

---

# The mathematical structure of theories of semigeostrophic type

I. Roulstone and M. J. Sewell

*Phil. Trans. R. Soc. Lond. A* 1997 **355**, 2489-2517

doi: 10.1098/rsta.1997.0144

---

## Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click [here](#)

---

To subscribe to *Phil. Trans. R. Soc. Lond. A* go to: <http://rsta.royalsocietypublishing.org/subscriptions>

---

# The mathematical structure of theories of semigeostrophic type

BY I. ROULSTONE AND M. J. SEWELL†

*Numerical Weather Prediction Development Division, Meteorological Office,  
London Road, Bracknell, Berkshire RG12 2SZ, UK*

## Contents

	PAGE
1. Introduction	2490
2. Thermodynamics and constitutive properties	2492
(a) Atmospheric theory	2492
(b) Shallow-water theory	2494
3. Transformation theory	2495
(a) Jacobian fluxes	2495
(b) Lift transformations	2496
(c) Legendre transformation	2497
(d) Examples of lift transformations	2498
(e) Augmented lift and Legendre transformations	2500
4. Mass balance	2502
(a) Shallow-water theory	2503
(b) Atmospheric theory	2503
5. Dynamics	2504
(a) Exact equations	2504
(b) Semigeostrophic approximation	2505
(c) Résumé of semigeostrophic equations	2505
6. Generalized semigeostrophic equations	2505
(a) Generalized Hamiltonian equations	2505
(b) Shallow-water theory on an $f$ -plane	2506
(c) Generalized augmented Hamiltonian equations	2507
(d) Atmospheric theory on an $f$ -plane	2508
(e) Potential vorticity and the Monge–Ampère equation	2509
7. Variable Coriolis parameter	2510
(a) Planetary semigeostrophic equations	2510
(b) Pseudo-plane approximations	2512
(c) Other lift transformations	2513
8. Concluding remarks	2515
References	2515

In the past twenty years, semigeostrophic equations have become a prominent model for describing certain atmospheric motions on a synoptic scale, including the presence

† Permanent address: Department of Mathematics, University of Reading, Whiteknights, P.O. Box 220, Reading RG6 6AF, UK.

of fronts. Theoretical studies of them have revealed Hamiltonian features, and novel numerical methods, motivated by the need to improve weather forecasts, have been explored. A shallow-water theory analogue has been used as a paradigm for some aspects.

This paper sets out to uncover the mathematical structure of the semigeostrophic equations that has been essential to finding solutions and developing numerical techniques. We study the shallow-water and atmospheric theories side-by-side, and we introduce a generalized form which encapsulates the differences between them. When the Coriolis parameter,  $f$ , is a constant, it is found that a lift transformation is at the heart of the theory, and the consequences of this are developed. When  $f$  is not a constant, the role of the lift transformation is, in some respects, looser; we explore the extent to which it still offers a worthwhile guide. In particular, it can be viewed as motivating a generalization of the geostrophic momentum transformation for planetary semigeostrophic equations.

The paper is broadly self-contained, and it takes account of several different strands in the existing literature.

---

## 1. Introduction

The semigeostrophic equations are an approximation to Newton's second law for a rotating fluid, in which the acceleration is replaced by the time derivative (following the particle) of the so-called geostrophic velocity. The hydrostatic approximation is used in the vertical direction. These equations are regarded as a good approximation for certain two- and three-dimensional atmospheric motions on a synoptic scale, such as warm and cold fronts, and solutions can be continued in time beyond the point of the formation of a discontinuity modelling a front.

A novel geometrical technique for numerically integrating the semigeostrophic equations has been used to study not only the formation of fronts, but also land-sea breezes, flow over mountains, and convection. For the construction of this model thus far it has been essential to have the gradient transformation property of the semigeostrophic equations, and a convexity principle which states that the solution represents a sequence of convectively and inertially stable equilibrium states of the atmosphere. A self-contained connected account of the transformation properties of these equations has not appeared before, despite the current interest in Hamiltonian structure, and in generalizations of semigeostrophic theory to incorporate a wider class of flow regimes. One of the aims of this paper is to present a unified account of the mathematical structure of these equations and to indicate possible future developments of the theory that share the invariant structure of the basic equations.

Physical arguments, and in some respects rather detailed ones, are naturally required to justify this 'geostrophic momentum' approximation, as it was called by Hoskins (1975, §3) in his account of them for the atmospheric case. He seems to have been the first to use the adjective 'semigeostrophic' to describe the resulting equations (contained in his (10)) for constant Coriolis parameter, and he remarked that they had also been introduced by Eliassen (1948) and Fjørtoft (1962).

Hoskins (1975, §4) and (in one dimension) Eliassen (1962) also introduced a change of independent variables, from cartesian coordinates of the particle to 'geostrophic coordinates'. Hoskins showed that their time derivatives following the particle are the

geostrophic velocity components. It was the system of equations expressed, after use of the geostrophic momentum approximation, in terms of these new variables that Hoskins originally called semigeostrophic. Conservation of energy applies. For adiabatic flow, potential temperature is conserved. When the Boussinesq approximation is also made, potential vorticity is conserved too.

In the course of constructing solution strategies for such equations, the work of Cullen and co-workers on the so-called geometric method (see, for example, Cullen & Purser 1984; Purser & Cullen 1987; Shutts & Cullen 1987) began to reveal features reminiscent of Hamiltonian mechanics, such as energy-minimization techniques and duality, in addition to conservation properties. Hamilton's equations themselves are not stated, and the connection with the first derivation of the semigeostrophic equations for shallow-water theory from Hamilton's principle by Salmon (1983, 1985) remained obscure. In the atmospheric context the connection with Hamilton's equations was viewed at first more as folklore rather than fact. Chynoweth *et al.* (1988) and Chynoweth & Sewell (1989, 1991), made explicit the connection of the semigeostrophic equations with the Legendre dual transformation and Hamilton's equations. Purser (1988) related these approaches to that of Salmon (1983, 1985). The Hamiltonian is an unknown function of the geostrophic coordinates. It needs to be found by solving a Monge–Ampère problem, in which the potential vorticity reciprocal is viewed as an assigned function of the geostrophic coordinates at each time, in the manner reviewed by Cullen *et al.* (1991).

Sewell & Roulstone (1993, 1994) began a study of certain basic aspects of transformation theory in Hamiltonian dynamics, namely canonical, contact and lift transformations. This was with the long-term object of contributing to an understanding of the semigeostrophic equations, having noted a remark by Blumen (1981) about a contact transformation property of them. It emerged that *the geostrophic coordinate transformation is part of a lift transformation.* (The contact property is secondary.)

In this paper we use that fact as a basis for a study of the mathematical structure of semigeostrophic equations in several contexts. The object is to complement the physical reasons for their adoption with the mathematical properties which flow from the hypothesis of a lift transformation as a starting point. We introduce concrete and abstract notations to run side-by-side, as a way of deliberately detaching the mathematics from the physics, and then bringing them together again. In particular, this allows us to define and work with generalized semigeostrophic equations. The method conveys a very clear indication of how generic features of  $f$ -plane semigeostrophic theory follow as a consequence of the lift transformation. It also allows us to go some way towards an understanding of what to expect about the structure of non- $f$ -plane theories. Among other results, we show how the planetary semigeostrophic equations of Shutts (1989) follow from a non- $f$ -plane generalization of the geostrophic coordinate transformation. Non- $f$ -plane shallow-water theories were studied by Roulstone & Sewell (1996) using a different approach.

There is a sense in which the approach *via* lift transformations proposed here starts at the opposite end from that represented by hypothesizing a Hamilton principle, and exploring its consequences, which is a more traditional method. Of course, ours is not the only approach which may be informative, but it has not been exploited before.

The plan of this paper is as follows. In §2 we describe the thermodynamic properties of an air parcel or fluid particle in any motion, to which is added the hydrostatic and adiabatic assumptions. In §3 the lift and Legendre transformation properties of

the equations are described in detail, including the different requirements of two- and three-dimensional motion. This serves as a basis for the discussion of more general theories in §§ 6 and 7. The topics of compressibility and incompressibility for shallow-water and atmospheric versions of semigeostrophic theory are reviewed in § 4, as a consequence of mass conservation. This enables us to present features of both two- and three-dimensional dynamics side-by-side in the remainder of the paper. A résumé of the exact and approximated Newtonian equations of motion are given in § 5. Then, in § 6, we describe how transformation theory unifies the description of momentum balance given in § 5 with the Hamiltonian structure of the semigeostrophic equations. In particular, we show that the lift transformation is at the heart of the theory in both the shallow-water and atmospheric cases. In § 7 we demonstrate how some variable Coriolis parameter models may be constructed using a lift transformation. Concluding remarks are made in § 8.

## 2. Thermodynamics and constitutive properties

In this section we describe properties which a representative particle is deemed to have in any motion.

There are two distinct contexts which we have particularly in mind from the outset, in both of which the semigeostrophic approximation is often applied. The definition of a material particle is different in the two cases.

One context is the mid-latitude motion of the atmosphere on a synoptic scale, which includes phenomena of direct interest to the meteorologist. The other context is the motion of shallow water over a rotating bed, which is often regarded as a useful mathematical paradigm of the atmospheric problem, but which lacks the ability to describe certain vertical effects, and which also replaces incompressibility by compressibility.

### (a) *Atmospheric theory*

In this context the particle is a parcel of air which is free to move in three dimensions. A cartesian reference frame fixed to the Earth is used as the standpoint from which to describe particle motions. At time  $t$  the particle has height  $h$  above a local tangent plane to the Earth's surface, and horizontal distances  $x$  and  $y$  (for example, east and north) measured parallel to that plane. Particle labels  $a$ ,  $b$  and  $c$  are defined as the values of  $x$ ,  $y$  and  $h$  at some particular time, say  $t = 0$ , and they identify the particle at all subsequent times.

A motion of the atmosphere is defined by three equations

$$x = x(a, b, c, t), \quad y = y(a, b, c, t), \quad h = h(a, b, c, t), \quad (2.1)$$

relating the values of  $x$ ,  $y$  and  $h$  on the left to functions, usually unknown in advance, of  $a$ ,  $b$ ,  $c$  and  $t$  on the right.

Standard thermodynamics endows every particle, for all time and for every motion, with the same function  $H(\eta, p)$  representing enthalpy per unit mass, such that

$$\tau = \frac{\partial H}{\partial \eta}, \quad \frac{1}{\rho} = \frac{\partial H}{\partial p}, \quad (2.2)$$

where  $\eta$ ,  $p$ ,  $\tau$  and  $\rho$  are entropy, pressure, absolute temperature and density, respectively. Thus the functions  $H(\eta, p)$ ,  $\tau(\eta, p)$  and  $\rho(\eta, p)$  do not depend explicitly on  $t$  or on  $a$ ,  $b$ ,  $c$  as well, so that their values can only vary when  $\eta$  or  $p$  do so.

For dry air the ideal gas model is usually assumed, in which  $p = k\rho^\gamma$  emerges from (2.2)<sub>2</sub> when

$$H(\eta, p) = \sigma(\eta)\zeta(p) \quad (2.3)$$

is the particular product function in which

$$\zeta(p) = (\gamma/(\gamma - 1))p^{(\gamma-1)/\gamma}, \quad \sigma(\eta) = k^{1/\gamma}, \quad k(\eta) = R \exp(\eta - d)/c_v,$$

where  $\gamma$ ,  $R$ ,  $d$ ,  $c_v$  and  $c_p$  are given constants related by  $R = c_p - c_v$  and  $\gamma = c_p/c_v$ .

For many years, e.g. since the work of Bjerknes in 1910, the pressure, or a function of it, has been a more convenient vertical coordinate for some purposes than the true particle height  $h$ . To define such a pseudo-height here, we first select arbitrary reference values  $\eta_0$  and  $p_0$  (say) of  $\eta$  and  $p$ , which imply reference values  $H(\eta_0, p_0) = H_0$  (say) and  $\rho(\eta_0, p_0) = \rho_0$  (say) from (2.2). Let  $g$  denote the combined acceleration due to gravity and the vertical effect of the Earth's rotation. Let  $z$ , defined by

$$z = \left(1 - \frac{H(\eta, p)}{H(\eta, p_0)}\right) \frac{H_0}{g} = \left(1 - \left(\frac{p}{p_0}\right)^{(\gamma-1)/\gamma}\right) \frac{\gamma p_0}{(\gamma - 1)\rho_0 g}, \quad (2.4)$$

denote the pseudo-height, in terms of a general enthalpy function in the first expression (Chynoweth & Sewell 1991), and of the ideal gas in the second (Hoskins 1971).

