

On Formal Stability of Stratified Shear Flows

by

Hirofumi SAKUMA and Yasuhide FUKUMOTO

Abstract

A novel linear stability criterion is established for the equilibria of general three-dimensional (3D) rotating flows of an ideal gas satisfying Boyle–Charles’ law by a newly refined energy–Casimir convexity (ECC) method that can exploit a larger class of Casimir invariants. As the conventional ECC method cannot be applied directly to stratified shear flows, in our new approach, rather than checking the local convexity of a Lyapunov functional $L \equiv E + C_E$ defined as a sum of the total energy and a certain Casimir, we seek the condition for nonexistence of unstable manifolds: orbits (physically realisable flow in phase space) on the leaves of invariants including L as well as other Casimirs connecting a given equilibrium point O and other points in the neighbourhood of it. We argue that the separatrices of the second variation of L ($\delta^2 L = 0$) generally consist of such unstable manifolds as well as pseudo unstable ones for which either the total energy or Casimirs actually serve as a barrier for escaping orbits. The significance of the new method lies in the fact that it eliminates the latter so as to derive a condition for O being an isolated equilibrium point in terms of orbital connections.

2010 Mathematics Subject Classification: Primary 76E20; Secondary 37B25.

Keywords: ECC method, stratified shear flows, 3D steady states.

§1. Introduction

The stability of a given steady state of a perfect fluid motion possessing Hamiltonian (H) structure is an intensively studied problem in fluid mechanics. Actually, the analyses of stability so far reported are innumerable; various methods are employed. From a purely mathematical point of view, the so-called Lyapunov direct (or second) method is a conceptually simple and elegant way to deal with the sta-

Communicated by M. Yamada. Received April 2, 2014. Revised February 25, 2015 and May 19, 2015.

H. Sakuma: Yokohama Institute for Earth Sciences, JAMSTEC,
Yokohama, Kanagawa 236-0001, Japan;
e-mail: sakuma@jamstec.go.jp

Y. Fukumoto: Institute of Mathematics for Industry, Kyushu University,
744 Motooka, Nishi-ku, Fukuoka 819-0395, Japan;
e-mail: yasuhide@imi.kyushu-u.ac.jp

bility of a given dynamical system. A central problem associated with this method is the construction of a Lyapunov function (or functional for infinite-dimensional systems) which behaves as a “generalised energy barrier” preventing orbits Γ_i ($i = 1, 2, \dots$) from leaving a neighbourhood of a given equilibrium point O . For inviscid fluid motions, we consider below how the invariants associated with the H structure of the system contribute to the construction of a Lyapunov functional.

Although the modern description of the H structure in its general form is given in abstract Lie-group-theory terms, a particular application to hydrodynamical systems needs some physical expertise. Hence, we recall these terms for those not familiar with the field. For simplicity, we start from a dynamical system of finite dimension. Consider the dynamics of fields u^i defined on a Banach space P (Poisson manifold) admitting the Poisson bracket operation $\{ \ , \ }$ that determines its Lie algebra:

$$(1.1) \quad \dot{u}^i = F^i(u^1, \dots, u^N), \quad i = 1, \dots, N,$$

where \dot{u}^i and F^i denote respectively, the time derivative of u^i and (nonlinear) operators mapping a domain in P to P . Additionally, we assume that there exists a conserved function $H(u^i)$ called the Hamiltonian, usually the total energy of a given system, satisfying $\dot{H}(u^i) = 0$ for any C^1 solution of (1.1). We say that the given system possesses an *H structure* if the equations of motion (1.1) can be written in the form

$$(1.2) \quad \dot{u}^i = \{u^i, H\}.$$

The formal properties of the Poisson bracket are the skew-symmetry,

$$(1.3) \quad \{f, g\} = -\{g, f\},$$

the Jacobi identity

$$(1.4) \quad \{f, \{g, h\}\} + \{g, \{h, f\}\} + \{h, \{f, g\}\} = 0,$$

and the chain rule

$$(1.5) \quad \{f, g(u^1, \dots, u^N)\} = \sum_{j=1}^N \{f, u^j\} \frac{\partial g}{\partial u^j},$$

with similar formulae holding if f , or both f and g , depend on a collection of other functions. The above properties of the Poisson bracket are the classical analogs of the well-known relations for commutators in quantum mechanics.

Historically, the H structure was intensively studied first for systems consisting of particles where canonical variables (or coordinates) associated with generalised positions (q 's) and canonically conjugate momenta (p 's) are employed as independent coordinates, by means of which any function u^i , including H on P , is expressed. Notice that the total number of variables (q 's and p 's) in this case is even. From a physical point of view, however, the use of canonical variables for quite a few dynamical systems is not advantageous because some of them are either unphysical or seemingly redundant. The widely used 3D Eulerian representation of hydrodynamics that we shall consider gives a typical example of such noncanonical variables, in which a complete set of physical variables describing the motion of a perfect fluid consists of the 3D velocity vector \vec{v} and a couple of thermodynamical variables, say, density ρ and specific entropy η in a given physical domain. These variables are regarded as functions of time t and spatial (Cartesian) coordinates x , y and z , as for example for geophysical fluids relative to Earth. Recall that hydrodynamics admits an alternative Lagrangian representation where, instead of x , y and z , the labels (or coordinates) a^i ($i = 1, 2, 3$) of a given infinitesimally small "fluid particle" are chosen as the unique comoving (i.e., $\dot{a}^i = 0$) coordinates of this representation. Using the modern mathematical terminology of differential geometry, the system of 3D Lagrangian representations is right invariant on the cotangent bundle of the group of mass-preserving diffeomorphisms, which can be described by a canonical formulation of the H structure having eight independent variables. Note that the number of independent variables in the Eulerian representation is five, which is smaller than that for the canonical Lagrangian representation.

The existence of such multiple dynamical representations for the same system suggests that a conventional symplectic structure generated by canonical variables (q 's and p 's) is not an essential ingredient describing the H structure. Indeed, Morrison and Green [18], [19] (cf. also Marsden [16]) made an important contribution in formulating the noncanonical H structure for various fluid systems by deriving the relation between the existence of Casimir function(al)s and the singularity of the Poisson bracket. The point of their argument can be concisely illustrated using a finite-dimensional system (1.2). To do so, we present here a lucid explanation of Littlejohn [14] starting with an introduction of a second rank tensor J^{ij} called the Poisson tensor given by

$$(1.6) \quad J^{ij}(u) \equiv \{u^i, u^j\}.$$

Then, for any $f(u^1, \dots, u^N)$ and $g(u^1, \dots, u^N)$, using (1.5), one finds

$$(1.7) \quad \{f, g\} = \sum_{i,j} \frac{\partial f}{\partial u^i} J^{ij} \frac{\partial g}{\partial u^j},$$

so that we readily see that (1.2) assumes the form

$$(1.8) \quad \dot{u}^i = \sum_{j=1}^N J^{ij} \frac{\partial H}{\partial u^j}.$$

If u^i are canonical coordinates, i.e., $(u^1, \dots, u^N) = (q^1, \dots, q^L, p^1, \dots, p^L)$, with $N = 2L$, then J^{ij} can be represented by its partition into four $L \times L$ matrices:

$$(1.9) \quad J^{ij} = \begin{pmatrix} 0 & I \\ -I & 0 \end{pmatrix},$$

with $\det(J^{ij}) = 1$. We see that $\dot{u}^i = 0$ corresponds to $\partial H / \partial u^i = 0$. For J^{ij} singular, let R and K denote respectively the rank and corank of J^{ij} with $N = R + K$. Then there exist K linearly independent covariant vectors $\xi_i^{(k)}$ satisfying

$$(1.10) \quad \sum_{j=1}^N J^{ij} \xi_j^{(k)} = 0, \quad k = 1, \dots, K.$$

In particular, if there exist $C^{(k)}$ such that $\xi_i^{(k)} = \partial C^{(k)} / \partial u^i$, then directly from (1.7), for any f , we obtain

$$(1.11) \quad \{f, C^{(k)}\} = 0, \quad k = 1, \dots, K.$$

We call $C^{(k)}$ a *Casimir function* because it commutes with any function. An important consequence of this result is that the Hamiltonian H is not unique. Indeed, we readily see that the equations of motion remain unaltered after the replacement of H by $H + \sum \lambda_{(k)} C^{(k)}$ where $\lambda_{(k)}$ are some constants:

$$(1.12) \quad \left\{ u^i, H + \sum \lambda_{(k)} C^{(k)} \right\} = \dot{u}^i + \sum \lambda_{(k)} \{u^i, C^{(k)}\} = \dot{u}^i,$$

although the extremal points of the Hamiltonian may change under this transformation. Note also that any Casimir function C is necessarily a constant of motion because

$$(1.13) \quad \dot{C} = \{C, H\} = 0,$$

and hence an orbit on P representing some physical motion is confined to any iso-Casimir leaf. The above is an outline of how Casimir invariants are related to the singularity of the Poisson bracket. One can find a formal theorem on the existence of Casimir functions, for instance, in the book by Sudarshan and Mukunda [23].

Because fluid dynamical systems are infinite-dimensional, the Hamiltonian formulation (1.2) for finite-dimensional systems must be revised accordingly. Gen-

eral aspects of the modification, such as the replacement of derivatives by their functional counterparts, are given in standard textbooks, such as Olver [20], on the application of Lie group theory. However, the determination of the associated Poisson brackets, especially noncanonical ones, is not straightforward. It was not until the seminal work of Morrison and Green that we eventually have them for fluid dynamical systems. As the number of independent variables in the Lagrangian representation with canonical variables is larger than in the noncanonical Eulerian representation, a similar singular structure for the Poisson bracket, briefly illustrated in (1.10) and (1.11) for finite-dimensional systems, arises as a result of the reduction of variables from Lagrangian to Eulerian descriptions (Holm et al. [9]); the Casimir functionals associated with it are to be determined by the condition corresponding to (1.11).

Returning to the main issue of stability analysis, the ECC method as an extension of Arnol'd's will be briefly reviewed in the following. A prototype of the ECC method was proposed first by Arnol'd [3], [4] for 2D incompressible flows and, in the 1980s, the method was systematically developed by a group of mathematical physicists, notably by Holm et al. [10] and Abarbanel et al. [1] for various 2D and 3D fluid and plasma systems. The usefulness of the ECC method for hydrodynamics lies in the following. First, the stability of a given equilibrium point (a steady state) is studied (in almost all practical cases) using an Eulerian reference frame, and second, the given steady state of interest is nontrivial, that is, not quiescent. Recall from (1.12) that by adding $\sum \lambda_{(k)} C^k$ to H , the extremal points can change. Actually, it can be shown that we can find a certain Casimir functional C_E such that a newly defined generalised Hamiltonian H ($H \equiv E + C_E$) has an extremal point for a given (arbitrary) steady state, where E denotes the total energy in the usual sense. After introducing $E + C_E$ as a Lyapunov functional, Arnol'd derived a condition for its convexity which limits the departure of finite perturbations from a given equilibrium point, and thus he established nonlinear stability.

A preliminary analysis for obtaining the local convexity of a given Lyapunov functional $L = E + C_E$ in an ϵ -neighbourhood of O is called *formal stability* (FS) *analysis* ([10]), where the positive (or negative) definiteness of the second variation of L is checked; the FS of various dynamical systems including fluids and plasma has been studied by numerous researchers. For finite dimensions, FS implies (nonlinear) stability, but it is not generally so for infinite-dimensional cases. Using examples from elasticity theory, Ball and Marsden [5] showed that an equilibrium solution having a positive definite second variation of energy permits an infinite number of unstable directions. Such a situation occurs in the infinite-dimensional

case because the Hessian associated with $\delta^2 L$ may have eigenvalues asymptotically approaching 0. Hence, for FS in infinite-dimensional systems, we must show that $\delta^2 L$ is coercive, as for $\delta^2 L$ studied by Arnol'd [3]. In any case, FS is considered to be a step towards nonlinear stability, as it deals with perturbations with infinitesimal amplitudes.

