary like the spherical Earth (Tribbia 1984), no boundary conditions for (A $9a, b$) are necessary, otherwise they are assumed to be invariant under $\mathscr{S}$.

Equations (A $9a, b$) together yield

$$g^{\frac{1}{2}} v^j [(v_{2,1} - v_{1,2}) / g^{\frac{1}{2}}]_{,j} = 0, \tag{A 10a}$$

which is equivalent to

$$J(\psi^{(\mu)}, \zeta^{(\mu)}) = 0. \tag{A 10b}$$

Here

$$\zeta^{(\mu)} \equiv \frac{v_{2,1} - v_{1,2}}{g^{\frac{1}{2}}}, \quad g^{\frac{1}{2}} v^1 = -\psi^{(\mu)}_{,2}, \quad g^{\frac{1}{2}} v^2 = \psi^{(\mu)}_{,1}, \quad \zeta^{(\mu)} = \Delta_\mu \psi^{(\mu)}, \tag{A 11a–d}$$

where Δ_μ is the Laplace–Beltrami operator on $\mathscr{M}$. $\psi^{(\mu)}$ and $\zeta^{(\mu)}$ are the modified stream function and vorticity in Σ_μ. Equation (A $10b$) implies that, for steady states, $\zeta^{(\mu)}$ is a function F_μ of $\psi^{(\mu)}$:

$$\zeta^{(\mu)} = F_\mu(\psi^{(\mu)}). \tag{A 12}$$

The Arnol'd stability criterion, in the given geometry, is

$$F'_\mu \equiv \mathrm{d}F_\mu / \mathrm{d}\psi^{(\mu)} > 0. \tag{A 13}$$

In fact, a straightforward calculation of the coordinate transformation yields

$$\zeta^{(\mu)} = g^{\frac{1}{2}}\zeta, \quad \psi_{,i} = \psi^{(\mu)}_{,i}, \quad i = 1, 2. \tag{A 14a, b}$$

where ζ and ψ are the vorticity and stream function in Cartesian coordinates x^1, x^2. Owing to the symmetry, the partial derivative of g_{ij} with respect to $t = \mu^1$ vanishes. In the case of cylindrical symmetry, for instance, choosing $\theta = \mu^1$ yields $g_{11} = r^2$, $g_{22} = 1$ and $g_{12} = 0$, so that $g_{ij,\theta} = 0$.

Differentiation of (A 14*a*) with respect to μ^1 gives

$$F'_\mu \psi^{(\mu)}_{,1} = g^{\frac{1}{2}} F' \psi_{,1}. \tag{A 15a}$$

Substituting (A 14*b*) into (A 15*a*) yields

$$F'_\mu = g^{\frac{1}{2}} F'. \tag{A 15b}$$

Since $g^{\frac{1}{2}} > 0$, it follows that F' and F'_μ have the same sign.

Differentiation of (A 11*a*) with respect to μ^1 and the use of (A 12) gives

$$\zeta^{(\mu)}_{,1} = (v_{2,1,1} - v_{1,2,1})/g^{\frac{1}{2}} = F'_\mu \psi^{(\mu)}_{,1}.$$

Multiplying both sides by $\psi^{(\mu)}_{,1}$ and using (A 11*b, c*) yields

$$\psi^{(\mu)}_{,1}(v_{2,1,1} - v_{1,2,1})/g^{\frac{1}{2}} = v^2(v_{2,1,1} - v_{1,2,1}) = F'_\mu(\psi^{(\mu)}_{,1})^2. \tag{A 16}$$

After multiplying the left-hand side of (A 16) by $g^{\frac{1}{2}}$ and rewriting as

$$g^{\frac{1}{2}} v^2(v_{2,1,1} - v_{1,2,1}) = (g^{\frac{1}{2}} v^2 v_{2,1})_{,1} - (g^{\frac{1}{2}} v^2 v_{1,1})_2 - g^{\frac{1}{2}} v^2_{,1} v_{2,1} + g^{\frac{1}{2}} v^2_{,2} v_{1,1},$$

the use of (A 9*b*) and of appropriate boundary conditions in the μ^2-direction leads to

$$-\iint [(v^1_{,1} v_{1,1} + v^2_{,1} v_{2,1}] g^{\frac{1}{2}} \,\mathrm{d}\mu^1 \,\mathrm{d}\mu^2 = \iint F'_\mu (\psi^{(\mu)}_{,1})^2 g^{\frac{1}{2}} \,\mathrm{d}\mu^1 \,\mathrm{d}\mu^2. \tag{A 17}$$

By definition

$$v_1 = g_{1j} v^j, \quad v_2 = g_{2j} v^j. \tag{A 18a, b}$$

Substituting (A 18) into (A 17), we obtain

$$-\iint g_{ij} v^i_{,1} v^j_{,2} g^{\frac{1}{2}} \,\mathrm{d}\mu^1 \,\mathrm{d}\mu^2 = \iint F'_\mu (\psi^{\mu}_{,1})^2 g^{\frac{1}{2}} \,\mathrm{d}\mu^1 \,\mathrm{d}\mu^2. \tag{A 19}$$

By assumption, the quadratic form $g_{ij} v^i_{,1} v^j_{,2}$, which is equal to the square of the line element $\mathrm{d}s^2$ divided by the square of the time increment $\mathrm{d}t^2$, is positive definite and hence $g > 0$ as well.

