GENERALIZED SEMIGEOSTROPHIC THEORY

by

Gudrun Magnusdottir

Department of Atmospheric Science
Colorado State University
Fort Collins, Colorado

Colorado State University

Department of Atmospheric Science

Paper No. 459

GENERALIZED SEMIGEOSTROPHIC THEORY

by

Gudrun Magnusdottir

Department of Atmospheric Science
Colorado State University
Fort Collins, CO 80523

This research was supported by the National Science Foundation
under Grant No. ATM–8814541

Atmospheric Science Paper No. 459

November 1989

ABSTRACT

The isentropic/vortex coordinate version of semigeostrophic theory is developed on the β-plane and on the sphere. In both cases this approach results in a simple mathematical form in which the horizontal ageostrophic velocities are implicit and the entire dynamics reduce to a predictive equation for the potential pseudodensity and an invertibility relation. Linearized versions of both theories lead to a generalized Charney-Stern theorem for barotropic/baroclinic instability and to Rossby wave solutions with a meridional structure which for the β-plane case is different from that in quasi-geostrophic theory and for the spherical case is different from spherical harmonics. In both cases the equator represents a singular point, so that a more accurate term for the spherical theory is hemispheric theory.

By applying Hamilton's principle to a Lagrangian which approximates the wind by the geostrophic wind and to which a coordinate transformation has been applied, it is shown how an entire class of approximate balanced models can be generated which differ only in the way in which the balance conditions together with the horizontal coordinate transformations are defined. The equations of motion for the long wave approximation are derived from a Lagrangian which neglects the meridional wind. These models share an essential characteristic — they all conserve total energy and potential vorticity.

A limitation inherent in semigeostrophic theory involves the geostrophic momentum approximation. It neglects curvature vorticity compared to shear vorticity and assumes geostrophic balance, which breaks down on the equator. In an attempt to move towards a globally balanced theory a two dimensional zonally symmetric model in gradient wind balance is used to study the Hadley cell and the conditions which are created by heating alone and which are favorable for the barotropic/baroclinic instability described by the Charney-Stern theorem.

ACKNOWLEDGEMENTS

I would like to thank my adviser, Professor Wayne H. Schubert, for his guidance and support during my studies of atmospheric dynamics. Many thanks also go to the other members of my graduate advisory committee, Professors Duane E. Stevens, Elmar R. Reiter and Gerald D. Taylor, for their careful review of this research. I have benefitted greatly from discussions with many past and present graduate students at Colorado State whom I hereby thank collectively. Additionally, a special note of thanks goes to Paul Ciesielski for much help throughout my studies. Finally, I thank Jenny Martin and Gail Cordova for help with preparing the manuscript.

This work was supported by the National Science Foundation, under grant ATM–8814541.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

The set of primitive equations has been successfully used in numerical weather prediction for a number of years. The primitive equations are derived from the Euler equations of compressible fluid motion with only one filtering approximation, the assumption of hydrostatic balance, in addition to the traditional approximation (Phillips, 1966) for a shallow atmosphere on a sphere. Thus the primitive equations have a wide spectrum of allowable solutions ranging from the horizontally propagating Lamb waves to gravity modes to the slow rotational modes. Only the slow modes are of importance for the dynamical evolution of the larger scale flow with which numerical weather prediction is concerned. Therefore, the prediction models currently used are working "harder" than needed by predicting modes of atmospheric motions which are not necessary for the development of the large scale flow. More importantly the analysis of observations with which the time integrations are started must be balanced so as not to produce spurious accelerations. This last process is called initialization. Richardson (1922), who was the first to try his hand at numerical forecasting and who used the set of primitive equations, failed partially on account of improper initialization. Aside from practicalities associated with numerical weather prediction it is essential to be able to treat the slow modes of atmospheric motions separately for physical interpretations and understanding of the underlying physical principles without the complexities the full primitive equation set brings with it.

Rossby (1939) introduced the first approximate dynamical model of large scale flow, the barotropic vorticity model. It assumes that the vertical component of vorticity is conserved in two dimensional horizontal motion on a β-plane. This was later generalized to the sphere by Haurwitz (1940). Ertel (1942) proved the fundamental conservation principle that potential vorticity, P, in a stratified, adiabatic, frictionless fluid is conserved. The assumption of hydrostatic balance was not required in his proof. The question remained as to how to simplify the primitive equations while maintaining this conservation principle. Charney (1948) introduced the widely used and deservedly celebrated quasi-geostrophic system using scale analysis. The classical quasi-geostrophic system (as developed by Charney and Stern, 1962) provided the necessary theory for understanding baroclinic instability and other midlatitude large scale processes. This system of equations is balanced in that they can neither describe sound waves nor gravity waves. The quasi-geostrophic equations are energy conserving and possess a potential vorticity principle which is different from Ertel's. The advecting wind is horizontal and geostrophic, the static stability is horizontally uniform and the nonlinear stretching and twisting of vorticity is neglected. Thus, the quasi-geostrophic system can neither describe areas of large vorticity (compared to the Coriolis parameter) nor areas of rapidly changing static stability; however, the system has been used with a fair measure of success in situations where its assumptions seem to be violated.

Eliassen (1948) proposed approximating momentum in the primitive equations by geostrophic momentum, the so called geostrophic momentum approximation. Note that in contrast to quasi-geostrophic theory the advecting wind contains both geostrophic and ageostrophic components; the only difference from the primitive equation momentum equations is that what is being advected is geostrophic wind. This is analogous to the hydrostatic approximation in which the vertical component of momentum is neglected but vertical advection is retained. These equations are balanced, are energy conserving and have a potential vorticity principle. The definitions of energy and potential vorticity are analogous to the corresponding definitions for the primitive equations except that wherever the wind appears it is replaced by its geostrophic value. A drawback of the system is that it contains two prognostic equations which are not independent, and the ageostrophic wind is only implied. The system first blossomed out with the definition of the coordinate transformations of semigeostrophic theory — the geostrophic coordinates. In geostrophic coordinates the horizontal ageostrophic advection becomes implicit, regions of high vorticity are stretched, regions of low vorticity are shrunk, and potential vorticity takes on the role of static stability in the quasi-geostrophic system. The semigeostrophic equations as given by Hoskins (1975) and Hoskins and Draghici (1977) are almost as simple as the quasi-geostrophic equations, but are not limited by their restrictive assumptions and therefore apply to more general physical situations where large shear vorticity and nonuniform static stability are present, i.e., such as in fronts, jets, occluding baroclinic waves, etc. The two dimensional version of semigeostrophic theory with a height independent deformation field (Hoskins, 1971; Hoskins and Bretherton, 1972) simulates quite realistically both surface (uniform potential vorticity assumed) and upper level (discontinuous potential vorticity assumed) frontogenesis. The three dimensional theory with a uniform potential vorticity jet has produced unstable baroclinic waves evolving into the nonlinear regime with fronts and an occluding warm sector (Hoskins, 1976; Hoskins and West, 1979; Hoskins and Heckley, 1980). A nonuniform potential vorticity jet produces more realistic upper tropospheric frontogenesis (Heckley and Hoskins, 1982). Schubert (1985) proposed a computationally more convenient form of three dimensional semigeostrophic theory, the "geopotential tendency" form, which allows time integrations of non-Boussinesq, nonuniform potential vorticity flows to be performed almost as easily as for the Boussinesq, uniform potential vorticity case. More recently, Schubert et al. (1989) developed what may be the most elegant and concise version of semigeostrophic theory — that version which makes simultaneous use of geostrophic and isentropic coordinates. With these coordinates, semigeostrophic theory reduces to two equations: a predictive equation for the potential pseudodensity (or inverse potential vorticity) and a diagnostic equation (or invertibility principle) whose solution yields the balanced wind and mass fields from the potential pseudodensity. In vortex and isentropic coordinates the divergent part of the circulation remains entirely implicit. In simple situations the prognostic equation can be solved analytically offering valuable physical insight.

Two limitations of semigeostrophic theory are that it does not include a variable Coriolis parameter and that it proves inadequate in physical situations where the curvature vorticity is as large as the shear vorticity. This work is concerned with mending the first of the two limitations. An attempt is made to offer suggestions for direction of future work in mending the second, which would not be semigeostrophic theory since it would not include the geostrophic momentum approximation; rather it would be some different and more general approximation to momentum with more general balance equations. A two dimensional (axisymmetric) balanced theory for highly curved flows is the well known set of Eliassen's (1952) balanced vortex equations. Schubert and Hack (1983) derived a simpler version of it for studying tropical cyclones on an f-plane by introducing a coordinate transformation (potential radius) which results in similar simplifications as does the transformation to geostrophic coordinates in semigeostrophic theory, i.e., the stretching of high vorticity regions and potential vorticity taking on the role of static stability. Schubert and Alworth (1987) derived the version of this model in isentropic coordinates. Similar to semigeostrophic theory, using potential temperature as the vertical coordinate reduces the theory to two equations — a predictive equation for the potential pseudodensity (or inverse potential vorticity) and a diagnostic equation (or invertibility principle) whose solution yields the balanced wind and mass fields from the potential pseudodensity. In potential radius and isentropic coordinates the divergent part of the circulation remains entirely implicit. By neglecting friction and assuming a simple midtropospheric heating, Schubert and Alworth (1987) solved the prognostic equation analytically and inverted the potential pseudodensity field to produce quite a realistic vortex at around five days. The value of their study lies not least in the fact that it describes the evolution of potential vorticity induced by a tropical heat source (Haynes and McIntyre, 1987). More recently, Hack et al. (1989) have generalized the two dimensional f-plane theory of Schubert and Hack (1983) to the sphere, to study the low latitude zonally symmetric circulation or the Hadley cell. An isentropic coordinate version of this model is derived in the present study (chapter 6). Closely related to the balanced vortex models is the long wave approximation model developed by Gill (1980) and later extended (Heckley and Gill, 1984; Gill and Phlips, 1986; Stevens et al., 1989). The most general form of this model is as presented in Stevens et al. (1989); it is derived in the present work by applying Hamilton's principle to a Lagrangian which neglects the meridional wind (section 4.3). Consequently, the meridional wind neither enters the definition of kinetic energy nor potential vorticity which have exactly the same form as for the balanced vortex model on the sphere (section 6.1). In this way the long wave approximation model is two dimensional. However, it is not zonally symmetric and in that way it is three dimensional.

In figure 1.1 an attempt is made to summarize and clarify the discussion above. It shows what McWilliams and Ghent (1980) called intermediate models, or those whose accuracy lies somewhere between the quasi-geostrophic equations and the primitive equations. Obviously the figure is not completely comprehensive; it shows only some of the models considered to be of major importance in the development and generalization of balanced theory. The column on the left indicates two dimensional models, the one on the right three dimensional models. As you go up the page in this figure the models become more general. We start at the bottom with quasi-geostrophic theory and end on top with the primitive equations. The balanced models become more general as you go up the page in two ways. First, the earth's geometry is better represented, progressing from f-plane to β-plane to the full spherical representation. Secondly, the assumed balance becomes more general as we go from geostrophic balance to gradient wind balance. Each box in the figure indicates a particular intermediate model. The first line in the box indicates the type of balance, the second its application and the third when and where it first appeared. Globally valid three dimensional balanced theory, represented by the top right-hand side box, has not yet been discovered. In particular the diagram shows that the present work fills two boxes, the semigeostrophic theory on the β-plane and the semigeostrophic theory on the sphere. The methodology in deriving those two models was that of Salmon (1983, 1985, 1988). We started by writing out the Lagrangian of the system, made analytical approximations to it while conserving its symmetries and transformed the horizontal coordinates to almost canonical form. The resulting momentum equations, derived by Hamilton's principle, are almost canonical. This method is powerful in two respects. By conserving the symmetries in the Lagrangian we are guaranteed by Noether's theorem that the corresponding conservation laws are maintained and the coordinate transformation which is so important in semigeostrophic theory, arises quite naturally as exactly that one which renders the momentum equations almost canonical. Shutts (1988) has

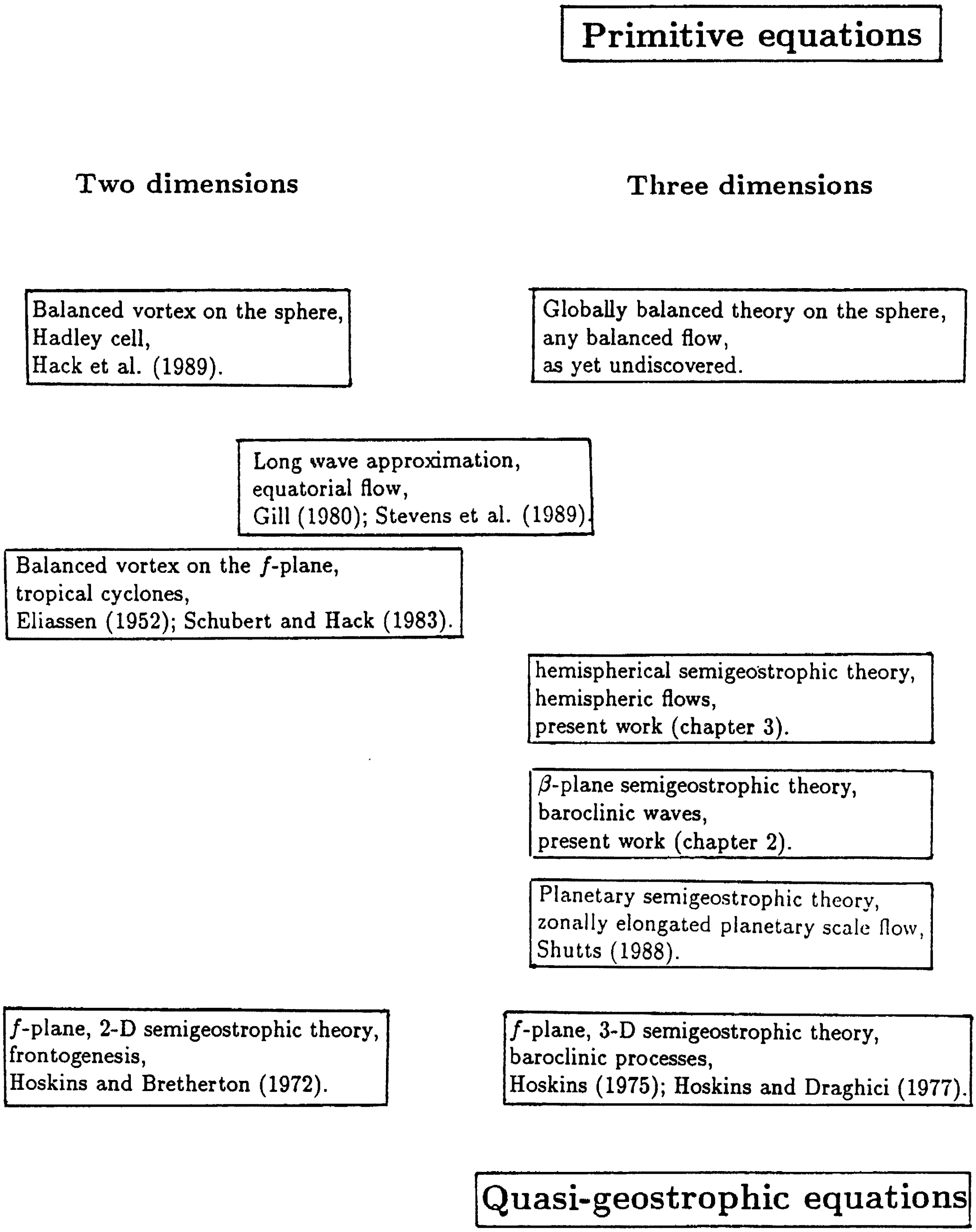

Figure 1.1: Development of the intermediate models considered to be of major importance in the theory of balanced flows. Each box represents a model and the higher it is on the diagram the more general it is. Listed in each box is the type of balance employed, the weather phenomena studied, and where and when the model was developed.

derived a three dimensional geostrophically balanced model using the Hamiltonian framework slightly differently. However, the validity of his model is somewhat restricted since the kinetic energy equation neglects the component of the geostrophic wind parallel to the earth's axis of rotation.

Chapter 2 introduces semigeostrophic theory on the β-plane and chapter 3 introduces it on the sphere in vortex and isentropic coordinates. The development is very similar in both cases. As in the f-plane case using those coordinates (Schubert et al., 1989), the dynamics reduce to one prognostic equation in potential pseudodensity (which is related to the inverse potential vorticity) and an invertibility principle. We show how the conservation principles of the primitive equations carry over to these approximate equations and consider the linear dynamics.

Chapter 4 addresses Hamilton's principle, the method used to arrive at the generalized geostrophic momentum approximation and the accompanying coordinate transformations of chapter 3. It is shown how an entire class of approximate models can be generated which differ only in the way in which the balance condition together with the horizontal coordinate transformations are defined. These models share an essential characteristic — they all conserve total energy and potential vorticity. Finally, the equations of the long wave approximation model are derived.

Chapter 5 proves the Charney-Stern theorem for both semigeostrophic systems. The proof is more general than the original proof of Charney and Stern (1962).

Chapter 6 discusses balanced models in the top line of the diagram in figure 1.1, i.e., models more general than the semigeostrophic and long wave approximation models. In particular, we derive the isentropic coordinate version of the theory developed by Hack et al. (1989). It is then used to study the evolution of potential vorticity on isentropic surfaces and the breakdown of the ITCZ. Finally, we offer some speculations on the direction of and the framework for future work on deriving a three dimensional globally valid balanced theory. This two dimensional theory would be a special case of the full three dimensional theory.

Chapter 7 contains some concluding remarks.

Chapter 2

SEMIGEOSTROPHIC THEORY ON THE β-PLANE

Here we explore the use of the geometric approximation of the β-plane within the realm of semigeostrophic theory. The Coriolis parameter is approximated by a linear function, such that

$$f(y) = f_0 + \beta y. \tag{2.1}$$

If $\beta = 0$ we have the usual f-plane approximation; if $f_0 = 0$ we are on the equatorial β-plane; if both f_0 and β are non-zero we are on the midlatitude β-plane.