The algorithm of the above EEC method for nonlinear stability analyses of fluids is mathematically elegant from the physical point of view; however, we can safely say that, except for symmetric stability, it provides either too strong (not sharp in physical sense) criteria for relatively simple 2D flows, or no information at all for 2D stratified and 3D general shear flows. Indeed, we know that even FS cannot be shown for the latter flows. The failure to prove FS, however, does not necessarily mean that the steady state under consideration is always unstable. Actually, for 2D stratified shear flows, we know that sufficiently strong stable stratification makes them stable. The failure in proving FS means that there exist separatrices (nodal directions) of $\delta^2 L$ along which we have $\delta^2 L|_{(s)} = \delta^2 E|_{(s)} + \delta^2 C_E|_{(s)} = 0$. Nevertheless, this does not necessarily mean that $\delta E|_{(s)} = \delta C_E|_{(s)} = 0$ although we have $\delta E|_{(s)} + \delta C_E|_{(s)} = 0$ by the choice of C_E . If $\delta E|_{(s)} = -\delta C_E|_{(s)} \neq 0$, then such a separatrix is not related to the direction of an unstable manifold (an orbit Γ_j^o in P starting from the equilibrium point O), because nonzero $\delta E|_{(s)}$ acts as an “energy barrier” for that direction. The crucial difference between unstable manifolds and such virtual separatrices is that the former are on the leaves in P defined as the constant level sets of all the invariants, namely, the total energy as well as the Casimirs associated with the reduction of variables. Hence, it is expected that certain Casimirs C_R other than C_E are used to distinguish them. Successful elimination of such virtual separatrices from the original $\delta^2 L$ may result in a new (coercive) norm which can provide a refined formal stability criterion for general rotating 3D baroclinic flows. The goal of this paper is to show that this is actually true.

To our knowledge, no notable advances in FS analyses have been made during the past decades. It appears that only a minor extension has been achieved in the 3D stratified shear instability problem by Mahalov et al. [15], although the original 2D version for spectral (or normal mode) instability analysis for incompressible fluid is well-documented in terms of the Richardson-number criterion with the threshold value of 1/4 [8], [17], [12]. The previous FS analyses suffer from a limitation in evaluating the second-order terms arising from C_E employed in a conventional approach; this will be briefly recapitulated in the following section.

The organization of this paper is as follows: As a necessary preliminary step for the new analysis, we focus in Section 2 on concrete examples, the governing

equations of a given hydrodynamical system as well as key expressions to be used in the ECC analyses. The essential idea of the new approach will be explained in Section 3; the main result and consequent remarks follow in Sections 4 and 5, respectively.

§2. Theoretical settings for a 3D perfect fluid system

§2.1. Governing system of equations and its H structure

Let us consider an Eulerian form of the governing system of equations defined on a bounded physical domain $\Omega = S^1 \times \mathbb{R}^2 \subset \mathbb{R}^3$ equipped with Cartesian coordinates (x, y, z) . For simplicity, in this subsection, we assume that the domain provides an inertial reference frame, and the orientation of (x, y, z) is such that x corresponds to a cyclic direction, while y and z are bounded by rigid boundaries respectively at $y = y_1, y_2$ and $z = z_1, z_2$ where the corresponding components of the velocity vector vanish. As we already mentioned, in the Eulerian representation, the complete set of dependent field variables is composed of the 3D velocity field $\vec{v} \equiv (u, v, w)$ and the independent thermodynamical variables: density ρ and either pressure p or specific entropy η . (Thermodynamics tells us that any thermodynamical variable can be expressed as a certain function of a couple of arbitrarily selected variables.) The governing system of equations is

$$(2.1) \quad \frac{\partial \vec{v}}{\partial t} = -(\vec{v} \cdot \vec{\nabla})\vec{v} - \frac{\vec{\nabla} p}{\rho} = -\vec{\nabla} \left(\frac{\vec{v} \cdot \vec{v}}{2} + h(\rho, \eta) \right) - (\vec{\nabla} \times \vec{v}) \times \vec{v} + T(\rho, \eta) \vec{\nabla} \eta,$$

$$(2.2) \quad \frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot (\rho \vec{v}),$$

$$(2.3) \quad \frac{\partial \eta}{\partial t} = -\vec{v} \cdot \vec{\nabla} \eta.$$

(2.1) gives the equation of motion with different (but equivalent) vector expressions for $-(\vec{v} \cdot \vec{\nabla})\vec{v}$ and for pressure gradient force $-\vec{\nabla} p / \rho$ in terms of specific enthalpy h , absolute temperature T and η , whereas (2.2) and (2.3) are conservation equations for the mass and entropy. The Hamiltonian functional as the total energy E is given by

$$(2.4) \quad H[\vec{v}, \rho, \eta] = E[\vec{v}, \rho, \eta] = \left\langle \frac{1}{2} \rho \vec{v} \cdot \vec{v} + \rho e(\rho, \eta) \right\rangle,$$

where $e(\rho, \eta)$ is the internal energy density per unit mass, the symbol $\langle \bullet \rangle$ denotes the volume integral over the entire domain Ω , and the noncanonical Poisson bracket of Morrison and Greene [18], [19] is

$$(2.5) \quad \{F, G\} \equiv - \left\langle \left(\frac{\delta F}{\delta \rho} \vec{\nabla} \cdot \frac{\delta G}{\delta \vec{v}} - \frac{\delta G}{\delta \rho} \vec{\nabla} \cdot \frac{\delta F}{\delta \vec{v}} \right) + \left(\frac{\vec{\nabla} \times \vec{v}}{\rho} \cdot \frac{\delta G}{\delta \vec{v}} \times \frac{\delta F}{\delta \vec{v}} \right) + \frac{\vec{\nabla} \eta}{\rho} \cdot \left(\frac{\delta F}{\delta \eta} \frac{\delta G}{\delta \vec{v}} - \frac{\delta G}{\delta \eta} \frac{\delta F}{\delta \vec{v}} \right) \right\rangle,$$

where the notation $\delta\phi/\delta\vec{v} \equiv (\delta\phi/\delta u, \delta\phi/\delta v, \delta\phi/\delta w)$ is employed. From (2.4), we get

$$(2.6) \quad \frac{\delta H}{\delta \vec{v}} = \rho \vec{v}, \quad \frac{\delta H}{\delta \rho} = \frac{1}{2} \vec{v} \cdot \vec{v} + \frac{\partial}{\partial \rho} [\rho e(\rho, \eta)] = \frac{1}{2} \vec{v} \cdot \vec{v} + h(\rho, \eta), \quad \frac{\delta H}{\delta \eta} = \rho \frac{\partial e}{\partial \eta}.$$

Using this, we find that (2.1)–(2.3) are equivalent to

$$(2.7) \quad \frac{\partial \vec{v}}{\partial t} = \{\vec{v}, H\}, \quad \frac{\partial \rho}{\partial t} = \{\rho, H\}, \quad \frac{\partial \eta}{\partial t} = \{\eta, H\}.$$

The form of the Poisson bracket depends upon the choice of variational variables. For instance, if we use $(\vec{M} \equiv \rho \vec{v}, \rho, s \equiv \rho \eta)$ instead of (\vec{v}, ρ, η) , then (2.4) and (2.5) can be rewritten respectively as

$$(2.8) \quad H[\vec{M}, \rho, s] = \left\langle \frac{1}{2\rho} \vec{M} \cdot \vec{M} + \rho e\left(\rho, \frac{s}{\rho}\right) \right\rangle,$$

$$(2.9) \quad \{F, G\} \equiv - \left\langle M_i \left(\frac{\delta F}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta G}{\delta M_i} - \frac{\delta G}{\delta M_j} \frac{\partial}{\partial x^j} \frac{\delta F}{\delta M_i} \right) + \rho \left(\frac{\delta F}{\delta \vec{M}} \cdot \vec{\nabla} \frac{\delta G}{\delta \rho} - \frac{\delta G}{\delta \vec{M}} \cdot \vec{\nabla} \frac{\delta F}{\delta \rho} \right) + s \left(\frac{\delta F}{\delta \vec{M}} \cdot \vec{\nabla} \frac{\delta G}{\delta s} - \frac{\delta G}{\delta \vec{M}} \cdot \vec{\nabla} \frac{\delta F}{\delta s} \right) \right\rangle.$$

The Poisson bracket of (2.9), called the Lie–Poisson bracket, is characterised by the linearity in the variational variables given as the coefficients of the skew-symmetric operators in the above parentheses. The governing system of equations corresponding to (2.7) in this case turns out to be

$$(2.10) \quad \frac{\partial \vec{M}}{\partial t} = \{\vec{M}, H\}, \quad \frac{\partial \rho}{\partial t} = \{\rho, H\}, \quad \frac{\partial s}{\partial t} = \{s, H\},$$

which are alternative expressions of (2.1), (2.2) and (2.3); specifically,

$$(2.11) \quad \frac{\partial \rho v_i}{\partial t} = -\nabla_j (\rho v^j v_i + \delta_i^j p), \quad \frac{\partial \rho}{\partial t} = -\vec{\nabla} \cdot (\rho \vec{v}), \quad \frac{\partial s}{\partial t} = -\vec{\nabla} \cdot (s \vec{v}),$$

where δ_i^j is the Kronecker delta. From (2.5) we can see that the condition $\{F, C\} = 0$ for all F implies

$$(2.12) \quad \vec{\nabla} \cdot \frac{\delta C}{\delta \vec{v}} = 0, \quad \frac{\vec{\nabla} \eta}{\rho} \cdot \frac{\delta C}{\delta \vec{v}} = 0, \quad \vec{\nabla} \frac{\delta C}{\delta \rho} + \frac{\vec{\nabla} \times \vec{v}}{\rho} \times \frac{\delta C}{\delta \vec{v}} - \frac{\vec{\nabla} \eta}{\rho} \frac{\delta C}{\delta \eta} = 0.$$

It can be shown that $C_1[\rho, \eta] \equiv \langle \rho F(\eta) \rangle$ where F is an arbitrary function and the helicity $C_h[\vec{v}] \equiv \langle \vec{v} \cdot \vec{\nabla} \times \vec{v} \rangle$ with zero potential vorticity $\vec{\nabla} \eta \cdot \vec{\nabla} \times \vec{v} = 0$ are

solutions to (2.12). In addition to these, we also have Casimir invariants related to the nonzero potential vorticity conservation; the general form of the Casimir functional(s) including vorticity, given, for example, in Kuroda [13] and Salmon [22], will be introduced in the next section.

§2.2. Equations for ideal gas in a rotating system under the influence of gravity

The governing system of equations for an inertial reference frame discussed in the previous subsection can be readily extended to a rotating reference frame under the influence of gravity. The equations with the β -plane approximation traditionally employed in geophysical fluid dynamics (GFD) are one such extension where the cyclic longitudinal, bounded latitudinal and vertical coordinates are respectively denoted by Cartesian x , y and z . In a newly defined bounded domain Ω , the 3D Eulerian form of the momentum equation corresponding to (2.1) becomes

$$(2.13) \quad \frac{\partial \vec{v}}{\partial t} = -\vec{\nabla} B - \vec{\omega}_a \times \vec{v} + T \vec{\nabla} \eta,$$

where B is the Bernoulli function $B \equiv C_P T + \vec{v} \cdot \vec{v}/2 + gz$ defined as a trinomial expression of the specific enthalpy of an ideal gas (specific heat at constant pressure C_P multiplied by T), and the kinetic and gravitational potential energy densities; $\vec{\omega}_a$ denotes the absolute vorticity vector including the vertical component of Earth rotation expressed as a linearized Coriolis parameter f of the form $f \vec{k} = (f_0 + \beta y) \vec{k}$ with \vec{k} a vertical unit vector, i.e., $\vec{\omega}_a \equiv \vec{\nabla} \times \vec{v} + f \vec{k}$. The traditional approximation in this model system assumes the horizontal component for Earth's rotation is dropped.