In terms of  $z$  we can interpret  $p_0$  as the pressure at the bottom of the atmosphere, because  $z = 0$  when  $p = p_0$ . If  $H(\eta, 0) = 0$  as in the ideal gas, the pseudo-height of the atmosphere is  $H_0/g$  or  $\gamma p_0/(\gamma - 1)\rho_0 g$ , because  $z$  has that value when  $p = 0$ .

The height of the atmosphere is small compared with the radius of the Earth, so that, notwithstanding the possibility of gravity anomalies, we adopt the approximation that  $g$  is an absolute (positive) constant.

Then during any motion (2.1) with (2.3),  $z$  can only vary when  $p$  does so. In general,  $p$  will be a function  $p(x, y, h, t)$  which is unknown in advance. Vertical acceleration is neglected, so that vertical momentum balance is achieved by satisfying the hydrostatic equation

$$\frac{\partial p}{\partial h} + \rho g = 0. \quad (2.5)$$

This ensures that the function  $z(x, y, h, t)$  implied by (2.4)<sub>2</sub> has the property

$$\frac{\partial z}{\partial h} = \frac{\sigma_0}{\sigma} > 0, \quad (2.6)$$

whenever (2.3) applies with  $\sigma(\eta) > 0$ , where  $\sigma_0 = \sigma(\eta_0)$ , in particular for the ideal gas.

Therefore the inverse function theorem always permits the true height  $h$  to be expressed as a function  $h(x, y, z, t)$ , and the pseudo-height  $z$  can subsequently be used as an alternative vertical coordinate. The geopotential  $gh = \phi$  (say) per unit mass can then be expressed as a function  $\phi(x, y, z, t)$  such that

$$\frac{\partial \phi}{\partial z} = \frac{g\sigma}{\sigma_0}. \quad (2.7)$$

This is an alternative statement of vertical momentum balance, in the sense that (2.5) and (2.7) each imply the other, using the thermodynamic equations in (2.2)<sub>2</sub> and (2.3). It is a relation between  $\eta$  and  $x, y, z$  and  $t$ . For the ideal gas, (2.7) can be written explicitly as

$$\eta - \eta_0 = c_p \ln \left( \frac{1}{g} \frac{\partial \phi}{\partial z} \right), \quad (2.8)$$



which expresses  $\eta$  as a function  $\eta(x, y, z, t)$ .

The absolute temperature function  $\tau(\eta, p)$  defined by (2.2)<sub>1</sub> can be used to define a so-called potential temperature  $\tau(\eta, p_0) = \theta(\eta)$  (say), which is the temperature which the particle would have if it retained its current value of entropy, but replaced its pressure by the reference value  $p_0$ . Let  $\theta_0 = \theta(\eta_0) = \tau(\eta_0, p_0)$  denote the reference temperature. The exponential in the ideal gas law implies  $\sigma/\sigma_0 = \theta/\theta_0$ , in which case (2.7) can also be written

$$\frac{\partial\phi}{\partial z} = \frac{g\theta}{\theta_0}. \quad (2.9)$$

We shall use the superposed dot to denote time differentiation following the particle in a motion (2.1), i.e. partial differentiation with respect to  $t$  of any function of  $a, b, c$  and  $t$ .

Attention will be confined to motions (2.1) which are adiabatic. The energy balance equation for such flows is

$$\dot{\eta} = 0. \quad (2.10)$$

This means that  $\eta$  is a constant for each particle, and so are  $\sigma, \theta$  and  $\partial\phi/\partial z$ . These constants can vary from particle to particle, however, for example according to functions such as  $\eta(a, b, c)$ . When such a variation is monotonic, that parameter is available for use as an alternative particle label.

#### (b) *Shallow-water theory*

In this context the *ab initio* definition of a fluid particle which is free to move in three dimensions is transmuted, as described below, into a definition of a particle which is only free to move in two horizontal dimensions. We use the same cartesian reference frame as in §2*a*, with  $x$ - and  $y$ -axes horizontal and fixed to the Earth, but a different letter  $n$  (for ‘normal’) instead of  $h$  to denote the vertical coordinate. There is no thermodynamics in the previous sense, although a formal analogue of it is available.

The simplest model assigns the same constant density  $\rho$  to every particle, so that the fluid is incompressible (‘water’) and unstratified. The flow takes place between a bed  $n = 0$  (this assumption of a flat unshifting bed could be relaxed) and a free surface  $n = s(x, y, t)$ , where the function on the right is unknown in advance.

Vertical momentum balance is again expressed by the hydrostatic equation  $\partial p/\partial n + \rho g = 0$ , in which the unknown function  $p(x, y, n, t)$  is the pressure, now defined mechanically rather than thermodynamically, and  $g$  is the previous constant.

Vertical integration of the hydrostatic equation, at each fixed horizontal  $(x, y)$  station, gives  $p = \rho g(s - n)$  when  $p = 0$  at  $n = s$  is assumed (by analogy with the definition of the pseudo-height  $H_0/g$  of the atmosphere in §2*a*). A second vertical integration expresses the vertically aggregated pressure as

$$\int_0^s p \, dn = \frac{1}{2} \rho g s^2.$$

This well known analogue of  $p = k\rho^\gamma$  in §2*a* is the starting point for some exact cusp and swallowtail geometrical properties (Sewell & Porter 1980; Broad *et al.* 1994).

Shallow-water theory, when regarded as a lowest order approximation in a formal perturbation expansion of the governing equations (Stoker 1957, p. 30), is completed by the assumption that the fluid velocity is horizontal and independent of  $n$ .

Such a motion is defined by two equations

$$x = x(a, b, t), \quad y = y(a, b, t), \quad (2.11)$$

relating the values of  $x$  and  $y$  on the left to functions, usually unknown in advance, of  $a$ ,  $b$  and  $t$  on the right. Here the labels  $a$  and  $b$  are the values of  $x$  and  $y$  at  $t = 0$ . They identify the transmuted particle as representing the motion of all those real particles which belong to the same material vertical line, and supposedly located at a height  $\frac{1}{2}s = h$  (say, to be notationally consistent with §2*a* above), so that the geopotential per unit mass is  $gh = \phi$  again, which can be expressed as a function  $\phi(x, y, t)$ . The energy balance differential equation analogous to (2.10) is satisfied identically in this case (Broad *et al.* 1992, equation (39); see also Broad *et al.* 1997).

### 3. Transformation theory

In this section we step back from the particularities of the two contexts just introduced, and we describe some mathematical properties of transformations of variables which are valid more generally. We consider coordinate transformations, lift transformations (which include contact transformations) and Legendre transformations. The purpose of such transformations is to convert one system of partial differential equations into another, and more amenable, form. Hamilton's equations will emerge as an example here. Classic references representing the transformation theory approach include the books of Forsyth (1890, 1906) and Carathéodory (1982).

Let  $x_i(a_j, t)$ , for  $i = 1, 2, \dots, n$  and  $j = 1, 2, \dots, n$ , denote  $n$  differentiable scalar functions of  $n + 1$  scalar variables  $a_j$  and  $t$ . The equations

$$x_i = x_i(a_j, t), \quad (3.1)$$

can be regarded as a  $t$ -dependent smooth mapping, from a space spanned by the  $a_j$  to another space spanned by the  $x_i$ . Assume that (3.1) have a local inverse expressible as

$$a_i = a_i(x_j, t). \quad (3.2)$$

With the additional assumption

$$a_i = x_i(a_j, 0) \quad \text{or} \quad x_i = a_i(x_j, 0), \quad (3.3)$$

we shall regard the mapping (3.1) or (3.2) as a smooth kinematically possible *motion* of material particles whose initial position (at time  $t = 0$ ) has coordinates  $a_j$ , and whose current position (at time  $t$ ) has coordinates  $x_i$ . We also use the superposed dot to denote time differentiation following the particle in such motions, i.e. partial differentiation with respect to  $t$  of any function of the  $a_j$  and  $t$ . For example,  $\dot{x}_i$  denotes velocity.

Of course we shall be most concerned with cases of planar ( $n = 2$ ) and three-dimensional ( $n = 3$ ) motion. The easiest case is when the  $a_i$  and the  $x_i$  are both cartesian coordinates measured with respect to the same set of reference axes, as illustrated by (2.11) with  $a, b, x, y$  replacing  $a_1, a_2, x_1, x_2$ , and by (2.1) with  $a, b, c, x, y, h$  replacing  $a_1, a_2, a_3, x_1, x_2, x_3$ , respectively. Other replacements can be envisaged, however.

#### (a) Jacobian fluxes

We choose  $n = 3$  in this subsection. Let  $v_i$ , for  $i = 1, 2, 3$ , denote three differentiable scalar functions of the  $x_i$ , and not yet necessarily velocity components. Employing



the convention of summation over repeated suffixes, we define the operator

$$D = \frac{\partial}{\partial t} + v_i \frac{\partial}{\partial x_i}, \quad (3.4)$$

which can be applied to any differentiable function of the  $x_i$  and  $t$ .

Let  $b_i$ , for  $i = 1, 2, 3$ , denote another set of three variables, which may be distinct from the  $x_i$  and the  $a_i$ . Let  $b_i(x_j, t)$  denote three differentiable scalar functions of the four variables  $x_j$  and  $t$ , so that

$$b_i = b_i(x_j, t), \quad \text{with } x_i = x_i(b_j, t), \quad (3.5)$$

defines another  $t$ -dependent smooth mapping together with its inverse. The  $b_i$  need be neither Eulerian nor Lagrangian coordinates, but could be either. Denote the Jacobians of (3.5) by

$$J = \frac{\partial(b_i)}{\partial(x_j)}, \quad \text{with } \frac{1}{J} = \frac{\partial(x_i)}{\partial(b_j)}. \quad (3.6)$$

The local inversion is unique where  $J$  is finite and non-zero. If this fails at isolated places ('singularities'), there may still be a multivalued global inversion. Chynoweth & Sewell (1991, equation (B3)) showed that

$$\frac{DJ}{J} = \frac{\partial Db_i}{\partial b_i} - \frac{\partial v_i}{\partial x_i}. \quad (3.7)$$

On the left  $D$  is applied to the function  $J(x_j, t)$  which emerges from (3.5)<sub>1</sub>. On the right  $D$  is applied to (3.5)<sub>1</sub>, after which (3.5)<sub>2</sub> is used to express the  $Db_i$  as functions of the  $b_j$  and  $t$  before the  $b_i$ -differentiation.

The  $v_i$  may or may not depend on  $t$  as well, but such dependence has not been used thus far. As stated above, the  $v_i$  may or may not have the interpretation of velocity components. Thus  $D$  may or may not signify time differentiation following a particle, although this is one possible interpretation which is common in applications.

To obtain this most familiar special case we choose  $v_i = \dot{x}_i$ , which implies  $D\lambda = \dot{\lambda}$  for any  $\lambda$  by the chain rule. If we also choose  $b_i = a_i$ , so that  $Db_i = \dot{a}_i = 0$  by definition, then (3.7) becomes

$$\frac{\dot{J}}{J} = -\frac{\partial \dot{x}_i}{\partial x_i}. \quad (3.8)$$

Different choices of the  $b_i$  in (4.6) and (6.19) yield other examples (4.7) and (6.21), respectively, of (3.7).

#### (b) Lift transformations

Sewell & Roulstone (1994, theorem 8) derived a family of lift transformations which encompasses ones associated with both the Legendre transformation and the geostrophic transformation of meteorology. Here we develop some consequences of particular members of that family.

To begin with, we return to the case of general  $n$  in (3.1), but we restrict the range of suffixes to  $i = 1, 2, \dots, m \leq n$ . Later we need to distinguish between the cases  $m = n$  ( $= 2$  in shallow-water theory) and  $m < n$  ( $m = 2$  and  $n = 3$  in atmospheric flows).

