Isaac Newton Institute Preprint NI96014

Legendre-Transformable Semi-Geostrophic Theories

R. J. Purser

General Sciences Corporation, Laurel, Maryland

Manuscript Submitted ???

Corresponding author address: Dr. R.J. Purser, National Centers for Environmental Prediction, W/NP2 WWB Room 207, Washington, D.C. 20233

ABSTRACT

For semi-geostrophic (SG) theories derived from the Hamiltonian principles suggested by Salmon it is known that a duality exists between the physical coordinates and geopotential, on the one hand, and isentropic geostrophic momentum coordinates and geostrophic Bernoulli function, on the other hand. The duality is characterized geometrically by a "contact structure". This enables the idealized balanced dynamics to be represented by horizontal geostrophic motion in the dual coordinates while the mapping back to physical space is determined uniquely by requiring each instantaneous state to be the one of minimum energy with respect to volume-conserving rearrangements within the physical domain.

It is found that the generic contact structure permits the emergence of topological anomalies during the evolution of discontinuous flows. For both theoretical and computational reasons it is desirable to seek special forms of SG dynamics in which the structure of the contact geometry prohibits such anomalies. We prove that this desideratum is equivalent to the existence of a mapping of geographical position to a Euclidean domain, combined with some position-dependent additive modification of the geopotential, which results in the SG theory being manifestly Legendre-transformable from this alternative representation to its associated dual variables.

Legendre transformable representations for standard Boussinesq f-plane SG theory and for the axisymmetric gradient-balance version used to study the Eliassen vortex are already known and exploited in finite element algorithms. Here, we reexamine two other potentially useful classes of SG theory discussed in a recent paper by the author: (i) the *non*-axisymmetric f-plane vortex; (ii) hemispheric (variable-f) SG dynamics. We find that the imposition of the natural dynamical and geometrical symmetry requirements together with the requirement of Legendre-transformability makes the choice of the f-plane vortex theory unique. Moreover, with modifications to accommodate sphericity, this special vortex theory supplies what appears to be the most symmetrical and

consistent formulation of variable-f SG theory on the hemisphere. The Legendre-transformable representations of these theories appear superficially to violate the original symmetry of rotation about the vortex axis. But, remarkably, this symmetry *is* preserved provided we interpret the metric of the new representation to be a pseudo-Euclidean "Minkowski" metric. Rotation-invariance of the dynamical formulation in physical space becomes a formal "Lorentz-invariance" of the dynamics in its Legendre-transformable representation.

---

1. Introduction

In order to further our understanding of atmospheric and oceanic dynamics it is desirable to possess a set of idealized and simplified equations whose solutions can be obtained with great precision (either numerically or analytically) while realistically treating the dynamical features of interest. In this way the idealized system can provide insights into the *essential* balanced dynamics which a more complete model might often obscure with numerical or gravity-wave "noise". When it comes to the study of discontinuous phenomena, such as atmospheric and oceanic fronts, then the semi-geostrophic (SG) systems (Hoskins and Bretherton ,1972; Hoskins 1975), which extended the earlier studies of vortex and frontal balance of Eliassen (1951, 1962), seem uniquely suited to the study of the "slow" or "balanced" components of these discontinuities. Being filtered systems of equations, they automatically exclude the obscuring gravity wave components, yet, in their Lagrangian form, they are expressible as a set of parcel-conservation laws requiring no evaluations of spatial derivative. As first pointed out by Cullen (1983), they are therefore able to tolerate contact discontinuities within the fluid and can be integrated using fully-Lagrangian finite elements.

The theory of the Lagrangian finite element "geometric model" form of SG dynamics was put on a firm foundation by Cullen and Purser (1984) and the method was applied to a variety of highly idealized f-plane situations by Cullen et al. (1987a, 1987b), Shutts (1987) and by Chynoweth (1987) who constructed the first general geometric model algorithms. Shutts et al. (1988) were also able to demonstrate the applicability of the geometric model, suitably modified, to Eliassen's axisymmetric balanced vortex in which gradient balance replaces geostrophic balance radially.

Independently, Salmon (1983, 1985), with mainly ocean simulations as his objective, rediscovered and greatly extended SG dynamics by starting with a Hamiltonian (variational) prescription. The filtering assumptions were introduced in a careful way that ensured retention of analogs of the important conservation laws of mass, energy and potential vorticity. The most significant feature of this alternative derivation was that it admitted a general spatial variation of the Coriolis parameter which, hitherto, had not been accomplished without violating one or more of the conservation principles. A global SG model based on similar pronciples was proposed by Shutts (1980). The existence of a contact structure in SG theory was first recognized by Blumen (1982). In an attempt to illuminate the common geometrical features of Salmon's Hamiltonian model and the conventional SG models Purser (1993) (henceforth P93) paid particular attention to the contact structure and showed how the SG models could be extended in a way consistent with both the Hamiltonian and geometric model formulations to non-axisymmetric vortex dynamics on either an f-plane or, with proper treatment of the varying Coriolis parameter and sphericity, to the hemisphere. A brief description of contact structure as it relates to Hamiltonian SG dynamics is given in section 2. For further mathematical details, the reader may consult Sewell and Roulstone (1994). Hamiltonian techniques in geophysical fluids are reviewed in Shepherd (1990). A recent interesting development, generalizing the SG models, is provided in McIntyre and Roulstone (1996).

Associated with the contact structure is a duality relation between the physical solution (comprising a geopotential function of physical coordinates) and the "dual solution" (comprising a Bernoulli function of isentropic geostrophic momentum coordinates, or a variant of these). The "graph" (hypersurface in the extended space of physical coordinates $\mathbf{x}$ augmented by an extra coordinate measuring geopotential ϕ) of the physical solution is, in a well defined sense, the envelope of a continuous family of what we call "neutral energy" (generating) surfaces coexisting in the same extended space. Each neutral energy surface is itself labeled by a dual coordinate $\mathbf{X}$ and the dual (Bernoulli) potential Φ, and therefore the entire surface can be identified by a single *point* in a dual extended space. The term, "duality" recognizes the fact that roles can be reversed. Thus, from the totality of neutral energy surfaces one can consider the subset which pass through a given point $\mathbf{x}$ and geopotential ϕ and notice that the locus of their "labels" $\mathbf{X}$ and Φ now constitutes a single new "generating surface", in extended dual space, whose own labels can be taken to be $(\mathbf{x},\phi)$. Now, just as the physical solution can be regarded as the envelope "above" (when the sense of increasing ϕ is taken as "up") the neutral energy surfaces labeled by the dual quantities $(\mathbf{X}, \Phi)$, so the dual solution can itself be regarded as the envelope "below" those dual generating surfaces labeled by the quantities $(\mathbf{x}, \phi)$ present in the original physical solution. In this interpretation, we find that the class of finite element solutions are simply those solutions $\phi(\mathbf{x})$ whose graphs are each constructed as envelopes of only *finitely* many neutral energy surfaces. Each element is therefore characterized by a physical volume (which it conserves), a single value of $\mathbf{X}$ and a single value of Φ.