A proof of the existence of solutions to (2.13) in a mathematically rigorous sense is beyond the scope of this paper, and hence in what follows we assume a priori that, for our present analyses, sufficiently smooth solutions exist together with the conservation of total energy and Casimirs arising from the inhomogeneous distribution of η and a vortical quantity q called Ertel's potential vorticity defined below. Under these assumptions, we can claim that η and q given by

$$(2.14) \quad q \equiv \frac{1}{\rho} (\vec{\omega}_a \cdot \vec{\nabla} \eta)$$

are two independent advective quantities along a fluid particle's trajectory in Ω :

$$(2.15) \quad \frac{D\eta}{Dt} \equiv \frac{\partial \eta}{\partial t} + \vec{v} \cdot \vec{\nabla} \eta = 0, \quad \frac{Dq}{Dt} \equiv \frac{\partial q}{\partial t} + \vec{v} \cdot \vec{\nabla} q = 0,$$

where $D(\bullet)/Dt$ denotes the Lagrangian time derivative. For 2D incompressible fluid motions analysed by Arnol'd, the vorticity $\omega = \partial v/\partial x - \partial u/\partial y$ alone behaves

like an advective scalar quantity, whereas in 3D problems, vorticity becomes a vector quantity $\vec{\omega}$ which is no longer advective for baroclinic (or nonisentropic, meaning that η varies from point to point) flows under consideration. That is, in 3D baroclinic problems, a vorticity-related advection quantity is given by the pseudo-scalar in (2.14).

For a steady state of (2.13), we have

$$(2.16) \quad \vec{\nabla} B + \vec{\omega}_a \times \vec{v} = T \vec{\nabla} \eta,$$

and using (2.15), we readily get

$$(2.17) \quad \vec{v} \cdot \vec{\nabla} B = 0 \Rightarrow B = B(\eta, q).$$

Multiplying (2.16) respectively by $\times \vec{\nabla} \eta$ and $\cdot \vec{\omega}_a$ after expanding B in terms of η and q , we obtain

$$(2.18) \quad \rho q \vec{v} = B_q \vec{\nabla} \eta \times \vec{\nabla} q, \quad B_q \vec{\nabla} q \cdot \vec{\omega}_a = (T - B_\eta) \rho q,$$

where the subscripts η and q on B denote the respective variable of the partial derivative. The energy and a general form of the Casimir invariants for (2.13) are given by

$$(2.19) \quad E \equiv \left\langle \frac{\rho}{2} \vec{v} \cdot \vec{v} + \rho e(\rho, \eta) + \rho g z \right\rangle, \quad C_F \equiv \langle \rho F(\eta, q) \rangle,$$

where $F(\eta, q)$ is an arbitrary function of its arguments (refer to Kuroda [13] for C_F). As we have mentioned in the introduction, the leaves of constant total energy and Casimir(s) of a given H system prevent orbits Γ_i in the phase space P from crossing over to other leaves of total energy and other Casimir invariants.

A nontrivial equilibrium point O in the fluid dynamical phase space P of our concern corresponds to a steady state of nonvanishing fluid motion given in (2.16) or the alternative expression (2.18), for which the first variations of the total energy E and other invariants evaluated at O do not vanish. In general, their specific superposition does (cf. the statement next to (1.12)). One of the key algebraic elements in the ECC method is to find whether, for a given total energy E of an arbitrary steady state, there exists an associated Casimir invariant C_E for which the first variation evaluated at O exactly cancels out that of E ([3], [24]), specifically,

$$(2.20) \quad \delta(E + C_E) = 0,$$

where the infinitesimal variation is calculated with respect to independent variational variables of $\vec{v} \rightarrow \vec{v} + \delta\vec{v}$, $\rho \rightarrow \rho + \delta\rho$ and $\eta \rightarrow \eta + \delta\eta$. Note that the

arbitrariness of $F(\eta, q)$ in the second identity of (2.19) can be exploited to impose the above condition. After deriving (2.20), we proceed to the evaluation of the second variation of $L = E + C_E$. If we can show the sign-definiteness of $\delta^2 L$, that is,

$$(2.21) \quad \delta^2(E + C_E) > 0 \quad \text{or} \quad \delta^2(E + C_E) < 0,$$

then the given equilibrium point O is said to be *formally stable*. Thus FS is equivalent to a linear Lyapunov stability of the system. However, there are cases in which a straightforward application of this procedure does not work. The stratified shear flow problem involving the potential vorticity q is one exception. The quadratic term $\delta\vec{\omega} \cdot \vec{\nabla}\delta\eta$ to appear in the second variation arising from (2.14) cannot, by any means, be transformed into a tractable form from which the sign-definiteness of the second variation of L can be derived. Indeed, as will be shown below, the second variation of E is readily expressible using quadratic terms with $\delta\vec{v}$, $\delta\rho$ and $\delta\eta$ being the independent variational variables, whereas that of any Casimir inevitably involves $\delta\vec{\omega} \cdot \vec{\nabla}\delta\eta$ which cannot be transformed into a tractable quadratic form similar to the ones derived for $\delta^2 E$. It seems inevitable that this notorious obstacle always appears as long as we consider \vec{v} , ρ and η as the independent variational variables in our FS analysis.

As a preparation for the main analyses in the subsequent section, we give below the final forms of $\delta(E + C_F)$ and a partially completed form of $\delta^2(E + C_F)$. The derivations are lengthy but straightforward; the details are given in Appendix A. In their derivation, we note that the first and second variations of q with respect to the independent variational variables are

$$(2.22) \quad \begin{aligned} \rho\Delta^{(2)}q &\equiv \rho(\delta q + \delta^2 q), & \rho\delta q &\equiv \vec{\omega}_a \cdot \vec{\nabla}\delta\eta + \vec{\nabla}\eta \cdot \delta\vec{\omega} - q\delta\rho, \\ & & \rho\delta^2 q &\equiv -\delta q\delta\rho + \delta\vec{\omega} \cdot \vec{\nabla}\delta\eta. \end{aligned}$$

With this expression for δq , we obtain

$$(2.23) \quad \begin{aligned} \delta(E + C_F) &= \langle (F - qF_q + B)\delta\rho + (\rho\vec{v} - \vec{\nabla}\eta \times \vec{\nabla}F_q) \cdot \delta\vec{v} \\ &\quad + (\rho T + \rho F_\eta - \vec{\omega}_a \cdot \vec{\nabla}F_q)\delta\eta \rangle + \langle \vec{\nabla} \cdot [F_q(\vec{\omega}_a\delta\eta + \delta\vec{v} \times \vec{\nabla}\eta)] \rangle, \end{aligned}$$

$$(2.24) \quad \begin{aligned} \delta^2(E + C_F) &= \left\langle \frac{\rho}{2}\delta\vec{v} \cdot \delta\vec{v} + \vec{v} \cdot \delta\vec{v}\delta\rho + \frac{(c_s)^2}{2\rho}(\delta\rho)^2 + \left(\frac{C_{PT}}{C_V} + F_\eta\right)\delta\rho\delta\eta \right. \\ &\quad \left. + \frac{1}{2}\left(\frac{\rho T}{C_V} + \rho F_{\eta\eta}\right)(\delta\eta)^2 + \frac{\rho}{2}F_{qq}(\delta q)^2 + [\rho F_{\eta q}\delta\eta\delta q + F_q\delta\vec{\omega} \cdot \vec{\nabla}\delta\eta] \right\rangle, \end{aligned}$$

where the subscripts on F , as in F_q and F_η , denote partial derivatives, whereas C_V and $c_s \equiv \sqrt{C_P RT/C_V}$ are respectively the specific heat at constant volume and the propagation speed of the acoustic wave. After integrating by parts, we

further rewrite the last two terms in the above square bracket as

$$(2.25) \quad \rho F_{\eta q} \delta \eta \delta q + F_q \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta = \vec{\nabla} \cdot [F_q \delta \eta \delta \vec{\omega} + \frac{1}{2} F_{\eta q} \vec{\omega}_a (\delta \eta)^2] - q F_{\eta q} \delta \eta \delta \rho \\ - \frac{1}{2} [\vec{\omega}_a \cdot \vec{\nabla} F_{\eta q}] (\delta \eta)^2 - F_{qq} \delta \eta \vec{\nabla} q \cdot \delta \vec{\omega},$$

which, as will be shown shortly, makes the form of $\delta^2(E + C_E)$ tractable in the sense that we can express it in terms of the Bernoulli function and its derivatives after identifying C_E by (2.20). Substituting this into (2.24), we get

$$(2.26) \quad \delta^2(E + C_F) = \left\langle \frac{\rho}{2} \delta \vec{v} \cdot \delta \vec{v} + \vec{v} \cdot \delta \vec{v} \delta \rho + \frac{(c_s)^2}{2\rho} (\delta \rho)^2 + \left(\frac{C_P T}{C_V} + F_\eta - q F_{\eta q} \right) \delta \rho \delta \eta \right. \\ \left. + \frac{1}{2} \left(\frac{\rho T}{C_V} + \rho F_{\eta \eta} - \vec{\omega}_a \cdot \vec{\nabla} F_{\eta q} \right) (\delta \eta)^2 + F_{qq} \left[\frac{\rho}{2} (\delta q)^2 - \delta \eta \vec{\nabla} q \cdot \delta \vec{\omega} \right] + \vec{\nabla} \cdot \vec{\zeta}_b \right\rangle,$$

where $\vec{\zeta}_b \equiv \delta \eta (F_q \delta \vec{\omega} + F_{\eta q} \vec{\omega}_a \delta \eta / 2)$.

At this point, we apply the first step of the EEC method to obtain C_E that satisfies (2.20), which is achieved by determining $F(\eta, q)$ from

$$(2.27) \quad F - q F_q + B = 0.$$

After combining (2.27) with (2.18), each term, except for the last divergence term, on the right-hand side of (2.23) vanishes. Regarding (2.23) and (2.26), we introduce here an important boundary condition on $\delta \eta$:

$$(2.28) \quad \delta \eta_b = 0,$$

where $\delta \eta_b$ denotes the boundary value of $\delta \eta$. With (2.28), $\langle \vec{\nabla} \cdot [F_q \vec{\omega}_a \delta \eta] \rangle$ in (2.23) and $\langle \vec{\nabla} \cdot \vec{\zeta}_b \rangle$ in (2.26) vanish. As to the remaining term $\delta I_b \equiv \langle \vec{\nabla} \cdot [F_q (\delta \vec{v} \times \vec{\nabla} \eta)] \rangle$ in (2.23), we first note that a general solution to (2.27) consists of a homogeneous solution $F^{(h)}$ satisfying $F^{(h)} - q F_q^{(h)} = 0$ and an inhomogeneous one, $F^{(i)}$, with $F^{(i)} - q F_q^{(i)} = -B$, so that we have $F = D(\eta)q + F^{(i)}$ where $D(\eta)$ is an arbitrary function of η . Using Gauss' divergence theorem, we rewrite δI_b in (2.23) as

$$(2.29) \quad \delta I_b = \{F_q (\delta \vec{v} \times \vec{\nabla} \eta) \cdot \vec{n}\} = \{(D(\eta) + F_q^{(i)}) \delta \pi\}, \quad \delta \pi \equiv (\delta \vec{v} \times \vec{\nabla} \eta) \cdot \vec{n},$$

where $\{\bullet\}$ and \vec{n} denote respectively the boundary integral and the outward normal on the boundary. As any 2D boundary surface is densely populated by the streamlines on it along which the values of η as well as q do not change, we see that $\vec{\nabla} \eta$ and $\vec{\nabla} q$ are degenerate. The degeneracy on the boundary means that $F_q^{(i)}(\eta, q)$ assumes the form $F_q^{(i)}(\eta, q) = F_q^{(i)}(\eta, q(\eta))$; specifically, on the boundary $F_q^{(i)}$ can be regarded as a function of η alone. Hence, we can choose $D(\eta)$ such that $D(\eta) = -F_q^{(i)}(\eta, q(\eta))$, which makes δI_b vanish. Thus, with (2.27) and (2.28), we have (2.20) together with

$$(2.30) \quad \delta^2(E + C_E) = \left\langle \frac{\rho}{2} \delta \vec{v} \cdot \delta \vec{v} + \vec{v} \cdot \delta \vec{v} \delta \rho + \frac{(c_s)^2}{2\rho} (\delta \rho)^2 + \left(\frac{C_P T}{C_V} - B_\eta \right) \delta \rho \delta \eta \right. \\ \left. + \frac{1}{2} \left(\frac{\rho T}{C_V} - \rho B_{\eta\eta} - \frac{B_{\eta q}}{q} (\vec{\omega}_a \cdot \vec{\nabla} q) \right) (\delta \eta)^2 + \frac{B_q}{q} \left[\frac{\rho}{2} (\delta q)^2 - \delta \eta \vec{\nabla} q \cdot \delta \vec{\omega} \right] \right\rangle.$$

To conclude this section, with regard to the balanced momentum equation (2.18), we impose the following set of basic conditions that constitute the core assumptions in our stability analysis:

$$(2.31) \quad F_{qq} = \frac{B_q}{q} \neq 0, \quad \vec{\nabla} \eta \times \vec{\nabla} q \neq 0, \quad q \neq 0.$$

The last condition is a well-known condition for symmetric stability discussed, say, by Charney [6]. We recall that the existence of the indefinite term $-B_q \delta \eta \vec{\nabla} q \cdot \delta \vec{\omega} / q$ in (2.30) is an inextricable issue in conventional FS analysis.