Salmon (1985) extended semigeostrophic theory to a variable Coriolis parameter in Cartesian coordinates. However, his equation set was for a shallow water system and although his set can be considered closed it was not in a form convenient for computations or physical interpretations. In this chapter the isentropic/vortex coordinate version of semigeostrophic theory is developed on the β-plane. This approach results in a simple mathematical form where the ageostrophic wind is entirely implicit. The complete dynamics reduce to a form reminiscent of nondivergent barotropic dynamics — the prognostic equation becomes an equation predicting the potential pseudodensity and a diagnostic equation where a Laplacian-like operator is inverted.

The following section of the chapter reviews the primitive equations on a β-plane and their conservation relations. The next section defines the geostrophic momentum approximation generalized to the β-plane; conservation relations are derived and compared to the primitive equations results. Coordinate transformations are defined so as to make the ageostrophic wind implicit. This results in one very simple prognostic equation. Next the invertibility principle is derived for this system. Finally the linear dynamics of the system are considered.

2.1 The primitive equations

Using potential temperature as the vertical coordinate, the primitive equations on the β-plane can be written

$$\frac{Du}{Dt} - f(y)\,v + \frac{\partial M}{\partial x} = F, \tag{2.2}$$

$$\frac{Dv}{Dt} + f(y)\,u + \frac{\partial M}{\partial y} = G, \tag{2.3}$$

$$\frac{\partial M}{\partial \theta} = \Pi, \tag{2.4}$$

$$\frac{D\sigma}{Dt} + \sigma\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \dot{\theta}}{\partial \theta}\right) = 0, \tag{2.5}$$

where u and v are the zonal and meridional components of the velocity, F and G the components of the frictional force per unit mass, $\Pi = c_p\,(p/p_{00})^\kappa$ the Exner function, $M = \theta\Pi + gz$ the Montgomery potential, $\sigma = -\partial p/\partial\theta$ the pseudodensity, and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \dot{\theta}\frac{\partial}{\partial \theta} \tag{2.6}$$

the total derivative.

2.1.1 Conservation relations

The kinetic energy equation can be derived by combining (2.2) and (2.3) to get

$$\frac{DK}{Dt} + u\frac{\partial M}{\partial x} + v\frac{\partial M}{\partial y} = uF + vG, \tag{2.7}$$

where $K \equiv \frac{1}{2}(u^2 + v^2)$ is the kinetic energy per unit mass. Combining (2.7) with the continuity equation (2.5) gives

$$\frac{\partial(\sigma K)}{\partial t} + \frac{\partial(\sigma u K)}{\partial x} + \frac{\partial(\sigma v K)}{\partial y} + \frac{\partial(\sigma\dot{\theta}K)}{\partial \theta} + \sigma u\frac{\partial M}{\partial x} + \sigma v\frac{\partial M}{\partial y} = \sigma\,(uF + vG)\,. \tag{2.8}$$

Using the continuity and hydrostatic equations the kinetic energy equation can be written

$$\frac{\partial}{\partial t}(\sigma K) + \frac{\partial}{\partial x}\Big(\sigma u(K+\phi)\Big) + \frac{\partial}{\partial y}\Big(\sigma v(K+\phi)\Big) + \frac{\partial}{\partial \theta}\Big(\sigma\dot{\theta}(K+\phi) - \phi\frac{\partial p}{\partial t}\Big) = \sigma(uF + vG - \alpha\omega), \tag{2.9}$$

where $\omega = Dp/Dt$.

For deriving the thermodynamic energy equation, multiply (2.5) by c_pT to obtain

$$\frac{\partial}{\partial t}(\sigma c_p T) + \frac{\partial}{\partial x}(\sigma u c_p T) + \frac{\partial}{\partial y}(\sigma v c_p T) + \frac{\partial}{\partial \theta}(\sigma \dot{\theta} c_p T) = \sigma(Q + \alpha\omega), \tag{2.10}$$

where $Q = \Pi\dot{\theta}$. Adding (2.9) and (2.10) we obtain the total energy equation

$$\frac{\partial}{\partial t}\Big(\sigma(K + c_pT)\Big) + \frac{\partial}{\partial x}\Big(\sigma u(K+M)\Big) + \frac{\partial}{\partial y}\Big(\sigma v(K+M)\Big) + \frac{\partial}{\partial \theta}\Big(\sigma\dot{\theta}(K+M) - \phi\frac{\partial p}{\partial t}\Big)$$
$$= \sigma(uF + vG + Q). \tag{2.11a}$$

Before integrating (2.11a), we adopt an idea which has proved useful in such contexts as the definition of available potential energy (Lorenz, 1955), the analysis of baroclinic instability (Bretherton, 1966; Hoskins et al., 1985; Hsu, 1988), and the finite amplitude Eliassen-Palm theorem (Andrews, 1983). The idea involves what happens when an isentropic surface intersects the earth's surface. We can regard such an isentrope as continuing just under the earth's surface with a pressure equal to the surface pressure. At any horizontal position where two distinct isentropic surfaces run just under the earth's surface (and hence have the same pressure), there is no mass trapped between them, so that $\sigma = 0$ there. Let us regard the bottom isentropic surface θ_B as the largest value of θ which remains everywhere below the earth's surface. Assuming the top boundary θ_T is both an isentropic and isobaric surface and assuming no topography and vanishing $\dot{\theta}$ at the top and bottom, we can integrate (2.11a) over the entire atmosphere to obtain

$$\frac{\partial}{\partial t}\iiint (K + c_pT)\sigma dx dy d\theta = \iiint (uF + vG + Q)\sigma dx dy d\theta. \tag{2.11b}$$

The equation for the absolute isentropic vorticity ζ can be derived from (2.2) and (2.3). It takes the form

$$\frac{D\zeta}{Dt} + \zeta\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) = \left(\xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y}\right)\dot{\theta} + \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}, \tag{2.12}$$

where

$$(\xi, \eta, \zeta) = \left(-\frac{\partial v}{\partial \theta}, \frac{\partial u}{\partial \theta}, f(y) + \frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right). \tag{2.13}$$

Equation (2.12) can be written in the alternative flux form

$$\frac{\partial(\sigma P)}{\partial t} + \frac{\partial\left(u\sigma P - \xi\dot{\theta} - G\right)}{\partial x} + \frac{\partial\left(v\sigma P - \eta\dot{\theta} + F\right)}{\partial y} = 0, \tag{2.14}$$

where $P = \zeta/\sigma$ is the Rossby-Ertel potential vorticity. The equivalence of (2.12) and (2.14) follows easily from the fact that

$$\frac{\partial \xi}{\partial x} + \frac{\partial \eta}{\partial y} + \frac{\partial \zeta}{\partial \theta} = 0.$$

The significance of (2.14) has recently been discussed by McIntyre (1987) and Haynes and McIntyre (1987), who emphasize the fact that for the primitive equations in the isentropic coordinate the flux form (2.14) leads directly to the notion that even when mass is crossing isentropic surfaces the potential vorticity flux is exactly isentropic (i.e., this flux is *along* the isentropic surface). Thus, the Haynes-McIntyre theorem states that even with diabatic heating and frictional forces "there can be no net transport of Rossby-Ertel potential vorticity across any isentropic surface" and that "potential vorticity can neither be created nor destroyed within a layer bounded by two isentropic surfaces". In this sense, an isentropic surface is impermeable to potential vorticity, and the potential vorticity in a layer between two isentropic surfaces is indestructible as long as the layer does not meet a boundary such as the earth's surface. Creation or destruction of potential vorticity within this layer can only occur at the ground.

We can now eliminate the horizontal divergence between (2.5) and (2.12) to obtain the potential vorticity equation,

$$\frac{DP}{Dt} = \frac{1}{\sigma}\left[\xi\frac{\partial \dot{\theta}}{\partial x} + \eta\frac{\partial \dot{\theta}}{\partial y} + \zeta\frac{\partial \dot{\theta}}{\partial \theta} + \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right]. \tag{2.15}$$

In the absence of friction and heating the potential vorticity P is conserved. Alternatively, we can derive an equation involving the inverse potential vorticity.

$$\frac{D\sigma^*}{Dt} + \frac{\sigma^*}{\zeta}\left[\left(\xi\frac{\partial}{\partial x} + \eta\frac{\partial}{\partial y} + \zeta\frac{\partial}{\partial \theta}\right)\dot{\theta} + \frac{\partial G}{\partial x} - \frac{\partial F}{\partial y}\right] = \frac{\sigma^*}{f(y)}\beta v, \tag{2.16}$$

where

$$\sigma^* = \sigma\frac{f(y)}{\zeta} \tag{2.17}$$

is the potential pseudodensity. The potential pseudodensity on an f-plane was discussed by Schubert et al. (1989). In that case the right hand side of (2.16) vanishes so that σ^* is the pseudodensity the fluid element would acquire if ζ were changed to the constant f

under a frictionless and adiabatic rearrangement process. In this case of a variable Coriolis parameter we have a β-term to contend with. However, in the semigeostrophic system with the appropriate coordinate transformations, the geostrophic potential pseudodensity equation takes on a very simple form. In fact as we shall see from the simplicity of (2.43) σ_g^* seems to be a more convenient variable than the more commonly used potential vorticity.

2.2 The semigeostrophic equations on the β-plane

As approximations to the isentropic coordinate version of the primitive equations on the β-plane let us consider

$$\frac{Du_g}{Dt} - \Big(f(Y)v + \beta(y-Y)v_g\Big) + \frac{\partial M}{\partial x} = F, \tag{2.18}$$

$$\frac{Dv_g}{Dt} + \Big(f(Y)u + \beta(y-Y)u_g\Big) + \frac{\partial M}{\partial y} = G, \tag{2.19}$$

where (u_g, v_g) are the geostrophic wind components, given by

$$\left(f(Y)v_g, -f(Y)u_g\right) = \left(\frac{\partial M}{\partial x}, \frac{\partial M}{\partial y}\right), \tag{2.20}$$

and Y is defined by

$$Y = y - \frac{u_g}{f(Y)}. \tag{2.21}$$

Note that (2.18) and (2.19) revert to the primitive equations if u_g, v_g and Y are replaced by u, v and y, and to the f-plane geostrophic momentum approximation when $\beta = 0$ (Eliassen, 1948; Hoskins, 1975). Now the definition of the geostrophic wind involves evaluating the Coriolis term at the transformed latitude, Y. The β terms in (2.18)–(2.19) can be regarded as corrections for the fact that f is taken at Y rather than y. The generalized geostrophic momentum approximation preserves important conservation principles of the primitive equations. Additionally, an accompanying coordinate transformation will lead to a simple prognostic equation.

2.2.1 Conservation relations

Derivation of the kinetic energy equation proceeds along exactly the same lines as for the primitive equations. Combining (2.18) and (2.19) and now defining

$$K \equiv \frac{1}{2}\left(u_g^2 + v_g^2\right) \tag{2.22}$$

we obtain (2.7). Thus, (2.11) is valid even when the generalized geostrophic momentum approximation has been made provided K is defined as in (2.22).

The equation for the absolute isentropic geostrophic vorticity ζ_g can be derived from (2.18) and (2.19). It takes the form

$$\begin{aligned}\frac{D\zeta_g}{Dt} + \zeta_g\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) - \left(\xi_g\frac{\partial}{\partial x} + \eta_g\frac{\partial}{\partial y}\right)\dot{\theta} \\ = \frac{\partial}{\partial x}\left[\frac{\partial X}{\partial y}F + \frac{\partial Y}{\partial y}G\right] - \frac{\partial}{\partial y}\left[\frac{\partial X}{\partial x}F + \frac{\partial Y}{\partial x}G\right],\end{aligned} \tag{2.23}$$

where

$$(\xi_g, \eta_g, \zeta_g) = f(Y)\left(\frac{\partial(X,Y)}{\partial(y,\theta)}, \frac{\partial(X,Y)}{\partial(\theta,x)}, \frac{\partial(X,Y)}{\partial(x,y)}\right). \tag{2.24}$$

with X defined below in (2.27). Equation (2.23) can be written in the potential vorticity form

$$\begin{aligned}\frac{\partial(\sigma P_g)}{\partial t} + \frac{\partial}{\partial x}\left(u\sigma P_g - \xi_g\dot{\theta} - \frac{\partial X}{\partial y}F - \frac{\partial Y}{\partial y}G\right) \\ + \frac{\partial}{\partial y}\left(v\sigma P_g - \eta_g\dot{\theta} + \frac{\partial X}{\partial x}F + \frac{\partial Y}{\partial x}G\right) = 0,\end{aligned} \tag{2.25}$$

where $P_g = \zeta_g/\sigma$ is the geostrophic Rossby-Ertel potential vorticity. The equivalence of (2.23) and (2.25) follow easily from the identity

$$\frac{\partial \xi_g}{\partial x} + \frac{\partial \eta_g}{\partial y} + \frac{\partial \zeta_g}{\partial \theta} = 0.$$

Equation (2.25) is the semigeostrophic equivalent of equation (2.14). From the above we conclude that the primitive equation result of Haynes and McIntyre also holds when we make the geostrophic momentum approximation.

We can now eliminate the horizontal divergence between (2.5) and (2.23) to obtain

$$\frac{D(\sigma/\zeta_g)}{Dt} + \frac{(\sigma/\zeta_g)}{\zeta_g}\left(\xi_g\frac{\partial}{\partial x} + \eta_g\frac{\partial}{\partial y} + \zeta_g\frac{\partial}{\partial \theta}\right)\dot{\theta} = \frac{(\sigma/\zeta_g)}{f(Y)}\left[\frac{\partial F}{\partial Y} - \frac{\partial G}{\partial X}\right]. \tag{2.26}$$

This is an equation for the geostrophic potential pseudodensity. The coordinate transformation in the next subsection will greatly simplify this equation in that it makes the ageostrophic advection totally implicit.

2.2.2 Coordinate transformation

Hoskins and Draghici (1977) first pointed out the duality between the use of geostrophic coordinates in the horizontal and the isentropic coordinate in the vertical. This duality has been further discussed by Gill (1981) and Heckley and Hoskins (1982). The combined use of geostrophic and isentropic coordinates has been discussed theoretically by McWilliams and Gent (1980) and has found application in the two dimensional upper tropospheric frontogenesis study of Buzzi et al. (1981) and the two dimensional response to squall lines study of Schubert et al. (1989). Similar to the f-plane study of Schubert et al. (1989), we are concerned with the simultaneous use of vortex coordinates — which on the f-plane are the geostrophic coordinates — and isentropic coordinates, because it will lead to an elegant version of the geostrophic potential pseudodensity equation. Thus, let us introduce the vortex coordinates X and Y which are in fact Salmon's generalized geostrophic coordinates

$$(X,Y) = \left(x + \frac{v_g}{f(Y)}, y - \frac{u_g}{f(Y)}\right). \tag{2.27}$$

Derivatives in (x, y, θ, t) space are then related to derivatives in (X, Y, Θ, T) space by

$$\frac{\partial}{\partial t} = \frac{\partial X}{\partial t}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial t}\frac{\partial}{\partial Y} + \frac{\partial}{\partial T}, \tag{2.28}$$

$$\frac{\partial}{\partial x} = \frac{\partial X}{\partial x}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial x}\frac{\partial}{\partial Y}, \tag{2.29}$$

$$\frac{\partial}{\partial y} = \frac{\partial X}{\partial y}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial y}\frac{\partial}{\partial Y}, \tag{2.30}$$

$$\frac{\partial}{\partial \theta} = \frac{\partial X}{\partial \theta}\frac{\partial}{\partial X} + \frac{\partial Y}{\partial \theta}\frac{\partial}{\partial Y} + \frac{\partial}{\partial \Theta}, \tag{2.31}$$

since $T = t$ and $\Theta = \theta$. Inverting (2.29) and (2.30) to obtain

$$\frac{\partial(X,Y)}{\partial(x,y)}\frac{\partial}{\partial X} = \frac{\partial Y}{\partial y}\frac{\partial}{\partial x} - \frac{\partial Y}{\partial x}\frac{\partial}{\partial y} \tag{2.32}$$

$$\frac{\partial(X,Y)}{\partial(x,y)}\frac{\partial}{\partial Y} = -\frac{\partial X}{\partial y}\frac{\partial}{\partial x}+\frac{\partial X}{\partial x}\frac{\partial}{\partial y}, \tag{2.33}$$

and applying (2.28), (2.31), (2.32) and (2.33) to the Bernoulli function

$$M^* = M + \frac{1}{2}\left(u_g^2 + v_g^2\right), \tag{2.34}$$

it can be shown that

$$\left(\frac{\partial M}{\partial x},\frac{\partial M}{\partial y},\frac{\partial M}{\partial \theta},\frac{\partial M}{\partial t}\right) = \left(\frac{\partial M^*}{\partial X},\frac{\partial M^*}{\partial Y} - \beta\frac{(u_g^2+v_g^2)}{f(Y)},\frac{\partial M^*}{\partial \Theta},\frac{\partial M^*}{\partial T}\right). \tag{2.35}$$

We can also use (2.32) and (2.33) in (2.34), along with (2.24), to obtain

$$\zeta_g\frac{\partial}{\partial\Theta} = \xi_g\frac{\partial}{\partial x} + \eta_g\frac{\partial}{\partial y} + \zeta_g\frac{\partial}{\partial\theta}, \tag{2.36}$$

which shows that $\partial/\partial\Theta$ is actually the derivative along the vorticity vector and thus the name "vortex coordinates" for (X,Y). Note that (2.36) leads to considerable simplifications. Combining the vorticity and continuity equations one obtains the geostrophic potential vorticity equation,

$$\frac{DP_g}{Dt} - P_g\frac{\partial\dot{\theta}}{\partial\Theta} = \frac{P_g}{f(Y)}\left(\frac{\partial G}{\partial X} - \frac{\partial F}{\partial Y}\right), \tag{2.37}$$

where $P_g = \zeta_g/\sigma$.

