J. Fluid Mech. (1994), *vol.* 272, *pp.* 211–233

A Hamiltonian structure with contact geometry for the semi-geostrophic equations

By I. ROULSTONE[1] AND J. NORBURY[2]

[1] Forecasting Research Division, Meteorological Office, London Road, Bracknell, Berkshire, RG12 2SZ, UK

[2] Mathematical Institute, 24–29 St. Giles, Oxford, OX1 3LB, UK

(Received 9 November 1992 and in revised form 11 February 1994)

A canonical Hamiltonian structure for the semi-geostrophic equations is presented and from this a reduced non-canonical Hamiltonian structure is derived, providing a fully nonlinear version of the approximate linearized vorticity advection representation. The structure of this model is described naturally within the framework of contact geometry. A Hamiltonian approach leading to a symplectic algorithm for calculating solutions to the equations of motion is formulated. Basic necessary functional methods are introduced and the Lagrangian and Eulerian kinematic structures are discussed, together with their relevance to symplectic integrating algorithms.

1. Introduction

Salmon (1985) has offered a new approach to semi-geostrophic theory using approximations to Hamilton's principle within the framework of Lagrangian particle labelling and has subsequently showed that the method may be understood in terms of generalized Hamiltonian systems (Salmon 1988*b*). In this paper we derive a symplectic approach to the semi-geostrophic equations with a view to using the results in numerical Hamiltonian mechanics, a relatively recent and promising development in numerical analysis (see Feng 1991; also McLachlan 1993).

The use of Hamiltonian methods in fluid mechanics has become the fashion in recent years as this enables one to use methods of modern geometry and algebra, in particular the symplectic structure and associated Poisson bracket, in formulating the equations of motion, giving insight, in particular, into the relationship between the underlying kinematics and the dynamics of the fluid. We use the word kinematics in the sense of volume-preserving rearrangements of the fluid particles in the physical domain and reserve the term dynamics for the description of the evolution of the flow.

The Hamiltonian approach may be especially important in atmospheric dynamics for a variety of reasons. The Lagrangian versus the Eulerian formulation of the kinematics, the stable long-time numerical integration of the equations of motion and the existence of a relationship between symmetries and conservation laws (Noether's theorem) are all subjects of significant interest to the theorist and modeller. Studies using *filtered* or *balanced* models (which attempt to focus the solution effort on stable fluid behaviour) are essential for describing flow regimes that dominate certain spatial or temporal scales. In these cases the Hamiltonian approach offers a systematic method for studying, for example, the conservation laws and related symmetries of the model which are preserved exactly even though approximations to the full equations have been used.

In this paper we are concerned with a three-dimensional version of the semi-

geostrophic (SG) equations (Hoskins & Draghici 1977). Hoskins & Bretherton (1972) showed that the SG equations may be expressed in terms of Lagrangian conservation laws and Cullen *et al.* (1987) have developed a numerical model (the so-called *geometric model*) which represents the fluid as a finite collection of parcels with conserved momentum and thermodynamic labels. A stable manifold within the dynamical system of the atmosphere is defined by using a convexity principle to minimize the energy. An extra advantage of this principle is that it applies to variables which have discontinuities. It is not our purpose here to discuss the solution structure of either SG theory or the geometric model, but rather the kinematics and the generation of the consistent dynamics. Salmon (1985) and Shutts (1989) have obtained versions of the SG equations from Hamilton's Principle – a statement that encompasses Newton's laws of motion. Their approach is based on Lagrangian kinematics (in the sense of position/velocity space), whereas for the purposes of studying numerical Hamiltonian dynamics and identifying the *symplectic algorithms*, the basis of our study needs to be fully Hamiltonian.

Here we are concerned with Lagrangian (in the sense of particle labelling coordinates) and Eulerian formulations of the equations. Recent developments (Channell & Scovel 1990; Feng 1991; McLachlan 1993) in the numerical integration of Hamiltonian systems have emphasized the importance of preserving kinematic structure in solution strategies for these models. We will pay particular attention to this structure as the SG equations of motion may be written either as an infinite set of coupled ordinary differential equations or, as in the formulation of the geometric model, an advection equation based on a *nonlinear vorticity/streamfunction system.* In this latter case the nonlinear Monge–Ampère equations replaces the usual 'vorticity equals the negative of Laplacian of streamfunction' relationship. A natural description of this system is found within the framework of *contact geometry* and is discussed in §4.2 of this paper. The contact geometry together with convex analysis and their application to atmospheric flows offers a new and crucially different perspective on both balanced systems (as they are known in the meteorological literature), and vorticity/streamfunction theories (see, for example, Shepherd 1990 for a review of the latter).

It is natural to develop the structure of infinite-dimensional systems in analytical mechanics from those of point particles. This approach is adopted by Salmon (1988*a*) and Shepherd (1990), and therefore we will not expand upon these matters here. It is sufficient for our present purposes to note that there are applications of finite-dimensional systems in fluid dynamics that form a useful approximation to the continua and models that exploit these ideas have been developed. Of particular interest are those with structures familiar from analytical mechanics (see, for example, Purser 1988 and also Shutts & Cullen 1987). A brief review of functional methods and symplectic geometry is given in the Appendix. Sewell & Roulstone (1993*a*, *b*) discuss some of the relevant features of canonical, contact and Legendre transformations. See Salmon (1988*a*) and Shepherd (1990) for introductions to Hamiltonian methods in fluids and Marsden's (1992) lectures, for a discussion of symplectic methods in numerical models.

The paper is organized as follows. In §2 we give a brief review of the definitions of Hamiltonian systems. In §3 the SG model and the Legendre transformation, under which the equations take a particularly simple form, are discussed. We then proceed in §4 to describe the Hamiltonian structure of the equations and discuss the symmetries and conservation laws. In particular, we show how the symplectic structure (the conserved geometry) of a Hamiltonian system, the crucially locally conserved nonlinear

balance, is represented by conservation of vorticity on fluid particles. Finally, in §§5 and 6 we summarize with a discussion of the applications of this work in numerical modelling. An Appendix contains an introduction to some of the mathematical methods that are used in this paper. Index or boldface vector notation, and the summation convention will be used throughout.

2. Hamiltonian systems

We begin with a definition of a Hamiltonian system, to make it clear that our approach requires more than the identification of an appropriate energy functional as a Hamiltonian.

DEFINITION 1. *A Hamiltonian system consists of a Poisson manifold, that is a manifold* $\mathscr{P}$ *together with the bilinear operation* $\{,\}$ *on* $\mathscr{P}$, *known as the Poisson bracket, and a function* $\mathscr{H}$, *on* $\mathscr{P}$, *called the Hamiltonian that generates the evolution of an observable* z^i, *of the system according to*

$$\dot{z}^i = \{z^i, \mathscr{H}\} = J^{ij}\frac{\delta \mathscr{H}}{\delta z^j}. \tag{1}$$

(See the Appendix for the definition and discussion of the terms used here.) In particular, if the system is finite-dimensional and the matrix J^{ij} takes the form

$$J^{ij} = \begin{pmatrix} 0^n & 1^n \\ -1^n & 0^n \end{pmatrix}, \tag{2}$$

with 0^n and 1^n being the $n \times n$ zero and identity matrices, then the system is said to be *canonical*. A transformation of the phase-space coordinates z^i that preserves the form (2) is called a *canonical transformation* or *symplectic map*. To make this clear, consider a change of coordinates $z^i \mapsto f^i(z^j)$; then the equations satisfied by f^i are

$$\dot{f}^i = A^i_j J^{jk}\frac{\partial \mathscr{H}}{\partial z^k} = A^i_j J^{jk} A^{lT}_k \frac{\partial \mathscr{H}(z(f))}{\partial f^l}, \tag{3}$$

where A^i_j is the matrix of derivatives $\partial f^i/\partial z^j$ and A^{iT}_j is the transpose of A^i_j. The new equations will be Hamiltonian iff $\boldsymbol{AJA}^T = \boldsymbol{J}$, that is, the transformation must preserve the canonical (cosymplectic) structure. For further discussion of canonical transformations, see Sewell & Roulstone (1993*a*). A symplectic numerical scheme is a finite-dimensional Hamiltonian system with the transformation from data at a time level t_0 to data at time level t_1 being symplectic or canonical.

In this paper we consider systems of partial differential equations that may be written in the form

$$\frac{\partial z}{\partial t} = \mathscr{J}\Delta(\mathscr{H}),$$

where $\mathscr{J}$ is, in general, a skew adjoint matrix of (pseudo-) differential operators, $\mathscr{H}[z]$ is the Hamiltonian functional and Δ denotes the variational derivative with respect to z, $\delta/\delta z$. These considerations are formal as the Sobolev subspace of L^2 under consideration will not be specified. If we assume that $\mathscr{J}$ has an inverse $\mathscr{J}^{-1}$, and that all operators are defined in the sense of the inner product $\langle \cdot | \cdot \rangle$ (see the Appendix for further detail), than (3) becomes, again formally,

$$\dot{f}^i = \left\langle A^i_j \middle| J^{jk}\frac{\delta \mathscr{H}}{\delta z^k} \right\rangle = \left\langle A^i_j \middle| J^{jk} A^{lT}_k \frac{\delta \mathscr{H}[z(f)]}{\delta f^l} \right\rangle \equiv A^i_j J^{jk} A^{lT}_k \frac{\delta \mathscr{H}[z(f)]}{\delta f^l}.$$

In addition, $\mathscr{J}$ explicitly depends on the function z and its derivatives, so it is not

canonical. Therefore the closure condition on the associated symplectic structure, which is the infinite-dimensional analogue of Darboux' theorem (Arnol'd 1989) giving the necessary and sufficient conditions for a variable skew-symmetric matrix to be equivalent to the standard symplectic matrix (2), is not easy to prove. In infinite dimensions, the degeneracies of $\mathscr{J}$ make it hard to determine an easily defined change of coordinates to the standard form. Note that Salmon's (1985) work is for canonical systems. A non-canonical example is discussed by Olver (1982) for the Euler equations, where the operator equivalent to $\mathscr{J}$ is shown to have an inverse. A further related difficulty that reflects the difference between Eulerian and Lagrangian variables is that $\partial z/\partial t$ is not the same as the derivative with respect to the Lagrangian time label τ (see §4 for an explanation of the notation). Our aim is to show that (1) is the appropriately reduced, non-canonical Hamiltonian system for the SG equations, when z^i, $\{,\}$ and $\delta/\delta z^i$ are appropriately identified.