§3. New approach that uses a reference iso-Casimir leaf

As explained in the introduction, the time evolution of a given fluid system in the phase space P is described by an orbit Γ_i which lies on the leaves of constant E and any Casimir. The notion of leaves on which a given invariant takes a constant value is important in our stability analysis so that, as a first step, we introduce the definition of a reference Casimir $C_R (\neq C_E)$ mentioned in Section 1:

Definition 3.1. A *reference iso-Casimir leaf* $\mathfrak{C}_R^{(w)}$ is the hyper-surface on which $C_R = w$, where w is a real number.

Geometrically speaking, the conditions (2.20) and (2.21) imply that the equilibrium point O in the phase space is isolated; that is, no orbit Γ_i connects O and a nearby point. Now suppose that there exist orbits Γ_i^o emanating from O . Because any such orbit Γ_i^o is on $\mathfrak{C}_R^{(0)}$ (without loss of generality we may set $w = 0$ for this case) as well as on other leaves such as $E = 0$ and $E + C_E = 0$, the quantity $\delta^2(E + C_E)$ evaluated on the submanifold Γ_i^o of $\mathfrak{C}_R^{(0)}$ cannot be sign-definite as Γ_i^o provides a separatrix of the Lyapunov functional $L = E + C_E$ on the leaf of $\mathfrak{C}_R^{(0)}$. This observation tells us that if we can show the sign-definiteness of $\delta^2(E + C_E)$ along Γ_i^o on $\mathfrak{C}_R^{(0)}$, then we can claim that no such Γ_i^o exists in the phase space. Notice that the conventional FS analysis requires the sign-definiteness of $\delta^2(E + C_E)$ for arbitrary variations of independent variables, whereas the condition we have just mentioned here requires the conditional sign-definiteness of $\delta^2(E + C_E)$ on the submanifold Γ_i^o of $\mathfrak{C}_R^{(0)}$, which is a key ingredient of our new stability analysis.

As we have already argued towards the end of the Introduction, conventional FS provides a mathematically stronger (physically less sharp) criterion than FS2.

As a well-known example, let us consider the FS criterion of 2D incompressible flows derived by Arnol'd; it assumes the form $d\psi(\omega)/d\omega > 0$ where ψ and ω respectively denote the stream function and vorticity (cf. (5.3)). From a practical viewpoint of GFD applications, where ψ and ω are negatively correlated through the geostrophic balance, the above criterion is not helpful for those who are interested in distinguishing linearly stable geostrophic flows from unstable ones. The major motivation in introducing the notion of FS2 comes from such applicational concerns. We readily see that FS2 is essentially related to the notion of the hyperbolic point for linearised dynamical systems and, in this sense, FS2 would cover linear stability analyses so far attempted employing, say, spectral (or normal modes) approaches. Nevertheless, the fact that FS2 provides a weaker condition than FS suggests a possibility that it may overlook some (presumably quite complex) unstable modes. Indeed, the topology of a constant hyper-surface of the invariants such as E , C_E and C_R of an infinite-dimensional system could become complicated (even pathological). There may be some cases for which our dichotomic classification on a couple of different separatrices of $\delta^2 L$ (physical unstable manifolds vs. virtual separatrices mentioned at the end of the Introduction) is too naive to cover such cases appropriately. Considering the present status of not having a practical stability criterion for 3D stratified shear flows, we believe that the attainability of FS2 would be of considerable value even if we leave this subtle issue unresolved. As will be seen, a substantial advantage of FS2 over FS is that we can use additional information coming from C_R . Hence, as a refined definition of FS given in this paper, we have

Definition 3.2. A given equilibrium point O is said to be stable *in the sense of FS of the second kind (FS2 for short)* if there exists no orbit Γ_i^o on $\mathfrak{C}_R^{(0)}$ connecting O to another point in the neighbourhood of it. A reference Casimir C_R associated with $\mathfrak{C}_R^{(0)}$ is now specifically defined as

$$(3.1) \quad C_R \equiv C_1 + C_2, \quad C_1 \equiv \langle \rho K(q) \rangle, \quad C_2 \equiv \frac{1}{2} \langle \rho q^2 \eta \rangle,$$

where $K(q)$ is a function of q satisfying

$$(3.2) \quad K''' \equiv \frac{d^3 K}{dq^3} \neq 0.$$

By straightforward calculations similar to those given in Appendix A, we obtain, up to second-order accuracy,

$$(3.3) \quad \delta C_R = (K - qK' - \frac{1}{2}\eta q^2)\delta\rho + (K' + \eta q)\vec{\nabla}\eta \cdot \delta\vec{\omega} - [(K'' + \eta)(\vec{\omega}_a \cdot \vec{\nabla}q) + \frac{1}{2}\rho q^2]\delta\eta,$$

$$(3.4) \quad \delta^2 C_R = \langle (K'' + \eta)[\frac{1}{2}\rho(\delta q)^2 - \delta\eta\vec{\nabla}q \cdot \delta\vec{\omega}] - \frac{1}{2}(\vec{\omega}_a \cdot \vec{\nabla}q)(\delta\eta)^2 - \frac{1}{2}q^2\delta\rho\delta\eta \rangle.$$

In deriving (3.3) and (3.4), we have used the boundary condition (2.28). To evaluate $\delta^2(E + C_E)$ on the submanifold Γ_i^o of $\mathfrak{C}_R^{(0)}$, we now introduce restricted variations: $[\delta\vec{v}_{(R)}, \delta\rho_{(R)}, \delta\eta_{(R)}]$ on the tangent hyper-plane $\mathfrak{T}\mathfrak{C}_R^{(0)}$ of $\mathfrak{C}_R^{(0)}$. By definition, we have

$$(3.5) \quad \delta C_R|_{\mathfrak{T}\mathfrak{C}_R^{(0)}} \equiv \delta C_R\{\delta\vec{v}_R, \delta\rho_R, \delta\eta_R\} = 0,$$

so that any unstable manifold Γ_j^o , if it exists, must lie on $\mathfrak{T}\mathfrak{C}_R^{(0)}$. Furthermore, the existence of Γ_j^o indicates that there is $[\delta\vec{v}_{(R)}^*, \delta\rho_{(R)}^*, \delta\eta_{(R)}^*] \in [\delta\vec{v}_{(R)}, \delta\rho_{(R)}, \delta\eta_{(R)}]$ such that

$$(3.6) \quad \delta^2 C_R\{\delta\vec{v}_R^*, \delta\rho_R^*, \delta\eta_R^*\} = 0.$$

At the end of Section 2, we pointed out a term in (2.30) that represents a stumbling block in the conventional FS analysis. Notice that the same factor appears in (3.4). The reason why we introduce C_R is that we shall use it to eliminate the difficulty in the proverbial sense that *one nail (difficulty) drives out another*. To do so, let us introduce a vector field \vec{A} satisfying

$$(3.7) \quad \vec{A} \equiv \vec{\nabla}\chi + \vec{\alpha}, \quad \vec{\nabla} \cdot \vec{A} = \nabla^2\chi = \frac{1}{2}\rho(\delta q_R^*)^2 - \delta\eta_R^*(\vec{\nabla}q \cdot \delta\vec{\omega}_R^*), \quad \vec{\nabla} \cdot \vec{\alpha} = 0.$$

For given $\rho(\delta q_R^*)^2/2 - \delta\eta_R^*(\vec{\nabla}q \cdot \delta\vec{\omega}_R^*)$, χ can be determined from the second expression of (3.7) using the Neumann boundary condition,

$$(3.8) \quad \vec{\nabla}\chi|_b \cdot \vec{n} = 0,$$

whereas the nondivergent $\vec{\alpha}$ is a free parameter; this situation can be regarded as exactly opposite to the electromagnetic vector potential “ \vec{A}_{em} ” for which the rotational part possessing $U(1)$ gauge symmetry has a physical meaning. To see the important role played by the vector field, consider the following identity derived by integrating by parts:

$$(3.9) \quad \langle (K'' + \eta)\vec{\nabla} \cdot \vec{A} \rangle = -\langle \vec{A} \cdot \vec{\nabla}(K'' + \eta) \rangle + \langle \vec{\nabla} \cdot [(K'' + \eta)\vec{A}] \rangle,$$

and observe that the left-hand side of (3.9) depends on χ but not on the free parameter $\vec{\alpha}$ in (3.7), although the latter formally appears in its right-hand side. Substituting (3.9) into (3.6) (while referring to its concrete form in (3.4)), we obtain

$$(3.10) \quad \begin{aligned} 0 &= \delta^2 C_R\{\delta\vec{v}_R^*, \delta\rho_R^*, \delta\eta_R^*\} \\ &= \langle (K'' + \eta) [\frac{1}{2}\rho(\delta q_R^*)^2 - \delta\eta_R^*\vec{\nabla}q \cdot \delta\vec{\omega}_R^*] - \frac{1}{2}(\vec{\omega}_a \cdot \vec{\nabla}q)(\delta\eta_R^*)^2 - \frac{1}{2}q^2\delta\rho_R^*\delta\eta_R^* \rangle \\ &= \langle (K'' + \eta)\vec{\nabla} \cdot \vec{A} - \xi_R^* \rangle = -\langle \vec{A} \cdot \vec{\nabla}(K'' + \eta) + \xi_R^* \rangle + \langle \vec{\nabla} \cdot [(K'' + \eta)\vec{A}] \rangle \\ &= -\langle (K'' + \eta)\vec{\nabla}q \cdot \vec{A} \rangle + \langle \vec{\nabla}\eta \cdot \vec{A} + \xi_R^* \rangle + \langle \vec{\nabla} \cdot [(K'' + \eta)\vec{A}] \rangle, \end{aligned}$$

where $\xi_R^* \equiv \delta\eta_R^*[(\vec{\omega}_a \cdot \vec{\nabla}q)(\delta\eta_R^*) + q^2\delta\rho_R^*]/2$. Using the arbitrariness in $\vec{\alpha}$ we make a specific choice such that

$$(3.11) \quad \vec{\nabla}q \cdot (\vec{\nabla}\chi + \vec{\alpha}) = 0, \quad \vec{\nabla}\eta \cdot (\vec{\nabla}\chi + \vec{\alpha}) + \xi_R^* = 0,$$

which is now free from the nonzero undetermined function K''' (cf. (3.2)). In justifying (3.11), we first note that the substantial degree of freedom of $\vec{\alpha}$ is not three but two. Indeed, the condition $\vec{\nabla} \cdot \vec{\alpha} = 0$ allows us to introduce $\vec{\lambda}$ satisfying

$$(3.12) \quad \vec{\alpha} = \vec{\nabla} \times \vec{\lambda},$$

with an arbitrary auxiliary constraint on $\vec{\lambda}$ to reduce the degree of freedom by one. Second, from the nondegeneracy condition on $\vec{\nabla}\eta$ and $\vec{\nabla}q$ given in (2.31), we see that (3.11) gives a couple of independent equations for determining $\vec{\alpha}$; this is consistent with the fact that its number of substantial degrees of freedom is two. The geometrical meaning of (3.11) is quite simple: specifically, $\vec{\alpha}$ can be determined by specifying its components projected to a couple of independent vectors $\vec{\nabla}\eta$ and $\vec{\nabla}q$ together with an additional auxiliary constraint.

