Phil. Trans. R. Soc. Lond. A **318**, 349–409 (1986) [349]
Printed in Great Britain

NONLINEAR STABILITY ANALYSIS OF STRATIFIED FLUID EQUILIBRIA

By H. D. I. ABARBANEL[1], D. D. HOLM[2], J. E. MARSDEN[3], and T. S. RATIU[4]

[1] *Scripps Institution of Oceanography and Institute for Nonlinear Science, University of California, San Diego, La Jolla, California* 92093, *U.S.A.*
[2] *Theoretical Division, MS B*284, *and Centre for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico* 87545, *U.S.A.*
[3] *Department of Mathematics, University of California, Berkeley, California* 94720, *U.S.A.*
[4] *Department of Mathematics, University of Arizona, Tucson, Arizona* 85721, *U.S.A.*

(*Communicated by T. B. Benjamin, F.R.S. – Received* 7 *January* 1985)

CONTENTS

Nonlinear stability is analysed for stationary solutions of incompressible inviscid stratified fluid flow in two and three dimensions. Both the Euler equations and their Boussinesq approximations are treated. The techniques used were initiated by Arnold around 1965. These techniques combine energy methods, conserved quantities and convexity estimates. The resulting nonlinear stability criteria involve standard quantities, such as the Richardson number, but they differ from the linearized stability criteria. For example, the full three-dimensional problem has nonlinearly stable stationary solutions with Richardson number greater than unity, provided the gradients of the variations in density satisfy explicitly given bounds. Specific examples and the associated Hamiltonian structures for the problems are given.

1. Introduction

Stability of a stratified, or heterogeneous, shear flow is a classical problem of hydrodynamics whose extensive history is reviewed, for example, in Drazin & Reid (1981). Our aim here is to provide sufficient conditions for nonlinear Liapunov stability of equilibria of ideal, stratified flows in both two and three dimensions, as described by either the Euler equations, or their Boussinesq approximation.

The Liapunov method for proving stability of an equilibrium point of a conservative dynamical system depends on finding a constant of the motion with a local maximum or minimum at the equilibrium point. Usually this constant is taken to be the energy, perhaps combined with some additional conserved quantity arising from a symmetry of the system. These energy methods are a common tool in the study of fluid and plasma stability problems (see, for example, Bernstein *et al.* 1958).

An important contribution to the applicability of the Liapunov method to fluid systems is made in Arnold (1965, 1969). This method, as we now view it, is summarized below and in Appendix A. The method depends on adding special conserved quantities to the energy. The reason the Euler equations for incompressible fluids of uniform density possess these additional conserved quantities is due to degeneracy of the Poisson bracket underlying the Hamiltonian formulation of these equations. In general, degeneracy of the Poisson bracket for a dynamical system means that there exist certain quantities – the so-called 'Casimirs' – which are constants of the motion for *any* Hamiltonian expressible in terms of the given dynamical variables (see, for example, Weinstein (1984) for an overview). Once determined from the Poisson structure, the Casimirs provide an important family of conserved quantities for implementing the Liapunov stability method. For fluid systems, the Casimirs are physically relevant quantities, such as (generalized) enstrophy, in the case of planar incompressible flow with uniform density. Degeneracy of the Poisson brackets for fluids and the associated existence of Casimirs in the Eulerian fluid representation is due to particle-exchange symmetries of the action principle for the fluid in the Lagrangian representation (see Seliger & Whitham (1968), Bretherton (1970), and Marsden & Weinstein (1983). Lagrangian relabelling transformations are also discussed in detail in §7 and Appendix B.).

Before turning to our main topics, we recall some terminology regarding stability. Consider an evolution equation

$$\mathrm{d}u/\mathrm{d}t = F(u)$$

for elements, u, in the domain of a nonlinear-operator F defined in a normed linear space. Equilibrium, or *stationary* solutions u_e satisfy $F(u_e) = 0$. Such a solution is called *Liapunov stable*, when any solution $u(t)$ beginning near u_e at $t = 0$ remains near u_e for all time. Formally, this means for every $\epsilon > 0$, there is a $\delta > 0$, such that when the norm $\|u(0) - u_e\| < \delta$, then $\|u(t) - u_e\| < \epsilon$ for $-\infty < t < \infty$.

Assuming F is differentiable (from its domain to the containing space), the *linearized equation* at u_e is given by the linear evolution equation

$$\mathrm{d}\delta u/\mathrm{d}t = \mathrm{D}F(u_e)\cdot\delta u,$$

which describes the evolution of an infinitesimal disturbance, δu, of the original equilibrium solution, u_e. If the solution of this linearized equation is (Liapunov) stable, then the equilibrium solution u_e for the original problem is said to be *linearized stable*, or *linearly stable*.

Most investigations of stability of fluid dynamics have concentrated on linear stability, i.e. on the evolution of infinitesimal perturbations of a given flow. The monographs of Chandrasekhar (1961) and Drazin & Reid (1981) contain excellent examples of methods for studying linear stability of a stationary solution. This is usually done by examining the spectrum of the linearized operator $DF(u_e)$. A necessary condition for linearized stability is that this spectrum lie on the imaginary axis. (This is also sufficient under some extra hypotheses such as the absence of multiple eigenvalues.). This situation is called *neutral stability*, or *spectral stability*. Drazin & Reid (1981) also mention (in §53.2) the method of Arnold (1965, 1969) for the study of nonlinear stability of fluid flows allowing finite deviations from a given stationary solution. Certain steps in that method involving second variation arguments have been employed by Blumen (1971), (other authors referred to in Drazin & Reid (1981), and Benzi *et al.* (1982). For some model equations of hydrodynamical type, nonlinear stability has been proven by the Liapunov method, as in the proof of stability of the single soliton solution of the KdV equation by Benjamin (1972) and Bona (1975), and in the Benjamin–Ono equation modelling internal solitary waves, by Bennett *et al.* (1983). Recently, Holm, Marsden *et al.* (1983, 1985, 1986) have used Arnold's contribution to the Liapunov stability method in a number of fluid and plasma problems. Convexity estimates are employed for the proof of nonlinear stability and the Casimirs for fluid dynamics play an essential role. As we shall explain, second variation arguments do not suffice to prove nonlinear stability.

The stability method carried out in specific cases in the body of the present paper proceeds as follows (see Appendix A for a detailed summary). For a particular arrangement, one adds to the energy functional, H, another conserved functional, C, depending upon the dynamical variables of the problem (these are the Casimirs in the present case, but any other conserved quantities should also be included), thereby obtaining a conserved quantity, $H+C = H_C$. The quantity C is selected so that the first variation of H_C vanishes at an equilibrium of interest. Definiteness of the second variation of the functional H_C at the equilibrium state is sufficient for linear stability and is a formal indication of nonlinear stability. If this definiteness of the second variation holds, we say our equilibrium is *formally stable*. Formal stability implies linearized Liapunov stability, but it need not imply nonlinear stability. Nonlinear stability is proved via certain convexity estimates, explained in Appendix A (see Arnold 1969; Holm, Marsden *et al.* 1983, 1985). Formal stability was studied earlier in plasma physics; see, for example, Kruskal & Oberman (1958), Newcomb (1960) and Rosenbluth (1964), and in meteorology, see Eliassen & Kleinschmidt (1957) and Fjörtoft (1946). However, in general, formal stability need not even imply that H_C has a local minimum or maximum at the equilibrium. For example, Ball & Marsden (1984) have constructed equilibria in nonlinear elasticity at which the second variation of the energy is positive definite, yet in spaces reasonable to the problem, the equilibrium is not even a local minimum of the energy. (The deformations which lower the energy do involve high wavenumbers, a situation compatible with the present treatment.) Therefore, the convexity estimates are an important step that is necessary to complete the argument for (nonlinear) stability.

In this paper, we employ the Arnold method to study stability of incompressible, inviscid flows of a stratified fluid in a gravitational field; i.e. the buoyancy force is $g\rho(\boldsymbol{x}, t)\,\hat{\boldsymbol{z}}$, where ρ is the density field, $\hat{\boldsymbol{z}}$ is the unit vector directed upward along the vertical axis, and g is the acceleration due to gravity. Buoyant flows are central to atmospheric and oceanic dynamics (see, for example, Pedlosky 1979 and Gill 1982), as well as being important in laboratory studies of the transition to turbulence, for example, in Rayleigh–Benard convection. We analyse both two- and three-dimensional situations and consider in each instance two descriptions: the Euler equations for heterogeneous incompressible flow, and the Boussinesq approximation to the Euler equations. Discussions of the Boussinesq approximation appear in Chandrasekhar (1961), and in oceanic terms, in Phillips (1977).

As we have remarked, one of the steps in the stability procedure is to choose the Casimir C such that H_C has a critical point at the equilibrium solution of interest. We shall see that this condition is closely related to the Long equation (Long 1953) for two-dimensional Boussinesq solutions. Using this procedure for the three-dimensional case and for the Euler equations, without the Boussinesq approximation, leads systematically to modifications and extensions of the Long equation. We note in passing that solutions of the Long equation depend sensitively on boundary conditions, for example, significant differences in behaviour occur for fixed, free, or radiative boundaries, as indicated by the work of Lilly & Klemp (1979), and Leonov *et al.* (1979).

The stability results we present for the Euler and Boussinesq descriptions differ from each other in detail, but they are qualitatively very similar. In each description, $\delta^2 H_C$, the second variation of H_C, is *indefinite* for sufficiently high wavenumbers of the density variation. This result contrasts with the 'Richardson number criterion', which states that the spectrum of the linearized problem for a two-dimensional stratified shear flow is purely imaginary if the Richardson number is greater than one-quarter; see Synge (1933), Chandrasekhar (1961), Miles (1961) and Howard (1961).

Keep in mind that $\delta^2 H_C$ can be indefinite and the equilibrium point be unstable, even when the linearized problem has a purely imaginary spectrum. This can occur already in finite dimensional problems due to nonlinear interactions via resonances and Arnold diffusion; compare with Lichtenberg & Lieberman (1982), Holmes & Marsden (1982, 1983) and Vivaldi (1984). Consistent with indefiniteness of the second variation we *conjecture that all stratified, incompressible, ideal fluid flows for sufficiently high wave numbers of the density variation are nonlinearly unstable.* As mentioned above, the indefiniteness of the second variation $\delta^2 H_C$ for such ideal fluids is due to density perturbations with unbounded wavenumber. Accordingly, the conjectured nonlinear instability is generated at high wave numbers. This means however, that lower modes may also become unstable via nonlinear coupling among the modes, unless some nonlinear saturation process (or dissipation) intervenes. This instability mechanism seems to be related to the one discussed by Davis & Acrivos (1967).

Thus, instability is indicated by indefiniteness of the second variation $\delta^2 H_C$, for density variations of sufficiently high wavenumbers or, equivalently, with sufficiently large gradients. Still, it is possible to estimate the range of *stable* wavenumbers of density variations, for which a certain equilibrium solution will be linearly, and even nonlinearly, Liapunov stable.

The main purpose of this paper is to show that *as long as the gradient of the density displacement away from equilibrium* $\nabla[\rho(\boldsymbol{x}, t) - \rho_e(\boldsymbol{x})]$ *is bounded in magnitude by the following inequality*

$$|\nabla[\rho(\boldsymbol{x}, t) - \rho_e(\boldsymbol{x})]|^2 < k_+^2[\rho(\boldsymbol{x}, t) - \rho_e(\boldsymbol{x})]^2, \tag{1.1}$$

where k_+^2 is explicitly determined by equilibrium state quantities, then there exists a set of stability criteria, under which the stratified, incompressible, ideal fluid flows will be Liapunov stable, even for finite-sized initial *values of the disturbance away from the the equilibrium state.* To show this property and to find the stability criteria explicitly, we establish nonlinear convexity estimates for H_C. These nonlinear estimates express Liapunov stability against disturbances of small, but finite amplitude for the Euler and Boussinesq description in both two and three dimensions, for as long as the solutions continue to exist and satisfy (1.1). As an illustration of the technique, we show that for a class of shear flows in three dimensions, stability holds if the Richardson number exceeds unity, i.e. $Ri > 1$. In this case, the quantity k_+^2 in (1.1) is approximately proportional to $(Ri-1)$. In two dimensions, a stratified Boussinesq shear flow is linearly stable in the norm given by $\delta^2 H_C$ in equation (2.40) of §2 if condition (2.50) is satisfied, and nonlinearly stable in the norm defined by (2.62) if conditions (2.59–2.60) hold.

For purposes of readability, we have not included precise differentiability hypotheses on the various functions entering our analysis, nor have we discussed questions of existence and uniqueness (see, for example, Marsden (1976) and Temam (1983)). However, the norms in which we express nonlinear stability criteria are given explicitly.

The plan of the paper follows. In §§2 and 3, we investigate first two-dimensional, then three-dimensional Boussinesq flows of inviscid, incompressible, stratified fluids. In §4, we discuss examples of the stability conditions for three-dimensional shear Boussinesq flows. §§5 and 6 are devoted to the Euler equations in two and three dimensions, respectively, without the Boussinesq approximation.

In each of §§2, 3, 5 and 6, we establish conditions for formal stability of equilibrium flows and then use convexity arguments to determine conditions for their Liapunov stability. In Appendix A we summarize the general steps one follows in proving stability by the energy-Casimir method. In §7 (and Appendix B) we present the Hamiltonian structures for the types of flows under discussion. The Poisson brackets and associated conservation laws underlying the method are described, as well as the derivation of these Poisson brackets by reduction from canonical Poisson brackets in the Lagrangian (material) picture, and their relationship to Clebsch representations. The Hamiltonian structures of the linearized equations are studied in Appendix C, where we show, that $\delta^2 H_C$, being the Hamiltonian for the linearized dynamics, is preserved, and, thus, when definite, provides the Liapunov functional needed for proving linearized stability.

The applicability of the inviscid Euler equations rather than the Navier–Stokes equations is limited to the study of flows whose wavenumbers (or spatial gradients) are sufficiently small (in practice, in regions away from boundaries and in the absence of discontinuities). Motivated by this consideration, we show in §8 that a modified fluid model which 'filters' the energy carried by the high wave numbers can be nonlinearly stable in two dimensions. The 'filtered' two-dimensional Boussinesq equations in §8 remove the high wavenumbers from the stream function ψ entirely and prevent their development, thereby enabling the modified equations to be nonlinear stable. Thus, §8 departs in philosophy from the rest of the paper, by modifying the starting equations. The rest of the paper treats the unmodified Boussinesq and Euler equations, and shows how to determine the nonlinearly stable range of wavenumbers in the density variation for a given initial equilibrium solution. One can consider 'monitoring' the development of disturbances from an initial equilibrium state, expecting nonlinear instability after the wavenumbers (or gradients) in density variation leave the stable range for that initial state as determined in the present work.

§9 briefly presents the conclusions of this work.

2. Two-dimensional Boussinesq flow

In a Boussinesq fluid, the density varies little, yet buoyancy drives the motion. Thus, the variation of density $\rho(\boldsymbol{x}, t)$ from a constant reference density ρ_* is negligible everywhere except in buoyancy effects. The Boussinesq equations describe the dynamics of an ideal, slightly stratified, and rotating fluid, moving with velocity $\boldsymbol{v}(\boldsymbol{x}, t)$ under the action of buoyancy, in combination with pressure forces and Coriolis forces. The Boussinesq equations are, in three dimensions,

$$\partial \boldsymbol{v}/\partial t = -(\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} - (1/\rho_*)\,\nabla p - (\rho g/\rho_*)\,\nabla z + \boldsymbol{v}\times\boldsymbol{f}, \tag{2.1}$$

$$\partial \rho/\partial t = -\boldsymbol{v}\cdot\nabla\rho, \tag{2.2}$$

$$\operatorname{div}\boldsymbol{v} = 0. \tag{2.3}$$

The additional notation is as follows: p is pressure (determined from the constraint $\operatorname{div}\boldsymbol{v} = 0$), g is gravitational acceleration along ∇z, and $\boldsymbol{f} = f(x, y)\,\nabla z$ is twice the frequency of rotation about the z-axis. Rewriting the Euler equation (2.1) by using the identity $-(\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} = \boldsymbol{v}\times\operatorname{curl}\boldsymbol{v} - \frac{1}{2}\nabla\,|\boldsymbol{v}|^2$ gives

$$\partial \boldsymbol{v}/\partial t = \boldsymbol{v}\times(\boldsymbol{\omega}+\boldsymbol{f}) - \nabla(p/\rho_* + \tfrac{1}{2}|\boldsymbol{v}|^2) - (\rho g/\rho_*)\,\nabla z, \tag{2.4}$$

where $\boldsymbol{\omega} = \operatorname{curl}\boldsymbol{v}$ is the vorticity. Taking the curl of (2.4) produces the Boussinesq vorticity equation,

$$\partial \boldsymbol{\omega}/\partial t = \operatorname{curl}\,[\boldsymbol{v}\times(\boldsymbol{\omega}+\boldsymbol{f})] + (g/\rho_*)\,\nabla z\times\nabla\rho. \tag{2.5}$$

In this section, we will consider two-dimensional motion in the (x, z) plane, with the z-axis oriented upward and the x-axis in the horizontal direction. As is customary, $\hat{\boldsymbol{x}}$ and $\hat{\boldsymbol{y}}$ will denote the unit vectors along the Ox and Oy axes. Since $\operatorname{div}\boldsymbol{v} = 0$ and the velocity is tangential on the boundary, one can show that there is a stream function $\psi(x, z, t)$, with $\boldsymbol{v} = \hat{\boldsymbol{y}}\times\nabla\psi = (\partial\psi/\partial z, -\partial\psi/\partial x)$ and $\boldsymbol{\omega} = \operatorname{curl}\boldsymbol{v} = \hat{\boldsymbol{y}}\omega(x, z, t)$. For this model, the Coriolis force is assumed to be zero, in order that the motion be strictly two-dimensional. The vorticity and stream function in a certain domain are related by $\omega = \nabla^2\psi$, subject to ψ being constant on each connected component of the boundary; i.e. $\nabla\psi\times\hat{\boldsymbol{n}} = \boldsymbol{v}\cdot\hat{\boldsymbol{n}} = 0$, with $\hat{\boldsymbol{n}}$ the unit normal vector on the boundary. In unbounded domains, we demand that ω and ψ vanish appropriately at infinity (precise function spaces for which the Laplacian ∇^2 is then invertible are known; see, for example, Cantor (1979)).

For two-dimensional motion, the Boussinesq equations reduce to

$$\partial\omega/\partial t = -\boldsymbol{v}\cdot\nabla\omega + \frac{g}{\rho_*}\frac{\partial\rho}{\partial x}, \tag{2.6}$$

$$\partial\rho/\partial t = -\boldsymbol{v}\cdot\nabla\rho, \tag{2.7}$$

or, in terms of the stream function ψ,

$$\partial\omega/\partial t = \{\omega, \psi\} + \{gz/\rho_*, \rho\}, \tag{2.8}$$

$$\partial\rho/\partial t = \{\rho, \psi\}, \tag{2.9}$$

with $\{,\}$ the Jacobian (or the canonical Poisson bracket) defined by

$$\{g, h\} = \frac{\partial g}{\partial z}\frac{\partial h}{\partial x} - \frac{\partial g}{\partial x}\frac{\partial h}{\partial z}, \tag{2.10}$$

for functions $g(x, z)$, $h(x, z)$. The sense in which (2.8) and (2.9) form a Hamiltonian system will be explained in §7 (see also Appendix C).

In view of (2.8), (2.9) and the properties of the Jacobian, certain functional dependences exist among equilibrium quantities, ψ_e, ρ_e, and ω_e. By the continuity equation (2.9), for stationary flows one has $\{\rho_e, \psi_e\} = 0$, so ψ_e and ρ_e are functionally related. Assume that

$$\psi_e = \bar{\psi}(\rho_e), \tag{2.11}$$

for a function $\bar{\psi}$. The use of this functional dependence in the vorticity equation (2.8) for stationary flows leads to

$$\left\{\omega_e + \frac{1}{\rho_*}\frac{d\rho_e}{d\psi_e} gz, \psi_e\right\} = 0, \tag{2.12}$$

so that the quantities

$$\left(\omega_e + \frac{1}{\rho_*}\frac{d\rho_e}{d\psi_e} gz\right)$$

and ψ_e are also functionally related. Assume that

$$\omega_e + \frac{1}{\rho_*}\frac{d\rho_e}{d\psi_e} gz = L(\psi_e), \tag{2.13}$$

for a function L (which we call the Long function). Equation (2.13) is called the Long equation for a planar Boussinesq fluid; compare with Dubreil-Jacotin (1935), Long (1953), Yih (1980).

As mentioned in the introduction, investigation of stability of stationary flows by the Arnold method relies on the construction of conserved quantities that are extremalized by the stationary flows. There are two types of conserved quantities for the two-dimensional Boussinesq equations

(i) the energy (normalized by ρ_*),

$$H(\boldsymbol{v}, \rho) = \int_D dx\, dz\, (\tfrac{1}{2}|\boldsymbol{v}|^2 + \rho gz/\rho_*), \tag{2.14}$$

which is verified to be conserved, using the boundary condition that $\boldsymbol{v}\cdot\hat{\boldsymbol{n}}$ vanishes on the fixed boundary ∂D of the domain D; and

(ii) a family of conserved quantities $C_{F,G}$ parametrized by two real-valued functions F and G of a single variable, given by

$$C_{F,G} = \int_D dx\, dz\, [\omega F(\rho) + G(\rho)], \tag{2.15}$$

where $F(\rho)$ is assumed to satisfy the condition

$$\int_D dx\, dz \frac{\partial \rho}{\partial x} F(\rho) = 0 \tag{2.16}$$

for all ρ.

The condition (2.16) holds, for example, if the domain D is periodic in x, or if ρ is constant on connected components of ∂D (isopychnal boundary conditions). The geometric meaning of $C_{F,G}$ is discussed in §7, in connection with the Hamiltonian structure of the Boussinesq equations.

Denote by $H_C(\boldsymbol{v}, \rho)$ the sum

$$H_C(\boldsymbol{v}, \rho) = H(\boldsymbol{v}, \rho) + C_{F,G}(\boldsymbol{v}, \rho) + \lambda \int_D dx\, dz\, \omega, \tag{2.17}$$

where λ is a constant. We will show that the quantity $H_C(\boldsymbol{v}, \rho)$ has an extremal point for stationary states $\boldsymbol{v}_e$, ρ_e, provided the functions F and G satisfy certain conditions. Explicitly, we are taking

$$H_C(\boldsymbol{v}, \rho) = \int_D \mathrm{d}x\, \mathrm{d}z\, [\tfrac{1}{2}|\boldsymbol{v}|^2 + \rho g z/\rho_* + \omega F(\rho) + G(\rho) + \lambda\omega], \tag{2.18}$$

where F, G and the constant λ are to be determined. The first variation

$$\delta H_C := \mathrm{D}H_C(\boldsymbol{v}, \rho)\cdot(\delta\boldsymbol{v}, \delta\rho) \tag{2.19}$$

is given by

$$\delta H_C = \int_D \mathrm{d}x\,\mathrm{d}z\,\{[\boldsymbol{v} + \nabla F(\rho)\times\hat{\boldsymbol{y}}]\cdot\delta\boldsymbol{v} + [gz/\rho_* + \omega F'(\rho) + G'(\rho)]\,\delta\rho\} + \lambda\int_D \mathrm{d}x\,\mathrm{d}z\,\mathrm{div}\,(\hat{\boldsymbol{y}}\times\delta\boldsymbol{v}) \tag{2.20}$$

$$= \int_D \mathrm{d}x\,\mathrm{d}z\,\{[\delta\boldsymbol{v}\cdot\hat{\boldsymbol{y}}\nabla(\psi - F(\rho))] + [gz/\rho_* + \omega F'(\rho) + G'(\rho)]\,\delta\rho\} + \oint_{\partial D}(\lambda + F(\rho))\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l}, \tag{2.21}$$

where $\mathrm{d}\boldsymbol{l}$ is the infinitesimal line element on the boundary and prime, $'$, refers to derivative with respect to the stated argument. For a given stationary flow satisfying the functional relations (2.11) and (2.13), the first variation δH_C vanishes, provided $F(\rho_e)$, $G(\rho_e)$, and λ are determined as follows:

$$F(\rho_e) = \bar{\psi}(\rho_e) \tag{2.22}$$

(absorbing an additive constant into $\bar{\psi}$),

$$-G'(\rho_e) = (gz/\rho_*) + \omega_e\bar{\psi}'(\rho_e) \tag{2.23a}$$

$$= \bar{\psi}'(\rho_e)\left(\omega_e + \frac{1}{\rho_*}\frac{\mathrm{d}\rho_e}{\mathrm{d}\psi_e}gz\right) \tag{2.23b}$$

and

$$\lambda + F(\rho_e)|_{\partial D} = 0. \tag{2.24}$$

Comparison of (2.23*b*) with the Long equation (2.13) shows that $G(\rho_e)$ is determined by

$$-G'(\rho_e) = \bar{\psi}'(\rho_e)\,L(\psi_e), \tag{2.25}$$

which is easily solved, given the stationary flow functions $\psi_e = \bar{\psi}(\rho_e)$ and $L(\psi_e)$, as

$$G(\psi_e) = -\int^{\psi_e} L(s)\,\mathrm{d}s. \tag{2.26}$$

Thus, with F, λ, and G determined by (2.22), (2.24) *and* (2.26), *respectively, a stationary solution of the two-dimensional Boussinesq equations is an extremum of H_C.*

This extremum is in a formal sense locally a minimum, a maximum, or a saddle, depending on whether the second variation of H_C when evaluated at the critical point is, respectively, positive definite, negative definite, or indefinite (the example of Ball & Marsden (1984) shows that caution is needed in making this correspondence). The second variation of H_C,

$$\delta^2 H_C := \mathrm{D}^2 H_C(\boldsymbol{v}_e, \rho_e)\cdot(\delta\boldsymbol{v}, \delta\rho)^2 \tag{2.27}$$

is given by
$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\,[|\delta \boldsymbol{v}|^2 + 2F'(\rho_e)\,\delta\rho\,\delta\omega + A(\delta\rho)^2], \tag{2.28}$$

where the coefficient A of $(\delta\rho)^2$ is

$$A := \omega_e F''(\rho_e) + G''(\rho_e) \tag{2.29}$$

$$= -\frac{gz}{\rho_*}\frac{\bar{\psi}''(\rho_e)}{\bar{\psi}'(\rho_e)} - (\bar{\psi}'(\rho_e))^2 \frac{\mathrm{d}L(\psi_e)}{\mathrm{d}\psi_e} \tag{2.30}$$

$$= -\frac{\mathrm{d}\psi_e}{\mathrm{d}\rho_e}\frac{\mathrm{d}\omega_e}{\mathrm{d}\rho_e} - \frac{g}{\rho_*}\frac{\mathrm{d}z}{\mathrm{d}\rho_e}, \tag{2.31}$$

as one finds, upon using the stationary flow relations. If $\delta^2 H_C$ in (2.28) has any definite sign, it must be positive, as we see by taking $\delta\rho = 0$. Then, taking $\delta\psi = 0$, we see that the quantity A in (2.29) must be positive, as well. Were there no background flow and if ρ_e depended on z alone, then by (2.31) the positivity of A would mean stable stratification, i.e. $\mathrm{d}\rho_e(z)/\mathrm{d}z < 0$.

Completing the square in the first two terms of $\delta^2 H_C$ in (2.28) leads to

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z[|\nabla(\delta\psi - \bar{\psi}'(\rho_e)\,\delta\rho)|^2 - |\nabla(\bar{\psi}'(\rho_e)\,\delta\rho)|^2 + A(\delta\rho)^2] + 2\oint_{\partial D} \bar{\psi}_e'(\rho)\,\delta\rho\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} \tag{2.32}$$

$$= \int_D \mathrm{d}x\,\mathrm{d}z\,[|\nabla(\delta\psi - \bar{\psi}'(\rho_e)\,\delta\rho)|^2 - (\bar{\psi}(\rho_e))^2|\nabla\delta\rho|^2 + (\bar{\psi}'(\rho_e)\,\nabla^2\bar{\psi}'(\rho_e) + A)\,(\delta\rho)^2] + 2\bar{\psi}'(\rho_e)|_{\partial D}\oint_{\partial D}\delta\rho\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l}, \tag{2.33}$$

where we have integrated by parts and chosen ρ_e to be constant on the boundary. The boundary integral vanishes if either $\delta\rho = 0$ on the boundary, or $\delta\rho\,|_{\partial D} =$ const. and $\oint_{\partial D}\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} = 0$. Now consider variations satisfying $\delta\psi - \bar{\psi}'(\rho_e)\,\delta\rho = 0$, so that the first term vanishes in (2.33), and allow $\delta\rho$ to have large gradients, say, $|\nabla\delta\rho|^2 = \sigma^2(\delta\rho)^2$, $\sigma^2 \gg 1$. Then the second variation $\delta^2 H_C$ will be negative if the quantity

$$[A - \bar{\psi}'(\rho_e)(-\nabla^2 + \sigma^2)\,\bar{\psi}'(\rho_e)] \tag{2.34}$$

is negative, with A given in (2.31) in terms of equilibrium quantities. Since σ^2 can be made arbitrarily large, $\delta^2 H_C$ is indefinite. This leads to the following.

Conjecture. Nonlinear stability can be lost in a stratified Boussinesq shear flow in a vertical plane via the creation of density variations at high wavenumber.

This conjecture is plausible, since indefiniteness of $\delta^2 H_C$ signals that the high wavenumber modes could interact strongly enough via nonlinear terms to cause the solution to drift away from equilibrium values, perhaps through resonances and/or the mechanism of Arnold diffusion.

In §8 we shall consider the idea suggested in the introduction of limiting the maximum wavenumber to a finite value, k_{max}. In the presence of such a cut-off in allowed wavenumber, flows for which

$$A > \bar{\psi}'(\rho_e)(-\nabla^2 + k_{max}^2)\,\bar{\psi}'(\rho_e) \tag{2.35}$$

throughout the domain D will be shown to be nonlinearly stable. The results of §8 lend added credence to the conjecture above that the physical mechanism for loss of nonlinear stability is via the development of disturbances at high wavenumber. Such a mechanism is, of course, consonant with ideas from linear spectral analysis.

The range of stable wavenumbers for a certain equilibrium solution can be estimated, as follows. First, we rewrite H_C in (2.18) in terms of velocity $\boldsymbol{v}$ and density ρ, as

$$H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{\tfrac{1}{2}|\boldsymbol{v}|^2 + (gz/\rho_*)\rho + [F(\rho)+\lambda]\,\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\boldsymbol{v} + G(\rho)\}. \tag{2.36}$$

Consequently, the first variation δH_C is given by

$$\delta H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{\boldsymbol{v}\cdot\delta\boldsymbol{v} + (gz/\rho_*)\delta\rho + [F(\rho)+\lambda]\,\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\delta\boldsymbol{v} + [F'(\rho)\,\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\boldsymbol{v} + G'(\rho)]\,\delta\rho\}. \tag{2.37}$$

Taking the second variation evaluated at the equilibrium state $(\rho_e, \boldsymbol{v}_e)$ now gives

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{|\delta\boldsymbol{v}|^2 + 2F'(\rho_e)\,\delta\rho\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\delta\boldsymbol{v} + [F''(\rho_e)\,\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\boldsymbol{v}_e + G''(\rho_e)]\,(\delta\rho)^2\}. \tag{2.38}$$

Integrating the second term in (2.38) by parts and taking $\delta\rho\,|_{\partial D} = \text{const.}$, as before, leads to

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{|\delta\boldsymbol{v}|^2 + 2\delta\boldsymbol{v}\cdot\mathrm{curl}\,(\delta\rho F'(\rho_e)\,\hat{\boldsymbol{y}}) + A(\delta\rho)^2\}, \tag{2.39}$$

where the quantity A is defined in (2.31). Thus, upon completing squares in (2.39) we have

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{|\delta\boldsymbol{v} + \nabla(F'(\rho_e)\,\delta\rho)\times\hat{\boldsymbol{y}}|^2 + A(\delta\rho)^2 - |\nabla(F'(\rho_e)\,\delta\rho)|^2\}. \tag{2.40}$$

Let us define a quantity τ with dimensions of time by setting

$$|\nabla(F'(\rho_e)\,\delta\rho)|^2 = (g\tau/\rho_*)^2\,(\delta\rho)^2, \tag{2.41}$$

where $F'(\rho_e) = \mathrm{d}\psi_e/\mathrm{d}\rho_e$, by (2.22). Note that $\delta^2 H_C$ in (2.40) will be positive definite, provided

$$A > (g\tau/\rho_*)^2, \tag{2.42}$$

where A is given in (2.31).

For a given equilibrium flow, the inequality (2.42) determines the maximum value of τ in (2.41) for which linearized Liapunov stability is assured by the present method. For each equilibrium flow, then, one can consider 'monitoring' the quantities A in (2.31) and $(g\tau/\rho_*)^2$ in (2.41) as the flow develops, knowing that as long as $A > (g\tau/\rho_*)^2$ the flow will be linearized Liapunov stable with conserved stability norm given by $\delta^2 H_C$ in (2.40).