3. Semi-geostrophic dynamics

3.1. *Equations of motion*

SG theory is an attempt to model atmospheric flows that vary on scales of typical synoptic patterns with characteristic timescales of at least several hours. Current operational forecast models, however, use a system of equations which describe a wider class of solutions than these, which include fast-moving inertia–gravity waves and sometimes sound waves. There are a number of summaries of the SG equations, their derivation and application. Here, we will follow the argument of Cullen *et al.* (1987) and sketch a derivation of the equations to motivate our basic p.d.e. model of the equations of motion. We do this to show that energy minimization is a key concept underlying this model.

Recall that the Boussinesq hydrostatic equations for the motion of a dry atmosphere can be written in terms of local Cartesian coordinates (x, y, z) on some domain D of a hemisphere, with y pointing north polewards and z as the vertical. The equations of motion for the two horizontal components of the velocity $\boldsymbol{v} \equiv (u, v)$ are

$$\frac{\mathrm{D}\boldsymbol{v}}{\mathrm{D}t} + f\boldsymbol{k} \wedge \boldsymbol{v} + \nabla_x \hat{\phi} = 0,$$

which are taken together with hydrostatic balance $\partial\hat{\phi}/\partial z = g\theta/\theta_0$, incompressibility $\nabla_x \cdot \boldsymbol{u} = 0$, and conservation of potential temperature $\mathrm{D}\theta/\mathrm{D}t = 0$. The material derivative is defined in terms of the locally Cartesian coordinates as

$$\frac{\mathrm{D}}{\mathrm{D}t} \equiv \frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla_x;$$

$\boldsymbol{u} \equiv (u, v, w)$ are respectively the three components of velocity and f is the Coriolis parameter (this amounts to a tangent plane approximation in that the effect of rotation is accounted for but not the variation of the Coriolis effect with latitude); and $\hat{\phi}$ is the geopotential. The boundary conditions at the ground surface and top of the model atmosphere are taken as $w = 0$ at $z = 0, H$.

Cullen *et al.* identify a balanced model by the requirement that motion be constrained so that an invariant energy of the system be minimized at all times. The expression for the energy of the full Euler equations is usually approximated by

$$\hat{E} = \int_D \left\{ \tfrac{1}{2}(u^2 + v^2) - \frac{g\theta z}{\theta_0} \right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4}$$

The equilibrium configuration is defined by requiring that the energy integral be stationary with respect to any virtual displacement of the fluid. The flow will be in geostrophic and hydrostatic balance, when the energy integral is stationary, if the horizontal velocity field is parallel to the $\hat{\phi}$-contours. Equilibrium means that the horizontal acceleration is small compared to the balancing terms. In the meteorological literature this would be referred to as neglecting the *ageostrophic* acceleration (see Hoskins 1975), and it suggests defining a geostrophic horizontal velocity field (u_g, v_g) in balance with the horizontal pressure gradient.

We seek ϕ, u_g and v_g to be approximations to $\hat{\phi}$, u and v and define them as follows. The evolving equilibrium model corresponds to the solutions of (see Hoskins & Draghici 1977; also Purser & Cullen 1987) the following approximations to the Boussinesq hydrostatic equations for the motion of a dry atmosphere:

$$\left.\begin{aligned} \frac{\mathrm{D}u_g}{\mathrm{D}t} - fv + \frac{\partial \phi}{\partial x} &= \frac{\mathrm{D}u_g}{\mathrm{D}t} - f(v - v_g) = 0, \\ \frac{\mathrm{D}v_g}{\mathrm{D}t} + fu + \frac{\partial \phi}{\partial y} &= \frac{\mathrm{D}v_g}{\mathrm{D}t} + f(u - u_g) = 0, \end{aligned}\right\} \tag{5}$$

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = 0, \tag{6}$$

$$\nabla_x \cdot \boldsymbol{u} = 0, \tag{7}$$

with ϕ, u_g and v_g related by

$$\nabla_x \phi \equiv \left(f v_g, -f u_g, \frac{g\theta}{\theta_0} \right).$$

Here, $\boldsymbol{u}$ means the velocity of the fluid particle at $\boldsymbol{x}$, although now the particle evolves according to the above equations.

Defining transformed coordinates

$$\boldsymbol{X} \equiv (X, Y, Z) \equiv \left[x + \frac{v_g}{f}, y - \frac{u_g}{f}, \frac{g\theta}{f^2 \theta_0} \right] \tag{8}$$

we find that (5) and (6) may be replaced by (using $\mathrm{D}\boldsymbol{x}/\mathrm{D}t \equiv \boldsymbol{u}$)

$$\frac{\mathrm{D}\boldsymbol{X}}{\mathrm{D}t} = \boldsymbol{u}_g \equiv (u_g, v_g, 0); \tag{9}$$

that is, the motion in these transformed coordinates is exactly geostrophic. In these transformed coordinates a consistent approximation to the energy integral can be written as

$$\begin{aligned} E &= \int_D \left(\tfrac{1}{2}\{u_g^2 + v_g^2\} - \frac{g\theta z}{\theta_0} \right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &= f^2 \int_D \left(\tfrac{1}{2}\{(y - Y)^2 + (X - x)^2\} - Zz \right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \end{aligned} \tag{10}$$

When written in this form, the condition for the integral to be stationary under the virtual displacements that we have considered may be expressed in terms of a convexity condition on a streamfunction for the geopotential (Cullen *et al.* 1987). We will introduce such a function in the next section and refer to the expression for the energy integral (10) in §4.5.

The Jacobian

$$q = \frac{\partial(X, Y, Z)}{\partial(x, y, z)} \tag{11}$$

defines a consistent form of the Ertel potential vorticity in SG theory, satisfying

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0, \tag{12}$$

as a lengthy calculation shows. The conservation of vorticity and potential vorticity arises naturally in mathematical terms from the particle labelling structure in the Lagrangian framework (see Salmon 1988*a*) and equivalently, in an explicit Hamiltonian sense, from the invariance of the *symplectic form* (see §4.4 and Abarbanel & Holm 1987).

3.2. *The Legendre transform*

We now focus on the general theorem that (12) represents, that is that potential vorticity is conserved on particles as they move in physical space, and attempt to reformulate this principle in an explicitly Hamiltonian way. Equations (5) and (9) have a particular duality structure. The vector $\boldsymbol{X}$ may be expressed as the gradient of some scalar function $P(\boldsymbol{x})$

$$\boldsymbol{X} = \nabla_x P. \tag{13}$$

Within an arbitrary additive constant this function is uniquely defined by, from (8),

$$P = \frac{1}{f^2}\phi + \tfrac{1}{2}(x^2 + y^2). \tag{14}$$

Define $$\boldsymbol{Q}(\boldsymbol{x}) \equiv \nabla_x \boldsymbol{X},$$

then $\boldsymbol{Q}$ is the Hessian of P with respect to $\boldsymbol{x}$: $\boldsymbol{Q} = \nabla_x \nabla_x P \equiv \mathrm{Hes}(P)$. $\boldsymbol{Q}$ is symmetric, so when it is non-singular its inverse exists, and $\boldsymbol{Q}^{-1} = \nabla_X \boldsymbol{x}$, the inverse Jacobian, where

$$\nabla_X \equiv \left(\frac{\partial}{\partial X}, \frac{\partial}{\partial Y}, \frac{\partial}{\partial Z}\right),$$

is symmetric also, implying that $\boldsymbol{x}$ is the gradient of some scalar function $R(\boldsymbol{X})$:

$$\boldsymbol{x} = \nabla_X R. \tag{15}$$

To determine R, note that

$$\mathrm{d}R = \boldsymbol{x}\cdot\mathrm{d}\boldsymbol{X} = \mathrm{d}(\boldsymbol{x}\cdot\boldsymbol{X}) - \boldsymbol{X}\cdot\mathrm{d}\boldsymbol{x} = \mathrm{d}(\boldsymbol{x}\cdot\boldsymbol{X} - P)$$

so that R is given to within an additive constant by

$$R(\boldsymbol{X}) = \boldsymbol{x}\cdot\boldsymbol{X} - P(\boldsymbol{x}),$$

which is the expression for the *Legendre transform* between R and P. From (7) and (12) we infer that the motion is non-divergent in $\boldsymbol{X}$-space $\nabla_X \cdot \boldsymbol{u}_g = 0$, and therefore, being constrained to Z-surfaces since (9) has $\dot{Z} = 0$, is expressible in terms of a streamfunction Ψ by

$$(u_g, v_g, 0) = \frac{1}{f}\left(-\frac{\partial\Psi}{\partial Y}, \frac{\partial\Psi}{\partial X}, 0\right).$$

From (8) and (15)

$$u_g = f\left(\frac{\partial R}{\partial Y} - Y\right), \quad v_g = -f\left(\frac{\partial R}{\partial X} - X\right) \tag{16}$$

and hence the simplest form for Ψ is given by

$$\Psi = f^2(\tfrac{1}{2}(X^2 + Y^2) - R(X)). \tag{17}$$

The particular duality structure described above is but one realization of a quartet of Legendre transformations described in Chynoweth & Sewell (1989).

3.3. *A solution strategy*

The Lagrangian variable formulation is central to the use of this model in simulating flows with discontinuities. We outline the numerical method as discussed by Purser & Cullen (1987) and Cullen, Norbury & Purser (1991) to show that the above formulation leads to a natural phase-space description.