**Theorem 3.1.** *Let  $x_i, \psi, z_i$  be one set of  $2m + 1$  scalar variables, and let  $X_i, \Psi, Z_i$  be a second set of  $2m + 1$  scalar variables, each for  $i = 1, 2, \dots, m$ . Let  $c_{ij}$  be the*

typical component of an  $m \times m$  symmetric non-singular matrix, which might depend on the  $x_i, \psi, z_i$  (and perhaps on other assignable parameters, for example  $t$ ), but not separately on the  $X_i, \Psi, Z_i$ , and whose inverse has typical component  $c_{ij}^{-1}$ .

Then the transformation

$$X_i = c_{ij}^{-1}x_j + z_i, \quad \Psi = \psi + \frac{1}{2}c_{ij}z_iz_j, \quad Z_i = c_{ij}z_j, \quad (3.9)$$

is a mapping from the space spanned by the first set of variables into the space spanned by the second set, with differentials which satisfy

$$d\Psi - Z_i dX_i = d\psi - z_i dx_i + \left(\frac{1}{2}z_i + c_{ik}^{-1}x_k\right)z_j dc_{ij}, \quad (3.10)$$

when all the quantities are allowed to vary.

*Proof.* This is a straightforward manipulation, which uses the Kronecker delta  $\delta_{ij} = c_{ik}c_{kj}^{-1}$ . ■

**Theorem 3.2.** When

$$\left(\frac{1}{2}z_i + c_{ik}^{-1}x_k\right)z_j dc_{ij} = 0, \quad (3.11)$$

and in particular when the  $c_{ij}$  are held fixed, so that  $dc_{ij} = 0$ ,

$$d\Psi - Z_i dX_i = d\psi - z_i dx_i. \quad (3.12)$$

*Proof.* This is immediate from (3.10). ■

Sewell & Roulstone (1994, theorem 8) show that (3.9) with (3.11) is not the only transformation which satisfies (3.12), but it is adequate for our needs here when the Coriolis parameter  $f$  is constant. It will emerge that the special choice

$$c_{ij} = \mu\delta_{ij}, \quad (3.13)$$

with  $m = 2$  and with constant scalar  $\mu$  (especially  $\mu = 1$ ) is particularly appropriate to  $f$ -plane semigeostrophic theory, but for the present we retain the more general choice of  $c_{ij}$  to exhibit the mathematical structure of the theory.

We call (3.9) with (3.11) a *lift transformation* in the sense that (3.12) is satisfied. It is not necessary that both sides of (3.12) be zero, although that possible consequence is the motivation for the terminology.

It will turn out to be the analytical structure of the lift transformation, rather than its geometrical attributes, which is more important in guiding our theory in this paper.

### (c) Legendre transformation

For simplicity suppose, in this and the next subsection, that  $m = n$ . Introduce the new hypotheses that there exist relations

$$\psi = \psi(x_i, t), \quad c_{ij} = c_{ij}(t), \quad (3.14)$$

between the stated variables, where the functions on the right may be either known, or unknown and to be sought from further considerations. Define the new function

$$P(x_i, t) = \frac{1}{2}c_{ij}^{-1}x_ix_j + \psi. \quad (3.15)$$

A straightforward manipulation using (3.9)<sub>1</sub> and (3.9)<sub>2</sub> shows that, in value,

$$\Psi = P - x_iX_i + \frac{1}{2}c_{ij}X_iX_j. \quad (3.16)$$

**Theorem 3.3.** Introduce the new hypothesis that

$$z_i = \frac{\partial\psi}{\partial x_i}. \quad (3.17)$$

Then (3.9)<sub>1</sub> can be written

$$X_i = \frac{\partial P}{\partial x_i}. \quad (3.18)$$

This gradient version is the starting point for generating a Legendre transformation to a dual function  $S(X_i, t)$  with the properties

$$x_i = \frac{\partial S}{\partial X_i}, \quad P + S = x_i X_i, \quad \frac{\partial P}{\partial t} + \frac{\partial S}{\partial t} = 0. \quad (3.19)$$

Then there exists a function

$$\Psi(X_i, t) = \frac{1}{2} c_{ij} X_i X_j - S(X_i, t) \quad (3.20)$$

such that (3.9)<sub>3</sub> implies

$$\frac{\partial \Psi}{\partial X_i} = Z_i. \quad (3.21)$$

*Proof.* Inserting (3.17) into (3.9)<sub>1</sub> gives (3.18) using (3.15). Applying the inverse function theorem, wherever it is locally valid, to (3.18) allows the  $x_i$  to be expressed as functions  $x_i(X_i, t)$  at each fixed  $t$ . The chain rule shows that this has the form (3.19)<sub>1</sub> where  $S(X_i, t) = X_i x_i(X_j, t) - P(x_i(X_j, t), t)$ . Then (3.20) follows from (3.16), and (3.21) from (3.9). Releasing the passive  $t$  and using the chain rule again gives (3.19)<sub>3</sub>. ■

Local singularities, i.e. where the inverse function theorem locally fails, can often be allowed. They can permit multivaluedness of either  $P(x_i, t)$  or  $S(X_i, t)$ , including swallowtail forms like those in figure 2, as illustrated by Sewell (1987), and by Chynoweth & Sewell (1989) in the context of atmospheric fronts.

It can also be shown that, at each  $t$ , wherever  $P(x_i)$  is convex, so is  $S(X_i)$ , and vice versa.

#### (d) Examples of lift transformations

It is helpful at this point to give a geometrical picture of what is being achieved by marrying lift and Legendre transformations in the way established in theorem 3.3.

The simplest case is when  $t$  is not explicitly present in the functions on the right of (3.14) and (3.17). Then a surface  $\psi = \psi(x_i)$  lying in  $x_i, \psi$  space is lifted into another surface in  $x_i, \psi, z_i$  space by defining  $z_i = \partial\psi/\partial x_i$ , as shown (for  $n = 1$ ) in figure 1, and having the property  $d\psi - z_i dx_i = 0$ . (Intersection of lifts then represents contact of their projections.)

The combined effect of the lift and Legendre transformations is to produce a surface  $\Psi = \Psi(X_i)$ , lying in  $X_i, \Psi$  space, which is itself lifted into another surface in  $X_i, \Psi, Z_i$  space with the property  $Z_i = \partial\Psi/\partial X_i$ , so that  $d\Psi - Z_i dX_i = 0$ . (In lift geometry there is a sense in which a hypothesis like  $\psi = \psi(x_i)$  confines discussion to regular points only (see (5) of Sewell & Roulstone 1994) where auxiliary parameters are not needed to describe the surface. Non-regular points where  $d\psi = z_i dx_i$  still holds can be of intrinsic interest, but we shall not digress to discuss them here.)

Any kinematically possible motion (3.1) will generate, for each particle, a curve and

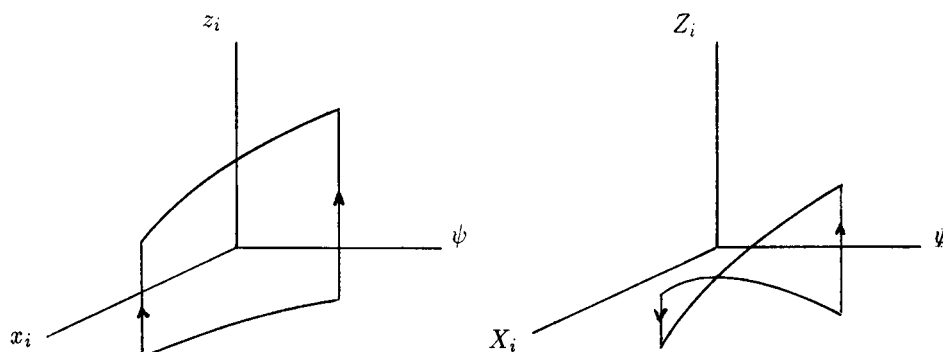


Figure 1. Example of a lift transformation.

a lifted curve, each parametrized by  $t$  on the respective surfaces in  $x_i, \psi$  and  $x_i, \psi, z_i$  spaces. It will map into another curve and another lifted curve on the corresponding surfaces in  $X_i, \Psi$  and  $X_i, \Psi, Z_i$  spaces.

When  $t$  is explicitly present on the right of (3.14) and (3.17), there is a *family* of surfaces  $\psi = \psi(x_i, t)$  lying in  $x_i, \psi$  space, one for each value of  $t$ , each of which may be said to be lifted into another surface in  $x_i, \psi, z_i$  space by using  $z_i = \partial\psi/\partial x_i$  as the defining property of lift, but now having the property

$$d\psi - z_i dx_i = \frac{\partial\psi}{\partial t} dt \quad (3.22)$$

when  $t$  is allowed to vary. In particular, differentiation following a motion (3.1) implies  $\dot{\psi} - z_i \dot{x}_i = \partial\psi/\partial t$ . Such a motion can still be pictured as generating space curves in both  $x_i, \psi$  and  $x_i, \psi, z_i$  space, but the second curve is no longer the lift of the first, and neither lies on a single  $t = \text{const.}$  surface.

The combined effect of the transformations produces a family of surfaces  $\Psi = \Psi(X_i, t)$  defined by (3.20) and lying in  $X_i, \Psi$  space, one for each  $t$ , each of which may be said to be lifted into another surface in  $X_i, \Psi, Z_i$  by using (3.21) as the defining property of lift, but now having the property

$$d\Psi - Z_i dX_i = \frac{\partial\Psi}{\partial t} dt \quad (3.23)$$

when  $t$  is allowed to vary. In particular, differentiation following a motion (3.1) implies  $\dot{\Psi} - Z_i \dot{X}_i = \partial\Psi/\partial t$ . Such a motion generates space curves in  $X_i, \Psi$  and  $X_i, \Psi, Z_i$  spaces, but again the second is not the lift of the first.

**Theorem 3.4.** *The partial  $t$  derivatives of (3.14) and (3.20) are related by*

$$\frac{\partial\Psi}{\partial t} = \frac{\partial\psi}{\partial t} + (\frac{1}{2}z_i + c_{ik}^{-1}x_k)z_j \dot{c}_{ij}. \quad (3.24)$$

*Proof.* This follows from theorem 3.1 with (3.22) and (3.23).

Alternatively, differentiating (3.20) and using (3.19)<sub>3</sub> and (3.15) gives

$$\frac{\partial\Psi}{\partial t} = \frac{1}{2}\dot{c}_{ij}X_iX_j + \frac{1}{2}\dot{c}_{ij}^{-1}x_ix_j + \frac{\partial\psi}{\partial t}.$$

The result then follows from (3.9)<sub>1</sub> and the time derivatives of  $c_{ik}c_{kj}^{-1} = \delta_{ij}$ . ■

It is evident from theorems 3.2 and 3.4 that the hypothesis of constant  $c_{ij}$  is

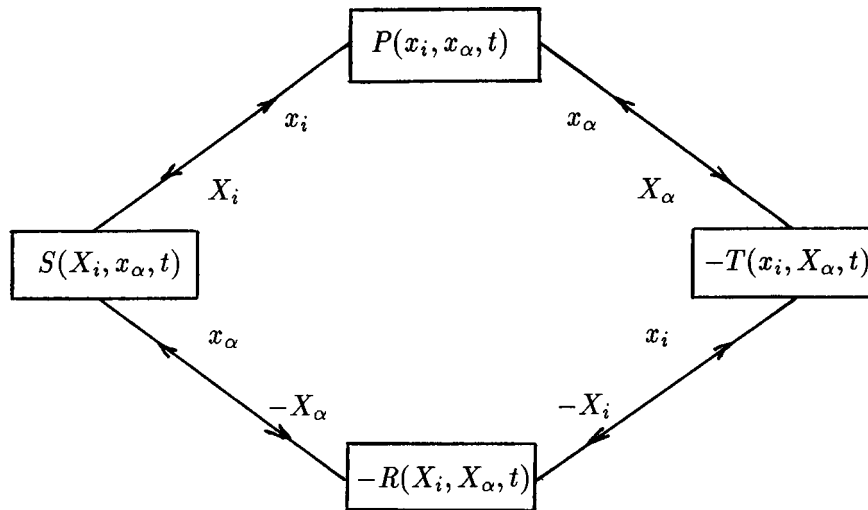


Figure 2. Quartet of Legendre transformations.

sufficient to ensure that

$$d\Psi - Z_i dX_i = d\psi - z_i dx_i \quad \text{and} \quad \frac{\partial \Psi}{\partial t} = \frac{\partial \psi}{\partial t}. \quad (3.25)$$

These are satisfied in the geophysical applications in view, where  $c_{ij} = \delta_{ij}$  is common. Although  $(3.25)_1$  is conventionally used in the formal definition of a lift transformation, as described after theorem 3.2, it is evident that the transformation of a lift into another lift becomes diffuse when  $t$  is present explicitly in  $\psi(x_i, t)$ , even when it is absent from  $c_{ij}$ .