The left- and right-hand sides of (A 19) can be reconciled, given Arnol'd stability (A 13), only if

$$\psi^{(\mu)}_{t} \equiv 0, \tag{A 20}$$

i.e. if the solution is constant along the trajectories of the symmetry group.

Q.E.D.

With respect to GFD applications, we notice that the highest symmetry on the sphere is given by rotations about an arbitrary axis. This yields $\psi \equiv 0$ on the sphere, i.e. a state of no motion. Rotation about an axis, i.e. a latitude-dependent Coriolis parameter $f = f(\theta)$, breaks this very high symmetry, leaving only the corresponding axisymmetry. Thus zonal flows can possess Arnol'd–Blumen stability, cf. Andrews (1984).

Finally, the Arnol'd criterion (A 3) for symmetric solutions implies nonlinear (Lyapunov) stability with respect to arbitrary perturbations, not just symmetric ones (compare Palais 1979).

Appendix B. Time independence of ϕ_i for steady states

In a steady state, local time changes of every observable quantity vanish, e.g. $\partial\pi_i/\partial x_0 = 0$ in such states. Whether $\partial\phi_i/\partial x_0 = 0$ holds or not in a steady state is less obvious since $\boldsymbol{\phi}$ is a potential which cannot be observed directly. It is consistent with the gauge condition (2.10) on $\vec{\boldsymbol{\phi}}$ to assume that ϕ_0 is a function of Q:

$$\phi_0 = \phi_0(Q). \tag{B 1}$$

From this we easily get

$$\frac{\partial\phi_0}{\partial x_0} = \frac{\mathrm{d}\phi_0}{\mathrm{d}Q}\frac{\partial Q}{\partial x_0} = 0, \tag{B 2}$$

since Q is an observable quantity. By differentiating (2.3b,c) with respect to time and using (B 2), we have

$$\frac{\partial^2\phi_i}{\partial x_0^2} = 0, \quad i = 1,2, \tag{B 3}$$

which yields

$$\frac{\partial\phi_i}{\partial x_0} = w_i(x_1,x_2), \quad i = 1,2. \tag{B 4}$$

Substitution of (B 4) into the steady-state continuity equation $\partial h/\partial x_0 = 0$ gives

$$\frac{\partial h}{\partial x_0} = \frac{1}{c}\frac{\partial}{\partial x_0}\left(\frac{\partial\phi_2}{\partial x_1} - \frac{\partial\phi_1}{\partial x_2}\right) = \frac{1}{c}\left(\frac{\partial w_2}{\partial x_1} - \frac{\partial w_1}{\partial x_2}\right) = 0.$$

Therefore $\boldsymbol{w}$ is a gradient,

$$w_i(x_1,x_2) = \frac{\partial W(x_1,x_2)}{\partial x_i}. \tag{B 5}$$

Applying the gauge condition (2.6) in the form (2.8) to steady states yields

$$\begin{aligned} hv_1\frac{\partial\phi_1}{\partial x_0} + hv_2\frac{\partial\phi_2}{\partial x_0} &= \left(\frac{\partial\phi_0}{\partial x_2} - \frac{\partial\phi_2}{\partial x_0}\right)\frac{\partial\phi_1}{\partial x_0} + \left(\frac{\partial\phi_1}{\partial x_0} - \frac{\partial\phi_0}{\partial x_1}\right)\frac{\partial\phi_2}{\partial x_0} \\ &= \frac{\partial\phi_0}{\partial x_2}\frac{\partial\phi_1}{\partial x_0} - \frac{\partial\phi_0}{\partial x_1}\frac{\partial\phi_2}{\partial x_0} = 0. \end{aligned} \tag{B 6}$$

Using (B 4) and (B 5), (B 6) becomes

$$\frac{\partial\phi_0}{\partial x_2}\frac{\partial W}{\partial x_1} - \frac{\partial\phi_0}{\partial x_1}\frac{\partial W}{\partial x_2} = \frac{\partial(W,\phi_0)}{\partial(x_1,x_2)} = 0.$$