In order to solve the equations of motion we seek a solution strategy based on the dual space representation. At each timestep define a *distribution* of

$$\rho(X) = q^{-1}(x) = \det(\boldsymbol{Q}^{-1})$$

and solve the nonlinear elliptic or parabolic (Monge–Ampère) equation for R:

$$\det|\mathrm{Hes}(R)| = \rho \tag{18}$$

subject to the boundary conditions implied by the requirement

$$\nabla_X R = x \in D,$$

where D is the physical domain in x-space whose total measure μ must equal the integral of ρ in X-space. In general however, the boundary conditions imposed with respect to the domain in which the fluid motion takes place need to be translated into conditions on ∇R. For R convex and t a tangent vector to any line through the components, $t \cdot \nabla R$ must be monotonic as we move along the line. This implies that x is monotonic as X moves along the line. Internal component boundaries need continuity conditions, while the external boundary of the components corresponds to the boundary in the x-domain.

The next step is to use Ψ defined in (17) in terms of R given by (18) to update the conserved density ρ, according to the advection equation

$$\frac{\partial \rho}{\partial t} = -\boldsymbol{u}_g \cdot \nabla_X \rho \equiv \frac{1}{f}\frac{\partial(\rho, \Psi)}{\partial(X, Y)}. \tag{19}$$

Hence, the principle behind this strategy is to construct a solution via a sequence of time-independent elliptic problems. Thus we find the $\partial/\partial t$ operator on the left-hand side. It is commonplace to find in the literature the equations of motion written in terms of vorticity/streamfunction variables. The example that we have here could be thought of as a nonlinear version of those models. A detailed examination of the application of these techniques but with an analytical approach can be found in the work of Abarbanel *et al.* (1986) and Shepherd (1990). The model is three-dimensional despite the structure of the advection equation (19). As mentioned, at each timestep the Monge–Ampère equation is solved, providing the three-dimensional coupling between

the advection steps. It would therefore appear to be consistent to think of the model as essentially isentropic in the sense outlined in §3.2. This picture leads us to the main result of the paper (§4.5): a Hamiltonian approach to (19).

4. Hamiltonian structure

4.1. *Lagrangian kinematics*

Introduce Lagrangian particle labelling coordinates $(a, b, c) \equiv \boldsymbol{a}$, assigned at $\tau = 0$:

$$\boldsymbol{X}(\boldsymbol{a}, \tau) : \boldsymbol{X}(\boldsymbol{a}, 0) = \boldsymbol{a}, \quad \boldsymbol{a} \in \Gamma \subset \boldsymbol{R}^3.$$

Within this framework time is labelled τ; $\partial/\partial t$ means that $\boldsymbol{X}$ are held fixed while $\partial/\partial\tau$ means that $\boldsymbol{a}$ are held fixed. This gives $\partial/\partial\tau = \mathrm{D}/\mathrm{D}t$. The Lagrangian mass label has been chosen to incorporate the mass density $\tilde{\rho}$ via

$$\tilde{\rho}\,\mathrm{d}\boldsymbol{X} = \mathrm{d}\,(\text{mass}) = \mathrm{d}\boldsymbol{a}.$$

With this choice $\tilde{\rho}$ becomes the Jacobian of the inverse of the map $(\boldsymbol{a}) \mapsto (\boldsymbol{X})$, that is

$$\tilde{\rho}(\boldsymbol{X}, \tau) = \frac{\partial(\boldsymbol{a})}{\partial(\boldsymbol{X}(\boldsymbol{a}, \tau))}.$$

This map gives the fluid flow. In §3.1 (and in the literature referred to therein) the potential vorticity q, and its inverse ρ, were given by the ratio of the measures in $\boldsymbol{x}$-space and (dual) $\boldsymbol{X}$-space. Thereby it is implicit in that description that there is a constant mass-weighting given by

$$\frac{\partial(\boldsymbol{x})}{\partial(\boldsymbol{a})} \tag{20}$$

and set equal to unity. Thus $\tilde{\rho} \equiv \rho$. In this paper (and as in Shutts 1989) the dependence on the particle label coordinates will be explicit and henceforth we just refer to ρ. In addition, we take ρ to have compact support in Γ to avoid later difficulties with boundary conditions.

4.2. *Contact geometry and semi-geostrophic theory*

The subject of contact structures in classical mechanics has a long history. Carathèodory (1982) discusses the development of these ideas and Sewell & Roulstone (1993*b*) have elucidated the relationships between lift, contact and Legendre transformations, by studying a new family of contact transformations and illustrating these features by providing a number of explicit examples. Blumen (1981) appears to have been the first to identify the contact nature of the geostrophic momentum transformation and this is an example discussed by Sewell & Roulstone (1993*b*). More recently, Purser (1993) has described a general contact structure of the energy minimization approach to SG theory. In this subsection we describe the SG phase space in terms of contact geometry and show how these ideas enable us to write down variational derivatives for the functionals that arise in the Hamiltonian system. For clarity we confine our discussion to the lowest-dimensional case.

The basic notion of Legendre duality is depicted in figure 1. Using the notation from §3.2 and considering the lowest-dimensional case, we observe that a point in the (P, x)-space, which we shall denote by (P_α, x_α), defines a unique line R_α in (R, X)-space. Here α is to be interpreted as an independent parameter (that is, specifying α picks out a particular point and thus a particular line). The equation of this line is $R_\alpha = x_\alpha X - P_\alpha$.

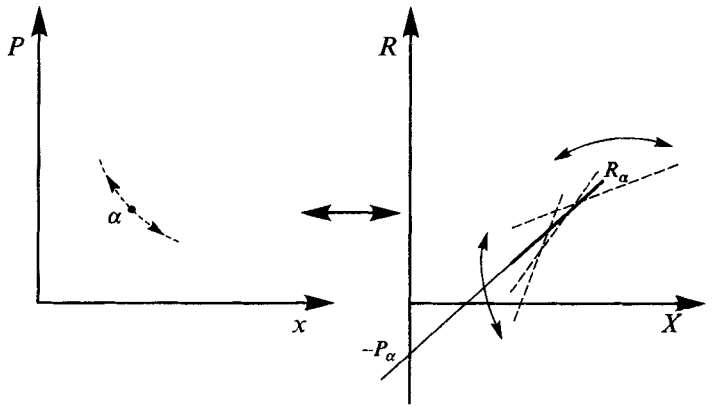

FIGURE 1. Legendre duality.

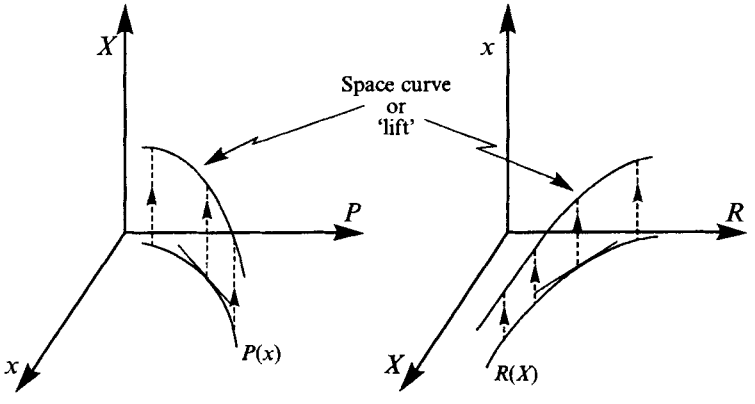

FIGURE 2. A representation of contact manifolds.

Thus $-P_\alpha$ is the intercept on the R-axis of the line which has a slope given by x_α. When the point is allowed to move (continuously) along a curve in (P, x)-space a family of lines envelope a curve in (R, X)-space. Each member of the family (labelled by the (continuous) variable α) of lines is in contact with this surface. We claim that a solution R of the Monge–Ampère equation (18) in our model is the *infimum* of a family of *generating lines* $\{R_\alpha\}$, each member of which is specified by a unique α. We remark that the dual solution surface P is the *supremum* of a family of generating lines each specified by a unique X (see also Purser 1993, p. 1451).

It is convenient to represent the duality depicted in figure 1 by augmenting the dimension of these spaces by one, which is achieved if we introduce the dual variables x, X into the spaces (R, X) and (P, x), respectively. This is illustrated in figure 2. We now impose the gradient relationships on the variables x and X: thus $\boldsymbol{x} = \nabla_X R$ and $\boldsymbol{X} = \nabla_x P$. The spaces depicted in figure 2 are realizations of so-called *contact manifolds* (see Arnol'd 1989, appendix 4; Burke 1985, pp. 99–108; Sewell & Roulstone 1993*b*). That is, to each curve in the (P, x)- or (R, X)-plane, there corresponds a space curve (known as a *lifted curve*) whose elevation above the base plane is given by the slope of the plane curve (its contact line) at the 'base point'. The mapping between the three variables (P, x, X) and (R, X, x) is a particular example of a lift transformation (see Sewell & Roulstone 1993*b* for details). We avoid the use of the word tangent here as it has a precise meaning in differential geometry and this is the reason why the adjective contact is used.

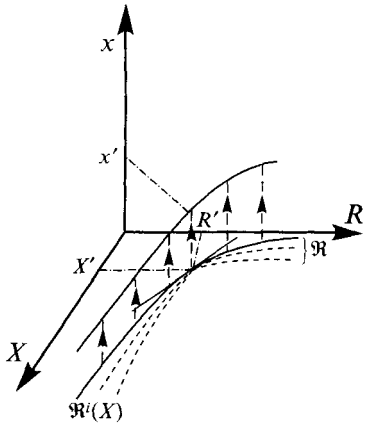

FIGURE 3. A family of solution surfaces.