(e) *Augmented lift and Legendre transformations*

Now we suppose that  $m < n$ , i.e. that the kinematically possible motions (3.1) contain  $n - m$  more variables than appear in the lift transformations (3.9). To distinguish between the two sets of variables, we confine Latin suffixes to range from  $i = 1, 2, \dots, m$  and we introduce Greek suffixes to range from  $\alpha = m + 1, \dots, n$ .

In place of (3.14) and (3.17) we introduce the hypotheses

$$\psi = \psi(x_i, x_\alpha, t), \quad z_i = (\partial\psi/\partial x_i), \quad c_{ij} = c_{ij}(t), \quad (3.26)$$

and definitions

$$P(x_i, x_\alpha, t) = \frac{1}{2}c_{ij}^{-1}x_ix_j + \psi, \quad X_\alpha = \partial P/\partial x_\alpha. \quad (3.27)$$

Then (3.16) still applies, and theorem 3.3 is augmented as follows.

**Theorem 3.5.** *This function  $P(x_i, x_\alpha, t)$  is the starting point for generating the four Legendre transformations summarized schematically in figure 2, in terms of three other functions  $S(X_i, x_\alpha, t)$ ,  $R(X_i, X_\alpha, t)$  and  $T(x_i, X_\alpha, t)$ , whose values and gradients satisfy the following equations:*

$$P + S = x_i X_i, \quad X_i = \frac{\partial P}{\partial x_i}, \quad x_i = \frac{\partial S}{\partial X_i}, \quad (3.28)$$

$$S - R = -x_\alpha X_\alpha, \quad X_\alpha = -\frac{\partial S}{\partial x_\alpha}, \quad x_\alpha = \frac{\partial R}{\partial X_\alpha}, \quad (3.29)$$

The mathematical structure of theories of semigeostrophic type 2501

$$R + T = X_i x_i, \quad x_i = \frac{\partial R}{\partial X_i}, \quad X_i = \frac{\partial T}{\partial x_i}, \quad (3.30)$$

$$T - P = -X_\alpha x_\alpha, \quad x_\alpha = -\frac{\partial T}{\partial X_\alpha}, \quad X_\alpha = \frac{\partial P}{\partial x_\alpha}, \quad (3.31)$$

$$\frac{\partial P}{\partial t} + \frac{\partial S}{\partial t} = \frac{\partial S}{\partial t} - \frac{\partial R}{\partial t} = \frac{\partial R}{\partial t} + \frac{\partial T}{\partial t} = \frac{\partial T}{\partial t} - \frac{\partial P}{\partial t} = 0. \quad (3.32)$$

*Proof.* This is a straightforward extension of the proof of theorem 3.3, as explained by Sewell (1987) and Chynoweth & Sewell (1989, 1991), where it is also shown how convexity and concavity properties can be transmitted round the quartet of functions. ■

Each of (3.28)<sub>2</sub>, (3.28)<sub>3</sub>, (3.30)<sub>2</sub> and (3.30)<sub>3</sub> is an alternative version of (3.9)<sub>1</sub> with (3.26)<sub>2</sub>, which in meteorology can be identified with the geostrophic transformation.

Figure 3 shows two simple examples of how the Legendre transformation works when the  $x_i$  and  $X_i$  are active, and the  $x_\alpha$  and  $t$  are passive. These are adapted from figures 2 and 3 of Chynoweth & Sewell (1989), where more general examples are also described. Suppose that  $t$  is fixed, and that we also fix all but one of the  $x_\alpha$ , and all but one of the  $X_i$  and of the associated  $x_i$ . The function

$$S(X_i, x_\alpha) = \frac{1}{12} X_i^4 + \frac{1}{2} x_\alpha X_i^2, \quad (3.33)$$

of the remaining variables, then has the shape shown in figure 3*a*. The long dashes indicate the locus of  $S$ -minima at each fixed  $x_\alpha$ , and the short dashes indicate the locus of inflections. The Legendre transformation maps (3.33) into the function  $P(x_i, x_\alpha)$  shown in figure 3*b*. The inflections have mapped into the edges of the swallowtail, and the minima have mapped into the self-intersection line. The convexification of (3.33) is the weakly convex function  $S(X_i, x_\alpha)$  shown in figure 3*c*, and the Legendre transform of that is the convex function  $P(x_i, x_\alpha)$  shown in figure 3*d*, in which the previous self-intersection has become a gradient discontinuity, such as may appear at an atmospheric front. The gradient jumps there represent jumps in temperature and wind velocity.

**Theorem 3.6.** *The values of the function  $P(x_i, x_\alpha, t)$  are related to the values of  $\Psi$  in (3.9)<sub>2</sub> by (3.16) as before. Therefore the values of  $S(X_i, x_\alpha, t)$  and  $R(X_i, X_\alpha, t)$  are related to those of  $\Psi$  by*

$$\Psi = \frac{1}{2} c_{ij} X_i X_j - S = \frac{1}{2} c_{ij} X_i X_j - R + X_\alpha (\partial R / \partial X_\alpha). \quad (3.34)$$

The first and second expressions here thereby define two different functions having the same value, which we write as  $\Psi_S(X_i, x_\alpha, t)$  and  $\Psi_R(X_i, X_\alpha, t)$ , respectively. When (3.9)<sub>3</sub> is presumed as well, their gradients are

$$\frac{\partial \Psi_S}{\partial X_i} = Z_i, \quad \frac{\partial \Psi_S}{\partial x_\alpha} = X_\alpha, \quad (3.35)$$

$$\frac{\partial \Psi_S}{\partial t} = \frac{\partial \psi}{\partial t} + (\frac{1}{2} z_i + c_{ik}^{-1} x_k) z_j \dot{c}_{ij}, \quad (3.36)$$

$$\frac{\partial \Psi_R}{\partial X_i} = Z_i + X_\beta \frac{\partial^2 R}{\partial X_\beta \partial X_i}, \quad \frac{\partial \Psi_R}{\partial X_\alpha} = X_\beta \frac{\partial^2 R}{\partial X_\beta \partial X_\alpha}, \quad (3.37)$$

$$\frac{\partial \Psi_R}{\partial t} = \frac{\partial \Psi_S}{\partial t} + X_\beta \frac{\partial^2 R}{\partial X_\beta \partial t}. \quad (3.38)$$



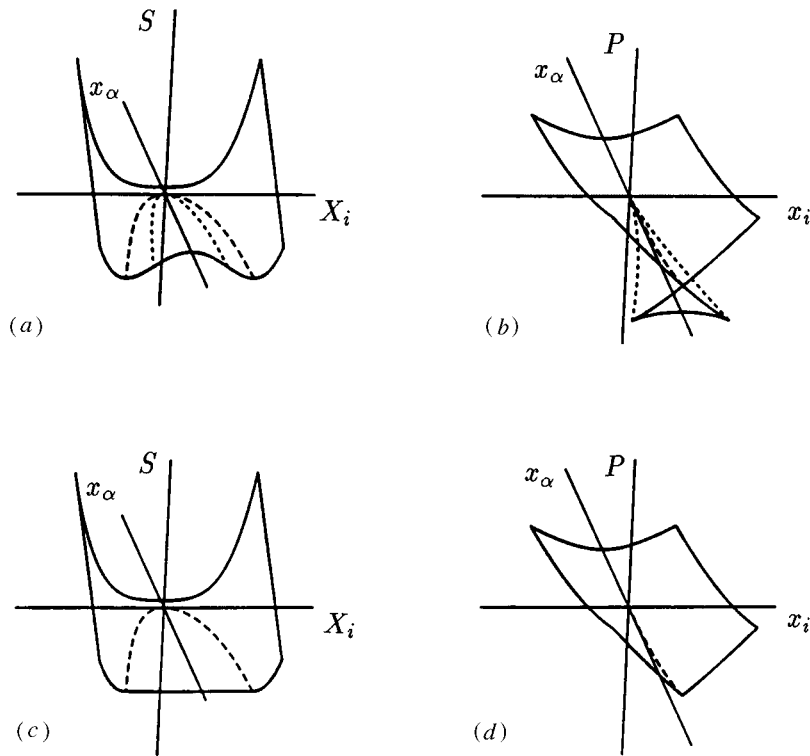


Figure 3. Examples of Legendre dual functions.

*Proof.* This follows from (3.10), (3.26) and (3.27), which imply

$$\left(\frac{\partial \Psi_S}{\partial X_i} - Z_i\right) dX_i + \frac{\partial \Psi_S}{\partial x_\alpha} dx_\alpha + \frac{\partial \Psi_S}{\partial t} dt = \frac{\partial \psi}{\partial x_\alpha} dx_\alpha + \left(\frac{\partial \psi}{\partial t} + \left(\frac{1}{2}z_i + c_{ik}^{-1}x_k\right)z_j \dot{c}_{ij}\right) dt, \quad (3.39)$$

which delivers (3.35) and (3.36). A similar calculation using  $d\Psi_R$  delivers (3.37) and (3.38). ■

We now have an augmented version of (3.22), namely

$$d\psi - z_i dx_i = X_\alpha dx_\alpha + \frac{\partial \psi}{\partial t} dt, \quad (3.40)$$

which leads to augmented versions of (3.23), for example

$$d\Psi_S - Z_i dX_i = X_\alpha dx_\alpha + (\partial \Psi_S / \partial t) dt. \quad (3.41)$$

The expressions in the last two equations are equal when the  $c_{ij}$  are constant. A third expression could be adjoined to (3.34) by defining a function  $\Psi_T(x_i, X_\alpha, t)$  in terms of  $T(x_i, X_\alpha, t)$ .

#### 4. Mass balance

The law of mass balance at places where the relevant functions have enough smoothness is a differential equation, but a different one in the two contexts of § 2, finally expressing compressibility in one case and incompressibility in the other.

## (a) Shallow-water theory

In a motion (2.11) the incompressibility of a column of water of height  $s = 2h$  and geopotential  $\phi = gh$  requires  $\phi$  to be a function  $\phi(a, b, t)$  with the property

$$\frac{\phi(a, b, 0)}{\phi(a, b, t)} = \frac{\partial(x, y)}{\partial(a, b)}, \quad (4.1)$$

where the Jacobian on the right is that of the mapping (2.11). For brevity we write  $\phi(a, b, 0) = \phi_0$  (say). It is not a serious restriction to suppose that  $\phi_0$  is an absolute constant, i.e. the same for every particle. The inverse  $a = a(x, y, t)$ ,  $b = b(x, y, t)$  of (2.11) can express (4.1) in terms of another function

$$\phi(x, y, t) = \frac{\partial(a, b)}{\partial(x, y)} \phi_0. \quad (4.2)$$

Applying (3.8) (with  $z = c$ , a constant, to make  $n = 3$ ) in this context leads to

$$\frac{\dot{\phi}}{\phi} + \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} = 0. \quad (4.3)$$

## (b) Atmospheric theory

The atmosphere is not incompressible, but the hydrostatic and Boussinesq approximations lead to it being regarded as such in adiabatic flow, as follows. Mass conservation expresses the current to initial density ratio as

$$\frac{\rho(a, b, c, t)}{\rho(a, b, c, 0)} = \frac{\partial(a, b, c)}{\partial(x, y, h)}, \quad (4.4)$$

where the Jacobian on the right is that of the inverse of (2.1). Applying (3.8) to this leads to the usual equation of continuity in the form

$$\frac{\dot{\rho}}{\rho} + \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{h}}{\partial h} = 0. \quad (4.5)$$

However, this is in terms of the real height, and we need to express it in terms of the pseudo-height  $z$  introduced in (2.4).