The transformation relations (2.28)–(2.31) also imply that the operator (2.6) can be written as

$$\frac{D}{Dt} = \frac{\partial}{\partial T} + \frac{DX}{Dt}\frac{\partial}{\partial X} + \frac{DY}{Dt}\frac{\partial}{\partial Y} + \dot{\Theta}\frac{\partial}{\partial\Theta}. \tag{2.38}$$

Writing (2.26) in terms of σ_g^* where we define the geostrophic potential pseudodensity as

$$\sigma_g^* = \frac{f(Y)}{\zeta_g}\sigma, \tag{2.39}$$

and using (2.37) we get

$$\frac{\partial\sigma_g^*}{\partial T} + \frac{DX}{Dt}\frac{\partial\sigma_g^*}{\partial X} + \frac{DY}{Dt}f(Y)\frac{\partial}{\partial Y}\left(\frac{\sigma_g^*}{f(Y)}\right) + \frac{\partial}{\partial\Theta}\left(\dot{\theta}\sigma_g^*\right)$$
$$+\frac{\sigma_g^*}{f(Y)}\left[\frac{\partial G}{\partial X} - \frac{\partial F}{\partial Y}\right] = 0. \tag{2.40}$$

Using (2.18), (2.19) , (2.20), (2.27) and (2.35) it can easily be shown that

$$f(Y)\frac{DY}{Dt} = \frac{\partial M^*}{\partial X} - F, \tag{2.41}$$

$$-f(Y)\frac{DX}{Dt} = \frac{\partial M^*}{\partial Y} - G. \tag{2.42}$$

A major advantage of the transformation from (x, y, θ, t) space to (X, Y, Θ, T) space is the change from advection by (u, v) to advection by $(DX/Dt, DY/Dt)$ given in (2.41) and (2.42). The total time derivative does not contain any ageostrophic advection, thus (2.40), the geostrophic potential pseudodensity equation, does not contain any ageostrophic advection.

Using the canonical momentum equations (2.41) and (2.42) we can now write the geostrophic potential pseudodensity equation (2.40) in flux form as

$$\frac{\partial \sigma_g^*}{\partial T} + \frac{\partial}{\partial X}\left(\sigma_g^*\frac{DX}{Dt}\right) + \frac{\partial}{\partial Y}\left(\sigma_g^*\frac{DY}{Dt}\right) + \frac{\partial}{\partial \Theta}\left(\sigma_g^*\dot{\theta}\right) = 0. \tag{2.43}$$

The horizontal flux terms can be written in the Jacobian form

$$\frac{\partial \sigma_g^*}{\partial T} + \frac{\partial\left(M^*, \left(\sigma_g^*/f(Y)\right)\right)}{\partial (X, Y)} + \frac{\partial}{\partial X}\left(\frac{\sigma_g^*}{f(Y)}G\right)$$

$$+\frac{\partial}{\partial Y}\left(-\frac{\sigma_g^*}{f(Y)}F\right) + \frac{\partial}{\partial \Theta}\left(\sigma_g^*\dot{\theta}\right) = 0, \tag{2.44}$$

which serves as the fundamental predictive equation of the model.

2.2.3 Invertibility principle

The geostrophic potential pseudodensity σ_g^* is a combination of the mass field σ and the geostrophic wind field $\partial(X, Y)/\partial(x, y)$. However, since σ is related to M^* through hydrostatic balance and $\partial(X, Y)/\partial(x, y)$ is related to M^* through (2.24) and geostrophic balance, σ_g^* depends only on M^*. Thus, everything can be obtained from σ_g^* if we can somehow invert it to obtain M^*. The relation between M^* and σ_g^* is derived as follows. From the definition of σ_g^* we have

$$\frac{f(Y)}{\zeta_g}\frac{\partial \Pi}{\partial \theta} + \Gamma\sigma_g^* = 0, \tag{2.45}$$

where $\Gamma = d\Pi/dp = \kappa\Pi/p$. Using (2.24) this last equation can be written

$$\frac{\partial(x,y,\Pi)}{\partial(X,Y,\Theta)} + \Gamma\sigma^* = 0. \tag{2.46}$$

Using the geostrophic, hydrostatic and coordinate transformation relations, we can express x, y and Π in terms of M^* as

$$(x,y,\Pi) = \left(X - f^{-2}M^*_X\,,\, Y - \hat{f}^{-2}(M^*_Y - \beta f^{-3}M^{*\,2}_X)\,,\, M^*_\Theta\right), \tag{2.47}$$

where we have used the shorthand notation $\hat{f}^2 = f(Y)(2f(Y) - f(y))$ and $f = f(Y)$. Substituting (2.47) into (2.46), we obtain

$$\frac{1}{f^2\hat{f}^2}\begin{vmatrix} M^*_{XX} - f^2 & \hat{f}^2\left(\hat{f}^{-2}(M^*_Y - \beta f^{-3}M^{*\,2}_X)\right)_X & M^*_{\Theta X} \\ f^2\,(f^{-2}M^*_X)_Y & \hat{f}^2\left(\hat{f}^{-2}(M^*_Y - \beta f^{-3}M^{*\,2}_X)\right)_Y - \hat{f}^2 & M^*_{\Theta Y} \\ M^*_{X\Theta} & \hat{f}^2\left(\hat{f}^{-2}(M^*_Y - \beta f^{-3}M^{*\,2}_X)\right)_\Theta & M^*_{\Theta\Theta} \end{vmatrix} + \Gamma\sigma^* = 0, \tag{2.48a}$$

which expresses the invertibility principle in terms of the determinant of a Hessian-type matrix. Shutts and Cullen (1987) have discussed in detail the relation of hydrodynamic stability and the positive definiteness of such matrices. The upper boundary is assumed to be an isentropic and isobaric surface with potential temperature Θ_T and pressure p_T, thus the upper boundary condition for (2.48a) is simply

$$M^*_\Theta = \Pi(p_T) \quad \text{at} \quad \Theta = \Theta_T. \tag{2.48b}$$

Since we are neglecting the effects of topography and assuming that the lower boundary is the constant height surface $z = 0$ and the isentropic surface $\Theta = \Theta_B$, then $M = \Theta\Pi$ at $\Theta = \Theta_B$. Written in terms of M^*, this lower boundary condition becomes

$$\Theta M^*_\Theta - M^* + \frac{1}{2f^2}\left[M^{*\,2}_X + \left(M^*_Y + 2\beta f^{-1}\left(\Theta M^*_\Theta - M^*\right)\right)^2\right] = 0 \quad \text{at} \quad \Theta = \Theta_B. \tag{2.48c}$$

The lateral boundary conditions depend on the particular application, but typically might consist of a zonally periodic midlatitude region with $v_g = 0$ on the southern and northern boundaries (or $v_g \to 0$ as $Y \to \infty$). In any event, for a given σ^*_g, we can regard (2.48) as a nonlinear second order problem in M^*. Note that Γ and $\hat{f}$ both depend on M^*. The mathematical problem (2.48) is a generalization of the f-plane case discussed by Schubert et al. (1989). In particular, when f is assumed to be a constant, (2.48) reduces to their (2.14), while if we further assume that $\partial/\partial Y = 0$, the middle elements of the first and third rows and the first and third columns vanish, in which case (2.48) reduces to their (3.2). An efficient multigrid solver for the two-dimensional f-plane case has been developed by Fulton (1989).

Equations (2.44) and (2.48) form a closed system for the prediction of σ_g^* and the diagnosis of M^*. Since the problem of isentropes intersecting the earth's surface has been addressed by adopting the massless region approach outlined in section 2.1.1, the system (2.44) and (2.48) can in principle handle surface frontogenesis. Since $\sigma = 0$ in the massless region, $\sigma_g^* = 0$ there also. Thus, the prediction of σ_g^* by (2.44) includes predicting the movement of the $\sigma_g^* = 0$ region. This procedure is consistent with Bretherton's (1966) notion that "any flow with potential temperature variations over a horizontal rigid plane boundary may be considered equivalent to a flow without such variations, but with a concentration of potential vorticity very close to the boundary". We have simply replaced Bretherton's thin sheet of infinite potential vorticity with a thin sheet of zero potential pseudodensity and chosen to predict the evolution of the entire σ_g^* field (including this zero potential pseudodensity region) with (2.44). Of course, such a procedure has implications for the numerical methods used to solve (2.44) and (2.48) since we must cope with discontinuities in σ_g^*. However, workable schemes do exist. For example, recently Arakawa and Hsu (see Chapter V of Hsu, 1988), in the context of solving (2.5) in a primitive equation model, have proposed a finite difference scheme which has very small dissipation and computational dispersion and which guarantees positive definiteness.

2.2.4 Linear dynamics

For simplicity let us consider adiabatic, frictionless flow for the near Boussinesq case in which Γ is set equal to the constant Γ_0, where $\Gamma_0 = R/p_B$. Then, linearizing about a basic state of rest with $\sigma_0 = (p_B - p_T)/(\Theta_T - \Theta_B)$, the potential pseudodensity equation

(2.44) becomes

$$f^2(Y)\frac{\partial}{\partial T}\left(\frac{\sigma_g^* - \sigma_0}{\sigma_0}\right) = \beta\frac{\partial M^*}{\partial X}, \tag{2.49}$$

while the invertibility relation (2.48a) becomes

$$\frac{\partial^2 M^*}{\partial X^2} + f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial M^*}{\partial Y}\right) + \frac{f^2(Y)}{\Gamma_0\sigma_0}\frac{\partial^2 M^*}{\partial \Theta^2} + f^2(Y)\left(\frac{\sigma_g^* - \sigma_0}{\sigma_0}\right) = 0. \tag{2.50}$$

Eliminating σ_g^* between (2.49) and (2.50) we obtain

$$\frac{\partial}{\partial T}\left\{\frac{\partial^2 M^*}{\partial X^2} + f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial M^*}{\partial Y}\right) + \frac{f^2(Y)}{\Gamma_0\sigma_0}\frac{\partial^2 M^*}{\partial \Theta^2}\right\} + \beta\frac{\partial M^*}{\partial X} = 0. \tag{2.51}$$

Subtracting the basic state $\overline{M}$, where $M^* = \overline{M} + \mathcal{M}$ we obtain

$$\frac{\partial}{\partial T}\left\{\frac{\partial^2 \mathcal{M}}{\partial X^2} + f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial \mathcal{M}}{\partial Y}\right) + \frac{f^2(Y)}{\Gamma_0\sigma_0}\frac{\partial^2 \mathcal{M}}{\partial \Theta^2}\right\} + \beta\frac{\partial \mathcal{M}}{\partial X} = 0, \tag{2.52a}$$

with the vertical boundary conditions ((2.48b) and linearized (2.48c))

$$\frac{\partial \mathcal{M}}{\partial \Theta} = 0 \qquad \text{at} \qquad \Theta = \Theta_T, \tag{2.52b}$$

$$\Theta\frac{\partial \mathcal{M}}{\partial \Theta} - \mathcal{M} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{2.52c}$$

Let us assume that we are on a midlatitude β-plane in the northern hemisphere and that it is infinite in the positive direction (northward) and reaches Y_0 in the negative direction (southward). We impose the following conditions

$$\mathcal{M} = 0 \qquad \text{when} \qquad Y = Y_0, \tag{2.52d}$$

$$\mathcal{M} \to 0 \qquad \text{as} \qquad Y \to \infty, \tag{2.52e}$$

where $Y_0 > -f_0/\beta$, which simply means that the β-plane does not cross the equator. Note that $f^2(Y) = 0$, which corresponds to the equator, represents a singular point for (2.52). Thus, there can be no waves travelling across the equator. In appendix A, vertical, zonal and meridional transforms are defined to solve (2.52). An explanation for this particular choice of meridional limits can be found therein. For now, let us simply remark that the development for the vertical and zonal transforms is identical to the one for standard quasi-geostrophic β-plane theory (Lindzen, 1967). The meridional transform is different. In quasi-geostrophic β-plane theory the three $f^2(Y)$ factors in (2.52) are replaced by the constant f_0^2, which results in a trigonometric variation of $\mathcal{M}$ in the meridional direction. By contrast, the solutions of (2.52) are of the form

$$\mathcal{M} \sim e^{-\mathcal{Y}^2/2}\mathcal{Y}^{\alpha+3/2}L_n^{(\alpha)}(\mathcal{Y}^2)\cos\left[c_l^{-1}(\Gamma_0\sigma_0)^{1/2}(\theta_T-\theta)\right]e^{i(mX+\nu T)}, \tag{2.53a}$$

where $L_n^{(\alpha)}$ are the generalized Laguerre polynomials (Magnus et al., 1966, pages 239–249), $\alpha = \pm 3/2$, $\mathcal{Y} = (\beta c_l)^{-1/2}(f_0 + \beta Y)$, and l is the index of the vertical mode. The linearized lower boundary condition (2.52c) is satisfied if the constants c_l are the solutions of the transcendental equation $c_l^{-1}(\Gamma_0\sigma_0)^{1/2}\theta_B \tan\left[c_l^{-1}(\Gamma_0\sigma_0)^{1/2}(\theta_T - \theta_B)\right] = 1$. The Laguerre polynomials can be expressed in terms of Hermite polynomials which makes (2.53a) somewhat less compact. The dispersion relation associated with the solution is

$$\nu = \frac{\beta m}{m^2 + \frac{\beta}{c_l}(4n + 2\alpha + 2)}. \tag{2.53b}$$

Thus, the midlatitude quasi-geostrophic and semigeostrophic Rossby wave solutions differ essentially only in meridional structure.

Chapter 3

SEMIGEOSTROPHIC THEORY ON THE HEMISPHERE

Our final extension of the isentropic/vortex coordinate version of semigeostrophic theory is to spherical geometry. The basic theoretical structure is the same as for the f-plane study of Schubert et al. (1989) and for the β-plane study in the previous chapter. As before this approach results in a simple mathematical form for the prognostic equation (σ_g^* or the geostrophic potential pseudodensity equation), where the ageostrophic velocities are entirely implicit, and an invertibility relation for obtaining the potential field (M^*) from σ_g^*. Thus, the complete dynamics reduce to a form reminiscent of nondivergent barotropic dynamics; the prognostic equation having become an equation predicting the geostrophic potential pseudodensity with an invertibility principle where a Laplacian-like operator has to be inverted. This dynamical structure has been successfully used to study the evolution of potential vorticity and wind in tropical cyclones (Schubert and Alworth, 1989) and it will be used in chapter 6 to study a Hadley cell problem. This general approach is probably the simplest way to look at all types of balanced flows.

It should be emphasized that semigeostrophic theory on the sphere is essentially a hemispheric theory since it cannot handle crossequatorial flow. The extension of semigeostrophic theory from the β-plane makes the study of extensive baroclinic waves more realistic not to mention the benefits to stratospheric studies. Matsuno (1970 and 1971) used quasi-geostrophic theory on a hemisphere to study stratospheric sudden warmings. Semigeostrophic theory is more suitable to study that problem since it allows for variable static stability; however the mathematical problem of solution is quite similar. The study by Hoskins et al. (1977) of energy dispersion in a barotropic atmosphere showed clearly that for phenomena of a scale somewhat less than planetary, even when the local dynamics are well represented on a β-plane, latitudinal tilts and the propagation of energy are strongly influenced by a variation in β. Moreover, realistic zonal flows did not produce much response in the other hemisphere. Thus, it is speculated that the present semigeostrophic theory will prove to be quite useful.

As in the previous chapter the fundamental prognostic variable is the potential pseudodensity, which is closely related to the potential vorticity and gives an instantaneous view of the total balanced mass and windfields. The power of analyzing potential vorticity on isentropic surfaces has been well documented by Hoskins et al. (1985) who state that "it is found that time sequences of isentropic potential vorticity and surface potential temperature charts — which succinctly summarize the combined effects of vorticity advection, thermal advection, and vertical motion field — lead to a very clear and complete picture of the dynamics". However, the emphasis of their study was different from the present one. Their study was concerned with the diagnosis of data in terms of potential vorticity rather than with using it as a concept for modelling. For their purposes it was quite sufficient to present the invertibility principle in physical coordinates, since it is no harder to solve in those coordinates than in the geostrophic ones. They did not concern themselves with a prognostic equation and how it might best be presented. The real importance of geostrophic coordinates is revealed when we need to predict. Then it becomes imperative to get rid of the ageostrophic advection because otherwise we are forced to solve diagnostic equations (equivalent to the ω equation in quasi-geostrophic theory) for each component of the ageostrophic wind.

The first section of this chapter will present the primitive equations on the sphere and derive their conservation principles. The next section introduces the generalized geostrophic momentum approximation along with the conditions for geostrophic balance. The corresponding conservation principles are derived for the approximate momentum equations. The true power of the methodology is revealed with the coordinate transformation which makes the ageostrophic circulation completely implicit. Next, the invertibility principle is derived. Finally, the linear dynamics of the approximate system are considered.

3.1 The primitive equations

Using potential temperature as the vertical coordinate, the quasi-static primitive equations with the "traditional approximation" (Phillips, 1966) can be written

$$\cos\phi\frac{Du}{Dt} - 2\Omega\sin\phi\, v\cos\phi - \frac{uv\sin\phi}{a} + \frac{\partial M}{a\partial\lambda} = F\cos\phi, \tag{3.1}$$

$$\cos\phi\frac{Dv}{Dt} + 2\Omega\sin\phi\, u\cos\phi + \frac{u^2\sin\phi}{a} + \cos\phi\frac{\partial M}{a\partial\phi} = G\cos\phi, \tag{3.2}$$

$$\frac{\partial M}{\partial\theta} = \Pi, \tag{3.3}$$

$$\frac{D\sigma}{Dt} + \sigma\left(\frac{\partial u}{a\cos\phi\partial\lambda} + \frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial\dot\theta}{\partial\theta}\right) = 0, \tag{3.4}$$

where u and v are the zonal and meridional components of the velocity, F and G the components of the frictional force per unit mass, $\Pi = c_p\,(p/p_{00})^{\kappa}$ the Exner function, $M = \theta\Pi + gz$ the Montgomery potential, $\sigma = -\partial p/\partial\theta$ the pseudodensity, and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{a\cos\phi\partial\lambda} + v\frac{\partial}{a\partial\phi} + \dot\theta\frac{\partial}{\partial\theta} \tag{3.5}$$

the total derivative.

3.1.1 Conservation relations

The zonal momentum equation, (3.1), can be written in terms of absolute angular momentum as

$$\frac{D}{Dt}\left(u\cos\phi + a\Omega\cos^2\phi\right) + \frac{\partial M}{a\partial\lambda} = F\cos\phi. \tag{3.6}$$

In the absence of external torques the absolute angular momentum per unit mass, $u\cos\phi + a\Omega\cos^2\phi$, is conserved.