Cullen and Purser (1984) showed that a trivial transformation of the geopotential of Hoskins' (1975) constant Coriolis form of SG theory (the "standard theory") enabled this version to be treated by the methods of Legendre duality. Based on a less obvious mapping to Legendre-transformable form, Shutts et al. (1988) extended the geometrical model to the SG form of the axisymmetric Eliassen vortex model described by Shutts and Thorpe (1978) and by Schubert and Hack (1983). Geometrical implications of Legendre duality were discussed by Purser and Cullen (1987), Chynoweth et al.(1988) and by Chynoweth and Sewell (1989, 1991). Primarily, the practical significance of Legendre-transformability is that it leads to the simplest algorithms for the finite element solutions (computations of intersections among curved surfaces is incomparably harder than the equivalent computations for intersecting hyperplanes). However, as we show in section 3, there is another reason for preferring a Legendre-transformable SG theory that is related to the topological structure of the connections between distinct fluid parcels in finite element solutions or in originally smooth solutions during frontal formation. In theories possessing a generic (not Legendre-transformable) contact structure it can be shown that topological "anomalies" can occur whereby elements may either spontaneously split into disconnected pieces, or else may achieve multiple contact (at disconnected interfaces) with the same neighboring elements. Apart from the obvious computational difficulties implied for finite element algorithms, the possibility of such anomalies makes the uniqueness and regularity of the "weak" (discontinuous) solutions resulting from initially continuous data questionable. This possibility thus undermines the supreme purported virtue of SG dynamics – its ability to accommodate discontinuous solutions. We therefore regard any SG theory *not* possessing a Legendre-transformable representation as structurally deficient.

Two such questionable theories were proposed in P93; one, a non-axisymmetric generalization of the f-plane vortex model; the other a natural extension of this model to the hemisphere (see Craig, 1991, and Magnusdottir and Schubert, 1991, for alternative SG treatments not obviously exhibiting the form of contact structure we have described). In sections 4 and 5, we re-examine the necessary properties of the contact structures of the Hamiltonian SG theories introduced in P93 and propose very minor modifications to the particular formulations suggested there in order to make the modified formulations exactly Legendre-transformable. The Legendre transformable representations of the vortex models appear superficially not to preserve the angular symmetry since the concentric circles of the vortex are mapped to sectors of concentric hyperbolae in each transformed horizontal plane. The symmetry-breaking is illusory, however; the new representation *does* preserve the symmetry — provided the space of points $\hat{\mathbf{x}}$ of the Legendre-transformable representation is regarded as being furnished with the pseudo-Euclidean metric of a "Minkowski" space in place of a true Euclidean metric. The operation of rotation by some angle about the axis of the vortex is then represented by a proportionate "Lorentz boost", with cosines and sines of azimuth in the components of the rotation operator being replaced by hyperbolic-cosines and sines of the corresponding Lorentz angles. In section 6 we make some remarks concerning the numerical implementation of the proposed new formulations and we conclude in section 7 with some more general suggestions about the potential applications, both in numerical weather prediction and in oceanography, of the type of methods we are advocating.

2. Contact structure in SG theories and Hamiltonian dynamics

We shall adopt most of the notational convention of P93. Thus, $\mathbf{x}$ and $\mathbf{X}$ are physical and dual coordinates, μ and ν are physical and dual measures of the respective coordinate volumes, ϕ and Φ are physical and dual geopotentials, η is the pseudo-density (mass per unit μ) and ρ the potential density (mass per unit ν). One essential feature of SG theories is that the dynamics is specified by the distribution of geopotential (and boundary constraints) alone. Therefore, the Hamiltonian is expressible in terms only of the geopotential distribution.

Let us define the specific energy of a parcel with physical and dual coordinates $\mathbf{x}$ and $\mathbf{X}$ to be $\mathcal{E}(\mathbf{x}, \mathbf{X})$. In general, we need not require the domain of $\mathbf{X}$ to be isometric to the domain of $\mathbf{x}$; as we shall see in section 5, it is sometimes more appropriate that the dual space differs from the physical space (for example, to enable the dual coordinates to be made formally "canonical").

We postulate that, at each instant, the collective disposition of the $\mathbf{x}$ associated with each $\mathbf{X}$ is such that the energy integral,

$$\mathcal{H} = \int \mathcal{E}(\mathbf{x}, \mathbf{X})\, \rho\, d\nu, \tag{2.1}$$

with respect to local rearrangements that conserve their $\mathbf{X}$ and mass ρ on material parcels, and that conserve pseudo-density η and stay within the domain in physical space, is minimized. The valid solution is then one associated with a scalar function ϕ that we identify as the geopotential and which satisfies:

$$\phi(\mathbf{x}) \; = \sup \;\; {\phi'}_{\mathbf{X}}(\mathbf{x}), \tag{2.2}$$

where, for each suffix $\mathbf{X}$, $\phi'_{\mathbf{X}}(\mathbf{x})$ denotes the "neutral energy" function (or its graph, the "neutral energy surface"):

$$\phi'_{\mathbf{X}}(\mathbf{x}) = \Phi(\mathbf{X}) - \mathcal{E}(\mathbf{x}, \mathbf{X}). \tag{2.3}$$

If the solution ϕ follows a neutral energy surface throughout some finite volume, then according to the precepts of SG theory, a prompt $\mathbf{X}$-conserving lateral or vertical displacement of any constituent parcel can be achieved with a net change in the total energy of the system. This idea of a neutral energy surface therefore serves to extend the one-dimensional concept of a neutral stratification to the horizontal dimensions also. In the same way that a vertical stable stratification is convex relative to the neutral profiles tangent to it, a symmetrically stable distribution of ϕ is (three-dimensionally) convex relative to the neutral energy surfaces tangent to it.

The dual potentials $\Phi(\mathbf{X})$ are defined implicitly to be those such that, for each set Σ of $\mathbf{X}$–space of measure ν, the corresponding set σ of $\mathbf{x}$–space of measure μ (conserved by potential rearrangements) is obtained as the volume of actual contact:

$$\mu_\sigma = \int_\sigma d\mu, \tag{2.4a}$$

$$\sigma = \{\mathbf{x} : \exists\ \mathbf{X}{\in}\Sigma,\ \phi(\mathbf{x}) = \Phi(\mathbf{X}) - \mathcal{E}(\mathbf{x}, \mathbf{X})\}. \tag{2.4b}$$

A more complete discussion of this idea is presented in P93, where it is shown that this prescription provides a definition for the theory's inherent "contact structure" and determines the basis for the geometrical duality between the physical solution $\phi(\mathbf{x})$ and the dual solution $\Phi(\mathbf{X})$. Properties of this contact structure are:

- $\mathbf{X}$ is regarded as a function jointly of $\mathbf{x}$ and $\nabla\phi$, and therefore $\mathcal{E}(\mathbf{x}, \mathbf{X})$ is also.
- If different solutions ϕ_1 and ϕ_2 make tangential contact at $\mathbf{x}$, then their duals, Φ_1 and Φ_2 make tangential contact at the image, $\mathbf{X}$, of $\mathbf{x}$.