To do this, we first need an example of (3.7) in which  $v_i = \dot{x}_i$  so that  $D\lambda = \dot{\lambda}$  for any  $\lambda$ , but the  $b_i$  are *not* chosen to be the particle labels  $a_i$ . Instead we choose (3.5) to be

$$b_1 = x, \quad b_2 = y, \quad b_3 = z(x, y, h, t), \quad (4.6)$$

(writing  $x_1 = x$ ,  $x_2 = y$ ,  $x_3 = h$ ) and denote *this*  $J$  as  $\alpha = \partial z / \partial h$  so that

$$\frac{\dot{\alpha}}{\alpha} = \left( \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} \right) - \left( \frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{h}}{\partial h} \right). \quad (4.7)$$

But  $\alpha = \sigma_0 / \sigma$  from (2.6), so that  $\dot{\alpha} = -\tau \dot{\eta} \alpha / H$  using (2.2) and (2.3), and

$$\frac{\tau \dot{\eta}}{H} = \frac{\partial \dot{h}}{\partial h} - \frac{\partial \dot{z}}{\partial z}. \quad (4.8)$$

Then we can rewrite (4.5) as

$$\frac{\partial}{\partial t} \left( \frac{\rho}{\alpha} \right) + \frac{\partial}{\partial x} \left( \frac{\rho \dot{x}}{\alpha} \right) + \frac{\partial}{\partial y} \left( \frac{\rho \dot{y}}{\alpha} \right) + \frac{\partial}{\partial z} \left( \frac{\rho \dot{z}}{\alpha} \right) = 0. \quad (4.9)$$

However, (2.2)<sub>2</sub> and (2.3) imply  $\alpha/\rho = \sigma_0 d\zeta/dp$ , so that  $\rho/\alpha$  depends only on  $z$  via the inverse of (2.4), and therefore

$$\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = -\dot{z} \frac{d}{dz} \left( \ln \frac{\rho}{\alpha} \right), \quad (4.10)$$

as shown by Hoskins & Bretherton (1972, equation 2.7, for the ideal gas) and Chynoweth & Sewell (1991, equation 42).

Hoskins & Bretherton also comment on circumstances in which the Boussinesq approximation is valid. It neglects the right side of (4.10), and thereby reduces the equation of continuity to

$$\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{z}}{\partial z} = 0. \quad (4.11)$$

In adiabatic flow (2.10) and (4.8) show that

$$\frac{\partial \dot{h}}{\partial h} = \frac{\partial \dot{z}}{\partial z}. \quad (4.12)$$

Evidently the hydrostatic, Boussinesq and adiabatic assumptions together imply  $\dot{\rho} = 0$  in (4.5). The resulting incompressibility approximation for the atmosphere can be expressed as a zero divergence of either vector  $\dot{x}, \dot{y}, \dot{h}$  or  $\dot{x}, \dot{y}, \dot{z}$  in the respective spaces.

These are finally different in structure from (4.3), which expresses incompressibility in shallow-water theory as if  $\phi$  were the density in a two dimensional compressible flow.

## 5. Dynamics

Now we have the perspective to add the laws of momentum balance to those of mass (in §4) and energy (in §2), and also to consider generalized forms (in the sense of §3) and approximations of the resulting dynamics.

### (a) Exact equations

The exact, or so-called primitive, differential equations of momentum balance in *shallow-water theory* are

$$\ddot{x} + 2 \frac{\partial \phi}{\partial x} - \dot{y}f = 0, \quad \ddot{y} + 2 \frac{\partial \phi}{\partial y} + \dot{x}f = 0, \quad (5.1)$$

in terms of the unknown geopotential function  $\phi(x, y, t)$  in (4.2). Here the Coriolis parameter  $f$  is the spin of the bed, for example the local vertical component of the Earth's rotation. Common approximations treat  $f$  as either a given constant, or as a given function of position, particularly  $f(y)$  when  $y$  is northward distance.

The exact equations of horizontal momentum balance in *atmospheric theory* are

$$\ddot{x} + \frac{\partial \phi}{\partial x} - \dot{y}f = 0, \quad \ddot{y} + \frac{\partial \phi}{\partial y} + \dot{x}f = 0, \quad (5.2)$$

in terms of the geopotential function  $gh = \phi(x, y, z, t)$  which appears in the alternative vertical momentum balance (hydrostatic) equations (2.7) and (2.9). Again here the Coriolis quantity  $f$  may be either a given constant, or a function of position, for example  $f(y)$ .

(b) *Semigeostrophic approximation*

The vector defined by

$$u_g = -\frac{\kappa}{f} \frac{\partial \phi}{\partial y}, \quad v_g = \frac{\kappa}{f} \frac{\partial \phi}{\partial x}, \quad (5.3)$$

where  $\kappa = 2$  in shallow-water theory and  $\kappa = 1$  in atmospheric theory, is called the geostrophic velocity or geostrophic wind. At this stage it is not more than a *notional* velocity, as it is not yet demonstrated to be the time derivative of another vector.

The semigeostrophic approximation replaces the true accelerations in (5.1) and (5.2) by the time derivatives of (5.3), thus giving

$$\dot{u}_g + \kappa \frac{\partial \phi}{\partial x} - \dot{y}f = 0, \quad \dot{v}_g + \kappa \frac{\partial \phi}{\partial y} + \dot{x}f = 0, \quad (5.4)$$

with  $\kappa = 2$  or 1 for shallow-water theory or atmospheric theory, respectively. This can be thought of as a *substitute problem*, which replaces the exact problem specified by (5.1) or completed by (5.2).

(c) *Résumé of semigeostrophic equations*

The semigeostrophic shallow-water problem seeks to find a motion (2.1) whose inverse generates a geopotential (4.2) satisfying

$$-\frac{2}{f} \frac{\partial \dot{\phi}}{\partial y} + 2 \frac{\partial \phi}{\partial x} - \dot{y}f = 0, \quad \frac{2}{f} \frac{\partial \dot{\phi}}{\partial x} + 2 \frac{\partial \phi}{\partial y} + \dot{x}f = 0. \quad (5.5)$$

The semigeostrophic atmospheric problem seeks to find a motion

$$x = x(a, b, c, t), \quad y = y(a, b, c, t), \quad z = z(a, b, c, t) \quad (5.6)$$

obtained via (2.1) and  $z = z(p(x, y, h, t))$  via (2.3) and (2.4), whose geopotential  $\phi(x, y, z, t)$  satisfies (2.7) with  $\dot{\eta} = 0$ , (4.11) and

$$-\frac{1}{f} \frac{\partial \dot{\phi}}{\partial y} + \frac{\partial \phi}{\partial x} - \dot{y}f = 0, \quad \frac{1}{f} \frac{\partial \dot{\phi}}{\partial x} + \frac{\partial \phi}{\partial y} + \dot{x}f = 0. \quad (5.7)$$

Note that

$$\frac{\partial \dot{\phi}}{\partial z} = \frac{\partial \phi}{\partial z} \frac{\tau \dot{\eta}}{H} \quad (5.8)$$

is zero when  $\dot{\eta} = 0$ .

## 6. Generalized semigeostrophic equations

Here we use the notation and theory developed in §3 to construct pertinent generalizations and transformations of the semigeostrophic equations summarized in §5c.

(a) *Generalized Hamiltonian equations*

Write  $n = m$  in (3.1), as in §§3c and 3d, and suppose that  $c_{ij}$  is constant, so that (3.25) applies. Let  $J_{ij}$  be the typical component of a constant anti-symmetric matrix (unrelated to the  $J$  in §3a). Hypothesize the dynamical equations

$$\dot{X}_i = J_{ij}(\partial \Psi / \partial X_j), \quad (6.1)$$

in the  $X_i$  space, where the gradients of the function (3.20) are on the right. These are a generalization of Hamilton's equations in which, for example,  $m$  is not required to be even. Solutions of them have the form

$$X_i = X_i(a_j, t). \quad (6.2)$$

The implication is that the functions on the right are the result of inserting suitable functions (3.1) into (3.9) with (3.17), as specified by the following result.

**Theorem 6.1.** *The problem (6.1) in  $X_i$  space is equivalent to the problem*

$$\frac{\dot{\partial\psi}}{\partial x_i} - J_{ij}c_{jk} \frac{\partial\psi}{\partial x_k} + c_{ij}^{-1} \dot{x}_j = 0 \quad (6.3)$$

in  $x_i$  space.

*Proof.* Suppose that (3.1) is a solution of (6.3), and that inserting it into (3.9) with (3.17) delivers (6.2). From (3.9) and (3.17) we have

$$\frac{\partial\Psi}{\partial X_i} = c_{ij} \frac{\partial\psi}{\partial x_j} \quad \text{and} \quad \dot{X}_i = c_{ij}^{-1} \dot{x}_j + \frac{\partial\psi}{\partial x_i}. \quad (6.4)$$

Using the first of these, and the time derivatives of the second following the particle, shows that each of (6.1) and (6.3) implies the other. ■

**Theorem 6.2.** *Solutions of (6.1) or (6.3) have the properties*

$$\frac{\partial\dot{X}_i}{\partial X_i} = 0, \quad \dot{X}_i \frac{\partial\Psi}{\partial X_i} = 0, \quad \dot{\Psi} = \dot{\psi} - \dot{x}_i \frac{\partial\psi}{\partial x_i}. \quad (6.5)$$

*Proof.* These are immediate consequences of the antisymmetry of  $J_{ij}$ , and of (3.25)<sub>1</sub>. ■

Next we identify circumstances in which (5.5) is an example of (6.3).

(b) *Shallow-water theory on an  $f$ -plane*

Here we specialize to the case in which  $n = m = 2$ , and establish a correspondence of the general theory with the semigeostrophic shallow-water theory of (5.3) and (5.5) in the case when  $f$  is a given constant.

Suppose that  $c_{ij} = \mu\delta_{ij}$  with constant  $\mu$ , and that

$$J_{ij} = \begin{bmatrix} 0 & -\lambda \\ \lambda & 0 \end{bmatrix}, \quad (6.6)$$

with constant  $\lambda$  in (6.1).

So far we have not needed to identify coordinates and variables in the general theory with those in either of the two physical illustrations. Now we do this by introducing constant scale factors  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $m$  such that

$$x_1 = \alpha x, \quad x_2 = \beta y, \quad \psi = 2\gamma\phi/g \quad (6.7)$$

and

$$m\dot{X}_1 = u_g, \quad m\dot{X}_2 = v_g. \quad (6.8)$$

The last condition requires the (previously notional) geostrophic velocity to be an actual time derivative (or at least proportional to one, like momentum if  $m$  is thought

of as mass) in  $X_i$  space. This scale factor  $m$  is quite distinct from the  $m$  introduced as a number of variables in §3*b*, and no confusion need arise. Likewise the  $\alpha$  here is unrelated to that defined in §4*b*.

**Theorem 6.3.** *Equations (6.3) (and thus (6.1)) are equivalent to the semi-geostrophic equations (5.5), and (6.8) are satisfied, if and only if the scale factors  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $m$  satisfy*

$$\alpha = \beta, \quad \lambda\mu = f, \quad \frac{\gamma\mu}{\alpha^2} = \frac{g}{f^2}, \quad \mu = m\alpha. \quad (6.9)$$

*Proof.* Substituting (6.7) with (6.6) into (6.3), with  $n = 2$ , delivers (5.5) if the first three relations in (6.9) hold (and also if  $\alpha = -\beta$  and  $\lambda\mu = -f$  instead).