For a stratified Boussinesq shear flow, with $\rho_e = \bar{\rho}(z)$, $\boldsymbol{v}_e = \bar{v}(z)\,\hat{\boldsymbol{x}}$, we have $\psi_e = \bar{\psi}(z)$, $\omega_e = \bar{\omega}(z)$, and inequality (2.42) becomes, with $N^2 = -(g/\rho_*)\,(\mathrm{d}\bar{\rho}/\mathrm{d}z)$ defined to be the Brunt–Väisälä frequency,

$$A = \frac{(g/\rho_*)}{N^2}\left(-\frac{\mathrm{d}\bar{\psi}}{\mathrm{d}z}\frac{\mathrm{d}\bar{\omega}}{\mathrm{d}z}\frac{1}{N^2} + 1\right) > (g\tau/\rho_*)^2. \tag{2.43}$$

Equivalently, we may write this as a quadratic inequality in N^2,

$$\tau^2(N^2)^2 - N^2 + \bar{v}(z)\,\bar{v}''(z) < 0. \tag{2.44}$$

For $(4\tau^2)^{-1} > \bar{v}(z)\,\bar{v}''(z)$, condition (2.44) is satisfied for real values of N^2 in the range

$$\lambda_- < N^2 < \lambda_+, \tag{2.45}$$

where

$$\lambda_\pm = 1/2\tau^2[1 \pm (1 - 4\tau^2 \bar{v}(z)\,\bar{v}''(z))^{\frac{1}{2}}]. \tag{2.46}$$

Note that for $\bar{v}(z)\,\bar{v}''(z) < 0$, stability is possible even for negative values of N^2. Also, for $\bar{v}(z)\,\bar{v}''(z) < 0$ the values of τ^2 are not bounded above by the requirement that the discriminant of (2.44) be positive, although the range of N^2 satisfying (2.45) collapses to zero as τ^2 becomes large. For $\tau^2 = 0$, the stability inequality (2.44) becomes $A \leqslant 0$, or

$$N^2 > \bar{v}(z)\,\bar{v}''(z), \tag{2.47}$$

which is reminiscent of the so-called Richardson number criterion due to Synge (1933), Miles (1961) and Howard (1961), namely,

$$Ri = N^2/(\bar{v}'(z))^2 \geqslant \tfrac{1}{4}, \tag{2.48}$$

for spectral stability of the linearized equations for planar Boussinesq shear flows. In terms of the Richardson number Ri in (2.48), the condition $A > 0$ in (2.47) becomes

$$Ri > \frac{\mathrm{d}\ln \bar{v}'(z)}{\mathrm{d}\ln \bar{v}(z)}, \tag{2.49}$$

which involves a logarithmic derivative of the velocity profile. Finally, in terms of N^2 and $v(z)$ the inequality (2.44) determines the stable range of density-variation gradients to be

$$\tau^2 = \left(\frac{\rho_*}{g}\right)^2 \left|\nabla\left(\frac{\bar{v}(z)\,\delta\rho}{\bar{\rho}'(z)}\right)\right|^2 \Big/ (\delta\rho)^2 < \frac{N^2 - \bar{v}(z)\,\bar{v}''(z)}{(N^2)^2}, \tag{2.50}$$

where $|\tau| = |\nabla(\bar{v}(z)\,\delta\rho/-N^2)|/|\delta\rho|$, by (2.41) and (2.22). *So long as inequality* (2.50) *is satisfied, the stratified Boussinesq shear flow* $\rho_e = \bar{\rho}(z)$, $\boldsymbol{v}_e = \bar{v}(z)\,\hat{\boldsymbol{x}}$ *will remain linearized Liapunov stable in two dimensions with conserved stability norm given in* (2.40). *However, stability can be lost by the development of high wavenumber density variations violating* (2.50), *subsequent to some initial perturbation with wavenumbers in the stable range of* (2.50).

A Hamiltonian formulation of the linearized equations and its relation to spectral theory is discussed in Appendix C. In particular, $\delta^2 H_C$ is shown to be a conserved Hamiltonian for the linearized equations, and the Taylor–Goldstein equation which governs spectral stability is derived in a manner parallel to the present analysis. Again, we emphasize that spectral stability in general does not imply linearized stability. However, if $\delta^2 H_C$ is definite, then its conservation by the linearized equation does establish linearized stability.

Nonlinear stability for the two-dimensional Boussinesq equations is determined by examining the conserved quantity,

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = H_C(\boldsymbol{v}_e + \Delta\boldsymbol{v}, \rho_e + \Delta\rho) - H_C(\boldsymbol{v}_e, \rho_e) - \mathrm{D}H_C(\boldsymbol{v}_e, \rho_e)\cdot(\Delta\boldsymbol{v}, \Delta\rho) \tag{2.51}$$

$$= \int_D \mathrm{d}x\,\mathrm{d}z\,\{\tfrac{1}{2}|\Delta\boldsymbol{v}|^2 + \Delta F(\rho_e)\,\hat{\boldsymbol{y}}\cdot\mathrm{curl}\,\Delta\boldsymbol{v} + \hat{G}(\Delta\rho) + \omega_e \hat{F}(\Delta\rho)\}, \tag{2.52}$$

where the symbol Δ stands for the finite difference, for example,

$$\Delta F(\rho_e) = F(\rho_e + \Delta\rho) - F(\rho_e), \tag{2.53}$$

and the symbol ^ means, for example,

$$\hat{F}(\Delta\rho) = F(\rho_e+\Delta\rho)-F(\rho_e)-F'(\rho_e)\,\Delta\rho$$
$$= \Delta F(\rho_e)-F'(\rho_e)\,\Delta\rho. \tag{2.54}$$

In a Taylor series around ρ_e, we would have

$$F(\rho_e+\Delta\rho) = F(\rho_e)+F'(\rho_e)\,\Delta\rho+\tfrac{1}{2}F''(\rho_e)(\Delta\rho)^2+\dots, \tag{2.55}$$

so the symbol ^ stands for the second-order term *plus the remainder* of the Taylor series (2.55). The factor of $\frac{1}{2}$ in $\hat{H}_C$ in (2.52) relative to $\delta^2 H_C$ in (2.38) appears in the infinitesimal limit as $(\Delta\boldsymbol{v}, \Delta\rho)\to(\delta\boldsymbol{v}, \delta\rho)$ via the Taylor series convention,

$$H_C = H_{C|e}+\delta H_C+\tfrac{1}{2}\delta^2 H_C+\dots. \tag{2.56}$$

Thus, $\hat{H}_C$ in (2.51) is the second-order term *plus the remainder* of the Taylor series (2.56). Just as for $\delta^2 H_C$ in (2.38), we first integrate H_C in (2.52) by parts using the boundary conditions $\Delta\rho\,|_{\partial D}$ = const. and $\oint_{\partial D}\Delta\boldsymbol{v}\cdot d\boldsymbol{l} = 0$. Then we regroup by completing a square to find the desired form (compare with (2.38)),

$$\hat{H}_C = \int_D dx\,dz\{\tfrac{1}{2}|\Delta\boldsymbol{v}+\nabla(\Delta F(\rho_e))\times\hat{\boldsymbol{y}}|^2-\tfrac{1}{2}|\nabla(\Delta F(\rho_e))|^2+[\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)]\}. \tag{2.57}$$

Note that $\nabla(\Delta F(\rho_e))$ depends on $\nabla\Delta\rho$, the *gradient* of the density variation and that $\nabla(\Delta F(\rho_e))\to\nabla(F'(\rho_e)\,\Delta\rho)$ as $\Delta\rho\to 0$. Next, we:

(i) define a quantity $\hat{\tau}$ with dimensions of time by setting

$$|\nabla(\Delta F(\rho_e))|^2 = (g\hat{\tau}/\rho_*)^2\,(\Delta\rho)^2; \tag{2.58}$$

(ii) assume the following convexity condition on the Casimir functions F and G,

$$0<\tfrac{1}{2}\alpha(\Delta\rho)^2 \leqslant [\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)] \leqslant \tfrac{1}{2}\bar{\alpha}(\Delta\rho)^2<\infty; \tag{2.59}$$

(iii) note that for
$$\alpha > (g\hat{\tau}/\rho_*)^2, \tag{2.60}$$

with $\hat{\tau}$ defined in (2.58), we have from (2.57) and the second inequality in (2.59),

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \geqslant \|(\Delta\boldsymbol{v}, \Delta\rho)\|^2. \tag{2.61}$$

Here, the quantity

$$\|(\Delta\boldsymbol{v}, \Delta\rho)\|^2 = \int_D dx\,dz\{\tfrac{1}{2}|\Delta\boldsymbol{v}+\nabla(\Delta F(\rho_e))\times\hat{\boldsymbol{y}}|^2+\tfrac{1}{2}[\alpha-(g\hat{\tau}/\rho_*)^2](\Delta\rho)^2\} \tag{2.62}$$

defines a norm in the space of pairs $(\Delta\boldsymbol{v}, \Delta\rho)$, upon inserting (2.58) into (2.62) and using the lower bound on $\hat{\tau}$, (2.60). Given the upper bound in (2.59), we may also assume that $\hat{H}_C\to 0$ as $(\Delta\boldsymbol{v}, \Delta\rho)\to 0$. Consequently,

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \leqslant R\|(\Delta\boldsymbol{v}, \Delta\rho)\|^2, \tag{2.63}$$

for some positive constant R.

If the convexity hypotheses (2.59–2.60) hold, then the equilibrium state $(\boldsymbol{v}_e, \rho_e)$ will be nonlinearly stable in the norm defined by (2.62). This stability is expressed by bounding the growth of disturbances $(\Delta \boldsymbol{v}, \Delta\rho)$ at time t in terms of disturbances $(\Delta \boldsymbol{v}_0, \Delta\rho_0)$ at time 0, by using the conservation of $\hat{H}_C$ and the two bounds (2.61) and (2.63). Namely, we have

$$\|(\Delta \boldsymbol{v}, \Delta\rho)\|^2 \leqslant \hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) = \hat{H}_C(\Delta \boldsymbol{v}_0, \Delta\rho_0) \leqslant R\|(\Delta \boldsymbol{v}_0, \Delta\rho_0)\|^2. \tag{2.64}$$

As long as (2.60) continues to hold for $\hat{r}$ defined in (2.58), the inequalities (2.64) provide bounding norms for $\hat{H}_C$. Thus, we have proved the following.

THEOREM. *Given an equilibrium flow* $(\boldsymbol{v}_e, \rho_e)$ *with constant density on each component of the boundary, if* (2.59–2.60) *hold, then* $(\boldsymbol{v}_e, \rho_e)$ *is nonlinearly stable in the norm defined by* (2.62), *as long as* (2.60) *continues to hold for* $\hat{r}$ *defined in* (2.58).

3. Three-dimensional Boussinesq flow

The stability analysis in the previous section applies to planar flows in the (x, z) plane. Planar analysis is appropriate for the linearized equations, since the Squire transformation reduces the *linearized* three-dimensional spectral stability problem for Boussinesq shear flows to an equivalent two-dimensional one, in the absence of rotation. Thus, the linearized stability properties of a non-rotating Boussinesq fluid in three dimensions can be understood from theoretical studies of plane parallel flows via the Taylor–Goldstein equation (see, for example, Drazin & Reid (1981), §44). For example, Howard (1961) and Miles (1961) show spectral stability (that is, purely imaginary spectrum) for a Taylor–Goldstein flow, provided the local Richardson number is everywhere greater than or equal to one-quarter, i.e. $Ri \geqslant \frac{1}{4}$, where, for equilibrium quantities $\rho_e = \bar{\rho}(z)$, $\boldsymbol{v}_e = \bar{v}(z)\,\hat{\boldsymbol{x}}$, with $\hat{\boldsymbol{x}}$ the unit vector in the x-direction,

$$Ri := -\frac{g}{\rho_*}\frac{\mathrm{d}\bar{\rho}(z)}{\mathrm{d}z}\bigg/\left(\frac{\mathrm{d}\bar{v}}{\mathrm{d}z}\right)^2. \tag{3.1}$$

However, geophysical flows are often not planar, even in a first approximation and, thus, will not admit the Squire transformation; for example, Coriolis force causes fluid motion out of the (x, z) plane.

We now consider nonlinear stability for Boussinesq flows in three dimensions. We will find nonlinear stability conditions, modulo certain assumptions on bounding the gradients of the density variations that develop from an initial perturbation. In §4 these stability conditions will be shown to hold for a certain class of shear flows if $Ri > 1$, with a Richardson number defined relative to surfaces of constant density.

We recall the Boussinesq equations in three dimensions:

$$\partial \boldsymbol{v}/\partial t = \boldsymbol{v}\times\boldsymbol{\Omega} - \nabla(\tfrac{1}{2}|\boldsymbol{v}|^2 + p/\rho_* + \rho g z/\rho_*) + gz\nabla\rho/\rho_*, \tag{3.2}$$

$$\partial\rho/\partial t = -\boldsymbol{v}\cdot\nabla\rho, \tag{3.3}$$

$$\operatorname{div}\boldsymbol{v} = 0, \tag{3.4}$$

where $\boldsymbol{\Omega} := (\operatorname{curl}\boldsymbol{v}) + \boldsymbol{f}$. As a consequence of the vorticity equation (2.5) written in terms of $\boldsymbol{\Omega}$ as

$$\partial\boldsymbol{\Omega}/\partial t = \operatorname{curl}(\boldsymbol{v}\times\boldsymbol{\Omega}) + (g/\rho_*)\,\nabla z\times\nabla\rho, \tag{3.5}$$

and the continuity equation (3.3), the *potential vorticity* $q := \boldsymbol{\Omega}\cdot\nabla\rho$ is conserved along flow lines, that is,

$$\partial q/\partial t = -\boldsymbol{v}\cdot\nabla q. \tag{3.6}$$

Discussion of potential vorticity and references to its role in geophysical stability analysis are to be found in Gill (1982), Pedlosky (1979), LeBlond & Mysak (1978), and Thomson & Stewart (1977).

Since ρ, the variation of density from its reference value, and q, the potential vorticity, are both conserved along flow lines, the quantity

$$C_{\phi,\lambda} = \int_{D} [\phi(\rho, q) + \lambda q]\, d^3x \tag{3.7}$$

is a constant of the motion for any real-valued function ϕ of two real variables and constant λ. The term λq is separated out for later convenience. Constancy of $C_{\phi,\lambda}$ is readily verified upon using the condition that fluid velocity be tangential to the fixed boundary ∂D of the domain D in which the flow takes place. Likewise, the total energy is a constant of the motion, namely

$$H = \int_{D} (\tfrac{1}{2}|\boldsymbol{v}|^2 + \rho g z/\rho_*)\, d^3x, \tag{3.8}$$

after normalizing by ρ_*.

The states of equilibrium $(\rho_e, \boldsymbol{v}_e)$ of the dynamical system (3.2), (3.3) are the stationary three-dimensional Boussinesq flows. For such stationary flows, there are three 'streamline relations,'

$$\boldsymbol{v}_e \cdot \nabla \rho_e = 0, \tag{3.9}$$

$$\boldsymbol{v}_e \cdot \nabla q_e = 0, \tag{3.10}$$

$$\boldsymbol{v}_e \cdot \nabla (\tfrac{1}{2}|\boldsymbol{v}_e|^2 + p_e/\rho_* + \rho_e g z/\rho_*) = 0. \tag{3.11}$$

The first two of these relations follows by conservation of ρ in (3.3) and q in (3.6), while the last one is the Bernoulli Law, which follows by taking the scalar product of $\boldsymbol{v}_e$ with (3.2) and by using $\boldsymbol{v}_e \cdot \nabla \rho_e = 0$ for stationary solutions.

At points where $\boldsymbol{v}_e \neq 0$, the streamline relations (3.9–3.11) imply that the quantities ρ_e, q_e, and $(\tfrac{1}{2}|\boldsymbol{v}_e|^2 + p_e/\rho_* + \rho_e g z/\rho_*)$ are functionally dependent. We *assume*, in fact, that

$$\tfrac{1}{2}|\boldsymbol{v}_e|^2 + p_e/\rho_* + \rho_e f z/\rho_* = K(\rho_e, q_e), \tag{3.12}$$

where $K(\rho_e, q_e)$ is called the *Bernoulli function*. We also *assume* that $\nabla\rho_e \times \nabla q_e \neq 0$, so that level surfaces of ρ_e, q_e are not mutually tangential anywhere in D. We will now show that if $q_e \neq 0$, then

$$\boldsymbol{v}_e = q_e^{-1} K_q(\rho_e, q_e)\, \nabla\rho_e \times \nabla q_e, \tag{3.13}$$

which automatically satisfies (3.9) and (3.10). In (3.13), subscript notation denotes partial derivative, for example, $K_q = \partial K/\partial q$. The motion equation (3.2) for stationary flows and the relation (3.12) lead to

$$\boldsymbol{v}_e \times \boldsymbol{\Omega}_e = \nabla K(\rho_e, q_e) - (gz/\rho_*)\, \nabla \rho_e. \tag{3.14}$$

Vector multiplication of this by $\nabla\rho_e$ produces

$$\boldsymbol{v}_e(\boldsymbol{\Omega}_e \cdot \nabla\rho_e) - \boldsymbol{\Omega}_e(\boldsymbol{v}_e \cdot \nabla\rho_e) = K_q(\rho_e, q_e)\, \nabla\rho_e \times \nabla q_e. \tag{3.15}$$

Relation (3.13) follows, since the scalar product $\boldsymbol{v}_e \cdot \nabla\rho_e$ vanishes and $\boldsymbol{\Omega}_e \cdot \nabla\rho_e = q_e$.

Another useful relation for stationary flows arises by scalar multiplication of (3.14) by $\boldsymbol{\Omega}_e$, yielding

$$\rho_*^{-1} g z - (\boldsymbol{\Omega}_e \cdot \nabla q_e)\, K_q(\rho_e, q_e)/q_e - K_\rho(\rho_e, q_e) = 0. \tag{3.16}$$

This replaces the Long equation, (2.13). Relations (3.13) and (3.16) will be useful in the development of a variational principle for stationary flows in three dimensions.

Stationary flows will now be sought as extrema of the sum of the conserved quantities $H+C_{\phi,\lambda}$ in (3.8) and (3.7). Let

$$H_C(\rho, \boldsymbol{v}) := H + C_{\phi,\lambda} = \int_D [\tfrac{1}{2}|\boldsymbol{v}|^2 + \rho g z/\rho_* + \phi(\rho, q) + \lambda q]\, d^3x, \tag{3.17}$$

where λ is a constant. Let δH_C be the first variation of H_C, i.e.

$$\delta H_C := DH_C(\rho, \boldsymbol{v}) \cdot (\delta\rho, \delta\boldsymbol{v}). \tag{3.18}$$

After integration by parts and use of the divergence theorem, δH_C is expressible as

$$\delta H_C = \int_D d^3x\, [(gz/\rho_* + \phi_\rho - \boldsymbol{\Omega}\cdot\nabla\phi_q)\,\delta\rho + (\boldsymbol{v} - \phi_{qq}\nabla\rho\times\nabla q)\cdot\delta\boldsymbol{v}] + \oint_{\partial D} ds(\phi_q + \lambda)(\delta\rho\boldsymbol{\Omega} - \nabla\rho\times\delta\boldsymbol{v})\cdot\hat{\boldsymbol{n}}, \tag{3.19}$$

where $\hat{\boldsymbol{n}}$ is the outward unit vector normal to the boundary ∂D and dS is its area element.

Consequently, the first variation δH_C *in* (3.19) *vanishes for stationary flows, provided* $\phi(\rho_e, q_e)$ *is determined from* $K(\rho_e, q_e)$ *by*

$$K(\rho_e, q_e) = -\phi(\rho_e, q_e) + q_e\phi_q(\rho_e, q_e) \tag{3.20}$$

in the interior of D, *and*

$$\lambda = -\phi_q(\rho_e, e_e)|_{\partial D} \tag{3.21}$$

on the boundary ∂D.

For stationary flows, the streamline relations $\boldsymbol{v}_e\cdot\nabla\rho_e = 0$ and $\boldsymbol{v}_e\cdot\nabla q_e = 0$ on the boundary together imply that $\phi_q(\rho_e, q_e)|_{\partial D}$ is a constant. Thus, if (3.21) holds, the boundary integral in δH_C will vanish. In view of (3.16), the velocity relation (3.13), and the definition $q_e = \boldsymbol{\Omega}_e\cdot\nabla\rho_e$, the coefficients of $\delta\rho$ and $\delta\boldsymbol{v}$ in (3.19) will also vanish, provided relation (3.20) holds in D. Solving this gives

$$\phi(\rho_e, q_e) = q_e\left[\int^{q_e} (ds/s^2)\, K(\rho_e, s) + \kappa'(\rho_e)\right], \tag{3.22}$$

where $\kappa'(\rho_e)$ is an arbitrary function of ρ_e; note that $q_e\kappa'(\rho_e)$ is a divergence, which contributes $\oint_{\partial D}\kappa(\rho_e)\,\hat{\boldsymbol{n}}\cdot\boldsymbol{\Omega}\, dS$ to $C_{\phi,\lambda}$.

At the equilibrium point $(\rho_e, \boldsymbol{v}_e)$, the second variation of H_C, defined by

$$\delta^2 H_C = D^2 H_C(\rho_e, \boldsymbol{v}_e)\cdot(\delta\rho, \delta\boldsymbol{v})^2, \tag{3.23}$$

is expressible as a matrix formula

$$\delta^2 H_C = \int_D d^3x \left\{|\delta\boldsymbol{v}|^2 + (\delta q, \delta\rho)\begin{bmatrix}\phi_{qq} & \phi_{q\rho}\\ \phi_{q\rho} & \phi_{\rho\rho}\end{bmatrix}\begin{bmatrix}\delta q\\ \delta\rho\end{bmatrix} + \begin{bmatrix}\phi+\lambda\\ q\end{bmatrix}\delta^2 q\right\}, \tag{3.24}$$

where the element of the 2×2 matrix are to be evaluated at ρ_e, $\boldsymbol{v}_e$ and $\delta^2 q = 2\,\mathrm{curl}\,\delta\boldsymbol{v}\cdot\nabla\delta\rho$. As mentioned in the introduction, positivity of $\delta^2 H_C$ implies linearized stability of the equilibrium solutions ρ_e, $\boldsymbol{v}_e$. A *sufficient* condition for positivity of $\delta^2 H_C$ is determined, as follows.

Collecting the first and last terms in (3.24) and integrating by parts gives

$$\int_D d^3x\{|\delta \boldsymbol{v}|^2+2(\phi_q+\lambda)\operatorname{curl}\delta\boldsymbol{v}\cdot\nabla\delta\rho\} = \int_D d^3x\{|\delta\boldsymbol{v}|^2+2\delta\boldsymbol{v}\cdot\nabla\phi_q\times\nabla\delta\rho\}$$

$$= \int_D d^3x\{|\delta\boldsymbol{v}+\nabla\phi_q\times\nabla\delta\rho|^2+(\nabla\phi_q\cdot\nabla\delta\rho)^2-|\nabla\phi_q|^2|\nabla\delta\rho|^2\}, \quad (3.25)$$

where we have used (3.21) to eliminate the boundary term and completed squares using the vector identity $|\boldsymbol{\alpha}\times\boldsymbol{\beta}|^2 = |\boldsymbol{\alpha}|^2|\boldsymbol{\beta}|^2-(\boldsymbol{\alpha}\cdot\boldsymbol{\beta})^2$ for any vectors $\boldsymbol{\alpha}$, $\boldsymbol{\beta}$. We control the last (negative) term in (3.25) by setting

$$|\nabla\delta\rho|^2 = |\boldsymbol{k}|^2(\delta\rho)^2 \quad \text{with} \quad |\boldsymbol{k}|^2 < k_+^2, \quad (3.26)$$

where k_+^2 will be determined in terms of equilibrium flow quantities. Thus, (3.24) becomes by virtue of (3.25–3.26),

$$\delta^2 H_C = \int_D d^3x\left\{|\delta\boldsymbol{v}+\nabla\phi_q\times\nabla\delta\rho|^2+(\nabla\phi_q\cdot\nabla\delta\rho)^2+(\delta\rho,\delta q)\begin{bmatrix}\phi_{\rho\rho}-|\boldsymbol{k}|^2|\nabla\phi_q|^2 & \phi_{q\rho}\\ \phi_{q\rho} & \phi_{qq}\end{bmatrix}\begin{bmatrix}\delta\rho\\ \delta q\end{bmatrix}\right\}. \quad (3.27)$$

Sufficient for positivity of $\delta^2 H_C$ is that the 2×2 matrix in (3.27) has only positive eigenvalues or, equivalently (by the Sylvester theorem), each subdeterminant be positive. This requires

$$\phi_{qq}(\rho_e, q_e)\ [\text{by } (3.20)] = \frac{1}{q_e}K_q(\rho_e, q_e)\ [\text{by } (3.13)] = \frac{\boldsymbol{v}_e\cdot\nabla\rho_e\times\nabla q_e}{|\nabla\rho_e\times\nabla q_e|^2} > 0 \quad (3.28)$$

and $|\boldsymbol{k}|^2 < k_+^2$, with

$$k_+^2 := \left[\frac{\phi_{\rho\rho}-\phi_{\rho q}^2/\phi_{qq}}{|\nabla\phi_q|^2}\right]_e > 0, \quad (3.29)$$

where subscript e in (3.29) means evaluated at (ρ_e, q_e) and we have used positivity of the 2×2 determinant also to define k_+^2. The requirements (3.28–3.29) for positivity of $\delta^2 H_C$ will provide conditional linearized stability criteria for three-dimensional Boussinesq flows, i.e. sufficient conditions for linearized stability, so long as the condition

$$|\boldsymbol{k}| := |\nabla\delta\rho|\,/\,|\delta\rho| < |k_+| \quad (3.30)$$

is satisfied. Thus, just as in the previous case of Boussinesq flow in a vertical plane, stability can be lost upon development of sufficiently large gradients in the density variation. With conditions (3.28–3.29) for formal stability in mind, we consider nonlinear stability.

The nonlinear stability argument uses convexity of the function $\phi(\rho, q)$ in combination with conservation of the following quantity

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) := H_C(\boldsymbol{v}_e+\Delta\boldsymbol{v}, \rho_e+\Delta\rho)-H_C(\boldsymbol{v}_e, \rho_e)-DH_C(\boldsymbol{v}_e, \rho_e)\cdot(\Delta\boldsymbol{v}, \Delta\rho). \quad (3.31)$$

From $\hat{H}_C$ we shall obtain nonlinear estimates that establish Liapunov stability in a certain norm (See Appendix A for an overview of the general procedure.). Here $\Delta\boldsymbol{v}$ and $\Delta\rho$ are considered to be *finite* velocity and density disturbances at a certain time t, which have the values $\Delta\boldsymbol{v}_0$ and $\Delta\rho_0$ at time zero and evolve according to the Boussinesq equations. The quantity $\hat{H}_C$ in (3.31) is conserved, since: $H_C(\boldsymbol{v}_e+\Delta\boldsymbol{v}, \rho_e+\Delta\rho)$ is conserved for any $\Delta\boldsymbol{v}$, $\Delta\rho$; $H_C(\boldsymbol{v}_e, \rho_e)$ is merely a constant real number and $DH_C(\boldsymbol{v}_e, \rho_e)$ vanishes.

Definition (3.31) now yields the formula

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) = \int \mathrm{d}^2x\,\{\tfrac{1}{2}|\Delta\boldsymbol{v}|^2 + \phi(\rho_e+\Delta\rho, q_e+\Delta q) - \phi(\rho_e, q_e) \\ -\phi_\rho(\rho_e, q_e)\,\Delta\rho - [\phi_q(\rho_e, q_e)+\lambda](\Delta q - \Delta\boldsymbol{\Omega}\cdot\nabla\Delta\rho)\}, \quad (3.32)$$

in which Δq is defined as

$$\Delta q = (\boldsymbol{\Omega}_e + \Delta\boldsymbol{\Omega})\cdot\nabla(\rho_e+\Delta\rho) - \boldsymbol{\Omega}_e\cdot\nabla\rho_e \quad (3.33)$$

$$= \boldsymbol{\Omega}_e\cdot\nabla(\Delta\rho) + \Delta\boldsymbol{\Omega}\cdot\nabla\rho_e + \Delta\boldsymbol{\Omega}\cdot\nabla(\Delta\rho). \quad (3.34)$$

Combining the first and last terms in (3.32), integrating by parts, and completing a square as in (3.25) gives

$$\frac{1}{2}\int_D \mathrm{d}^3x\,\{|\Delta\boldsymbol{v}|^2 + 2(\phi_q+\lambda)\,\mathrm{curl}\,\Delta\boldsymbol{v}\cdot\nabla\Delta\rho\} \\ = \frac{1}{2}\int_D \mathrm{d}^3x\,\{|\Delta\boldsymbol{v}+\nabla\phi_q\times\nabla\Delta\rho\,|^2 + (\nabla\phi_q\cdot\nabla\Delta\rho)^2 - |\nabla\phi_q|^2|\nabla\Delta\rho|^2\}. \quad (3.35)$$

The factor of one-half in (3.35) relative to (3.25) appears since in a Taylor series we have

$$H_C = H_C|_e + \delta H_C + \tfrac{1}{2}\delta^2 H_C + \ldots \quad (3.36)$$

and $\hat{H}_C$ in (3.31) is the second-order term *plus the remainder* of this Taylor series. Define also

$$\hat{\phi}(\Delta\rho, \Delta q) := \phi(\rho_e+\Delta\rho, q_e+\Delta q) - \phi(\rho_e, q_e) - \mathrm{D}\phi(\rho_e, q_e)\;(\Delta\rho, \Delta q), \quad (3.37)$$

so that, by (3.35) and (3.37) equation (3.32) becomes

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta q) = \frac{1}{2}\int_D \mathrm{d}^2x\,\{|\Delta\boldsymbol{v}+\nabla\phi_q\times\nabla\Delta\rho|^2 + (\nabla\phi_q\cdot\nabla\Delta\rho)^2 - |\nabla\phi_q|^2|\nabla\Delta\rho|^2 + 2\hat{\phi}(\Delta\rho, \Delta q)\}. \quad (3.38)$$

To find convexity conditions for $\hat{H}_C$ in (3.38) we first set, as in (3.26),

$$|\nabla\Delta\rho|^2 = |\boldsymbol{k}|^2(\Delta\rho)^2 < k_+^2(\Delta\rho)^2, \quad (3.39)$$

where k_+^2 is to be determined. Next, we define

$$\Psi(\rho, q) := \phi(\rho, q) - \tfrac{1}{2}|\nabla\phi_q(\rho_e, q_e)|^2|\boldsymbol{k}^2|\rho^2, \quad (3.40)$$

where ρ, q are *any* arguments of Ψ, but $|\nabla\phi_q(\rho_e, q_e)|^2$ is a *fixed* function of $\boldsymbol{x}$ determined by the equilibrium flow. This means, in particular, that

$$\begin{aligned}\Psi_{\rho\rho}(\rho_e, q_e) &= \phi_{\rho\rho}(\rho_e, q_e) - |\nabla\phi_q(\rho_e, q_e)|^2\,|\boldsymbol{k}^2| \\ &\geqslant \phi_{\rho\rho}(\rho_e, q_e) - |\nabla\phi_q(\rho_e\, q_e)\,|^2\,k_+^2, \\ \Psi_{qq}(\rho_e, q_e) &= \phi_{qq}(\rho_e, q_e), \\ \Psi_{\rho q}(\rho_e, q_e) &= \phi_{\rho q}(\rho_e, q_e),\end{aligned} \quad (3.41)$$

for the partial derivatives of Ψ. Then the use of (3.39–3.40) in (3.38) gives

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta q) \geqslant \frac{1}{2}\int_D \mathrm{d}^3x\,\{|\Delta\boldsymbol{v}+\nabla\phi_q\times\nabla\Delta\rho|^2 + (\nabla\phi_q\cdot\nabla\Delta\rho)^2 + 2\hat{\Psi}(\Delta\rho, \Delta q)\}, \quad (3.42)$$

where $\hat{\Psi}(\Delta\rho, \Delta q)$ is defined by replacing ϕ in (3.37) by Ψ from (3.40). Finally, we strengthen the positivity conditions for $\delta^2 H_C$ in (3.28–3.39) to read

$$0 < \alpha \leqslant \Psi_{qq}(\rho_e, q_e) < \infty, \tag{3.43}$$

$$0 < \begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}^t \begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix} \leqslant \begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}^t \begin{bmatrix}\Psi_{\rho\rho}(\rho_e, q_e) & \Psi_{\rho q}(\rho_e, q_e) \\ \Psi_{\rho q}(\rho_e, q_e) & \Psi_{qq}(\rho_e, q_e)\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix} < \infty, \tag{3.44}$$

which are convexity conditions for the functional $\int d^3x\, \hat{\Psi}(\Delta\rho, \Delta q)$, in the range of values for ρ_e, q_e, where α, β, and γ are finite positive constants, satisfying $\alpha > 0$, $\gamma\alpha - \beta^2 > 0$. For given choices of α, β, and γ, the second inequality in (3.44) determines k_+^2 in (3.39) as (with the use of (3.41)),

$$0 < k_+^2 = [\phi_{qq}\phi_{\rho\rho} - (\phi_{\rho q}^2 + \gamma\alpha - \beta^2)]/\phi_{qq}|\nabla\phi_q|^2, \tag{3.45}$$

compare with equation (3.29). Inequalities (3.43) and (3.44) lead to a lower bound

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \geqslant \int d^3x \left\{\tfrac{1}{2}|\Delta\boldsymbol{v} + \nabla\phi_q(\rho_e, q_e) \times \nabla\Delta\rho|^2 + \tfrac{1}{2}(\Delta\rho, \Delta q)\begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}\right\} > 0. \tag{3.46}$$

Let

$$\|(\Delta\boldsymbol{v}, \Delta\rho, \Delta q)\|^2 = \int d^3x \left\{\tfrac{1}{2}|\Delta\boldsymbol{v} + \nabla\phi_q(\rho_e, q_e) \times \nabla\Delta\rho|^2 + \tfrac{1}{2}(\Delta\rho, \Delta q)\begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}\right\}, \tag{3.47}$$

which by (3.43) and (3.44) defines a norm on the space of triples $(\Delta\boldsymbol{v}, \Delta\rho, \Delta q)$.

Let us also assume that $\hat{H}_C \to 0$ as $(\Delta\boldsymbol{v}, \Delta\rho) \to 0$ in the norm (3.47). To be specific, assume that

$$\Psi_{qq}(\rho_e, q_e) \leqslant \bar{\alpha}, \tag{3.48}$$

and

$$\Psi_{qq}(\rho_e, q_e)\,\Psi_{\rho\rho}(\rho_e, q_e) - \Psi_{q\rho}^2(\rho_e, q_e) \leqslant \bar{\alpha}\bar{\gamma} - \bar{\beta}^2 \tag{3.49}$$

in the range of ρ_e, q_e, for finite positive constants $\bar{\alpha}$, $\bar{\beta}$, $\bar{\gamma}$. It follows that

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \leqslant R\|(\Delta\boldsymbol{v}, \Delta\rho, \Delta q)\|^2 \tag{3.50}$$

for some positive constant R.

If the hypotheses (3.43), (3.44), (3.48) and (3.49) hold, then $(\boldsymbol{v}_e, \rho_e)$ is nonlinearly stable in the norm (3.47), so long as (3.39) continues to hold, with k_+^2 given in (3.45). Indeed, letting $(\Delta\boldsymbol{v}_0, \Delta\rho_0, \Delta q_0)$ denote the initial value of a disturbance $(\Delta\boldsymbol{v}, \Delta\rho, \Delta q)$ at time zero, we have

$$\|(\Delta\boldsymbol{v}, \Delta\rho, \Delta q)\|^2 \leqslant \hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = \hat{H}_C(\Delta\boldsymbol{v}_0, \Delta\rho_0) \leqslant R\|(\Delta\boldsymbol{v}_0, \Delta\rho_0, \Delta q_0)\|^2, \tag{3.51}$$

which bounds the growth of future disturbances in terms of disturbances at $t = 0$, and thus gives conditional nonlinear stability (see Appendix A), so long as (3.39) holds, with k_+^2 given in (3.45), in terms of the constants α, β, γ, and equilibrium quantities.