The quantity $\mathcal{H}$ of (2.1) is the Hamiltonian, which directly prescribes the evolution of the flow in $\mathbf{X}$–space and, indirectly (from the rearrangement result), the flow in physical space also. The variational method for these problems is developed by Salmon (1983, 1985). During the period $[t_1, t_2]$ an action-integral is extremized:

$$\delta \int_{t_1}^{t_2} \mathcal{L}_s \, dt = 0, \tag{2.5}$$

where the Lagrangian $\mathcal{L}_s$ is defined:

$$\mathcal{L}_s = \int \mathbf{A} \cdot \frac{D\mathbf{X}}{Dt} \rho \, d\nu - \mathcal{H} \tag{2.6}$$

The horizontal vector field $\mathbf{A}$ is a time-independent function of the horizontal dual coordinates $\mathbf{X}$ such that its integral C:

$$C = \oint \mathbf{A} \cdot d\mathbf{X}, \tag{2.7}$$

in a circuit of constant Z measures an absolute circulation associated with the effective Coriolis function,

$$f^*(\mathbf{X}) = \frac{\partial A_y}{\partial X} - \frac{\partial A_x}{\partial Y}. \tag{2.8}$$

Variations of the action integral with respect to X and Y, subject to the constraint that parcel values of Z and mass remain constant, imply "geostrophic" dynamics in $\mathbf{X}$–space:

$$f^* \frac{D\mathbf{X}}{Dt} = \left(-\frac{\partial \mathcal{H}}{\partial Y}, \frac{\partial \mathcal{H}}{\partial Y}, 0 \right)^T . \tag{2.9}$$

Any circulation integral C defined by (2.9) is now a materially conserved quantity. We note that a transformation of dual coordinates, $\mathbf{X} \longrightarrow \mathbf{X}'$ accompanied by a circulation-preserving re-definition of the effective Coriolis function,

$$f^* \longrightarrow f'^* \equiv f^* \frac{\partial(X, Y)}{\partial(X', Y')}, \tag{2.10}$$

leaves the form of (2.9) unchanged. As discussed by Roulstone and Sewell (1996), this enables a choice for X' and Y' to be made such that the new effective Coriolis function f'^* is constant, whereupon, the X' and Y' of each material parcel become "canonical coordinates" of the Hamiltonian description of the dynamics.

3. Legendre-duality

As discussed in Schubert (1985) and P93, it is generally possible in SG theory to express the dynamics for ϕ or its dual, Φ, in terms of some linear elliptic "tendency equation" and it is tempting therefore to think that standard numerical methods, involving some form of "relaxation" procedure, will automatically supply a practical way to integrate the time dependent solutions of interest. However, very frequently the solutions of primary interest in SG studies are of a singular character, such as those describing fronts. Here, the standard grid-point methods, which rely heavily on the use of spatial differencing, can become severely compromised by the numerical difficulties associated with evaluating derivatives near the modelled discontinuities or by the spontaneous emergence of perceived non-elliptic regions at these places.

As noted in the introduction, an alternative numerical procedure designed specifically to handle these otherwise intractable problems in SG theory is the Lagrangian finite-element "geometric method" proposed by Cullen (1983) and further elaborated by Cullen and Purser (1984), Cullen et al. (1987a, b), Chynoweth (1987) and Chynoweth et al. (1988). The finite elements of this method each conserve their mass and (in adiabatic dynamics, at least) a value of potential temperature that is assumed uniform throughout the element. The horizontal dual coordinates X and Y are also assumed uniform throughout the element, but subject to change in time according to the dynamics implied by the Hamiltonian, which is evaluated by summing the contributions from each polyhedral element. The total energy associated with each element can be computed as an expression involving the "moments" of that element and simple functions of its $\mathbf{X}$, but at no time is it necessary to evaluate spatial derivatives. As a consequence, the finite element mode of computation stands aloof from the kinds of numerical problems associated with frontal discontinuities that seriously beset other methods of calculation.

In principle, the finite element methods should apply to any SG theory possessing the contact structure described in the previous section. However, as a practical matter, actual implementations of the geometric method have been restricted to the special class of SG dynamics for which a representation (possibly via a nontrivial spatial mapping) exists in which the neutral energy generating surfaces become hyperplanes in the "extended" physical space $(\hat{\mathbf{x}}, \hat{\phi})$ of this representation. Only in this case do the geometrical calculations involving the surfaces, edges and vertices of intersections among the various generating surfaces become sufficiently simple to be feasible. The Boussinesq standard f-plane SG theories in two and three dimensions have simple Legendre-transformable representations, as exploited by Cullen and Purser (1984) and discussed in detail in Purser and Cullen (1987). Also, Shutts et al. (1988) discovered that the axisymmetric (two-dimensional) variant of SG theory on the f-plane (Shutts and Thorpe, 1978), in which the radial component is balanced in the "gradient" sense proposed by Eliassen (Eliassen and Kleinschmidt, 1957), possesses a Legendre-transformable representation once the radial coordinate of the vortex has been suitably mapped. This enables the geometric method to be applied to the investigation of thermally forced solutions in the context of idealized axisymmetric tropical cyclones.

Other potentially useful extensions of SG theory have been formulated, but they do not possess obvious Legendre-transformable representations. These include various f-plane and hemispheric non-axisymmetric (three-dimensional) generalizations of vortex models (Craig, 1991; Magnusdottir and Schubert, 1991; P93) and the variable-f form of Salmon's (1985) $\mathcal{L}_s$–dynamics. While it is obviously desirable, from the computational point of view, to find Legendre-transformable variants of these non-axisymmetric vortex and variable-f theories, we claim that such SG variants are to be preferred also on theoretical grounds. We base this assertion on the following observations concerning the possible forms of finite element solutions (or other singular solutions, such as fronts, which can occur spontaneously from initially smooth data).

The intersections, in a horizontal surface at some fixed elevation, of the neutral energy surfaces associated with two neighboring dual-space labels, $\mathbf{X}_1$ and $\mathbf{X}_2$, constitute a family of non-intersecting curves (contours at this horizontal surface of the difference of their respective energy functions $\mathcal{E}$) that cover the area in physical space where both of these particular generating surfaces come into play. If we select one such curve say $S_{1,2}$, together with some particular point $\mathbf{x}$ belonging to it, then we can generally find a third neighboring dual-space label, say $\mathbf{X}_3$, for which the corresponding family of curves formed by all possible intersections, at this same elevation, of neutral energy surfaces ${\phi'}_2$ and ${\phi'}_3$ includes one member, the curve $S_{2,3}$, which is tangent to $S_{1,2}$ at $\mathbf{x}$ but which fails to coincide elsewhere in the immediate neighborhood of $\mathbf{x}$. Note, however, that tangency without coincidence of the two curves becomes impossible whenever the contact structure is transformable into one in which the duality takes the special Legendre form in which the inersecting surfaces are all planes. Assuming the label order $(1, 2, 3)$ is monotonic in the sense of the gradients of their respective neutral energy surfaces at $\mathbf{x}$, then the two generic possibilities for the general contact structure are:

- (i) the curves of intersection $S_{1,2}$ and $S_{2,3}$ curl outward at $\mathbf{x}$ leaving a pinched off "bowtie"-shaped portion of ${\phi'}_2$ able to form part of the solution surface ϕ, but now in two virtually separate pieces (schematically depicted in figute 1a);
- (ii) the curves curl inward at $\mathbf{x}$ so as to exclude the exposure [under the *sup*–operation of (2.2)] of any finite portion of fluid element–2 beyond the single locus of contact, $\mathbf{x}$, itself.