The kinetic energy equation can be derived by combining (3.1) and (3.2) to get

$$\frac{DK}{Dt} + u\frac{\partial M}{a\cos\phi\partial\lambda} + v\frac{\partial M}{a\partial\phi} = uF + vG, \tag{3.7}$$

where $K \equiv \frac{1}{2}(u^2+v^2)$ is the kinetic energy per unit mass. Combining (3.7) with the continuity equation (3.4) gives

$$\frac{\partial(\sigma K)}{\partial t} + \frac{\partial(\sigma u K)}{a\cos\phi\partial\lambda} + \frac{\partial(\sigma v\cos\phi K)}{a\cos\phi\partial\phi} + \frac{\partial(\sigma\dot\theta K)}{\partial\theta} + \sigma u\frac{\partial M}{a\cos\phi\partial\lambda} + \sigma v\frac{\partial M}{a\partial\phi} = \sigma\,(uF+vG). \tag{3.8}$$

Using the continuity and hydrostatic equations the kinetic energy equation can be written

$$\frac{\partial(\sigma K)}{\partial t}+\frac{\partial(\sigma u(K+gz))}{a\cos\phi\partial\lambda}+\frac{\partial(\sigma v\cos\phi(K+gz))}{a\cos\phi\partial\phi}+\frac{\partial\left(\sigma\dot{\theta}(K+gz)\right)}{\partial\theta}-\frac{\partial}{\partial\theta}\left(gz\frac{\partial p}{\partial t}\right)$$
$$=\sigma\left(uF+vG-\alpha\omega\right), \tag{3.9}$$

where $\omega = Dp/Dt$.

For deriving the thermodynamic energy equation, multiply (3.4) by c_pT to obtain

$$\frac{\partial}{\partial t}\left(\sigma c_pT\right)+\frac{\partial\left(\sigma u c_pT\right)}{a\cos\phi\partial\lambda}+\frac{\partial\left(\sigma v\cos\phi c_pT\right)}{a\cos\phi\partial\phi}+\frac{\partial}{\partial\theta}\left(\sigma\dot{\theta}c_pT\right)=\sigma\left(Q+\alpha\omega\right), \tag{3.10}$$

where $Q = \Pi\dot{\theta}$. Adding (3.9) and (3.10) we obtain the total energy equation

$$\frac{\partial}{\partial t}\Big(\sigma(K+c_pT)\Big)+\frac{\partial}{a\cos\phi\partial\lambda}\Big(\sigma u(K+M)\Big)+\frac{\partial}{a\cos\phi\partial\phi}\Big(\sigma v\cos\phi(K+M)\Big)$$
$$+\frac{\partial}{\partial\theta}\Big(\sigma\dot{\theta}(K+M)-\phi\frac{\partial p}{\partial t}\Big)=\sigma(uF+vG+Q). \tag{3.11a}$$

The lower boundary will be regarded in the same way as it was in the previous chapter. This way of viewing isentropes crossing the earth's surface has proved useful for the definition of available potential energy (Lorenz, 1955), the analysis of baroclinic instability (Bretherton, 1966; Hoskins et al., 1985; Hsu, 1988), and the finite amplitude Eliassen-Palm theorem (Andrews, 1983). An isentrope crossing the surface of the earth is assumed to continue just under the earth's surface with a pressure equal to the surface pressure. At any horizontal position where two distinct isentropic surfaces run just under the earth's surface (and hence have the same pressure), there is no mass trapped between them, so that $\sigma = 0$ there. The bottom isentropic surface θ_B is the largest value of θ which remains everywhere below the earth's surface. We can integrate (3.11a) over the entire atmosphere to obtain

$$\frac{\partial}{\partial t}\iiint(K+c_pT)\sigma a^2\cos\phi\,d\lambda d\phi d\theta=\iiint(uF+vG+Q)\sigma a^2\cos\phi\,d\lambda d\phi d\theta, \tag{3.11b}$$

where we have assumed that the top boundary θ_T is both an isentropic and isobaric surface, there is no topography with a vanishing $\dot{\theta}$ at the top and bottom.

The equation for the absolute isentropic vorticity ζ can be derived from (3.1) and (3.2). It's form is

$$\frac{D\zeta}{Dt} + \zeta\left(\frac{\partial u}{a\cos\phi\partial\lambda} + \frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi}\right) = \left(\xi\frac{\partial}{a\cos\phi\partial\lambda} + \eta\frac{\partial}{a\partial\phi}\right)\dot{\theta} + \frac{\partial G}{a\cos\phi\partial\lambda} - \frac{\partial(F\cos\phi)}{a\cos\phi\partial\phi}, \tag{3.12}$$

where

$$(\xi,\eta,\zeta) = \left(-\frac{\partial v}{\partial\theta}, \frac{\partial u}{\partial\theta}, 2\Omega\sin\phi + \frac{\partial v}{a\cos\phi\partial\lambda} - \frac{\partial(u\cos\phi)}{a\cos\phi\partial\phi}\right). \tag{3.13}$$

Equation (3.12) can also be written in the flux form

$$\frac{\partial(\sigma P)}{\partial t} + \frac{\partial\left(u\sigma P - \xi\dot{\theta} - G\right)}{a\cos\phi\partial\lambda} + \frac{\partial\left((v\sigma P - \eta\dot{\theta} + F)\cos\phi\right)}{a\cos\phi\partial\phi} = 0, \tag{3.14}$$

where $P = \zeta/\sigma$ is the potential vorticity. The equivalence of (3.12) and (3.14) follows easily from the fact that the curl of the vector vorticity vanishes, or

$$\frac{\partial\xi}{a\cos\phi\partial\lambda} + \frac{\partial(\eta\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial\zeta}{\partial\theta} = 0.$$

The relation expressed in (3.14) is the spherical equivalent of the β-plane result in (2.14), the Haynes-McIntyre theorem. Thus, even in the face of diabatic effects the total potential vorticity between two isentropic surfaces does not change, it is simply redistributed so that a deficiency in one area is met with a surplus in another one bounded by the same isentropic surfaces. This idea seems to be closely related to the idea of a reference state in Hoskins et al. (1985), the existence of which was one of the three conditions for inversion of the potential vorticity field to obtain the complete flow field. It may be more correct to think of the reference state as expressing the mass distribution of potential vorticity rather than the mass distribution of potential temperature which obviously changes when diabatic effects are included. A case in point, Schubert and Alworth (1987) solved the invertibility principle and proved the uniqueness of their solution when heating played the major role in forcing the flow. In the next section, we shall show that the Haynes-McIntyre theorem is valid for the semigeostrophic equations on the sphere.

The potential vorticity equation is derived by eliminating the horizontal divergence between (3.4) and (3.12) to obtain

$$\frac{DP}{Dt} = \frac{1}{\sigma}\left[\xi\frac{\partial\dot{\theta}}{a\cos\phi\partial\lambda} + \eta\frac{\partial\dot{\theta}}{a\partial\phi} + \zeta\frac{\partial\dot{\theta}}{\partial\theta} + \frac{\partial G}{a\cos\phi\partial\lambda} - \frac{\partial(F\cos\phi)}{a\cos\phi\partial\phi}\right]. \tag{3.15}$$

In the absence of friction and heating the potential vorticity P is conserved. Defining the potential pseudodensity as $\sigma^* = 2\Omega \sin\phi\, \sigma/\zeta$ we can just as easily derive an equation for σ^* since $\sigma^* = 2\Omega \sin\phi/P$. The potential pseudodensity equation takes the form

$$\frac{D\sigma^*}{Dt} + \frac{\sigma^*}{\zeta}\left[\left(\xi\frac{\partial}{a\cos\phi\partial\lambda} + \eta\frac{\partial}{a\partial\phi} + \zeta\frac{\partial}{\partial\theta}\right)\dot{\theta} + \frac{\partial G}{a\cos\phi\partial\lambda} - \frac{\partial(F\cos\phi)}{a\cos\phi\partial\phi}\right]$$
$$= \frac{\sigma^*}{2\Omega\sin\phi}\frac{D}{Dt}2\Omega\sin\phi. \tag{3.16}$$

The physical meaning of σ^* on the f-plane and on the β-plane was discussed in Schubert et al. (1989) and in the previous chapter, respectively. Here it must suffice to say that in the semigeostrophic system on the sphere with the appropriate coordinate transformations the geostrophic potential pseudodensity equation takes on a very simple form. Indeed, (3.45) is so simple that σ_g^* seems to be a more natural variable than the more commonly used potential vorticity.

3.2 The semigeostrophic equations on the sphere

As approximations to the primitive equations (3.1) and (3.2) let us consider

$$\cos\Phi\frac{Du_g}{Dt} - 2\Omega\Big[\sin\Phi\, v\cos\phi + (\sin\phi - \sin\Phi)\, v_g\cos\Phi\Big]$$
$$-\frac{u_g v_g\sin\Phi}{a} + \frac{\partial M}{a\partial\lambda} = F\cos\Phi, \tag{3.17}$$
$$\cos\phi\frac{Dv_g}{Dt} + 2\Omega\Big[\sin\Phi\, u\cos\Phi + (\sin\phi - \sin\Phi)\, u_g\cos\phi\Big]$$
$$+\frac{u_g^2\sin\Phi\cos\phi}{a\cos\Phi} + \cos\Phi\frac{\partial M}{a\partial\phi} = G\cos\phi, \tag{3.18}$$

where (u_g, v_g) is the geostrophic wind on the sphere, given by

$$u_g\frac{\cos\phi}{\cos\Phi}2\Omega\sin\Phi + \frac{\partial M}{a\partial\phi} = 0, \tag{3.19}$$

$$-2\Omega\sin\Phi\cos\Phi\, v_g + \frac{\partial M}{a\partial\lambda} = 0, \tag{3.20}$$

and where Φ is defined as

$$a(\sin\phi - \sin\Phi)\,2\Omega\sin\Phi = u_g\cos\Phi. \tag{3.21}$$

Comparing the approximate momentum equations (3.17) and (3.18) to the momentum equations of the primitive equations set (3.1) and (3.2) reveals several interesting differences. Approximating momentum by the geostrophic momentum involves making selective changes to the latitude as well as changing (u, v) to (u_g, v_g) everywhere except in the "Coriolis term". The latter now splits into two terms which can be interpreted as coming from the linear expansion of $\sin\phi$ around $\sin\Phi$. The transformed latitude also enters our definition of the geostrophic wind (u_g, v_g) in (3.19) and (3.20). The geostrophic momentum approximation generalized to the sphere maintains important conservation principles of the primitive equations. In addition, an accompanying coordinate transformation will lead to a simple prognostic equation.

3.2.1 Conservation relations

Writing (3.17) in terms of approximate angular momentum gives

$$\frac{D}{Dt}\left(a\Omega\cos^2\Phi\right) + \frac{\partial M}{a\partial\lambda} = F\cos\Phi. \tag{3.22}$$

This is equivalent to

$$\frac{D}{Dt}\left(u_g\cos\Phi\frac{(\sin\phi + \sin\Phi)}{2\sin\Phi} + a\Omega\cos^2\phi\right) + \frac{\partial M}{a\partial\lambda} = F\cos\Phi, \tag{3.23}$$

which makes apparent the approximation to the angular momentum in (3.6) and shows that Φ is closely related to an approximate angular momentum coordinate.

Derivation of the kinetic energy equation proceeds along exactly the same lines as for the primitive equations. Combining (3.17) and (3.18) and now defining

$$K \equiv \frac{1}{2}\left(u_g^2 + v_g^2\right) \tag{3.24}$$

we obtain (3.7). Thus, (3.11a) and (3.11b) are valid with the generalized geostrophic momentum approximation provided K is defined as in (3.24).

The equation for the absolute isentropic geostrophic vorticity ζ_g can be derived from (3.17) and (3.18). We can write

$$\frac{D\zeta_g}{Dt} + \zeta_g\left(\frac{\partial u}{a\cos\phi\partial\lambda} + \frac{\partial\,(v\cos\phi)}{a\cos\phi\partial\phi}\right) - \left(\xi_g\frac{\partial}{a\cos\phi\partial\lambda} + \eta_g\frac{\partial}{a\partial\phi}\right)\dot{\theta}$$
$$= \frac{\partial}{a\cos\phi\partial\lambda}\left[\frac{\partial\Lambda}{\partial\phi}F\cos\Phi + \frac{\partial\Phi}{\partial\phi}G\right] - \frac{\partial}{a\cos\phi\partial\phi}\left[\frac{\partial\Lambda}{\partial\lambda}F\cos\Phi + \frac{\partial\Phi}{\partial\lambda}G\right], \tag{3.25}$$

where the geostrophic vorticity vector is

$$(\,\xi_g/(a\cos\phi),\ (\eta_g\cos\phi)/a,\ \zeta_g\,) = 2\Omega\sin\Phi\left(\frac{\partial(\Lambda,\sin\Phi)}{\partial(\sin\phi,\theta)},\frac{\partial(\Lambda,\sin\Phi)}{\partial(\theta,\lambda)},\frac{\partial(\Lambda,\sin\Phi)}{\partial(\lambda,\sin\phi)}\right), \tag{3.26}$$

with Λ defined below in (3.29). Equation (3.25) can be written in the potential vorticity form

$$\frac{\partial(\sigma P_g)}{\partial t} + \frac{\partial}{a\cos\phi\partial\lambda}\left(u\sigma P_g - \xi_g\dot{\theta} - \frac{\partial\Lambda}{\partial\phi}F\cos\Phi - \frac{\partial\Phi}{\partial\phi}G\right)$$
$$+\frac{\partial}{a\cos\phi\partial\phi}\left((v\sigma P_g - \eta_g\dot{\theta})\cos\phi + \frac{\partial\Lambda}{\partial\lambda}F\cos\Phi + \frac{\partial\Phi}{\partial\lambda}G\right) = 0, \tag{3.27}$$

where $P_g = \zeta_g/\sigma$ is the geostrophic Rossby-Ertel potential vorticity. Again, the equivalence of (3.25) and (3.27) follow easily from the fact that

$$\frac{\partial\xi_g}{a\cos\phi\partial\lambda} + \frac{\partial(\eta_g\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial\zeta_g}{\partial\theta} = 0.$$

Equation (3.27) is the semigeostrophic equivalence of equation (3.14) and we conclude that the primitive equation result of Haynes and McIntyre is maintained when we make the geostrophic momentum approximation generalized to the sphere.

Eliminating the horizontal divergence between (3.4) and (3.25), we obtain a form of the geostrophic potential pseudodensity equation

$$\frac{D(\sigma/\zeta_g)}{Dt} + \frac{(\sigma/\zeta_g)}{\zeta_g}\left(\xi_g\frac{\partial}{a\cos\phi\partial\lambda} + \eta_g\frac{\partial}{a\partial\phi} + \zeta_g\frac{\partial}{\partial\theta}\right)\dot{\theta}$$
$$= \frac{(\sigma/\zeta_g)}{2\Omega\sin\Phi}\left[\frac{\partial(F\cos\Phi)}{a\partial\sin\Phi} - \frac{\partial G}{a\cos\Phi\partial\Lambda}\right]. \tag{3.28}$$

When we make the coordinate transformation this equation simplifies greatly.

3.2.2 Coordinate transformation

The duality between the use of geostrophic coordinates in the horizontal and the isentropic coordinate in the vertical is well known. Hoskins and Draghici (1977), Gill (1981), Heckley and Hoskins (1982) all discussed it. The purpose of the coordinate transformation here is the same as before: to make the ageostrophic circulation completely implicit. Combining vortex coordinates, which are just two of the Clebsch potentials of the wind field, with the isentropic coordinate produces the desired result. Again this will lead to an elegant version of the geostrophic potential pseudodensity equation. Thus, let us introduce the vortex coordinates

$$(\Lambda\, , \sin\Phi\, , \Theta\, , T) = \left(\lambda + \frac{v_g}{a2\Omega\sin\Phi\cos\Phi}\, , \sin\phi - \frac{u_g\cos\Phi}{a2\Omega\sin\Phi}\, , \theta\, , t\right). \tag{3.29}$$

Derivatives in $(\lambda, \phi, \theta, t)$ space are then related to derivatives in $(\Lambda, \Phi, \Theta, T)$ space by

$$\frac{\partial}{\partial t} = \frac{\partial\Lambda}{\partial t}\frac{\partial}{\partial\Lambda} + \frac{\partial\Phi}{\partial t}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial T}, \tag{3.30}$$

$$\frac{\partial}{\partial\lambda} = \frac{\partial\Lambda}{\partial\lambda}\frac{\partial}{\partial\Lambda} + \frac{\partial\Phi}{\partial\lambda}\frac{\partial}{\partial\Phi}, \tag{3.31}$$

$$\frac{\partial}{\partial\phi} = \frac{\partial\Lambda}{\partial\phi}\frac{\partial}{\partial\Lambda} + \frac{\partial\Phi}{\partial\phi}\frac{\partial}{\partial\Phi}, \tag{3.32}$$

$$\frac{\partial}{\partial\theta} = \frac{\partial\Lambda}{\partial\theta}\frac{\partial}{\partial\Lambda} + \frac{\partial\Phi}{\partial\theta}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial\Theta}. \tag{3.33}$$

Inverting (3.31) and (3.32) to obtain

$$\frac{\partial(\Lambda,\Phi)}{\partial(\lambda,\phi)}\frac{\partial}{\partial\Lambda} = \frac{\partial\Phi}{\partial\phi}\frac{\partial}{\partial\lambda} - \frac{\partial\Phi}{\partial\lambda}\frac{\partial}{\partial\phi} \tag{3.34}$$

$$\frac{\partial(\Lambda,\Phi)}{\partial(\lambda,\phi)}\frac{\partial}{\partial\Phi} = -\frac{\partial\Lambda}{\partial\phi}\frac{\partial}{\partial\lambda} + \frac{\partial\Lambda}{\partial\lambda}\frac{\partial}{\partial\phi}, \tag{3.35}$$

and applying (3.30), (3.33), (3.34) and (3.35) to the Bernoulli function

$$M^* = M + \frac{1}{2}\left(u_g^2 + v_g^2\right), \tag{3.36}$$

it can be shown that

$$\left(\frac{\partial M}{\partial \lambda}, \frac{\partial M}{\partial \theta}, \frac{\partial M}{\partial t}\right) = \left(\frac{\partial M^*}{\partial \Lambda}, \frac{\partial M^*}{\partial \Theta}, \frac{\partial M^*}{\partial T}\right),$$
$$\frac{\partial M}{\partial \phi} = \frac{\cos\phi}{\cos\Phi}\left[\frac{\partial M^*}{\partial \Phi} - \frac{u_g^2 + v_g^2\left(\cos^2\Phi - \sin^2\Phi\right)}{\sin\Phi\cos\Phi}\right]. \tag{3.37}$$

We can also use (3.34) and (3.35) in (3.36), along with (3.26), to obtain

$$\zeta_g \frac{\partial}{\partial \Theta} = \xi_g \frac{\partial}{a\cos\phi\partial\lambda} + \eta_g \frac{\partial}{a\partial\phi} + \zeta_g \frac{\partial}{\partial\theta}, \tag{3.38}$$

which shows that (Λ, Φ) are vortex coordinates. Note that (3.38) leads to a considerable simplification. Combining the vorticity and continuity equations one obtains the geostrophic potential vorticity equation,

$$\frac{DP_g}{Dt} - P_g\frac{\partial\dot{\theta}}{\partial\Theta} = \frac{P_g}{2\Omega\sin\Phi}\left(\frac{\partial G}{a\cos\Phi\partial\Lambda} - \frac{\partial(F\cos\Phi)}{a\partial\sin\Phi}\right), \tag{3.39}$$

where $P_g = \zeta_g/\sigma$.