Now we shall consider variations of the generating lines and of the 'solution' or 'envelope' surfaces. In terms of the parameter α, R_α on the envelope may be varied with respect to X_α (the contact point between the line R_α and the envelope), x_α, and P_α as follows:

$$\delta R_\alpha = \delta x_\alpha X_\alpha + x_\alpha \delta X_\alpha - \delta P_\alpha. \tag{21}$$

We treat the gradient relationship (13) as a constraint in (21). Hence, by restricting the otherwise free variations δx_α and δX_α by the identity

$$X_\alpha = \left.\frac{\partial P}{\partial x}\right|_\alpha,$$

we find that the variation of R_α for given α is (using, since $P = P(x)$, the exact differential relationship $\mathrm{d}P = (\partial P/\partial x)\,\mathrm{d}x$)

$$\delta R_\alpha = x_\alpha \delta X_\alpha. \tag{22}$$

This is a result to be used later in §4.3 on the canonical Hamiltonian formulation of the equations of motion. In that section we identify α with the continuous particle label a and thus we may interpret the above as the variation of the potential $R(X(a))$ at fixed a. (Note that (22) is consistent (in one dimension) with (15).)

We may choose to fix the coordinate X and then consider (21) with (13) treated as a constraint. If this is carried out, we obtain

$$\delta R = X\delta x - \delta P = 0. \tag{23}$$

We have dropped the suffix α because in our application later the coordinate X is treated as the independent parameter. This result will be used in §4.5 on the reduced Hamiltonian formulation.

Finally, we consider variations of the solution surface R at fixed X and x. That is, referring to figure 2, at a fixed point in the three space (R, X, x). To distinguish a solution surface in the (R, X)-plane from its envelope of lines we shall denote such a surface by $\mathfrak{R}(X)$. We are at liberty to consider a *family of solution surfaces* $\{\mathfrak{R}\}$, each member of which passes through the fixed point (R', X') (see figure 3), and they all share a common lift x' at that point. (A curve passing through the point (R', X') but with a different value x' would not share the common contact line of the members of the family that *do* share a common value of x' at that point. That is, the curve would

'cross' the others at (R', X'), as opposed to just meeting them at that point.) The ith member $\mathfrak{R}^i$ of $\{\mathfrak{R}\}$ has a second derivative ρ^i at (R', X', x') given by

$$\rho^i = \left.\frac{\mathrm{d}^2\mathfrak{R}^i}{\mathrm{d}X^2}\right|_{X=X'} \tag{24}$$

We may therefore consider variations $\delta\rho$ and interpret them as varying $\mathfrak{R}^i$-amongst $\{\mathfrak{R}\}$. In fact, the label i is the parameter which is to be varied in (24). In general, these variations $\delta\rho$ will be non-zero, even though the variation of the contact line at (R', X'), δR as given by (23), is identically zero. We shall be led to consider such higher-order quantities in the formulation of the reduced dynamics.

4.3. *Hamiltonian dynamics*

The equations of motion as written in (9) possess a canonical Hamiltonian structure in the sense of Definition 1. For any differentiable function R of X for which we write the relationship (15), we can define the Hamiltonian functional

$$\mathscr{H}[X] = \int_\Gamma \mathrm{d}\gamma\, \Psi(X) = f^2 \int_\Gamma \mathrm{d}\gamma\, (\tfrac{1}{2}(X^2(\boldsymbol{a}) + Y^2(\boldsymbol{a})) - R(X(\boldsymbol{a}))), \tag{25}$$

where $\mathrm{d}\gamma$ is the measure over the Lagrangian particle labelling coordinates and Ψ is defined by (17). This functional is autonomous in our basic solution strategy of §3.3.

We now write our equations of motion (9) in Lagrangian phase-space variables $\boldsymbol{X}$ as a Hamiltonian system by means of a Poisson bracket defined as follows.

PROPOSITION 1. *The equations of motion* (9) *take the form on surfaces of constant* Z,

$$\frac{\partial X}{\partial \tau} = \{\mathscr{H}, X\}_C, \qquad \frac{\partial Y}{\partial \tau} = \{\mathscr{H}, Y\}_C, \tag{26}$$

where $\{,\}_C$ *is given by*

$$\{\mathscr{F}, \mathscr{G}\}_C = \int_\Gamma \mathrm{d}\gamma \left(\frac{\delta\mathscr{F}}{\delta X(\boldsymbol{a})}\frac{\delta\mathscr{G}}{\delta Y(\boldsymbol{a})} - \frac{\delta\mathscr{F}}{\delta Y(\boldsymbol{a})}\frac{\delta\mathscr{G}}{\delta X(\boldsymbol{a})} \right).$$

Proof. Calculating the functional derivatives of (25) we find (see the Appendix, (A 2) and (A 3), and recalling our remarks in the previous section), noting (16), that

$$\frac{\delta\mathscr{H}}{\delta X} = f^2\left(X - \frac{\partial R}{\partial X}\right) = v_g f,$$

$$\frac{\delta\mathscr{H}}{\delta Y} = f^2\left(Y - \frac{\partial R}{\partial Y}\right) = -u_g f.$$

Using (A 4) (see Appendix), we obtain the equations of motion (26)

$$\frac{\partial X}{\partial \tau} = -\frac{1}{f}\frac{\delta\mathscr{H}}{\delta Y} = u_g,$$

$$\frac{\partial Y}{\partial \tau} = \frac{1}{f}\frac{\delta\mathscr{H}}{\delta X} = v_g,$$

in agreement with (9). □.

One can show directly in these variables that the $\boldsymbol{X}$-space mass density ρ, is conserved, since

$$\{\rho, \mathscr{H}\}_C = 0,$$

as follows. Express ρ in functional form via the definition of the δ-function

$$\rho(\boldsymbol{X}(\boldsymbol{a})) = \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma}\, \delta(\boldsymbol{a}-\hat{\boldsymbol{a}}) \frac{\partial(\hat{\boldsymbol{a}})}{\partial(\boldsymbol{X})}.$$

With this form we calculate the variation of ρ with respect to X^j:

$$\delta\rho = \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma}\, \delta(\boldsymbol{a}-\hat{\boldsymbol{a}})(-1) \frac{(\partial\hat{\boldsymbol{a}})}{\partial(\boldsymbol{X})} \frac{\partial\hat{a}^i}{\partial X^j} \frac{\partial\delta X^j}{\partial\hat{a}^i},$$

where we have used

$$\delta_x\left(\frac{\partial(\boldsymbol{x})}{\partial(\boldsymbol{y})}\right) = \frac{\partial(\boldsymbol{x})}{\partial(\boldsymbol{y})} \frac{\partial y^i}{\partial x^j} \frac{\partial\delta x^j}{\partial y^i}.$$

It is useful to note the identity

$$\frac{\partial}{\partial y^i}\left(\frac{\partial(\boldsymbol{x})}{\partial(\boldsymbol{y})} \frac{\partial y^i}{\partial x^j}\right) \equiv 0,$$

known as the *Piola identity*, which has been used frequently in these calculations. Thus $\delta\rho/\delta X^j$ is given by

$$\frac{\delta\rho(\boldsymbol{X}(\boldsymbol{a}))}{\delta X^j(\hat{\boldsymbol{a}})} = \frac{\partial(\boldsymbol{X})}{\partial(\hat{\boldsymbol{a}})} \frac{\partial\hat{a}^i}{\partial X^j} \frac{\partial}{\partial\hat{a}^i}(\rho(\boldsymbol{X}(\hat{a}))^2 \delta(\boldsymbol{a}-\hat{\boldsymbol{a}})).$$

Using the shorthand notation for Jacobian terms and derivatives

$$\hat{\rho} \equiv \frac{\partial(\hat{\boldsymbol{a}})}{\partial(\boldsymbol{X})}, \quad \hat{\rho}^i_{X^j} \equiv \frac{\partial\hat{a}^i}{\partial X^j},$$

we find that the expression for the Poisson bracket is

$$\begin{aligned}
\{\rho, \mathscr{H}\} &= f^2 \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma} \left(\hat{\rho}^{-1}(\hat{\rho})^i_X \frac{\partial}{\partial\hat{a}^i}(\hat{\rho}^2\delta(\boldsymbol{a}-\hat{\boldsymbol{a}}))(Y-y) - \hat{\rho}^{-1}(\hat{\rho})^i_Y \frac{\partial}{\partial\hat{a}^i}(\hat{\rho}^2\delta(\boldsymbol{a}-\hat{\boldsymbol{a}}))(X-x) \right) \\
&= f^2 \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma} \left((-1)\hat{\rho}^{-1}(\hat{\rho})^i_X \hat{\rho}^2 \delta(\boldsymbol{a}-\hat{\boldsymbol{a}}) \frac{\partial}{\partial\hat{a}^i}(Y-y) + \hat{\rho}^{-1}(\hat{\rho})^i_Y \hat{\rho}^2 \delta(\boldsymbol{a}-\hat{\boldsymbol{a}}) \frac{\partial}{\partial\hat{a}^i}(X-x) \right) \\
&= f^2 \left(-\rho \frac{\partial a^i}{\partial X} \frac{\partial}{\partial a^i}(Y-y) + \rho \frac{\partial a^i}{\partial Y} \frac{\partial}{\partial a^i}(X-x) \right) = 0,
\end{aligned}$$

where we have integrated by parts and discarded the boundary terms (because of compact supports) which are of the form $\hat{\rho}^2\delta(\boldsymbol{a}-\hat{\boldsymbol{a}})|_\partial$ to go from the first line to the second and used volume conservation in the last step.

Conservation of potential vorticity may be established using the procedure outlined above by noting that

$$q(\boldsymbol{a}) = \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma}\, \delta(\boldsymbol{a}-\hat{\boldsymbol{a}})\hat{\rho}^{-1}$$

and thus

$$\{q, \mathscr{H}\} = \int_{\hat{\Gamma}} \mathrm{d}\hat{\gamma} \frac{\delta q}{\delta\hat{\rho}} \{\hat{\rho}, \mathscr{H}\} = 0. \tag{27}$$

Alternatively one may demonstrate by using the Lagrangian, and Noether's theorem (Arnol'd 1989) to relate symmetries and conservation laws, that invariance of the

Hamiltonian system with respect to the allocation of particle labels leads to the conservation law for potential vorticity. This calculation was given by Shutts (1989).