In the former case, the theory permits the spontaneous destruction of the integrity of finite elements. In frontal formation, it would seem to allow the impulsive distribution of potential vorticity associated with the resulting contact-discontinuity to be negative, and introduce some undesirable problems associated with guaranteeing uniqueness of solutions with respect to energy-minimizing local rearrangements of fluid elements along such a front. In the latter case, the formal difficulties are less severe but involve such anomalies as one "crescent"-shaped finite element being completely surrounded by only two of its neighbors (illustrated in figure 1b). The two outer elements in contact would then share an interface, possibly in many disconnected segments, on which there is also an impulsive distribution of potential vorticity now of positive sign. These theoretical complications are better avoided when the freedom in constructing generalized SG formulations allows one to do so. A result of some help in seeking such anomaly-free formulations is summarized in the following result (which is roughly analogous to the theorem of Darboux in symplectic geometry discussed in Arnold, 1980).

a. Theorem 1

In an arbitrarily differentiable semi-geostrophic contact structure, *bowtie* and *crescent* anomalies are impossible if and only if the semi-geostrophic solutions possess a representation which, in each neighborhood, is Legendre-transformable.

b. Remarks

A proof of this result is provided in the appendix. The virtue of this result is that it provides us with a criterion for judiciously modifying the existing non-axisymmetric vortex and variable-f SG theories in order to identify those special forms which are not only known to be anomaly-free, but which also promise the possession of Legendre-transformable representations (and hence, computationally feasible dynamics). The next section describes the minor modifications of the f-plane vortex theory of P93 required to render it Legendre-transformable, while section 5 extends this result to the corresponding hemispheric vortex model which, away from the pole, can be regarded essentially as a very minor modification of Salmon's variable-f $\mathcal{L}_s$–dynamics applied to the hemispheric domain.

4. The f-plane non-axisymmetric vortex

In this and the following section, we shall simplify the algebraic development by omitting the vertical dimension of the SG theories and therefore omit the associated potential energy contribution to the specific energy function $\mathcal{E}(\mathbf{x}, \mathbf{X})$ and to the

Hamiltonian. In every case, the potential energy contribution, $-Z\,z$, to the total remains unaltered by the various horizontal mappings we shall be considering.

First, we recall from P93 that, for the axisymmetric f-plane vortex at physical radius r from the axis and with "potential radius" R (to which the ring of fluid must be expanded or contracted conserving its angular momentum in order to bring it to rest in the rotating framework), the (kinetic component of the) energy function takes the form:

$$\mathcal{E}(r,R) = \frac{f^2\,R^4}{8r^2} + \frac{f^2\,r^2}{8} - \frac{f^2\,R^2}{4}. \tag{4.1}$$

If we relate such a vortex to an *unrotated* framework, then the specific energy is just the first term on the right, which is manifestly self-similar with respect to rescaling of either r or R. In P93 we argued that, in order to accommodate non-axisymmetric effects consistent with the appropriate (frame-relative) definition of "geostrophy" for first-order perturbations about any state of solid-body rotation, then it was necessary for this self-similarity to extend to the form of the energy function generalized in the azimuthal direction, and that the necessary geometrical constraint was that the Hessian of each neutral energy surface evaluated at vanishing relative azimuth ($\xi - \Xi = 0$) should have identical radial and tangential components, where ξ and Ξ are the physical and dual azimuth angles about the axis of the vortex. In section 6a of P93 we suggested one particular form of the new energy function, $\mathcal{E}(r,\xi;R,\Xi)$, satisfying this requirement. In the light of theorem 1 and the related discussion of section 3, it is worth reconsidering the exact choice for the azimuthal structure and seeking an alternative functional form for $\mathcal{E}$, equivalent to that proposed in P93 up to second-order in relative azimuth, but departing from that form at fourth-order in such a way as to avoid the occurrence of bowtie or crescent anomalies in the solution. As at least a necessary condition, we must find that the curves of intersection (at a horizontal level) of the desired neutral energy surfaces will collectively form a bi-parametric continuous family, just as the lines in a plane form such a family. If we write the energy function in a form that preserves the manifest self-similarity, with respect to radial rescaling, of the first right-hand terms:

$$\mathcal{E}(r,\xi;R,\Xi)=\frac{f^2\,R^4\,F(\xi-\Xi)}{8r^2}+\frac{f^2\,r^2}{8}-\frac{f^2\,R^2}{4} \tag{4.2}$$

for some yet to be determined function F, the isotropy of the Hessian at vanishing relative azimuth requires that F satisfies:

$$\begin{aligned} F(0)&=1,\\ F'(0)&=0,\\ F''(0)&=8. \end{aligned} \tag{4.3}$$

Now, since a limiting case of the Energy function obtained as $R\to 0$ gives,

$$\mathcal{E}_0(r,\xi)=\frac{f^2\,r^2}{8},$$

a necessary condition for obtaining the desired bi-parametric family of intersections is that F has the following property:

- Given arbitrary constants R_1, R_2, Ξ_1 and Ξ_2 then, except for sets in this four-dimensional parameter space of measure zero, a further pair, R_3 and Ξ_3, (possibly complex) can be found such that, for all ξ,

$$R_3F(\xi-\Xi_3)=R_1F(\xi-\Xi_1)-R_2F(\xi-\Xi_2). \tag{4.4}$$

The only solution of such a problem that also satisfies (4.3) is

$$F(\xi)=\cosh(\sqrt{8}\xi) \tag{4.5}$$

If we write $\hat{\xi}=(8)^{1/2}\xi$ and $\hat{\Xi}=(8)^{1/2}\Xi$, then we do indeed confirm that our choice for F leads to a Legendre-duality,

$$\hat{\phi}=\hat{\Phi}+\hat{\mathbf{x}}\cdot\hat{\mathbf{X}}, \tag{4.6}$$

in the following representation of the physical and dual variables:

$$\hat{x}=-\frac{1}{2f^2r^2}\cosh(\hat{\xi}), \tag{4.7a}$$

$$\hat{y}=+\frac{1}{2f^2r^2}\sinh(\hat{\xi}), \tag{4.7b}$$

$$\hat{\phi} = \phi + \frac{f^2\, r^2}{8}, \tag{4.7c}$$

$$\hat{X} = \frac{f^4\, R^4}{4} \cosh(\hat{\Xi}), \tag{4.7d}$$

$$\hat{Y} = \frac{f^4\, R^4}{4} \sinh(\hat{\Xi}), \tag{4.7e}$$

$$\hat{\Phi} = \Phi + \frac{f^2\, R^2}{4}. \tag{4.7f}$$

While superficially, it now seems that the angular symmetry in the original description of the dynamics has been destroyed by the intrusion of these *cosh* and *sinh* functions of azimuth, in fact, the underlying symmetry remains; in effect, the azimuth angles are subjected to a multiplicative scaling by a constant which happens to be the *imaginary* number, $(-8)^{1/2}$. Real values are recovered by recognizing the equivalence of such a scaling with a transformation from the horizontal Euclean plane to a two-dimensional "Minkowski" space, considered either to be a Euclidean space with one coordinate imaginary (the original application of this idea was to Special Relativity theory), or more conveniently, to be a space of real coordinates but with a *pseudo-Euclidean* metric,

$$d\hat{r}^2 = d\hat{x}^2 - d\hat{y}^2. \tag{4.8}$$

Then, the cyclic one-parameter group of axial rotations (generating displacements along circles) is now replaced by the one-parameter group of two-dimensional *Lorentz boosts* (generating displacements along hyperbolae). We note that, in the pseudo metric, the radial coordinate $\hat{r}$ is recovered from the components $\hat{x}$ and $\hat{y}$ by,

$$\hat{r}^2 = \hat{x}^2 - \hat{y}^2, \tag{4.9}$$

while scaled "Lorentz angles", $\hat{\xi}$ and $\hat{\Xi}$, are recovered using,

$$\hat{\xi} = -arctanh(\frac{\hat{y}}{\hat{x}}), \tag{4.10a}$$

$$\hat{\Xi} = +arctanh(\frac{\hat{Y}}{\hat{X}}). \tag{4.10b}$$

The group of Lorentz boosts is not cyclic. Therefore, transformed back into the original physical domain (x and y), the neutral energy surfaces formally wind repeatedly around the axis on a *Riemann surface*, but it is only portions possessing small relative azimuths, $|\xi - \Xi| \ll 0$ that will ever be of practical significance in constructing a solution.