Using the vector identity $\vec{\nabla}P \cdot \vec{\alpha} = \vec{\nabla} \cdot (\vec{\lambda} \times \vec{\nabla}P)$ and the fact that q as well as η are nonvanishing (see (2.31)), we can rewrite (3.11) as

$$(3.13) \quad \vec{\nabla} \cdot (\vec{\lambda} \times \vec{\nabla} \ln q) = -\vec{\nabla} \ln q \cdot \vec{\nabla}\chi, \quad \vec{\nabla} \cdot (\vec{\lambda} \times \vec{\nabla} \ln \eta) = -\vec{\nabla} \ln \eta \cdot \vec{\nabla}\chi - \frac{\xi_R^*}{\eta}.$$

As $\vec{\nabla}\eta$, $\vec{\nabla}q$ and \vec{v} are linearly independent vectors, they form a basis of a 3D vector space. Hence

$$(3.14) \quad \vec{\lambda} = \frac{Bq}{q} [\kappa_{(\eta)} \vec{\nabla} \ln \eta + \kappa_{(q)} \vec{\nabla} \ln q] + \kappa_{(v)} \vec{v},$$

where $\kappa_{(\eta)}$, $\kappa_{(q)}$ and $\kappa_{(v)}$ are scalar variables. Using the first equation of (2.18), and (3.14), we have

$$(3.15) \quad \vec{\lambda} \times \vec{\nabla} \ln q = \frac{\kappa_{(\eta)}}{\eta q} \rho \vec{v} + \kappa_{(v)} \vec{L}_q, \quad \vec{\lambda} \times \vec{\nabla} \ln \eta = -\frac{\kappa_{(q)}}{\eta q} \rho \vec{v} + \kappa_{(v)} \vec{L}_\eta,$$

where $\vec{L}_q \equiv \vec{v} \times \vec{\nabla} \ln q$ and $\vec{L}_\eta \equiv \vec{v} \times \vec{\nabla} \ln \eta$. Substituting (3.15) into (3.13), we obtain

$$(3.16) \quad \frac{\rho}{\eta q} \vec{v} \cdot \vec{\nabla} \kappa_{(\eta)} + \vec{L}_q \cdot \vec{\nabla} \kappa_{(v)} + (\vec{\nabla} \cdot \vec{L}_q) \kappa_{(v)} = -\vec{\nabla} \ln q \cdot \vec{\nabla}\chi,$$

$$(3.17) \quad -\frac{\rho}{\eta q} \vec{v} \cdot \vec{\nabla} \kappa_{(q)} + \vec{L}_\eta \cdot \vec{\nabla} \kappa_{(v)} + (\vec{\nabla} \cdot \vec{L}_\eta) \kappa_{(v)} = -\vec{\nabla} \ln \eta \cdot \vec{\nabla}\chi - \frac{\xi_R^*}{\eta}.$$

Adding (3.16) and (3.17) yields

$$(3.18) \quad \frac{\rho \vec{v}}{\eta q} \cdot \vec{\nabla} (\kappa_{(\eta)} - \kappa_{(q)}) + \vec{L} \cdot \vec{\nabla} \kappa_{(v)} + (\vec{\nabla} \cdot \vec{L}) \kappa_{(v)} = -(\vec{\nabla} \ln q + \vec{\nabla} \ln \eta) \cdot \vec{\nabla}\chi - \frac{\xi_R^*}{\eta},$$

where $\vec{L} \equiv \vec{L}_q + \vec{L}_\eta$. Because we can impose an auxiliary condition on $\vec{\lambda}$, we may set $\vec{v} \cdot \vec{\nabla}(\kappa_{(\eta)} - \kappa_{(q)}) = 0$ as such a condition. Hence (3.18) becomes a first-order linear partial differential equation for $\kappa_{(v)}$ and the existence of a solution is guaranteed by the Cauchy–Kovalevskaya theorem. $\kappa_{(\eta)}$ and $\kappa_{(q)}$ can be respectively determined in a similar fashion using (3.16) and (3.17) after specifying $\kappa_{(v)}$ via (3.18). We have shown that there exists $\vec{\alpha}$ satisfying (3.11).

Using (2.28), we see that ξ_R^* defined in (3.10) becomes zero at the boundaries, so that

$$(3.19) \quad \vec{\nabla}q \cdot \vec{A}|_b = 0, \quad \vec{\nabla}\eta \cdot \vec{A}|_b = 0.$$

Comparing (3.19) with the first identity of (2.18), we see that $\vec{A}|_b$ is parallel to $\vec{v}|_b$, which is perpendicular to the outward normal \vec{n} there. Hence, the last divergence term in (3.10) also vanishes, thereby proving that $\delta^2 C_R \{\delta\vec{v}_R^*, \delta\rho_R^*, \delta\eta_R^*\} = 0$ is successfully parameterised by $\vec{\alpha}$ introduced in (3.11).

As a last step, we apply the above parameterisation to the last term in (2.30) ($\equiv \delta^2 J_R^*$) evaluated on the submanifold Γ_i^o of $\mathfrak{C}_R^{(0)}$:

$$(3.20) \quad \begin{aligned} \delta^2 J_R^* &\equiv \left\langle \frac{B_q}{q} \left[\frac{\rho}{2} (\delta q_R^*)^2 - \delta\eta_R^* \vec{\nabla}q \cdot \delta\vec{\omega}_R^* \right] \right\rangle = \left\langle \frac{B_q}{q} \vec{\nabla} \cdot \vec{A} \right\rangle \\ &= - \left\langle \vec{\nabla} \left(\frac{B_q}{q} \right) \cdot \vec{A} \right\rangle = - \left\langle \frac{1}{q} \left(B_{qq} - \frac{B_q}{q} \right) \vec{\nabla}q \cdot \vec{A} + \frac{B_{q\eta}}{q} \vec{\nabla}\eta \cdot \vec{A} \right\rangle, \end{aligned}$$

where $\vec{A}|_b \perp \vec{n}$ is used. We see that the substitution of (3.11) into (3.20) makes $\delta^2 J_R^*$ tractable in our FS2 analysis.

§4. FS2 analysis for 3D stratified shear flows

Main Theorem 4.1. *A given steady state of a 3D stratified shear flow of (2.16) is stable in the sense of FS2 if the following three conditions hold:*

$$(4.1) \quad \begin{aligned} q &\neq 0, \quad (c_s)^2 - V^2 > 0, \\ ((c_s)^2 - V^2) \left(\frac{T}{C_V} - B_{\eta\eta} \right) - \left(\frac{C_P T}{C_V} - B_\eta + \frac{q}{2} B_{q\eta} \right)^2 &> 0, \end{aligned}$$

where $V^2 \equiv \vec{v} \cdot \vec{v}$.

Proof. Having prepared the parameterisation of Γ_i^o on $\mathfrak{C}_R^{(0)}$, we now apply it to (2.30). Directly from (2.30), the value of $\delta^2(E + C_E)$ evaluated on the submanifold Γ_i^o of $\mathfrak{C}_R^{(0)}$ is given by

$$\begin{aligned}
(4.2) \quad & \delta^2(E + C_E)|_{\mathfrak{e}_R^{(0)}} \\
& \equiv \left\langle \frac{\rho}{2} \delta \vec{v}_R^* \cdot \delta \vec{v}_R^* + \vec{v} \cdot \delta \vec{v}_R^* \delta \rho_R^* + \frac{(c_s)^2}{2\rho} (\delta \rho_R^*)^2 + \left(\frac{C_P T}{C_V} - B_\eta \right) \delta \rho_R^* \delta \eta_R^* \right. \\
& \quad \left. + \frac{1}{2} \left(\frac{\rho T}{C_V} - \rho B_{\eta\eta} - \frac{B_{\eta q}}{q} (\vec{\omega}_a \cdot \vec{\nabla} q) \right) (\delta \eta_R^*)^2 + \frac{B_q}{q} \left[\frac{\rho}{2} (\delta q_R^*)^2 - \delta \eta_R^* \vec{\nabla} q \cdot \delta \vec{\omega}_R^* \right] \right\rangle.
\end{aligned}$$

Substitution of (3.20) into (4.2) using (3.11) yields

$$\begin{aligned}
(4.3) \quad & \delta^2(E + C_E)|_{\mathfrak{e}_R^{(0)}} \\
& = \left\langle \frac{\rho}{2} \delta \vec{v}_R^* \cdot \delta \vec{v}_R^* + \vec{v} \cdot \delta \vec{v}_R^* \delta \rho_R^* + \frac{(c_s)^2}{2\rho} (\delta \rho_R^*)^2 + \left(\frac{C_P T}{C_V} - B_\eta \right) \delta \rho_R^* \delta \eta_R^* \right. \\
& \quad \left. + \frac{1}{2} \left(\frac{\rho T}{C_V} - \rho B_{\eta\eta} - \frac{B_{\eta q}}{q} (\vec{\omega}_a \cdot \vec{\nabla} q) \right) (\delta \eta_R^*)^2 + \frac{B_{\eta q}}{q} \xi_R^* \right\rangle,
\end{aligned}$$

which finally becomes

$$\begin{aligned}
(4.4) \quad & \delta^2(E + C_E)|_{\mathfrak{e}_R^{(0)}} \\
& = \left\langle \frac{\rho}{2} \delta \vec{v}_R^* \cdot \delta \vec{v}_R^* + \vec{v} \cdot \delta \vec{v}_R^* \delta \rho_R^* + \frac{(c_s)^2}{2\rho} (\delta \rho_R^*)^2 \right. \\
& \quad \left. + \left(\frac{C_P T}{C_V} - B_\eta + \frac{q}{2} B_{\eta q} \right) \delta \rho_R^* \delta \eta_R^* + \frac{\rho}{2} \left(\frac{T}{C_V} - B_{\eta\eta} \right) (\delta \eta_R^*)^2 \right\rangle \\
& \geq \left\langle \frac{\rho}{2V^2} (\delta u_R^*)^2 + \delta u_R^* \delta \rho_R^* + \frac{(c_s)^2}{2\rho} (\delta \rho_R^*)^2 + N_1 \delta \rho_R^* \delta \eta_R^* + \frac{\rho}{2} N_2 (\delta \eta_R^*)^2 \right\rangle,
\end{aligned}$$

where $\delta u_R^* \equiv \vec{v} \cdot \delta \vec{v}_R^*$, $N_1 \equiv C_P T / C_V - B_\eta + q B_{\eta q} / 2$ and $N_2 \equiv T / C_V - B_{\eta\eta}$. By applying Sylvester's Criterion, which states that a real symmetric matrix is positive definite if and only if all the leading principal minors of it are positive, to

$$(4.5) \quad M_{ij} = \begin{pmatrix} \rho/V^2 & 1 & 0 \\ 1 & (c_s)^2/\rho & N_1 \\ 0 & N_1 & \rho N_2 \end{pmatrix},$$

we get (4.1), which is the central result of this paper.