To require (6.8) to hold with (5.5) and (6.1), with  $\partial\Psi/\partial X_i = c_{ij}\partial\psi/\partial x_j$  from (3.9)<sub>3</sub>, (6.4) and (6.7), it is necessary and sufficient that  $\alpha = \beta$  and  $m\lambda\mu\gamma = \alpha g/f$ . With the second and third of (6.9), this delivers  $\mu = m\alpha$ . ■

Except for  $\lambda$  and  $\mu$ , any pair selected from  $\alpha$ ,  $\gamma$ ,  $\lambda$ ,  $\mu$  and  $m$  can be chosen arbitrarily to satisfy (6.9), with  $f$  and  $g$  regarded as given, and the remaining three are then determined, apart from possible duplicity of sign. Without loss of generality, this is equivalent to choosing arbitrarily any pair selected from  $\lambda$ ,  $\gamma$  and  $m$  and requiring the third parameter to satisfy

$$\lambda\gamma m^2 = g/f. \quad (6.10)$$

Then the corresponding

$$\alpha = \gamma m f^2/g. \quad (6.11)$$

For example, one might instinctively wish to choose  $\alpha = 1$  and  $\lambda = 1$  to achieve the simplest versions of space coordinates  $x_i$  and of Hamilton's equations (6.1), which would imply  $m = \mu = f$  and  $\gamma = g/f^2$ ; but other choices are possible.

In particular, if one wishes to call the  $X_i$  *geostrophic coordinates*, a justification of this is provided by (6.8) with the choice  $m = 1$ , for then the geostrophic velocity (5.3) with  $\kappa = 2$  is the rate of change of such geostrophic coordinates. With  $m = 1$ , however, we can only choose  $\alpha = 1$  or  $\lambda = 1$ , but not both. Any constant  $\lambda$  can be absorbed into the Hamiltonian to redefine it as  $\lambda\Psi$  in place of  $\Psi$  in (6.1), so these particular criteria offer

$$m = \alpha = 1 = \mu, \quad \gamma = g/f^2, \quad \lambda = f, \quad (6.12)$$

as an inviting choice of parameters, with  $\psi = 2\phi/f^2$ .

### (c) Generalized augmented Hamiltonian equations

Here we consider the case  $m < n$  first allowed for explicitly in §3*e*. The case  $m = 2$  with  $n = 3$  is required to describe the semigeostrophic approximation of the atmospheric problem summarized in (5.6)–(5.8). Latin suffixes here range only up to  $m$ , and Greek suffixes range from  $m + 1$  to  $n$ .

Now let  $c_{ij}$  and  $J_{ij}$  be constant  $m \times m$  matrices, with  $J_{ij}$  anti-symmetric. Use the  $X_i$  defined in (3.9) and the  $X_\alpha$  defined in (3.27) to hypothesize the dynamical equations

$$\dot{X}_i = J_{ij} \frac{\partial \Psi_S}{\partial X_j}, \quad \dot{X}_\alpha = 0, \quad (6.13)$$

in terms of the function  $\Psi_S(X_i, x_\alpha, t)$  defined in theorem 3.6, having gradients (3.35).



This a generalized augmented Hamiltonian problem, in which again it is not essential that  $m$  be even.

Bearing in mind that  $x_\alpha = \partial R / \partial X_\alpha$  from (3.24), the solutions of (6.13) for each material particle (i.e. for each set of constant values  $a_1, \dots, a_m, a_{m+1}, \dots, a_n$  of the labels in (3.1)) have the form

$$X_i = X_i(t), \quad X_\alpha = \text{const.} \quad (6.14)$$

The generalized version of incompressibility (4.11) with the Boussinesq approximation can be stated as

$$\frac{\partial \dot{x}_i}{\partial x_i} + \frac{\partial \dot{x}_\alpha}{\partial x_\alpha} = 0. \quad (6.15)$$

This acts as a constraint to be imposed on the solutions (6.14), with the aid of the expressions in theorems 3.5 and 3.6.

The problem (6.13) in  $X_i, X_\alpha$  space is equivalent to the problem (6.3) with

$$\overline{\partial P / \partial x_\alpha} = 0$$

in  $x_i, x_\alpha$  space. To show this requires the same manipulations as in the proof of theorem 6.1, together with the additional properties in (3.26) and (3.27). Solutions of (6.13) have the properties (6.5) but with  $\Psi$  there replaced by  $\Psi_S$ .

(d) *Atmospheric theory on an  $f$ -plane*

Let  $m = 2$  and  $n = 3$  in the previous subsection. Let  $c_{ij} = \mu \delta_{ij}$ , and  $J_{ij}$  be as in (6.6), with constant  $\mu$  and  $\lambda$ . Introduce constant scale factors  $\alpha, \beta$  and  $\gamma$  such that the variables  $x_1, x_2$  and  $\psi$  in the general theory begun in theorem 3.1 are related by

$$x_1 = \alpha x, \quad x_2 = \beta y, \quad \psi = \gamma \phi / g, \quad (6.16)$$

to the horizontal coordinates  $x$  and  $y$  and the geopotential  $\phi$  in (5.6) and (5.7).

The variables  $X_1$  and  $X_2$  are required to satisfy (3.9)<sub>1</sub> with (3.26)<sub>2</sub>. Introduce the choices

$$x_3 = \nu h, \quad X_3 = \eta, \quad \nu = \gamma / \eta. \quad (6.17)$$

Here  $\nu$  is a constant scale factor which relates the real height  $h$  to the only surviving member of the set of  $x_\alpha$  variables introduced in §3e. We shall seek to model only adiabatic flows, for which the choices in (6.17) are consistent with (6.13)<sub>2</sub>. (We could have chosen  $X_3$  to be an arbitrary function of  $\eta$ , but that yields no extra advantage.)

**Theorem 6.4.** *Equations (6.13) are equivalent to the semigeostrophic equations (5.7) and the adiabatic condition, together with the incompressibility condition*

$$\frac{\partial \dot{x}}{\partial x} + \frac{\partial \dot{y}}{\partial y} + \frac{\partial \dot{h}}{\partial h} = 0, \quad (6.18)$$

and (6.8) are satisfied, if and only if the scale factors satisfy (6.9), provided the choice of  $\nu$  satisfies (6.17).

*Proof.* The proof of theorem 6.3 applies as it stands to identify the  $X_1$  and  $X_2$  Hamiltonian equations in (6.13) with (5.7). The scale factors  $\alpha, \alpha, \nu$  associated with the three cartesian physical space coordinates appear in both numerator and denominator of (6.15) and therefore (6.18), and so cancel out. The adiabatic condition  $\dot{\eta} = 0$  is equivalent to  $\dot{X}_3 = 0$ . From (3.27) and (6.16)<sub>3</sub> we have

$$X_3 = \frac{\partial \psi}{\partial x_3} = \frac{\gamma}{g} \frac{\partial \phi}{\partial x_3}.$$

For adiabatic flow, this is consistent with the choices (6.17) because  $\phi = gh$ . ■

(e) *Potential vorticity and the Monge–Ampère equation*

The geostrophic momentum transformation, in the generalized notation, is a mapping

$$X_i = X_i(x_j, x_\beta, t), \quad X_\alpha = X_\alpha(x_j, x_\beta, t), \quad (6.19)$$

at each  $t$  from  $x_j, x_\beta$  space to  $X_i, X_\alpha$  space. It is defined by equations (3.9) with (3.17) if  $m = n$ , or by (3.28)<sub>2</sub> and (3.27)<sub>2</sub> if  $m < n$ . Let  $q$  denote the  $n \times n$  Jacobian of this mapping. Clearly  $q$  is also the Hessian, at each  $t$ , of the function  $P(x_i, t)$  in (3.15) or  $P(x_i, x_\alpha, t)$  in (3.27)<sub>1</sub>, respectively.

When  $n = 3$  and  $m = 2$ , we can regard (6.19) as a particular example of (3.5)<sub>1</sub>, in which case (3.7) becomes

$$\frac{Dq}{q} = \frac{\partial DX_i}{\partial X_i} + \frac{\partial DX_\alpha}{\partial X_\alpha} - \left( \frac{\partial v_i}{\partial x_i} + \frac{\partial v_\alpha}{\partial x_\alpha} \right). \quad (6.20)$$

If we also choose  $v_i = \dot{x}_i$ ,  $v_\alpha = \dot{x}_\alpha$  to be particle velocity, (6.20) becomes

$$\frac{\dot{q}}{q} = \frac{\partial \dot{X}_i}{\partial X_i} + \frac{\partial \dot{X}_\alpha}{\partial X_\alpha} - \left( \frac{\partial \dot{x}_i}{\partial x_i} + \frac{\partial \dot{x}_\alpha}{\partial x_\alpha} \right), \quad (6.21)$$

as in (20) of Chynoweth & Sewell (1991).

**Theorem 6.5.** *When (6.13) and (6.15) apply for  $n = 3$  and  $m = 2$ ,  $\dot{q} = 0$ .*

*Proof.* This is immediate from (6.21). ■

This result proves the conservation of  $q$  following the particle. In the atmospheric case  $q$  is identifiable with the potential vorticity (see Hoskins, 1975, §§ 3(ii) and 3(iii)). The structure of the conservation proof is clearly displayed *via* (6.21). Liouville's theorem makes  $\partial \dot{X}_i / \partial X_i = 0$ ; the adiabatic assumption makes  $\partial \dot{X}_\alpha / \partial X_\alpha = 0$ ; the Boussinesq approximation and incompressibility (4.11) justify (6.15).

At each  $t$ , the Jacobian of the inverse of (6.19) is  $q^{-1}$ . Because  $P(x_i, x_\alpha, t)$  and  $R(X_i, X_\alpha, t)$  are Legendre dual functions, by (3.28) and (3.29),  $q^{-1}$  is also the Hessian of  $R(X_i, X_\alpha, t)$ , i.e.

$$\det \left( \frac{\partial^2 R}{\partial (X_i, X_\alpha)} \right) = q^{-1}. \quad (6.22)$$

In developing solution strategies for the semigeostrophic equations in the atmosphere, Cullen *et al.* (1991) use the conservation of  $q$  as the starting point for an assumption that  $q$  can be regarded as an assigned function  $q(X_i, X_\alpha)$  at each  $t$ , in which case (6.22) becomes a Monge–Ampère equation for the determination of  $R(X_i, X_\alpha, t)$  regarded as an unknown.

When  $n = m = 2$ , instead of (6.21) we obtain

$$\frac{\dot{q}}{q} = \frac{\partial \dot{X}_i}{\partial X_i} - \frac{\partial \dot{x}_i}{\partial x_i}. \quad (6.23)$$

If the incompressibility assumption (6.15) is replaced by a compressibility assumption

$$\frac{\dot{\phi}}{\phi} + \frac{\partial \dot{x}_i}{\partial x_i} = 0, \quad (6.24)$$

consistent with (4.3), a quantity different from  $q$  is conserved.

**Theorem 6.6.** When (6.13)<sub>1</sub> and (6.24) apply for  $n = m = 2$ ,  $q/\phi$  is conserved following a particle.

*Proof.* Liouville's theorem and (6.24) imply  $\overline{q/\phi} = 0$  from (6.23). ■

This result would be one reason for calling  $q/\phi$ , instead of  $q$ , the potential vorticity in this context. A completely different proof is offered by Salmon (1985, equation (3.28)), building upon Salmon (1983, §2).

## 7. Variable Coriolis parameter

The problem of identifying what may be regarded as essential elements in the mathematical structure of semigeostrophic equations for cases in which the Coriolis parameter is not a fixed constant is broached here. We shall not offer a complete solution, but we shall use the rather clear cut structure exposed above for a constant Coriolis parameter to develop certain pointers.

### (a) Planetary semigeostrophic equations

Here we explore how the form of some equations obtained by Shutts (1989) for motion on the surface of a sphere can be recovered, by quite a different method to his.

We begin by hypothesizing Hamilton's equations in the abstract form (6.1), but in the particular case  $n = m = 2$  with (6.6), namely

$$\dot{X}_1 = -\lambda \frac{\partial \Psi}{\partial X_2}, \quad \dot{X}_2 = \lambda \frac{\partial \Psi}{\partial X_1}. \quad (7.1)$$

Here  $\Psi$  is a given function  $\Psi(X_1, X_2, t)$ , and  $\lambda$  is a constant.