4.4. *Symplectic structure and particle labels*

In this subsection we demonstrate that the conservation of vorticity may be related directly to the requirement of symplectic invariance with the choice of canonical coordinates X and Y. The reader who is unfamiliar with the notation used in symplectic geometry should refer to the Appendix for a brief introduction. Qualitatively, a Hamiltonian system possesses a volume element called the *symplectic form*, which we denote by Ω, and this is an invariant in the sense that its Lagrangian derivative along the flow vanishes. One can identify a natural dual to Ω, such that their inner produce is a scalar. We refer to this object as Θ, known as the *cosymplectic structure*. The Lagrangian derivative along the flow of the scalar vanishes and, following Schutz (1980, pp. 171–174, and references therein), we identify this behaviour with a conservation law. As in the case of the proof of the vanishing of the Poisson bracket (27) this approach is essentially Hamiltonian and thus the introduction of the Lagrangian functional is unnecessary. We express the form of the symplectic structure Ω, (A 8), and the cosymplectic structure Θ, (A 9), in terms of the canonical bases $\mathbf{d}l^i$ and $\partial/\partial l^i$. Here $l^i \leftrightarrow (a, b)$ are the particle labels on Z-surfaces by choosing $c = Z$, and the indices are taken to run over the values 1, 2. We write

$$\tfrac{1}{2}\langle \mathbf{d}X \wedge \mathbf{d}Y | \Theta \rangle = \int_\Gamma \mathrm{d}\gamma \left[\frac{\partial X}{\partial a}\frac{\partial Y}{\partial b} - a \leftrightarrow b \right], \tag{28}$$

where $\Theta = \frac{1}{2}J^{ij}(\partial/\partial l^i) \wedge (\partial/\partial l^j)$ and the notation $\leftrightarrow$ means interchange a and b. The above integrand may be written, using (8), as

$$\frac{\partial(x, y)}{\partial(a, b)} \frac{1}{f}\left[f + \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} + \frac{1}{f}\frac{\partial(u_g, v_g)}{\partial(x, y)} \right],$$

which is just the SG potential vorticity (as $\partial Z/\partial c = 1$). Equation (28) is, therefore, an important relationship between the dynamical invariant $q|_Z$ and the canonical kinematic structures Θ and Ω. The statement here is that the symplectic inner product on the particle labels (that corresponds to the dual-space mass density, or equivalently the inverse of the potential vorticity) is conserved. When we discuss numerical modelling of Hamiltonian systems in §5 the importance of preserving kinematic invariants will be emphasized. Thus at this stage we note that any consistent Hamiltonian numerical scheme for the SG model should be symplectic and potential vorticity preserving.

4.5. *A non-canonical Hamiltonian approach to the advection equation*

We make some brief remarks concerning the generalization of the results of §2 to *non-canonical systems*. Suppose that a system has densities $\mathscr{F}$, $\mathscr{G}$, which are functionals of canonical coordinates (observables) z^i. Thus the Poisson bracket for such a system is given by (A 7) with bilinear structure (2). It may be possible to write the Poisson brackets (and thus the equations of motion) in terms of non-canonical coordinates $\bar{z}^\alpha(\boldsymbol{e})$ where $\alpha < i$ (Greek indices will be used on functionals of Eulerian variables denoted $\boldsymbol{e}$), in the following manner:

$$\{\mathscr{F}, \mathscr{G}\}_E = \int \mathrm{d}\boldsymbol{e}_1 \int \mathrm{d}\boldsymbol{e}_2 \frac{\delta \mathscr{F}}{\delta \bar{z}_{1\alpha}} \{\bar{z}_{1\alpha}, \bar{z}_{2\beta}\}_C \frac{\delta \mathscr{G}}{\delta \bar{z}_{2\beta}}. \tag{29}$$

This will lead to an expression of the form (A 7) but with J^{ij} more general than (2). We will show that this structure leads to an appropriate generalization of (19).

In the following calculations the canonical coordinates z^i will be identified with the Lagrangian coordinates $X^i \leftrightarrow (X, Y)$, while the non-canonical variable $\bar{z}$ will be identified with $\rho(\bar{X}^i)$. It is the symmetry invariance (often referred to as gauge freedom in mathematical physics) with respect to the particle relabelling that enables the kinematics of the system to be described in terms of a single observable $\bar{z}$, together with a non-canonical structure. For further details of this point see §5 and Salmon (1988*a*).

We now seek a reduction of our canonical Hamiltonian system of Proposition 1. In order to proceed with a version of (29) we formulate the necessary Eulerian functionals. The essential idea is to describe the fluid motion using the mass-density ρ as the dynamical variable, which is a function of the variables X, Y, Z (which until now have been functions of the Lagrangian labelling coordinates). That is, the configuration space for the fluid motion is a region of $\boldsymbol{R}^3$ with coordinate surfaces labelled with the values of X, Y, Z. In the meteorological literature the (X = constant, Y = constant) coordinates are often referred to as *absolute momentum surfaces*. As discussed in §3.2, motion on the Z = constant surfaces is known as the isentropic description. To emphasize the new interpretation of the coordinates we will henceforth distinguish them with an overbar.

The phase-space mass density is written in functional form as

$$\rho(\bar{\boldsymbol{X}}) = \int \mathrm{d}\gamma\, \delta(\boldsymbol{a}(\bar{\boldsymbol{X}}) - \boldsymbol{a}) \frac{\partial(\boldsymbol{a})}{\partial(\bar{\boldsymbol{X}})}$$

$$= \int \mathrm{d}\gamma\, \delta(\bar{\boldsymbol{X}} - \boldsymbol{X}(\boldsymbol{a})) \tag{30}$$

$$= \int \mathrm{d}\boldsymbol{X} \rho(\boldsymbol{X})\, \delta(\bar{\boldsymbol{X}} - \boldsymbol{X}(\boldsymbol{a})), \tag{31}$$

where we have used the following result in (30):

$$\delta(f(x)) \frac{\partial f(x)}{\partial x} = \delta(x),$$

understood as

$$\int \delta(f)\, \mathrm{d}f = \int \delta(x)\, \mathrm{d}x = \int \delta(f(x)) \frac{\partial f}{\partial x} \mathrm{d}x.$$

With the functional form (30) we find the derivatives of ρ:

$$\frac{\delta \rho(\bar{\boldsymbol{X}})}{\delta X^i(\boldsymbol{a})} = \frac{\partial}{\partial \bar{X}^i} \delta(\bar{\boldsymbol{X}} - \boldsymbol{X}(\boldsymbol{a})), \tag{32}$$

and similarly (31) yields

$$\frac{\delta \rho(\bar{\boldsymbol{X}})}{\delta \rho(\boldsymbol{X})} = \delta(\bar{\boldsymbol{X}} - \boldsymbol{X}). \tag{33}$$

We use our previous Hamiltonian (25), but evaluated in phase space solely as a functional of ρ. The Hamiltonian is given by

$$\mathscr{H}[\rho] = f^2 \int \mathrm{d}\bar{\boldsymbol{X}} \rho(\bar{\boldsymbol{X}}) (\tfrac{1}{2}(\bar{X}^2 + \bar{Y}^2) - R(\bar{\boldsymbol{X}})), \tag{34}$$

where $R(\bar{X})$ is defined by ρ from (18) and the boundary conditions. To calculate the functional derivative consider the variation

$$\delta\mathscr{H}[\rho] = f^2 \int \mathrm{d}\bar{X}\,\delta\rho(\bar{X})\,(\tfrac{1}{2}(\bar{X}^2+\bar{Y}^2)-R(\bar{X})) + f^2 \int \mathrm{d}\bar{X}\,\rho(\bar{X})\,\delta(\tfrac{1}{2}(\bar{X}^2+\bar{Y}^2)-R(\bar{X})) \quad (35)$$

and examine the second term on the right-hand side. We consider variations at fixed $(\bar{X}, \bar{Y}, \bar{Z})$ (cf. figures 2 and 3), and thus the terms involving the dual space variables $(\bar{X}, \bar{Y})$ do not contribute. The remaining terms may be written as (cf. the material following (15) and (23))

$$\delta R = \delta(\boldsymbol{x}\cdot\boldsymbol{X}-P) = \delta\boldsymbol{x}\cdot\boldsymbol{X} - \frac{\partial P}{\partial \boldsymbol{x}}\cdot\delta\boldsymbol{x} \equiv \delta\boldsymbol{x}\cdot\boldsymbol{X} - \boldsymbol{X}\cdot\delta\boldsymbol{x} = 0.$$

Thus from the first term on the right-hand side of (35), we have

$$\frac{\delta\mathscr{H}}{\delta\rho} = f^2(\tfrac{1}{2}(\bar{X}^2+\bar{Y}^2)-R). \quad (36)$$

Hence we *can* treat ρ as the sole dynamical variable if we consider the variations here as a higher-dimensional version of those one-dimensional examples interpreted geometrically in §4.2 (i.e. (23) and (24)). In $2+1$ dimensions ρ is related to the *extrinsic curvature* κ, of the solution surface R,

$$\kappa = \frac{1}{\rho[1+X^2+Y^2]^2}$$

(see Purser & Cullen 1987 for further details). Note that, unlike the corresponding procedure in two-dimensional perfect fluids (Shepherd 1990; Marsden & Weinstein 1983), no integration by parts is necessary in evaluating the functional derivative of $\mathscr{H}$. It is the contact geometry that distinguishes this Hamiltonian system from that of the perfect fluid.

We note in passing that we can relate this calculation to the energy minimization principle. Consider variations for which the energy (10) is stationary, which is a consequence of the Monge–Ampère relation (18) in the definition of $\mathscr{H}$. The integrals of the quadratic terms are conserved under the displacements; thus we are left with

$$-f^2\delta\int_\Gamma \mathrm{d}\gamma\,(xX+yY+zZ), \quad (37)$$

and this is the expression for the energy functional that we *require* to be stationary for evolving equilibrium solutions and hence have taken in our definition of $\mathscr{H}$ indirectly by means of the Monge–Ampère equation.

We may now formulate the Hamiltonian structure of (19).