In summary, we have proved the following.

THEOREM. *Given an equilibrium flow* $(\boldsymbol{v}_e, \rho_e)$, *we form* $\phi(\rho_e, q_e)$ *satisfying* (3.20), (3.21). *If* (3.43–3.44) *and* (3.48–3.49) *hold, then* $(\boldsymbol{v}_e, \rho_e)$ *is nonlinearly stable in the norm given by* (3.47), *subject to condition* (3.39), *with* k_+^2 *given in* (3.45).

The next section shows how to implement this result in specific cases of three-dimensional flow.

4. A Richardson number criterion for nonlinear stability of shear flows

Here we present specific examples showing how to implement the sufficient conditions for stability given in the previous section.

The meaning of stability is roughly as follows. An initial condition near the equilibrium will produce a time-dependent nonlinear solution; this solution may be approximated by the linearized solution to first-order and by another equation (analogous to the Benjamin–Ono equation, describing two-dimensional flow, see Benjamin (1972)) to second-order. The second-order equation may represent either small-amplitude Rossby waves, or internal waves, depending on the stationary flow. Stability of the stationary flow means that this nonlinear solution remains close to the stationary flow for all time (while the solution exists and is C^1). This does not prove that the time-dependent solutions of the second-order equation are either periodic or stable. See Bennett *et al.* (1983) for results on stability of internal waves in a model equation.

Example I

We first present a simplified version of the basic example for which the analysis is relatively straightforward. Below, we shall generalize some features of it. The example uses a velocity field with a vertical shear, which is a function of y as well as of z and points in the x direction. The density profile is a function of z alone. In this example, there are no Coriolis forces and we take $\rho_* = 1$. Let

$$\boldsymbol{v}_{\mathrm{e}}(x, y, z) = (u(y, z), 0, 0) \tag{4.1}$$

and

$$\rho_{\mathrm{e}}(x, y, z) = \rho(z). \tag{4.2}$$

For $f = 0$, this is an equilibrium with pressure

$$p(z) = -g \int^{z} \rho(z)\,\mathrm{d}z. \tag{4.3}$$

The Bernoulli function is

$$K(\rho_{\mathrm{e}}, q_{\mathrm{e}}) = \tfrac{1}{2}u^2 + p + \rho g z. \tag{4.4}$$

Notice that

$$q_{\mathrm{e}} = -u_y \rho_z. \tag{4.5}$$

If $\rho_z \neq 0$, we can solve $\rho = \rho(z)$ for $z = z(\rho)$ and if $u_{yy} \neq 0$ we can solve $q_{\mathrm{e}} = -u_y \rho_z$ for $y = y(q, z(\rho))$. In this way, K becomes a function of ρ_{e} and q_{e}, which we shall just denote ρ, q for notational simplicity in the rest of this section.

Before proceeding, we observe that the stability condition (3.48), namely

$$0 < \alpha \leqslant \phi_{qq} = K_q/q \leqslant \bar{\alpha},$$

becomes

$$0 < \alpha \leqslant u/(u_{yy}[\rho_z]^2) \leqslant \bar{\alpha} \tag{4.6}$$

by implicit differentiation; note that $K_q = K_y(\partial y/\partial q) = u u_y(\partial y/\partial q)$ and from $q = -u_y \rho_z$, $1 = -u_{yy}(\partial y/\partial q)\,\rho_z$ so

$$\partial y/\partial q = -1/u_{yy}\rho_z. \tag{4.7}$$

Alternatively, one may simply substitute the equilibrium expressions (4.2) and (4.5) into (3.28).

Next we compute ϕ. To simplify the form of u and ρ so we can solve for $z = z(\rho)$ and $y = y(\rho, q)$ explicitly, let us take

$$\rho(z) = -rz \tag{4.8}$$

and

$$u(y, z) = \tfrac{1}{2}\lambda y^2 + \zeta z \tag{4.9}$$

where r, ζ and λ are constants. Then

$$z = -\rho/r \quad \text{and} \quad y = q/\lambda r. \tag{4.10}$$

Substitution into (4.4) gives

$$K(\rho, q) = \frac{1}{2}\left(\frac{1}{2}\frac{q^2}{\lambda r^2} - \frac{\zeta\rho}{r}\right)^2 - \frac{1}{2}\frac{\rho^2 g}{r} = \frac{1}{8}\frac{q^4}{\lambda^2 r^4} - \frac{\zeta\rho q^2}{2\lambda r^3} + \frac{\zeta^2 - gr}{2r^2}\rho^2. \tag{4.11}$$

Thus, using

$$\phi(\rho, q) = q\left[\int^q (\mathrm{d}s/s^2)\, K(\rho, s) + F(\rho)\right], \tag{4.12}$$

we find

$$\phi(\rho, q) = \frac{q^4}{24\lambda^2 r^4} - \frac{q^2\zeta\rho}{2\lambda r^3} - \frac{\zeta^2 - gr}{2r^2}\rho^2 + qF(\rho). \tag{4.13}$$

The condition (3.48) becomes, via (4.6),

$$0 < \alpha \leqslant u/\lambda r^2 \leqslant \bar{\alpha} < \infty, \tag{4.14}$$

which is valid if $u_{yy} = \lambda > 0$ and u is positive in, for example, the domain $0 < a \leqslant y \leqslant b < \infty$ and $0 < c \leqslant z \leqslant d$. This is also ensured if $\zeta/\lambda > 0$, as one sees by directly calculating ϕ_{qq} from (4.13). Namely, the quantity

$$\phi_{qq}(\rho_e, q_e) = \frac{q_e^2}{2\lambda^2 r^4} - \frac{\zeta\rho_e}{\lambda r^3} = \frac{y^2}{2r^2} + \frac{\zeta z}{\lambda r^2} \tag{4.15}$$

is positive for $\zeta/\lambda > 0$ in the above z-domain. Upon choosing $F(\rho) = 0$ in (4.13), equations (3.41) become

$$\begin{aligned} \Psi_{\rho\rho}(\rho_e, q_e) &= -(\zeta^2 - gr)/r^2 - (|\boldsymbol{k}|^2/r^2)\,[\zeta^2(y^2+z^2) + \lambda\zeta y^2 z + \tfrac{1}{4}\lambda^2 y^4], \\ \Psi_{qq}(\rho_e, q_e) &= y^2/2r^2 + \zeta z/\lambda r^2, \\ \Psi_{\rho q}(\rho_e, q_e) &= -y\zeta/r^2 \end{aligned} \tag{4.16}$$

and we find, at equilibrium,

$$\Psi_{qq}\Psi_{\rho\rho} - \Psi_{\rho q}^2 = \frac{u}{\lambda}\frac{\zeta^2}{r^4}\left[\left(\frac{gr}{\zeta^2} - 1\right) - |\boldsymbol{k}|^2(y^2+z^2)\right] - \frac{y^2\zeta^2}{r^2}\left(1 + |\boldsymbol{k}|^2\frac{uz}{\zeta}\right) - \lambda\,|\boldsymbol{k}|^2\frac{uy^4}{r^4}. \tag{4.17}$$

By (4.14) u/λ is positive, so expression (4.17) will be positive (given u and the domain of y and z), for sufficiently small, *but non zero* λ, provided the square-bracketed coefficient in the first term in (4.17) is positive. Looking at the definition of k_+^2 in (3.29) gives

$$k_+^2 = \frac{gr/\zeta^2 - 1}{y^2 + z^2} + O(\lambda). \tag{4.18}$$

Thus, to lowest order in λ [i.e. $O(\lambda^{-1})$], we have, from (4.17),

$$\Psi_{qq}\Psi_{\rho\rho}-\Psi^2_{\rho q}=\frac{u}{\lambda}\frac{\zeta^2}{r^2}\left(1-\frac{|k|^2}{k_+^2}\right)\left(\frac{gr}{\zeta^2}-1\right). \tag{4.19}$$

Consequently, for sufficiently small, but nonzero λ, the stability conditions (3.32) will be satisfied, provided

$$Ri = gr/\zeta^2 > 1. \tag{4.20}$$

We note that $Ri > 1$ is also the condition in this example for $K_{\rho\rho} < 0$, i.e. for the Bernoulli function to have negative curvature in ρ.

The stability inequality (4.20) confirms expectations that increase of buoyancy $r = -\mathrm{d}\rho/\mathrm{d}z > 0$ is stabilizing, while increase of shear $\zeta = u_z$ is destabilizing. Remarkably, though, (4.19) suggests also that a decrease of horizontal profile curvature $\lambda = u_{yy} > 0$ is another stabilizing influence. However, one must keep in mind that the norm (3.27) in which one obtains stability depends on λ. For this example, the coefficient $\phi_{qq} = u/r^2\lambda$ in (3.27) blows up for $\lambda = 0$. In other words, as $\lambda \to 0$ the norm that measures deviations from the stationary solution deteriorates and at $\lambda = 0$ this norm actually becomes infinite.

Example II

We generalize the previous example and give necessary and sufficient conditions for formal stability in a certain approximation. We consider the stability of a parallel shear flow

$$\boldsymbol{v}_\mathrm{e}(\boldsymbol{x}) = (u(y,z),0,0) = (\gamma(y)+U(z),0,0)$$

with density profile depending only on z

$$\rho_\mathrm{e}(\boldsymbol{x}) = \rho(z).$$

The linear stability of this arrangement in the planar case has been exhaustively studied when $\gamma(y) = 0$. In that case, it is a classic result due to Miles & Howard (see the discussion in Drazin & Reid (1981)) that when the Richardson number,

$$Ri = \frac{-g}{\rho_*}\frac{\mathrm{d}\rho}{\mathrm{d}z}\bigg/\left(\frac{\mathrm{d}U}{\mathrm{d}z}\right)^2, \tag{4.21}$$

is greater than, or equal to one-quarter, the fluid is neutrally stable (that is, the eigenvalues of the linearized equation lie on the imaginary axis).

For the three-dimensional situation, we are not treating the case with $\gamma(y)$ identically zero, since then $q_\mathrm{e} = 0$ and the equilibrium flows available as critical points of H_C are static, i.e. then $\boldsymbol{v}_\mathrm{e}(\boldsymbol{x}) = 0$. Furthermore, as we have seen, the corresponding two-dimensional problem has $\delta^2 H_C$ indefinite. To break this degenerate situation, as in example I, we allow a small, slow variation in the y direction on a length scale L (in units of the domain size) larger than any other length in the problem. To be specific, we imagine the flow to occur in a horizontally large box

$$-L_1 \leqslant x \leqslant L_1, \tag{4.22}$$

$$-L_2 \leqslant y \leqslant L_2, \tag{4.23}$$

and

$$-\mathrm{D} \leqslant z \leqslant 0. \tag{4.24}$$

The length L is to be taken larger than L_1, L_2 or D. The boundary conditions on $\boldsymbol{v}$ and ρ are that $\boldsymbol{v}\cdot\hat{\boldsymbol{n}} = 0$ on all sides of the box and ρ is constant on each side, but perhaps taking different values on different sides. In addition, we require periodicity in x and y, but imagine L_1 and L_2 large enough not to interfere with the intended investigation of z-variations.

We consider the restricted example with

$$\gamma(y) = V_0(1+y/L+\tfrac{1}{2}y^2/L^2) \tag{4.25}$$

and have in mind the further restriction $V_0 \ll |U(z)|$, to assure the smallness of the variation in y. The flow is then the usual parallel shear flow, for all physical intents and purposes. We are also going to work in the Boussinesq approximation, setting $\rho_* = 1$.

The analysis proceeds as in §3. We need to determine q and ρ as a function of y and z, and then invert these functions to calculate $\phi(\rho, q)$ at ρ_e and q_e. From its definition

$$q_e = (\operatorname{curl}\boldsymbol{v}_e)\cdot\hat{\boldsymbol{z}}\rho_z. \tag{4.26}$$

We take $\boldsymbol{f} = 0$, so the $\boldsymbol{v}_e$ above is a solution to the momentum equations. This gives

$$q_e = -\rho_z(V_0/L)(1+y/L), \tag{4.27}$$

and from (3.22) we integrate (dropping the subscript 'e')

$$K(\rho, q) = p+\rho g z+\tfrac{1}{2}(u(z)+\gamma(y))^2 \tag{4.28}$$

to arrive at

$$\phi(\rho, q) = -(p+\rho g z+\tfrac{1}{2}U(z)^2)+\phi_{qq}(\rho)\,\tfrac{1}{2}q^2+O(q^4). \tag{4.29}$$

The quantity ϕ_{qq} is given by

$$\phi_{qq} = \frac{\gamma(y)+U(z)}{\gamma_{yy}}\frac{1}{\rho_z^2} \tag{4.30}$$

and for small $\gamma(y)$ it is a function only of z (or equivalently ρ) to order $1/L^2$. Thus, for this example, $\phi(\rho, q)$ is essentially minus the Bernoulli function, plus corrections that are small.

To examine the necessary and sufficient conditions for formal stability, we need to express (3.24) for $\delta^2 H_C$ in terms of two independent components of $\delta\boldsymbol{v}$ and of $\delta\rho$. We choose the two independent components to be $v_3 = \delta\boldsymbol{v}\cdot\hat{\boldsymbol{z}}$ and $\omega_3 = (\operatorname{curl}\delta\boldsymbol{v})\cdot\hat{\boldsymbol{z}}$. Then $\delta^2 H_C$ has the form

$$\delta^2 H_C = \int_D \mathrm{d}^3x \left\{ v_3 \frac{\nabla^2}{\nabla_\perp^2} v_3 + \omega_3\left(-\frac{1}{\nabla_\perp^2}\right)\omega_3 + 2\delta\boldsymbol{v}\cdot(\boldsymbol{\nabla}\phi_q \times \boldsymbol{\nabla}\delta\rho) + (\delta q, \delta\rho)\begin{bmatrix}\phi_{qq} & \phi_{q\rho}\\ \phi_{q\rho} & \phi_{\rho\rho}\end{bmatrix}\begin{bmatrix}\delta q\\ \delta\rho\end{bmatrix}\right\}, \tag{4.31}$$

where $\nabla_\perp^2 = \partial_x^2+\partial_y^2$ and $\nabla^2 = \partial_z^2+\nabla_\perp^2$. The inverse of $\nabla_\perp^2$ can be expressed in terms of the Fourier transform, for variations that are periodic in x and y. For our example,

$$\delta q = \omega_3\rho_z + U_z\partial_y\delta\rho - \gamma_y\partial_z\delta\rho \tag{4.32}$$

and on the basis of our assumptions on γ_y we now choose γ_y so small that the last term in (4.32) is negligible. Under these assumptions and with our ordering scheme for small variations in y, the second variation (4.31) becomes

$$\delta^2 H_C = \int \mathrm{d}^3x\,(v_3, \omega_3, \delta\rho)\begin{bmatrix}\nabla^2/\nabla_\perp^2 & 0 & 0\\ 0 & -1/\nabla_\perp^2+\rho_z^2\phi_{qq} & \rho_z U_z\phi_{qq}\partial_y\\ 0 & -\rho_z U_z\phi_{qq}\partial_y & \phi_{\rho\rho}-U_z^2\phi_{qq}\partial_y^2\end{bmatrix}\begin{bmatrix}v_3\\ \omega_3\\ \delta\rho\end{bmatrix}. \tag{4.33}$$

In reaching this expression, we have dropped terms of higher order in $1/L$, commensurate with our approximations. In particular, the term $\delta \boldsymbol{v} \cdot (\nabla \phi_q \times \nabla \delta\rho)$, which proved so troublesome in the previous section, contributes only to the $(\omega_3, \delta\rho)$ matrix element in (4.33), to leading order in L. Further, its contribution is one order smaller than the term in ϕ_{qq} which we retain. The potential source of instability which it represents is thus absent here. Our task now is to establish the conditions under which the bilinear form (4.38) has definite sign. We do this by establishing conditions for the three-by-three matrix operator in (4.33) to have positive eigenvalues.

If (4.33) is definite, its sign must be positive. This can be seen by considering the variation $(v_3, 0, 0)$. The operator $\nabla^2/\nabla_\perp^2$ is positive definite operating on v_3 with boundary conditions $v_3 = 0$ at $z = 0, -\mathrm{D}$. The eigenfunctions of this operator with these boundary conditions are $(\tilde{\boldsymbol{x}} = (x, y))$

$$\xi_\lambda(\tilde{\boldsymbol{x}}, z) = \xi_0 \, \mathrm{e}^{\mathrm{i}\tilde{\boldsymbol{k}}\cdot\tilde{\boldsymbol{x}}} \sin\frac{\pi n z}{\mathrm{D}}; \quad n = 1, 2, \ldots. \tag{4.34}$$

and the eigenvalues are

$$\lambda = \pi^2 n^2/k_\perp^2 \mathrm{D}^2 + 1; \quad n = 1, 2, 3, \ldots, \tag{4.35}$$

which are clearly positive.

The functional $\delta^2 H_C$ must be positive for an *arbitrary* variation $(v_3, \omega_3, \delta\rho)$ and, in particular, for the variation $(0, \omega_3, \delta\rho)$. The necessary and sufficient condition for the positivity of $\delta^2 H_C$ under this latter choice is that the two-by-two submatrix in (4.33) have only positive eigenvalues. By Fourier transforming in $\tilde{\boldsymbol{k}} = (k_1, k_2)$ space we see this means that the matrix

$$\begin{bmatrix} 1/k_\perp^2 + \rho_z^2 \phi_{qq} & \mathrm{i}k_2 \rho_z U_z \phi_{qq} \\ -\mathrm{i}k_2 \rho_z U_z \phi_{qq} & \phi_{\rho\rho} + k_2^2 U_z^2 \phi_{qq} \end{bmatrix} \tag{4.36}$$

must have only positive eigenvalues for all values of $k_\perp$ and k_2. The conditions for this are

$$\phi_{qq} = u/u_{yy}(\rho_z)^2 > 0 \tag{4.37}$$

and

$$\phi_{\rho\rho} > \max_{(k_1, k_2)} \frac{-k_2^2 U_z^2 \phi_{qq}}{1 + k_\perp^2 \rho_z^2 \phi_{qq}}. \tag{4.38}$$

The first of these is the usual Rayleigh stability condition on the y variation of the velocity profile, and we assume it is satisfied. Its appearance here is expected, since there is no stratification in the y direction. The quantity $\phi_{\rho\rho}$ may be expressed as

$$\phi_{\rho\rho} = \frac{N^2(z)}{\rho_z^2} - \frac{1}{2}\frac{\partial^2}{\partial\rho^2} U(z)^2 + O\left(\frac{1}{L^2}\right), \tag{4.39}$$

where the buoyancy frequency

$$N^2(z) = -\frac{g}{\rho_*}\frac{\partial\rho}{\partial z} \tag{4.40}$$

has been introduced.

If $(U^2)_{\rho\rho}$ is positive, we can translate (4.38) and (4.39) into a condition on the *local Richardson number* defined by

$$Ri(z) = N^2(z)/\rho_z^2 (U^2/2)_{\rho\rho}; \tag{4.41}$$

namely

$$Ri(z) - 1 > \max_{(k_1, k_2)} \frac{-k_\perp^{-2} k_2^2 U_z^2 \phi_{qq}}{k_\perp^{-2} + \rho_z^2 \phi_{qq}} \frac{1}{(U^2/2)_{\rho\rho}} = 0. \tag{4.42}$$

$Ri(z)$ agrees with the usual Richardson number if U and ρ are replaced by their linearizations at z. The result (4.42) is complementary to the Miles–Howard condition; $Ri > \frac{1}{4}$ for linear stability. One also has nonlinear stability, as in example I. When $\min k_2 \neq 0$ and the disturbance is three-dimensional, we have the analogue of the Squire Theorem (see Drazin & Howard 1981) which says here that two-dimensional disturbances are stable for a smaller range of $Ri(z)$ (only $Ri(z) > 1$) than three-dimensional perturbations.

Up to small corrections commensurate with the present approximations, the stability condition (4.38) that $\phi_{\rho\rho} > 0$ can be expressed, by using (4.29), as

$$K_{\rho\rho} \approx (\tfrac{1}{2}U^2 + \rho g z/\rho_* + p/\rho_*)_{\rho\rho} < 0 \tag{4.43}$$

at equilibrium. Thus, for stability the Bernoulli function at equilibrium should have negative curvature as a function of ρ. This approximate stability condition is consistent with the first example, as mentioned after (4.18). We observe from (4.39) that when $(\tfrac{1}{2}U^2)_{\rho\rho}$ is sufficiently negative, an equilibrium flow can be stable even with $\rho_z > 0$, i.e. even when the stratification would be statically unstable. This stabilizing effect of negative curvature does not show up in the first example, since U and ρ are linear in z at equilibrium. The essence of this section has appeared in Abarbanel *et al.* (1984).

Inclusion of rotation

If we work in a rotating frame, taking $\boldsymbol{f} = f(x, y)\,\hat{\boldsymbol{z}}$, then the corresponding changes in the foregoing examples are minor. With $q = (\operatorname{curl} \boldsymbol{v} + \boldsymbol{f}) \cdot \nabla\rho$ we have by (3.13) and (3.16),

$$K(\rho_e, q_e) = \tfrac{1}{2}|\boldsymbol{v}_e|^2 + p_e/\rho_* + \rho_e g z/\rho_*, \tag{4.44}$$

$$\phi_{qq}(\rho_e, q_e) = \frac{K_q(\rho_e, q_e)}{q_e} = \frac{\boldsymbol{v}_e \cdot \nabla\rho_e \times \nabla q_e}{|\nabla\rho_e \times \nabla q_e|^2}, \tag{4.45}$$

$$K_\rho(\rho_e, q_e) = \frac{gz}{\rho_*} - (\boldsymbol{\Omega}_e \cdot \nabla q_e)\,\frac{\boldsymbol{v}_e \cdot \nabla\rho_e \times \nabla q_e}{|\nabla\rho_e \times \nabla q_e|^2}. \tag{4.46}$$

For the equilibrium flow given by (4.1) and (4.2), we have

$$\boldsymbol{\Omega}_e = U_z\,\hat{\boldsymbol{y}} + (f - u_y)\,\hat{\boldsymbol{z}}, \tag{4.47}$$

$$q_e = (f - u_y)\,\rho_z \tag{4.48}$$

and the stability condition (4.6) becomes

$$0 < \alpha \leqslant \phi_{qq}(\rho_e, q_e) = \frac{-u}{(f_y - u_{yy})(\rho_z)^2} \leqslant \bar{\alpha}. \tag{4.49}$$

Thus one of the stability conditions is that the expression $f_y - u_{yy}$ does not change sign in the domain considered. This is a slight generalization of the result of Kuo (1949), who showed that a necessary condition for linear instability of planar, incompressible, parallel flow in a rotating frame, with $f = f_0 + \bar{\beta} y$ for constant f_0, $\bar{\beta}$, is that $\bar{\beta} - u_{yy} = 0$ somewhere in the domain of flow.

This modification of example I to include rotation involves taking

$$f = f_0 + \tilde{\beta} y, \quad \rho(z) = -rz, \quad u(y, z) = \tfrac{1}{2}\lambda y^2 + \zeta z, \tag{4.50}$$

where f_0, $\tilde{\beta}$, r, ζ, and λ are constants, so that

$$z = -\rho/r, \quad y = \frac{q/r + f_0}{\lambda - \tilde{\beta}}.$$

Retracing the steps in example I with these modifications produces two stability conditions, analogous to (4.14) and (4.20), but more complicated in form (especially the second condition).

The first condition, $0 < \alpha \leqslant ur^{-2}/(\lambda-\tilde{\beta})$, simply reflects the Kuo (1949) correction to include rotation, while the second condition shows that the Richardson number criterion $gr/\xi^2 > 1$ for stability found in the previous examples is not essentially modified by the presence of rotation.

5. Two-dimensional inviscid incompressible stratified flow

We now release the Boussinesq approximation and study the stability of stratified, inviscid flow in two spatial dimensions as described by the Euler equations,

$$\partial \boldsymbol{v}/\partial t = -\nabla p/\rho - g\hat{\boldsymbol{z}} + \boldsymbol{v}\times\boldsymbol{\omega} - \nabla\tfrac{1}{2}|\boldsymbol{v}|^2. \tag{5.1}$$

In solving these equations, p is determined from conservation of the incompressibility condition,

$$\operatorname{div}\boldsymbol{v} = 0, \tag{5.2}$$

with ρ satisfying

$$\partial\rho/\partial t + \boldsymbol{v}\cdot\nabla\rho = 0, \tag{5.3}$$

in a fixed, bounded domain D in the xz plane, subject to the boundary conditions $\boldsymbol{v}\cdot\hat{\boldsymbol{n}} = 0$. The vorticity equation for $\omega = \hat{\boldsymbol{y}}\cdot\operatorname{curl}\hat{\boldsymbol{v}} = \nabla^2\psi$ reads

$$\begin{aligned}\partial\omega/\partial t + \boldsymbol{v}\cdot\nabla\omega &= (1/\rho^2)(\nabla\rho\times\nabla p)\cdot\hat{\boldsymbol{y}} \\ &= (1/\rho^2)(\partial_z\rho\,\partial_x p - \partial_z p\,\partial_x\rho)\end{aligned} \tag{5.4}$$

Denoting $\{f, g\} = \partial_z f\partial_x g - \partial_x f\partial_z g$, equations (5.5) and (5.3) become

$$\partial\omega/\partial t = \{\omega, \psi\} + \{p, 1/\rho\} \tag{5.5}$$

$$\partial\rho/\partial t = \{\rho, \psi\}. \tag{5.6}$$

The sense in which (5.5) and (5.6) form a Hamiltonian system is explained in §7.

From the continuity equation (5.6) we see that for a stationary solution (ω_e, ρ_e), the vectors $\nabla\psi_e$ and $\nabla\rho_e$ are collinear in the plane; since ρ_e, ψ_e must be functionally related for the bracket $\{\rho_e, \psi_e\}$ to vanish. Sufficient for this collinearity are the functional relationships

$$\psi_e = \bar{\psi}(\rho_e), \quad \rho_e = \bar{\rho}(\psi_e). \tag{5.7}$$

The Long equation is a condition equivalent to the vanishing of the right-hand side of (5.6). This equation characterizes stationary solutions and is derived as follows. Multiply equation (5.1) in the case of stationary solutions by ρ_e and take its curl to obtain

$$\operatorname{curl}(\rho_e\boldsymbol{v}_e\times\omega_e\hat{\boldsymbol{y}}) = \nabla\rho_e\times\nabla(\tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz). \tag{5.7$'$}$$

Now,

$$\begin{aligned}\operatorname{curl}(\rho_e\boldsymbol{v}_e\times\omega_e\hat{\boldsymbol{y}}) &= [-\rho_e(\boldsymbol{v}_e\cdot\nabla)\,\omega_e\hat{\boldsymbol{y}} - \omega_e\hat{\boldsymbol{y}}\operatorname{div}(\rho_e\boldsymbol{v}_e) + (\omega_e\hat{\boldsymbol{y}}\cdot\nabla)\,\rho_e\boldsymbol{v}_e + \rho_e\boldsymbol{v}_e\operatorname{div}(\omega_e\hat{\boldsymbol{y}})] \\ &= -\rho_e(\boldsymbol{v}_e\cdot\nabla)\,\omega_e\hat{\boldsymbol{y}};\end{aligned}$$

the last two terms in the square brackets are obviously zero and the second one vanishes for stationary solutions. By using (5.7), the dot product of (5.7′) with $\hat{\boldsymbol{y}}$ yields

$$\begin{aligned}0 &= \boldsymbol{v}_e\cdot\nabla\omega_e + (1/\rho_e)\{\rho_e, \tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz\} \\ &= \boldsymbol{v}_e\cdot\nabla\omega_e + \frac{1}{\rho_e}\frac{d\rho_e}{d\psi_e}(\boldsymbol{v}_e\cdot\nabla)(\tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz),\end{aligned} \tag{5.8}$$

Since $\boldsymbol{v}_{\mathrm{e}}\cdot\nabla\psi_{\mathrm{e}} = 0$ and $\boldsymbol{v}_{\mathrm{e}}\cdot\nabla\rho_{\mathrm{e}} = 0$, (5.8) becomes

$$\boldsymbol{v}_{\mathrm{e}}\cdot\nabla\left[\omega_{\mathrm{e}}+\frac{1}{\rho_{\mathrm{e}}}\frac{\mathrm{d}\rho_{\mathrm{e}}}{\mathrm{d}\psi_{\mathrm{e}}}\left(\tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2+gz\right)\right] = 0, \tag{5.9}$$

i.e. the quantity under the gradient is constant on streamlines. One can also reach the conclusion (5.9) rather more directly from (5.5). Just substitute ∇p_{e} using (5.1) in the case of stationary solutions into the right-hand side of (5.5) and use (5.7). Thus the quantities

$$\psi_{\mathrm{e}} \quad \text{and} \quad \left[\omega_{\mathrm{e}}+\frac{1}{\rho_{\mathrm{e}}}\frac{\mathrm{d}\rho_{\mathrm{e}}}{\mathrm{d}\psi_{\mathrm{e}}}\left(\tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2+gz\right)\right]$$

are functionally related: we assume that

$$\omega_{\mathrm{e}}+\frac{1}{\rho_{\mathrm{e}}}\frac{\mathrm{d}\rho_{\mathrm{e}}}{\mathrm{d}\psi_{\mathrm{e}}}\left(\tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2+gz\right) = L(\psi_{\mathrm{e}}). \tag{5.10}$$

This is known as the *Long equation* (Long (1953), Turner (1973), Yih (1980)).

The total energy, which is conserved, has the expression

$$H(\boldsymbol{v},\rho) = \int_{\mathrm{D}} \mathrm{d}x\,\mathrm{d}z\,\left(\tfrac{1}{2}\rho|\boldsymbol{v}|^2+\rho gz\right). \tag{5.11}$$

It is easy to see that when F and G are any real-valued functions of a real variable, ρ, the quantity

$$C_{F,G,\lambda}(\boldsymbol{v},\rho) = \int_{\mathrm{D}} \mathrm{d}x\,\mathrm{d}z(\omega F(\rho)+G(\rho))+\lambda\oint_{\partial\mathrm{D}} \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} \tag{5.12}$$

is also conserved, provided we assume ρ is constant on each connected component of the boundary, i.e. the boundary is 'isopychnal'. If this condition holds at $t = 0$, it holds for all t since ρ is conserved. The geometrical significance of the functional in (5.12) will be explained in §7. Here we mention, though, that $C_{F,G,\lambda}$ is also conserved if ω is replaced by $\hat{\boldsymbol{y}}\cdot\operatorname{curl}\rho\boldsymbol{v}$.

According to the methodology of Appendix A, we form the sum

$$\begin{aligned} H_C(\boldsymbol{v},\rho) &= H(\boldsymbol{v},\rho)+C_{F,G,\lambda}(\boldsymbol{v},\rho) \\ &= \int_{\mathrm{D}} \mathrm{d}x\,\mathrm{d}z\,\left[\tfrac{1}{2}\rho|\boldsymbol{v}|^2+\rho gz+\omega F(\rho)+G(\rho)\right]+\lambda\oint_{\partial\mathrm{D}} \boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} \end{aligned} \tag{5.13}$$

and study its first variation $\delta H_C = \mathrm{D}H_C(\boldsymbol{v},\rho)\cdot(\delta\boldsymbol{v},\delta\rho)$ where $\mathrm{D}H_C(\boldsymbol{v},\rho)\cdot(\delta\boldsymbol{v},\delta\rho)$ denotes the derivative of H_C at $(\boldsymbol{v},\rho)$ in the direction $(\delta\boldsymbol{v},\delta\rho)$. We find

$$\begin{aligned} \delta H_C = \int_{\mathrm{D}} \mathrm{d}x\,\mathrm{d}z\,\{&\rho\boldsymbol{v}\cdot\delta\boldsymbol{v}+F(\rho)\,\hat{\boldsymbol{y}}\cdot\operatorname{curl}\delta\boldsymbol{v} \\ &+\left[\tfrac{1}{2}|\boldsymbol{v}|^2+gz+\omega F'(\rho)+G'(\rho)\right]\delta\rho\}+\lambda\oint_{\partial\mathrm{D}} \delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} \end{aligned} \tag{5.14}$$

$$\begin{aligned} = \int_{\mathrm{D}} \mathrm{d}x\,\mathrm{d}z\,\{&\delta\boldsymbol{v}\cdot\hat{\boldsymbol{y}}\times(\rho\nabla\psi-\nabla F(\rho)) \\ &+\left[\tfrac{1}{2}|\boldsymbol{v}|^2+gz+\omega F'(\rho)+G'(\rho)\right]\delta\rho\}+\oint_{\partial\mathrm{D}} (F(\rho)+\lambda)\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l}. \end{aligned} \tag{5.15}$$

This variational derivative vanishes for a stationary flow $(\boldsymbol{v}_e, \rho_e)$ satisfying (5.1), (5.2), and (5.3), provided $F(\rho_e)$, $G(\rho_e)$, and λ are determined as follows,

$$\rho_e \bar{\psi}'(\rho_e) = F'(\rho_e), \tag{5.16}$$

$$\tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz + \omega_e F'(\rho_e) + G'(\rho_e) = 0, \tag{5.17}$$

$$\lambda + F(\rho_e)|_{\partial D} = 0. \tag{5.18}$$

In the last equality we recall that ρ_e is constant on the boundary. Comparing (5.17) with the Long equation (5.10) and using (5.16) determines $G(\rho_e)$ as

$$G'(\rho_e) = -\rho_e \bar{\psi}'(\rho_e)\, L(\bar{\psi}(\rho_e)). \tag{5.19}$$

Thus, *a stationary solution of the two-dimensional stratified fluid equations is an extremum of H_C, the sum of the energy H in* (5.11) *and the conserved quantity $C_{F,G,\lambda}$ in* (5.12).