Next we suppose that there is a pair of abstract variables  $x_1$  and  $x_2$ , and another function  $\psi(x_1, x_2, t)$ , these being related to the foregoing variables by the particular lift transformation

$$\left. \begin{aligned} X_1 &= x_1 + \frac{\partial \psi}{\partial x_1}, & X_2 &= x_2 + \frac{\partial \psi}{\partial x_2}, \\ \Psi &= \psi + \frac{1}{2} \left( \left( \frac{\partial \psi}{\partial x_1} \right)^2 + \left( \frac{\partial \psi}{\partial x_2} \right)^2 \right), & \frac{\partial \Psi}{\partial X_1} &= \frac{\partial \psi}{\partial x_1}, & \frac{\partial \Psi}{\partial X_2} &= \frac{\partial \psi}{\partial x_2}. \end{aligned} \right\} \quad (7.2)$$

This corresponds to the case  $\mu = 1$  in (3.13) and (3.9), with (6.4).

The differential equations (7.1) become

$$\dot{x}_1 + \frac{\partial \dot{\psi}}{\partial x_1} = -\lambda \frac{\partial \psi}{\partial x_2}, \quad \dot{x}_2 + \frac{\partial \dot{\psi}}{\partial x_2} = \lambda \frac{\partial \psi}{\partial x_1} \quad (7.3)$$

in the  $x_1, x_2$  space. The Legendre transformation in theorem 3.3 tells us that

$$\left. \begin{aligned} \Psi(X_1, X_2, t) &= \frac{1}{2}(X_1^2 + X_2^2) - S(X_1, X_2, t), \\ \psi(x_1, x_2, t) &= P(x_1, x_2, t) - \frac{1}{2}(x_1^2 + x_2^2) \end{aligned} \right\} \quad (7.4)$$

in terms of the Legendre dual functions  $S(X_1, X_2, t)$  and  $P(x_1, x_2, t)$ .

Now we hypothesize relations between such abstract variables, and the cartesian coordinates  $x$  and  $y$  representing local eastward and northward distances in the tangent plane at the Earth's surface, with a geopotential function  $\phi(x, y, t)$ . We

suppose that the Coriolis parameter  $f$  is a given function  $f(y)$ , and that a particular value  $f_0$  (say) of  $f(y)$  is present as well. The geostrophic velocity is now defined by (5.3) with  $\kappa = 1$  but with the *variable*  $f(y)$  in each denominator.

The required relations are set down as scaling relations like those in (6.7) or (6.16) with (6.8), but with the difference that some of the scale factors are now allowed to be certain functions of  $y$  at least, instead of the constants which they were when  $f$  was constant.

Firstly, however, we choose some of the scale factors to be constants again, namely

$$\lambda = \alpha = f_0, \quad \text{so that} \quad x_1 = f_0 x, \quad (7.5)$$

together with

$$\dot{X}_1 = f_0 u_g. \quad (7.6)$$

Instead of (6.7)<sub>2</sub> and (6.8)<sub>2</sub>, however, we choose

$$x_2 = \int_{y_0}^y f(u) du, \quad \text{so that} \quad \frac{dx_2}{dy} = f(y), \quad (7.7)$$

and

$$\dot{X}_2 = f(y) v_g. \quad (7.8)$$

Finally, in place of (6.7)<sub>3</sub> or (6.16)<sub>3</sub> we choose

$$\psi(x_1, x_2) = \phi(x, y). \quad (7.9)$$

**Theorem 7.1.** *The partly  $y$ -dependent scalings set down in (7.5)–(7.9) convert Hamilton's equations (7.1) into*

$$\dot{u}_g + \frac{\partial \phi}{\partial x} - f(y) \dot{y} = 0, \quad \frac{f(y)}{f_0} \overline{\left( \frac{f(y) v_g}{f_0} \right)} + \frac{\partial \phi}{\partial y} + f(y) \dot{x} = 0, \quad (7.10)$$

which are representations of zonal and meridional planetary semigeostrophic equations.

*Proof.* We have already shown that (7.1) are equivalent to (7.3) by using the lift transformation (7.2). The transformations (7.5) and (7.7) show that the choice (7.9) has the properties

$$\frac{\partial \phi}{\partial x} = f_0 \frac{\partial \psi}{\partial x_1}, \quad \frac{\partial \phi}{\partial y} = f(y) \frac{\partial \psi}{\partial x_2}. \quad (7.11)$$

Then (7.3) becomes

$$f_0 \dot{x} + \frac{1}{f_0} \overline{\frac{\partial \phi}{\partial x}} + \frac{f_0}{f(y)} \frac{\partial \phi}{\partial y} = 0, \quad f(y) \dot{y} + \frac{1}{f(y)} \overline{\frac{\partial \phi}{\partial y}} - \frac{\partial \phi}{\partial x} = 0. \quad (7.12)$$

In reverse order these are equivalent to (7.10), using the definitions

$$u_g = -\frac{1}{f(y)} \frac{\partial \phi}{\partial y}, \quad v_g = \frac{1}{f(y)} \frac{\partial \phi}{\partial x}. \quad (7.13)$$

Suppose now that  $\varphi$  and  $\Omega$  denote latitude and the spin of the Earth, respectively. If, in (7.12), we replace  $f(y)$  by  $2\Omega \sin \varphi$ , and  $f_0$  by  $2\Omega$ , we obtain the meridional and zonal equations of Shutts (1989, equations (77) and (78)). ■

(b) *Pseudo-plane approximations*

It is clear that the planetary semigeostrophic equations (7.10) are *not* obtained as in §5*b*, i.e. not just by replacing the acceleration in the exact equations by the time derivatives of the geostrophic velocity. This is not surprising.

There is an intermediate situation, between the planetary case and the case of constant  $f$ . This is when variable  $f$  is approximated as

$$f(y) = \beta y, \quad (7.14)$$

with constant  $\beta$  (quite distinct from the  $\beta$  used in (6.7), (6.9) and (6.16)). This is called the  $\beta$ -plane approximation, with northward  $y$ .

Again we might begin with two sets of abstract variables  $X_1, X_2, \Psi$  and  $x_1, x_2, \psi$  related by the lift transformation (7.2). Suppose that pseudo-Hamiltonian equations

$$\dot{X}_1 = -f(y) \frac{\partial \Psi}{\partial X_2}, \quad \dot{X}_2 = f(y) \frac{\partial \Psi}{\partial X_1} \quad (7.15)$$

are postulated. Here the variable  $f(y)$  takes the place of the constant scale factor  $\lambda$  in (7.1), so that Liouville's theorem will no longer apply in general. Even so, we can ask whether recognizable semigeostrophic equations can be recovered from (7.3), which holds with  $\lambda = f(y)$ , by postulating nonlinear relations between the abstract  $x_1, x_2$  and the cartesian position coordinates  $x, y$ .

If we repeat the choice (7.7) with the particular  $f(y) = \beta y$ ,

$$x_2 = \frac{1}{2}\beta(y^2 - y_0^2) \quad \text{and} \quad f = (2\beta x_2 + \beta^2 y_0^2)^{1/2}. \quad (7.16)$$

We use this to define an intermediate function  $\Phi(x_1, y)$ , say, by  $\Phi(x_1, y) = \psi(x_1, x_2)$ , in terms of which (7.15) implies

$$-\frac{1}{f(y)} \frac{\partial \Phi}{\partial y} + f(y) \frac{\partial \Phi}{\partial x_1} - y f(y) = 0, \quad \frac{\partial \Phi}{\partial x_1} + \frac{\partial \Phi}{\partial y} + \dot{x}_1 = 0, \quad (7.17)$$

from (7.3) in reverse order. These are formally close to the semigeostrophic equations (5.4). Whether they are equivalent to them, however, requires investigation of whether we can choose a mapping

$$x_1 = x_1(x, y), \quad \text{such that} \quad \phi(x, y) = \Phi(x_1, y), \quad (7.18)$$

whose properties,

$$\frac{\partial \phi}{\partial x} = \frac{\partial \Phi}{\partial x_1} \frac{\partial x_1}{\partial x}, \quad \frac{\partial \phi}{\partial y} = \frac{\partial \Phi}{\partial x_1} \frac{\partial x_1}{\partial y} + \frac{\partial \Phi}{\partial y}, \quad (7.19)$$

lead to the desired result.

The difficulty which emerges is that, while the properties

$$\frac{\partial x_1}{\partial x} = f(y), \quad \frac{\partial x_1}{\partial y} = 0 \quad (7.20)$$

would be sufficient to allow (7.17) to be written as (5.4) with  $\kappa = 1$  and  $f = f(y)$ , as may be easily verified, there is no mapping (7.18)<sub>1</sub> which could have the properties (7.20), for example because (7.20) would imply

$$\frac{\partial^2 x_1}{\partial y \partial x} \neq \frac{\partial^2 x_1}{\partial x \partial y}. \quad (7.21)$$

In other words, (7.20) are not mutually consistent, because if  $x_1$  depends only on  $x$ , then  $\partial x_1/\partial x$  would not depend on  $y$ .

One might suppose that associated difficulties can be circumvented by using so-called non-holonomic systems of equations, which in the present case would take the differential form

$$dx_1 = f(y) dx, \quad (7.22)$$

but such suggestions have not so far been explicit enough to carry conviction.

(c) *Other lift transformations*

The interest which can attach to problems in which the Coriolis parameter is not a constant prompts enquiry of whether lift transformations having a similar but different structure to those in theorem 3.1 can be devised, such that the new structure is useful in geophysical fluid dynamics. There are, after all, many different types of lift transformation. Sewell & Roulstone (1994) proved how those in (3.9) are particular members of a class deducible from a simple algebraic starting point, namely by requiring the  $X_i$  and  $\Psi$  to be quadratic functions of the  $x_i$ ,  $\psi$  and  $z_i$ . Here we explore briefly a different starting point.

Again let  $x_i, \psi, z_i$  and  $X_i, \Psi, Z_i$  be two sets of  $2m+1$  scalar variables. Let  $u(x_i, z_i)$  be a scalar function of the  $2m$   $x_i$  and  $z_i$ , to be found from criteria to be formulated below. The transformation

$$X_i = x_i + (z_i/u), \quad \Psi = \psi + (z_i z_i/2u), \quad Z_i = Z_i(x_j, \psi, z_j), \quad (7.23)$$

is a mapping from  $x_i, \psi, z_i$  space to  $X_i, \Psi, Z_i$  space, for any  $m+1$  scalar functions  $u(x_i, z_i)$  and  $Z_i(x_j, \psi, z_j)$  still to be specified. If  $u$  is not a constant, the functions  $X_i(x_j, z_j)$  and  $\Psi(x_i, \psi, z_i)$  will not be quadratic in general.

**Theorem 7.2.** *The mapping (7.23) is a lift transformation from  $x_i, \psi, z_i$  space to  $X_i, \Psi, Z_i$  space, in the sense that differentials satisfy*

$$d\Psi - Z_i dX_i = d\psi - z_i dx_i, \quad (7.24)$$

for all  $dx_i$  and  $dz_i$ , if and only if the functions  $u(x_i, z_i)$  and  $Z_i(x_j, z_j)$  satisfy

$$u \frac{\partial u}{\partial z_j} - \frac{\partial u}{\partial x_j} = 0, \quad (7.25)$$

and

$$\left( \delta_{ij} - \frac{z_j}{u} \frac{\partial u}{\partial z_i} \right) Z_j = \left( \delta_{ij} - \frac{z_j}{2u} \frac{\partial u}{\partial z_i} \right) z_j. \quad (7.26)$$

*Proof.* Direct calculation shows that differentials satisfy

$$\begin{aligned} d\Psi - Z_i dX_i &= d\psi - z_i dx_i + (z_i - Z_i) dx_i + \frac{(z_i - Z_i)}{u} dz_i + \frac{z_i}{u^2} (Z_i - \frac{1}{2} z_i) du \\ &= d\psi - z_i dx_i + \left( z_i - Z_i + \frac{z_j}{u^2} (Z_j - \frac{1}{2} z_j) \frac{\partial u}{\partial x_i} \right) dx_i \\ &\quad + \left( z_i - Z_i + \frac{z_j}{u} (Z_j - \frac{1}{2} z_j) \frac{\partial u}{\partial z_i} \right) \frac{dz_i}{u}. \end{aligned}$$

Hence (7.24) holds for any  $dx_i$  and  $dz_i$  if

$$Z_i - z_i = \frac{z_j}{u^2} (Z_j - \frac{1}{2} z_j) \frac{\partial u}{\partial x_i} = \frac{z_j}{u} (Z_j - \frac{1}{2} z_j) \frac{\partial u}{\partial z_i}.$$



Evidently  $Z_j = \frac{1}{2}z_j$  would not be self-consistent, and assuming  $z_j \neq 0$ , (7.25) and (7.26) then follow. ■

The problem is thereby reduced to solving the  $m$  partial differential equations (7.25) for  $u(x_i, z_i)$ , inserting these solutions into (7.26), and solving those linear algebraic equations to find functions  $Z_i(x_j, z_j)$ . There will be many solutions of the system (7.25), and we illustrate two approaches.