PROPOSITION 2. *The equation of motion* (19) *may be written in the following Hamiltonian form*

$$\frac{\partial\rho(\bar{X})}{\partial t} = \{\mathscr{H}, \rho(\bar{X})\}_E$$

with

$$\{\mathscr{F}, \mathscr{G}\}_E = \int_\Gamma \mathrm{d}\bar{X}\,\frac{\delta\mathscr{G}}{\delta\rho(\bar{X})}\left(\frac{\partial\left(\rho(\bar{X}), \dfrac{\delta\mathscr{F}}{\delta\rho(\bar{X})}\right)}{\partial(\bar{X},\bar{Y})}\right),$$

and Hamiltonian given by (34). *Γ is taken to be a union of bounded components in $\bar{X}$ in which the defining relation for R* (18) *holds.*

Proof. We integrate on constant-Z cross-sections of Γ. Thus we choose $c = Z$. (Note that ρ is then given by

$$\rho = \frac{\partial(a,b,c)}{\partial(x,y,Z)}\frac{\partial(x,y,Z)}{\partial(X,Y,Z)} = \frac{\partial(a,b,c)}{\partial(x,y,Z)}\det\begin{vmatrix} \dfrac{\partial^2 R}{\partial X^2} & \dfrac{\partial^2 R}{\partial X \partial Y} \\ \dfrac{\partial^2 R}{\partial Y \partial X} & \dfrac{\partial^2 R}{\partial Y^2} \end{vmatrix},$$

where we have used the (T, R) pair from Chynoweth & Sewell (1989), and the result (15).)

From (29) we have

$$\{\mathscr{F},\mathscr{G}\}_E = \int \mathrm{d}\bar{X}_1 \int \mathrm{d}\bar{X}_2 \frac{\delta\mathscr{F}}{\delta\rho(\bar{X}_1)}\{\rho(\bar{X}_1),\rho(\bar{X}_2)\}_C \frac{\delta\mathscr{G}}{\delta\rho(\bar{X}_2)}. \tag{38}$$

Substituting for $\delta\rho(\bar{X})/\delta X$, the canonical form of the bracket is

$$\begin{aligned}\{\rho(\bar{X}_1),\rho(\bar{X}_2)\}_C &= \int \mathrm{d}X \rho(X)\left(\frac{\partial\delta(\bar{X}_1 - X)}{\partial X}\frac{\partial\delta(\bar{X}_2 - X)}{\partial Y} - X \leftrightarrow Y\right) \\ &= -\frac{\partial\delta(\bar{X}_2 - \bar{X}_1)}{\partial\bar{Y}_1}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{X}_1} + \frac{\partial\delta(\bar{X}_2 - \bar{X}_1)}{\partial\bar{X}_1}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{Y}_1}.\end{aligned} \tag{39}$$

Substituting (39) into (38) gives

$$\begin{aligned}&\int \mathrm{d}\bar{X}_1 \int \mathrm{d}\bar{X}_2 \frac{\delta\mathscr{F}}{\delta\rho(\bar{X}_1)}\frac{\delta\mathscr{G}}{\delta\rho(\bar{X}_2)}\left(\frac{\partial\delta(\bar{X}_2 - \bar{X}_1)}{\partial\bar{X}_1}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{Y}_1} - \frac{\partial\delta(\bar{X}_2 - \bar{X}_1)}{\partial\bar{Y}_1}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{X}_1}\right) \\ &= \int \mathrm{d}\bar{X}_1 \int \mathrm{d}\bar{X}_2 \delta(\bar{X}_2 - \bar{X}_1)\frac{\delta\mathscr{G}}{\delta\rho(\bar{X}_2)}\left[\frac{\partial}{\partial\bar{Y}_1}\left(\frac{\delta\mathscr{F}}{\delta\rho(\bar{X}_1)}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{X}_1}\right) - \frac{\partial}{\partial\bar{X}_1}\left(\frac{\delta\mathscr{F}}{\delta\rho(\bar{X}_1)}\frac{\partial\rho(\bar{X}_1)}{\partial\bar{Y}_1}\right)\right] \\ &= \int_\Gamma \mathrm{d}\bar{X}\frac{\delta\mathscr{G}}{\delta\rho(\bar{X})}\left(\frac{\partial\left(\rho(\bar{X}), \dfrac{\delta\mathscr{F}}{\delta\rho(\bar{X})}\right)}{\partial(\bar{X},\bar{Y})}\right).\end{aligned} \tag{40}$$

As in the case of the canonical set of equations in §4.3, the boundary terms arising in the integration by parts that leads to (40), have the form $\delta(\bar{X}_2 - \bar{X}_1)|_\partial$, and are discarded. We may identify $\mathscr{F}$ with $\mathscr{H}$ and $\mathscr{G}$ with ρ and evaluate the bracket to obtain (19). The functional derivatives we require are given by (33) and (36). From (40), we evaluate

$$f^2 \int \mathrm{d}\bar{X}\,\mathrm{d}\bar{Y}\,\delta(\bar{X} - X)\left(\frac{\partial\rho(\bar{X})}{\partial\bar{X}}\frac{\partial(\frac{1}{2}(\bar{X}^2 + \bar{Y}^2) - R)}{\partial\bar{Y}} - \bar{X} \leftrightarrow \bar{Y}\right).$$

Integrating out the variables $\bar{X}$ leaves us immediately with

$$f^2\left(\frac{\partial\rho}{\partial X}\frac{\partial(\frac{1}{2}(X^2 + Y^2) - R)}{\partial Y} - X \leftrightarrow Y\right)$$

and thus, using (15) and (16), we have

$$-fu_g\frac{\partial\rho}{\partial X} - fv_g\frac{\partial\rho}{\partial Y}, \tag{41}$$

which is precisely the right-hand side of the advection equation (19). □

We note that the form of (40) means that the bracket of functionals depending on ρ alone will be in *involution* with respect to $\{,\}_E$ (i.e. their Poisson brackets all vanish). We identify the bilinear skew adjoint operator $\mathscr{J}$ from (40), and write it

$$\mathscr{J}^{\bullet\circ} = \circ\frac{\partial(\rho, \bullet)}{\partial(\bar{X}, \bar{Y})}. \tag{42}$$

5. Applications to the geometric model

There now exist numerical algorithms for integrating the SG equations. The geometric ideas originating from Cullen (1983) and developed in Cullen & Purser (1984) were subsequently used in the formulation of the geometric model (Chynoweth 1987). Recently (R. J. Purser, private communication) a much improved and efficient version of this model has been produced thus enabling an increased resolution in the modelling of atmospheric flows. The basic algorithm was described in §3.3, and §4 describes the (formal) Hamiltonian structure.

It is a goal of much current effort to develop algorithms that capture the details of the Hamiltonian structure in numerical models. These algorithms are referred to as *symplectic integrators*. A symplectic integrator is an evolutionary finite-difference algorithm which has the property that each iteration n is given by a canonical transformation (also known as a symplectic transformation) of the phase space. For details see, for example, Channell & Scovel (1990), Marsden (1992) and references contained therein. With regard to the geometric model, we have learned from §4 that the essential ρ-structure is non-canonical and therefore it is natural to ask how one may integrate such systems. This leads us to the theory of reduction and *reduced symplectic integrators*.

Reduction is the rigorous approach to the type of process involved in (29). The salient feature of reduction is that the essential dynamics of a system may be described in terms of trajectories on a manifold which has a lower dimension than the dimension of the problem's original phase space. Suppose there is a group G of symmetry transformations of $\mathscr{P}$ that transform $\mathscr{P}$ to itself by a canonical transformation. We use the symmetry group to generate a vector-valued conserved quantity denoted $\boldsymbol{M}$ and called the *moment map*. If the conserved quantity takes a given value then we consider the set of all points in phase space where $\boldsymbol{M}$ has that value. We call this set the *level set for* $\boldsymbol{M}$. Assign the value μ to $\boldsymbol{M}$. Then the reduced phase space, denoted $\mathscr{P}_\mu$, is constructed as follows.

DEFINITION 2. *$\mathscr{P}_\mu$ is the μ-level set for $\boldsymbol{M}$ on which any two points that can be transformed one to another by a group transformation are identified.*

The *reduction theorem* (Marsden & Weinstein 1974; also Marsden 1992) states that

> $\mathscr{P}_\mu$ inherits the symplectic (or Poisson bracket) structure from that of $\mathscr{P}$, so it can be used as a new phase space. Also, the dynamical trajectories of the Hamiltonian $\mathscr{H}$ on $\mathscr{P}$ determine new trajectories on the reduced phase space.

For further discussion see Marsden & Weinstein (1974, 1983), or for a less abstract approach with discussion of the bracket used in (29) see Salmon (1988*a*).

The basic structure of the geometric model is symplectic (in the sense of proposition 2) and the conserved quantity, associated with the reduction to the non-canonical form used in the solution strategy, is ρ. In order for the numerics to be symplectic the

timestepping algorithm used should preserve the canonical structure (2) in the following way. Let $\boldsymbol{A}$ be the matrix that leads from one timestep to another and corresponds to the Jacobian matrix

$$\boldsymbol{A} = \frac{\partial(X_1, Y_1)}{\partial(X_0, Y_0)},$$

where the subscripts 0 and 1 on the variables represent the functions at the two time levels. Then we require

$$\boldsymbol{A}\boldsymbol{J}\boldsymbol{A}^T = \boldsymbol{J}, \tag{43}$$

with $\boldsymbol{J}$ given by (2).

In order for the non-canonical algorithm to preserve the corresponding generalized form $\boldsymbol{J}_\rho$,

$$\boldsymbol{J}_\rho^{\bullet\circ} = \circ \det \begin{vmatrix} \dfrac{\partial \rho}{\partial X} & \dfrac{\partial \bullet}{\partial X} \\ \dfrac{\partial \rho}{\partial Y} & \dfrac{\partial \bullet}{\partial Y} \end{vmatrix} \tag{44}$$

(as outlined in §4.5), the geometric method should conserve ρ. As indicated by Chynoweth & Sewell (1991) a variety of methods may be used to generate the time–Δt map and therefore (43) needs to be imposed for each algorithm; conserving ρ in reduced algorithms is then sufficient for (44) to be correctly defined.