The fact that the norm given in (4.4) is coercive can be checked as follows. To make manipulations simple, we redefine nondimensional perturbations: $\delta \hat{u} \equiv \delta u_R^* / V^2$, $\delta \hat{\rho} \equiv \delta \rho_R^* / \rho$ and $\delta \hat{\eta} \equiv \delta \eta_R^* / C_P$. Using these, one can rewrite (4.4) as

$$\begin{aligned}
(4.6) \quad & \delta^2(E + C_E)|_{\mathfrak{e}_R^{(0)}} \\
& \geq \left\langle (\rho V^2) \left[\frac{1}{2} (\delta \hat{u})^2 + \delta \hat{u} \delta \hat{\rho} + \frac{W_1 + 1}{2} (\delta \hat{\rho})^2 + W_2 \delta \hat{\rho} \delta \hat{\eta} + \frac{W_3}{2} (\delta \hat{\eta})^2 \right] \right\rangle,
\end{aligned}$$

where $W_1 \equiv (c_s)^2/V^2 - 1$, $W_2 \equiv C_P N_1/V^2$ and $W_3 \equiv (C_P)^2 N_2/V^2$. Using the following three different expressions for the quantity in square brackets above:

$$\begin{aligned}
 (4.7) \quad \mathfrak{J} &\equiv \frac{1}{2}(\delta\hat{u})^2 + \delta\hat{u}\delta\hat{\rho} + \frac{W_1+1}{2}(\delta\hat{\rho})^2 + W_2\delta\hat{\rho}\delta\hat{\eta} + \frac{W_3}{2}(\delta\hat{\eta})^2 \\
 &= \frac{1}{2}(\delta\hat{u} + \delta\hat{\rho})^2 + \frac{W_1}{2}\left(\delta\hat{\rho} + \frac{W_2}{W_1}\delta\hat{\eta}\right)^2 + \frac{W_*}{2W_1}(\delta\hat{\eta})^2 \quad \left[\geq \frac{W_*}{2}\left(\frac{\delta\hat{\eta}}{\sqrt{W_1}}\right)^2 \right] \\
 &= \frac{1}{2}(\delta\hat{u} + \delta\hat{\rho})^2 + \frac{W_3}{2}\left(\delta\hat{\eta} + \frac{W_2}{W_3}\delta\hat{\rho}\right)^2 + \frac{W_*}{2W_3}(\delta\hat{\rho})^2 \quad \left[\geq \frac{W_*}{2}\left(\frac{\delta\hat{\rho}}{\sqrt{W_3}}\right)^2 \right] \\
 &= \frac{W_3}{2}\left(\delta\hat{\eta} + \frac{W_2}{W_3}\delta\hat{\rho}\right)^2 + \frac{W_4}{2W_3}\left(\delta\hat{\rho} + \frac{W_3}{W_4}\delta\hat{u}\right)^2 + \frac{W_*}{2W_4}(\delta\hat{u})^2 \quad \left[\geq \frac{W_*}{2}\left(\frac{\delta\hat{u}}{\sqrt{W_4}}\right)^2 \right],
 \end{aligned}$$

where $W_* \equiv W_1W_3 - (W_2)^2$, $W_4 \equiv W_* + W_3$, and the conditions $W_1 > 0$, $W_3 > 0$, $W_* > 0$, $W_4 > 0$, all derived from (4.1), we obtain

$$\begin{aligned}
 (4.8) \quad \delta^2(E + C_E)|_{\mathfrak{C}_R^{(0)}} &\geq \frac{1}{6}\left\langle (\rho V^2)W_* \left[\left(\frac{\delta\hat{u}}{\sqrt{W_4}}\right)^2 + \left(\frac{\delta\hat{\rho}}{\sqrt{W_3}}\right)^2 + \left(\frac{\delta\hat{\eta}}{\sqrt{W_1}}\right)^2 \right] \right\rangle > 0. \quad \square
 \end{aligned}$$

Remark. As we have explained, FS2 means a state-dependent topological nature of L which can be derived by looking into conditional local minima of L along submanifolds Γ_i^o of $\mathfrak{C}_R^{(0)}$. In this regard, we can say that the basic notion of FS2 is similar to that of Vallis et al. [24] in which conditional energy minima on the iso-vortical leaf were utilised to find stable solutions of 2D Euler equations. Actually, for a 2D incompressible flow, it is known that Arnol'd's formal stability condition (cf. (5.3)) remains substantially the same for the following two formulations: (1) checking the positive definiteness of $\delta^2(E + C_E)$ for arbitrary perturbations and (2) checking the same quantity for iso-vortical perturbations satisfying $\delta\langle\omega\rangle = 0$ where ω denotes vorticity. Of course, our 3D stability analysis is different from Arnol'd's original 2D method and it should be emphasised that our approach does not directly deal with arbitrary perturbations, which suggests that there may exist unstable solutions that satisfy the FS2 condition. Secondly, in some situations, as the well-known 2D stability norm of Arnol'd exemplifies, velocity and vorticity variations are treated as independent ones though they are not totally independent. The reason for such treatment is the simplicity of analysis but the price to pay is that this gives a stronger (physically less sharp) stability criterion. Thus, from the viewpoint of applications, it is quite favourable that the independent variational variables employed coincide with physically independent variables. Note that the FS2 analysis presented in this paper gives such an example.

In the context of GFD, for a steady state with zonal symmetry (x -independent), that is, $\vec{v} \rightarrow U(y, z)$, the steady momentum equation derived from (2.16) can be rewritten in the η - q coordinate system as

$$(4.9) \quad T = B_\eta - UU_\eta^{(f)}, \quad B_q = UU_q^{(f)},$$

where

$$(4.10) \quad U^{(f)} \equiv U(y, z) - \int_{y_0}^y f(y) dy$$

with $f(y)$ the ‘‘latitude’’ (y)-dependent Coriolis parameter. The derivation of (4.9) is given in Appendix B. By the use of (4.9), the third condition in (4.1) can be rewritten with the introduction of the stability index I_s as

$$(4.11) \quad I_s \equiv \frac{1}{C_P} \left(gz + \frac{U^2}{2} \right)_\eta + \frac{1}{C_P} UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta - \frac{q}{C_P} (UU_q^{(f)})_\eta - \frac{U^2}{C_P^2} \\ - M^2 \left[\frac{1}{C_P} \left(gz + \frac{U^2}{2} \right)_\eta - \frac{1}{C_P} UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta \right] \\ - M^2 \left(-U_\eta^{(f)} + \frac{q}{2} U_{q\eta}^{(f)} \right)^2 - MM_p \left(-U_\eta^{(f)} + \frac{q}{2} U_{q\eta}^{(f)} \right) U_\eta - \frac{1}{4} U_\eta^2 M_p^2 > 0,$$

where $M \equiv U/c_s$ and $M_p \equiv qU_q^{(f)}/c_s$ are the Mach and tentatively introduced ‘‘pseudo-Mach’’ numbers; a detailed derivation of (4.11) is given in Appendix C. If both numbers are much smaller than unity, then it can be shown that, for the potential temperature θ , $d\eta \equiv C_P d \ln \theta$ and q coordinates, the above reduces to

$$(4.12) \quad \left[\frac{1}{\theta} \left(gz + \frac{U^2}{2} \right)_\theta - U_\theta U_\theta^{(f)} - UU_{\theta\theta}^{(f)} \right] - \frac{q}{\theta} [UU_q^{(f)}]_\theta - \frac{U^2}{\theta^2} > 0,$$

where we have used the identity

$$(4.13) \quad UU_\eta^{(f)}/C_P - [UU_\eta^{(f)}]_\eta = -\theta^2 [UU_\theta^{(f)}]_\theta / C_P^2.$$

§5. Some remarks on the approximate criterion (4.12)

A caveat concerning (4.12) is in order. In our reference frame, the zonal velocity $U(y, z)$ is strongly related to the pressure gradient through the geostrophic balance, so that the dynamical meaning of the zonal symmetry in our case is not equivalent to that of the x -invariant symmetry, say, for 2D (in the x - z plane) stratified shear flow $U(z)$, for which the reference point can be freely chosen as the result of the invariance under Galilean transformations. Several new factors appear in (4.12) because the stability criteria known so far are for 2D or quasi-2D incompressible flows. The new terms are discussed in order of appearance in (4.12).

- (I) *The first represents a generalised Richardson number criterion including shear (the first derivatives of the background velocity) as well as curvature (its second derivatives) terms evaluated along a couple of independent directions in the y - z plane.*

A well-known stability (or instability) criterion dealing with y - z dependent flows is the Charney–Stern (CS) [7] barotropic-baroclinic instability theorem derived for a quasi-geostrophic (QG) system of equations, and is regarded as a direct extension of the Rayleigh–Fjørtoft (RF) theorem as a necessary condition for instability of 2D barotropic Rossby waves. The CS theorem states that the meridional (y) gradient of the pseudo-potential vorticity q_p (approximated by q for a QG system) changes sign in the domain (the inflection point theorem) if a given flow is unstable. It is a direct extension of the 2D RF theorem because the latter is recovered formally through replacing dq_p/dy in the CS theory by $d\omega/dy$ where $\omega \equiv -dU/dy$ is the background vorticity. Note that for 2D incompressible (non-divergent) flows, vorticity is an advective scalar quantity satisfying $D\omega/Dt = 0$, so that the general form of the Casimir of this system becomes $F^{(A)} = F^{(A)}(\omega)$. Although both RF and CS conclusions are derived by a conventional analysis employing normal modes (spectral instability), it must be noted that basically a similar *contrapositive* conclusion was obtained by Arnol'd [3] for 2D nondivergent flows in a nonrotating ($f = 0$) reference frame. This says that a given 2D steady state is Arnol'd stable (FS) if U_{yy} ($= -\omega_y$) is sign-definite. It is in his method that the prototype of the ECC method was first applied, and the original form of his FS criterion takes the form

$$(5.1) \quad \frac{d^2 F^{(A)}}{d\omega^2} > 0.$$

Vanishing of the first variation of the Lyapunov functional employed yields the condition

$$(5.2) \quad \frac{dF^{(A)}}{d\omega} = \psi(\omega),$$

where $\psi(\omega)$ is a stream function of the given steady flow, so that (5.1) can be alternatively written as

$$(5.3) \quad \frac{d\psi}{d\omega} > 0.$$

The aforementioned condition of the sign-definiteness of $d\omega/dy$ is readily derived from (5.3) by changing the independent variable, $\omega \rightarrow y$, that is,

$$(5.4) \quad \frac{d\psi}{d\omega} = \frac{d\psi}{dy} \frac{dy}{d\omega} = \frac{-U}{\omega_y} > 0,$$

because the sign-definiteness of $-U$ is maintained under Galilean transformations.

In the 2D meridional plane, the η and q coordinates correspond, roughly speaking, to the vertical and the latitudinal coordinates, respectively. Because the Coriolis parameter f in q increases with latitude, it is expected that a similar inflection-type property must be reflected in our new criterion (4.12). To see this, we focus on the term containing the “horizontal” q derivative, specifically, $-q[UU_q^{(f)}]_\theta/\theta$ in (4.12). For the 2D “horizontal” aspect of (4.12), let us consider an isentropic surface on which $\theta = \theta_c = \text{const}$. For simplicity, we consider the case for which the method of separation of variables is applicable to $B(\theta, q)$, and gives $B = D(\theta)\hat{B}(q)$. In view of the second equation of (4.9), the “stability on the isentropic surface” when comparing (4.12) with (5.1) or (5.3) is given in our context by

$$(5.5) \quad -\frac{q}{\theta}[UU_q^{(f)}]_\theta = -\frac{q}{\theta_c}(B_q)_\theta = -\frac{D'}{\theta_c}(q\hat{B}_q) = -\frac{D'}{\theta_c}(q^2\hat{F}_{qq}) > 0,$$

where (2.27) ($-qF_{qq} + B_q = 0$) has been used. Because D'/θ_c has a constant value on the isentropic surface and $q \neq 0$ by assumption, we see that (5.5) certainly assumes a non-inflection-type structure similar to (5.1) and (5.3), for which we have the following correspondence:

$$(5.6) \quad \frac{d^2F^{(A)}}{d\omega^2} > 0 \Leftrightarrow -\frac{D'}{\theta_c}(q^2\hat{F}_{qq}) > 0, \quad \frac{d\psi}{d\omega} > 0 \Leftrightarrow -\frac{D'}{\theta_c}(q\hat{B}_q) > 0.$$

The “vertical” counterpart consists of the three terms in the first square bracket in (4.12). Except for the explicit f dependence, it contains the familiar Richardson number expression $gz_\theta/\theta - U_\theta U_\theta^{(f)}$ given by Miles [17] and Howard [12] and its 3D extension by Mahalov [15] together with a couple of additional new terms of $(U^2)_\theta/2\theta$ and $UU_{\theta\theta}^{(f)}$. The latter is a curvature term missing in the conventional analyses employing normal modes, whereas the former, as is explained shortly in (III), stems implicitly from the geostrophic balance of a given steady state. Hence, in the sense mentioned above, (4.12) is a criterion combining CS (or RF) and the Miles–Howard-type criteria.

(II) *Compressibility of fluid*

The last term in (4.12), $-U^2/\theta^2$, clearly reflects a thermodynamical property of the fluid which is missing in the dynamics of an incompressible fluid. Generally speaking, because large θ corresponds to high acoustic-wave speed, we can say that this term plays the same role as $-U^2$ in the second condition of (4.1).