An elementary approach is to look for solutions  $u = v(x_i)w(z_j)$  in which the variables are separable, where  $v(x_i)$  and  $w(z_j)$  are two scalar functions, each of the indicated  $m$  variables, to be found.

**Theorem 7.3.** *Separable solutions of (7.25) must have the form*

$$u = \frac{a + c_j z_j}{b - c_i x_i}, \quad (7.27)$$

where  $a$ ,  $b$  and the  $c_i$  are  $m + 2$  arbitrary constants. Then (7.26) has the solution

$$Z_j = z_j + (c_j z_i z_i / 2a), \quad (7.28)$$

provided  $0 \neq a \neq -c_j z_j$ .

*Proof.* The separable assumption inserted into (7.25) implies  $w = 0$  or

$$\frac{\partial w}{\partial z_j} = -\frac{\partial v^{-1}}{\partial x_j} = c_j,$$

which gives (7.27). The matrix of coefficients on the left of (7.26) has determinant  $a/(a + c_j z_j)$ , whence the result (7.28), which can easily be verified directly. ■

The corresponding lift transformation can be written down by inserting (7.27) and (7.28) into (7.23).

More general solutions of (7.25) can be found by exploiting the idea of characteristics (e.g. in the case of  $m = 1$  at first). The following result emerges. Write  $s_i = z_i + u x_i$  and let  $F(s_i)$  denote an arbitrary differentiable function of  $m$  variables.

**Theorem 7.4.** *Any function  $u(x_i, z_i)$  which satisfies*

$$u = F(z_i + u x_i), \quad (7.29)$$

and

$$x_i \frac{\partial F}{\partial s_i} \neq 1 \quad (7.30)$$

will satisfy (7.25).

*Proof.* The partial derivatives of (7.29) combine to give

$$\left( u \frac{\partial u}{\partial z_j} - \frac{\partial u}{\partial x_j} \right) \left( 1 - x_i \frac{\partial F}{\partial s_i} \right) = 0. \quad (7.31)$$

The inverse function theorem shows that (7.30) is sufficient for (7.29) to have at least a local solution  $u(x_i, z_i)$ . The corresponding  $Z_i(x_j, z_j)$  are then to be found from (7.26).

It is clear that (7.27) is recoverable from (7.29) in the particular case of linear

$$F(s_i) = \frac{a + c_i s_i}{b}, \quad (7.32)$$

when  $0 \neq b \neq c_i x_i$ .

The meteorological problem associated with these results is the following. Can physical spatial variables be linked to the  $x_i$ , for example by a linear scaling in the case  $m = 2$ , in such a way that the arbitrariness of the function  $F(s_i)$  can be exploited so that the solution  $u(x_i, z_i)$  of (7.29) fits a pre-assigned function  $u(x_i)$  which represents, at least approximately, how the Coriolis parameter depends on position?

Put otherwise, suppose that  $u$  is an assigned function  $u(f)$  of the Coriolis parameter  $f$ , and that  $f$  is an assigned function  $f(x_i)$  of the  $x_i$ , so that  $u(x_i) = u[f(x_i)]$  is also assigned. Suppose also that functions  $z_i(x_j)$  can be hypothesized (representing the role served by (3.17) in the  $f$ -plane case). Then, if we refrain from assigning the function  $F(s_i)$ , equation (7.29) becomes an equation

$$u(x_i) = F[x_i u(x_i) + z_i(x_j)], \quad (7.33)$$

to be satisfied by  $F(s_i)$  for all  $x_i$ . Under what circumstances can this be solved to find  $F(s_i)$ ? Does (7.23) offer a helpful guide to a generalization of theorem 3.3 to the non- $f$ -plane case?

Dynamical systems associated with other generalizations of the geostrophic coordinate transformation have been considered in the literature, for example by Salmon (1985), Magnusdottir & Schubert (1990, 1991) and Purser (1988, 1993). Future work will be needed to study them in the context of the guidance offered by the lift structure explained here and in Sewell & Roulstone (1994).

## 8. Concluding remarks

This paper identifies the interconnection of ideas which make up the mathematical structure of the semigeostrophic equations. This structure has emerged from studies of this model which were motivated for a variety of different reasons. It is fundamental to the analytical and numerical methods for integrating the semigeostrophic equations, including the initiation and development in time of discontinuous features such as fronts. Of course, the semigeostrophic equations are only one approximation to Newton's second law, but semigeostrophic theory has had substantial successes.

The transformation properties and Hamiltonian structure described here have prompted investigations into alternative, more accurate, balanced models that possess similar mathematical features to semigeostrophic theory (McIntyre & Roulstone 1997). In particular, a general theory for the construction of canonical coordinates for balanced models analogous to the  $X_i$  of (3.9) emerges. Explicit examples of theories with such coordinates, including the semigeostrophic equations, have been shown to possess quaternionic structure (Rubtsov & Roulstone 1997). A widely accessible review of some properties of duality, including Legendre duality and others, and lift, contact and canonical transformations as they apply to mechanics, has recently been given by Sewell (1997).

## References

- Blumen, W. 1981 The geostrophic coordinate transformation. *J. Atmos. Sci.* **38**, 1100–1105.  
 Broad, A. S., Porter, D. & Sewell, M. J. 1992 A new approach to shallow flow over an obstacle. I. General theory. U.K. Meteorological Office Short Range Forecasting Division Scientific paper no. 11, pp. 54.

*Phil. Trans. R. Soc. Lond. A* (1997)

- Broad, A. S., Porter, D. & Sewell, M. J. 1994 Shallow flow over general topography with applications to monotonic mountains. *Proc. 4th IMA Conf. on Stably Stratified Flows* (ed. I. Castro & N. Rockliff), pp. 133–138. Oxford University Press.
- Broad, A. S., Porter, D. & Sewell, M. J. 1997 New shallow flows over an obstacle. *Q. Jl Appl. Math.* **50**, 625–653.
- Carathéodory, C. 1982 *Calculus of variations and partial differential equations of the first order*, 2nd edn (revised). New York: Chelsea.
- Chynoweth, S., Porter, D. & Sewell, M. J. 1988 The parabolic umbilic and atmospheric fronts. *Proc. R. Soc. Lond. A* **419**, 337–362.
- Chynoweth, S. & Sewell, M. J. 1989 Dual variables in semigeostrophic theory. *Proc. R. Soc. Lond. A* **424**, 155–186.
- Chynoweth, S. & Sewell, M. J. 1991 A concise derivation of the semigeostrophic equations. *Q. Jl R. Met. Soc.* **117**, 1109–1128.
- Cullen, M. J. P., Norbury, J. & Purser, R. J. 1991 Generalised Lagrangian solutions for atmospheric and oceanic flows. *SIAM J. Appl. Math.* **51**, 20–31.
- Cullen, M. J. P. & Purser, R. J. 1984 An extended Lagrangian theory of semigeostrophic frontogenesis. *J. Atmos. Sci.* **41**, 1477–1497.
- Eliassen, A. 1948 The quasi-static equations of motion with pressure as independent variable. *Geophys. Publ.* **17**, 1–44.
- Eliassen, A. 1962 On the vertical circulation in frontal zones. *Geophys. Publ.* **24**, 147–160.
- Fjørtoft, R. 1962 On the integration of a system of geostrophically balanced prognostic equations. *Proc. Intern. Symp. Numerical Weather Prediction*, pp. 153–159. Tokyo: Meteorological Society of Japan.
- Forsyth, A. R. 1890 *Theory of differential equations. Part I. Exact equations and Pfaff's problem*. Cambridge University Press.
- Forsyth, A. R. 1906 *Theory of differential equations. Part IV. Partial differential equations*. Cambridge University Press.
- Hoskins, B. J. 1971 Atmospheric frontogenesis models: some solutions. *Q. Jl R. Met. Soc.* **97**, 139–153.
- Hoskins, B. J. 1975 The geostrophic momentum approximation and the semigeostrophic 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.
- Magnusdottir, G. & Schubert, W. H. 1990 The generalization of semigeostrophic theory to the  $\beta$ -plane. *J. Atmos. Sci.* **47**, 1714–1720.
- Magnusdottir, G. & Schubert, W. H. 1991 Semigeostrophic theory on the hemisphere. *J. Atmos. Sci.* **48**, 1449–1456.
- McIntyre, M. E. & Roulstone, I. 1997 Hamiltonian balanced models: constraints, slow manifolds and velocity splitting. Forecasting Research Scientific Paper 41, Meteorological Office, UK. *J. Fluid Mech.* (Submitted.)
- Purser, R. J. 1988 Variational aspects of semigeostrophic theory. Meteorological Office Forecasting Research Scientific Note no. 5, pp. 20.
- Purser, R. J. 1993 Contact transformations and Hamiltonian dynamics in generalized semigeostrophic theories. *J. Atmos. Sci.* **50**, 1449–1468.
- Purser, R. J. & Cullen, M. J. P. 1987 A duality principle in semigeostrophic theory. *J. Atmos. Sci.* **44**, 3449–3468.
- Roulstone, I. & Sewell, M. J. 1996 Potential vorticities in semigeostrophic theory. *Q. Jl R. Met. Soc.* **122**, 983–992.
- Rubtsov, V. N. & Roulstone, I. 1997 Examples of quaternionic and Kähler structures in Hamiltonian models of nearly geostrophic flow. *J. Phys. A* **30**, L63–L68.
- Salmon, R. 1983 Practical use of Hamilton's principle. *J. Fluid Mech.* **132**, 431–444.
- Salmon, R. 1985 New equations for nearly geostrophic flow. *J. Fluid Mech.* **153**, 461–477.
- Sewell, M. J. 1987 *Maximum and minimum principles. A unified approach, with applications*. Cambridge University Press.

*The mathematical structure of theories of semigeostrophic type* 2517

- Sewell, M. J. 1997 Some applications of transformation theory in mechanics. UK Meteorological Office Forecasting Research Division Scientific Paper. In *Proc. Isaac Newton Institute Programme on the Mathematics of Atmosphere and Ocean Dynamics, Cambridge, July–December 1996*. Cambridge University Press.
- Sewell, M. J. & Porter, D. 1980 Constitutive surfaces in fluid mechanics. *Math. Proc. Camb. Phil. Soc.* **88**, 517–546.
- Sewell, M. J. & Roulstone, I. 1993 Anatomy of the canonical transformation. *Phil. Trans. R. Soc. Lond. A* **345**, 577–598.
- Sewell, M. J. & Roulstone, I. 1994 Families of lift and contact transformations. *Proc. R. Soc. Lond. A* **447**, 493–512.
- Shutts, G. J. 1989 Planetary semigeostrophic 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 semigeostrophic theory. *J. Atmos. Sci.* **44**, 1318–1330.
- Stoker, J. J. 1957 *Water waves*. New York: Interscience.

*Received 29 May 1996; revised 8 January 1997; accepted 10 June 1997*

MATHEMATICAL,  
PHYSICAL  
& ENGINEERING  
SCIENCES

THE ROYAL  
SOCIETY

PHILOSOPHICAL  
TRANSACTIONS  
OF

MATHEMATICAL,  
PHYSICAL  
& ENGINEERING  
SCIENCES

THE ROYAL  
SOCIETY

PHILOSOPHICAL  
TRANSACTIONS  
OF