The second variation $\delta^2 H_C = \mathrm{D}^2 H_C(\boldsymbol{v}_e, \rho_e)\cdot(\delta\boldsymbol{v}, \delta\rho)^2$ is given by

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\,\{\rho_e|\delta\boldsymbol{v}|^2 + 2\delta\boldsymbol{v}\cdot[\boldsymbol{v}_e\,\delta\rho + \nabla(F'(\rho_e)\,\delta\rho)\times\hat{\boldsymbol{y}}]$$
$$+ [\omega_e F''(\rho_e) + G''(\rho_e)]\,(\delta\rho)^2\} - 2\oint_{\partial D} F'(\rho_e)\,\delta\rho\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l} \tag{5.20}$$

$$= \int_D \mathrm{d}x\,\mathrm{d}z\,\{\rho_e|\nabla\delta\psi|^2 + 2\nabla\delta\psi\cdot[\delta\rho\nabla\psi_e - \nabla(F'(\rho_e)\,\delta\rho)]$$
$$+ [\omega_e F''(\rho_e) + G''(\rho_e)](\delta\rho)^2\} - 2F'(\rho_e)|_{\partial D}\oint_{\partial D}\delta\rho\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l}. \tag{5.21}$$

Completing the square gives

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\left\{\rho_e\left|\nabla\left(\delta\psi - \frac{F'(\rho_e)}{\rho_e}\,\delta\rho\right)\right|^2 - \rho_e\left|\nabla\left(\frac{F'(\rho_e)}{\rho_e}\,\delta\rho\right)\right|^2\right.$$
$$\left. + [\omega_e F''(\rho_e) + G''(\rho_e)](\delta\rho)^2\right\} - 2F'(\rho_e)|_{\partial D}\oint_{\partial D}\delta\rho\,\delta\boldsymbol{v}\cdot\mathrm{d}\boldsymbol{l}. \tag{5.22}$$

The boundary integral in (5.22) vanishes when $\delta\rho|_{\partial D}$ = const. for variations that preserve the circulation on the boundary.

If $\delta\rho = 0$, then $\delta^2 H_C$ is positive. If, however, $\delta\psi = F'(\rho_e)\,\delta\rho/\rho_e$, $\nabla\delta\psi$ is zero on the boundary, and $\delta\rho$ has small variations, but with very steep gradient, the sum of the last two terms in the first integral can be negative. This shows that $\delta^2 H_C$ is indefinite and, as in §2, we are led to the following.

Conjecture. Nonlinear stability can be lost for stationary, two-dimensional solutions with isopychnal boundary conditions of incompressible stratified flow via the creation of density variations at high wavenumber.

As in §2, we conjecture that Arnold diffusion will occur at large wavenumber and cause instability because of nonlinear coupling. In fact, unless systems are completely integrable, this is always expected to occur (see Arnold 1978, Appendix 8; Lichtenberg & Lieberman 1983; Holmes & Marsden 1983).

The range of stable wavenumbers for a given equilibrium solution can be estimated, as follows. Using (5.16), we rewrite (5.22) as

$$\delta^2 H_C = \int_D \mathrm{d}x\,\mathrm{d}z\{\rho_e|\nabla(\delta\psi-\bar{\psi}'(\rho_e)\,\delta\rho)|^2-\rho_e|\nabla(\bar{\psi}'(\rho_e)\,\delta\rho)|^2+A(\delta\rho)^2\}, \tag{5.23}$$

where

$$A := \omega_e F''(\rho_e)+G''(\rho_e). \tag{5.24}$$

Let us define a quantity τ with dimensions of time by setting

$$|\nabla(\bar{\psi}'(\rho_e)\,\delta\rho)|^2 = (g\tau/\rho_e)^2\,(\delta\rho)^2; \tag{5.25}$$

compare with (2.41). Note that $\delta^2 H_C$ in (5.23) will be positive definite, provided

$$A > (g\tau)^2/\rho_e, \tag{5.26}$$

where A is given in (5.24). The interpretation of A in (5.24) is obtained from (5.10), (5.16) and (5.19) as

$$\begin{aligned} A &= \omega_e F''(\rho_e)+G''(\rho_e) \\ &= [\omega_e - L(\bar{\psi}(\rho_e))]\,F''(\rho_e)-F'(\rho_e)\,L'(\bar{\psi}(\rho_e))\,\bar{\psi}'(\rho_e) \\ &= -\rho_e\bar{\psi}'(\rho_e)(\mathrm{d}\omega_e/\mathrm{d}\rho_e)-(\mathrm{d}/\mathrm{d}\rho_e)(\tfrac{1}{2}|\boldsymbol{v}_e|^2+gz). \end{aligned} \tag{5.27}$$

For a stratified Euler shear flow, with $\rho_e=\bar{\rho}(z)$, $\boldsymbol{v}_e=\bar{v}(z)\,\hat{\boldsymbol{x}}$, $=\bar{\psi}'(z)\,\hat{\boldsymbol{x}}$, we have $\omega_e=\bar{\omega}(z)=\bar{v}'(z)$ and $N^2 := -(g/\bar{\rho}(z))\,\bar{\rho}'(z)$, so that

$$\begin{aligned} A &= -\bar{\rho}(z)\,\frac{\bar{\psi}'(z)}{\bar{\rho}'(z)}\frac{\bar{\omega}'(z)}{\bar{\rho}'(z)}-\frac{1}{\bar{\rho}'(z)}\,(\bar{v}(z)\,\bar{v}'(z)+g) \\ &= \frac{g^2/\bar{\rho}(z)}{N^2}\left[1+\frac{1}{g}\,\bar{v}(z)\,\bar{v}'(z)-\frac{1}{N^2}\,\bar{v}(z)\,\bar{v}''(z)\right]. \end{aligned} \tag{5.28}$$

Thus the stability condition (5.26) becomes a quadratic inequality in the buoyancy frequency N^2,

$$\tau^2(N^2)^2-[1+g^{1}\bar{v}(z)\,\bar{v}'(z)]\,N^2+\bar{v}(z)\,\bar{v}''(z) < 0; \tag{5.29}$$

compare with (2.44). For positive discriminant, this quadratic has two real roots; so there is a range of stable values of N^2, including negative values when $\bar{v}(z)\,\bar{v}''(z)<0$. Now, in terms of the buoyancy frequency and other equilibrium quantities, we may express the stable range of density-variation gradients as

$$\tau^2 < \frac{[1+g^{1}\bar{v}(z)\,\bar{v}'(z)]\,N^2-\bar{v}(z)\,\bar{v}''(z)}{(N^2)^2}, \tag{5.30}$$

where, by the definition of τ in (5.25), we have

$$|\tau| = \frac{\bar{\rho}(z)}{g}\left|\nabla\!\left(\frac{\bar{v}(z)\,\delta\rho}{\bar{\rho}'(z)}\right)\right|\Big/|\delta\rho|; \tag{5.31}$$

compare with (2.50). So long as inequality (5.30) is satisfied, the stratified Euler flow $\rho_e=\bar{\rho}(z)$, $\boldsymbol{v}_e=\bar{v}(z)\,\hat{\boldsymbol{x}}$ will remain linearized Liapunov stable in two dimensions, with conserved stability norm given by $\delta^2 H_C$ in (5.23).

Nonlinear stability for the two-dimensional Euler equations is determined by examining the conserved quantity

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = H_C(\boldsymbol{v}_e+\Delta\boldsymbol{v}, \rho_e+\Delta\rho) - H_C(\boldsymbol{v}_e, \rho_e) - DH_C(\boldsymbol{v}_e, \rho_e)\cdot(\Delta\boldsymbol{v}, \Delta\rho) \tag{5.32}$$

$$= \int_D dx\, dz\{\tfrac{1}{2}(\rho_e+\Delta\rho)|\Delta\boldsymbol{v}|^2 + \Delta\rho\boldsymbol{v}_e\cdot\Delta\boldsymbol{v} + [\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)] + \Delta\omega\,\Delta F(\rho_e)\}, \tag{5.33}$$

where the operations Δ and $\hat{}$ are defined as in (2.53) and (2.54), respectively. Integration of the last term in (5.33) by parts and use of the boundary conditions $\Delta\rho|_{\partial D} = \text{const.}$ and $\oint_{\partial D}\Delta\boldsymbol{v}\cdot d\boldsymbol{l} = 0$ gives, upon substituting $\boldsymbol{v}_e = \hat{\boldsymbol{y}}\times\nabla\psi_e$ and $\Delta\boldsymbol{v} = \hat{\boldsymbol{y}}\times\nabla\Delta\psi$,

$$\hat{H}_C(\Delta v, \Delta\rho) = \int_D dx\, dz\{(\rho_e+\Delta\rho)|\Delta\nabla\psi|^2 + \nabla\Delta\psi\cdot[\Delta\rho\nabla\psi_e - \nabla(\Delta F(\rho_e)] + [\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)]\}. \tag{5.34}$$

Relation (5.16) implies

$$\frac{1}{\rho_e+\Delta\rho}[\Delta\rho\nabla\psi_e - \nabla(\Delta F(\rho_e))] = \frac{1}{\rho_e}\nabla F(\rho_e) - \frac{1}{\rho_e+\Delta\rho}\nabla F(\rho_e+\Delta\rho)$$

$$= -\Delta\left[\frac{1}{\rho_e}\nabla F(\rho_e)\right] = -\Delta(\nabla\bar{\psi}(\rho_e)) = -\nabla\Delta\bar{\psi}(\rho_e). \tag{5.35}$$

Hence,

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = \int_D dx\, dz\{\tfrac{1}{2}(\rho_e+\Delta\rho)|\nabla(\Delta\psi - \Delta\bar{\psi}(\rho_e))|^2 - \tfrac{1}{2}(\rho_e+\Delta\rho)|\nabla\Delta\bar{\psi}(\rho_e)|^2 + [\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)]\}. \tag{5.36}$$

Next, we:

(i) define a quantity $\hat{t}$ with dimensions of time by setting, as in (2.58),

$$|\nabla(\Delta\bar{\psi}(\rho_e))|^2 = (g\hat{t}/\rho_*)^2(\Delta\rho)^2; \tag{5.37}$$

(ii) assume the following convexity condition on the Casimir functions F and G,

$$0 < (\alpha/2)(\Delta\rho)^2 \leqslant [\hat{G}(\Delta\rho)+\omega_e\hat{F}(\Delta\rho)] \leqslant (\bar{\alpha}/2)(\Delta\rho)^2 < \infty; \tag{5.38}$$

(iii) observe that since ρ is conserved, we have

$$0 < \rho_{\min} \leqslant \rho_e+\Delta\rho \leqslant \rho_{\max} < \infty \tag{5.39}$$

for all time, if this is satisfied initially (which we assume). Consequently, provided

$$\alpha > \rho_{\max}(g\hat{t}/\rho_*)^2, \tag{5.40}$$

we have

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \geqslant \|(\Delta\boldsymbol{v}, \Delta\rho)\|^2, \tag{5.41}$$

where

$$\|(\Delta\boldsymbol{v}, \Delta\rho\|^2 := \int_D dx\, dz\{\tfrac{1}{2}\rho_{\min}|\Delta\boldsymbol{v} + \nabla(\Delta\bar{\psi}(\rho_e))\times\hat{\boldsymbol{y}}|^2 + \tfrac{1}{2}[\alpha - \rho_{\max}(g\hat{t}/\rho_*)^2](\Delta\rho)^2\} \tag{5.42}$$

defines a norm in the space of pairs $(\Delta\boldsymbol{v}, \Delta\rho)$ upon inserting (5.37) into (5.42) and using the lower bound on $\hat{r}$, (5.40). Given the upper bounds in (5.38) and (5.39), we also have

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) \leqslant \int_D \mathrm{d}x\,\mathrm{d}z\,\{\rho_{\max}|\Delta\boldsymbol{v}+\nabla(\Delta\bar{\psi}(\rho_e))\times\hat{\boldsymbol{y}}|^2+(\alpha-\rho_{\min}(g\hat{r}/\rho_*)^2)(\Delta\rho)^2\}. \quad (5.43)$$

Therefore, $\hat{H}_C \rightarrow 0$ as $(\Delta\boldsymbol{v}, \Delta\rho) \rightarrow 0$, and we may put

$$\hat{H}_C \leqslant Q\|(\Delta\boldsymbol{v}, \Delta\rho)\|^2, \quad (5.44)$$

for some positive constant Q.

If the hypotheses (5.38–5.40) hold and solutions exist for all time, then the equilibrium state $(\boldsymbol{v}_e, \rho_e)$ will be nonlinearly stable in the norm (5.42). Indeed, the estimates (5.41) and (5.44) provide bounds on the growth of disturbances $(\Delta\boldsymbol{v}, \Delta\rho)$ at time t in terms of disturbances $(\Delta\boldsymbol{v}_0, \Delta\rho_0)$ at time 0, using the conservation of $\hat{H}_C$ under the nonlinear evolution. Namely, we have

$$\|(\Delta\boldsymbol{v}, \Delta\rho)\|^2 \leqslant \hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = \hat{H}_C(\Delta\boldsymbol{v}_0, \Delta\rho_0) \leqslant Q\|(\Delta\boldsymbol{v}_0, \Delta\rho_0)\|^2, \quad (5.45)$$

so long as (5.40) holds, with $\hat{r}$ defined in (5.37). We have proved the following.

THEOREM. *Given an equilibrium flow* $(\boldsymbol{v}_e, \rho_e)$ *with constant density on each component of the boundary; if* (5.38–5.40) *hold, then* $(\boldsymbol{v}_e, \rho_e)$ *is nonlinearly stable in the norm* (5.42), *as long as* (5.40) *holds, with* $\hat{r}$ *defined in* (5.37).

Remarks. (A) In Long (1953), (5.10) emerges from another variational principle, in which density variations are taken to be functionally related to variations of the stream function by the equilibrium relation $\rho_e = \bar{\rho}(\psi_e)$, so that $\delta\rho = \bar{\rho}'(\psi_e)\,\delta\psi$. This restriction is consistent with the equilibrium flow being a critical point of the energy functional chosen in Long (1953), but it precludes establishing a proper stability condition by taking a second variation, since only this special direction in $(\delta\rho, \delta\psi)$ space is being tested.

(B) In §8 we shall examine how imposition on the two-dimensional Boussinesq equations of a maximum spatial wavenumber (so-called 'filtering') leads to formal, and rigorous nonlinear stability. A similar analysis of 'filtered' Euler equations could also be performed. The 'filtered' two-dimensional Boussinesq equations in §8 remove the high wavenumbers from the stream function ψ entirely and prevent their development, thereby leading to stability criteria for these modified equations. In contrast to 'filtering', in which the equations are altered in §8, the rest of this paper treats the unmodified Boussinesq and Euler equations, and determines the stable range in wavenumber or gradient modulus of the density variations for a given initial equilibrium solution. 'Monitoring' the magnitudes of the gradients of the density variations would determine whether a certain disturbance remains stable for a given equilibrium solution.

6. Three-dimensional inviscid incompressible stratified flow

The motion of an incompressible, inhomogeneous, rotating fluid under gravity is given by solutions to

$$\partial\boldsymbol{v}/\partial t = -\rho^{-1}\nabla p - \nabla(\tfrac{1}{2}|\boldsymbol{v}|^2+gz)+\boldsymbol{v}\times\boldsymbol{\Omega}, \quad (6.1)$$

$$\partial\rho/\partial t+\boldsymbol{v}\cdot\nabla\rho = 0, \quad (6.2)$$

$$\mathrm{div}\,(\boldsymbol{v}) = 0, \quad (6.3)$$

where $\boldsymbol{\Omega} = \operatorname{curl} \boldsymbol{v} + \boldsymbol{f}$, and $\boldsymbol{f} = f(x, y)\,\hat{\boldsymbol{z}}$ is twice the time independent angular velocity about the z-axis, with unit vector $\hat{\boldsymbol{z}}$ in the opposite direction than that of gravity. At the boundary $\boldsymbol{v} \cdot \hat{\boldsymbol{n}} = 0$ where $\hat{\boldsymbol{n}}$ is the outward unit normal on the boundary of the domain D for the fluid. It is easy to check that the *potential vorticity*,

$$q = \boldsymbol{\Omega} \cdot \nabla \rho, \tag{6.4}$$

is conserved along fluid trajectories, i.e. q satisfies (6.2) with ρ replaced by q. Consequently, the functional

$$C_F = \int_{\mathrm{D}} \mathrm{d}^3 x \, F(\rho, q) \tag{6.5}$$

is conserved for an arbitrary function $F(\rho, q)$. The energy

$$H(\boldsymbol{v}, \rho) = \int_{\mathrm{D}} \mathrm{d}^3 x \, (\tfrac{1}{2}\rho|\boldsymbol{v}|^2 + \rho g z) \tag{6.6}$$

is also conserved. The sense in which (6.1), (6.2) and (6.3) form a Hamiltonian system with energy (6.6) and Casimirs (6.5) will be discussed in the next section.

We form the functional $H_C = H + C_F$ and vary $\boldsymbol{v}$ and ρ. The variation of $\boldsymbol{v}$ will be taken in the space of $\boldsymbol{v}$'s with $\operatorname{div}(\boldsymbol{v}) = 0$. Also, we vary $\boldsymbol{\omega} = \operatorname{curl} \boldsymbol{v}$ in the space of vector fields that are curls; in particular $\delta\boldsymbol{\omega} = \operatorname{curl} \delta\boldsymbol{v}$. We have

$$H_C(\boldsymbol{v}, \rho) = \int_{\mathrm{D}} \mathrm{d}^3 x \, [\tfrac{1}{2}\rho|\boldsymbol{v}|^2 + \rho g z + F(\rho, q) + \lambda q], \tag{6.7}$$

so that the first variation for $\lambda = \text{const.}$ equals

$$\delta H_C = \int_{\mathrm{D}} \mathrm{d}^3 x \, [(\tfrac{1}{2}|\boldsymbol{v}|^2 + g z + F_\rho - \boldsymbol{\Omega} \cdot \nabla F_q)\, \delta\rho + (\rho \boldsymbol{v} - F_{qq} \nabla\rho \times \nabla q) \cdot \delta \boldsymbol{v}] + \oint_{\partial\mathrm{D}} \mathrm{d}S (F_q + \lambda) [\boldsymbol{\Omega} \cdot \hat{\boldsymbol{n}} \delta\rho - (\nabla\rho \times \delta\boldsymbol{v}) \cdot \hat{\boldsymbol{n}}]. \tag{6.8}$$

Since $\boldsymbol{v}_{\mathrm{e}} \cdot \hat{\boldsymbol{n}} = 0$, q_{e} and ρ_{e} are constant on the boundary. Thus the boundary term in (6.8) vanishes for $(F_q + \lambda)|_{\partial\mathrm{D}} = 0$ and stationarity of $H_C(\boldsymbol{v}, \rho)$ at $\boldsymbol{v}_{\mathrm{e}}$, ρ_{e} is achieved, provided $F(\rho_{\mathrm{e}}, q_{\mathrm{e}})$ satisfies

$$\rho_{\mathrm{e}} \boldsymbol{v}_{\mathrm{e}} = F_{qq}(\rho_{\mathrm{e}}, q_{\mathrm{e}}) \, \nabla\rho_{\mathrm{e}} \times \nabla q_{\mathrm{e}} \tag{6.9}$$

and

$$\tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2 + g z + F_\rho - \boldsymbol{\Omega}_{\mathrm{e}} \cdot \nabla F_q = 0. \tag{6.10}$$

Since (6.1), (6.2) and (6.3) imply

$$\partial/\partial t (\tfrac{1}{2}\rho|\boldsymbol{v}|^2 + \rho g z) = -\operatorname{div}(\rho\boldsymbol{v}(\tfrac{1}{2}|\boldsymbol{v}|^2 + p/\rho + g z)), \tag{6.11}$$

stationary solutions must satisfy

$$\boldsymbol{v}_{\mathrm{e}} \cdot \nabla\rho_{\mathrm{e}} = \boldsymbol{v}_{\mathrm{e}} \cdot \nabla q_{\mathrm{e}} = \boldsymbol{v}_{\mathrm{e}} \cdot \nabla K_{\mathrm{e}} = 0, \tag{6.12}$$

where

$$K_{\mathrm{e}} = \tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2 + p_{\mathrm{e}}/\rho_{\mathrm{e}} + g z. \tag{6.13}$$

This means that the three gradient vectors $\nabla\rho_{\mathrm{e}}$, ∇q_{e}, ∇K_{e} are perpendicular to the streamlines. Sufficient conditions for this are

$$\boldsymbol{v}_{\mathrm{e}} = V(\boldsymbol{x}) \, \nabla\rho_{\mathrm{e}} \times \nabla q_{\mathrm{e}}, \tag{6.14}$$

$$\tfrac{1}{2}|\boldsymbol{v}_{\mathrm{e}}|^2 + p_{\mathrm{e}}/\rho_{\mathrm{e}} + g z = K(\rho_{\mathrm{e}}, q_{\mathrm{e}}) = K_{\mathrm{e}}, \tag{6.15}$$

where V is some real valued function defined on the domain. A comparison of (6.14) with (6.9) yields $\rho_e V(\boldsymbol{x}) = F_{qq}(\rho_e, q_e)$. We now determine F in terms of K by inserting (6.9) in the Euler equation for stationary solutions. Comparing coefficients of $\nabla\rho_e$ and ∇q_e leads us to

$$\rho_e K_q(\rho_e, q_e)/q_e = F_{qq}(\rho_e, q_e) \tag{6.16}$$

and

$$K(\rho_e, q_e) + \rho_e K_\rho(\rho_e, q_e) = \tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz - F_{qq}\boldsymbol{\Omega}_e \cdot \nabla q_e. \tag{6.17}$$

We satisfy (6.16) with

$$\rho_e K(\rho_e, q_e) = q_e F_q - F + \zeta(\rho_e), \tag{6.18}$$

where ζ is an arbitrary function of ρ_e only. Inserting (6.18) into (6.17) yields

$$\tfrac{1}{2}|\boldsymbol{v}_e|^2 + gz + (F-\zeta)_\rho - \boldsymbol{\Omega}_e \cdot \nabla F_q = 0. \tag{6.19}$$

Renaming the difference $F(\rho, q) - \zeta(\rho)$ as $F(\rho, q)$ turns (6.19) into (6.10). Consequently, the new function F is related to K by

$$\rho K(\rho, q) = qF_q(\rho, q) - F(\rho, q). \tag{6.20}$$

The second variation of H_C in (6.7) when evaluated at $(\boldsymbol{v}_e, \rho_e)$ yields

$$\delta^2 H_C = \int_D \mathrm{d}^3x \left\{\rho_e|\delta\boldsymbol{v}|^2 + 2\delta\rho\,\delta\boldsymbol{v}\cdot\boldsymbol{v}_e + (\delta\rho, \delta q)\begin{bmatrix} F_{\rho\rho} & F_{q\rho} \\ F_{q\rho} & F_{qq}\end{bmatrix}\begin{bmatrix}\delta\rho \\ \delta q\end{bmatrix} + (F_q + \lambda)\,\delta^2 q\right\} \tag{6.21}$$

$$= \int_D \mathrm{d}^3x \left\{\rho_e|\delta\boldsymbol{v} + \rho_e^{-1}\boldsymbol{v}_e\,\delta\rho + \rho_e^{-1}\nabla F_q \times \nabla\delta\rho|^2 \right.$$
$$\left. + (\delta\rho, \delta q)\begin{bmatrix} F_{\rho\rho} & F_{q\rho} \\ F_{q\rho} & F_{qq}\end{bmatrix}\begin{bmatrix}\delta\rho \\ \delta q\end{bmatrix} - \rho_e^{-1}|\boldsymbol{v}_e\,\delta\rho + \nabla F_q \times \nabla\delta\rho|^2\right\} \tag{6.22}$$

It is somewhat complicated to display the full set of conditions that are both necessary and sufficient for $\delta^2 H_C$ to be positive definite. However, *sufficient* conditions for formal stability are given by setting $\nabla\delta\rho = \boldsymbol{k}\delta\rho$ in (6.22), so that

$$\delta^2 H_C = \int_D \mathrm{d}^2x \left\{\rho_e|\delta\boldsymbol{v} + \rho_e^{-1}\boldsymbol{v}_e\,\delta\rho + \rho_e^{-1}\nabla F_q \times \nabla\delta\rho\,|^2 \right.$$
$$\left. + (\delta\rho, \delta q)\begin{bmatrix} F_{\rho\rho} - \rho_e^{-1}|\boldsymbol{v}_e + \nabla F_q \times \boldsymbol{k}|^2 & F_{\rho q} \\ F_{\rho q} & F_{qq}\end{bmatrix}\begin{pmatrix}\delta\rho \\ \delta q\end{pmatrix}\right\}. \tag{6.23}$$

Sufficient conditions for $\delta^2 H_C$ in (6.23) to be positive definite are

$$F_{qq} > 0 \tag{6.24a}$$

and

$$|\boldsymbol{v}_e + \nabla F_q \times \boldsymbol{k}\,|^2 > \rho_e[F_{\rho\rho}F_{qq} - F_{\rho q}^2]/F_{qq} > 0. \tag{6.24b}$$

In particular, (6.24*b*) requires $|\boldsymbol{k}| < |\boldsymbol{v}_e|/|\nabla F_q(\rho_e, q_e)|$.

Nonlinear stability is determined by examining the conserved quantity,

$$\hat{H}_C(\Delta\boldsymbol{v}, \Delta\rho) = H_C(\boldsymbol{v}_e + \Delta\boldsymbol{v}, \rho_e + \Delta\rho) - H_C(\boldsymbol{v}_e, \rho_e) - \mathrm{D}H_C(\boldsymbol{v}_e, \rho_e)\cdot(\Delta\boldsymbol{v}, \Delta\rho) \tag{6.25}$$

$$= \int_D \mathrm{d}^3x\,\{\tfrac{1}{2}(\rho_e + \Delta\rho)(|\boldsymbol{v}_e|^2 + 2\boldsymbol{v}_e\cdot\Delta\boldsymbol{v} + |\Delta\boldsymbol{v}|^2) - \tfrac{1}{2}\rho_e|\boldsymbol{v}_e|^2 - \tfrac{1}{2}|\boldsymbol{v}_e|^2\Delta\rho - \rho_e\boldsymbol{v}_e\cdot\Delta\boldsymbol{v}$$
$$+ F(\rho_e + \Delta\rho, q_e + \Delta q) - F(\rho_e, q_e) + \lambda\Delta q - F_\rho(\rho_e, q_e)\,\Delta\rho$$
$$- [F_q(\rho_e, q_e) + \lambda](\Delta q - \Delta\boldsymbol{\Omega}\cdot\nabla\Delta\rho)\} \tag{6.26}$$

$$= \int_D d^3x \{\tfrac{1}{2}(\rho_e+\Delta\rho)|\Delta \boldsymbol{v}|^2 + \Delta\rho \boldsymbol{v}_e \cdot \Delta \boldsymbol{v} + F(\rho_e+\Delta\rho, q_e+\Delta q) - F(\rho_e, q_e)$$
$$- F_\rho(\rho_e, q_e)\,\Delta\rho - F_q(\rho_e, q_e)\,\Delta q + [F_q(\rho_e, q_e)+\lambda]\,\Delta\boldsymbol{\Omega}\cdot\nabla\Delta\rho\} \tag{6.27}$$

(see Appendix A). Integrating the last term in (6.27) by parts and completing a square gives

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) = \int_D d^3x \{\tfrac{1}{2}(\rho_e+\Delta\rho)[\,|\Delta\boldsymbol{v} + (\boldsymbol{v}_e\,\Delta\rho + \nabla F_q \times \nabla\Delta\rho)/(\rho_e+\Delta\rho)|^2$$
$$- (\boldsymbol{v}_e\,\Delta\rho + \nabla F_q \times \nabla\Delta\rho)/(\rho_e+\Delta\rho)^2] + \hat{F}(\Delta\rho, \Delta q)\}, \tag{6.28}$$

where

$$\hat{F}(\Delta\rho, \Delta q) := F(\rho_e+\Delta\rho, q_e+\Delta q) - F(\rho_e, q_e) - F_\rho(\rho_e, q_e)\,\Delta\rho - F_q(\rho_e, q_e)\,\Delta q. \tag{6.29}$$

Then, defining

$$\frac{\Delta \boldsymbol{M}}{\rho_e+\Delta\rho} := \frac{(\rho_e+\Delta\rho)(\boldsymbol{v}_e+\Delta\boldsymbol{v}) - \rho_e \boldsymbol{v}_e}{\rho_e+\Delta\rho} = \Delta\boldsymbol{v} + \frac{\boldsymbol{v}_e\,\Delta\rho}{\rho_e+\Delta\rho} \tag{6.30}$$

gives

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) = \int_D d^3x \left\{\frac{1}{2(\rho_e+\Delta\rho)}[\,|\Delta\boldsymbol{M} + \nabla F_q \times \nabla\Delta\rho|^2 - |\boldsymbol{v}_e\,\Delta\rho + \nabla F_q \times \nabla\Delta\rho|^2] + \hat{F}(\Delta\rho, \Delta q)\right\}. \tag{6.31}$$

Since ρ is conserved, if ρ is bounded above and below throughout the domain initially, it will remain so for all time. Thus, we can assume

$$0 < \rho_{\min} \leqslant \rho \leqslant \rho_{\max} < \infty. \tag{6.32}$$

We then note, on replacing $\nabla\Delta\rho$ by $\boldsymbol{k}\Delta\rho$ as before, that by (6.31),

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) \geqslant \int_D d^3x \{|\Delta\boldsymbol{M} + \nabla F_q \times \nabla\Delta\rho|^2/2\rho_{\max}$$
$$- |\boldsymbol{v}_e + \nabla F_q \times \boldsymbol{k}|^2 (\Delta\rho)^2/2\rho_{\min} + \hat{F}(\Delta\rho, \Delta q)\}. \tag{6.33}$$

Next, we define

$$\Psi(\rho, q) := F(\rho, q) - |\boldsymbol{v}_e + \nabla F_q \times \boldsymbol{k}|^2 \rho^2/2\rho_{\min}, \tag{6.34}$$

as in §3, compare with (3.40). Consequently,

$$\begin{aligned}
\Psi_{\rho\rho}(\rho_e, q_e) &= F_{\rho\rho}(\rho_e, q_e) - |\boldsymbol{v}_e + \nabla F_q \times \boldsymbol{k}|^2/\rho_{\min} \\
&\geqslant F_{\rho\rho}(\rho_e, q_e) - (|\boldsymbol{v}_e| + |\nabla F_q||\boldsymbol{k}|)^2/\rho_{\min}, \\
\Psi_{qq}(\rho_e, q_e) &= F_{qq}(\rho_e, q_e), \\
\Psi_{\rho q}(\rho_e, q_e) &= F_{\rho q}(\rho_e, q_e),
\end{aligned} \tag{6.35}$$

for the partial derivatives of Ψ, and the $H_{\hat{C}}$ inequality (6.33) becomes

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) \geqslant \int d^3x \{|\Delta\boldsymbol{M} + \nabla F_q \times \nabla\Delta\rho|^2/2\rho_{\max} + \hat{\Psi}(\Delta\rho, \Delta q)\}, \tag{6.36}$$

where $\hat{\Psi}$ is defined as in (6.29), upon replacing F by Ψ. If Ψ satisfies the convexity conditions

$$0 < \alpha \leqslant \Psi_{qq}(\rho_e, q_e) < \infty \tag{6.37}$$

and
$$0 < \begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}^t \begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix} \leqslant \begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}^t \begin{bmatrix}\Psi_{\rho\rho}(\rho_e, q_e) & \Psi_{\rho q}(\rho_e, q_e) \\ \Psi_{\rho q}(\rho_e, q_e) & \Psi_{qq}(\rho_e, q_e)\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix} < \infty, \tag{6.38}$$

where α, β, γ are finite positive constants satisfying $\alpha > 0$, $\gamma\alpha - \beta^2 > 0$, then we conclude

$$\begin{aligned}\hat{H}_C(\Delta \boldsymbol{v}(\boldsymbol{x}, t), \Delta\rho(\boldsymbol{x}, t)) &= \hat{H}_C(\Delta \boldsymbol{v}(\boldsymbol{x}, 0), \Delta\rho(\boldsymbol{x}, 0)) \\ &\geqslant \int_D \mathrm{d}^3x \left\{\frac{1}{2\rho_{\max}} |\Delta \boldsymbol{M} + \nabla F_q \times \nabla\Delta\rho|^2 + (\Delta\rho, \Delta q)\begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}\right\} > 0. \end{aligned} \tag{6.39}$$

The second inequality in (6.38) determines the maximum value of $|\boldsymbol{k}| = |\Delta\nabla\rho|/|\Delta\rho|$ from

$$(|\boldsymbol{v}_e| + |\nabla F_q||\boldsymbol{k}|)^2 \leqslant \frac{F_{\rho\rho}F_{qq} - F_{\rho q}^2 - (\gamma\alpha - \beta^2)}{F_{qq}/\rho_{\min}}, \tag{6.40}$$

in terms of equilibrium state quantities. So long as (6.40) is satisfied, relation (6.39) will provide a lower bound for $\hat{H}_C$ in terms of a norm on the space of triples $(\Delta \boldsymbol{M}, \Delta\rho, \Delta q)$ given by

$$\|(\Delta \boldsymbol{M}, \Delta\rho, \Delta q)\|^2 = \int_D \mathrm{d}^3x \left\{\frac{1}{2\rho_{\max}} |\Delta \boldsymbol{M} + \nabla F_q \times \nabla\Delta\rho|^2 + (\Delta\rho, \Delta q)\begin{bmatrix}\gamma & \beta \\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\rho \\ \Delta q\end{bmatrix}\right\}. \tag{6.41}$$

Let us also assume that $\hat{H}_C \to 0$ as $(\Delta \boldsymbol{M}, \Delta\rho, \Delta q) \to 0$. For example, this is satisfied if

$$\Psi_{qq}(\rho_e, q_e) \leqslant \bar{\alpha} < \infty, \tag{6.42}$$

$$\Psi_{qq}(\rho_e, q_e)\,\Psi_{\rho\rho}(\rho_e, q_e) - \Psi_{\rho q}^2(\rho_e, q_e) \leqslant \bar{\alpha}\bar{\gamma} - \bar{\beta}^2 < \infty, \tag{6.43}$$

for finite positive constants $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$. Then

$$\hat{H}_C(\Delta \boldsymbol{v}, \Delta\rho) \leqslant Q\|(\Delta \boldsymbol{M}, \Delta\rho, \Delta q)\|^2, \tag{6.44}$$

for a positive constant Q. If the above hypotheses hold, then the equilibrium state $(\boldsymbol{v}_e, \rho_e)$ is nonlinearly stable in the norm (6.41). Indeed, $\hat{H}_C$ satisfies

$$\begin{aligned}\|(\Delta \boldsymbol{M}(\boldsymbol{x}, t), \Delta\rho(\boldsymbol{x}, t), \Delta q(\boldsymbol{x}, t))\|^2 &\leqslant \hat{H}_C(\Delta \boldsymbol{v}(\boldsymbol{x}, t), \Delta\rho(\boldsymbol{x}, t)) \\ &= \hat{H}_C(\Delta \boldsymbol{v}(\boldsymbol{x}, 0), \Delta\rho(\boldsymbol{x}, 0)) \\ &\leqslant Q\|(\Delta \boldsymbol{M}(\boldsymbol{x}, 0), \Delta\rho(\boldsymbol{x}, 0), \Delta q(\boldsymbol{x}, 0))\|^2. \end{aligned} \tag{6.45}$$

In summary, we have proved the following.