The transformation relations (3.30)-(3.33) also imply that the operator (3.5) can be written as

$$\frac{D}{Dt} = \frac{\partial}{\partial T} + \frac{D\Lambda}{Dt}\frac{\partial}{\partial\Lambda} + \frac{D\Phi}{Dt}\frac{\partial}{\partial\Phi} + \dot{\Theta}\frac{\partial}{\partial\Theta}. \tag{3.40}$$

Writing (3.28) in terms of σ_g^* where we define the geostrophic potential pseudodensity as

$$\sigma_g^* = \frac{2\Omega\sin\Phi}{\zeta_g}\sigma, \tag{3.41}$$

and using (3.39) we get

$$\frac{\partial\sigma_g^*}{\partial T} + \frac{D\Lambda}{Dt}\frac{\partial\sigma_g^*}{\partial\Lambda} + \frac{D\sin\Phi}{Dt}\sin\Phi\frac{\partial}{\partial\sin\Phi}\left(\frac{\sigma_g^*}{\sin\Phi}\right) + \frac{\partial}{\partial\Theta}\left(\dot{\theta}\sigma_g^*\right)$$
$$+\frac{\sigma_g^*}{2\Omega\sin\Phi}\left[\frac{\partial G}{a\cos\Phi\partial\Lambda} - \frac{\partial(F\cos\Phi)}{a\partial\sin\Phi}\right] = 0. \tag{3.42}$$

Using (3.17), (3.18) , (3.19), (3.20), (3.29) and (3.37) it can easily be shown that

$$2\Omega a\sin\Phi\frac{D\sin\Phi}{Dt} = \frac{\partial M^*}{a\partial\Lambda} - F\cos\Phi, \tag{3.43}$$

$$2\Omega a\sin\Phi\frac{D\Lambda}{Dt} = -\frac{\partial M^*}{a\partial\sin\Phi} + \frac{G}{\cos\Phi}. \tag{3.44}$$

A major advantage of the transformation from $(\lambda, \phi, \theta, t)$ space to $(\Lambda, \Phi, \Theta, T)$ space is the change from advection by (u, v) to advection by $(D\Lambda/Dt, D\Phi/Dt)$ given in (3.43) and (3.44). The total time derivative does not contain any ageostrophic advection, thus (3.42), the potential pseudodensity equation, does not contain any ageostrophic advection. It is all implicit in the coordinate transformation.

Using the canonical momentum equations (3.43) and (3.44) we can now write the geostrophic potential pseudodensity equation (3.42) in flux form as

$$\frac{\partial \sigma_g^*}{\partial T} + \frac{\partial}{\partial \Lambda}\left(\sigma_g^* \frac{D\Lambda}{Dt}\right) + \frac{\partial}{\partial \sin\Phi}\left(\sigma_g^* \frac{D\sin\Phi}{Dt}\right) + \frac{\partial}{\partial \Theta}\left(\sigma_g^* \dot{\theta}\right) = 0. \tag{3.45}$$

The horizontal flux terms can be written in the Jacobian form

$$\frac{\partial \sigma_g^*}{\partial T} + \frac{1}{2\Omega a^2}\frac{\partial\left(M^*, \left(\sigma_g^*/\sin\Phi\right)\right)}{\partial\left(\Lambda, \sin\Phi\right)} + \frac{1}{2\Omega a}\frac{\partial}{\partial \Lambda}\left(\frac{\sigma_g^*}{\sin\Phi}\left(G/\cos\Phi\right)\right)$$

$$+\frac{1}{2\Omega a}\frac{\partial}{\partial \sin\Phi}\left(-\frac{\sigma_g^*}{\sin\Phi}\left(F\cos\Phi\right)\right) + \frac{\partial}{\partial\Theta}\left(\sigma_g^*\dot{\theta}\right) = 0, \tag{3.46}$$

which serves as the fundamental predictive equation of the model.

3.2.3 Invertibility principle

The geostrophic potential pseudodensity σ_g^*, the fundamental variable of the model, is a combination of the mass field σ and the geostrophic wind field ζ_g. However, all the balanced fields can be obtained from σ_g^* since it depends only on M^* as can be seen from the fact that σ is related to M^* through hydrostatic balance and ζ_g is related to M^* through geostrophic balance. What is needed is to invert σ_g^* to obtain M^* from which the balanced mass and wind fields follow. The relation between M^* and σ_g^* is derived as follows. From the definition of σ_g^* we have

$$\frac{2\Omega\sin\Phi}{\zeta_g}\frac{\partial\Pi}{\partial\theta} + \Gamma\sigma_g^* = 0, \tag{3.47}$$

where $\Gamma = d\Pi/dp = \kappa\Pi/p$. Using (3.26) this last equation can be written

$$\frac{\partial(\lambda, \sin\phi, \Pi)}{\partial(\Lambda, \sin\Phi, \Theta)} + \Gamma\sigma_g^* = 0. \tag{3.48}$$

Using the geostrophic ((3.19) and (3.20)), hydrostatic (3.3), and coordinate transformation relations ((3.29) and (3.37)), we can express λ, $\sin\phi$ and Π in terms of M^* as

$$\lambda = \Lambda - a^{-2}\cos^{-2}\Phi\, f^{-2} M^*_\Lambda, \tag{3.49}$$

$$\sin\phi = \sin\Phi - a^{-2}\cos^2\Phi\, \hat{f}^{-2}\left[M^*_{\sin\Phi} - a^{-2}\cos^{-2}\Phi\,(f\cos\Phi)_{\sin\Phi} f^{-3}\left(M^*_\Lambda\right)^2\right], \tag{3.50}$$

$$\Pi = M^*_\Theta. \tag{3.51}$$

The two functions f and $\hat{f}^2$ are introduced so as to allow us to write the invertibility relation in a more compact form. They are defined as follows

$$f = 2\Omega\sin\Phi, \tag{3.52}$$

$$\hat{f}^2 = (2\Omega\sin\Phi)^2 + 2\Omega\sin\Phi\,(2\Omega\sin\Phi - 2\Omega\sin\phi)\cos^{-2}\Phi. \tag{3.53}$$

Thus, f is the Coriolis parameter at the latitude Φ, whereas $\hat{f}^2$ is a certain combination of the Coriolis parameter at two latitudes, one of which depends on the other coordinates. Substituting (3.49)–(3.51) into (3.48), we obtain

$$\begin{vmatrix} \cos^{-2}\Phi\, M^*_{\Lambda\Lambda} - f^2a^2 & \hat{f}^2\left(\cos^2\Phi\,\hat{f}^{-2}(M^*_{\sin\Phi} - (a\cos\Phi)^{-2}(f\cos\Phi)_{\sin\Phi} f^{-3} M^{*2}_\Lambda)\right)_\Lambda & M^*_{\Theta\Lambda} \\ f^2\left(\cos^{-2}\Phi\, f^{-2}M^*_\Lambda\right)_{\sin\Phi} & \hat{f}^2\left(\cos^2\Phi\,\hat{f}^{-2}(M^*_{\sin\Phi} - (a\cos\Phi)^{-2}(f\cos\Phi)_{\sin\Phi} f^{-3} M^{*2}_\Lambda)\right)_{\sin\Phi} - \hat{f}^2a^2 & M^*_{\Theta\sin\Phi} \\ \cos^{-2}\Phi\, M^*_{\Lambda\Theta} & \hat{f}^2\left(\cos^2\Phi\,\hat{f}^{-2}(M^*_{\sin\Phi} - (a\cos\Phi)^{-2}(f\cos\Phi)_{\sin\Phi} f^{-3} M^{*2}_\Lambda)\right)_\Theta & M^*_{\Theta\Theta} \end{vmatrix}$$

$$+a^4 f^2 \hat{f}^2 \Gamma \sigma^*_g = 0, \tag{3.54a}$$

which expresses the invertibility principle in terms of the determinant of a Hessian-type matrix. The upper boundary is an isentropic and isobaric surface with potential temperature Θ_T and pressure p_T. Thus, the upper boundary condition for (3.54a) is

$$M^*_\Theta = \Pi(p_T) \quad \text{at} \quad \Theta = \Theta_T. \tag{3.54b}$$

Neglecting the effects of topography and assuming that the lower boundary is the constant height surface $z = 0$ and the isentropic surface $\Theta = \Theta_B$, we get that $M = \Theta\Pi$ at $\Theta = \Theta_B$. Written in terms of M^*, this lower boundary condition becomes

$$\Theta M^*_\Theta - M^* + \frac{1}{2}a^{-2}f^{-2}\Big[\cos^{-2}\Phi\, M^{*2}_\Lambda$$

$$+\cos^2\Phi\Big(M^*_{\sin\Phi} + 2\sin^{-1}\Phi\,\cos^{-2}\Phi\,\left(\Theta M^*_\Theta - M^* + a^{-2}f^{-2}\tan^2\Phi\, M^{*2}_\Lambda\right)\Big)^2\Big] = 0$$

$$\text{at} \quad \Theta = \Theta_B. \qquad (3.54c)$$

For lateral boundary conditions we note that symmetry at the poles requires

$$\frac{\partial M^*}{\partial \Phi} = 0 \quad \text{at} \quad \Phi = \pm\frac{\pi}{2}, \qquad (3.54d)$$

and we simply have cyclic boundary conditions in Λ. For a given σ^*_g, we can regard (3.54) as a nonlinear second order problem in M^*. Note that Γ and $\hat{f}$ both depend on M^*.

Equations (3.45) and (3.54) form a closed system for the prediction of σ^*_g and the diagnosis of M^*. The problem of isentropes intersecting the earth's surface has been addressed by adopting the massless region approach outlined in section 3.1.1 and in the previous chapter. Since $\sigma = 0$ in the massless region, $\sigma^*_g = 0$ there also. This discontinuity in σ^*_g on an isentropic surface puts strict requirements both on the numerical procedure used to solve the invertibility relation (3.54) and also, and more importantly, on the procedure for predicting σ^*_g using (3.45) since one might expect a ripple effect from the discontinuity. However, a problem similar to this one has been solved by Arakawa and Hsu. For solving (3.4) in a primitive equation model they proposed a finite difference scheme which has very small dissipation and computational dispersion and which guarantees positive definiteness (see Chapter V of Hsu, 1988). This picture of the lower boundary which is consistent with that of Bretherton (1966) seems to be more useful than that of Eliassen and Raustein (1968, 1970) who had to make predictions at imaginary underground gridpoints.

3.2.4 Linear dynamics

Consider adiabatic, frictionless flow for the near Boussinesq case in which Γ is set equal to the constant Γ_0, where $\Gamma_0 = R/p_B$. Then, linearizing about a basic state of rest with $\sigma_0 = (p_B - p_T)/(\Theta_T - \Theta_B)$, the potential pseudodensity equation (3.45) becomes

$$2\Omega a^2 \sin^2\Phi\, \frac{\partial}{\partial T}\left(\frac{\sigma^*_g - \sigma_0}{\sigma_0}\right) = \frac{\partial M^*}{\partial \Lambda}, \qquad (3.55)$$

while the invertibility relation (3.54a) becomes

$$\frac{\partial^2 M^*}{\partial \Lambda^2} + (2a\Omega \sin\Phi \cos\Phi)^2 \frac{\partial}{\partial \sin\Phi}\left\{\left(\frac{\cos\Phi}{2a\Omega\sin\Phi}\right)^2 \frac{\partial M^*}{\partial \sin\Phi}\right\}$$

$$+\frac{(2a\Omega\sin\Phi\cos\Phi)^2}{\Gamma_0\sigma_0}\frac{\partial^2 M^*}{\partial\Theta^2} + (2a\Omega\sin\Phi\cos\Phi)^2\left(\frac{\sigma_g^* - \sigma_0}{\sigma_0}\right) = 0. \qquad (3.56)$$

Eliminating σ_g^* between (3.55) and (3.56) we obtain

$$\frac{\partial}{\partial T}\left\{\frac{\partial^2 M^*}{a^2\cos^2\Phi\,\partial\Lambda^2} + \sin^2\Phi\frac{\partial}{a\partial\sin\Phi}\left\{\frac{\cos^2\Phi}{\sin^2\Phi}\frac{\partial M^*}{a\partial\sin\Phi}\right\} + \frac{(2\Omega\sin\Phi)^2}{\Gamma_0\sigma_0}\frac{\partial^2 M^*}{\partial\Theta^2}\right\}$$

$$+\frac{2\Omega}{a}\frac{\partial M^*}{a\partial\Lambda} = 0. \qquad (3.57)$$

Let us compare the horizontal part of (3.57) to the linearized barotropic vorticity equation (Longuet-Higgins, 1964)

$$\frac{\partial}{\partial t}\left\{\frac{\partial^2\psi}{\cos^2\phi\,\partial\lambda^2} + \frac{\partial}{\partial\sin\phi}\left(\cos^2\phi\frac{\partial\psi}{\partial\sin\phi}\right)\right\} + 2\Omega\frac{\partial\psi}{\partial\lambda} = 0, \qquad (3.58)$$

where ψ is the streamfunction. The primary difference between the two is in the meridional part; the wave solutions to (3.58) can propagate freely, but (3.57) has a singular point on the equator so that now there can be no wave propagation across the equator. The meridional structure equation that results from (3.57) is somewhat related to the differential equation for associated Legendre functions. Attempts to solve it analytically have been unsuccessful. Matsuno (1970, 1971) obtained the same meridional structure equation in his hemispheric studies of the stratosphere.

Chapter 4

DERIVATION OF THE EQUATIONS OF MOTION FROM HAMILTON'S PRINCIPLE

Hamiltonian mechanics have been widely and successfully used in classical and quantum mechanics for decades (Goldstein, 1980; Lanczos, 1970; Merzenbacher, 1970). Only relatively recently have these powerful principles found application in fluid mechanics (Salmon, 1988, and references therein) with quite promising results. The power of the Hamiltonian method arises from several of its properties. First, Hamilton's principle is a very succinct statement of dynamics. One need only make approximations to one functional, the Lagrangian, from which the approximate dynamical equations can be derived by Hamilton's principle. Secondly, conservation principles correspond to symmetries in the Hamiltonian by Noether's theorem. This has two important advantages. Conservative quantities can readily be identified by examining the symmetries in the Hamiltonian, and when making approximations to the Hamiltonian, not disturbing symmetries of its original form will guarantee the existence of conservation principles, albeit of an approximate form. Thirdly, Hamiltonian mechanics are coordinate independent and in fact the Hamiltonian method suggests transformations to coordinates where the mathematics of the problem are the simplest.

There are two primary schools of thought on how to present Hamilton's principle for a perfect fluid. The first corresponds to variational methods in particle physics, i.e. the positions and momenta of marked fluid particles are varied at fixed times (Eckart, 1960; Salmon, 1983, 1985). This is the approach used here. The second (Lin, 1963; Seliger and Whitham, 1968) involves variations at fixed locations and times of entropy and four scalar potentials (the Clebsch potentials (Lamb, 1932, page 248)), two of which are exactly the vortex coordinates. The fluid velocity can be written in terms of the four potentials and entropy (see (6.50)). This approach is more Eulerian-like and physical interpretation of the Lagrangian structure is not as straightforward. In fact the simplicity of the dynamical description of perfect fluids in terms of Eulerian variables seems to have delayed the acceptance of the Hamiltonian view in fluid mechanics, which has its natural representation in Lagrangian coordinates. Van Saarloos (1981) proved that the two representations are related by a canonical transformation, the particle representation being more general than it need be.

To briefly review the basic concepts of Hamiltonian mechanics, let's consider a conservative system of N discrete particles. The Lagrangian of this system is

$$L(\mathbf{q},\dot{\mathbf{q}}) = \sum_i \frac{1}{2} m_i \dot{\mathbf{q}}_i \cdot \dot{\mathbf{q}}_i - V(\mathbf{q}_1, \cdots, \cdots, \mathbf{q}_N), \tag{4.1}$$

where m_i is the mass of particle i and $\mathbf{q}_i(\tau)$ is its location at time τ. V is the potential energy of the system. The action integral is the time integral of the Lagrangian,

$$\int L d\tau. \tag{4.2}$$

The dynamical equations can be derived from Hamilton's principle which states that

$$\delta \int_{\tau_1}^{\tau_2} L d\tau = 0, \tag{4.3}$$

where δ corresponds to independent variations of $\mathbf{q}_i(\tau)$ and $\delta\mathbf{q}_i(\tau_1) = \delta\mathbf{q}_i(\tau_2) = 0$. Defining the conjugate momenta

$$\mathbf{p}_i = \partial L / \partial \dot{\mathbf{q}}_i, \tag{4.4}$$

the extended form of Hamilton's principle corresponds to independent variations of $\mathbf{q}_i, \mathbf{p}_i$ which are again zero at τ_1, τ_2 and

$$L(\mathbf{p},\mathbf{q}) = \sum_i \mathbf{p}_i \cdot \dot{\mathbf{q}}_i - H(\mathbf{p},\mathbf{q}). \tag{4.5}$$

$H(\mathbf{p},\mathbf{q})$ is the Hamiltonian which in most cases of physical interest is to be identified with the total energy. From (4.3) we get the two canonical equations,

$$\dot{\mathbf{q}}_i = \partial H / \partial \mathbf{p}_i, \qquad \dot{\mathbf{p}}_i = -\partial H / \partial \mathbf{q}_i. \tag{4.6}$$

Salmon (1985) demonstrated how Hamilton's principle could be used to derive approximate equations of motion for a shallow water system. He considered the adiabatic, frictionless case and approximated the velocity u, v (or equivalently the momenta per unit mass) by the geostrophic velocity u_g, v_g which is directly related to the mass field. Preserving symmetries in the Hamiltonian secured the conservation principles. Transformation to canonical coordinates made the mathematics of the problem simpler. He showed that for a constant Coriolis parameter the approximate equations and the new coordinates were equivalent to semigeostrophic theory. Here, there were semigeostrophic equations for the shallow water system with a variable Coriolis parameter, i.e. $f = f(x, y)$. In this chapter Salmon's work will be extended to the three dimensional case on the sphere. Still, only conservative forces are allowed. Diabatic forcing and friction can be included after Hamilton's principle has been applied.