It is instructive to see the shapes implied by this construction of the neutral energy and dual generating surfaces. Figure 2a shows some of the curves, passing through a fixed point, formed by intersections of pairs of generating surfaces. Note that some of these curves ("*cosh-type*") avoid the axis while others ("*sinh-type*") intersect it, according to their orientation relative to radial lines. (This topological distinction is exactly analogous to that in two-dimensional relativity theory between "*space-like*" and "*time-like*" lines). The picture corresponding to Fig. 2a for a focus of intersections at some other distance from the axis is essentially no different apart from a trivial change of scale.

Figure 2b shows uniformly spaced contours in physical space (x, y) of the "speed", u, that one would associate with the kinetic energy function, that is, $u = (2\mathcal{E})^{1/2}$, at the fixed $\mathbf{X}$ shown by the symbol. Note that near-circularity and even spacing of these contours at small amplitudes give way to distorted loops of progressively uneven spacing only when their scale becomes commensurate with the distance to the axis. Figure 2c shows corresponding contours plotted in the dual plane, (X, Y), when $\mathbf{x}$ is kept fixed. Radial cross-sections corresponding to figures 2b and 2c and at vanishing relative azimuth can be seen in figures 3a and 3b of P93.

5. Hemispheric SG theory

We continue to omit from our discussion the vertical components and associated potential energy, but we further simplify the algebra for the hemispheric development by choosing units of time and horizontal distance that make the polar value of the Coriolis parameter and the radius of the earth both unity. Thus, in these units, the Coriolis parameter at latitude λ is,

$$f(r) = (1 - r^2)^{1/2}, \tag{5.1}$$

where

$$r = cos(\lambda).$$

As discussed in P93, the neutral energy generating surfaces labeled by the fluid element's "potental radius", R, must possess a vanishing gradient and a horizontally isotropic Hessian of magnitude $-f^2$ at the location on the earth where $r = R$ and where physical and dual azimuth (or longitude) angles ξ and Ξ are the same. This ensures that, to first order, the SG dynamics reduces to geostrophy. It follows that the radial and tangential components of the Hessian of each equatorially-projected neutral energy function are, in our convenient units,

$$\frac{\partial^2 \phi'_R}{\partial r^2} = -1 \tag{5.2a}$$

$$\frac{1}{r^2}\frac{\partial^2 \phi'_R}{\partial \xi^2} = -(1 - r^2) \tag{5.2b}$$

at $r = R$. (The angle Λ for which $cos(\Lambda) = R$ is referred to as the "potential latitude" by Schubert and Hack, 1983).

In order to obtain a Legendre-transformable representation we seek a mapping of the physical coordinates and geopotential,

$$\hat{r} = \hat{r}(r), \tag{5.3a}$$

$$\hat{\xi} = \alpha\,\xi, \tag{5.3b}$$

$$\hat{\phi} = \phi + \Delta\phi(r), \tag{5.3c}$$

such that the corresponding neutral energy surfaces $\hat{\phi}'$ are linear functions of $\hat{x}$ and $\hat{y}$ defined by,

$$\hat{x} = \hat{r}\; cos(\hat{\xi}), \tag{5.4a}$$

$$\hat{y} = \hat{r}\; sin(\hat{\xi}). \tag{5.4b}$$

¿From the Hessian conditions (5.2a) and (5.2b) it is then possible to derive the identities,

$$\left(\frac{dr}{d\hat{r}}\right)^2 = \frac{d^2\Delta\phi}{d\hat{r}^2}, \tag{5.5a}$$

$$\frac{1}{\hat{r}}\frac{d\Delta\phi}{d\hat{r}} = \frac{r^2}{\hat{r}^2}\frac{(1-r^2)}{\alpha^2}, \tag{5.5b}$$

and hence, by eliminating the derivatives of $\Delta\phi$, to find that the logarithmic derivative,

$$p\equiv\frac{d\,\log\hat{r}}{d\,\log r}, \tag{5.6}$$

satisfies the quadratic equation,

$$p^2 - 2\beta p + \gamma = 0. \tag{5.7}$$

The β and γ in (5.7) are the following functions of r:

$$\beta = \frac{(1-2r^2)}{(1-r^2)}, \tag{5.8a}$$

$$\gamma = \frac{\alpha^2}{(1-r^2)}. \tag{5.8b}$$

The appearance of solutions in pairs corresponds to the fact that the isotropic Hessian condition for generating function ϕ' implies a Hessian condition for the dual generating function Φ' of identical form except for a sign change. From the formula for the two possible solutions, p,

$$p^{\pm} = \beta\pm(\beta^2-\gamma)^{1/2}, \tag{5.9}$$

and the definitions (5.8a) and (5.8b) it is apparent that, in order to obtain a real valued logarithmic derivarive p over the whole hemispheric range, $0 < r < 1$, the quantity α^2 cannot be positive. As we have seen, the paradox of imaginary angular scaling is nicely resolved with the aid of a pseudo-Euclidean mapped domain. It is natural, then, to select the *same* angular scaling as used in the f-plane vortex so that the hemispheric theory most closely corresponds with the previous development. This judicious choice also appears to lead to the hemispheric theory of greatest formal simplicity and symmetry. Thus, choosing $\alpha^2 = -8$, our solutions for p are,

$$p^- = \frac{-2}{1-r^2}, \tag{5.10a}$$

$$p^+ = 4. \tag{5.10b}$$

The first solution, p^-, is the one corresponding to the generalized vortex theory, which gives $\hat{r}$ within an arbitrary multiplicative constant and $\Delta\phi$ as follows:

$$\hat{r} = -\frac{(1-r^2)}{2r^2}, \tag{5.11a}$$

$$\Delta\phi = \frac{r^2}{8}. \tag{5.11b}$$

As in the f-plane vortex, it is convenient to redefine the transformed angular variables so that they become real, that is, replace the definition (5.3b) with $\hat{\xi} = (+8)^{1/2}\xi$ while simultaneously replacing the circular functions in (5.4a) and (5.4b) by their hyperbolic counterparts. The exact form for the horizontal part of the energy function compatible with this geometry and with a vanishing kinetic energy when $(r,\xi) = (R,\Xi)$ is,

$$\mathcal{E}(r,\xi;R,\Xi) = \frac{R^4}{8}(\frac{1}{r^2} - 1)\cosh(\sqrt{8}(\xi - \Xi)) + \frac{r^2}{8} - \frac{R^2}{4} + \frac{R^4}{8}, \tag{5.12}$$

and, when,

$$\hat{R} = \frac{R^4}{4}, \tag{5.13a}$$

$$\hat{\Phi} = \Phi + \frac{R^2}{4} - \frac{R^4}{8}, \tag{5.13b}$$

together with

$$\hat{x} = \hat{r}\ \cosh(\xi), \tag{5.14a}$$

$$\hat{y} = -\hat{r}\ \sinh(\xi), \tag{5.14b}$$

$$\hat{X} = \hat{R}\ \cosh(\Xi), \tag{5.14c}$$

$$\hat{Y} = \hat{R}\ \sinh(\Xi), \tag{5.14d}$$

we reproduce the conditions (4.5) for Legendre-duality to hold. Note that (5.13a) confirms that the other solution, p^+, of (5.10b) provides the dual radial transform. As noted in P93, the representations of the dual radial variables $\hat{R}$ in the f-plane and hemispheric vortex models are identical if we identify the polar value of Coriolis of the hemispheric case with the f used in the f-plane vortex model. Both vortex SG theories possess a form of "frame invariance" in the sense that the choice of rotation rate (and hence, Coriolis parameter) used to define the frame of reference can be changed without fundamentally changing the physical content of the theory itself (such a change involves a consistent change in the definition of geopotential and of kinetic energy, of course; the relevant transformations are discussed in P93). This frame invariance appears only to hold in the hemispheric Legendre-transformable case when we adopt our present choice, $(-8)^{1/2}$, for the α in (5.3b).