THEOREM: *Given an equilibrium flow* $(\boldsymbol{v}_e, \rho_e)$ *with constant density on each component of the boundary, we form* Ψ *in* (6.34) *by solving* (6.20). *If* (6.37), (6.38), (6.42) *and* (6.43) *hold, then* $(\boldsymbol{v}_e, \rho_e)$ *is nonlinearly stable in the norm* (6.41), *so long as* (6.40) *continues to hold.*

7. Hamiltonian structure and Casimirs

As explained in Arnold (1978) (Appendix 2 and references therein) and Ebin & Marsden (1970), the hydrodynamics of an incompressible, inviscid, constant density fluid can be understood as motion along geodesics in the group of volume-preserving diffeomorphisms of the domain of fluid flow. The same situation for stratified incompressible flows was found in Marsden (1976). (The metric used to define the geodesics in that case depends on the initial density). The principle of least action in any of these cases implies that the motion of the fluid is described by geodesics in the metric given by the kinetic energy. This geodesic motion is a manifestation of the Hamiltonian structure of the fluid theory written in material (Lagrangian) coordinates. Elimination, for example, via symmetry methods of the particle labels results in the noncanonical Hamiltonian structure (or Poisson bracket structure) for ideal fluids in terms of the velocity written in spatial (Eulerian) coordinates. In the case of spatially constant density taken to be unity, the Euler equations are

$$\begin{aligned} \partial \boldsymbol{v}/\partial t + (\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} &= -\nabla p, \\ \operatorname{div}(\boldsymbol{v}) &= 0, \\ \boldsymbol{v}\cdot\hat{\boldsymbol{n}} &= 0, \\ \boldsymbol{v}(\boldsymbol{x},0) &= \text{given function on D}, \end{aligned} \tag{7.1}$$

where $\hat{\boldsymbol{n}}$ is the outward unit normal to the boundary ∂D. The pressure is determined from $\boldsymbol{v}$ by $\nabla^2 p = -\operatorname{div}(\boldsymbol{v}\cdot\nabla\boldsymbol{v})$ and $\partial p/\partial n = -\hat{\boldsymbol{n}}\cdot[(\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}]$. Equations (7.1) are Hamiltonian, i.e. for any real valued function of a real variable F, one has

$$\partial/\partial t\, F(\boldsymbol{v}) = \{F, H\}(\boldsymbol{v}), \tag{7.2}$$

with Poisson bracket (essentially due to Arnold (1966, 1967))

$$\{F, G\}(\boldsymbol{v}) = \int_{\mathrm{D}} \boldsymbol{v}\cdot\left[\left(\frac{\partial G}{\partial \boldsymbol{v}}\cdot\nabla\right)\frac{\partial F}{\partial \boldsymbol{v}} - \left(\frac{\partial F}{\partial \boldsymbol{v}}\cdot\nabla\right)\frac{\partial G}{\partial \boldsymbol{v}}\right]\mathrm{d}^3x, \tag{7.3}$$

and Hamiltonian

$$H(\boldsymbol{v}) = \frac{1}{2}\int_{\mathrm{D}} |\boldsymbol{v}|^2\,\mathrm{d}^3x, \tag{7.4}$$

where the functional derivative $\delta F/\delta \boldsymbol{v}$, is the divergence-free vector field defined by

$$\mathrm{D}F(\boldsymbol{v})\cdot\delta\boldsymbol{v} = \int_{\mathrm{D}} \frac{\delta F}{\delta \boldsymbol{v}}\cdot\delta\boldsymbol{v}\,\mathrm{d}^3x =: \left\langle \frac{\delta F}{\delta \boldsymbol{v}}, \delta\boldsymbol{v}\right\rangle \tag{7.5}$$

for all divergence-free vector-fields $\delta\boldsymbol{v}$. Let $\mathscr{X}(\mathrm{D})$ denote the space of all vector fields on D and $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ the divergence-free vector fields that are tangent to the boundary. Formula (7.5) uses the weakly non-degenerate pairing

$$\begin{aligned} \langle\,,\rangle &: \mathscr{X}_{\mathrm{div}}(\mathrm{D})\times\mathscr{X}_{\mathrm{div}}(\mathrm{D})\to\mathbb{R} \\ \langle \boldsymbol{u}, \boldsymbol{v}\rangle &:= \int_{\mathrm{D}} \boldsymbol{u}\cdot\boldsymbol{v}\,\mathrm{d}^3x \end{aligned} \tag{7.6}$$

which identifies $\mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$, the dual of $\mathscr{X}_{\mathrm{div}}(\mathrm{D}) := \{\boldsymbol{u}\in\mathscr{X}(\mathrm{D})\,|\,\operatorname{div}(\boldsymbol{u}) = 0,\ \boldsymbol{u}\text{ tangent to }\partial\mathrm{D}\}$ with itself. Thus, the Eulerian velocity field $\boldsymbol{v}$ in (7.1) is thought of as an element of $\mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$. A direct proof that the Euler equations (7.1) are Hamiltonian with respect to the Poisson bracket (7.3) with Hamiltonian (7.4) proceeds as follows. If $\dot{\boldsymbol{v}}$ denotes $\partial\boldsymbol{v}/\partial t$, we get, on the one hand,

$$\partial F(\boldsymbol{v})/\partial t = \mathrm{D}F(\boldsymbol{v})\cdot\dot{\boldsymbol{v}} = \langle \partial F/\partial \boldsymbol{v}, \dot{\boldsymbol{v}}\rangle \tag{7.7}$$

and on the other, since $\delta H/\delta \boldsymbol{v} = \boldsymbol{v}$,

$$\partial F(\boldsymbol{v})/\partial t = \{F, H\}(\boldsymbol{v}) \tag{7.8}$$

$$= \int_{\mathrm{D}} [\boldsymbol{v}\cdot(\boldsymbol{v}\cdot\nabla)\,\delta F/\delta\boldsymbol{v} - \boldsymbol{v}\cdot(\delta F/\delta\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}]\,\mathrm{d}^3x \tag{7.9}$$

$$= \int_{\mathrm{D}} [-(\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}\cdot\delta F/\delta\boldsymbol{v} - (\delta F/\delta\boldsymbol{v}\cdot\nabla)\,(\tfrac{1}{2}|\boldsymbol{v}|^2)]\,\mathrm{d}^3x. \tag{7.10}$$

By the Helmholtz (or Hodge) decomposition theorem, any vector field can be uniquely written as an L^2 orthogonal sum of a divergence-zero vector field parallel to the boundary and the gradient of a function. Let $\boldsymbol{P}$ denote the projection of a vector field onto its divergence-free part parallel to the boundary. Then the above integral becomes

$$\int \boldsymbol{P}((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v})\cdot\delta F/\delta\boldsymbol{v}\,\mathrm{d}^3x, \tag{7.11}$$

so we get from (7.7) and (7.11)

$$\dot{\boldsymbol{v}} = -\boldsymbol{P}((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}). \tag{7.12}$$

Now writing, again by the Helmholtz decomposition, $\boldsymbol{v} = \boldsymbol{u} - \nabla p$, with $\operatorname{div}(\boldsymbol{u}) = 0$, $\boldsymbol{u}$ parallel to $\partial\mathrm{D}$, we get

$$\nabla^2 p = -\operatorname{div}((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}), \quad \partial p/\partial n = -\hat{\boldsymbol{n}}\cdot(\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}, \tag{7.13}$$

i.e. p is the pressure, since it satisfies the appropriate Neumann problem. Thus, $\boldsymbol{P}((\boldsymbol{v}\cdot\nabla)\cdot\boldsymbol{v}) = (\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} + \nabla p$ and (7.12) becomes

$$\partial\boldsymbol{v}/\partial t + (\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} = -\nabla p, \tag{7.13$'$}$$

which is Euler's equation (7.1).

The Poisson bracket (7.3) is a special case of a general bracket, called the Lie–Poisson bracket, which operates on real-valued functions defined on the dual of any Lie algebra. The general construction will be reviewed in Appendix B. It is natural to hope that Lie–Poisson brackets and the Lie group theory underlying it can be extended to other fluid theories, as additional physical effects are taken into account, such as stratification, compressibility, magnetic fields, and so forth. Indeed this is possible, and the Hamiltonian structures for a variety of fluid theories have been systematically uncovered and associated with Lie algebras and Lie groups; see, for example, Holm & Kupershmidt (1983), Marsden & Weinstein (1983), Marsden *et al.* (1983), and Marsden *et al.* (1984*a*, *b*), for examples and earlier references. We hope that understanding the Lie algebraic nature of these Poisson brackets will lead to new perspectives and analytic tools for the investigation of ideal fluid dynamics, as in the study of stability.

The purpose of this section is to explain the Hamiltonian structure of stratified Boussinesq and Euler flow in two and three dimensions. First, we shall just quote the Poisson bracket for three-dimensional stratified Boussinesq flow. We then study the Hamiltonian structure of three-dimensional stratified Euler flow in a systematic manner. This explanation takes quite a bit of space, but we have included all the details needed to facilitate understanding. As in the case of homogeneous flow, this approach is based on the map from Lagrangian to Eulerian quantities. The explanation of the crucial properties of the Lagrangian to Eulerian map is provided in a short summary in Appendix B.

The Poisson brackets for two-dimensional Boussinesq and Euler flow will first be deduced from the three-dimensional brackets by restriction to a plane, and then discussed from an alternative viewpoint using the more traditional method of Clebsch variables.

To help the reader interested primarily in the results, we mention that all the brackets in terms of velocity and density are the same for the systems we consider and are given in (7.15) and (7.62). For the two-dimensional systems, there is also a bracket (7.69) in terms of vorticity and density. In each case, the Hamiltonian equations of motion have the form $\dot{F} = \{F, H\}$, where F is a functional of the physical dynamical variables. Even though the stability results do not explicitly require the Hamiltonian approach explained here, this geometrical structure underlies the success of the energy-Casimir convexity method. To facilitate the presentation, in the three-dimensional cases we set $\boldsymbol{f} = 0$ (that is, no rotation), and in the Boussinesq cases $\rho_* = 1$. The brackets we present retain their form when velocity is measured relative to a rotating frame and the Hamiltonian is appropriately altered. (One could also keep the same Hamiltonian, but modify the brackets.)

(a) *Three-dimensional stratified Boussinesq flow*

We consider the equations of motion (3.1) with $\rho_* = 1$ and $\boldsymbol{f} = 0$ and seek a bracket analogous to (7.3) for which these equations are Hamiltonian, with Hamiltonian function given

$$H(\boldsymbol{v}, \rho) = \int_{\mathrm{D}} \mathrm{d}^3x\left(\tfrac{1}{2}|\boldsymbol{v}|^2 + \rho g z\right). \tag{7.14}$$

The Poisson bracket for which (7.1) is Hamiltonian has the following expression

$$\{F, G\}(\boldsymbol{v}, \rho) = \int_{\mathrm{D}} \boldsymbol{v}\cdot\left[\left(\frac{\delta G}{\delta \boldsymbol{v}}\cdot\nabla\right)\frac{\delta F}{\delta \boldsymbol{v}} - \left(\frac{\delta F}{\delta \boldsymbol{v}}\cdot\nabla\right)\frac{\delta G}{\delta \boldsymbol{v}}\right]\mathrm{d}^3x + \int_{\mathrm{D}} \rho\left[\frac{\delta G}{\delta \boldsymbol{v}}\cdot\nabla\frac{\delta F}{\delta \rho} - \frac{\delta F}{\delta \boldsymbol{v}}\cdot\nabla\frac{\delta G}{\delta \rho}\right]\mathrm{d}^3x, \tag{7.15}$$

where the functional derivative with respect to ρ is defined using the L^2-pairing, i.e. if $\mathrm{D}_\rho F$ denotes the partial Fréchet derivative of $F: \mathscr{X}_{\mathrm{div}}(\mathrm{D}) \times \mathscr{F}(\mathrm{D}) \rightarrow \mathbb{R}$, then $\delta F/\delta\rho \in \mathscr{F}(\mathrm{D})$ is the unique function on D satisfying

$$\mathrm{D}_\rho F(\rho)\cdot\delta\rho = \int_{\mathrm{D}} \frac{\delta F}{\delta \rho}\,\delta\rho\,\mathrm{d}^3x =: \left\langle\frac{\delta F}{\delta\rho}, \delta\rho\right\rangle \tag{7.16}$$

for any $\delta\rho \in \mathscr{F}(\mathrm{D})$. The proof that the equations (3.1) are equivalent to $\dot{F} = \{F, H\}$ follows the steps indicated before. First, by taking the divergence of (3.1) and the dot product with $\hat{\boldsymbol{n}}$, we get the Neumann problem

$$\nabla^2 p = -\operatorname{div}\left((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}\right) - g\partial\rho/\partial z, \tag{7.17}$$

$$\partial p/\partial n = -\hat{\boldsymbol{n}}\cdot\left((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v}\right) - g\rho\hat{\boldsymbol{n}}\cdot\nabla z, \tag{7.18}$$

which determines p up to a constant. Second, since $\delta H/\delta\boldsymbol{v} = \boldsymbol{v}$ and $\delta H/\delta\rho = gz$, the Poisson bracket of F and H becomes

$$\{F, H\} = \int_{\mathrm{D}} \boldsymbol{v}\cdot\left[(\boldsymbol{v}\cdot\nabla)\frac{\delta F}{\delta \boldsymbol{v}} - \left(\frac{\delta F}{\delta \boldsymbol{v}}\cdot\nabla\right)\boldsymbol{v}\right]\mathrm{d}^3x + \int_{\mathrm{D}} \rho\left[\boldsymbol{v}\cdot\nabla\frac{\delta F}{\delta\rho} - \frac{\delta F}{\delta \boldsymbol{v}}\cdot\nabla gz\right]\mathrm{d}^3x \tag{7.19}$$

$$= \int_{\mathrm{D}} \left[(\boldsymbol{v}\cdot\nabla)\boldsymbol{v}\cdot\frac{\delta F}{\delta \boldsymbol{v}} - \left(\frac{\delta F}{\delta \boldsymbol{v}}\cdot\nabla\right)\left(\tfrac{1}{2}|\boldsymbol{v}|^2\right) - (\boldsymbol{v}\cdot\nabla\rho)\frac{\delta F}{\delta\rho} - \frac{\delta F}{\delta \boldsymbol{v}}\cdot\rho g\nabla z\right]\mathrm{d}^3x \tag{7.20}$$

$$= \int\left[-\boldsymbol{P}((\boldsymbol{v}\cdot\nabla)\,\boldsymbol{v} + g\rho\nabla z)\cdot\frac{\delta F}{\delta \boldsymbol{v}} - (\boldsymbol{v}\cdot\nabla\rho)\,\frac{\delta F}{\delta \boldsymbol{v}}\right]\mathrm{d}^3x, \tag{7.21}$$

where, as before, $\boldsymbol{P}$ is the projection of a vector field onto its divergence-free part tangent to the boundary as given by the Helmholtz–Hodge decomposition. Third, applying the definitions of the pairings (7.6) and (7.16) we arrive at

$$\frac{\partial}{\partial t} F(\boldsymbol{v}, \rho) = \mathrm{D}_{\boldsymbol{v}} F(\boldsymbol{v}, \rho) \cdot \dot{\boldsymbol{v}} + \mathrm{D}_{\rho} F(\boldsymbol{v}, \rho) \cdot \dot{\rho} \tag{7.22}$$

$$= \left\langle \frac{\delta F}{\delta \boldsymbol{v}}, \dot{\boldsymbol{v}} \right\rangle + \left\langle \frac{\delta F}{\delta \rho}, \dot{\rho} \right\rangle . \tag{7.23}$$

Thus the Hamilton equations become, by comparing (7.21) with (7.23)

$$\dot{\boldsymbol{v}} + \boldsymbol{P}((\boldsymbol{v} \cdot \nabla)\, \boldsymbol{v} + g\rho \nabla z) = 0 \tag{7.24}$$

$$\dot{\rho} + \boldsymbol{v} \cdot \nabla \rho = 0. \tag{7.25}$$

Equation (7.25) is already identical to the density advection equation, whereas equation (7.24) will become the momentum equation. For this, we write again by the Helmholtz decomposition

$$(\boldsymbol{v} \cdot \nabla)\, \boldsymbol{v} + g\rho \nabla z = \boldsymbol{u} - \nabla p. \tag{7.26}$$

Taking the divergence and the dot product with $\hat{\boldsymbol{n}}$ of this relation, yields the Neumann problem (7.17) and (7.18) and, thus, p is the pressure. Moreover, by (7.26),

$$\boldsymbol{u} = \boldsymbol{P}((\boldsymbol{v} \cdot \nabla)\, \boldsymbol{v} + g\rho \nabla z) = (\boldsymbol{v} \cdot \nabla)\, \boldsymbol{v} + g\rho \nabla z + \nabla p, \tag{7.27}$$

so equation (7.24) becomes the momentum equation in (3.1).

Remark. Formula (7.15) is the Lie–Poisson bracket on the dual of the semidirect product Lie algebra $\mathscr{X}_{\mathrm{div}}(\mathrm{D}) \circledS \mathscr{F}(\mathrm{D})$, where $\mathscr{F}(\mathrm{D})$ denotes the smooth functions on D, and the action of the first factor on the second is minus the Lie derivative. (See Appendix B).

(*b*) *Three-dimensional stratified Euler flow*

Let D be a region in Euclidean three space with smooth boundary ∂D, filled with an inviscid incompressible inhomogeneous fluid, moving under the influence of gravity. We assume that the motion of the fluid is such that the velocity field and density are, at least once, continuously differentiable. We shall denote by capital letters $\boldsymbol{V}, \boldsymbol{X}, \ldots$, quantities written in the *material* (or *Lagrangian*) picture.

The configuration space of the fluid motion is determined in the following way. A given fluid particle that was at the point $\boldsymbol{X}$ at a time $t = 0$ will occupy a position $\eta_t(\boldsymbol{X}) = \boldsymbol{x}(\boldsymbol{X}, t)$ at time t, called the *spatial*, or *Eulerian* position. Since no two distinct fluid elements can occupy the same position and since cavitation is excluded, the map $\eta_t\colon \mathrm{D} \to \mathrm{D}$ is required to be one-to-one and onto. We shall also require certain differentiability hypotheses on η_t and its inverse η_t^{-1} to ensure that the Eulerian velocity is, at least once, continuously differentiable; we refer the reader to Marsden (1976) for the correct choice of Sobolev, or Hölder differentiability classes. Thus, the motion of the fluid is completely characterized by the set of diffeomorphisms η_t. Since the fluid is assumed to be incompressible, the diffeomorphisms η_t cannot alter the volume element d^3X of D, i.e. $\eta_t^*(\mathrm{d}^3X) = \mathrm{d}^3X$, or $J_{\eta_t} = 1$, where η^* denotes the pull-back operation and J_{η_t} is the determinant of the Jacobian matrix $\mathrm{d}\boldsymbol{x}/\mathrm{d}\boldsymbol{X}$ of η_t. Since diffeomorphisms preserve boundaries, we have $\eta_t(\partial \mathrm{D}) = \partial \mathrm{D}$.

Given the mass density $\rho_0(\boldsymbol{X})$ in Lagrangian coordinates, the conservation of mass equation

(5.3) is equivalent to

$$\rho(\boldsymbol{x}, t) = \rho_0(\boldsymbol{X}) \quad \text{or} \quad \eta_t^*(\rho_t(\boldsymbol{x})\, \mathrm{d}^3x) = \rho_0(\boldsymbol{X})\, \mathrm{d}^3X, \tag{7.28}$$

where $\rho_0(\boldsymbol{x}, t) = \rho_t(\boldsymbol{x})$ is the density in Eulerian coordinates at time t and $\rho_0(\boldsymbol{X})$ the density in Lagrangian coordinates. Relations (7.7) imply that once the Lagrangian density $\rho_0(\boldsymbol{X})$ is given (an initial condition), its subsequent change in Eulerian coordinates is completely determined by the motion. Hence, *the configuration space of incompressible inhomogeneous fluid flow with a given mass density in the reference configuration is the group of volume preserving diffeomorphisms* $\mathrm{Diff_{vol}(D)}$ *of* D. *Consequently, we choose the phase space to be the cotangent bundle* $\mathrm{T^*(Diff_{vol}(D))}$.

To fix notation, let us recall the definitions of Lagrangian and Eulerian velocity. For a motion $\boldsymbol{x}(\boldsymbol{X}, t) = \eta_t(\boldsymbol{X})$, the quantity

$$\boldsymbol{V}(\boldsymbol{X}, t) := \boldsymbol{V}_t(\boldsymbol{X}) := \partial\eta(\boldsymbol{X}, t)/\partial t \tag{7.29}$$

is the *material* or *Lagrangian velocity*, whereas

$$\boldsymbol{v}(\boldsymbol{x}, t) := \boldsymbol{v}_t(\boldsymbol{x}) := \boldsymbol{V}(\boldsymbol{X}, t) \tag{7.30}$$

is the *spatial*, or *Eulerian velocity*. Note that the relation between $\boldsymbol{V}$ and $\boldsymbol{v}$ is given by

$$\boldsymbol{v}_t \circ \eta_t = \boldsymbol{V}_t \tag{7.31}$$

and that both $\boldsymbol{V}_t$ and $\boldsymbol{v}_t$ are tangential to D at $\boldsymbol{x} = \eta_t(\boldsymbol{X})$. This means that $\boldsymbol{v}_t$ is a standard, time-dependent, vector field on D, whereas $\boldsymbol{V}_t$ is a *vector field over* η_t *on* D. This concept is quite common and shows up in the geometric structure of $\mathrm{Diff_{vol}(D)}$ to which we turn next.

If D is compact, the group $\mathrm{Diff_{vol}(D)}$ is made into a Lie group in two inequivalent ways: by taking the C^∞ Fréchet space topology (Leslie (1967)), or by taking inverse limits of Sobolev, or Hölder class topologies (Ebin & Marsden (1970)). For the purposes of this paper, it suffices to know that $\mathrm{Diff_{vol}(D)}$ is a Lie group in the sense that the usual Lie group operations are allowed.

We first determine the tangent space $\mathrm{T}_\eta(\mathrm{Diff_{vol}(D)})$ at η. Let $t \mapsto \eta_t$ be a smooth curve with $\eta_0 = \eta$. Then $(\mathrm{d}\eta_t/\mathrm{d}t)|_{t=0}$ is, by definition, a tangent vector at η to $\mathrm{Diff_{vol}(D)}$. If $\boldsymbol{X} \in \mathrm{D}$, then $t \mapsto \eta_t(\boldsymbol{X})$ is a smooth curve in D through $\eta(\boldsymbol{X})$ and, thus,

$$(\mathrm{d}\eta_t(\boldsymbol{X})/\mathrm{d}t)|_{t=0} \in \mathrm{T}_{\eta(\boldsymbol{X})}\mathrm{D}, \tag{7.32}$$

where $\mathrm{T}_{\eta(\boldsymbol{X})}\mathrm{D}$ is the tangent space to D at $\eta(\boldsymbol{X})$. Moreover, if $\boldsymbol{X} \in \partial\mathrm{D}$, then $(\mathrm{d}\eta_t(\boldsymbol{X})/\mathrm{d}t)|_{t=0}$ is tangent to $\partial\mathrm{D}$ at $\eta(\boldsymbol{X})$, since η_t preserves $\partial\mathrm{D}$. Consequently, we have a map $\boldsymbol{X} \in \mathrm{D} \mapsto (\mathrm{d}\eta_t(\boldsymbol{X})/\mathrm{d}t)|_{t=0} \in \mathrm{T}_{\eta(\boldsymbol{X})}\mathrm{D}$, i.e. $(\mathrm{d}\eta_t/\mathrm{d}t)|_{t=0}$ is a vector field over η tangent to $\partial\mathrm{D}$. We can now define $\boldsymbol{V}_\eta$ to be a *vector field over* η, if $\boldsymbol{V}_\eta$ is a smooth map from D to the tangent bundle TD such that $\boldsymbol{V}_\eta(\boldsymbol{X})$ is a tangent vector at $\eta(\boldsymbol{X})$. Thus, in particular, if $\eta = e =$ identity, $\mathrm{T_e(Diff_{vol}(D))}$ consists of vector fields on D which are tangent to $\partial\mathrm{D}$. However, each of our diffeomorphisms η_t is volume preserving, so that if $t \mapsto \eta_t$ is a curve through the identity, by the relation between Lie derivatives and flows, we have

$$0 = \mathrm{d}/\mathrm{d}t|_{t=0}\, \eta_t^*(\mathrm{d}^3X) = \mathrm{L}_{\boldsymbol{V}}(\mathrm{d}^3X) = \mathrm{div}\,(\boldsymbol{V})\, \mathrm{d}^3X, \tag{7.33}$$

where $\boldsymbol{V} = (\mathrm{d}\eta_t/\mathrm{d}t)|_{t=0}$, $\mathrm{L}_{\boldsymbol{V}}$ is the Lie derivative and $\mathrm{div}\,(\boldsymbol{V})$ is the divergence of $\boldsymbol{V}$ with respect to the volume element d^3X. Thus, $\mathrm{T_e(Diff_{vol}(D))} := \mathscr{X}_{\mathrm{div}}(\mathrm{D}) := \{\boldsymbol{V}: \mathrm{D} \to \mathrm{TD} \mid \boldsymbol{V}$ is a vector field on D tangent to $\partial\mathrm{D}$ and $\mathrm{div}\,(\boldsymbol{V}) = 0\}$. Now, if $t \mapsto \eta_t$ is a curve such that $\eta_0 = \eta$, applying the prior reasoning to $\eta_t \circ \eta^{-1}$ and letting $\boldsymbol{V}_\eta = (\mathrm{d}\eta_t/\mathrm{d}t)|_{t=0}$, we get $\mathrm{T}_\eta(\mathrm{Diff_{vol}(D)}) := \{\boldsymbol{V}_\eta: \mathrm{D} \to \mathrm{TD} \mid \boldsymbol{V}_\eta(\boldsymbol{X}) \in \mathrm{T}_{\eta(\boldsymbol{X})}\mathrm{D}$, $\boldsymbol{V}_\eta$ is tangential to $\partial\mathrm{D}$ and $\mathrm{div}\,(\boldsymbol{V}_\eta \circ \eta^{-1}) = 0\}$.

In coordinates, if $\boldsymbol{x} = \eta(\boldsymbol{X})$, then

$$\boldsymbol{V}_\eta(\boldsymbol{X}) = V^i_\eta(\partial/\partial x^i). \tag{7.34}$$

Since $\mathrm{Diff_{vol}(D)}$ is a Lie group, its tangent space at the identity $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ is its Lie algebra. It turns out that the Lie algebra bracket of $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ is *minus* the usual Lie bracket for vector fields, which in components is $[U, V]^i = U^j(\partial V^i/\partial X^j) - V^j(\partial U^i/\partial X^j)$, where $\boldsymbol{U}, \boldsymbol{V} \in \mathscr{X}_{\mathrm{div}}(\mathrm{D})$.

To determine the dual of $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ and the cotangent bundle $\mathrm{T^*(Diff_{vol}(D))}$, we take a geometric point of view. Instead of considering the functional-analytic dual of all linear continuous functionals on $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$, we will be content to find another vector space $\mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$ and a weakly non-degenerate pairing

$$\langle\, ,\rangle : \mathscr{X}_{\mathrm{div}}(\mathrm{D})^* \times \mathscr{X}_{\mathrm{div}}(\mathrm{D}) \rightarrow \mathbb{R}; \tag{7.35}$$

this means that $\langle\, ,\rangle$ is a bilinear mapping, such that, if $\langle \boldsymbol{M}, \boldsymbol{V}\rangle = 0$ for all $\boldsymbol{V} \in \mathscr{X}_{\mathrm{div}}(\mathrm{D})$, then $\boldsymbol{M} = 0$, and if $\langle \boldsymbol{M}, \boldsymbol{V}\rangle = 0$ for all $\boldsymbol{M} \in \mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$, then $\boldsymbol{V} = 0$. Clearly, $\mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$ is a subspace of the functional-analytic dual. In order to make the exposition simpler, we will search for the 'geometric' dual only among smooth objects. Thus, for example, we will exclude point vortices or vortex patches. They can be dealt with in a similar manner by relaxing our point of view; see Marsden & Weinstein (1983). There is an obvious pairing between $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ and itself as we saw at the beginning of this section, namely the L^2 pairing

$$\langle \boldsymbol{U}, \boldsymbol{V}\rangle = \int_{\mathrm{D}} \boldsymbol{U}\cdot \boldsymbol{V}\, \mathrm{d}^3 X. \tag{7.36}$$

This pairing is weakly nondegenerate, and, thus, we identify $\mathscr{X}_{\mathrm{div}}(\mathrm{D})^*$ with $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$. Consequently, we have

$$\mathrm{T}^*_\eta(\mathrm{Diff_{vol}(D)}) := \mathrm{T}_\eta(\mathrm{Diff_{vol}(D)}). \tag{7.37}$$

Certain operations on $\mathrm{Diff_{vol}(D)}$ will be useful later. Left and right translations are defined by

$$\mathrm{L}_\eta : \mathrm{Diff_{vol}(D)} \rightarrow \mathrm{Diff_{vol}(D)}, \quad \mathrm{L}_\eta(\phi) = \eta \circ \phi, \tag{7.38}$$

$$\mathrm{R}_\eta : \mathrm{Diff_{vol}(D)} \rightarrow \mathrm{Diff_{vol}(D)}, \quad \mathrm{R}_\eta(\phi) = \phi \circ \eta, \tag{7.39}$$

for $\eta, \phi \in \mathrm{Diff_{vol}(D)}$. Both are diffeomorphisms of the Lie group $\mathrm{Diff_{vol}(D)}$. By considering curves of diffeomorphisms, their derivatives have the following expressions

$$\mathrm{T}_\phi \mathrm{L}_\eta : \mathrm{T}_\phi(\mathrm{Diff_{vol}(D)}) \rightarrow \mathrm{T}_{\eta\circ\phi}(\mathrm{Diff_{vol}(D)}), \tag{7.40}$$

$$\mathrm{T}_\phi \mathrm{L}_\eta(\boldsymbol{V}_\phi) = \mathrm{T}\eta \circ \boldsymbol{V}_\phi \tag{7.41}$$

and

$$\mathrm{T}_\phi \mathrm{R}_\eta : \mathrm{T}_\phi(\mathrm{Diff_{vol}(D)}) \rightarrow \mathrm{T}_{\phi\circ\eta}(\mathrm{Diff_{vol}(D)}), \tag{7.42}$$

$$\mathrm{T}_\phi \mathrm{R}_\eta(\boldsymbol{V}_\phi) = \boldsymbol{V}_\phi \circ \eta, \tag{7.43}$$

for $\boldsymbol{V}_\phi \in \mathrm{T}_\phi(\mathrm{Diff(D)})$. The physical interpretation of these formulae is the following. Think of ϕ as a relabelling, or rearrangement of the particles in D, and think of η as a motion. Then (7.43) says that the material derivative of the motion η, followed by the relabelling ϕ, equals $\mathrm{T}\eta \circ \boldsymbol{V}_\phi$. In local coordinates, if $\phi(\boldsymbol{X}) = \boldsymbol{Y}$ and $\eta(\boldsymbol{Y}) = \boldsymbol{y}$, then $\boldsymbol{V}_\phi(\boldsymbol{X}) = V^i(\boldsymbol{X})(\partial/\partial Y^i)$ and

$$(\mathrm{T}\eta \circ \boldsymbol{V}_\phi)^i(\boldsymbol{X}) = (\partial x^i/\partial Y^j)(\boldsymbol{Y})\, V^j(\boldsymbol{X})(\partial/\partial y^i). \tag{7.44}$$

On the other hand, (7.32) says that the material derivative of the relabelling ϕ followed by the motion η equals $V_\phi \circ \eta$. In local coordinates, if

$$\eta(X) = x, \quad \phi(x) = y \quad \text{then} \quad V_\phi(X) = V^i(X)(\partial/\partial Y^i) \tag{7.45}$$

and

$$(V_\phi \circ \eta)^i(X) = (V^i \circ \eta)(X)(\partial/\partial y^i). \tag{7.46}$$

Note that by (7.19) the material velocity V_t is the right translate of the spatial velocity v_t by η_t.

If $V \in \mathscr{X}_{\text{div}}(\mathrm{D})$, a diffeomorphism $\eta \in \mathrm{Diff}_{\text{vol}}(\mathrm{D})$ acts on V by the adjoint action, the analogue of conjugation for matrices. Its definition, combined with (7.40–7.43) gives

$$\begin{aligned} \mathrm{Ad}_\eta V &:= \mathrm{T_e}(\mathrm{L}_\eta \circ \mathrm{R}_{\eta^{-1}})(V) \\ &= \mathrm{T}_{\eta^{-1}} \mathrm{L}_\eta(\mathrm{T_e R}_{\eta^{-1}}(V)) \\ &= \mathrm{T}\eta \circ V \circ \eta^{-1} && (7.47) \\ &= \eta_* V, && (7.48) \end{aligned}$$

i.e. the adjoint action of η on V is the push-forward of vector fields:

$$\mathrm{Ad}_\eta V = \eta_* V. \tag{7.49}$$

Using the pairing (7.36), a change-of-variables argument shows that the coadjoint action $\mathrm{Ad}^*_{\eta^{-1}}$ of η on $U \in \mathscr{X}_{\text{div}}(\mathrm{D})^* = \mathscr{X}_{\text{div}}(\mathrm{D})$, defined by

$$\langle \mathrm{Ad}^*_{\eta^{-1}} U, V \rangle = \langle U, \mathrm{Ad}_{\eta^{-1}} V \rangle, \tag{7.50}$$

has the expression

$$\mathrm{Ad}^*_{\eta^{-1}} U = (\mathrm{T}\eta^{-1})^+ \circ U \circ \eta^{-1}, \tag{7.51}$$

where $(\mathrm{T}\eta^{-1})^+$ denotes the adjoint of the linear map $\mathrm{T}_X \eta^{-1}$ with respect to the dot product for every $X \in \mathrm{D}$.