6. Summary

Salmon (1985) identified a consistent set of approximations to the shallow-water equations that were implemented within Hamilton's Principle and defined a canonical Lagrangian form of the SG equations together with the appropriate Hamiltonian functional. The material presented in this paper exemplifies the geometric emphasis of modern theoretical mechanics. When studying Hamiltonian dynamics from a geometric perspective, it is essential, when formulating a model, to distinguish the features that depend on the Hamiltonian function from those that depend only on the properties of the structure of phase space. We have now set the SG equations within a formal but explicit canonical Hamiltonian framework. We have also derived a reduced non-canonical Hamiltonian representation of the phase-space mass density/potential vorticity advection equation. The equation of motion for the phase-space mass density is written in terms of the 'vorticity bracket' which has many other applications (see for example Marsden & Weinstein 1983). However, the important difference between the other systems that may be described in terms of this bracket (e.g. the two-dimensional Euler equations) and SG theory is that the dynamics as encoded in the Hamiltonian used here is based upon the Monge–Ampère equation, or equivalently the energy minimization and Legendre duality. It is not necessary in this model to integrate by parts to go from a velocity representation of the equations to the vorticity representation. Instead the Legendre duality plays a key role as described in the section on contact geometry.

Understanding this framework is essential to a numerical model that seeks to capture *exactly* the Hamiltonian nature of the mechanics, via say a symplectic (or canonical) integrating algorithm.

The authors would like to thank Professor M. J. Sewell for many useful discussions

and J.N. acknowledges financial support from a US Navy contract (No. N00014/91/J/1163).

Appendix. Hamiltonian methods and symplectic geometry

A.1. *Preliminaries*

The equations of fluid dynamics whether written in terms of Eulerian coordinates or Lagrangian labels represent infinite-dimensional dynamical systems. In what follows we will be content with the practice of functional analysis rather than the absolute rigour necessary to prove our application in modern measure spaces. The reason for adopting such an approach is twofold. First the subject matter: Lagrangian SG theory is applicable to flows with discontinuities and the problem of extending the usual analysis to such solutions presents a considerable challenge. Secondly, our primary aim is to establish the existence of a structure that in practice will be studied within the context of a finite-dimensional numerical model. (All the analytical difficulties arise in taking the infinite-dimensional limit.) However, it is appropriate to remind ourselves of the basic functional methods and *formal* structure at this stage.

We begin with a discussion of our typical evolution problem. Consider the following system of autonomous evolution (in time t) equations for a fluid with the configuration represented by a generalized coordinate $\boldsymbol{x}$, and velocity $\boldsymbol{u}$:

$$\frac{\partial u^i(t,\boldsymbol{x})}{\partial t} = F^i(\boldsymbol{u},\boldsymbol{x}).$$

Here, $\boldsymbol{x}\in\Gamma\subset\boldsymbol{R}^n$ for some integer n. The F^i are general nonlinear (partial) differential or integral operators on $\boldsymbol{u}$. We consider the F^i that arise in (26). Here we will not be concerned with necessary conditions for the existence and uniqueness of solutions, but suppose solutions exist and are elements of some vector space $\mathscr{V}$ (usually referred to as phase space), over $\boldsymbol{R}^n$, that is equipped with the inner product

$$\langle f|g\rangle = \int_\Gamma f(\boldsymbol{x})\,g(\boldsymbol{x})\,\mathrm{d}\gamma \quad \text{for} \quad f,g\in\mathscr{V}.$$

In Hamiltonian fluid dynamics certain functionals arise from *densities* defined on $\mathscr{V}$. To this end we define another vector space Ξ, over $\boldsymbol{R}^n$, of differentiable functionals that have the form, for the density $G\in\mathscr{V}$,

$$\mathscr{G}[\boldsymbol{u}] = \int_\Gamma G(\boldsymbol{u},\boldsymbol{x})\,\mathrm{d}\gamma,$$

where $\mathscr{G}$ is an operator on $\mathscr{V}$. Functional differentiation is often defined by the usual Gateaux derivative

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\mathscr{G}[u^i+\epsilon\varpi]\bigg|_{\epsilon=0} = \left\langle \frac{\delta\mathscr{G}}{\delta u^i}\,\middle|\,\varpi\right\rangle, \tag{A 1}$$

where the variation δ is taken with respect to functions ϖ that satisfy some boundary condition on $\partial\Gamma$. Equation (A 1) defines the functional derivatives $\delta\mathscr{G}/\delta u^i$, which, in general, is nonlinear in the u^i.

Let us now work through some of the essential steps of the proof of Proposition 1. The Hamiltonian is given by

$$\mathscr{H}[X] = f^2\int_\Gamma \mathrm{d}\gamma\,(\tfrac{1}{2}(X^2+Y^2)-R(X))$$

and the variation is

$$\delta\mathscr{H} = f^2 \int_\Gamma \mathrm{d}\gamma \left(X^i \delta X^i - \frac{\partial R}{\partial X^i} \delta X^i \right), \tag{A 2}$$

where δX^i is interpreted in the sense of (22) in §4.2. Fixing γ in this calculation is equivalent (except for dimensionality) to fixing α in §4.2. The expression (A 2) yields the derivatives: for example on constant-Z surfaces

$$\frac{\delta\mathscr{H}}{\delta X} = f^2 \left(X - \frac{\partial R}{\partial X} \right). \tag{A 3}$$

If we write the Lagrangian variables as

$$X^i(\boldsymbol{a}) = \int_{\Gamma'} \mathrm{d}\gamma' \, \delta(\boldsymbol{a} - \boldsymbol{a}') \, X^i(\boldsymbol{a}'),$$

then we find

$$\frac{\delta X^i(\boldsymbol{a})}{\delta X^j(\boldsymbol{a}')} = \delta^{ij} \delta(\boldsymbol{a} - \boldsymbol{a}'), \tag{A 4}$$

where δ^{ij} is the Kronecker delta. This may be used to establish the following result:

$$\begin{aligned} \{X(\boldsymbol{a}), \mathscr{H}\}_C &= f^2 \int_{\Gamma'} \mathrm{d}\gamma' \, \delta(\boldsymbol{a} - \boldsymbol{a}') \left(Y(\boldsymbol{a}') - \frac{\partial R(\boldsymbol{X}(\boldsymbol{a}'))}{\partial Y(\boldsymbol{a}')} \right) \\ &= f^2 \left(Y(\boldsymbol{a}) - \frac{\partial R(\boldsymbol{X}(\boldsymbol{a}))}{\partial Y(\boldsymbol{a})} \right). \end{aligned}$$

A.2. *Structure of phase space*

The material presented in this subsection is a review of the structure of differential equations of the Hamiltonian type. We have used these results in §4.4.

The evolution of a classical continuous physical system (for example, a fluid) can be described in terms of a curve $z(t)$ (parameterized by some continuous variable t) in phase space. There are two structures that enter the definition of phase space and encode the *kinematics* of the system. First, the phase or the configuration space is described by a differentiable manifold $\mathscr{P}$ with an appropriate topology. Secondly, there is an antisymmetric bilinear operation $\{,\}$ on the space of C^∞ functions from $\mathscr{P}$ to $\boldsymbol{R}$, denoted by $\ell(\mathscr{P})$, such that the operation (defined for all $\mathscr{F}$, $\mathscr{H} \in \ell(\mathscr{P})$)

$$L_{\mathscr{H}}(\mathscr{F}) : \ell(\mathscr{P}) \otimes \ell(\mathscr{P}) \mapsto \ell(\mathscr{P})$$

written as

$$L_{\mathscr{H}}(\mathscr{F}) = \{\mathscr{F}, \mathscr{H}\} \tag{A 5}$$

where $\{\mathscr{F}, \mathscr{G}\}$ is defined for any functionals $(\mathscr{F}, \mathscr{G})$, by the following properties:

(i) $\{\mathscr{F}, \mathscr{G}\}$ is bilinear in $\mathscr{F}$ and $\mathscr{G}$.

(ii) $\{\mathscr{F}, \mathscr{G}\} = -\{\mathscr{G}, \mathscr{F}\}$ for $\mathscr{F}, \mathscr{G} \in \ell(\mathscr{P})$.

(iii) the Jacobi condition,

$$\{\mathscr{E}, \{\mathscr{F}, \mathscr{G}\}\} + \{\mathscr{F}, \{\mathscr{G}, \mathscr{E}\}\} + \{\mathscr{G}, \{\mathscr{E}, \mathscr{F}\}\} = 0 \text{ for every } \mathscr{E}, \mathscr{F}, \mathscr{G}, \in \ell(\mathscr{P}).$$

(iv) the Leibnitz' rule,

$$\{\mathscr{E}\mathscr{F}, \mathscr{G}\} = \mathscr{E}\{\mathscr{F}, \mathscr{G}\} + \mathscr{F}\{\mathscr{E}, \mathscr{G}\}.$$

The bracket operation makes Ξ a Lie algebra. That is, conditions (i)–(iii) make $(\ell(\mathscr{P}), \{,\})$ into a Lie algebra. Conditions (i)–(iv) define the pair $(\mathscr{P}, \{,\})$ as a *Poisson*

manifold; that is, a manifold with a Lie algebra structure on the space of differentiable functions. This structure is preserved in the following sense. The vector field generated by the operation $L_{\mathscr{H}}$, will be denoted $X_{\mathscr{H}}$. Then the causal maps φ, for the flow of $X_{\mathscr{H}}$ (Abraham & Marsden 1985, p. 60), satisfy

$$\{\mathscr{F}\circ\varphi, \mathscr{G}\circ\varphi\} = \{\mathscr{F}, \mathscr{G}\}\circ\varphi.$$

The flow satisfies the differential equations determined by $L_{\mathscr{H}}$:

$$\dot{z}(t) = X_{\mathscr{H}} = L_{\mathscr{H}}(x) = \{z, \mathscr{H}\},$$

where $z(t)$ represents an observable. This flow is called the *Hamiltonian flow*, and $\mathscr{H}$ is called its *generating function* or *Hamiltonian*. The Hamiltonian is the map $\mathscr{H}: \mathscr{P} \mapsto \boldsymbol{R}$, and encodes the *dynamics*. The evolution equations may be written in terms of this structure as