Figure 3a plots the curves formed by various pairs of intersecting neutral energy surfaces for this hemispheric vortex model. These curves are plotted in transverse Mercator conformal projections in order to minimize the distortion of the shapes of small figures located near the principal meridian. Here we show two foci, since the patterns in physical space are no longer invariant under scaling of r. Note that oblique curves never intersect the equator. Figure 3b plots the contours of the energy function, again, for several values of R. Fig. 3c includes the case of a value R slightly exceeding the earth's radius; since the theory allows values of R (but obviously not r) larger than the earth's radius, the implicitly restrictive "potential latitude" is *not* an appropriate dual coordinate in general applications. This point is further reinforced when we plot the energy contours, for the cases in which r is fixed, directly on the projected dual equatorial plane. Figure 3d shows a selection of these contours, together with the projection of lines of latitude and longitude (dotted). From the fact that some contours cross the circle of the projected equator we deduce that this circle represents no intrinsic barrier in the dual domain (recall that dual flow is tangent to these contours of the dual generating surface at the point of its contact with the dual solution).

6. Numerical considerations

For a finite element implementation, it is natural to perform almost all the calculations in terms of the Legendre-dual coordinates and geopotentials ("hatted" physical and dual variables). As we have seen, the dynamics can be expressed in a convenient (but non-canonical) form when we have both the Hamiltonian and effective Coriolis parameter defined in the system of dual coordinates. In our case, the computation of the Hamiltonian and its first derivative components with respect to Legendre-dual space variables requires the intermediate computations of certain "moments" of each contributing element. A generic element takes its simplest form in these Legendre-transformable physical coordinates (it then has the form of some convex polyhedron) but its internal distribution of effective density $\hat{\eta}$ and of specific energy $\mathcal{E}$ are both non-linear functions of these coordinates. Fortunately, both functions are smooth away from the singularity representing the equator. The element's total energy and its mass can therefore be approximated to any desired accuracy using an expansion in terms of successive moments of the element combined with the first few Taylor series coefficients, about the same point, of the specific energy-density and of the mass-density. We shall not pursue these technical issues in great detail, but a brief outline of the basic idea, exemplified by the problem of estimating the mass in an element, is instructive.

Consider the example of the f-plane vortex, with uniform effective density η in the physical space mapping to a non-uniform counterpart $\hat{\eta}$ in the Legendre-transformable domain via the Jacobian:

$$\hat{\eta} = -\frac{r}{\hat{r}}\frac{\partial(r,\xi)}{\partial(\hat{r},\hat{\xi})}\eta. \tag{6.1a}$$

$$= \frac{\eta}{4\,(8)^{1/2}\,\hat{r}^3}. \tag{6.1b}$$

The total mass $\mathcal{I}_\sigma(\hat{\eta})$ of a finite element σ is therefore the integral,

$$\mathcal{I}_\sigma(\hat{\eta})\equiv\int_\sigma \hat{\eta}\, d\hat{\mu}, \tag{6.2}$$

where,

$$d\hat{\mu}\equiv d\hat{x}\, d\hat{y}\, d\hat{z}\equiv d\hat{x}^1\, d\hat{x}^2\, d\hat{x}^3.$$

We can expand $\hat{\eta}$ about some point $\hat{\mathbf{x}}_\sigma$ near element σ as a Taylor series:

$$\hat{\eta}(\mathbf{x}) = \hat{\eta}_\sigma + \sum_i \frac{\partial \hat{\eta}_\sigma}{\partial \hat{x}^i}(\hat{x}^i - \hat{x}^i_\sigma) + \frac{1}{2!}\sum_{i,j} \frac{\partial^2 \hat{\eta}_\sigma}{\partial \hat{x}^i \partial \hat{x}^j}(\hat{x}^i - \hat{x}^i_\sigma)(\hat{x}^j - \hat{x}^j_\sigma)\cdots, \tag{6.3}$$

Whence:

$$\mathcal{I}_\sigma(\hat{\eta}) = \hat{\eta}_\sigma \mathcal{I}_\sigma(1) + \sum_i \frac{\partial \hat{\eta}_\sigma}{\partial \hat{x}^i}\mathcal{I}_\sigma(\hat{x}^i - \hat{x}^i_\sigma) + \frac{1}{2!}\sum_{i,j} \frac{\partial^2 \hat{\eta}_\sigma}{\partial \hat{x}^i \partial \hat{x}^j}\, \mathcal{I}_\sigma\left((\hat{x}^i - \hat{x}^i_\sigma)(\hat{x}^j - \hat{x}^j_\sigma)\right)\cdots \tag{6.4}$$

The mass, and more generally, mass-weighted moments (of which the parcel-integrated energy is an example), are therefore expressed in terms of the parcel's *ordinary* moments in $\hat{\mathbf{x}}$-space. The latter are relatively straight-forward to compute, since each element σ comprises a polyhedron.

This technique of employing moment expansions obviously applies to the evaluation of the Hamiltonian as well as to verifying the mass of each element, but it can also be shown that the *derivatives* of such mass-weighted moments, with respect to variations of the dual extended-coordinates, $(\hat{\mathbf{X}}_\tau, \hat{\Phi}_\tau)$, for τ and σ either identical or adjacent, can be evaluated using analogous expansions with moments associated with the *interfaces* between elements. Such methods should enable the bulk of the computations, associated with both solution generation and the computation of its instantaneous trajectory, to be carried out in the spaces of $\hat{\mathbf{x}}$ and $\hat{\mathbf{X}}$.

Optimal convergence will presumably be obtained when the location of each $\hat{\mathbf{x}}_\sigma$ at which the partial derivatives are evaluated is close to the center of the corresponding element σ. Then a reasonable resolution will ensure that the variations of $\hat{\eta}$, or of the product of this density with the energy function, will be small enough across the element to make even a short moment expansion an extremely accurate estimate.

7. Conclusions

We have shown that the contact structure of a generic form of the Hamiltonian SG theory generalizing the work of Salmon (1985) may imply "anomalies" in the connectivity of neighboring fluid elements and that these anomalies can only be completely eliminated by ensuring that the contact structure is of the special class that admits a Legendre-transformable representation. We have proceeded to re-examine the SG vortex theories on the f-plane and on the hemisphere proposed by this author in a recent paper (P93) and determined the necessary minor modifications required to render them Legendre-transformable, and hence, anomaly-free. The Legendre-transformable representations exhibit the curious feature of preserving angular symmetry only when we interpret the effective metric as being pseudo-Euclidean; rotational invariance of the original theory takes the form of Lorentz invariance in the new representation and concentric circles in the vortex are mapped to concentric hyperbolae.