Having studied the phase space, we must express the total energy, the Hamiltonian of the stratified fluid motion on $\mathrm{T}^*(\mathrm{Diff}_{\text{vol}}(\mathrm{D})) = \mathrm{T}(\mathrm{Diff}_{\text{vol}}(\mathrm{D}))$, in Lagrangian coordinates. We have, after performing the change of variables $x = \eta_t(X)$ and taking into account (7.17) and (7.24),

$$H = \frac{1}{2} \int_\mathrm{D} \rho_t(x) \| v_t(x) \|^2 \, \mathrm{d}^3 x + g \int_\mathrm{D} \rho_t(x) \, z \, \mathrm{d}^3 x \tag{7.52}$$

$$= \frac{1}{2} \int_\mathrm{D} \rho_0(X) \| V_{\eta_t}(X) \|^2 \mathrm{d}^3 X + g \int_\mathrm{D} \rho_0(X) \, \eta_t^3(X) \, \mathrm{d}^3 X, \tag{7.53}$$

where $\eta_t^i(X)$ is the ith component of $\eta_t(X)$, $i = 1, 2, 3$. The first term, the kinetic energy, is the quadratic norm of the following weak Riemannian metric on $\mathrm{Diff}_{\text{vol}}(\mathrm{D})$:

$$\langle\langle U_\eta, V_\eta \rangle\rangle = \int_\mathrm{D} \rho_0(X) \, U_\eta(X) \cdot V_\eta(X) \, \mathrm{d}^3 X, \tag{7.54}$$

where $\cdot$ denotes the dot product in $\mathbb{R}^3$. This metric induces a bundle metric on

$$\mathrm{T}^*(\mathrm{Diff}_{\text{vol}}(\mathrm{D})) = \mathrm{T}(\mathrm{Diff}_{\text{vol}}(\mathrm{D})),$$

via the pairing (7.36), which coincides with (7.54), where

$$U_\eta, V_\eta \in \mathrm{T}^*_\eta(\mathrm{Diff}_{\text{vol}}(\mathrm{D})) = \mathrm{T}_\eta(\mathrm{Diff}_{\text{vol}}(\mathrm{D})).$$

Thus, the Hamiltonian (6.6) has the following expression on $T^*(\mathrm{Diff}_{\mathrm{vol}}(D)) = T(\mathrm{Diff}_{\mathrm{vol}}(D))$:

$$H_{\rho_0}(\boldsymbol{V}_\eta) = \tfrac{1}{2}\langle\langle \boldsymbol{V}_\eta, \boldsymbol{V}_\eta\rangle\rangle + g\int_D \rho_0(\boldsymbol{X})\,\eta^3(\boldsymbol{X})\,\mathrm{d}^3X. \tag{7.55}$$

We study the invariance properties of H under right translations; i.e. we replace η by $\eta\circ\phi$ and right translate the argument of H by ϕ. Since operations on the cotangent bundle are defined by duality, it suffices to work with (7.54) And apply (7.42) and (7.43). The right translated Hamiltonian has the expression

$$\frac{1}{2}\int_D \rho_0(\boldsymbol{X})|\boldsymbol{V}_\eta(\phi(\boldsymbol{X}))|^2\,\mathrm{d}^3X + g\int_D \rho_0(\boldsymbol{X})(\eta\circ\phi)^3\,(\boldsymbol{X})\,\mathrm{d}^3X \tag{7.56}$$

$$= \frac{1}{2}\int_D \rho_0(\phi^{-1}(\boldsymbol{Y})|\,\boldsymbol{V}_\eta(\boldsymbol{Y})\,|^2\,\mathrm{d}^3Y + g\int_D \rho_0(\phi^{-1}(\boldsymbol{Y}))\,\eta^3(\boldsymbol{Y})\,\mathrm{d}^3Y, \tag{7.57}$$

after changing variables $\boldsymbol{Y} = \phi(\boldsymbol{X})$. It is apparent now that H_{ρ_0} satisfies

$$H_{\rho_0} = H_{\rho_0\circ\phi^{-1}}\circ T^*R_\phi \tag{7.58}$$

and *is invariant under the subgroup*

$$\mathrm{Diff}_{\mathrm{vol}}(D)_{\rho_0} = \{\phi\in\mathrm{Diff}_{\mathrm{vol}}(D)\,|\,\rho_0\circ\phi = \rho_0\}. \tag{7.59}$$

The cotangent bundle $T^*(\mathrm{Diff}_{\mathrm{vol}}(D))$ has the canonical Poisson bracket which can be explicitly written in a chart. Consider the Lagrangian to Eulerian map (explained in Appendix B)

$$J: T^*(\mathrm{Diff}_{\mathrm{vol}}(D))\times\mathscr{F}(D)\to\mathscr{X}_{\mathrm{div}}(D)\,\circledS\,\mathscr{F}(D))^* = \mathscr{X}_{\mathrm{div}}(D)\times\mathscr{F}(D) \tag{7.60}$$
$$J(\boldsymbol{V}_\eta, \rho_0) = (\boldsymbol{V}_\eta\circ\eta^{-1}, \rho),\quad \rho = \rho_0\circ\eta^{-1},$$

where $\mathscr{F}(D)$ denotes functions on D. With respect to the L^2 pairing,

$$\langle f, g\rangle = \int_D f(\boldsymbol{X})\,g(\boldsymbol{X})\,\mathrm{d}^3X \tag{7.61}$$

$\mathscr{F}(D)$ is its own dual; so $\boldsymbol{v}$ and ρ denote the Eulerian velocity and density.

With these notations, the canonical bracket of F and G in $T^*(\mathrm{Diff}_{\mathrm{vol}}(D))$, becomes, via the map J,

$$\{F, G\}(\boldsymbol{v}, \rho) = \int_D \boldsymbol{v}\cdot\left[\left(\frac{\delta G}{\delta\boldsymbol{v}}\cdot\nabla\right)\frac{\delta F}{\delta\boldsymbol{v}} - \left(\frac{\delta F}{\delta\boldsymbol{v}}\cdot\nabla\right)\frac{\delta G}{\delta\boldsymbol{v}}\right]\mathrm{d}^3x + \int_D \rho\left[\frac{\delta G}{\delta\boldsymbol{v}}\cdot\nabla\frac{\delta F}{\delta\rho} - \frac{\delta F}{\delta\boldsymbol{v}}\cdot\nabla\frac{\delta G}{\delta\rho}\right]\mathrm{d}^3x, \tag{7.62}$$

which coincides with the formula (7.15), i.e. the Boussinesq approximation and the full Euler equations have the *same* Poisson bracket, but *different* Hamiltonians. The fact that the map J sends canonical Poisson brackets in Lagrangian coordinates to Lie–Poisson brackets in the dual of the semidirect product Lie algebra $\mathscr{X}_{\mathrm{div}}(D)\,\circledS\,\mathscr{F}(D)$ (action is by minus the Lie derivative) in Eulerian coordinates is a special case of a general theorem discussed in Appendix B. As in 7A it can be checked directly (if one wishes to do so) that the Hamilton equations $\dot{F} = \{F, H\}$ with the Hamiltonian given by (6.6) and the bracket given by (7.6) yield the equations of motion (6.1), (6.2) and (6.3).

It can also be checked directly that the functions C_F given by (6.5) commute with any other function on $(\mathscr{X}_{\mathrm{div}}(D)\,\circledS\,\mathscr{F}(D))^* = \mathscr{X}_{\mathrm{div}}(D)\times\mathscr{F}(D)$ in the Poisson bracket (7.62). Such functions characterizing the degeneracy of the Poisson bracket are called *Casimir functions.* We shall give a different criterion for checking whether a given function is a Casimir in Appendix B, using the coadjoint action in (7.44).

One can also impose the condition ρ = constant on ∂D in the foregoing discussion, with only minor modifications.

(c) *Two-dimensional stratified Boussinesq flow*

The Poisson bracket in this case remains (7.15) and the Hamiltonian is (7.14) in terms of density and velocity, with the following provisions: all vector fields have identically zero component in the y-direction and all functions and vector fields are independent of y. Under the additional restrictions that the domain of flow in the (x, y) plane is simply connected, and the stream function vanishes on the boundary, a Hamiltonian structure expressible in terms of density and *vorticity* can be derived either as above, or by using the Clebsch procedure (See, for example, Seliger & Whitham (1968); Henyey (1982); Morrison (1982); Holm & Kupershmidt (1983); Marsden & Weinstein (1983) and Benjamin (1984) for other discussions of these procedures).

For density $\rho(x, z, t)$ and vorticity $\boldsymbol{\omega} = \hat{\boldsymbol{y}}\omega(x, z, t)$, the Boussinesq equations describing a two-dimensional stratified flow in a region of the (x, z) plane are

$$\partial\omega/\partial t = \{\omega, \psi\} + \{gz/\rho*, \rho\}, \tag{7.63}$$

$$\partial\rho/\partial t = \{\rho, \psi\}, \tag{7.64}$$

where $\psi = (\nabla^2)^{-1}\omega$ is the stream function, and $\{h, k\} := \partial_z h \partial_x g - \partial_x h \partial_z g$. The Clebsch procedure for deriving the Hamiltonian structure of these equations is based on the following proposition.

PROPOSITION. *Equations* (7.63) *and* (7.64) *result from a constrained action principle* $\delta S = 0$ *with*

$$S = \int \mathrm{d}x\, \mathrm{d}z\, \mathrm{d}t[-\tfrac{1}{2}\omega(\nabla^2)^{-1}\omega - \rho gz/\rho* + \alpha(\partial\rho/\partial t + \{(\nabla^2)^{-1}\omega, \rho\}) + \gamma_i(\partial l^i/\partial t + \{(\nabla^2)^{-1}\omega, l^i\})]. \tag{7.65}$$

Here α, γ_i are Lagrange multipliers that impose, respectively, the constraints of mass conservation, and preservation of 'particle identity.' The maps $l^i(\boldsymbol{x}, t)$ are the components of η_t^{-1}, the inverse of the map η_t in the previous subsection. Summation on repeated spatial indices i is implied.

The proof of the proposition given here is by direct computation. Independent variations of (7.65) give the following relations,

$$\begin{aligned}
\delta\omega&: \quad \omega - \{\rho, \alpha\} - \{l^i, \gamma_i\} = 0,\\
\delta\rho&: \quad gz/\rho_* + \partial\alpha/\partial t + \{\psi, \alpha\} = 0,\\
\delta\alpha&: \quad \partial\rho/\partial t + \{\psi, \rho\} = 0,\\
\delta\gamma_i&: \quad \partial l^i/\partial t + \{\psi, l^i\} = 0,\\
\delta l^i&: \quad \partial\gamma_i/\partial t + \{\psi, \gamma_i\} = 0,
\end{aligned} \tag{7.66}$$

where boundary integrals have been set equal to zero when integrating by parts. From (7.66), the vorticity equation (7.63) may be reconstructed as

$$\frac{\partial\omega}{\partial t} = \left\{\frac{\partial\rho}{\partial t}, \alpha\right\} + \left\{\rho, \frac{\partial\alpha}{\partial t}\right\} + \left\{\frac{\partial l^i}{\partial t}, \gamma_i\right\} + \left\{l^i, \frac{\partial\gamma_i}{\partial t}\right\} = \left\{\frac{gz}{\rho_*}, \rho\right\} + \{\{\rho, \alpha\} + \{l^i, \gamma_i\}, \psi\}, \tag{7.67}$$

upon using the Jacobi identity. This proves the proposition.

Since the action principle for the Lagrangian (7.65) contains time derivatives only linearly, the Clebsch representation for the vorticity in (7.66) can be expressed as

$$\omega = \sum_s \{q^s, p_s\}, \tag{7.68}$$

where (q^s, p_s) are conjugate pairs $(\rho, \alpha), (l^i, \gamma_i)$ with respect to the canonical Poisson bracket. Under the Clebsch map (7.68), the canonical Poisson bracket induces, by direct computation (see Holm & Kupershmidt (1983) for discussions of this type of calculation)

$$\{F, G\}(\omega, \rho) = \int \mathrm{d}x\,\mathrm{d}z \left[\omega \left\{\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta \omega}\right\} + \rho \left\{\frac{\delta F}{\delta \rho}, \frac{\delta G}{\delta \omega}\right\} + \rho \left\{\frac{\delta F}{\delta \omega}, \frac{\delta G}{\delta \rho}\right\}\right] \tag{7.69}$$

as the Poisson bracket on the space of functionals of vorticity and density. In this expression, $\delta F/\delta\omega$ must be interpreted with care, see Marsden & Weinstein (1983) and Lewis *et al.* (1985). The corresponding Hamiltonian for Boussinesq flow is

$$H(\omega, \rho) = -\int \mathrm{d}x\,\mathrm{d}z\,[\tfrac{1}{2}\omega(\nabla^2)^{-1}\omega - \rho g z/\rho^*]. \tag{7.70}$$

As can be verified easily using the Poisson bracket (7.69), the functionals

$$C_{F,G}(\omega, \rho) = \int \mathrm{d}x\,\mathrm{d}z[\omega F(\rho) + G(\rho)] \tag{2.15}$$

commute with all other functionals of (ω, ρ), provided F satisfies $\int_{\mathrm{D}} (\partial\rho/\partial x)\, F(\rho)\,\mathrm{d}x\,\mathrm{d}z = 0$. These functionals play an important role in the stability analysis of §2.

(*d*) *Two-dimensional stratified Euler flow*

The Poisson bracket is given by (7.62) and the Hamiltonian by (6.6), upon restriction to planar (x, z) dependence only. It can be verified again as in the Boussinesq case that $C_{F,G}$ in (2.15) commutes, using the bracket (7.62) restricted to the plane, with all other functionals, provided $F(\rho)$ in $C_{F,G}$ satisfies $\int_{\mathrm{D}} \partial\rho/\partial x\, F(\rho)\,\mathrm{d}x\,\mathrm{d}z = 0$. For simply connected domains, the Poisson bracket in terms of vorticity and density is again given by (7.69).

8. High frequency cut-off for two-dimensional Boussinesq flow

In this section, we study stability of two-dimensional stationary solutions for a modified inviscid model which 'cuts off' the wavenumbers k at a maximum value, k_{M}. The modification is obtained by replacing the Laplacian in the Hamiltonian by a bounded operator whose spectrum truncates that of the Laplacian.

One motivation for this model is the fact that inviscid models governed by the Euler equations have a range of validity limited to flows whose wave number (or gradients) are small enough so that dissipation is negligible compared to inertial accelerations. In the flow of any physical fluid, the requirement that the velocity vanish at undriven boundaries induces boundary layers in which gradients of $\boldsymbol{v}(\boldsymbol{x}, t)$ are important and viscosity essential. As described in great detail by Pedlosky (1979) for oceanic and atmospheric applications, it is through these boundary layers that the motion in the effectively inviscid interior is determined. It is natural then to clarify our sense of what constitutes an inviscid flow by allowing only dynamics with wavenumbers bounded by a maximum value; call it k_{M}.

For illustration, consider the Ekman layer created in a fluid rotating at angular velocity Ω, with effective viscosity A_V in the vertical direction. The Ekman depth $(A_V/\Omega)^{\frac{1}{2}}$ gives a scale to k_M

$$k_M \approx (\Omega/A_V)^{\frac{1}{2}}.$$

For an ocean rotating at the inertial frequency $\approx 10^{-4}$/s and $A_V \approx 10^{-2}$ m²/s we would have $k_M \approx 0.1$/m. Restricting our attention to oceanic scales $\gtrsim 10$ m would then define the realm in which we work.

As in section 7C we shall assume that the domain D is simply connected and, thus, any divergence free velocity field tangent to the boundary admits a stream function vanishing on ∂D. In this case the Hamiltonian of stratified Boussinesq fluid flow has the expression

$$H(\omega, \rho) = -\int_D (\tfrac{1}{2}\omega(\nabla^2)^{-1}\omega - (\rho/\rho_*)\, gz)\, dx\, dz. \tag{8.1}$$

The dynamic variables can be taken as ω and ρ, and the equation of motion are

$$\partial_t \omega + \boldsymbol{v}\cdot\nabla\omega = \frac{g}{\rho_*}\frac{\partial\rho}{\partial x}, \tag{8.2a}$$

$$\partial_t \rho + v\cdot\nabla\rho = 0. \tag{8.2b}$$

These equations are Hamiltonian with respect to the Lie–Poisson bracket (7.69), namely

$$\{F, G\}(\omega, \rho) = \int_D \left[\omega\left\{\frac{\delta F}{\delta\omega}, \frac{\delta G}{\delta\omega}\right\} + \rho\left\{\frac{\delta F}{\delta\omega}, \frac{\delta G}{\delta\rho}\right\} - \rho\left\{\frac{\delta G}{\delta\omega}, \frac{\delta F}{\delta\rho}\right\}\right] dx\, dz \tag{8.3}$$

on the dual of the semidirect product of functions on D vanishing on ∂D and functions on D; the action is by the minus (z, x)-Poisson bracket, $\{,\}$. The Hamiltonian function is given by (8.1). Using the stream function $\psi = (\nabla^2)^{-1}\omega$ (with zero boundary conditions) the system (8.2) becomes

$$\left.\begin{aligned} \partial_t \omega &= \{\omega, \psi\} - \{\rho, gz/\rho_*\}, \\ \partial_t \rho &= \{\rho, \psi\}, \end{aligned}\right\} \tag{8.4}$$

which can also be easily verified using the Lie–Poisson bracket (8.5), since $\delta H/\delta\omega = -\psi$ and $\delta H/\delta\rho = gz/\rho_*$.

We turn to a modification of the planar stratified Boussinesq flow. Define the cutoff Laplace operator ∇^2_M as follows. If we are in a bounded domain, consider the eigenvalue problem for the Laplacian ∇^2 with the eigenfunctions $\phi_q(\tilde{\boldsymbol{x}})$ ($\tilde{\boldsymbol{x}} = (x, z)$) required to vanish on the domain boundaries. Thus, we get a set of eigenfunctions with eigenvalues q^2 solving

$$\nabla^2\phi_q(\tilde{\boldsymbol{x}}) = -q^2\phi_q(\tilde{\boldsymbol{x}}). \tag{8.5}$$

If ω has the orthonormal expansion

$$\omega(\tilde{\boldsymbol{x}}) = \sum_q C_q\phi_q(\tilde{\boldsymbol{x}}), \tag{8.6}$$

then

$$\nabla^2\omega(\tilde{\boldsymbol{x}}) = -\sum_q q^2 C_q\phi_q(\tilde{\boldsymbol{x}}). \tag{8.7}$$

Now set, for fixed k_M,

$$\nabla^2_M\,\omega(\tilde{\boldsymbol{x}}) = -\sum_q \min(q^2, k^2_M)\, C_q\phi_q(\tilde{\boldsymbol{x}}). \tag{8.8}$$

For inviscid Boussinesq flow we modify the Hamiltonian (8.1) to a new Hamiltonian function H^M, given by

$$h^{\mathrm{M}}(\omega, \rho) = \int (-\tfrac{1}{2}\omega(\nabla_{\mathrm{M}}^2)^{-1}\omega + (\rho/\rho_*)\,gz)\,\mathrm{d}x\,\mathrm{d}z \tag{8.9}$$

defined on all (scalar curls) ω and ρ. The Poisson structure, however, is unaltered, so the Casimir functions available to us remain unchanged. The evolution equations associated with H^M via the bracket (8.3) are given as follows:

$$\left.\begin{aligned} \frac{\partial \omega}{\partial t} + \boldsymbol{v}_{\mathrm{M}}\cdot\nabla\omega &= \frac{g}{\rho_*}\frac{\partial}{\partial x}\rho, \\ \partial\rho/\partial t + \boldsymbol{v}_{\mathrm{M}}\cdot\nabla\rho &= 0, \end{aligned}\right\} \tag{8.10}$$

where $$\boldsymbol{v}_{\mathrm{M}} = (\partial\psi_{\mathrm{M}}/\partial z, \partial\psi_{\mathrm{M}}/\partial z), \quad \psi_{\mathrm{M}} = (\nabla_{\mathrm{M}}^2)^{-1}\omega. \tag{8.11}$$

Note that $\psi_{\mathrm{M}} = (\nabla_{\mathrm{M}}^2)^{-1}\nabla^2\psi$, so ψ_{M} is obtained from ψ by a modification of the identity that 'cuts off' the high wavenumbers in ψ.

By construction we have

$$\int \mathrm{d}x\,\mathrm{d}z\,\Delta\omega(-\nabla_{\mathrm{M}}^2)^{-1}\Delta\omega \geqslant (1/k_{\mathrm{M}}^2)\int \mathrm{d}x\,\mathrm{d}z\,(\Delta\omega)^2, \tag{8.12}$$

where $\Delta\omega$ is a finite perturbation of ω. Let

$$H_C^M = H^M + \int \mathrm{d}x\,\mathrm{d}z\,[\omega F(\rho) + G(\rho)] \tag{8.13}$$

and thus $\delta^2 H_C^M$ is definite when

$$1/k_{\mathrm{M}}^2 > \int \mathrm{d}x\,\mathrm{d}z\,(F'(\rho_{\mathrm{e}})^2/A)(\Delta\omega)^2 \Big/ \int \mathrm{d}x\,\mathrm{d}z\,(\Delta\omega)^2 \tag{8.14}$$

and then *formal* stability holds.

To establish nonlinear stability for those cut-off flows that are formally stable, we form the quantity

$$\hat{H}^M(\Delta\psi, \Delta\rho) = H_C^M(\psi_{\mathrm{e}}+\Delta\psi, \rho_{\mathrm{e}}+\Delta\rho) - H_C^M(\psi_{\mathrm{e}}, \rho_{\mathrm{e}}) - \mathrm{D}H_C^M(\psi_{\mathrm{e}}, \rho_{\mathrm{e}})\cdot(\Delta\psi, \Delta\rho), \tag{8.15}$$

where $\Delta\psi$ and $\Delta\rho$ are finite deviations from ψ_{e} and ρ_{e}. Conservation of

$$\hat{H}^M(\Delta\omega, \Delta\rho) = H_C^M(\omega_{\mathrm{e}}+\Delta\omega, \rho_{\mathrm{e}}+\Delta\rho) - H_C^M(\omega_{\mathrm{e}}, \rho_{\mathrm{e}}) - \mathrm{D}H_C(\omega_{\mathrm{e}}, \rho_{\mathrm{e}})\cdot(\Delta\omega, \Delta\rho) \tag{8.16}$$

follows directly from conservation of $H_C^M(\omega, \rho)$, as before. Noting that the maximum allowed wavenumber in $\Delta\psi$ is k_{M}, we have

$$\hat{H}^M(\Delta\psi, \Delta\rho) \geqslant \int \mathrm{d}x\,\mathrm{d}z\,[\phi(\omega_{\mathrm{e}}+\Delta\omega, \rho_{\mathrm{e}}+\Delta\rho) - \phi(\omega_{\mathrm{e}}, \rho_{\mathrm{e}}) - \partial_\rho\phi(\rho_{\mathrm{e}}, \omega_{\mathrm{e}})\,\Delta\rho - \partial_\omega\phi(\rho_{\mathrm{e}}, \omega_{\mathrm{e}})\,\Delta\omega], \tag{8.17}$$

where $$\phi(\omega, \rho) = \tfrac{1}{2}\omega^2/k_{\mathrm{M}}^2 + \omega F(\rho) + G(\rho) + \rho g z \tag{8.18}$$

and $\omega_{\mathrm{e}} = \nabla^2\psi_{\mathrm{e}}$.

If we now require

$$\frac{\partial^2}{\partial\rho^2}\phi(\rho) = \frac{\partial\omega_{\mathrm{e}}}{\partial\rho_{\mathrm{e}}}\frac{\partial F}{\partial\rho_{\mathrm{e}}} - g\frac{\partial z}{\partial\rho_{\mathrm{e}}} =: A > \alpha \tag{8.19}$$

and
$$\frac{\partial^2\phi}{\partial\omega^2}\frac{\partial^2\phi}{\partial\rho^2}-\left(\frac{\partial^2\phi}{\partial\rho\,\partial\omega}\right)^2 = \frac{A}{k_{\mathrm{M}}^2}-F'(\rho_{\mathrm{e}})^2 > \alpha\beta-\gamma^2 > 0, \tag{8.20}$$

for finite α, β, and γ, then we may conclude that

$$\hat{H}^M(\Delta\omega, \Delta\rho) \geqslant \frac{1}{2}\int \mathrm{d}x\,\mathrm{d}z\,(\Delta\omega, \Delta\rho)\begin{bmatrix}\gamma & \beta\\ \beta & \alpha\end{bmatrix}\begin{bmatrix}\Delta\omega\\ \Delta\rho\end{bmatrix}. \tag{8.21}$$

Conservation of $\hat{H}^M(\Delta\omega, \Delta\rho)$ means $\hat{H}^M(\Delta\omega(\tilde{\boldsymbol{x}}, t), \Delta\rho(\tilde{\boldsymbol{x}}, t)) = \hat{H}^M(\Delta\omega(\tilde{\boldsymbol{x}}, 0), \Delta\rho(\tilde{\boldsymbol{x}}, 0))$. So $\hat{H}^M(\Delta\omega(\tilde{\boldsymbol{x}}, 0), \Delta\rho(\tilde{\boldsymbol{x}}, 0))$ bounds the size of the excursions $\Delta\omega(\tilde{\boldsymbol{x}}, t), \Delta\rho(\tilde{\boldsymbol{x}}, t)$ for all time by virtue of (8.21). For $\Delta\omega(\tilde{\boldsymbol{x}}, 0)$, $\Delta\rho(\tilde{\boldsymbol{x}}, 0)$ giving finite $\hat{H}^M$, we conclude that meeting the conditions (8.10) and (8.20), as well as the corresponding upper bounds guarantees stability since $\Delta\omega$ and $\Delta\rho$ are bounded in the norm provided by (8.21).

As an example of stabilization by applying this wavenumber cut-off, we consider equilibrium solutions for which the Long equation (2.13)

$$\nabla^2\psi_{\mathrm{e}}+\frac{gz}{\rho_*}\bar{\rho}'(\psi_{\mathrm{e}}) = L(\psi_{\mathrm{e}}) \tag{8.22}$$

is linear. In particular, take $L(\psi_{\mathrm{e}}) = -k^2\psi_{\mathrm{e}}$ and $\bar{\rho}(\psi_{\mathrm{e}}) = (\rho_*\,a^2)/2g\,\psi_{\mathrm{e}}$. Then calculation of $F(r_{\mathrm{e}})$ by (2.22) and $G(\rho_{\mathrm{e}})$ by (2.26) gives

$$F(\rho_{\mathrm{e}}) = \bar{\psi}(\rho_{\mathrm{e}}) = \left(\frac{2g}{\rho_*\,a^2}\,\rho_{\mathrm{e}}\right)^{\frac{1}{2}}, \tag{8.23}$$

$$G(\rho_{\mathrm{e}}) = \left(\frac{k^2 g}{a^2\rho_*}\right)\rho_{\mathrm{e}}+\text{constant}. \tag{8.24}$$

Consequently, by (2.29)

$$A = \omega_{\mathrm{e}}\,F''(\rho_{\mathrm{e}})+G''(\rho_{\mathrm{e}}) = (k^2+a^2z)\,\frac{g}{\rho*\,a^2}\,\rho_{\mathrm{e}}^{-1} > 0. \tag{8.25}$$

Completing the square in (2.28) for the second variation gives

$$\delta^2 H_C = \int_{\mathrm{D}}\mathrm{d}x\,\mathrm{d}z\left[|\,\delta\boldsymbol{v}\,|^2-\frac{(F'(\rho_{\mathrm{e}}))^2}{A}\,(\delta\omega)^2+A((\delta\rho+\frac{F'(\rho_{\mathrm{e}}^2)}{A}\,\delta\omega))^2\right]. \tag{8.26}$$

Now, imposing the wavenumber cut-off and substituting (8.23) and (8.25) leads to

$$\delta^2 H_C^M \geqslant \int_{\mathrm{D}}\mathrm{d}x\,\mathrm{d}z\left[\left(\frac{1}{k_{\mathrm{M}}^2}-\frac{2}{k^2+a^2z}\right)(\delta\omega)^2+(k^2+a^2z)\,\frac{g}{\rho_*\,a^2}\,\rho_{\mathrm{e}}^{-1}\left(\delta\rho+\frac{(\rho_*\,a^2\rho_{\mathrm{e}}/2g)^{\frac{1}{2}}}{k^2+a^2z}\,\delta\omega\right)^2\right]. \tag{8.27}$$

Thus $\delta^2 H_C^M$ is definite and formal stability occurs when

$$\frac{1}{k_{\mathrm{M}}^2} > \int_{\mathrm{D}}\mathrm{d}x\,\mathrm{d}z\,\frac{2(\delta\omega)^2}{k^2+a^2z}\bigg/\int_{\mathrm{D}}\mathrm{d}x\,\mathrm{d}z\,(\delta\omega)^2, \tag{8.28}$$

in agreement with (8.14). In particular, this holds if the domain is such that we can sensibly choose throughout

$$k_{\mathrm{M}}^2 < \tfrac{1}{2}(k^2+a^2z). \tag{8.29}$$

9. Conclusions

In this paper, we have used a constrained energy method for the investigation of nonlinear stability of stationary solutions of stratified, incompressible, ideal flows. Specific stability criteria involving a Richardson number are developed in examples of parallel shear flow.

The method, initiated by Arnold around 1965 and developed by others since, relies on the Hamiltonian structure of the ideal fluid equations augmented by a convexity analysis of the energy, H, plus a conserved quantity, C, of a special type called Casimir. The Hamiltonian structure is fully explained in the main body of the paper, as is the convexity analysis of $H_C = H + C$.

The first step in the method is to find C such that the first variation of H_C vanishes at the equilibrium flow in question. For two-dimensional flows, the Long equation results from this step, and in three dimensions the generalization of the Long equation emerges. The second step is to examine the second variation of H_C at the equilibrium to determine whether it is of definite sign for all variations. If the flow passes this test successfully, we call it formally stable. Indeed, formal stability allows us to conclude the Liapunov stability of the linearized equations of motion at the equilibrium state. Appendix C is devoted to a discussion of this issue.

Formal stability is prerequisite to the use of convexity estimates on H_C to establish the full nonlinear stability of the flow. For unstratified shear flow, this leads to the classical Rayleigh criterion, as shown by Arnold. For stratified flows in two and three dimensions we demonstrate both under the Boussinesq approximation and for the full Euler equations that Liapunov stability is not achieved, in general. However, the ingredient causing the failure is the development of density perturbations with high wavenumber, and we determine the range of nonlinearly stable wavenumbers in terms of equilibrium-state quantities for each of these incompressible fluid models. The development of wavenumber controls on the streamfunction instead of the density variations and an instructive example in two dimensions are discussed in §8.

The failure of nonlinear stability is just as interesting as its success. It *suggests* that the full nonlinear problem is actually unstable. If the linearized flow is spectrally stable (no normal modes are unstable), we expect the nonlinear problem to have a slowly developing (algebraic) instability when formal stability fails. In the two-dimensional context, this remark would apply especially to flows with Richardson number exceeding one-quarter according to the usual criterion. A discussion of the Hamiltonian nature of the linearized equations and of the two-dimensional Taylor–Goldstein equations is given in Appendix C.

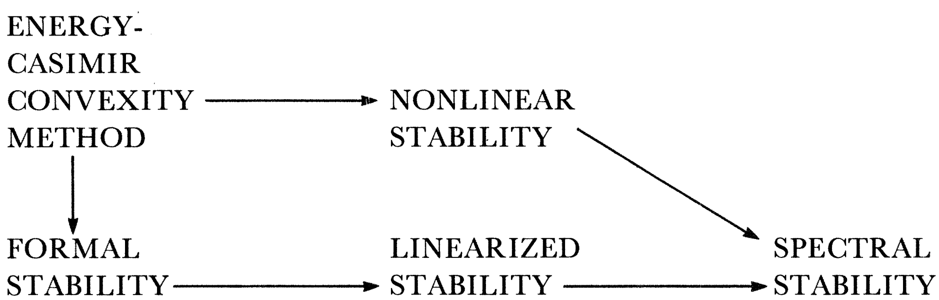
Three-dimensional stratified flow is a richer field of inquiry and has a richer class of available Casimirs than in two dimensions. We give both sufficient, and necessary and sufficient conditions for the formal stability of these flows, in the stable range of density-variation gradients for given equilibrium. We then proceed to use the needed convexity arguments to establish criteria for nonlinear stability (provided density-variation gradients stay in the stable range). In a key example treated in §4, we show that for parallel shear flows $\boldsymbol{u}_e = (u(y, z), 0, 0)$, $\rho_e = \rho(z)$ with a small variation of velocity in the y dimension, nonlinear stability is achieved, provided the Richardson number for variations across density surfaces exceeds unity. In addition, our stability criterion in this case indicates that *statically unstable* configurations, namely $\rho_z > 0$, can be stabilized by appropriate shears.

H.D.I.A. was supported by U.S. Department of Energy and Office of Naval Research, Code 422 PD; D.D.H was supported by D.O.E. contract no. W-7405-Eng-36 and Office of Basic Energy Sciences, Department of Applied Mathematics; J.E.M. was partly supported by D.O.E. contract no. AT03-82ER12097; T.S.R. was supported by an N.S.F. postdoctoral fellowship and an A.P. Sloan Foundation Fellowship while at the University of California, Berkeley.

Appendix A. The energy-Casimir convexity method

In this Appendix, we summarize the general steps one follows in proving stability as formulated in Holm *et al.* (1985). We include it here for the convenience of the reader. This method, illustrated concisely in Arnold (1969), is followed in the main text. The method is given here for the reader's convenience and to link the ideas with the more recent literature.

In the introduction, we defined the terms 'nonlinear stability', 'formal stability', 'linearized stability' and 'spectral stability'. These concepts and the convexity method are logically interrelated as in the diagram, where each arrow means 'implies'.



The algorithm that is used in each of the examples proceeds in a step by step manner with certain optional, but sometimes useful, steps discussed under the heading of 'Remarks'.

The stability algorithm

A. *Equations of motion and Hamiltonian*

Choose a (*Banach*) *space* P *of fields u and write the equation of motion in first order form as*

$$\dot{u} = X(u) \tag{A 1}$$

for a (*nonlinear*) *operator X mapping a domain in* P *to* P. *Find a conserved function H for* (1); *that is, a map* $H\colon \mathrm{P} \to \mathbb{R}$ *such that* $(\mathrm{d}/\mathrm{d}t)\, H(u) = 0$ *for any* C^1 *solution u of* (1). (Usually, H is the energy of the system.)

Remark A. Often P is a Poisson space, i.e. a linear space (or more generally a manifold) admitting a Poisson bracket operation $\{,\}$ on the space of real valued functions on P which makes them into a Lie algebra and which is a derivation in each variable. There are systematic procedures for obtaining such brackets; these procedures are reviewed in Appendix B. The equations (A 1) are often Hamiltonian for such a bracket structure:

$$\dot{F} = \{F, H\} \tag{A 2}$$

where H is the energy, F is any function of $u \in \mathrm{P}$, and $\dot{F}$ is its time derivative through the dependence of u on t.