(III) *Effect of Coriolis parameter f*

Another important new factor in (4.12) is the dependence of f on shear instability. Presumably, a stability criterion in which f appears explicitly is the well-

known inertial stability/instability criterion [11] that implies that a zonal flow in the northern hemisphere where $f > 0$ is inertially stable when the absolute vorticity $f - U_y$ is positive. In the vorticity equation for 2D or quasi-2D horizontal flows, the magnitude of vorticity (absolute vorticity if f is included) appears as a product of (absolute) vorticity and divergence terms, so that this term disappears if the given flow is nondivergent. In a QG system in GFD, the flow is assumed to be quasi-nondivergent, whereas in the argument of inertial instability, it is not. For 2D nondivergent flows, to which we referred in explaining the result of Arnol'd, what matters is U_{yy} , so that it is not f but its meridional derivative β that enters the stability criterion of geophysical flows. Hence, we can say that the explicit appearance of f in our divergent flows is expected from the argument of inertial instability, although that effect is missing in prevailing stability theorems for 2D nondivergent flows in GFD.

To clarify the meaning of $(U^2)_\theta/2\theta$ in (4.12), we have to look back at its derivation and find its origin. From the calculations covering (5.27), (5.28) and (5.29) in Appendix C, we see that a couple of terms $(U^2)_\eta/2C_P$ and $-UU_\eta^{(f)}/C_P$ arise from splitting $-U(fy_\eta)/C_P$ using the η -differentiated version of (4.10). This splitting can be done under the conditions that $f \neq 0$ and $y_\eta \neq 0$, which are generally satisfied for geostrophic balances.

As final remarks, we point out a couple of studies relating to the theme of this paper. The first is that of Abarbanel et al. [2]. To get a 2D nonlinear generalisation of the Richardson number criterion, they discussed an asymptotic approach in a dimensional reduction from 3D to 2D of the equation for a Boussinesq fluid with nonvanishing potential vorticity field. Although the background assumptions for their system are different from ours, under the assumption that the curvature terms are negligible, the criterion derived by their asymptotic EEC method gives the same result $R_i > 1$ as ours, which is contrasted with the spectral stability criterion $R_i > 1/4$ derived by Miles [17]. The second is the study by Sakuma and Ghil [21], the result of which on moving coherent vortices may be generalised with the use of the present result if such 3D stable steady states actually exist.

Appendix A

(I) *Calculation of the first and second variations of the total energy E given in (2.19).* We start from the first law of thermodynamics for an ideal gas:

$$(5.7) \quad de = Td\eta + \frac{RT}{\rho}d\rho, \quad e = C_V T,$$

from which we obtain

$$(5.8) \quad e_\eta = T, \quad e_\rho = \frac{RT}{\rho},$$

where the subscript denotes partial derivative. Based on (5.7) and (5.8), we can calculate the second derivatives of e :

$$(5.9) \quad e_{\eta\eta} = \frac{T}{C_V}, \quad e_{\rho\eta} = \frac{RT}{C_V\rho}, \quad e_{\rho\rho} = \frac{R}{\rho^2} \left(\frac{RT}{C_V} - T \right).$$

Up to second-order accuracy, the expansion of the first and third terms of E given in (2.19) defined as $\Delta^{(2)}E_{I,III}$ becomes

$$(5.10) \quad \begin{aligned} \Delta^{(2)}E_{(I,III)} &= \left\langle \frac{1}{2}(\rho + \delta\rho)(\vec{v} + \delta\vec{v}) \cdot (\vec{v} + \delta\vec{v}) - \frac{1}{2}\rho\vec{v} \cdot \vec{v} + gz\delta\rho \right\rangle \\ &= \left\langle \left[\left(\frac{1}{2}\vec{v} \cdot \vec{v} + gz \right) \delta\rho + \rho\vec{v} \cdot \delta\vec{v} \right] + \left[\frac{\rho}{2}\delta\vec{v} \cdot \delta\vec{v} + \delta\rho\vec{v} \cdot \delta\vec{v} \right] \right\rangle. \end{aligned}$$

Similarly, the expansion of the second term $\Delta^{(2)}E_{II}$ becomes

$$(5.11) \quad \begin{aligned} \Delta^{(2)}E_{(II)} &= \langle (\rho + \delta\rho)e(\rho + \delta\rho, \eta + \delta\eta) - \rho e(\rho, \eta) \rangle \\ &= \langle (\rho + \delta\rho)(e + e_\rho\delta\rho + e_\eta\delta\eta + \frac{1}{2}e_{\rho\rho}(\delta\rho)^2 + e_{\rho\eta}\delta\rho\delta\eta + \frac{1}{2}e_{\eta\eta}(\delta\eta)^2 + \dots) - \rho e \rangle \\ &= \left\langle [(\rho e_\rho + e)\delta\rho + \rho e_\eta\delta\eta] + \left[\left(\frac{\rho}{2}e_{\rho\rho} + e_\rho \right) (\delta\rho)^2 + (\rho e_{\rho\eta} + e_\eta)\delta\rho\delta\eta + \frac{\rho}{2}e_{\eta\eta}(\delta\eta)^2 \right] \right\rangle. \end{aligned}$$

Using (5.7) and (5.8), we have

$$(5.12) \quad \Delta^{(2)}E_{(II)} = \left\langle [C_P T \delta\rho + \rho T \delta\eta] + \left[\frac{c_s^2}{2\rho} (\delta\rho)^2 + \frac{C_P T}{C_V} \delta\rho\delta\eta + \frac{\rho T}{2C_V} (\delta\eta)^2 \right] \right\rangle,$$

and adding (5.10) and (5.12), we finally obtain

$$(5.13) \quad \begin{aligned} \Delta^{(2)}E &= \left\langle (B\delta\rho + \rho\vec{v} \cdot \delta\vec{v} + \rho T \delta\eta) \right. \\ &\quad \left. + \left[\frac{\rho}{2}\delta\vec{v} \cdot \delta\vec{v} + \delta\rho\vec{v} \cdot \delta\vec{v} + \frac{(c_s)^2}{2\rho} (\delta\rho)^2 + \frac{C_P T}{C_V} \delta\rho\delta\eta + \frac{\rho T}{2C_V} (\delta\eta)^2 \right] \right\rangle. \end{aligned}$$

(II) *Calculation of the first and second variations of C_F given in (2.19).* The “formal” variation of C_F denoted by ΔC_F for which $\Delta^{(2)}q$ in (2.22) is temporarily considered to be an independent variable becomes

$$(5.14) \quad \begin{aligned} \Delta C_F &\equiv \langle (\rho + \delta\rho)F(\eta + \delta\eta, q + \Delta^{(2)}q) - \rho F(\eta, q) \rangle \\ &= \langle (\rho + \delta\rho) \left[F + F_\eta\delta\eta + F_q(\Delta^{(2)}q) + \frac{1}{2}F_{\eta\eta}(\delta\eta)^2 + F_{\eta q}\delta\eta(\Delta^{(2)}q) \right. \right. \\ &\quad \left. \left. + \frac{1}{2}F_{qq}(\Delta^{(2)}q)^2 + \dots \right] - \rho F(\eta, q) \right\rangle \\ &= \left\langle \rho F_\eta\delta\eta + F\delta\rho + \rho F_q(\Delta^{(2)}q) + \frac{\rho}{2}F_{\eta\eta}(\delta\eta)^2 + \rho F_{\eta q}\delta\eta(\Delta^{(2)}q) + \frac{\rho}{2}F_{qq}(\Delta^{(2)}q)^2 \right. \\ &\quad \left. + F_\eta\delta\eta\delta\rho + F_q\delta\rho(\Delta^{(2)}q) \right\rangle \end{aligned}$$

$$\begin{aligned}
 &= \left\langle \rho F_\eta \delta \eta + F \delta \rho + \rho F_q \left(\delta q - \delta q \frac{\delta \rho}{\rho} + \frac{1}{\rho} \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right) + \frac{\rho}{2} F_{\eta\eta} (\delta \eta)^2 + F_\eta \delta \eta \delta \rho \right. \\
 &\quad + \rho F_{\eta q} \delta \eta \left(\delta q - \delta q \frac{\delta \rho}{\rho} + \frac{1}{\rho} \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right) + F_q \delta \rho \left(\delta q - \delta q \frac{\delta \rho}{\rho} + \frac{1}{\rho} \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right) \\
 &\quad \left. + \frac{\rho}{2} F_{qq} \left(\delta q - \delta q \frac{\delta \rho}{\rho} + \frac{1}{\rho} \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right)^2 \right\rangle.
 \end{aligned}$$

Hence, up to the second-order accuracy, we have

$$\begin{aligned}
 (5.15) \quad \Delta^{(2)} C_F &= \left\langle \rho F_\eta \delta \eta + F \delta \rho + \rho F_q \left(\delta q - \delta q \frac{\delta \rho}{\rho} + \frac{1}{\rho} \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right) \right. \\
 &\quad \left. + \frac{\rho}{2} F_{\eta\eta} (\delta \eta)^2 + F_\eta \delta \eta \delta \rho + \rho F_{\eta q} \delta \eta \delta q + F_q \delta \rho \delta q + \frac{\rho}{2} F_{qq} (\delta q)^2 \right\rangle.
 \end{aligned}$$

If we apply the definition of δq given in (2.22) to the first-order term only, the above becomes

$$\begin{aligned}
 (5.16) \quad \Delta^{(2)} C_F &= \left\langle \rho F_\eta \delta \eta + F \delta \rho + F_q (\vec{\omega}_a \cdot \vec{\nabla} \delta \eta + \vec{\nabla} \eta \cdot \delta \vec{\omega} - q \delta \rho) \right. \\
 &\quad \left. + \frac{\rho}{2} F_{\eta\eta} (\delta \eta)^2 + F_\eta \delta \eta \delta \rho + \rho F_{\eta q} \delta \eta \delta q + \frac{\rho}{2} F_{qq} (\delta q)^2 + F_q \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right\rangle.
 \end{aligned}$$

After using the vector identities

$$(5.17) \quad F_q (\vec{\omega}_a \cdot \vec{\nabla} \delta \eta) = F_q \vec{\nabla} \cdot (\vec{\omega}_a \delta \eta) = -(\vec{\omega}_a \cdot \vec{\nabla} F_q) \delta \eta + \vec{\nabla} \cdot [F_q \vec{\omega}_a \delta \eta],$$

$$\begin{aligned}
 (5.18) \quad F_q (\vec{\nabla} \eta \cdot \delta \vec{\omega}) &= F_q \vec{\nabla} \cdot (\delta \vec{v} \times \vec{\nabla} \eta) = -(\delta \vec{v} \times \vec{\nabla} \eta) \cdot \vec{\nabla} F_q + \vec{\nabla} \cdot [F_q (\delta \vec{v} \times \vec{\nabla} \eta)] \\
 &= -(\vec{\nabla} \eta \times \vec{\nabla} F_q) \cdot \delta \vec{v} + \vec{\nabla} \cdot [F_q (\delta \vec{v} \times \vec{\nabla} \eta)],
 \end{aligned}$$

equation (5.16) becomes

$$\begin{aligned}
 (5.19) \quad \Delta^{(2)} C_F &= \left\langle [(F - q F_q) \delta \rho + (\rho F_\eta - \vec{\omega}_a \cdot \vec{\nabla} F_q) \delta \eta - (\vec{\nabla} \eta \times \vec{\nabla} F_q) \cdot \delta \vec{v}] \right. \\
 &\quad + \left[\frac{\rho}{2} F_{\eta\eta} (\delta \eta)^2 + F_\eta \delta \eta \delta \rho + \rho F_{\eta q} \delta \eta \delta q + \frac{\rho}{2} F_{qq} (\delta q)^2 + F_q \delta \vec{\omega} \cdot \vec{\nabla} \delta \eta \right] \\
 &\quad \left. + \vec{\nabla} \cdot [F_q (\delta \vec{v} \times \vec{\nabla} \eta + \vec{\omega}_a \delta \eta)] \right\rangle.
 \end{aligned}$$

Appendix B

For the (x -independent) zonally symmetric flow field with $U = U(y, z)$ and $V = W = 0$, the steady-state momentum equation (2.16) becomes

$$(5.20) \quad B_y + U(-U_y + f) = T \eta_y, \quad B_z - U U_z = T \eta_z,$$

where subscripts denote partial derivatives and f is the Coriolis parameter. (5.20) becomes slightly simpler with the introduction of $U^{(f)} \equiv U - \int_{y_0} f(y) dy$,

$$(5.21) \quad \begin{pmatrix} B_y \\ B_z \end{pmatrix} = U \begin{pmatrix} U_y^{(f)} \\ U_z^{(f)} \end{pmatrix} + T \begin{pmatrix} \eta_y \\ \eta_z \end{pmatrix}.$$

Multiplying (5.21) by $\begin{pmatrix} y_\eta & z_\eta \\ y_q & z_q \end{pmatrix}$, we obtain

$$(5.22) \quad \begin{pmatrix} B_\eta \\ B_q \end{pmatrix} = U \begin{pmatrix} U_\eta^{(f)} \\ U_q^{(f)} \end{pmatrix} + T \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

because $\eta_\eta = 1$ and $\eta_q = 0$.