In the first section of this chapter the primitive equations are derived from Hamilton's principle and conservation relations are derived from symmetries in the Hamiltonian. In the second section approximations are made to the Lagrangian and a whole spectrum of approximate momentum equations result, each characterized by a certain balance condition of the approximate wind and a certain coordinate transformation in the horizontal. Symmetries in the Hamiltonian are preserved and the conservation principles are thereby protected. The whole dynamics reduce to a prognostic equation for the generic potential pseudodensity and a generic invertibility principle. Finally a particular balanced model is specified. The last section of the chapter derives the equations of motion corresponding to the long wave approximation from a Lagrangian which neglects the meridional wind. The corresponding conservation relations are examined.

4.1 Derivation of the primitive equations from Hamilton's principle

Using potential temperature as the vertical coordinate, the quasi-static, frictionless, primitive equations with the "traditional approximation" (Phillips, 1966) can be written

$$\cos\phi \frac{Du}{Dt} - 2\Omega \sin\phi \, v \cos\phi - \frac{uv \sin\phi}{a} + \frac{\partial M}{a \partial \lambda} = 0, \tag{4.7}$$

$$\cos\phi\frac{Dv}{Dt} + 2\Omega\sin\phi\, u\cos\phi + \frac{u^2\sin\phi}{a} + \cos\phi\frac{\partial M}{a\partial\phi} = 0, \tag{4.8}$$

$$\frac{\partial M}{\partial\theta} = \Pi, \tag{4.9}$$

$$\frac{D\sigma}{Dt} + \sigma\left(\frac{\partial u}{a\cos\phi\partial\lambda} + \frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial\dot{\theta}}{\partial\theta}\right) = 0, \tag{4.10}$$

where u and v are the zonal and meridional components of the velocity, $\Pi = c_p\,(p/p_{00})^\kappa$ the Exner function, $M = \theta\Pi + gz$ the Montgomery potential, $\sigma = -\partial p/\partial\theta$ the pseudodensity, and

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + u\frac{\partial}{a\cos\phi\partial\lambda} + v\frac{\partial}{a\partial\phi} + \dot{\theta}\frac{\partial}{\partial\theta} \tag{4.11}$$

the total derivative.

4.1.1 Conservation of mass – the continuity equation

Let the longitude, latitude, and potential temperature of marked fluid parcels be denoted by $\lambda(\lambda_0, \phi_0, \theta_0, \tau)$, $\phi(\lambda_0, \phi_0, \theta_0, \tau)$, $\theta(\lambda_0, \phi_0, \theta_0, \tau)$ where $\lambda_0, \phi_0, \theta_0$ are the labelling coordinates and τ is time. The labelling coordinates can be thought of as the initial positions, and since these remain fixed following a parcel, we can interpret $\partial/\partial\tau$ as the total time derivative or the derivative following the motion. The conservation of mass principle can be expressed by $\sigma\cos\phi d\lambda d\phi d\theta = \sigma_0\cos\phi_0 d\lambda_0 d\phi_0 d\theta_0$, or

$$\frac{\sigma_0}{\sigma} = \frac{\partial\,(\lambda, \sin\phi, \theta)}{\partial\,(\lambda_0, \sin\phi_0, \theta_0)}, \tag{4.12}$$

where σ_0 can be interpreted as the constant initial pseudodensity. Alternatively, conservation of mass can be expressed in terms of density ρ in the z coordinate system,

$$\frac{\rho_0}{\rho} = \frac{\partial\,(\lambda, \sin\phi, z)}{\partial\,(\lambda_0, \sin\phi_0, z_0)}, \tag{4.13}$$

where ρ_0 is the constant initial density and the labelling coordinates are λ_0, ϕ_0, z_0.

To derive the familiar form of the continuity equation we take $\partial/\partial\tau$ of (4.12) to obtain

$$\frac{\sigma_0}{\sigma^2}\frac{\partial\sigma}{\partial\tau} + \frac{\partial(\dot{\lambda}, \sin\phi, \theta)}{\partial(\lambda_0, \sin\phi_0, \theta_0)} + \frac{\partial(\lambda, \dot{\phi}\cos\phi, \theta)}{\partial(\lambda_0, \sin\phi_0, \theta_0)} + \frac{\partial(\lambda, \sin\phi, \dot{\theta})}{\partial(\lambda_0, \sin\phi_0, \theta_0)} = 0, \tag{4.14}$$

where $\dot{\lambda} = \partial\lambda/\partial\tau$, etc. The second term in (4.14) can be written

$$\frac{\partial(\dot{\lambda}, \sin\phi, \theta)}{\partial(\lambda_0, \sin\phi_0, \theta_0)} = \frac{\partial(\dot{\lambda}, \sin\phi, \theta)}{\partial(\lambda, \sin\phi, \theta)}\frac{\partial\,(\lambda, \sin\phi, \theta)}{\partial(\lambda_0, \sin\phi_0, \theta_0)} = \frac{\sigma_0}{\sigma}\frac{\partial\dot{\lambda}}{\partial\lambda}, \tag{4.15}$$

with similar expressions for the third and fourth terms in (4.14). Using these we can rewrite (4.14) as

$$\frac{\partial \sigma}{\partial \tau} + \sigma \left(\frac{\partial \dot{\lambda}}{\partial \lambda} + \frac{\partial (\dot{\phi} \cos \phi)}{\cos \phi \partial \phi} + \frac{\partial \dot{\theta}}{\partial \theta} \right) = 0. \tag{4.16}$$

Defining $u = a \cos \phi \dot{\lambda}$ and $v = a\dot{\phi}$ in the usual fashion, (4.16) can be written in the more familiar form (4.10).

4.1.2 Hydrostatic and horizontal momentum equations

The quasi-static, adiabatic and frictionless equations of motion can now be derived from Hamilton's principle

$$\delta \int L d\tau = 0, \tag{4.17}$$

where

$$L = \iiint \left[a^2 \cos^2 \phi \left(\dot{\lambda} + \Omega \right) \frac{\partial \lambda}{\partial \tau} + a^2 \dot{\phi} \frac{\partial \phi}{\partial \tau} \right] \rho_0 \cos \phi_0 d\lambda_0 d\phi_0 dz_0 - H \tag{4.18}$$

is the Lagrangian,

$$H = \iiint \left[\frac{1}{2} \left(a^2 \cos^2 \phi \, \dot{\lambda}^2 + a^2 \dot{\phi}^2 \right) + E(\alpha, \theta) + gz \right] \rho_0 \cos \phi_0 \, d\lambda_0 d\phi_0 dz_0 \tag{4.19}$$

is the Hamiltonian, and $E(\alpha, \theta)$ the internal energy per unit mass, which is a function of specific volume, α, and the entropy. The δ stands for independent variations $\delta\lambda, \delta\phi, \delta z, \delta\dot{\lambda}, \delta\dot{\phi}$ in the three dimensional fluid particle locations and in the horizontal particle velocities. The latter two variations yield $\dot{\lambda} = \partial\lambda/\partial\tau$ and $\dot{\phi} = \partial\phi/\partial\tau$. The variation δz yields

$$\int d\tau \iiint \left[\left(\frac{\partial E}{\partial \alpha} \right)_\theta \alpha_0 \frac{\partial(\lambda, \sin\phi, \delta z)}{\partial(\lambda_0, \sin\phi_0, z_0)} + g\delta z \right] \rho_0 \cos \phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.20}$$

From thermodynamics we have $p = -\left(\partial E/\partial \alpha\right)_\theta$, so that integration by parts allows (4.20) to be written

$$\int d\tau \iiint \delta z \left[\alpha_0 \frac{\partial(\lambda, \sin\phi, p)}{\partial(\lambda_0, \sin\phi_0, z_0)} + g \right] \rho_0 \cos \phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.21}$$

Using (4.13) and noting that δz is arbitrary, we obtain the hydrostatic equation

$$\alpha \frac{\partial p}{\partial z} + g = 0. \tag{4.22}$$

The variation $\delta\lambda$ yields

$$\int d\tau \iiint \left[a^2 \cos^2\phi \left(\dot{\lambda} + \Omega\right) \frac{\partial \delta\lambda}{\partial \tau} - \left(\frac{\partial E}{\partial \alpha}\right)_\theta \alpha_0 \frac{\partial(\delta\lambda, \sin\phi, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.23}$$

Using integration by parts this can be written

$$\int d\tau \iiint \delta\lambda \left\{\frac{\partial}{\partial \tau}\left[a^2 \cos^2\phi \left(\dot{\lambda} + \Omega\right)\right] + \alpha_0 \frac{\partial(p, \sin\phi, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right\} \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.24}$$

Using (4.13) in the last term and noting that $\delta\lambda$ is arbitrary, we obtain

$$\frac{\partial}{\partial \tau}\left[a^2 \cos^2\phi \left(\dot{\lambda} + \Omega\right)\right] + \alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, z)} = 0. \tag{4.25}$$

Since

$$\alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, z)} = \alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, \theta)} \frac{\partial(\lambda, \sin\phi, \theta)}{\partial(\lambda, \sin\phi, z)} = \alpha \left[\left(\frac{\partial p}{\partial \lambda}\right)_\theta \frac{\partial z}{\partial \theta} - \left(\frac{\partial z}{\partial \lambda}\right)_\theta \frac{\partial p}{\partial \theta}\right] \frac{\partial \theta}{\partial z}$$
$$= \alpha \left(\frac{\partial p}{\partial \lambda}\right)_\theta + g \left(\frac{\partial z}{\partial \lambda}\right)_\theta = \frac{\partial M}{\partial \lambda},$$

we can rewrite (4.25) as

$$\frac{\partial}{\partial \tau}\left[a \cos^2\phi \left(\dot{\lambda} + \Omega\right)\right] + \frac{\partial M}{a \partial \lambda} = 0, \tag{4.26}$$

which is the absolute angular momentum equation. Similarly, the variation $\delta\phi$ yields

$$a \frac{\partial \dot{\phi}}{\partial \tau} + \left(2\Omega + \dot{\lambda}\right) \sin\phi \, a \cos\phi \, \dot{\lambda} + \frac{\partial M}{a \partial \phi} = 0. \tag{4.27}$$

Equations (4.26) and (4.27) can also be written in the more familiar Eulerian forms (4.7) and (4.8).

4.1.3 Conservation relations

By Noether's theorem conservation relations can be found by considering symmetries in the Lagrangian, or equivalently, in the Hamiltonian of the dynamical system. If the Lagrangian does not explicitly contain a generalized coordinate then the corresponding canonical momentum is conserved, be it linear momentum in the case of translation along a coordinate axis, or angular momentum in the case of rotation by a coordinate angle about an axis. The term symmetry is used since the lack of explicit dependence of the Lagrangian on the coordinate implies that the system can be translated along the coordinate axis in question (translational symmetry) or rotated about the coordinate angle in question (rotational symmetry) without affecting the action integral and hence the dynamics. Energy conservation corresponds to invariance of the Lagrangian or the Hamiltonian with respect to translation in time.

Noether's theorem is, however, considerably more comprehensive and powerful than indicated above since it applies equally well to dependent variables or field quantities. Thus, every parameter associated with infinitesimal transformations which leave the action integral (4.2) invariant leads to a conservation law. Note that the theorem involves continuous transformations and it assumes form-invariance of the Lagrangian, i.e., it assumes that the Lagrangian has the same functional form in terms of the transformed variables as it had in terms of the original ones.

Examining the Hamiltonian (4.19) we see that if α is independent of λ, which by (4.13) implies that the flow is zonally symmetric, then we find that angular momentum is conserved, which is exactly what is expressed in (4.26) for zonally symmetric flow. Similarly, since (4.19) does not contain an explicit time dependence the total energy (which is just the Hamiltonian) is conserved.

Salmon (1982, 1983, 1988) has shown that potential vorticity conservation corresponds to the symmetry property of the labelling coordinates entering the Lagrangian only through the Jacobian (4.12). Let us now adopt the θ coordinate system, i.e., the labelling coordinates are $\lambda_0, \phi_0, \theta_0$. For adiabatic flow, $\theta(\lambda_0, \phi_0, \theta_0, \tau) = \theta_0$. Consider variations in the labels which leave the Jacobian unchanged, i.e.,

$$\delta\left(\frac{\partial(\lambda_0, \sin\phi_0, \theta_0)}{\partial(\lambda, \sin\phi, \theta)}\right) = 0. \tag{4.28}$$

We assume that $\delta\theta_0 = 0$ so that fluid particles are relabelled within isentropic surfaces. Equation (4.28) then becomes

$$\frac{\partial(\delta\lambda_0, \sin\phi_0, \theta_0)}{\partial(\lambda_0, \sin\phi_0, \theta_0)}\frac{\partial(\lambda_0, \sin\phi_0, \theta_0)}{\partial(\lambda, \sin\phi, \theta)} + \frac{\partial(\lambda_0, \delta\sin\phi_0, \theta_0)}{\partial(\lambda_0, \sin\phi_0, \theta_0)}\frac{\partial(\lambda_0, \sin\phi_0, \theta_0)}{\partial(\lambda, \sin\phi, \theta)} = 0, \tag{4.29}$$

or

$$\left(\frac{\partial \delta\lambda_0}{\partial\lambda_0} + \frac{\partial\left(\delta\phi_0 \cos\phi_0\right)}{\cos\phi_0 \partial\phi_0}\right)\frac{\sigma}{\sigma_0} = 0. \tag{4.30}$$

Thus, $\delta\lambda_0$ and $\delta\phi_0$ must be related by

$$\delta\lambda_0 \ \cos\phi_0 = -\frac{\partial\delta\psi}{\partial\phi_0}, \qquad \delta\phi_0 \ \cos\phi_0 = \frac{\partial\delta\psi}{\partial\lambda_0}, \tag{4.31}$$

where $\delta\psi$ is arbitrary. From (4.17) we then have

$$\int d\tau \iiint \left[\cos^2\phi\left(\dot\lambda+\Omega\right)\delta\left(\frac{\partial\lambda}{\partial\tau}\right) + \dot\phi\,\delta\left(\frac{\partial\phi}{\partial\tau}\right)\right]\sigma_0 \cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0, \tag{4.32}$$

where the Lagrangian is expressed in the θ-coordinate system. Using

$$\delta\left(\frac{\partial\lambda}{\partial\tau}\right) = -\frac{\partial\lambda}{\partial\lambda_0}\frac{\partial\delta\lambda_0}{\partial\tau} - \frac{\partial\lambda}{\partial\phi_0}\frac{\partial\delta\phi_0}{\partial\tau} \tag{4.33}$$

$$\delta\left(\frac{\partial\phi}{\partial\tau}\right) = -\frac{\partial\phi}{\partial\lambda_0}\frac{\partial\delta\lambda_0}{\partial\tau} - \frac{\partial\phi}{\partial\phi_0}\frac{\partial\delta\phi_0}{\partial\tau} \tag{4.34}$$

in (4.32) and integrating by parts in time, we obtain

$$\int d\tau \iiint \left\{\delta\lambda_0 \frac{\partial}{\partial\tau}\left[\cos^2\phi\left(\dot\lambda+\Omega\right)\frac{\partial\lambda}{\partial\lambda_0} + \dot\phi\frac{\partial\phi}{\partial\lambda_0}\right]\right.$$
$$\left. + \delta\phi_0\frac{\partial}{\partial\tau}\left[\cos^2\phi\left(\dot\lambda+\Omega\right)\frac{\partial\lambda}{\partial\phi_0} + \dot\phi\frac{\partial\phi}{\partial\phi_0}\right]\right\}\sigma_0\cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0. \tag{4.35}$$

Using (4.31), integrating by parts with respect to λ_0 and ϕ_0, and recalling that $\delta\psi$ is arbitrary, we obtain from (4.35)

$$\frac{\partial}{\partial\tau}\left\{\frac{\partial\left(\cos^2\phi\left(\dot\lambda+\Omega\right),\lambda,\theta\right)}{\partial\left(\lambda_0,\phi_0,\theta_0\right)} + \frac{\partial\left(\dot\phi,\phi,\theta\right)}{\partial\left(\lambda_0,\phi_0,\theta_0\right)}\right\} = 0. \tag{4.36}$$

Using (4.12), this can be written

$$\frac{\partial}{\partial\tau}\left\{\frac{\sigma_0}{\sigma}\left[\frac{\partial\dot\phi}{\cos\phi\partial\lambda} - \frac{\partial}{\cos\phi\partial\phi}\left(\cos^2\phi(\dot\lambda+\Omega)\right)\right]\right\} = 0, \tag{4.37}$$

or

$$\frac{D}{Dt}\left\{\frac{\sigma_0}{\sigma}\left[2\Omega\sin\phi + \frac{\partial v}{a\cos\phi\partial\lambda} - \frac{\partial\left(u\cos\phi\right)}{a\cos\phi\partial\phi}\right]\right\} = 0, \tag{4.38}$$

which is the Rossby-Ertel potential vorticity equation on the sphere.

4.2 Derivation of the semigeostrophic equations on the sphere from Hamilton's principle

It has been shown in the previous section that the continuity equation is implicit in the Lagrangian coordinates and that the equations of motion and the hydrostatic equation can be derived from Hamilton's principle. This makes the principle an extremely succinct statement of dynamics. Also the conservation relations of the dynamical system can easily be deduced from symmetry properties of its Lagrangian or Hamiltonian. A third property that will be further explored in this section concerns the fact that Hamiltonian methods are not restricted to a particular coordinate system.