B. *Constants of motion*

Find a family of constants of the motion for (A 1). *That is, find a collection of functions C on* P *such that* $\mathrm{d}/\mathrm{d}t\, C(u) = 0$ *for any* C^1 *solution u of* (A 1).

Remark B. Unless a sufficiently large family of constants of motion is found, the ensuing step (C) may not be possible. A good way to find such functions is to use the Hamiltonian formalism in *Remark A* to find F's such that $\{F, H\} = 0$ and to find Casimir functions for the Poisson structure; that is, C's such that $\{C, G\} = 0$ for *all* G.

C. *First variation*

Relate an equilibrium solution u_e *of* (A 1) *i.e.* $X(u_e) = 0$ (so that $\mathrm{d}/\mathrm{d}t\, u_e = 0$) *and a constant of the motion C by requiring that* $H_C := H + C$ *has a critical point at* u_e. *Note*: C may or may not be uniquely determined at this stage. Keep C as general as possible, as any freedom may be useful in step (D).

Remark C. If *Remarks A* and *B* are followed, then, in principle, such a C always exists, at least locally. Indeed, level sets of the Cs define the 'symplectic leaves' of the Poisson structure $\{,\}$ and equilibrium solutions are critical points of H restricted to such leaves. Thus, by the Lagrange multiplier theorem, $H+C$ has a critical point at u_e for an appropriate Casimir function C. (Because of technical problems, one cannot guarantee that Casimirs can be explicitly found in all cases).

D. *Convexity estimates*

Find quadratic forms Q_1 *and* Q_2 *on* P *such that*

$$Q_1(\Delta u) \leqslant H(u_e + \Delta u) - H(u_e) - \mathrm{D}H(u_e) \cdot \Delta u \tag{A 3}$$

and

$$Q_2(\Delta u) \leqslant C(u_e + \Delta u) - C(u_e) - \mathrm{D}C(u_e) \cdot \Delta u \tag{A 4}$$

for all Δu *in* P. *Require that*

$$Q_1(\Delta u) + Q_2(\Delta u) > 0 \tag{A 5}$$

for all Δu *in* P, $\Delta u \neq 0$.

Remark D. Formal stability-second variation. As a prelude to checking (A 3), (A 4), and (A 5) it is often convenient to see whether the second variation $\mathrm{D}^2 H_C(u_e) \cdot (\Delta u)^2$, is definite, or when feasible, whether $\mathrm{D}^2 H_C(u_e)$ restricted to the symplectic leaf through u_e is definite. This is a prerequisite for step (D) to work, but it is not sufficient (see also remark 2 below).

E. *A priori estimates*

If steps (A) *to* (D) *have been carried out, then for any solution u of* (A 1) *we have the following estimate on* $\Delta u = u - u_e$:

$$Q_1(\Delta u(t)) + Q_2(\Delta u(t)) \leqslant H_C(u(0)) - H_C(u_e) \tag{A 6}$$

(This is proved below.)

F. *Nonlinear stability*

Suppose steps (A) *to* (D) *have been carried out. Then if we set*

$$\|\boldsymbol{v}\|^2 = Q_1(\boldsymbol{v}) + Q_2(\boldsymbol{v}) > 0 \;(\text{for}\, \boldsymbol{v} \neq 0), \tag{A 7}$$

so $\|\boldsymbol{v}\|$ *defines a norm on* P, *and if* H_C *is continuous in this norm at* u_e *and provided solutions to* (A 1) *exist, then* u_e *is nonlinearly stable. Should solutions to* (A 1) *be unknown to exist for all time, we still have 'conditional stability:' stability while* C^1 *solutions exist. A sufficient condition for continuity of* H_C *is the existence of positive constants* C_1 *and* C_2 *such that*

$$H(u_e+\Delta u)-H(u_e)-\mathrm{D}H(u_e)\cdot\Delta u \leqslant C_1\|\Delta u\|^2 \tag{A 8}$$

$$C(u_e+\Delta u)-C(u_e)-\mathrm{D}C(u_e)\cdot\Delta u \leqslant C_2\|\Delta u\|. \tag{A 9}$$

In this case, one gets the stability estimate

$$\begin{aligned}\|\Delta u(t)\|^2 := Q_1(\Delta u(t))+Q_2(\Delta u(t)) &\leqslant C_1\,Q_1(\Delta u(0))+C_2\,Q_2(\Delta u(0)) \\ &\leqslant (C_1+C_2)\,\|\Delta u(0)\|^2.\end{aligned} \tag{A 10}$$

(These assertions are proved below.)

Proof of a priori estimate (A 6). Adding (A 8) and (A 9) gives

$$\begin{aligned}Q_1(\Delta u)+Q_2(\Delta u) &\leqslant H_C(u_e+\Delta u)-H_C(u_e)-\mathrm{D}H_C(u_e)\cdot\Delta u \\ &= H_C(u_e+\Delta u)-H_C(u_e)\end{aligned} \tag{A 11}$$

since $\mathrm{D}H_C(u_e)=0$ by step (C). Because H_C is a constant of the motion, $H_C(u_e+\Delta u)-H_C(u_e)$ equals its value at $t=0$, which is (A 6). □

Proof of the assertion in step (F). We prove nonlinear (Liapunov) stability of u_e as follows. Given $\epsilon>0$, find a δ such that $\|u-u_e\|<\delta$ implies $|H_C(u)-H_C(u_e)|<\epsilon$. Thus, if $\|u(0)-u_e\|<\delta$, then (E) gives

$$\|u(t)-u_e\| \leqslant |H_C(u(0))-H_C(u_e)|<\epsilon. \tag{A 12}$$

Thus, $u(t)$ never leaves the ϵ-ball about u_e if it starts in the δ-ball, so u_e is nonlinearly stable. To see that (A 8) and (A 9) suffice for continuity of H_C at u_e add them to give, as in the proof of (A 6),

$$|H_C(u_e+\Delta u)-H_C(u_e)| \leqslant (C_1+C_2)\,\|\Delta u\|^2 \tag{A 13}$$

which implies that H_C is continuous at u_e. □

Remark

(1) In many examples, Q_1 and Q_2 are each positive (so H and C are individually convex). Then (A 5) is automatic. However, as already noted by Arnold (1969), there are some interesting examples where Q_1 is positive, Q_2 is negative, and yet Q_1 'beats' Q_2 and (A 5) is valid. If Q_2 'beats' Q_1 so Q_1+Q_2 is negative, then one can apply analogous procedures with $H+C$ replaced by $-(H+C)$.

(2) It has been presumed that P carries a Banach space topology (although one could merely assume P is Fréchet space) relative to which the symbols $\dot{u}$ and $\mathrm{D}H(u_e)$ are defined, and the conditions (A), (B), and (C) are verified. The norm $\|\cdot\|$ found in step (F) is usually not complete, and relative to it, the functions H and F need to be differentiable. (This fact is related to the difficulty one encounters when trying to prove nonlinear stability from formal stability). A sufficient condition for (A 8) is that

$$Q_1(v) \leqslant \mathrm{D}^2H(u)\cdot(v,v) \tag{A 14}$$

holds for all u and v in P. The sufficiency of (A 14) follows from the mean value theorem. There are similar assertions for C and H_C. Note that $\|v\|^2 \leqslant D^2H_C(u)(v, v)$ is considerably stronger than formal stability: $D^2H_C(u_e)(v, v) \geqslant 0$. Indeed, it is a global convexity condition which reflects the additional hypothesis involved in step (D).

(3) In examples where solutions form shocks, the solutions leave the space P and the stability algorithm may apply only up to the first shock time. Shocks may form, for example, in compressible flow; see Holm *et al.* (1983), (1985), for discussion of conditional stability for cases in which shocks may form.

(4) More delicate analytic techniques than those employed in the examples here are sometimes needed to obtain the convexity estimates. This occurs in the stability of the circular vortex patch in two-dimensional incompressible flow that was proved by Wan & Pulverenti (1984).

(5) As already noted, in systems with a finite number of degrees of freedom, formal stability implies nonlinear stability. This fact was used by Arnold (1966) to reproduce the well known results on stability of rigid body motion. See Marsden & Weinstein (1974) for the relationship of the formal stability ideas to the stability of relative equilibria and reduction.

(6) For Hamiltonian systems with additional symmetries, there will be additional constants of the motion besides Casimirs. These are to be incorporated into C in step B. This is needed in fluid examples with a translational symmetry, for example, and in the stability analysis of a heavy top. (See Holm *et al.* (1985).)

(7) For two-dimensional incompressible flow, the appropriate Casimir function is the generalized enstrophy. This suggests, as mentioned in Bretherton & Haidvogel (1976), that the Casimir functions may play a role in the 'selective decay hypothesis' when dissipation is added.

(8) As already noted in the main body of the paper, it is often necessary to define the norm on more variables than the original dynamical ones. For example, the expressions (3.34) and (6.34) define norms on $(\Delta \boldsymbol{M}, \Delta\rho, \Delta q)$, while the same expressions regarded only as functions of $\Delta \boldsymbol{M}$ and $\Delta\rho$ do *not* define a norm.

Appendix B. Semidirect products and the Lagrange-to-Euler map

Having seen how the examples in §7 are connected to semidirect products, we shall present here a brief summary of the theory in Ratiu (1982), Marsden *et al.* (1983), Marsden *et al.* (1984*a*, *b*), and Holm & Kupershmidt (1983) and Holm, Kupershmidt & Levermore (1983).

Let there be given a Lie group G and a representation ϕ of G on a vector space V, i.e. ϕ is a smooth group homomorphism from G to the automorphism group of V. In our three-dimensional examples, G was $\mathrm{Diff}_{\mathrm{vol}}(\mathrm{D})$, V was $\mathscr{F}(\mathrm{D})$, and the representation of G on V was given by push-forward. In the two-dimensional examples, G was the group of diffeomorphisms of D and the representation is again push-forward. Let $\mathscr{g}$ denote the left Lie algebra of G. As we remarked in §7, the left Lie algebra of $\mathrm{Diff}_{\mathrm{vol}}(\mathrm{D})$ is $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ endowed with *minus* the usual Lie bracket of vector fields. In the two-dimensional examples, $\mathscr{g}$ is $\mathscr{F}(\mathrm{D})$ with *minus* the (z, x)-Poisson bracket. Taking derivatives, ϕ induces a Lie algebra representation on V, i.e. a bracket-preserving linear map ϕ' from $\mathscr{g}$ to the space of all linear maps on V. In the three-dimensional examples, ϕ' is given by minus the Lie derivative and in the two-dimensional ones by minus the (z, x)-Poisson bracket with the argument on the left.

Now we form the semidirect product group G Ⓢ V defined as the Lie group with the underlying manifold G × V and composition law

$$(g_1, u_1)\,(g_2, u_2) = (g_1\, g_2,\, u_1 + \phi(g_1)\, u_2), \tag{B 1}$$

for $g_1, g_2 \in G$, $u_1, u_2 \in \mathrm{V}$. The identity element and inverse are given by (1, 0) and

$$(g, u)^{-1} = (g^{-1} - \phi(g^{-1})\, u) \tag{B 2}$$

The Lie algebra of G Ⓢ V is the semidirect product Lie algebra $\mathscr{g}$ Ⓢ V with underlying vector space $\mathscr{g} \times \mathrm{V}$ and Lie bracket

$$[(\zeta_1, v_1), (\zeta_2, v_2)] = ([\zeta_1, \zeta_2], \phi'(\zeta_1)\, v_2 - \phi'(\zeta_2)\, v_1). \tag{B 3}$$

In our examples, the relevant Lie brackets are (7.15) and (7.69).

The dual $\mathscr{h}^*$ of *any* Lie algebra $\mathscr{h}$ has a Poisson bracket, called the *Lie–Poisson bracket*, defined by

$$\{F, H\}(\mu) = \langle \mu, [\delta F/\delta\mu, \delta H/\delta\mu]\rangle \tag{B 4}$$

where $\mu \in \mathscr{h}^*$, $F, H\colon \mathscr{h}^* \to \mathbb{R}$ and the functional derivative is just the usual derivative (Jacobian matrix) of F regarded as an element of $\mathscr{h}$ rather than $\mathscr{h}^*$, i.e.

$$\langle \delta F/\delta\mu, \delta\mu\rangle := \mathrm{D}F(\mu)\cdot\delta\mu, \tag{B 5}$$

where the increment $\delta\mu \in \mathscr{h}^*$ and $\langle\, ,\rangle$ denotes the pairing between $\mathscr{h}^*$ and $\mathscr{h}$. Explicit pairings are given by (7.6) and (7.16). (As is explained in, for example, Marsden *et al.* (1983), the Lie–Poisson bracket comes from the canonical bracket on T*G by reduction.) The underlying vector space of $\mathscr{g}$ Ⓢ V is just $\mathscr{g} \times \mathrm{V}$, whose dual is $\mathscr{g}^* \times \mathrm{V}^*$, which has the Lie–Poisson bracket given by

$$\{F, H\}(\mu, a) = \left\langle \mu, \left[\frac{\delta F}{\delta\mu}, \frac{\delta H}{\delta\mu}\right]\right\rangle + \left\langle a, \phi'\left(\frac{\delta F}{\delta\mu}\right)\frac{\delta H}{\delta a}\right\rangle - \left\langle a, \phi'\left(\frac{\delta H}{\delta\mu}\right)\frac{\delta F}{\delta a}\right\rangle, \tag{B 6}$$

where $\mu \in \mathscr{g}^*$, $a \in \mathrm{V}^*$, $\delta F/\delta\mu$ and $\delta H/\delta\mu \in \mathscr{g}$, and $\delta F/\delta a$, $\delta H/\delta a \in \mathrm{V}$. An inspection of the brackets (7.15), (7.69) shows that they are Lie–Poisson on the duals of the indicated semidirect products. It should also be noted that the Lie–Poisson bracket is degenerate, i.e. there exist functions C, called *Casimir functions* which commute with all other functions.

A useful criterion for checking whether a given function is a Casimir, is its invariance under the coadjoint action of the Lie group on the dual of its Lie algebra. The general formula for the coadjoint action in $(\mathscr{g}$ Ⓢ $\mathrm{V})^* = \mathscr{g}^* \times \mathrm{V}^*$ is

$$\mathrm{Ad}^*_{(g,u)^{-1}}(\nu, a) = (\mathrm{Ad}^*_{g^{-1}}\,\nu + (\phi'_u)^*(\phi(g^{-1})^* a), \quad \phi(g^{-1})^* a), \tag{B 7}$$

where $g \in \mathrm{G}$, $u \in \mathrm{V}$, $\nu \in \mathscr{g}^*$, $a \in \mathrm{V}^*$, $\mathrm{Ad}^*_{g^{-1}}$ is the usual coadjoint action of G on $\mathscr{g}^*$, the upper star on maps denotes the dual maps induced between the dual vector spaces, and $\phi'_u\colon \mathscr{g} \to \mathrm{V}$ is the map given by $\phi'_u(\zeta) = \phi'(\zeta)\, u$, for $\zeta \in \mathscr{g}$. Explicitly, the coadjoint action on $\mathscr{X}_{\mathrm{div}}(\mathrm{D})$ Ⓢ-$\mathscr{F}(\mathrm{D}))^* = \mathscr{X}_{\mathrm{div}}(\mathrm{D}) \times \mathscr{F}(\mathrm{D})$ is given by (see 7.30)

$$\mathrm{Ad}^*_{(\eta, f)^{-1}}(\boldsymbol{v}, \rho) = ((T\eta^{-1})^\dagger \circ \boldsymbol{v} \circ \eta^{-1} + \boldsymbol{P}(\eta_* \rho \nabla f), \eta_* \rho), \tag{B 8}$$

where $\boldsymbol{P}$ denotes projection of a vector field into its component with zero divergence, tangent to the boundary, in the Helmholtz–Hodge decomposition, and † denotes adjoint.

With the aid of (B 8) one can easily verify that $C_{F,G}$ and C_F given respectively by (2.15) and (3.7) are invariant under the coadjoint action and hence are Casimir functions.

Remark. Since the underlying vector space for $\mathscr{X}_{\text{div}}(\mathrm{D})$ Ⓢ $\mathscr{F}(\mathrm{D})$ is $\mathscr{X}_{\text{div}}(\mathrm{D}) \times \mathscr{F}(\mathrm{D})$ and $\mathscr{X}_{\text{div}}(\mathrm{D})^*$ is identified with $\mathscr{X}_{\text{div}}(\mathrm{D})$ by (7.6) and $\mathscr{F}(\mathrm{D})^*$ with $\mathscr{F}(\mathrm{D})$ by the L^2 pairing (see (7.16)), it follows that $(\mathscr{X}_{\text{div}}(\mathrm{D})$ Ⓢ $\mathscr{F}(\mathrm{D}))^*$ is identified with $\mathscr{X}_{\text{div}}(\mathrm{D}) \times \mathscr{F}(\mathrm{D})$.

Returning to the general case, consider the map

$$J: \mathrm{T^*G} \times \mathrm{V^*} \rightarrow (\mathfrak{g} \circledS \mathrm{V})^*,$$

$$J(\alpha_g, a) = (\mathrm{T}_{\mathrm{e}}^* R_g(\alpha_g), \quad \phi(g^{-1})^* \boldsymbol{a}), \tag{B 9}$$

where $\mathrm{T_e} R_g: \mathrm{T_e G} = \mathfrak{g} \rightarrow \mathrm{T}_g \mathrm{G}$ denotes the derivative of the right translation $R_g: \mathrm{G} \rightarrow \mathrm{G}$, defined by $R_g(h) = hg$, and $\mathrm{T}_{\mathrm{e}}^* R_g: \mathrm{T}_{\mathrm{g}}^* \mathrm{G} \rightarrow \mathfrak{g}^*$ is its dual; $\mathrm{T_g G}$ and $\mathrm{T}_g \mathrm{G}$ are the tangent and cotangent spaces of G at $g \in \mathrm{G}$. A crucial general fact is that J is *canonical*, i.e. *preserves Poisson brackets*, where $\mathrm{T^*G} \times \mathrm{V^*}$ has the canonical Poisson bracket on the T*G part and $(\mathfrak{g} \circledS \mathrm{V})^*$ has the Lie–Poisson bracket. This map transforms Lagrangian variables, i.e. variables in $\mathrm{T^*G} \times \mathrm{V^*}$ (for example, material velocity and material density function) to Eulerian variables (for example, $\boldsymbol{v}$, ρ; see formula (7.60)). The canonical nature of the map J is proved in two different ways. First, it can be shown that J is the momentum map of a natural action of the semidirect product G Ⓢ V on $\mathrm{T^*G} \times \mathrm{V^*}$. Second, the map J is obtained via reduction by the subgroup V of G Ⓢ V of the momentum map of right translation of G Ⓢ V on its cotangent bundle $\mathrm{T^*(G \circledS V)} = \mathrm{T^*G} \times \mathrm{V} \times \mathrm{V^*}$. We mention that in some applications the V variable, while cyclic, is useful as, for example, in Schutz (1970). The canonical nature of J is proved in Marsden *et al.* (1984*b*) as a consequence of the results of Marsden *et al.* (1984*a*). This canonical nature is shown explicitly for several, more general cases (including superfluids and nonabelian Yang–Mills plasmas) by Holm, Kupershmidt & Levermore (1983). The Hamiltonian $H(\alpha_g, a)$ in Lagrangian variables (α_g, a) induces a Hamiltonian $H_{\mathrm{E}}(\mu, a)$ in Eulerian variables (μ, a), provided that $H_a(\alpha_g) := H(\alpha_g, a)$ is compatible with the G-action, i.e.

$$H_{\phi(g)^* a} \circ \mathrm{T}^* R_{g^{-1}} = H_a, \tag{B 10}$$

for all $g \in \mathrm{G}$, $a \in \mathrm{V^*}$. Thus, in particular, $H(\alpha_g, a)$ is invariant under the stabilizer $\mathrm{G}_a = \{g \in \mathrm{G} \mid \phi(g)^* a = a\}$ (for example, see (7.55)). The relation between H and H_{E} is given by

$$H(\alpha_g, a) = H_{\mathrm{E}}(\mathrm{T}_{\mathrm{e}}^* R_g(\alpha_g), \quad \phi(g^{-1})^* a). \tag{B 11}$$

The equations of motion $\dot{F} = \{F, H\}$ in Lagrangian coordinates imply those for Eulerian coordinates: $\dot{F} = \{F, H_{\mathrm{E}}\}_{\text{Lie–Poisson}}$. This follows from the invariance of H_{E} under G_a and general principles.

We conclude this Appendix with some general remarks. In many examples, one is given the phase space T*G, but it is not obvious *a priori* what V and ϕ should be. The phase space T*G is interpreted as 'material' or 'Lagrangian' coordinates. This means that the Hamiltonian might be given directly on a space of the form $\mathfrak{g}^* \times \mathrm{V^*}$, where the evolution of the V* variable is by 'dragging along' or 'Lie transport,' i.e. it is of the form $t \mapsto \phi(g(t)^{-1})^* a$, where $a \in \mathrm{V}$ and $g(t)$ is the solution curve in the configuration space G. This evolution determines the representation ϕ. The parameter $a \in \mathrm{V^*}$ often appears in the form of an initial condition on some physical variable of the given problem. In our examples, the role of the parameter, a, is played by the initial density configuration, or equivalently, by the Lagrangian mass density.

Appendix C. The Hamiltonian structure of the linearized equations and the Taylor–Goldstein equation

In this Appendix, we will show that the equations linearized about an equilibrium solution of a Lie–Poisson system (such as the ideal fluid equations) are Hamiltonian with respect to a 'constant coefficient' Lie–Poisson bracket. The Hamiltonian for these linearized equations is $\frac{1}{2}\delta^2 H_{C|e}$, the quadratic functional obtained by taking one-half of the second variation of the Hamiltonian plus conserved quantities and evaluating it at the equilibrium solution where the first variation δH_C vanishes. An immediate consequence is that the linearized dynamics preserves $\delta^2 H_{C|e}$. We will also show that formal stability of the stationary solution implies its linear (Liapunov) stability. Finally, the Taylor–Goldstein equation will be derived using this Hamiltonian formalism. This equation concerns the spectrum of the linearized equations. It will be compared to the condition for the positivity of the second variation.

For a Lie algebra $\mathfrak{g}$, the Lie–Poisson bracket is defined on $\mathfrak{g}^*$, the dual of $\mathfrak{g}$ with respect to a weakly non-degenerate pairing $\langle\,,\rangle$ between $\mathfrak{g}^*$ and $\mathfrak{g}$ by

$$\{F, G\}(\mu) = \langle \mu, [\delta F/\delta\mu, \delta G/\delta\mu]\rangle \tag{C 1}$$

where $\delta F/\delta\mu \in \mathfrak{g}$ is determined by

$$\mathrm{D}F(\mu)\cdot\delta\mu = \langle \delta\mu, \delta F/\delta\mu\rangle, \tag{C 2}$$

when such an element $\delta F/\delta\mu$ exists, for any μ, $\delta\mu \in \mathfrak{g}^*$. The equations of motion are easily seen to be

$$\mathrm{d}\mu/\mathrm{d}t = -\mathrm{ad}\,(\delta H/\delta\mu)^*\mu \tag{C 3}$$

where $H:\mathfrak{g}^* \rightarrow \mathbb{R}$ is the Hamiltonian, $\mathrm{ad}(\xi):\mathfrak{g}\rightarrow\mathfrak{g}$ is the adjoint action, $\mathrm{ad}(\xi)\cdot\eta = [\xi, \eta]$ for $\xi, \eta \in \mathfrak{g}$ and $\mathrm{ad}(\xi)^*:\mathfrak{g}^*\rightarrow\mathfrak{g}^*$ is its dual. Let $\mu_e \in \mathfrak{g}^*$ be an equilibrium solution of (C 3). The linearized equations of (C 3) at μ_e are obtained by expanding all quantities in a Taylor expansion with small parameter ϵ and taking $\mathrm{d}/\mathrm{d}\epsilon|_{\epsilon=0}$ of the resulting equations. For $\mu = \mu_e + \epsilon\delta\mu$, using the Taylor theorem gives

$$\frac{\delta H}{\delta\mu} = \frac{\delta H}{\delta\mu_e} + \epsilon\,\mathrm{D}\left(\frac{\delta H}{\delta\mu}\right)(\mu_e)\cdot\delta\mu + O(\epsilon^2), \tag{C 4}$$

where $\langle \delta H/\delta\mu_e, \delta\mu\rangle := \mathrm{D}H(\mu_e)\cdot\delta\mu$, and the derivative $\mathrm{D}(\delta H/\delta\mu)\,(\mu_e)\cdot\delta\mu$ is easily seen to equal the linear functional

$$\nu \in \mathfrak{g}^* \mapsto \mathrm{D}^2 H(\mu_e)\cdot(\delta\mu, \nu) \in \mathbb{R} \tag{C 5}$$

by using the definition (C 2). Since $\delta^2 H := \mathrm{D}^2 H(\mu_e)\cdot(\delta\mu, \delta\mu)$, it follows that the functional (C5) equals $\frac{1}{2}\delta(\delta^2 H)/\delta(\delta\mu)$. Consequently, (C 4) becomes

$$\frac{\delta H}{\delta\mu} = \frac{\delta H}{\delta\mu_e} + \tfrac{1}{2}\epsilon\frac{\delta(\delta^2 H)}{\delta(\delta\mu)} + O(\epsilon^2) \tag{C 6}$$

and the Lie–Poisson equations (C 3) yield

$$\frac{\mathrm{d}\mu_e}{\mathrm{d}t} + \epsilon\frac{\mathrm{d}(\delta\mu)}{\mathrm{d}t} = -\mathrm{ad}\left(\frac{\delta H}{\delta\mu_e}\right)^*\mu_e - \tfrac{1}{2}\epsilon\left[\mathrm{ad}\left(\frac{\delta(\delta^2 H)}{\delta(\delta\mu)}\right)^*\mu_e - \mathrm{ad}\left(\frac{\delta H}{\delta\mu_e}\right)^*\delta\mu\right] + O(\epsilon^2).$$

Thus, the linearized equations are

$$\frac{\mathrm{d}(\delta\mu)}{\mathrm{d}t} = -\tfrac{1}{2}\mathrm{ad}\left(\frac{\delta(\delta^2 H)}{\delta(\delta\mu)}\right)^* \mu_{\mathrm{e}} - \mathrm{ad}\left(\frac{\delta H}{\delta\mu_{\mathrm{e}}}\right)^* \delta\mu. \tag{C 7}$$

Now, if H is replaced by $H_C := H + C$, with the function C satisfying $\delta H_C/\delta\mu_{\mathrm{e}} = 0$, we get $\mathrm{ad}\,(\delta H_C/\delta\mu_{\mathrm{e}})^*\mu_{\mathrm{e}} = 0$, and so

$$\frac{\mathrm{d}(\delta\mu)}{\mathrm{d}t} = -\tfrac{1}{2}\mathrm{ad}\left(\frac{\delta(\delta^2 H_C)}{\delta(\delta\mu)}\right)^* \mu_{\mathrm{e}}. \tag{C 8}$$

This equation is Hamiltonian with respect to the Poisson bracket

$$\{F, G\}(\mu) = \langle \mu_{\mathrm{e}}, [\delta F/\delta\mu, \delta G/\delta\mu]\rangle. \tag{C 9}$$

That this bracket satisfies the Jacobi identity is proved in Guillemin & Sternberg (1981), Ratiu (1982) and Weinstein (1984); the second paper also interprets it in terms of a Lie–Poisson structure of a loop extension of $\mathfrak{g}$. The Poisson bracket (C 9) differs from the Lie–Poisson bracket (C 1) in that it is *constant* in μ. With respect to the Poisson bracket (C 9), Hamilton's equations given by $\delta^2 H_C$ are (C 8), as an easy verification shows. Note that the critical points of $\delta^2 H_C$ are stationary solutions of the linearized equation (C 8), i.e. they are *neutral modes* for (C 8).

Finally, note that if $\delta^2 H_C$ is definite, then either $\delta^2 H_C$, or $-\delta^2 H_C$ is positive definite and, hence, defines a norm on the space of perturbations $\delta\mu$ (which is $\mathfrak{g}^*$). Being twice the Hamiltonian function for (C 8), $\delta^2 H_C$ is conserved. So, any solution of (C 8) starting on an energy surface of $\delta^2 H_C$ (i.e. on a sphere in this norm) stays on it and, hence, the zero solution of (C 8) is (Liapunov) stable. *Thus, formal stability*, (i.e. $\delta^2 H_C$ definite) *implies linearized stability*. It should be noted, however, that the conditions for definiteness of $\delta^2 H_C$ are entirely different from the conditions for 'normal mode stability,' i.e. that the operator acting on $\delta\mu$ given by (C 8) have purely imaginary spectrum. In particular, having purely imaginary spectrum for the linearized equation does *not* produce Liapunov stability of the linearized equations. The difference between $\delta^2 H_C$ and the operator in (C 8) can be made explicit, as follows. Assume that the pairing $\langle\,,\rangle$ identifies the dual $\mathfrak{g}^*$ with $\mathfrak{g}$ itself, i.e. there is a weak invariant metric $\langle\langle\,,\rangle\rangle$ on $\mathfrak{g}$. Then

$$\delta^2 H_C = \langle\langle \delta\mu, L\delta\mu \rangle\rangle, \tag{C 10}$$

for $L: \mathfrak{g} \to \mathfrak{g}$ a linear operator, symmetric with respect to the metric $\langle\langle\,,\rangle\rangle$, i.e. $\langle\langle \alpha, L\beta\rangle\rangle = \langle\langle L\alpha, \beta\rangle\rangle$ for all $\alpha, \beta \in \mathfrak{g}$. Then the linear operator in (C 8) becomes

$$\delta\mu \mapsto [L\delta\mu, \mu_{\mathrm{e}}] \tag{C 11}$$

which of course, differs from L, in general. However, note that the kernel of L is included in the kernel of the linear operator (C 11), i.e. the zero eigenvalues of L give rise to 'neutral modes' in the spectral analysis of (C 11). There is remarkable coincidence of the zero-eigenvalue equations for these operators in fluid mechanics: for the Rayleigh equation describing plane-parallel shear flow in an inviscid homogeneous fluid, taking normal modes makes the zero-eigenvalue equations corresponding to L and to (C 11) coincide (see Holm *et al.* (1985)).