Appendix C

(1) *Calculation of N_1 .* We have

$$(5.23) \quad \begin{aligned} N_1 &= \frac{C_V + R}{C_V} T - B_\eta + \frac{q}{2} B_{q\eta} \\ &= \frac{RT}{C_V} + (T - B_\eta) + \frac{q}{2} B_{q\eta} = \frac{RT}{C_V} - UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta, \end{aligned}$$

where the first and the second rows of (5.22) were used in deriving the second line.

(2) *Calculation of N_2 .* Using the identity $1/C_V = 1/C_P + (C_P - C_V)/C_P C_V = 1/C_P + R/C_P C_V$, we have

$$(5.24) \quad N_2 = \left(\frac{T}{C_P} - B_{\eta\eta} \right) + \frac{RT}{C_P C_V},$$

and the quantity in parentheses can be evaluated as follows. The balanced equation (5.20) can be rewritten in more familiar form of geostrophic and hydrostatic balances:

$$(5.25) \quad fU = -C_P T_y + T \eta_y, \quad -C_P T_z + T \eta_z - g = 0.$$

Applying the same transformation (5.21) \rightarrow (5.22), we have

$$(5.26) \quad \begin{pmatrix} C_P T_\eta \\ C_P T_q \end{pmatrix} = \begin{pmatrix} T - g z_\eta - U(f y_\eta) \\ -g z_q - U(f y_q) \end{pmatrix}.$$

Differentiating the first row of (5.22) with respect to η , we obtain

$$(5.27) \quad B_{\eta\eta} = T_\eta + (UU_\eta^{(f)})_\eta,$$

and substituting the first row of (5.26) into (5.27) yields

$$(5.28) \quad B_{\eta\eta} = \frac{1}{C_P} [T - gz_\eta - U(fy_\eta)] + (UU_\eta^{(f)})_\eta.$$

Using the definition of $U^{(f)}$, namely $fy_\eta = U_\eta - U_\eta^{(f)}$, we can rewrite (5.28) as

$$(5.29) \quad \frac{T}{C_P} - B_{\eta\eta} = \frac{1}{C_P} \left(gz + \frac{U^2}{2} \right)_\eta - \frac{1}{C_P} UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta.$$

Having derived N_1 and N_2 , we now expand the third condition in (4.1). Dividing (4.1) by $(c_s)^2$, we have

$$(5.30) \quad \begin{aligned} I_s &\equiv (1 - M^2) \left(\frac{T}{C_P} - B_{\eta\eta} + \frac{RT}{C_P C_V} \right) - \frac{1}{(c_s)^2} \left(\frac{RT}{C_V} - UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right)^2 \\ &= (1 - M^2) \left(\frac{T}{C_P} - B_{\eta\eta} \right) + \frac{RT}{C_P C_V} - M^2 \frac{RT}{C_P C_V} \\ &\quad - \frac{1}{(c_s)^2} \left[\frac{R^2 T^2}{C_V^2} + \frac{2RT}{C_V} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right) + \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right)^2 \right], \end{aligned}$$

where $M \equiv U/c_s$ is the Mach number. If we now apply $1/(c_s)^2 = C_V/(C_P RT)$ to the third, fourth and fifth terms in (5.30), I_s becomes

$$(5.31) \quad \begin{aligned} I_s &= (1 - M^2) \left(\frac{T}{C_P} - B_{\eta\eta} \right) + \frac{RT}{C_P C_V} - \frac{U^2}{C_P^2} \\ &\quad - \frac{RT}{C_P C_V} - \frac{2}{C_P} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right) - \frac{1}{(c_s)^2} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right)^2 \\ &= (1 - M^2) \left(\frac{T}{C_P} - B_{\eta\eta} \right) - \frac{U^2}{C_P^2} - \frac{2}{C_P} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right) \\ &\quad - \frac{1}{(c_s)^2} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right)^2. \end{aligned}$$

Substitution of (5.29) into (5.31) yields

$$(5.32) \quad \begin{aligned} I_s &= (1 - M^2) \left[\frac{1}{C_P} \left(gz + \frac{U^2}{2} \right)_\eta - \frac{1}{C_P} UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta \right] - \frac{U^2}{(C_P)^2} \\ &\quad - \frac{2}{C_P} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right) - \frac{1}{(c_s)^2} \left(-UU_\eta^{(f)} + \frac{q}{2} (UU_q^{(f)})_\eta \right)^2. \end{aligned}$$

The last term can be rewritten further as

$$\begin{aligned}
 (5.33) \quad & -\frac{1}{(c_s)^2}[-UU_\eta^{(f)} + \frac{q}{2}(UU_q^{(f)})_\eta]^2 \\
 & = -\frac{1}{(c_s)^2}\left[U\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right) + \frac{1}{2}U_\eta(qU_q^{(f)})\right]^2 \\
 & = -M^2\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)^2 - M\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)U_\eta\left(\frac{qU_q^{(f)}}{c_s}\right) - \frac{1}{4}U_\eta^2\left(\frac{qU_q^{(f)}}{c_s}\right)^2.
 \end{aligned}$$

As the dimension of $qU_q^{(f)}$ is the same as that of U , if we introduce the pseudo-Mach number M_p defined as $M_p \equiv qU_q^{(f)}/c_s$, then the right-hand side of (5.33) reads

$$(5.34) \quad -M^2\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)^2 - M\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)U_\eta M_p - \frac{1}{4}U_\eta^2 M_p^2.$$

On using this expression, I_s finally becomes

$$\begin{aligned}
 (5.35) \quad I_s & = \frac{1}{C_P}\left(gz + \frac{U^2}{2}\right)_\eta + \frac{1}{C_P}UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta - \frac{q}{C_P}(UU_q^{(f)})_\eta - \frac{U^2}{C_P^2} \\
 & \quad - M^2\left[\frac{1}{C_P}\left(gz + \frac{U^2}{2}\right)_\eta - \frac{1}{C_P}UU_\eta^{(f)} - (UU_\eta^{(f)})_\eta\right] \\
 & \quad - M^2\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)^2 - MM_p\left(-U_\eta^{(f)} + \frac{q}{2}U_{q\eta}^{(f)}\right)U_\eta - \frac{1}{4}U_\eta^2 M_p^2.
 \end{aligned}$$

Acknowledgements

We thank Profs. D. D. Holm and M. V. Kurgansky for their encouraging and useful comments. Our appreciation is extended to the anonymous reviewers whose comments were quite helpful in improving the quality of our paper. Y. F. was supported in part by the Joint Research Project between the Japan Society of the Promotion of Science and the Royal Society of London.

References

- [1] H. D. I. Abarbanel, D. D. Holm, J. E. Marsden and T. S. Ratiu, Richardson number criterion for the nonlinear stability of three-dimensional stratified flow, *Phys. Rev. Lett.* **52** (1984), 2352–2355. [MR 0748636](#)
- [2] ———, Nonlinear stability analysis of stratified fluid equilibria, *Philos. Trans. Roy. Soc. London Ser. A* **318** (1986), 349–409. [Zbl 0637.76119](#) [MR 0854437](#)
- [3] V. I. Arnol'd, Conditions for nonlinear stability of stationary plane curvilinear flows of an ideal fluid, *Dokl. Akad. Nauk SSSR* **162** (1965), 975–978 (in Russian); English transl.: *Soviet Math. Dokl.* **6** (1965), 773–777. [Zbl 0141.43901](#) [MR 0180051](#)

- [4] ———, On an a priori estimate in the theory of hydrodynamical stability, *Izv. Vyssh. Uchebn. Zaved. Mat.* **54** (1966), 3–5 (in Russian); English transl.: *Amer. Math. Soc. Transl.* (2) **79** (1969), 267–269.
- [5] J. M. Ball and J. E. Marsden, Quasiconvexity at the boundary, positivity of the second variation and elastic stability, *Arch. Ration. Mech. Anal.* **86** (1984) 251–277. [Zbl 0552.73006](#) [MR 0751509](#)
- [6] J. G. Charney, *Planetary fluid dynamics*, Reidel, 1973. [Zbl 0297.76036](#)
- [7] J. G. Charney and M. E. Stern, On the stability of internal baroclinic jets in a rotating atmosphere, *J. Atmospheric Sci.* **19** (1962), 159–172.
- [8] F. V. Dolzhansky, *Fundamentals of geophysical hydrodynamics*, Springer, 2013. [Zbl 1286.86001](#) [MR 2987817](#)
- [9] D. D. Holm, B. A. Kuperschmidt and C. D. Levermore, Canonical maps between Poisson brackets in Eulerian and Lagrangian descriptions of continuum mechanics, *Phys. Lett. A* **98** (1983), 389–395. [MR 0726505](#)
- [10] D. D. Holm, J. E. Marsden, T. S. Ratiu and A. Weinstein, Nonlinear stability of fluid and plasma equilibria, *Phys. Rep.* **123** (1985), 1–116. [Zbl 0717.76051](#) [MR 0794110](#)
- [11] J. R. Holton, *An introduction to dynamic meteorology*, 4th ed., Elsevier, 2004.
- [12] L. N. Howard, Note on a paper of John W. Miles, *J. Fluid Mech.* **10** (1961), 509–512. [Zbl 0104.20704](#) [MR 0128233](#)
- [13] Y. Kuroda, Symmetries and Casimir invariants for perfect fluid, *Fluid Dynam. Res.* **5** (1990), 273–287.
- [14] R. G. Littlejohn, Singular Poisson tensors, in *Mathematical methods in hydrodynamics and integrability in dynamical systems* (La Jolla, CA, 1981), AIP Conf. Proc. 88, Amer. Inst. Phys., 1982, 47–66. [Zbl 0537.70014](#) [MR 0695502](#)
- [15] A. Mahalov, M. Moustaoui and B. Nicolaenko, Three-dimensional instabilities in non-parallel shear stratified flows, *Kinetic Related Models* **2** (2009), 215–229. [Zbl 05643198](#) [MR 2472157](#)
- [16] J. E. Marsden and P. J. Morrison, Noncanonical Hamiltonian field theory and reduced MHD, in *Fluids and plasmas: geometry and dynamics* (Boulder, CO, 1983), *Contemp. Math.* 28, Amer. Math. Soc., 1984, 133–150. [Zbl 0546.58024](#) [MR 0751979](#)
- [17] J. W. Miles, On the stability of heterogeneous shear flow, *J. Fluid Mech.* **10** (1961), 496–508. [Zbl 0101.43002](#) [MR 0128232](#)
- [18] P. J. Morrison and J. M. Greene, Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics, *Phys. Rev. Lett.* **45** (1980), 790–794. [MR 0581650](#)
- [19] ———, Noncanonical Hamiltonian density formulation of hydrodynamics and ideal magnetohydrodynamics addition, *Phys. Rev. Lett.* **48** (1982), 569.
- [20] P. J. Olver, *Applications of Lie groups to differential equations*, Springer, 1986. [Zbl 0588.22001](#) [MR 0836734](#)
- [21] H. Sakuma and M. Ghil, Stability of propagating modons for small-amplitude perturbations, *Phys. Fluids A* **3** (1991), 408–414. [Zbl 0850.76239](#) [MR 1093586](#)
- [22] R. Salmon, *Lectures on geophysical fluid dynamics*, Oxford Univ. Press, 1998. [MR 1718369](#)
- [23] E. C. G. Sudarshan and N. Mukunda, *Classical dynamics: A modern perspective*, Interscience, New York, 1974. [Zbl 0329.70001](#) [MR 0434047](#)
- [24] G. K. Vallis, G. F. Carnevale and W. Young, Extremal energy properties and construction of stable solutions of the Euler equations, *J. Fluid Mech.* **207** (1989), 133–152. [MR 1023732](#)