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathscr{F}\circ\varphi) = \{\mathscr{F}, \mathscr{H}\}\circ\varphi$$

or, as the time derivative commutes with the time-t map,

$$\frac{\mathrm{d}\mathscr{F}}{\mathrm{d}t} = \{\mathscr{F}, \mathscr{H}\}. \tag{A 6}$$

The formal definition of a Hamiltonian system is given in §2, we now examine the structure of the system developed thus far. The conditions (i) and (iv) imply the existence of a tensor $\boldsymbol{J}$ on $\mathscr{P}$, assigning to each point $z \in \mathscr{P}$ a linear map between the cotangent space (denoted $\mathscr{T}^*$) and the tangent space ($\mathscr{T}$)

$$\sharp: \mathscr{T}_z^* \mathscr{P} \mapsto \mathscr{T}_z \mathscr{P}$$

such that

$$\{\mathscr{F}, \mathscr{H}\}(z) = \langle \mathbf{d}\mathscr{F}(z), \boldsymbol{J}(z)\cdot\mathbf{d}\mathscr{H}(z)\rangle.$$

Here $\langle , \rangle$ denotes the natural pairing between vectors and covectors (forms), $\mathbf{d}$ the exterior derivative (Arnol'd 1989) generating forms from functions $\mathscr{F}$, and the dot is contraction. As we are concerned with a tensor operator we have indicated the existence of the map at each point z. Because of (ii), $\boldsymbol{J}(z)$ is antisymmetric. Let z^i, $i = 1, \ldots, n$ denote the coordinates on $\mathscr{P}$, then (A 5) becomes

$$\{F, H\} = \frac{\partial F}{\partial z_i} J^{ij} \frac{\partial H}{\partial z_j}$$

or in the infinite-dimensional case

$$\{\mathscr{F}, \mathscr{G}\} = \int_\Gamma \frac{\delta \mathscr{F}}{\delta z_i} J^{ij} \frac{\delta \mathscr{G}}{\delta z_j} \mathrm{d}\gamma = \left\langle \frac{\delta \mathscr{F}}{\delta z_i} \middle| J^{ij} \frac{\delta \mathscr{G}}{\delta z_j} \right\rangle. \tag{A 7}$$

DEFINITION 3. LET *$\mathscr{P}$ be a manifold and Ω a two-form, on $\mathscr{P}$. The pair $(\mathscr{P}, \Omega)$ is called a symplectic manifold if Ω satisfies*: (*i*) $\mathbf{d}\Omega = 0$ (*i.e. Ω is closed*), (*ii*) *Ω is non-degenerate.*

(See Arnol'd 1989 and Olver 1982 for definition and further discussion.) If $(\mathscr{P}, \Omega)$ is a symplectic manifold, define the Poisson bracket operation $\{,\}$ by

$$\{\mathscr{F}, \mathscr{H}\} = \Omega(X_{\mathscr{F}}, X_{\mathscr{H}})$$

where $X_{\mathscr{F}}$ is the vector field generated by the function $\mathscr{F}$ on $\mathscr{P}$. This construction makes $(\mathscr{P}, \{,\})$ into a Poisson manifold. Thus, in other words:

PROPOSITION 3. *Every symplectic manifold is Poisson.*

The converse is not true; for example the zero (degenerate) bracket makes any manifold Poisson. In our application we need to distinguish between Poisson and symplectic manifolds as the former arise naturally in the Eulerian kinematics of fluids.

The symplectic two-form Ω may be expressed in terms of a (local) coordinate basis of one-forms $\mathbf{d}z^i$:

$$\Omega = \tfrac{1}{2}K_{ij}\,\mathbf{d}z^i \wedge \mathbf{d}z^j, \tag{A 8}$$

where $\wedge$ is the exterior product. $\boldsymbol{K}$ is the inverse of $\boldsymbol{J}$, $\boldsymbol{K} = \boldsymbol{J}^{-1}$ (see Olver 1982 for further details). The cosymplectic tensor field written in terms of the vector basis $\partial/\partial z^i$ is

$$\Theta = \tfrac{1}{2}J^{ij}\frac{\partial}{\partial z^i} \wedge \frac{\partial}{\partial z^j} \tag{A 9}$$

and may be regarded as the fundamental object determining a Poisson structure, rather than the Poisson bracket, which is easily recovered from Θ by

$$\{\mathscr{F}, \mathscr{G}\} = \langle \mathbf{d}F \wedge \mathbf{d}G \,|\, \Theta \rangle,$$

where $\mathbf{d}F = \Omega(X_{\mathscr{F}})$ (and similarly for $\mathbf{d}G$), is the one-form which corresponds, under the map $\flat : \mathscr{T}\mathscr{P} \mapsto \mathscr{T}^*\mathscr{P}$, determined by Ω, to the vector field $X_{\mathscr{F}}$. Thus we recover (A 7).

REFERENCES

ABARBANEL, H. D. I. & HOLM, D. D. 1987 Nonlinear stability analysis of inviscid flows in three dimensions: Incompressible fluids and barotropic fluids. *Phys. Fluids* **30**, 3369–3382.

ABARBANEL, H. D. I., HOLM, D. D., MARSDEN, J. E. & RATIU, T. 1986 Nonlinear stability analysis of stratified fluid equilibria. *Phil. Trans. R. Soc. Lond.* A **318**, 349–409.

ABRAHAM, R. & MARSDEN, J. E. 1985 *Foundations of Mechanics*. Addison Wesley.

ARNOL'D, V. I. 1989 *Mathematical Methods of Classical Mechanics*. Springer

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

BURKE, W. L. 1985 *Applied Differential Geometry*. Cambridge University Press.

CARATHÈODORY, C. 1992 *Calculus of Variations*, 2nd edn. Chelsea.

CHANNELL, S. & SCOVEL, C. 1990 Symplectic integration of Hamiltonian systems. *Nonlinearity* **3**, 231–259.

CHYNOWETH, S. 1987 The semi-geostrophic equations and the Legendre transform. PhD thesis, University of Reading, UK.

CHYNOWETH, S. & SEWELL, M. J. 1989 Dual variables in semi-geostrophic theory. *Proc. R. Soc. Lond.* A **424**, 155–186.

CHYNOWETH, S. & SEWELL, M. J. 1991 A concise derivation of the semigeostrophic equations. *Q. J. R. Met. Soc.* **117**, 1109–1128.

CULLEN, M. J. P. 1983 Solutions to a model of a front forced by deformation. *Q. J. R. Met. Soc.* **109**, 565–573.

CULLEN, M. J. P., NORBURY, J. & PURSER, R. J. 1991 Generalized Lagrangian solutions for atmospheric and oceanic flows. *SIAM J. Appl. Maths.* **51**, 20–31.

CULLEN, M. J. P., NORBURY, J., PURSER, R. J. & SHUTTS, G. J. 1987 Modelling the quasi-equilibrium dynamics of the atmosphere. *Q. J. R. Met. Soc.* **113**, 735–757.

CULLEN, M. J. P. & PURSER, R. J. 1984 An extended Lagrangian theory of semi-geostrophic frontogenesis. *J. Atmos. Sci.* **41**, 1477–1497.

FENG, K. 1991 The Hamiltonian way for computing Hamiltonian dynamics. In *The State of the Art of Industrial and Applied Mathematics* (ed. R. Spigler). Kluwer.

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

HOSKINS, B. J. & BRETHERTON, F. P. 1972 Atmospheric frontogenesis models: mathematical formulation and solution. *J. Atmos. Sci.* **29**, 11–37.

HOSKINS, B. J. & DRAGHICI, I. 1977 The forcing of ageostrophic motion according to the semi-geostrophic equations and in an isentropic coordinate model. *J. Atmos. Sci.* **34**, 1859–1867.

MARSDEN, J. E. 1992 *Lectures on Mechanics*. Cambridge University Press.

MARSDEN, J. E. & WEINSTEIN, A. 1974 Reduction of symplectic manifolds with symmetry. *Rep. Math. Phys.* **5**, 121–130.

MARSDEN, J. E. & WEINSTEIN, A. 1983 Coadjoint orbits, vortices and Clebsch variables for incompressible fluids. *Physica* **7D**, 305–323.

MCLACHLAN, R. I. 1993 Explicit Lie–Poisson integration and the Euler equations. *Phys. Rev. Lett.* **71**, 3043–3046.

OLVER, P. J. 1982 A nonlinear Hamiltonian structure for the Euler equations. *J. Math. Anal. Appl.* **89**, 233–250.

PURSER, R. J. 1988 Variational aspects of semi-geostrophic theory. *Met. O.* 11 *Sci. Note* 5. Meteorological Office, UK.

PURSER, R. J. 1993 Contact transformations and Hamiltonian dynamics in generalized semi-geostrophic theories. *J. Atmos. Sci.* **50**, 1449–1468.

PURSER, R. J. & CULLEN, M. J. P. 1987 A duality principle in semi-geostrophic theory. *J. Atmos. Sci.* **44**, 3449–3468.

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

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

SALMON, R. 1988*b* Semi-geostrophic theory as a Dirac bracket projection. *J. Fluid Mech.* **196**, 345–358.

SCHUTZ, B. F. 1980 *Geometrical Methods of Mathematical Physics*. Cambridge University Press.

SEWELL, M. J. & ROULSTONE, I. 1993*a* Anatomy of the canonical transformation. *Phil. Trans. R. Soc. Lond.* A **345**, 577–598.

SEWELL, M. J. & ROULSTONE, I. 1993*b* Families of contact transformations. *Forecasting Res. Div. Sci. Notes* 20. Meteorological Office, UK.

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.

SHUTTS, G. J. & CULLEN, M. J. P. 1987 Parcel stability and its relation to semi-geostrophic theory. *J. Atmos. Sci.* **44**, 1318–1330.