In this section we will derive a whole set of balanced models by approximating the horizontal wind (u, v) by a balanced wind (u_b, v_b) while at the same time defining a coordinate transformation which makes the Lagrangian attain a particularly simple form. In fact the new coordinates are almost canonical. Care is taken along the way not to disturb the important symmetry properties of the Hamiltonian. This approach allows us to derive the most general balanced models while preserving conservation properties of the original system and while casting the physics into their simplest mathematical form. The form is that of one very simple prognostic equation for the potential pseudodensity (4.68) and an invertibility relation (4.69) which allows us to diagnose the balanced wind field from the potential pseudodensity field. It is then a simple matter to include diabatic effects as in (4.70). Defining the balance to be the geostrophic balance of chapter 3 ((3.19) and (3.20)) and the coordinates to be the corresponding vortex coordinates (3.29) allows us to derive the equations of the geostrophic momentum approximation generalized to the sphere.

4.2.1 The approximate Lagrangian

Consider applying Hamilton's principle to the approximate Lagrangian

$$\delta \int L_b d\tau = 0, \tag{4.39}$$

where

$$L_b = \iiint a^2 \Omega \cos^2 \Phi \frac{\partial \Lambda}{\partial \tau} \sigma_0 \cos\phi_0 \, d\lambda_0 d\phi_0 d\theta_0 - H_b \tag{4.40}$$

is the approximate Lagrangian,

$$H_b = \iiint \left[\frac{1}{2}\left(a^2\cos^2\Phi\,\dot{\Lambda}^2 + a^2\dot{\Phi}^2\right) + E(\alpha,\theta) + gz\right]\sigma_0\cos\phi_0\,d\lambda_0 d\phi_0 d\theta_0 \qquad (4.41$$

is the approximate Hamiltonian, and δ now stands for independent variations $\delta\Lambda, \delta\Phi$ ir the transformed fluid parcel locations. As for the transformed coordinates, Λ, Φ, Θ, T $(\Theta, T) = (\theta, t)$ and (Λ, Φ) will be defined later in terms of (λ, ϕ) and $(\dot{\Lambda}, \dot{\Phi})$, which in turr will be related to the M field. The variations $\delta\Lambda$ and $\delta\Phi$ now give

$$2\Omega\sin\Phi\, a\frac{\partial\sin\Phi}{\partial\tau} = \frac{\delta H_b}{a\delta\Lambda}, \qquad (4.42)$$

$$-2\Omega\sin\Phi\, a\frac{\partial\Lambda}{\partial\tau} = \frac{\delta H_b}{a\delta\sin\Phi}, \qquad (4.43)$$

where the notation on the right hand sides stands for functional derivatives.

The above follows the same approach as did Salmon (1985). The coordinate transformation is defined exactly such that L_b takes the form (4.40) and the resulting momentum equations take the form (4.42) and (4.43). In section 4.2.3 it will be shown that the functional derivatives of the approximate Hamiltonian are precisely equal to the corresponding partial derivatives of a potential function M^* in which case the two momentum equations are almost canonical. They would be canonical if $2\Omega\sin\Phi$ were constant.

4.2.2 Conservation relations

The observations made in section 4.1.3 on the conservation of angular momentum in the zonally symmetric case, the conservation of total energy and the conservation of potential vorticity still apply since we were careful not to destroy any symmetries in the Hamiltonian when making the approximations. In this case the angular momentum corresponds to $a\Omega\cos^2\Phi$ and the kinetic energy corresponds to $(u_b^2 + v_b^2)/2$. For potential vorticity, as in section 4.1.3 we consider variations in the particle labels which leave the Jacobian (4.12) unchanged. From (4.39) we have

$$\int d\tau \iiint \Omega\cos^2\Phi\;\delta\left(\frac{\partial\Lambda}{\partial\tau}\right)\sigma_0\cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0. \qquad (4.44)$$

Using

$$\delta\left(\frac{\partial\Lambda}{\partial\tau}\right) = -\frac{\partial\Lambda}{\partial\lambda_0}\frac{\partial\delta\lambda_0}{\partial\tau} - \frac{\partial\Lambda}{\partial\phi_0}\frac{\partial\delta\phi_0}{\partial\tau} \qquad (4.45)$$

in (4.44) and integrating by parts in time, we obtain

$$\int d\tau \iiint \left\{ \delta\lambda_0 \frac{\partial}{\partial\tau}\left[\Omega\cos^2\Phi \frac{\partial\Lambda}{\partial\lambda_0}\right] + \delta\phi_0 \frac{\partial}{\partial\tau}\left[\Omega\cos^2\Phi\frac{\partial\Lambda}{\partial\phi_0}\right]\right\} \sigma_0 \cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0. \tag{4.46}$$

This equation is equivalent to (4.35) for the unapproximated case. Using (4.31), integrating by parts with respect to λ_0 and ϕ_0, and recalling that $\delta\psi$ is arbitrary, we obtain

$$\frac{\partial}{\partial\tau}\left[\frac{\partial\left(\Omega\cos^2\Phi, \Lambda, \Theta\right)}{\partial\left(\lambda_0, \phi_0, \theta_0\right)}\right] = 0, \tag{4.47}$$

which can also be written as

$$\frac{\partial}{\partial\tau}\left[\frac{\partial\left(\Omega\cos^2\Phi,\Lambda,\Theta\right)}{\partial\left(\Lambda,\Phi,\Theta\right)}\frac{\partial\left(\Lambda,\Phi,\Theta\right)}{\partial\left(\lambda,\phi,\theta\right)}\frac{\partial\left(\lambda,\sin\phi,\theta\right)}{\partial\left(\lambda_0,\sin\phi_0,\theta_0\right)}\frac{\cos\phi_0}{\cos\phi}\right] = 0. \tag{4.48}$$

Using (4.12) for the last Jacobian and noting that the first Jacobian is $2\Omega\cos\Phi\sin\Phi$, we obtain

$$\frac{\partial}{\partial\tau}\left[\frac{\sigma_0}{\sigma}2\Omega\sin\Phi\frac{\cos\Phi}{\cos\phi}\frac{\partial\left(\Lambda,\Phi\right)}{\partial\left(\lambda,\phi\right)}\right] = 0, \tag{4.49}$$

which is a statement of the conservation of the approximate potential vorticity P_b, defined as

$$P_b = \frac{1}{\sigma}2\Omega\sin\Phi\frac{\partial\left(\Lambda,\sin\Phi\right)}{\partial\left(\lambda,\sin\phi\right)}. \tag{4.50}$$

4.2.3 The approximate potential pseudodensity equation and the invertibility principle

We will now proceed to derive the two fundamental equations of the balanced model, the approximate potential pseudodensity equation and the invertibility relation. Let us define the approximate potential pseudodensity σ_b^* as

$$\sigma_b^* = \sigma\left(\frac{\partial\left(\Lambda,\sin\Phi\right)}{\partial\left(\lambda,\sin\phi\right)}\right)^{-1}, \tag{4.51}$$

so that $P_b\sigma_b^* = 2\Omega\sin\Phi$. Using (4.12), (4.51) can be written as

$$\sigma_b^* = \sigma_0\left(\frac{\partial\left(\Lambda,\sin\Phi\right)}{\partial\left(\lambda_0,\sin\phi_0\right)}\right)^{-1}. \tag{4.52}$$

Taking $\partial/\partial\tau$ of (4.52) we obtain

$$\frac{\partial\sigma_b^*}{\partial\tau}+\sigma_0\left[\frac{\partial(\Lambda,\sin\Phi)}{\partial(\lambda_0,\sin\phi_0)}\right]^{-2}\left[\frac{\partial\Lambda}{\partial\lambda_0}\frac{\partial}{\partial\sin\phi_0}\left(\frac{\partial\sin\Phi}{\partial\tau}\right)+\frac{\partial}{\partial\lambda_0}\left(\frac{\partial\Lambda}{\partial\tau}\right)\frac{\partial\sin\Phi}{\partial\sin\phi_0}\right.$$
$$\left.-\frac{\partial\Lambda}{\partial\sin\phi_0}\frac{\partial}{\partial\lambda_0}\left(\frac{\partial\sin\Phi}{\partial\tau}\right)-\frac{\partial}{\partial\sin\phi_0}\left(\frac{\partial\Lambda}{\partial\tau}\right)\frac{\partial\sin\Phi}{\partial\lambda_0}\right]=0. \tag{4.53}$$

Derivatives in $(\lambda_0,\sin\phi_0,\theta_0,\tau)$ coordinates are related to derivatives in $(\Lambda,\sin\Phi,\Theta,T)$ coordinates by

$$\frac{\partial}{\partial\tau} = \frac{\partial}{\partial T}+\frac{\partial\Lambda}{\partial\tau}\frac{\partial}{\partial\Lambda}+\frac{\partial\sin\Phi}{\partial\tau}\frac{\partial}{\partial\sin\Phi}, \tag{4.54}$$

$$\frac{\partial}{\partial\lambda_0} = \frac{\partial\Lambda}{\partial\lambda_0}\frac{\partial}{\partial\Lambda}+\frac{\partial\sin\Phi}{\partial\lambda_0}\frac{\partial}{\partial\sin\Phi}, \tag{4.55}$$

$$\frac{\partial}{\partial\sin\phi_0} = \frac{\partial\Lambda}{\partial\sin\phi_0}\frac{\partial}{\partial\Lambda}+\frac{\partial\sin\Phi}{\partial\sin\phi_0}\frac{\partial}{\partial\sin\Phi}. \tag{4.56}$$

Inverting (4.55) and (4.56) we obtain

$$\frac{\partial(\Lambda,\sin\Phi)}{\partial(\lambda_0,\sin\phi_0)}\frac{\partial}{\partial\Lambda} = \frac{\partial\sin\Phi}{\partial\sin\phi_0}\frac{\partial}{\partial\lambda_0}-\frac{\partial\sin\Phi}{\partial\lambda_0}\frac{\partial}{\partial\sin\phi_0}, \tag{4.57}$$

$$\frac{\partial(\Lambda,\sin\Phi)}{\partial(\lambda_0,\sin\phi_0)}\frac{\partial}{\partial\sin\Phi} = -\frac{\partial\Lambda}{\partial\sin\phi_0}\frac{\partial}{\partial\lambda_0}+\frac{\partial\Lambda}{\partial\lambda_0}\frac{\partial}{\partial\sin\phi_0}. \tag{4.58}$$

Using (4.54), (4.57) and (4.58) in (4.53) we obtain the approximate potential pseudodensity equation

$$\frac{\partial\sigma_b^*}{\partial T}+\frac{\partial}{\partial\Lambda}\left(\sigma_b^*\frac{\partial\Lambda}{\partial\tau}\right)+\frac{\partial}{\partial\sin\Phi}\left(\sigma_b^*\frac{\partial\sin\Phi}{\partial\tau}\right)=0, \tag{4.59}$$

which is the fundamental prognostic equation of the model. We will now simplify it.

Let us turn to the approximate potential vorticity equation (4.49), which was derived in the last subsection. Using (4.12), (4.52) and (4.54) we can express it as

$$\frac{\partial\sigma_b^*}{\partial T}+2\Omega\sin\Phi\left[\frac{\partial\Lambda}{\partial\tau}\frac{\partial}{\partial\Lambda}\left(\frac{\sigma_b^*}{2\Omega\sin\Phi}\right)+\frac{\partial\sin\Phi}{\partial\tau}\frac{\partial}{\partial\sin\Phi}\left(\frac{\sigma_b^*}{2\Omega\sin\Phi}\right)\right]=0. \tag{4.60}$$

Taking the difference between (4.59) and (4.60) we obtain

$$\frac{\partial}{\partial\Lambda}\left(2\Omega\sin\Phi\frac{\partial\Lambda}{\partial\tau}\right)+\frac{\partial}{\partial\sin\Phi}\left(2\Omega\sin\Phi\frac{\partial\sin\Phi}{\partial\tau}\right)=0, \tag{4.61}$$

which implies that

$$2\Omega \sin\Phi\, a \frac{\partial \sin\Phi}{\partial \tau} = \frac{\partial M^*}{a\partial \Lambda}, \tag{4.62}$$

$$-2\Omega \sin\Phi\, a \frac{\partial \Lambda}{\partial \tau} = \frac{\partial M^*}{a\partial \sin\Phi}, \tag{4.63}$$

for some function M^* related to M. The factor a, the earth's radius, in (4.62) and (4.63) is inserted for dimensional consistency. These are just the two momentum equations and they are almost canonical in form. In fact they can be further transformed so that they are canonical (Salmon, 1985), however (4.62) and (4.63) are already so simple that it is not worth the effort. Comparing (4.62) and (4.63) with (4.42) and (4.43) we note that

$$\frac{\delta H_b}{\delta \Lambda} = \frac{\partial M^*}{\partial \Lambda} \tag{4.64}$$

and

$$\frac{\delta H_b}{\delta \sin\Phi} = \frac{\partial M^*}{\partial \sin\Phi}. \tag{4.65}$$

Additionally, we will want the hydrostatic equation (4.9) to look the same in the transformed coordinates and the time derivative of M^* to maintain the same form as in physical space. Thus, we require

$$\frac{\partial M}{\partial t} = \frac{\partial M^*}{\partial T}, \tag{4.66}$$

$$\frac{\partial M}{\partial \theta} = \frac{\partial M^*}{\partial \Theta}. \tag{4.67}$$

Using (4.62) and (4.63) we can write (4.59) as

$$\frac{\partial \sigma_b^*}{\partial T} + \frac{\partial}{\partial \Lambda}\left(-\frac{\sigma_b^*}{2\Omega a^2 \sin\Phi}\frac{\partial M^*}{\partial \sin\Phi}\right) + \frac{\partial}{\partial \sin\Phi}\left(\frac{\sigma_b^*}{2\Omega a^2 \sin\Phi}\frac{\partial M^*}{\partial \Lambda}\right) = 0, \tag{4.68}$$

which is the fundamental predictive equation of the model.

Our final task is to find the invertibility relation, i.e., the relation between M^* and σ_b^*. To do this we first note that (4.51) can be written

$$\sigma \frac{\partial(\lambda, \sin\phi)}{\partial(\Lambda, \sin\Phi)} = \sigma_b^*. \tag{4.69}$$

Now we can substitute for λ and ϕ in terms of $\Lambda, \Phi, \dot{\Lambda}$ and $\dot{\Phi}$, where the latter two can in turn be expressed in terms of $\partial M/\partial\lambda$ and $\partial M/\partial\phi$ which can then be related to horizontal derivatives of M^* in capital space. Thus, (4.69) becomes a partial differential equation which allows us to diagnose M^* from σ_b^*. The invertibility relation (4.69) is the fundamental diagnostic equation of the model. Numerical integration of the model is performed entirely in (Λ, Φ, Θ) space by predicting new values of σ_b^* using (4.68) and by inverting (4.69) at each time step. The transformation of the results to (λ, ϕ, θ) space need only be done when output maps are desired.

We can regard (4.68) and (4.69) as the governing equations of the generic balanced model or as the equations for a class of approximate models. A particular member of this class results from a particular choice of balance conditions (i.e., relations between $\Lambda, \Phi, \dot{\Lambda}, \dot{\Phi}$ and $\lambda, \phi, \partial M/\partial\lambda, \partial M/\partial\phi$) and coordinate transformations (i.e., relations between $\Lambda, \Phi, \lambda, \phi, \dot{\Lambda}, \dot{\Phi}$). In this sense (4.39)–(4.41) generates a class of approximations with the generic potential pseudodensity equation (4.68) and the generic invertibility condition (4.69). Note that (4.64)–(4.67) put restrictions and compatibility requirements on the definitions of M^*, the balance conditions and the coordinate transformation.

If heating is included an extra term is added to (4.68) corresponding to a flux of σ_b^* across isentropic surfaces. Thus, with heating (4.68) becomes

$$\frac{\partial \sigma_b^*}{\partial T} + \frac{\partial}{\partial \Lambda}\left(-\frac{\sigma_b^*}{2\Omega a^2 \sin\Phi}\frac{\partial M^*}{\partial \sin\Phi}\right) + \frac{\partial}{\partial \sin\Phi}\left(\frac{\sigma_b^*}{2\Omega a^2 \sin\Phi}\frac{\partial M^*}{\partial \Lambda}\right) + \frac{\partial}{\partial \Theta}\left(\sigma_b^*\dot{\theta}\right) = 0, \quad (4.70)$$

The invertibility principle can be written in terms of a three dimensional Jacobian. The hydrostatic equation (4.9) provides the link between the σ and the M^* field. Thus, (4.69) becomes

$$\frac{\partial(\lambda, \sin\phi, \Pi)}{\partial(\Lambda, \sin\Phi, \Theta)} + \Gamma\sigma_b^* = 0, \quad (4.71)$$

where $\Gamma = d\Pi/dp = \kappa\Pi/p$.