Hence

$$W(x_1,x_2) = W(\phi_0). \tag{B 7}$$

From (2.3b,c) and (B 7), we get

$$hv_1 = \frac{\partial\phi_0}{\partial x_2} - \frac{\partial\phi_2}{\partial x_0} = \frac{\partial(\phi_0 - W(\phi_0))}{\partial x_2} = \frac{\partial\phi_0^{\dagger}(Q)}{\partial x_2}, \tag{B 8a}$$

$$hv_2 = \frac{\partial\phi_1}{\partial x_0} - \frac{\partial\phi_0}{\partial x_1} = -\frac{\partial(\phi_0 - W(\phi_0))}{\partial x_1} = -\frac{\partial\phi_0^{\dagger}(Q)}{\partial x_1}, \tag{B 8b}$$

where

$$\phi_0^{\dagger}(Q) \equiv \phi_0(Q) - W(\phi_0(Q)).$$

Equations (B $8a, b$) show that it is always possible to use $\phi_0^\dagger$ as ϕ_0 without violating the gauge condition (2.24). So for steady states,

$$\frac{\partial \phi_i^\dagger}{\partial x_0} = 0, \quad i = 1, 2, \tag{B 9}$$

as required for (2.30).

Thus, in particular, ($2.3b, c$) show that we can, at steady state, set

$$\phi_0 = -\Psi \tag{B 10}$$

where Ψ is the momentum stream function.

Appendix C. An illustrative example for the proof of Statement B

Let us consider one-dimensional motion of a point mass under the influence of a conservative force $k_1 x + k_2 x^3$, $k_1 > 0$, $k_2 > 0$. The equation of motion for and the Hamiltonian of this simple system are

$$m\ddot{x} = k_1 x + k_2 x^3, \tag{C 1}$$

$$H = \frac{P^2}{2m} - \tfrac{1}{2}k_1 x^2 - \tfrac{1}{4}k_2 x^4, \tag{C 2}$$

where the momentum is $p = m\dot{x}$.

The origin $O \equiv (x = 0, p = 0)$ of the phase space is an unstable equilibrium point. The unstable and (asymptotically) stable manifolds of O are given by the separatrices of H, i.e. by

$$H \equiv \frac{1}{2m}\{p - x[m(k_1 + \tfrac{1}{2}k_2 x^2)]^{\frac{1}{2}}\}\{p + x[m(k_1 + \tfrac{1}{2}k_2 x^2)]^{\frac{1}{2}}\} = 0.$$

The unstable and stable manifolds are

$$p \mp x[m(k_1 + \tfrac{1}{2}k_2 x^2)]^{\frac{1}{2}} = 0, \tag{C $3a, b$}$$

respectively. Differentiation of (C 3) with respect to time t and multiplication of both sides by $\mathrm{d}t$ yields

$$\mathrm{d}p = \pm\left\{[m(k_1 + \tfrac{1}{2}k_2 x^2)]^{\frac{1}{2}} + \frac{k_2 x^2}{2[m(k_1 + \tfrac{1}{2}k_2 x^2)]^{\frac{1}{2}}}\right\}\mathrm{d}x$$

In the limit $x \to 0$

$$\mathrm{d}p = \pm(mk_1)^{\frac{1}{2}}\,\mathrm{d}x. \tag{C $4a, b$}$$

Each sign corresponds to the equation of a straight line through the origin, even though $\dot{p} = \dot{x} = 0$ precisely at O. Equations (2.36) are the analogue of (C 4) for the conjugate momenta (ϕ_i, π_i), given by field variables $\phi_i(x_0, x_1, x_2)$, $\pi_i(x_0, x_1, x_2)$, $i = 1, 2$, rather than the scalars (x, p).

REFERENCES

Abarbanel, H. D. I. 1985 Hamiltonian description of almost geostrophic flow. *Geophys. Astrophys. Fluid Dyn.* **33**, 145–171.

Abarbanel, H. D. I, Holm, D. D., Marsden, J. E. & Ratiu, T. S. 1986 Nonlinear stability analysis of stratified fluid equilibria. *Phil. Trans. R. Soc. Lond.* A **318**, 349–409.

Andrews, D. G. 1984 On the existence of nonzonal flows satisfying sufficient conditions for stability. *Geophys. Astrophys. Fluid Dyn.* **28**, 143–256.

Arnol'd, V. I. 1965 Conditions for nonlinear stability of stationary plane curvilinear flows of an ideal fluid. *Sov. Maths Dokl.* **6**, 773–776.

Arnol'd, V. I. 1969 On an a priori estimate in the theory of hydrodynamical stability. *Am. Math. Soc. Trans.* Ser. 2, **79**, 267–269.

ARNOL'D, V. I. 1978 *Mathematical Methods of Classical Mechanics*. Springer.

BENZI, R., PIERINI, S., SALUSTI, E. & VULPIANI, A. 1982 On nonlinear hydrodynamic stability of planetary vortices. *Geophys. Astrophys. Fluid Dyn.* **20**, 293–306.