The new versions of the SG vortex theories will enable the "geometric model" techniques of Cullen (1983) and Cullen and Purser (1984) to be extended to fully three-dimensional simulations for which these assumptions of approximately gradient balance are valid. But, unlike the more restrictive model of Shutts et al. (1988), we are now able to handle azimuthally varying components in such solutions, for example, in a simulation of the internal structure of a developing tropical cyclone. A part of one sector of this model can also be adapted to simulations of significantly curved fronts, largely overcoming the defects of standard SG in this area reported by Gent et al. (1991). In the hemispheric case, the distinction between gradient and geostrophic balance of a zonal wind is virtually insignificant (except very close to the pole itself) and so we can legitimately regard the hemispheric vortex model as a minor variant of Salmon's $\mathcal{L}_s$-dynamics. We can therefore look forward to future implementations of Salmon's powerful generalization of SG theory in both the atmosphere and oceans, using the geometric method. In the oceanic case, this will allow simulation of a variety of discontinuous phenomena, such as boundary current separation, outcropping of subsurface layers and the evolution of unsteady currents where strong thermodynamic and momentum gradients come into play. In the atmospheric case, we shall be able to perform idealized simulations of the entire life-cycle of fully nonlinear baroclinic wave in the total absence of artificial viscosity or thermal diffusion, since the geometric method is not compromised by the formation of discontinuities. The new hemispheric model is formally self-consistent even in tropical latitudes. It is therefore a candidate for the study of quasi-steady monsoon circulations. As discussed in P93, it should be feasible in practice to combine a pair of hemispheric models of this form with dual coordinates matched at the equator, and allow exchange of mass (conserving this dual coordinate) between the hemispheres, thereby obtaining a fully global model.

Other methods based on generalized "balance" but with consistent analogs of energy conservation and circulation invariants have been proposed recently (e.g., Allen et al. 1990; Allen and Holm 1996; Shapiro and Montgomery 1993) but it is unclear how well these methods are able to accommodate the formation of frontal contact discontinuities. The SG models, while perhaps formally less accurate, *do* possess this ability to handle the formation and evolution of contact discontinuities without difficulty. Because of this, they can be used to generate a number of valuable "bench-mark" tests in which such discontinuities are prominent, and against which the more conventional methods of spatial discretization employed by operational forecasting and climate models can be compared and improved. Thus, the methods proposed here could, indirectly, have an impact on the technical development of operational forecast and climate models, now that the typical resolution of such models is beginning to make the proper handling of frontal details a relevant consideration.

Acknowledgements. This work was carried out during the author's participation in the programme, "Atmosphere and Ocean Dynamics" at the Isaac Newton Institute (INI), Cambridge, UK. The author thanks the staff of the INI and the programme organizers for their assistance. The work benefitted from stimulating discussions with Drs M. J. P. Cullen, I. Roulstone, D. D. Holm, R. Salmon, J. Norbury, I. Sunyogh and Profs. E. Titi, M. Sewell and T. Ratiu.

APPENDIX A

Proof of theorem 1

Write

$$\mathcal{E}_{\alpha\beta} = \mathcal{E}(\mathbf{x}, \mathbf{X}_\alpha) - \mathcal{E}(\mathbf{x}, \mathbf{X}_\beta), \tag{A.1}$$

and use,

$$\hat{\mathbf{x}}(\mathbf{x}) = -\left(\frac{\partial}{\partial X_\alpha}, \frac{\partial}{\partial Y_\alpha}\right) \mathcal{E}_{\alpha 0}|_{\hat{\mathbf{X}}_\alpha = \hat{\mathbf{X}}_0}\,, \tag{A.2}$$

to define new coordinates $\hat{\mathbf{x}} \equiv (\hat{x}, \hat{y})$ where the contact structure is "generic", or non-degenerate in the sense that,

$$\frac{\partial(\hat{x}, \hat{y})}{\partial(x, y)} \neq 0. \tag{A.3}$$

By construction, the curvature of the contours of the difference, $\mathcal{E}_{\alpha\beta}$, vanishes everywhere in $\hat{\mathbf{x}}$-space as we take the limit (in any direction), $\mathbf{X} \to \mathbf{X}_0$. Therefore, prohibition of the bowtie or crescent anomalies described in section 3 requires, for *all* pairs $\mathbf{X}_\alpha$, $\mathbf{X}_\beta$, that the contours of the general difference, $\mathcal{E}_{\alpha\beta}$, are also straight relative to the new coordinates $\hat{\mathbf{x}}$. Thus we have a generic formula describing the contour:

$$\hat{a} + \hat{b}\hat{x} + \hat{c}\hat{y} = 0, \tag{A.4}$$

where the three coefficients of this affine equation are functions jointly of $\mathcal{E}_{\alpha\beta}$, $\mathbf{X}_\alpha$ and $\mathbf{X}_\beta$. Consider the first-order variations of these coefficients as we change $\mathbf{X}_\beta$ in the vicinity of $\mathbf{X}_0$. Denoting gradient with respect to $\mathbf{X}_\beta$ at $\mathbf{X}_0$ by ∇_β, and partial derivatives of $\hat{b}$ and $\hat{c}$ with respect $\mathcal{E}_{\alpha 0}$, also at $\mathbf{X}_\beta = \mathbf{X}_0$, by $\hat{b}_{\mathcal{E}}$ and $\hat{c}_{\mathcal{E}}$, then we find that,

$$\nabla_\beta \hat{a} + (\nabla_\beta \hat{b})\hat{x} + (\nabla_\beta \hat{c})\hat{y} + \hat{\mathbf{x}}(\hat{b}_{\mathcal{E}}\, \hat{x} + \hat{c}_{\mathcal{E}}\, \hat{y}) = 0. \tag{A.5}$$

Maintaining straight-line contours necessitates that the terms quadratic in $\hat{x}$ and $\hat{y}$ vanish. Hence,

$$\hat{b}_{\mathcal{E}} = \hat{c}_{\mathcal{E}} = 0, \tag{A.6}$$

and, since this is equivalent to saying that the contours of $\mathcal{E}_{\alpha 0}$ are mutually parallel in $\hat{x}$-space, we can therefore alway express each energy difference of the type, $\mathcal{E}_{\alpha 0}$ as a scalar function of a dot product:

$$\mathcal{E}_{\alpha 0} \equiv b_\alpha(\hat{\mathbf{x}} \cdot \hat{\mathbf{U}}_\alpha), \tag{A.7}$$

for some unit covector $\hat{\mathbf{U}}_\alpha$ independent of $\hat{\mathbf{x}}$. Finally, we write the general energy difference $\mathcal{E}_{\alpha\beta} \equiv \mathcal{E}_{\alpha 0} - \mathcal{E}_{\beta 0}$ and substitute the form (A.7):

$$\mathcal{E}_{\alpha\beta} = b_\alpha(\hat{\mathbf{x}} \cdot \hat{\mathbf{U}}_\alpha) - b_\beta(\hat{\mathbf{x}} \cdot \hat{\mathbf{U}}_\beta), \tag{A.8}$$

to infer that an infinitesimal displacement $d\hat{\mathbf{x}}$ parallel to a contour of this difference must satisfy,

$$d\hat{\mathbf{x}} \cdot (\hat{\mathbf{U}}_\alpha \, b'_\alpha - \hat{\mathbf{U}}_\beta \, b'_\beta) = 0. \tag{A.9}$$

In the generic case, $\hat{\mathbf{U}}_\alpha$ will not be parallel to $\hat{\mathbf{U}}_\beta$ and so both derivatives b'_α and b'_β must vanish in order for (A.9) to hold over any finite area. Thus, not only are the contours parallel, but they are also uniformly spaced, implying that *all* the energy differences $\mathcal{E}_{\alpha\beta}$ are affine functions of the new coordinates $\hat{\mathbf{x}}$. By an additive modification of each neutral energy function,

$$\begin{aligned} \hat{\phi}'_\alpha &= \phi'_\alpha + \mathcal{E}(\mathbf{x}, \mathbf{X}) \\ &\equiv \Phi_\alpha - \mathcal{E}_{\alpha 0}, \end{aligned} \tag{A.10}$$

and regarding $\hat{\phi}'_\alpha$ as a function of $\hat{\mathbf{x}}$, we find now that this modified generating function is expressible in manifestly Legendre-transformable form as asserted:

$$\begin{aligned} \hat{\phi}'_\alpha &= (\Phi_\alpha - b_\alpha(0)) + \hat{\mathbf{x}} \cdot \left(-\hat{U}_\alpha \, b'_\alpha\right) \\ &\equiv \hat{\Phi}_\alpha + \hat{\mathbf{x}} \cdot \hat{\mathbf{X}}_\alpha. \end{aligned} \tag{A.11}$$

REFERENCES

Allen, J. S., Barth, J. A., and P. A. Newberger, 1990: On intermediate models for barotropic continental shelf and slope flow fields. Part I: Formulation and comparison of exact solutions. *J. Phys. Oceanog.*, **20,** 1017–1043.