We shall devote the rest of this Appendix to the derivation of the Taylor–Goldstein equation for the two-dimensional stratified Boussinesq flow. In order to simplify certain computations, it is convenient to set $\psi|_{\partial\mathrm{D}} = 0$ and to use, for such flows, the Hamiltonian formulation in terms of ω and ρ, with $\omega = \nabla^2\psi$. The condition $\psi|_{\partial\mathrm{D}} = 0$ can always be achieved for plane parallel flows, in a channel invariant under translation in the x-direction, by moving to a frame in which the total mass flux across a vertical line is zero. (In this case, we are dealing with travelling wave solutions that are steady in a certain frame.) It is shown in §7 C that if $\psi|_{\partial D} = 0$ the vorticity equations for stratified Boussinesq flow

$$\partial\omega/\partial t = \{\omega, \psi\} + \{gz/\rho_*, \rho\} \qquad \text{(C 12)}$$
$$\partial\rho/\partial t = \{\rho, \psi\},$$

where $\{f, g\} = \partial_z f \partial_x g - \partial_x f \partial_z g$, are Hamiltonian (see section 3C) with Lie–Poisson bracket given by equation (7.69) on the dual of the semidirect product of functions on D vanishing on ∂D with functions on D

$$\{F, G\}(\omega, \rho) = \int_D \left[\omega\left\{\frac{\delta F}{\delta\omega}, \frac{\delta G}{\delta\omega}\right\} + \rho\left\{\frac{\delta F}{\delta\rho}, \frac{\delta G}{\delta\omega}\right\} - \rho\left\{\frac{\delta G}{\delta\rho}, \frac{\delta F}{\delta\omega}\right\}\right] \mathrm{d}x\,\mathrm{d}z \qquad \text{(C 13)}$$

and Hamiltonian

$$H(\omega, \rho) = \int_D \tfrac{1}{2}|\nabla\psi|^2\, \mathrm{d}x\,\mathrm{d}z + \int (\rho/\rho_*)\, gz\, \mathrm{d}x\,\mathrm{d}z. \qquad \text{(C 14)}$$

The Casimir functions are,

$$C_{F,G}(\omega, \rho) = \int_D (\omega F(\rho) + G(\rho))\, \mathrm{d}x\,\mathrm{d}z, \qquad \text{(C 15)}$$

as discussed before, in equation (2.15). The equilibrium conditions imply vanishing of the first variation of $H_C := H + C_{F,G} + \lambda \int_D \omega\, \mathrm{d}x\,\mathrm{d}z$ at a stationary solution (ω_e, ρ_e). These conditions are

$$F(\rho_e) = \psi_e, \qquad \text{(C 16)}$$

$$gz/\rho_* + \omega_e F'(\rho_e) + G'(\rho_e) = 0, \qquad \text{(C 17)}$$

$$\lambda + F(\rho_e)|_{\partial D} = 0. \qquad \text{(C 18)}$$

With respect to the linearized Poisson bracket

$$\{F, G\}(\nu, \sigma) = \int\left[\omega_e\left\{\frac{\delta F}{\delta\nu}, \frac{\delta G}{\delta\nu}\right\} + \rho_e\left\{\frac{\delta F}{\delta\sigma}, \frac{\delta G}{\delta\nu}\right\} - \rho_e\left\{\frac{\delta G}{\delta\sigma}, \frac{\delta F}{\delta\nu}\right\}\right] \mathrm{d}x\,\mathrm{d}z, \qquad \text{(C 19)}$$

the Hamilton equations for the linearized motion at ω_e, ρ_e, are

$$\left.\begin{aligned}\partial\nu/\partial t &= \{\delta h/\delta\nu, \omega_e\} + \{\delta h/\delta\sigma, \rho_e\},\\ \partial\sigma/\partial t &= \{\delta h/\delta\nu, \rho_e\},\end{aligned}\right\} \qquad \text{(C 20)}$$

where $h(\nu, \sigma) = \frac{1}{2}\delta^2 H = \frac{1}{2}D^2 H(\omega_e, \rho_e)(\nu, \sigma)^2$. Since $\psi|_{\partial D} = 0$ and $\nu = \delta\omega$, we conclude that $\delta\psi = (\nabla^2)^{-1}\nu$ and $\delta\psi|_D = 0$. Thus,

$$\delta h/\delta\nu = -(\nabla^2)^{-1}\nu + F'(\rho_e)\,\sigma := -\bar{\phi} + F'(\rho_e)\,\sigma, \qquad \text{(C 21)}$$

$$\delta h/\delta\sigma = F'(\rho_e)\,\nu + [\omega_e F''(\rho_e) + G''(\rho_e)]\,\sigma. \qquad \text{(C 22)}$$

Equations (C 20) become

$$\frac{\partial\nu}{\partial t} = -\frac{\partial\omega_e}{\partial z}\frac{\partial}{\partial x}(F'(\rho_e)\,\sigma - \bar{\phi}) - \frac{\partial\rho_e}{\partial z}\frac{\partial}{\partial x}[F'(\rho_e)\,\nu + (\omega_e F''(\rho_e) + G''(\rho_e))\,\sigma], \qquad \text{(C 23)}$$

$$\frac{\partial\sigma}{\partial t} = -\frac{\partial\rho_e}{\partial z}\frac{\partial}{\partial x}(F'(\rho_e)\,\sigma - \bar{\phi}). \qquad \text{(C 24)}$$

Setting

$$\left.\begin{aligned}\bar{\phi}(x,z) &= \mathrm{e}^{\mathrm{i}\alpha(x-ct)}\phi(z),\\ \sigma(x,z) &= \mathrm{e}^{\mathrm{i}\alpha(x-ct)}\chi(z),\end{aligned}\right\} \tag{C 25}$$

gives $$\nu = \nabla^2\bar{\phi} = (\phi''(z)-\alpha^2\phi)\,\mathrm{e}^{\mathrm{i}\alpha(x-ct)}.$$

Now solving for χ in (C 23) leads to

$$\chi = \frac{\partial\rho_e/\partial z}{(\partial\rho_e/\partial z)\,F'(\rho_e)-c}\,\phi. \tag{C 26}$$

By (C 15), we have

$$\frac{\partial\rho_e}{\partial z}F'(\rho_e) = \frac{\partial}{\partial z}(F(\rho_e)) = \frac{\partial\psi_e}{\partial z} = U(z), \tag{C 27}$$

since $\boldsymbol{v}_e = (-\partial\psi_e/\partial z, \partial\psi_e/\partial x)$ according to the definition of the stream function. Thus

$$\chi = \left[\frac{\partial\rho_e/\partial z}{U(z)-c}\right]\phi. \tag{C 28}$$

Equation (C 22) becomes

$$\begin{aligned} c[\phi''(z)-\alpha^2\phi(z)] = \frac{\partial\omega_e}{\partial z}(F'(\rho_e)\,\chi-\phi(z))+\frac{\partial\rho_e}{\partial z}F'(\rho_e)(\phi''(z)-\alpha^2\phi(z))\\ +\frac{\partial\rho_e}{\partial z}(\omega_e F''(\rho_e)+G''(\rho_e))\,\chi, \end{aligned} \tag{C 29}$$

i.e. $$0 = (\mathrm{U(z)-c})(\phi''(z)-\alpha^2\phi)-\frac{\partial\omega_e}{\partial z}\phi(z)+\left[\frac{\partial\omega_e}{\partial z}F'(\rho_e)+\frac{\partial\rho_e}{\partial z}(\omega_e F''(\rho_e)+G''(\rho_e))\right]\chi. \tag{C 30}$$

Since $\omega_e(z) = U'(z)$, $\partial\omega_e/\partial z = U''(z)$, and by (C 17)

$$\begin{aligned}\left[\frac{\partial\omega_e}{\partial z}F'(\rho_e)+\frac{\partial\rho_e}{\partial z}(\omega_e F''(\rho_e)+G''(\rho_e))\right]\chi &= \chi\frac{\mathrm{d}}{\mathrm{d}z}[\omega_e F'(\rho_e)+G'(\rho_e)]\\ &= \chi\frac{\mathrm{d}}{\mathrm{d}z}\left(-\frac{gz}{\rho_*}\right)\\ &= \frac{g}{\rho_*}\chi,\end{aligned} \tag{C 31}$$

we get from the prior relation (C 30) that

$$0 = (U(z)-c)\,(\phi''(z)-\alpha^2\phi(z))-U''(z)\,\phi(z)-\frac{(g/\rho_*)\,(\partial\rho_e/\partial z)}{U-c}\phi(z). \tag{C 32}$$

Denoting the Brunt–Väisälä frequency by

$$N^2(z) = -\frac{g}{\rho_*}\frac{\partial\rho_e}{\partial z}, \tag{C 33}$$

(C 32) becomes the *Taylor–Goldstein equation*

$$(U(z)-c)(\phi''(z)-\alpha^2\phi(z))-U''(z)\,\phi+\frac{N^2(z)}{U(z)-c}\phi = 0; \tag{C 34}$$

see, for example, equation (44.10) in Drazin & Reid (1981), p. 324. Note that

$$N^2(z)/U'(z)^2 = Ri$$

is the local Richardson number. The standard eigenvalue analysis of equation (C 34), as presented in Miles (1961) and Howard (1961), states that a necessary condition for the phase velocity c to have non-zero imaginary part is that $Ri < \frac{1}{4}$ somewhere in the field of flow. This criterion, involving the spectrum of the linearized equation, is a necessary condition for linearized instability. However, the opposite inequality $Ri \geqslant \frac{1}{4}$ everywhere in the field of flow does *not* imply Liapunov stability of the zero solution of the linearized equation. Indeed, the nonlinear analysis in the example of §4 finds $Ri > 1$ as a sufficient condition for non-linear stability. In the 'no-man's land' $Ri \in [\frac{1}{4}, 1]$ we conjecture that nonlinear instability occurs.

Let us compare the linearized equations (C 20) with the eigenvalue equation of the operator L given by $\delta^2 H_C$ via (C 10). A short calculation shows that L is the operator matrix

$$L = \begin{bmatrix} -\nabla^2 & \nabla^2 \alpha(z) \\ \alpha(z)\nabla^2 & \beta(z) \end{bmatrix}, \tag{C 35}$$

where

$$\alpha(z) = F'(\rho_e) = U(z)/U''(z),$$

$$\beta(z) = -\left[\frac{g}{\rho_*} + \frac{U''(z)\,U(z)}{\rho_e'(z)}\right] / \rho_e'(z).$$

In the same notation, the linearized equations (C 20) corresponding to (C 8) are

$$\partial_t \begin{bmatrix} \nabla^2 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \delta\psi \\ \delta\rho \end{bmatrix} = \partial_x \begin{bmatrix} \omega_e'(z) - \rho_e'(z)\,\alpha(z)\,\nabla^2 & -\omega_e'(z)\,\alpha(z) - \rho_e'(z)\,\beta(z) \\ \rho_e'(z) & \rho_e'(z)\,\beta(z) \end{bmatrix} \begin{bmatrix} \delta\psi \\ \delta\rho \end{bmatrix}. \tag{C 36}$$

According to (C 8) and (C 11), zero-eigenvectors of L in (C 35) are neutral modes, i.e. stationary states of (C 35). The converse, of course, is not necessarily true: not all neutral modes need be zero-eigenvectors of L.

References

Abarbanel, H. D. I., Holm, D. D., Marsden, J. E. & Ratiu, T. 1984 Richardson number criterion for the nonlinear stability of three dimensional stratified flow. *Phys. Rev. Lett.* **52**, 2552–2555.

Arnold, V. I. 1965 Conditions for nonlinear stability of the stationary plane curvilinear flows of an ideal fluid. *Dokl. Mat. Nauk.* **162 (5)**, 773–777.

Arnold, V. I. 1969 An a priori estimate in the theory of hydrodynamic stability. [English translation] *Am. math. Soc. Transl.* **19**, 267–269.

Arnold, V. I. 1978 *Mathematical methods of classical mechanics.* (*Graduate Texts in Mathematics* no. 60). Berlin, Heidelberg and New York: Springer-Verlag.

Arnold, V. I. 1966 Sur la géometrie differentielle des groupes de Lie de dimension infinie et ses applications a l'hydrodynamique des fluids parfaits. *Annls. Inst. Fourier, Grenoble* **16**, 319–361.

Ball, J. M. & Marsden, J. E. 1984 Quasiconvexity, second variations and nonlinear stability in elasticity. *Arch ration. Mech. Analysis* **86**, 251–277.

Benjamin, T. B. 1972 The stability of solitary waves. *Proc. R. Soc. Lond.* A **328**, 153–183.

Benjamin, T. B. 1984 Impulse, flow force, and variational principles. *IMA J. appl. Math.* **32**, 3–68.

Bennett, D. P., Brown, R. W., Stansfield, S. E., Stroughair, J. D. & Bona, J. L. 1983 The stability of internal solitary waves. *Math. Proc. Camb. phil. Soc.* **94**, 351–361.

Benzi, R., Pierini, S., Vulpiani, A. & Salusti, E. 1982 On nonlinear hydrodynamic stability of planetary vortices. *Geophys. astrophys. Fluid Dyn.* **20**, 293–306.

Bernstein, I. B., Frieman, E. A., Kurskal, M. D. & Kulsrud, R. M. 1958 An energy principle for hydromagnetic stability problems. *Proc. R. Soc. Lond.* A, **244**, 17–40.

Blumen, W. 1971 On the stability of plane flow with horizontal shear to three-dimensional nondivergent disturbances. *Geophys. Fluid Dyn.* **2**, 189–200.

Bona, J. 1975 On the stability theory of solitary waves. *Proc. R. Soc. Lond.* A **344**, 363–374.

Bretherton, F. P., 1970 A note on Hamilton's principle for perfect fluids. *J. Fluid Mech.* **44**, 19–31.

Bretherton, F. P. & Haidvogel, D. B. 1976 Two-dimensional turbulence above topography. *J. Fluid Mech.* **78**, 129–154.

Cantor, M. 1979 Some problems of global analysis on asymptotically simple manifolds. *Compositio Mathematica* **38**, 3–35.

Chandrasekhar, S. 1961 *Hydrodynamic and hydromagnetic stability*. Oxford. Reprinted by Dover, 1981.

Davis, R. E. & Acrivos, A. 1967 Solitary internal waves in deep water. *J. Fluid Mech.* **29**, 593–608; **30**, 723–736.

Drazin, P. G. & Reid, W. H. 1981 *Hydrodynamic stability*. Cambridge University Press.

Dubreil-Jacotin, M. L. 1935 Complement a une note antériure sur les ondes de type permanent dans les liquides hétérogènes. *Atti Accad. naz. Lincei Rc.* (*Cl. Sci. Fis. Mat. Nat.*) **21**, 344–346.

Ebin, D. & Marsden, J. 1970 Groups of diffeomorphisms and the motion of an incompressible fluid. *Ann. Math.* **92**, 102–163.

Eliassen, A. & Kleinschmidt, E. 1957 Dynamic meteorology. In *Handbuch der Physik*, vol. 48 (*Geophysik II*) (ed. J. Bartels), pp. 1–154. Berlin, Heidelberg and New York: Springer-Verlag.

Fjörtoft, R. 1946 On the frontogenesis and cyclogenesis in the atmosphere. *Geophys. Publs, Dubl.* **16**, 1–28.

Gill, A. E. 1982 *Atmosphere-ocean dynamics*. San Diego: Academic Press.

Guillemin, V. & Sternberg, S. 1983 On collective complete integrability according to the method of Thimm. *Ergodic Theory and dynam. Systems* **3**, 219–230.

Henyey, F. S. 1982 Hamiltonian description of stratified fluid dynamics. *Physics Fluids.* **26**, 40–47.

Holm, D. D. & Kupershmidt, B. A. 1983 Poisson brackets and Clebsch representations for magnetohydrodynamics, multifluid plasmas, and elasticity. *Physica* **6D**, 347–363.

Holm, D. D., Kupershmidt, B. A. & Levermore, C. D. 1983 Canonical maps between Poisson brackets in Eulerian and Lagrangian descriptions of continuum mechanics, *Physics Lett.* **98A**, 389–395.

Holm, D. D., Kupershmidt, B. A. & Levermore, C. D. 1985 Hamiltonian differencing of fluid dynamics. *Adv. appl. Math.* **6**, 52–84.

Holm, D. D., Marsden, J. E. & Ratiu, T. 1986 Nonlinear stability of the Kelvin–Stuart cat's eyes flow in *AMS Lects. appl. Math.* (*Proceedings of AMS–SIAM summer seminar, Santa Fe*).

Holm, D. D., Marsden, J. E., Ratiu, T. & Weinstein, A. 1983 Nonlinear stability conditions and a priori estimates for barotropic hydrodynamics. *Physics Lett.* **98A**, 15–21.

Holm, D. D., Marsden, J. E., Ratiu, T. & Weinstein, A. 1985 Nonlinear stability of fluid and plasma equilibria. *Physics Rep.* **123**, 1–116.

Holmes, P. J. & Marsden, J. E. 1982 Melnikov's method and Arnold diffusion for perturbations of integrable Hamiltonian systems. *J. math. Phys.* **23**, 523–544.

Holmes, P. J., Marsden, J. E. 1983 Horseshoes and Arnold diffusion for Hamiltonian systems on Lie groups. *Indiana Univ. math. J.* **32**, 273–310.

Howard, L. N. 1961 Note on a paper of John W. Miles. *J. Fluid Mech.* **10**, 509–512.

Kruskal, M. D. & Oberman, C. R. 1958 On the stability of a plasma in static equilibrium. *Physics Fluids*, **1**, 275–280.

Kuo, H. L. 1949 Dynamic instability of two-dimensional non-divergent flow in a barotropic atmosphere. *J. Met.* **6**, 105–122.

LeBlond, P. H. & Mysak, L. A. 1978 *Waves in the ocean*. New York: Elsevier.

Leonov, A. I., Miropolsky, Yu Z. & Tamsalu, R. E. 1979 Nonlinear stationary internal and surface waves in shallow sea. *Tellus* **31**, 150–160.

Leslie, J. 1967 On a differential structure for the group of diffeomorphisms. *Topology* **6**, 263–271.

Lewis, D., Marsden, J., Montgomery, R. & Ratiu, T. 1985 Hamiltonian structures for dynamic free boundary problems. *Physica* **D** (In the press.)

Lewis, D., Marsden, J. & Ratiu, T. 1986 Formal stability of liquid drops with surface tension, *Proc. ONR Conf. on Nonlinear Dynamics*, (ed. M. Schlesinger), *World Sci. Publ. Co.*

Lichtenberg, A. J. & Lieberman, M. A. 1982 Regular and stochastic motion. *Appl. math. Sciences.* New York: Springer.

Lilly, D. K. & Klemp, J. B. 1979 The effects of terrain on nonlinear hydrostatic mountain waves. *J. Fluid Mech.* **95**, 241–261.

Long, R. R. 1953 Some aspects of the flow of stratified fluids I. A theoretical investigation. *Tellus* **5**, 42–57.

Marsden, J. 1976 Well-posedness of the equations of a non-homogeneous perfect fluid. *Communs P.D.E.*, **1**, 215–230.

Marsden, J. E., Ratiu, T. & Weinstein, A. 1984*a* Reduction and Hamiltonian structures on duals of semidirect product Lie algebras. *Contr. Math. AMS.* **28**, 55–100.

Marsden, J. E., Ratiu, T. & Weinstein, A. 1984*b* Semidirect products and reduction in mechanics. *Trans. Am. math. Soc.* **281**, 147–177.

Marsden, J. E. & Weinstein, A. 1974 Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.* **5**, 121–130.

Marsden, J. E. & Weinstein, A. 1983 Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica* **7D**, 305–323.
Marsden, J. E., Weinstein, A., Ratiu, T., Schmidt, R. & Spencer, R. 1983 Hamiltonian systems with symmetry, coadjoint orbits and plasma physics, *Atti. Acc. Sci. Torino*, Suppl. **117**, 289–340.
Miles, J. W. 1961 On the stability of heterogeneous shear flows, *J. Fluid Mech.* **10**, 496–508.
Morrison, P. J. 1982 Poisson brackets for fluids and plasmas, in *Mathematical methods in hydrodynamics and integrability in related dynamical systems.* AIP Conf. Proc. La Jolla, California (ed M. Tabor).
Newcomb, W. A. 1960 Hydromagnetic stability of a diffuse linear pinch. *Annls of Phys.* **10**, 232–267.
Pedlosky, J. 1979 *Geophysical fluid dynamics.* New York: Springer.
Phillips, O. M. 1977 *The dynamics of the upper ocean.* Cambridge University Press.
Rosenbluth, M. N. 1964 In *Advanced Plasma Physics.* (ed. M. Rosenbluth), p. 137 Academic Press.
Ratiu, T. 1982 Euler–Poisson equations on Lie algebras and the N-dimensional heavy rigid body. *Am. J. Math.* **104**, 409–448.
Schutz, B. F. 1970 Perfect fluids in general relativity – velocity potentials and a variational principle. *Phys. Rev.* D**2**, 2762–2773.
Seliger, R. L. & Whitham, G. B. 1968 Variational principles in continuum mechanics. *Proc. R. Soc. Lond.* A **305**, 1–25.
Synge, J. L. 1933 The stability of heterogeneous liquids. *Trans. R. Soc. Can.* **27**, 1–18.
Temam, R. 1983 *Navier–Stokes equations and nonlinear functional analysis*, CBMS–SIAM, 1983.
Thompson, R. E. & Stewart, R. W. 1977 The balance and redistribution of potential vorticity within the ocean. *Dynam. Atmos. Oceans* **1**, 299–321.
Turner, J. S. 1973 *Buoyancy effects in fluids.* Cambridge University Press.
Vivaldi, F. M. 1984 Weak instabilities in many dimensional Hamiltonian systems. *Rev. mod. Phys.* **56**, 737–754.
Wan, Y. H. & Pulvirenti, F. 1985 Nonlinear stability of circular vortex patches. *Commun math. Phys.* **99**, 435–450.
Weinstein, A. 1984 Stability of Poisson–Hamilton equilibria. *Contr. Math. AMS*, **28**, 3–13.
Yih, C. S. 1980 *Stratified flows.* New York: Academic Press.
Yuen, H. C. & Lake, B. M. 1980 Instability of waves in deep water. *A. Rev. Fluid Mech.* **12**, 303–340.

Richardson Number Criterion for the Nonlinear Stability of Three-Dimensional Stratified Flow

Henry D. I. Abarbanel
Scripps Institution of Oceanography and Project in Nonlinear Science, University of California, San Diego, La Jolla, California 92093

and

Darryl D. Holm
Theoretical Division and Center for Nonlinear Studies, Los Alamos National Laboratory, Los Alamos, New Mexico 87545

and

Jerrold E. Marsden
Department of Mathematics, University of California, Berkeley, California 94720

and

Tudor Ratiu
Department of Mathematics, University of Arizona, Tucson, Arizona 85721
(Received 15 March 1984)

With use of a method of Arnol'd, we derive the necessary and sufficient conditions for the formal stability of a parallel shear flow in a three-dimensional stratified fluid. When the local Richardson number defined with respect to density variations is everywhere greater than unity, the equilibrium is formally stable under nonlinear pertrubations. The essential physical content of the nonlinear stability result is that the total energy acts as a "potential well" for deformations of the fluid across constant density surfaces; this well is required to have definite curvature to assure stability under these deformations.

PACS numbers: 47.20.+m, 03.40.Gc, 92.10.Dh

With use of a method of Arnol'd[1] and others,[2,3] we have investigated the nonlinear stability of two- and three-dimensional incompressible flows of an inviscid stratified fluid treated as a Hamiltonian system. In this note, we report on the application of this technique to the important case of a shear flow with velocity profile $U(z)$, and density profile $\rho(z)$. We do not present the full set of conditions for nonlinear stability of this flow, but do exhibit the necessary and sufficient conditions for the formal stability of the flow. Formal stability means that a certain functional of the flow fields is definite in sign. Given formal stability, nonlinear stability requires additional convexity estimates to be satisfied. These do not alter the physical implications of the conditions derived here.[3]

The two-dimensional analysis[4] of the stratified fluid equations linearized about a planar shear flow $U(z)$, $\rho(z)$, shows that neutral stability (purely imaginary spectrum) occurs provided the Richardson number is everywhere greater than $\frac{1}{4}$. Here we derive the analogous criterion for formal stability for three-dimensional nonlinear deformations of the flow. Our criterion is that the local Richardson number defined with respect to variations across constant-density surfaces must be greater than 1. This focuses attention on the realm between $\frac{1}{4}$ and 1 for intensive theoretical and experimental investigation.

We treat stability in the Boussinesq approximation[4] for incompressible flow. See Ref. 2 for the treatment of nonlinear stability for compressible flows, and Ref. 3. for incompressible, stratified, non-Boussinesq flows. We address solutions of the momentum equation

$$\frac{\partial}{\partial t}\vec{u} + (\vec{u}\cdot\nabla)\vec{u} = -\nabla p - \rho g\hat{z}, \tag{1}$$

along with

$$\frac{\partial}{\partial t}\rho + \vec{u}\cdot\nabla\rho = 0 \quad \text{and} \quad \nabla\cdot\vec{u} = 0, \tag{2}$$

in a domain on whose boundary the normal component of the velocity $\vec{u}$ must vanish and the density ρ must be constant. In (1) and (2), p is the pressure and g is the constant gravitational acceleration in the $-\hat{z}$ direction. The constant reference density multiplying the acceleration in (1) has been set equal to unity.

Solutions to these equations conserve the energy

$$\int d^3x\,[\tfrac{1}{2}|\vec{u}|^2+\rho g z]. \quad (3)$$

Both ρ and the potential vorticity

$$q=(\nabla\times\vec{u})\cdot\nabla\rho \quad (4)$$

are conserved along fluid particle trajectories. Thus, for an arbitrary function $G(q,\rho)$,

$$A(\vec{u},\rho) = \int d^3x\,[\tfrac{1}{2}|\vec{u}|^2+\rho g z + G(q,\rho)+\lambda q] \quad (5)$$

is conserved. The term λq in (5) is separated to cancel some boundary terms which arise below. The role of the function $G(q,\rho)$ is that of a familiar Lagrange multiplier expressing the constraints on the flow imposed by conservation of q and ρ.

We now examine the first variation of $A(\vec{u},\rho)$ and relate its critical points to stationary solutions $\vec{u}_e$. The first variation is

$$\delta A(\vec{u}_e,\rho_e)=\int d^3x\,\{\delta\vec{u}\cdot[\vec{u}_e-G_{qq}\nabla\rho_e\times\nabla q_e]+\delta\rho[gz+G_\rho(\nabla\times\vec{u}_e)\cdot\nabla G_q]\} + (\lambda+G_q)|_s\int ds\hat{n}\cdot\{\delta\rho\,\nabla\times\vec{u}_e-\nabla\rho_e\times\nabla\vec{u}_e\}, \quad (6)$$

where $G_\rho=\partial G/\partial\rho$ evaluated at $q_e\rho_e$, etc., S is the boundary surface of the domain of the flow, and $\hat{n}$ is the outward unit normal vector on S.

δA in (6) vanishes at $\vec{u}_e$, ρ_e satisfying

$$\vec{u}_e=G_{qq}\nabla\rho_e\times\nabla q_e, \quad (7)$$

$$gz+G_\rho=(\nabla\times\vec{u}_e)\cdot\nabla G_q \quad (8)$$

in the interior, and

$$\lambda=-G_q \quad (9)$$

on the boundary. Flows satisfying (7) and (8) can be verified to be stationary solutions of (1) and (2). Expression (7) implies the requirements $\vec{u}_e\cdot\nabla\rho_e=\vec{u}_e\cdot\nabla q_e=0$ for stationary flows; (8) is the three-dimensional analog of Long's equation.[5]

We use (7) and (8) to determine $G(q_e,\rho_e)$ in terms of the Bernoulli function

$$K(q_e,\rho_e)=p_e+\rho_e g z+\tfrac{1}{2}|\vec{u}_e|^2, \quad (10)$$

via

$$G(q_e,\rho_e)=q_e\int^{q_e}\frac{dx}{x^2}K(x,\rho_e)+q_e\gamma(\rho_e), \quad (11)$$

where $\gamma(\rho_e)$ is an arbitrary function of ρ_e.

An equilibrium flow is said to be *formally stable* if the second variation of $A(\vec{u},\rho)$ at the critical point $\vec{u}_e,\rho_e$ is definite in sign. Formal stability implies[3] linearized stability since definiteness of δ^2A gives a preserved norm for the linearized solutions. As noted, nonlinear stability requires both formal stability and some convexity conditions on the function $G(q,\rho)$. For the present case, we find

$$\delta^2A(\vec{u}_e,\rho_e) = \int d^3x\left\{|\delta\vec{u}|^2+(\delta q,\delta\rho)\begin{bmatrix}G_{qq} & G_{q\rho}\\ G_{q\rho} & G_{\rho\rho}\end{bmatrix}\begin{pmatrix}\delta q\\ \delta\rho\end{pmatrix}\right\}. \quad (12)$$

From this we see that a *sufficient* condition for formal stability is that the eigenvalues of the two-by-two matrix in (12) are positive; namely,

$$G_{qq}>0, \quad (13)$$

and

$$G_{qq}G_{\rho\rho}-G_{q\rho}{}^2>0. \quad (14)$$

We can sharpen these sufficient conditions, however, by noting that $\mathrm{div}\nabla\cdot\delta\vec{u}=0$, so there are only two independent components of $\delta\vec{u}$, which along with $\delta\rho$ allow us to cast the definiteness of δ^2A into a linear three-by-three operator eigenvalue condition, whose eigenvalues must then be either all positive or all negative. This condition is made explicit in the example we now discuss.

Our example is the parallel equilibrium flow

$$\vec{u}_e(\vec{x})=(u(y,z),0,0), \quad (15)$$

$$\rho_e(\vec{x})=\rho(z). \quad (16)$$

This is a standard configuration and application of the Arnol'd method to it provides insight into the value of the technique. The validity of the linearized results on this flow have been examined in laboratory and geophysical situations. Our nonlinear result will thus provide impetus for further experimental study of these important flows. We separate the y and z dependences in $u(y,z)$ into a small, slowly varying y dependence plus a general z dependence $U(z)$. Thus, we write

$$u(y,z)=f(y)+U(z). \quad (17)$$

The role of $f(y)$ is to break the $q_e=0$ degeneracy of the two-dimensional $f=0$ flow, which is the conventional setup. The physical situation we wish to describe is a shear flow $U(z)$ with a smooth, small $f(y)$ imposed upon it to give the three-dimensionality needed for $q_e \neq 0$. We wish to parametrize $f(y)$ by a velocity scale, f_0, which is much less than $U(z)$, and by a length scale L which is large compared to any other lengths in the problem. We choose

$$f(y) = f_0(y/L)^2; \quad f_0 << U(z); \tag{18}$$

and restrict the domain of y to be $|y| << L$. In what follows, we expand all quantities in L^{-1}, capturing the essence of the stability problem in the leading orders of L which are retained for L very large.

From the Bernoulli function, (10), we find (dropping the subscript e henceforth)

$$G(q,\rho) = -[p + \rho g z + \tfrac{1}{2}U^2(z)] + \tfrac{1}{2}G_{qq}q^2 + O(q^4), \tag{19}$$

with

$$G_{qq} = \frac{u}{u_{yy}\rho_z^2} = \frac{L^2 U(z)}{f_0 \rho_z^2}\left[1 + \frac{f(y)}{U(z)}\right], \tag{20}$$

and we drop the last term commensurate with our assumptions on $f(y)$. G_{qq} is now a function of z (or ρ) alone. q in our flow is

$$q = (f_0/L)(y/L)(-\rho_z). \tag{21}$$

Since q is small for $|y| << L$, the neglect of higher-order terms in q, wherever they occur, is an excellent approximation.

Now we choose the two independent components of $\delta\vec{u}$ in (12) from the vertical velocity $v_3(\vec{x},t) = \delta\vec{u}\cdot\hat{z}$ and the vorticity $\omega_3(\vec{x},t) = (\nabla\times\delta\vec{u})\cdot\hat{z}$. This choice is motivated by the observation that the only essential dependence on the equilibrium flow is on the vertical coordinate z.[6] To leading order in L^{-1} a calculation shows that $\delta^2 A(\vec{u}_e, \rho_e)$ is given by

$$\delta^2 A(\vec{u}_e,\rho_e) = \int d^3x\,(v_3,\omega_3,\delta\rho)\begin{bmatrix} \nabla^2/\nabla_\perp^2 & 0 & 0 \\ 0 & -\dfrac{1}{\nabla_\perp^2} + \rho_z^2 G_{qq} & \rho_z U_z G_{qq}\partial_y \\ 0 & -\rho_z U_z G_{qq}\partial_y & G_{\rho\rho} - U_z^2 G_{qq}\partial_y^2 \end{bmatrix}\begin{bmatrix} v_3 \\ \omega_3 \\ p\delta_\rho \end{bmatrix}, \tag{22}$$

with $\nabla_\perp^2 = \partial_x^2 + \partial_y^2$ and $\nabla^2 = \nabla_\perp^2 + \partial z^2$. Precise meaning to $(\nabla_\perp^2)^{-1}$ is given by imposing periodic boundary conditions in x and y for each of v_3, ω_3, and $\delta\rho$. A term $f_y\partial_z\delta\rho$ has been neglected relative to $U_z\partial_y\delta\rho$, which is retained. This ordering means our choice of L must be large enough to overcome any very large vertical wave numbers in $\delta\rho$. The arbitrary function $\gamma(\rho_e)$ in (11) is set to zero.

For formal stability, we demand that $\delta^2 A$ be of definite sign for all independent variations in $(v_3,\omega_3,\delta\rho)$ space. That sign must be positive, as we see by looking in the direction $(v_3,0,0)$. Then by looking in the direction $(0,\omega_3,\delta\rho)$ we learn that the necessary and sufficient conditions for formal stability are that the two-by-two submatrix operator in (22) have only positive eigenvalues. This requirement is most easily expressed by Fourier transforming in x and y to wave numbers k_1 and k_2. The two-by-two submatrix becomes algebraic, and positivity of its eigenvalues occurs if and only if

$$1/k_\perp^2 + \rho_z^2 G_{qq} > 0, \tag{23}$$

and

$$G_{\rho\rho}[1 + k_\perp^2\rho_z^2 G_{qq}] + k_2^2 U_z^2 G_{qq} > 0, \tag{24}$$

with $k_\perp^2 = k_1^2 + k_2^2$.

Since we allow arbitrary variations of v_3, ω_3, and $\delta\rho$, each of k_1 and k_2 can be as large as we like. This means that we must have

$$\rho_z^2 G_{qq} = u/u_{yy} > 0, \tag{25}$$

and

$$G_{\rho\rho} > \max_{(k_1,k_2)}\left[\frac{k_2^2 U_z^2 G_{qq}}{1 + k_\perp^2\rho_z^2 G_{qq}}\right] = 0. \tag{26}$$

The first of these is the usual Rayleigh criterion for stability of shear flows in y. Its presence here is expected since we have no stratification in the horizontal direction. Condition (26) is the desired Richardson-number criterion. Note that

$$G_{\rho\rho} = -g\,\partial z/\partial\rho - \partial^2[\tfrac{1}{2}U^2(z)]/\partial\rho^2. \tag{27}$$

When $(U^2)_{\rho\rho}$ is positive, we may define the generalization of the usual Richardson number to be

$$N_{\mathrm{Ri}}(z) = N(z)^2/\{\rho_z^2\partial^2[\tfrac{1}{2}U^2(z)]/\partial\rho^2\}, \tag{28}$$

with $N^2(z) = -g\,\partial\rho/\partial z$ the Brünt-Väisäla frequency in Boussinesq approximation. [N_{Ri} defined by (28) agrees locally with the standard gradient definition, if one uses the linearization of U and ρ (e.g., Ref. 3)]. The necessary and sufficient condition for formal stability then becomes

$$N_{\mathrm{Ri}}(z) > 1 \tag{29}$$

everwhere in the flow. This is our central result.

In addition, there are situations where ρ_z positive (a statically unstable configuration) may be stabilized by the shear flow. To exhibit this stabilization, we assume $\rho_z \neq 0$ and define the "inverse Richardson number"

$$a(z) = \{\partial^2[\tfrac{1}{2}U^2(z)]/\partial\rho^2\}(-\rho z/g). \tag{30}$$

When $\rho_z < 0$, that is for statically stable stratification, all flows with $a(z) < 1$ are formally stable. When $\rho_z > 0$, that is for *statically unstable stratification,* all flows with $a(z) > 1$ are formally stable. The first case is usually understood by saying that the kinetic energy acquired by a parcel of fluid crossing density surfaces is not sufficient to overcome the potential energy required to move the parcel. The second case is less familiar and is only possible if second derivatives of U are relatively large. In this case, the potential energy that would be gained by a fluid parcel in crossing density surfaces is not sufficient to overcome kinetic energy lost in the same traverse.

The essence of our argument in this note is that the negative of the Bernoulli function (10) acts as a "potential well" for stratified flow. This is seen in (19) where G is, for this heuristic discussion, $-(p+\rho gz+\frac{1}{2}|\vec{u}|^2)$. Our requirement that $G_{\rho\rho} > 0$ tells us that this potential well has positive curvature for crossing density surfaces, when the flow is formally stable. This note provides detailed demonstration of this notion, which itself was discussed as long ago as 1931 by Prandtl.[7]

One of us (H.D.I.A.) would like to thank J. W. Miles and W. H. Munk for illuminating discussions on the physical content of this work. This work was supported in part by the U. S. Department of Energy under Contracts No. W-7405-ENG-36 and No. AT-3-82ER12097, and by the Office of Basic Energy Sciences, Division of Applied Mathematical Sciences, and by the Office of Naval Research, Code 422PO. One of us (T.R.) was a National Science Foundation postdoctoral fellow while at the University of California, Berkeley.

[1]V. I. Arnol'd, Dokl. Nat. Nauk. **162**, 773 (1965), and Am. Math. Soc. Transl. **79**, 267 (1969).

[2]D. D. Holm, J. E. Marsden, T. Ratiu, and A. Weinstein, Phys. Lett. **98A**, 15 (1983), and to be published.

[3]H. D. I. Abarbanel, D. D. Holm, J. E. Marsden, and T. Ratiu, "Nonlinear Stability Analysis of Ideal Stratified, Incompressible Fluid Flow" (to be published).

[4]P. G. Drazin and W. H. Reid, *Hydrodynamic Stability* (Cambridge Univ. Press, Cambridge, England, 1981). Especially important for us is the work of Miles and Howard reported in Sect. 44.

[5]R. R. Long, Tellus **5**, 42 (1953). See also the monograph by C. S. Yih, *Stratified Flows* (Academic, New York, 1980).

[6]See, for example, the discussion of thermal convection in S. Chandrasekhar, *Hydrodynamic and Hydromagnetic Stability* (Cambridge Univ. Press, Cambridge, England, 1961).

[7]L. Prandtl, *Führer durch die Strömungslehre* (Vieweg, Braunschweig, 1931), Sec. V, 12(d).