BLUMEN, W. 1968 On the stability of quasigeostrophic flow. *J. Atmos. Sci.* **25**, 929–931.

CARNEVALE, G. F., VALLIS, G. K., PURINI, R. & BRISCOLINI, M. 1988 The role of initial conditions in flow stability with an application to modons. *Phys. Fluids* **31**, 2567–2572.

CHERN, S.-J. & MARSDEN, J. E. 1990 A note on symmetry and stability for fluid flows. *Geophys. Astrophys. Fluid Dyn.* **51**, 19–26.

FLIERL, G. R. 1987 Isolated eddy models in geophysics. *Ann. Rev. Fluid Mech.* **19**, 493–530.

FLIERL, G. R., LARICHEV, V. D., MCWILLIAMS, J. C. & REZNIK, G. M. 1980 The dynamics of baroclinic and barotropic solitary eddies. *Dyn. Atmos. Oceans* **5**, 1–41.

GHIL, M. & CHILDRESS, S. 1987 *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics*. Springer.

GOLDSTEIN, H. 1980 *Classical Mechanics*, 2nd edn. Addison-Wesley.

HOLM, D. D., MARSDEN, J. E., RATIU, T. S. & WEINSTEIN, A. 1983 Nonlinear stability conditions and a priori estimates for barotropic hydrodynamics. *Phys. Lett.* **98**A, 15–21.

HOLM, D. D., MARSDEN, J. E., RATIU, T. S. & WEINSTEIN, A. 1985 Nonlinear stability of fluid and plasma equilibria. *Phys. Rep.* **123**, 1–116.

INGERSOLL, P. I. & CUONG, P. G. 1981 Numerical model of long-lived Jovian vortices. *J. Atmos. Sci.* **38**, 2067–2076.

KUO, H. L. 1973 Quasi-geostrophic flows and instability theory. *Adv. Appl. Mech.* **13**, 247–330.

LAEDKE, E. W. & SPATSCHEK, K. H. 1986 Two-dimensional drift vortices and their stability. *Phys. Fluids* **29**, 133–142.

LANDAU, L. D. & LIFSHITZ, E. M. 1962 *The Classical Theory of Fields*, 2nd edn. Pergamon.

LIN, C. C. 1967 *The Theory of Hydrodynamic Stability*. Cambridge University Press.

MALANOTTE-RIZZOLI, P. 1982 Planetary solitary waves in geophysical flows. *Adv. Geophys.* **24**, 147–224.

MAXWORTHY, T. & REDEKOPP, L. G. 1976 New theory of the Great Red Spot from solitary waves in the Jovian atmosphere. *Nature* **260**, 509–511.

MCWILLIAMS, J. C. 1980 An application of equivalent modons to atmospheric blocking. *Dyn. Atmos. Oceans* **5**, 43–66.

MCWILLIAMS, J. C., FLIERL, G. R., LARICHEV, V. D. & REZNIK, G. M. 1981 Numerical studies of barotropic modons. *Dyn. Atmos. Oceans* **5**, 219–238.

PALAIS, R. S. 1979 The principle of symmetric criticality. *Commun. Math Phys.* **69**, 19–30.

PIERINI, S. 1985 On the stability of equivalent modons. *Dyn. Atmos. Oceans* **9**, 273–280.

REDEKOPP, L. G. 1977 On the theory of solitary Rossby waves. *J. Fluid Mech.* **82**, 725–745.

RIPA, P. 1983 General stability conditions for zonal flows in a one-layer model on the β-plane or the sphere. *J. Fluid Mech.* **126**, 463–489.

SAKUMA, H. 1989 A new stability criterion for two-dimensional coherent eddies in the atmosphere and oceans. Ph.D. thesis, University of California, Los Angeles.

SALMON, R. 1988 Hamiltonian fluid mechanics. *Ann. Rev. Fluid Mech.* **20**, 225–256.

STERN, M. E. 1975 Minimal properties of planetary eddies. *J. Mar. Res.* **33**, 1–33.

SWATERS, G. E. 1986 Stability conditions and a priori estimates for equivalent barotropic modons. *Phys. Fluids* **29**, 1419–1422.

TRIBBIA, J. J. 1984 Modons in spherical geometry. *Geophys. Astrophys. Fluid Dyn.* **30**, 131–168.

WAN, Y. H. & PULVIRENTI, M. 1985 Nonlinear stability of circular vortex patches. *Commun. Math. Phys.* **99**, 435–450.

ZABUSKY, N. J. & KRUSKAL, M. D. 1965 Interaction of 'solitons' in a collisionless plasma and the recurrence of initial states. *Phys. Rev. Lett.* **15**, 240–243.