——,and D. D. Holm, 1996: Extended-geostrophic Hamiltonian models for rotanting shallow water motion. *Physica D*, (in press)

Arnold, V. I., 1980: *Mathematical Methods of Classical Mechanics.* 2nd Ed. Springer Verlag. 508pp.

Blumen, W., 1981: The geostrophic coordinate transformation. *J. Atmos. Sci.*, **38,** 1100–1105.

Chynoweth, S., 1987: The semigeostrophic equations and the Legendre transformation. Ph.D. thesis, Department of Mathematics, University of Reading, 248 pp. [Available from the Department of Mathematics, University of Reading, Whiteknights, Reading, RG6 2AU, UK.]

——, D. Porter and M. J. Sewell, 1988: The parabolic umbilic and atmospheric fronts. *Proc. R. Soc. Lond.*, **A 419,** 337–362.

——, and M. J. Sewell, 1989: Dual variables in semigeostrophic theory. *Proc. R. Soc. Lond.*, **A 424,** 155–186.

Craig, G. C., 1991: A three-dimensional generalization of Eliassen's balanced vortex equations derived from Hamilton's principle. *Quart. J. Roy. Meteor. Soc.*,**117,** 435–448.

Cullen, M. J. P., 1983: Solutions to a model of a front forced by deformation. *Quart. J. Roy. Meteor. Soc.*, **109,**565–573.

——, and R. J. Purser, 1984: An extended Lagrangian theory of semi-geostrophic frontogenesis. *J. Atmos. Sci.*, **41,** 1477–1497.

——, S. Chynoweth and R. J. Purser, 1987a: On semi-geostrophic flow over synoptic scale topography. *Quart. J. Roy. Meteor. Soc.*, **109,**163–180.

——, J. Norbury, R. J. Purser and G. J. Shutts, 1987b: Modelling the quasi-equilibrium dynamics of the atmosphere. *Quart. J. Roy. Meteor. Soc.*, **109,** 735-758.

Eliassen, A., 1951: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norv.*, **5,** 19–60.

——, 1962: On the vertical circulation in frontal zones. *Geophys. Publ.*, **24,** 147–160.

——, and E. Kleinschmidt, 1957: *Dynamic Meteorology. Handbuch der Physik, Geophysik II*, Springer-Verlag. 154pp.

Gent, P. R., J. C. McWilliams, and C. Snyder, 1994: Scaling analysis of curved fronts: validity of the balance equations and semigeostrophy. *J. Atmos. Sci.*, **51,** 160–163.

Hoskins, B. J., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, **32,** 233–242.

——,and F. P. Bretherton, 1972: Atmospheric frontogenesis models: mathematical foundation and solution. *J. Atmos. Sci.*, **29,** 11–37.

Magnusdottir, G., and W. H. Schubert, 1991: Semigeostrophic theory on the hemisphere. *J. Atmos. Sci.*,**48,** 1449–1456.

McIntyre, M. E., and I. Roulstone, 1996: Hamiltonian balanced models: constraints, slow manifolds and velocity splitting. Meteorological Office Forecasting Research Scientific Paper No. 41. [Available from the Meteorological Office, London Road, Bracknell, RG12 2SZ, UK.]

Purser, R. J., 1993: Contact transformations and Hamiltonian dynamics in generalized semigeostrophic theories. *J. Atmos. Sci.*, **50,** 1449–1468.

——, and M. J. P. Cullen, 1987: A duality principle in semigeostrophic theory. *J. Atmos. Sci.*,**44,** 3449–3468.

Roulstone, I., and M. J. Sewell, 1996: Potential vorticity in semi-geostrophic theory. *Quart. J. Roy. Meteor. Soc.*,**122,** 983–992.

Salmon, R., 1983: Practical use of Hamilton's principle. *J. Fluid Mech.*,**132,** 431–444.

——,1985: New equations for nearly geostrophic flow. *J. Fluid Mech.*,**153,** 461–477.

Schubert, W. H., 1985: Semigeostrophic theory. *J. Atmos. Sci.*, **42,** 1770–1772.

——, and J. J. Hack, 1983: Transformed Eliassen balanced vortex model. *J. Atmos. Sci.*, **40,** 1561–1583.

Sewell, M. J., and I. Roulstone, 1994: Families of lift and contact transformations. *Proc. R. Soc. Lond.*, **A 447,** 493–512.

Shapiro, L. J., and M. T. Montgomery, 1993: A three dimensional balance theory for rapidly rotating vortices. *J. Atmos. Sci.*, **50,** 3322–3335.

Shepherd, T. G., 1990: Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Adv. Geophys.*, **32,** 287–338.

Shutts, G. J., 1989: Planetary semi-geostrophic equations derived from Hamilton's principle. *J. Fluid Mech.*, **208,** 545–573.

——, and A. J. Thorpe, 1978: Some aspects of vortices in rotating stratified flows. *Pure Appl. Geophys.*, **116,** 993–1006.

——, M. Booth and J. Norbury, 1988: A geometric model of balanced, axisymmetric flows with embedded penetrative convection. *J. Atmos. Sci.*, **45,** 2609–2621.

Figure Captions

FIG. 1. Schematic illustration of bowtie (a) and crescent (b) anomalies.

FIG. 2. Geometrical structures implied by the f-plane vortex model. (a) Curves through a point in the (x, y)-plane formed by intersecting pairs of neutral energy surfaces; (b) Contours of kinetic energy function $\mathcal{E}$ in the (x, y)-plane for two fixed values of $\mathbf{X}$; (c) Contours of $\mathcal{E}$ in the (X, Y)-plane for two fixed values of $\mathbf{x}$.

FIG. 3. Geometrical structures implied for the hemispheric vortex model. (a) Curves of intersection of pairs of neutral energy surfaces passing through the points at latitudes 45^o and 5^0, as they would appear in a transverse Mercator projection. (b) Contours of the energy function plotted in the same projection at intervals of $20ms^{-1}$ in equivalent speed, for fixed dual coordinates corresponding to locations on the central meridian at latitudes 75^o, 45^o and 15^o. (c) Energy contours as in (b), but for a dual coordinate at $R = 1.02$ not corresponding to a latitude. The largest contour value is $100ms^{-1}$; (d) Contours of the energy function at intervals of $20ms^{-1}$ in equivalent speed plotted by normal projection onto the equatorial plane for fixed values of $\mathbf{x}$ corresponding to latitudes 75^o, 45^o and 15^o. Note that the contours may cross the projected circle of the equator.