4.2.4 A particular balanced model

Let us now define the geostrophic wind on the sphere, (u_g, v_g), in terms of M as follows

$$-2\Omega\sin\Phi\cos\Phi\, v_g + \frac{\partial M}{a\partial\lambda} = 0, \quad (4.72)$$

$$2\Omega\sin\Phi\, u_g\frac{\cos\phi}{\cos\Phi} + \frac{\partial M}{a\partial\phi} = 0. \quad (4.73)$$

The transformed coordinates are

$$\Lambda = \lambda + \frac{v_g}{a2\Omega \sin\Phi \cos\Phi}, \tag{4.74}$$

$$\sin\Phi = \sin\phi - \frac{u_g \cos\Phi}{a2\Omega \sin\Phi}. \tag{4.75}$$

As before, $(\Theta, T) = (\theta, t)$. Derivatives in $(\lambda, \phi, \theta, t)$ space are related to derivatives in $(\Lambda, \Phi, \Theta, T)$ space by

$$\frac{\partial}{\partial t} = \frac{\partial \Lambda}{\partial t}\frac{\partial}{\partial \Lambda} + \frac{\partial \Phi}{\partial t}\frac{\partial}{\partial \Phi} + \frac{\partial}{\partial T}, \tag{4.76}$$

$$\frac{\partial}{\partial \lambda} = \frac{\partial \Lambda}{\partial \lambda}\frac{\partial}{\partial \Lambda} + \frac{\partial \Phi}{\partial \lambda}\frac{\partial}{\partial \Phi}, \tag{4.77}$$

$$\frac{\partial}{\partial \phi} = \frac{\partial \Lambda}{\partial \phi}\frac{\partial}{\partial \Lambda} + \frac{\partial \Phi}{\partial \phi}\frac{\partial}{\partial \Phi}, \tag{4.78}$$

$$\frac{\partial}{\partial \theta} = \frac{\partial \Lambda}{\partial \theta}\frac{\partial}{\partial \Lambda} + \frac{\partial \Phi}{\partial \theta}\frac{\partial}{\partial \Phi} + \frac{\partial}{\partial \Theta}. \tag{4.79}$$

Defining M^* as

$$M^* = M + \frac{1}{2}\left(u_g^2 + v_g^2\right), \tag{4.80}$$

it can be shown that (4.64)–(4.67) are valid. Now,

$$\frac{\partial M^*}{\partial \Lambda} = u_g \frac{\partial u_g}{\partial \Lambda} + v_g \frac{\partial v_g}{\partial \Lambda} + \frac{\partial M}{\partial \lambda}\frac{\partial \lambda}{\partial \Lambda} + \frac{\partial M}{\partial \phi}\frac{\partial \phi}{\partial \Phi},$$

$$\frac{\partial M^*}{\partial \Phi} = u_g \frac{\partial u_g}{\partial \Phi} + v_g \frac{\partial v_g}{\partial \Phi} + \frac{\partial M}{\partial \lambda}\frac{\partial \lambda}{\partial \Phi} + \frac{\partial M}{\partial \phi}\frac{\partial \phi}{\partial \Phi}.$$

Differentiating (4.74) and (4.75) with respect to Λ and Φ, and using (4.72) and (4.73) we can derive the balance conditions in transformed space. They can be expressed as

$$-2\Omega \sin\Phi \cos\Phi \, v_g + \frac{\partial M^*}{a\partial \Lambda} = 0, \tag{4.81}$$

$$2\Omega \sin\Phi \, u_g - \frac{u_g^2 + v_g^2(\cos^2\Phi - \sin^2\Phi)}{a \sin\Phi \cos\Phi} + \frac{\partial M^*}{a\partial \Phi} = 0, \tag{4.82}$$

or

$$\left(\frac{\partial M}{\partial \lambda}, \frac{\partial M}{\partial \theta}, \frac{\partial M}{\partial t}\right) = \left(\frac{\partial M^*}{\partial \Lambda}, \frac{\partial M^*}{\partial \Theta}, \frac{\partial M^*}{\partial T}\right),$$
$$\frac{\partial M}{\partial \phi} = \frac{\cos\phi}{\cos\Phi}\left[\frac{\partial M^*}{\partial \Phi} - \frac{u_g^2 + v_g^2\,(\cos^2\Phi - \sin^2\Phi)}{\sin\Phi\cos\Phi}\right]. \tag{4.83}$$

We can now derive in physical space the momentum equations with the geostrophic momentum approximation. Taking the total time derivative of (4.75), using (4.62) and (4.72) gives us the geostrophic zonal momentum equation,

$$\cos\Phi\frac{Du_g}{Dt} - 2\Omega\Big[\sin\Phi\, v\cos\phi + (\sin\phi - \sin\Phi)\, v_g\cos\Phi\Big]$$
$$-\frac{u_g v_g \sin\Phi}{a} + \frac{\partial M}{a\partial\lambda} = 0, \tag{4.84}$$

which is an approximation to (4.7) in section 4.1. Similarly, the total time derivative of (4.74), using (4.62), (4.63), (4.81), and (4.74) gives the geostrophic meridional momentum equation,

$$\cos\phi\frac{Dv_g}{Dt} + 2\Omega\Big[\sin\Phi\, u\cos\Phi + (\sin\phi - \sin\Phi)\, u_g\cos\phi\Big]$$
$$+\frac{u_g^2\sin\Phi\cos\phi}{a\cos\Phi} + \cos\Phi\frac{\partial M}{a\partial\phi} = 0, \tag{4.85}$$

which is an approximation to (4.8) in section 4.1. The approximate momentum equations (4.84) and (4.85) are identical to the geostrophic momentum approximation equations generalized to the sphere (3.17) and (3.18) which were introduced in chapter 3. Thus, the same comments as to their nature apply here as before.

4.3 Derivation of the equations of motion for the long wave approximation

The long wave approximation was proposed in the remarkable paper by Gill, (1980) in the context of a simple barotropic, linear model on the equatorial β-plane. In this theory the zonal wind is balanced, the meridional wind remains unbalanced, but the meridional acceleration is neglected, hence the meridional wind does not enter in the definitions of the conservative properties resulting in a filtered theory. The linear dynamics include the Kelvin wave and Rossby waves whose dispersion relation is tangent to the unapproximated dispersion relation at the origin of a frequency-wavenumber graph, (hence the name, long wave approximation).

In this section the adiabatic, frictionless, nonlinear equations of motion corresponding to the long wave approximation are derived by applying Hamilton's principle to a Larangian which neglects the meridional wind. The corresponding conservation relations are subsequently derived.

4.3.1 The approximate Lagrangian

Let us apply Hamilton's principle to the approximate Lagrangian

$$\delta \int L_l d\tau = 0, \tag{4.86}$$

where

$$L_l = \iiint a^2 \cos^2\phi \left(\dot{\lambda} + \Omega\right) \frac{\partial \lambda}{\partial \tau} \rho_0 \cos\phi_0 d\lambda_0 d\phi_0 dz_0 - H_l \tag{4.87}$$

is the Lagrangian corresponding to the adiabatic, frictionless long wave approximation,

$$H_l = \iiint \left[\frac{1}{2} a^2 \cos^2\phi \dot{\lambda}^2 + E(\alpha,\theta) + gz\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 \tag{4.88}$$

the corresponding Hamiltonian, and $E(\alpha,\theta)$ the internal energy per unit mass, which is a function of the specific volume, α, and the entropy. Now, the δ in (4.86) stands for independent variations $\delta\lambda, \delta\phi, \delta z, \delta\dot{\lambda}$ in the three dimensional fluid particle locations and in the particle zonal velocity. Note that the meridional velocity is no longer included. It is neglected in the definition of the approximate Lagrangian and thus does not enter the problem. This is entirely consistent with the treatment of vertical velocity when the hydrostatic approximation is implied both above and when deriving the primitive equations in section 4.1.2.

The variation $\delta\dot{\lambda}$ yields $\dot{\lambda} = \partial\lambda/\partial\tau$. The variation δz yields

$$\int d\tau \iiint \left[\left(\frac{\partial E}{\partial \alpha}\right)_\theta \alpha_0 \frac{\partial(\lambda, \sin\phi, \delta z)}{\partial(\lambda_0, \sin\phi_0, z_0)} + g\delta z\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.89}$$

Noting that from thermodynamics we have $p = -\left(\partial E/\partial\alpha\right)_\theta$, integrating (4.89) by parts and using (4.13), we obtain the hydrostatic equation

$$\alpha \frac{\partial p}{\partial z} + g = 0,$$

which is exactly the same as (4.22).

The variation $\delta\lambda$ yields

$$\int d\tau \iiint \left[a^2 \cos^2\phi \left(\dot{\lambda}+\Omega\right) \frac{\partial \delta\lambda}{\partial \tau} - \left(\frac{\partial E}{\partial \alpha}\right)_\theta \alpha_0 \frac{\partial(\delta\lambda, \sin\phi, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.91}$$

Using integration by parts this can be written

$$\int d\tau \iiint \delta\lambda \left\{\frac{\partial}{\partial \tau}\left[a^2\cos^2\phi\left(\dot{\lambda}+\Omega\right)\right] + \alpha_0 \frac{\partial(p, \sin\phi, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right\} \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.92}$$

Using (4.13) in the last term and noting that $\delta\lambda$ is arbitrary, we obtain

$$\frac{\partial}{\partial \tau}\left[a^2\cos^2\phi\left(\dot{\lambda}+\Omega\right)\right] + \alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, z)} = 0. \tag{4.93}$$

Since

$$\alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, z)} = \alpha \frac{\partial(p, \sin\phi, z)}{\partial(\lambda, \sin\phi, \theta)} \frac{\partial(\lambda, \sin\phi, \theta)}{\partial(\lambda, \sin\phi, z)} = \alpha \left[\left(\frac{\partial p}{\partial \lambda}\right)_\theta \frac{\partial z}{\partial \theta} - \left(\frac{\partial z}{\partial \lambda}\right)_\theta \frac{\partial p}{\partial \theta}\right] \frac{\partial \theta}{\partial z}$$
$$= \alpha \left(\frac{\partial p}{\partial \lambda}\right)_\theta + g\left(\frac{\partial z}{\partial \lambda}\right)_\theta = \frac{\partial M}{\partial \lambda},$$

we can rewrite (4.93) as

$$\frac{\partial}{\partial \tau}\left[a^2\cos^2\phi\left(\dot{\lambda}+\Omega\right)\right] + \frac{\partial M}{\partial \lambda} = 0,$$

which is the absolute angular momentum equation. It is identical to (4.26), as would be expected, since with the long wave approximation the zonal wind remains a predictive variable.

The variation $\delta\phi$ yields

$$\int d\tau \iiint \left[-2a^2\cos\phi\sin\phi\left(\dot{\lambda}+\Omega\right)\dot{\lambda}\,\delta\phi + a^2\cos\phi\sin\phi\,\dot{\lambda}^2\,\delta\phi\right.$$
$$\left. - \left(\frac{\partial E}{\partial \alpha}\right)_\theta \alpha_0 \frac{\partial(\lambda, \cos\phi\,\delta\phi, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.94}$$

Using integration by parts this can be written

$$\int d\tau \iiint \delta\phi \left[a^2\cos\phi\sin\phi\,\dot{\lambda}^2 + 2a^2\cos\phi\sin\phi\,\Omega\dot{\lambda}\right.$$
$$\left. +\alpha_0\cos\phi \frac{\partial(\lambda, p, z)}{\partial(\lambda_0, \sin\phi_0, z_0)}\right] \rho_0 \cos\phi_0 \, d\lambda_0 d\phi_0 dz_0 = 0. \tag{4.95}$$

Using (4.13) in the last term and noting that $\delta\phi$ is arbitrary, we obtain

$$2\Omega\sin\phi\, a\cos\phi\,\dot{\lambda} + \frac{\tan\phi}{a}\left(a\cos\phi\,\dot{\lambda}\right)^2 + \frac{\alpha\cos\phi}{a}\frac{\partial(\lambda,p,z)}{\partial(\lambda,\sin\phi,z)} = 0. \qquad (4.96)$$

As before the Jacobian term can be written in terms of a derivative of M, the Montgomery potential, and we get

$$\left(2\Omega\sin\phi + \frac{u\,\tan\phi}{a}\right)u + \frac{\partial M}{a\partial\phi} = 0. \qquad (4.97)$$

The system of equations, (4.16), (4.22), (4.26) and (4.97), is the same as (1a) - (1d) with the isentropic coordinate in Stevens et al. (1989).

4.3.2 Conservation relations

Again the observations made in section 4.1.3 on the conservation of angular momentum in the zonally symmetric case, the conservation of total energy and the conservation of potential vorticity still apply since we were careful not to destroy any symmetries in the Hamiltonian when making the approximations. In this case the angular momentum equation is the same as for the primitive equations, but the kinetic energy corresponds to $u^2/2$. For potential vorticity, as in section 4.1.3 we consider variations in the particle labels which leave the Jacobian (4.12) unchanged. From (4.86) we have

$$\int d\tau \iiint \cos^2\phi\,\left(\dot{\lambda}+\Omega\right)\,\delta\left(\frac{\partial\lambda}{\partial\tau}\right)\sigma_0\cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0. \qquad (4.98)$$

Using

$$\delta\left(\frac{\partial\lambda}{\partial\tau}\right) = -\frac{\partial\lambda}{\partial\lambda_0}\frac{\partial\delta\lambda_0}{\partial\tau} - \frac{\partial\lambda}{\partial\phi_0}\frac{\partial\delta\phi_0}{\partial\tau} \qquad (4.99)$$

in (4.98) and integrating by parts in time, we obtain

$$\int d\tau \iiint \left\{\delta\lambda_0\frac{\partial}{\partial\tau}\left[\cos^2\phi\left(\dot{\lambda}+\Omega\right)\frac{\partial\lambda}{\partial\lambda_0}\right]\right.$$
$$\left.+\delta\phi_0\frac{\partial}{\partial\tau}\left[\cos^2\phi\left(\dot{\lambda}+\Omega\right)\frac{\partial\lambda}{\partial\phi_0}\right]\right\}\sigma_0\cos\phi_0 d\lambda_0 d\phi_0 d\theta_0 = 0. \qquad (4.100)$$

This equation is equivalent to (4.35) for the unapproximated case. Using (4.31), integrating by parts with respect to λ_0 and ϕ_0, and recalling that $\delta\psi$ is arbitrary, we obtain

$$\frac{\partial}{\partial\tau}\left[\frac{\partial\left(\cos^2\phi\left(\dot{\lambda}+\Omega\right),\lambda,\theta\right)}{\partial\left(\lambda_0,\phi_0,\theta_0\right)}\right] = 0, \qquad (4.101)$$

which can also be written as

$$\frac{\partial}{\partial \tau}\left[\frac{\partial\left(\cos^2\phi\left(\dot{\lambda}+\Omega\right),\lambda,\theta\right)}{\partial\left(\lambda,\sin\phi,\theta\right)}\frac{\partial\left(\lambda,\sin\phi,\theta\right)}{\partial\left(\lambda_0,\sin\phi_0,\theta_0\right)}\right]=0. \tag{4.102}$$

Evaluating the first Jacobian and using (4.12) for the latter one we obtain

$$\frac{\partial}{\partial \tau}\left\{\frac{\sigma_0}{\sigma}\left[2\Omega\sin\phi-\frac{\partial(u\cos\phi)}{a\cos\phi\partial\phi}\right]\right\}=0, \tag{4.103}$$

which is a statement of the conservation of potential vorticity within the framework of the long wave approximation. Thus, we define the potential vorticity as

$$P_l=\frac{1}{\sigma}\left[2\Omega\sin\phi-\frac{\partial(u\cos\phi)}{a\cos\phi\partial\phi}\right]. \tag{4.104}$$

The discussion above demonstrates how naturally the definition of potential vorticity arises within the Hamiltonian framework.

Chapter 5

THE CHARNEY-STERN THEOREM

Semigeostrophic theory has proven to be particularly fruitful in the study of baroclinic wave processes, especially in the nonlinear development where the relative vorticity reaches large values and where static stability is highly variable (Hoskins, 1976). These are the conditions under which quasi-geostrophic theory breaks down; however one would hope for more physical insight than primitive equation results often provide. Here we will consider the necessary conditions for instability in a horizontally sheared baroclinic zonal current, the Charney-Stern theorem, for the system of equations derived in chapters 2 and 3, semigeostrophic theory on the β-plane and on the sphere, respectively. In fact the Charney-Stern theorem seems to be such a fundamental statement of balanced fluid flow that one would expect it to be valid for any consistent balanced theory. We will only consider linear disturbances; however the theory allows for time-integrations into the nonlinear regime.

The theorem was originally proved by Charney and Stern (1962) for linear disturbances on a three dimensional quasi-geostrophic flow with somewhat restricted vertical boundary conditions that were later generalized by Pedlosky (1964) and Bretherton (1966). Charney and Stern's approach was based on that of Rayleigh (1880) in that they assumed an exponential disturbance time dependence and integrated over the volume of the fluid to arrive at the necessary conditions for instability. Bretherton (1966) and Eliassen (1983) have shown how the requirement for exponential time dependence may be relaxed by considering the Lagrangian concept of particle displacements about the mean flow, originally considered by Taylor (1915). Hoskins (1976) proved the theorem for semigeostrophic theory with a constant Coriolis parameter using Charney and Stern's approach. Eliassen

(1983) proved the theorem for a linear set of equations in isentropic coordinates on tl β-plane using the geostrophic momentum approximation. However, he introduced the a proximation into the linearized equations. Thus his set of equations cannot be integrat into the nonlinear regime.

Here, we use Eliassen's approach to prove the Charney-Stern theorem for our sen geostrophic equations in transformed (geostrophic) space; on the β-plane in section 6 and on the sphere in section 6.2. The argument is identical in both cases. First, we u the linearized geostrophic potential pseudodensity equation (σ^*, where the subscript g h been dropped) to derive an equation for the geostrophic meridional eddy flux of σ^* in terr of the time derivative of an expression involving the meridional gradient of the zonal me σ^*/f and the mean of the geostrophic particle displacements squared. Secondly, we u the linearized invertibility relation to write the geostrophic eddy flux of σ^* in terms of tl divergence of a geostrophic Eliassen-Palm (E-P) flux. Combining the two expressions, v obtain the conservative (adiabatic and frictionless) form of an equation that Andrews ar McIntyre (1976, 1978), McIntyre (1980) and Andrews (1983) have termed the generaliz Eliassen-Palm relation. This equation in its general form can be written as follows

$$\frac{\partial A}{\partial t} + \nabla \cdot \mathbf{F} = D, \qquad (5.$$

where $\mathbf{F}$ is the E-P flux, A is the density of E-P wave activity and D is zero for conservati motion. Edmon et al. (1980) and McIntyre (1980) have discussed this equation with quasi-geostrophic theory. Andrews (1983) has discussed it for the primitive equatio on a β-plane in isentropic coordinates. Integrating our conservative form of (5.1) ov the volume of the fluid, we obtain that the time derivative of the volume integral of vanishes, since the boundary fluxes of the E-P flux vanish. Thus, for a growing disturban the meridional gradient of $\bar{\sigma}^*/f$ must have both signs, which proves the Charney-Ste theorem.

5.1 The Charney-Stern theorem generalized to semigeostrophic theory on the β-plane

Consider a zonal current on a cyclic β-plane in the absence of friction and diabatic effects. Using overbars to denote the dependent variables for the basic current, we have

$$\frac{\partial \bar{M}}{\partial y} = -f(Y)\bar{u}_g, \qquad \frac{\partial \bar{M}}{\partial \theta} = \Pi(\bar{p}), \tag{5.2}$$

or in geostrophic space

$$\frac{\partial \bar{M}^*}{\partial Y} = -f(Y)\bar{u}_g + \frac{\bar{u}_g^2}{f(Y)}\beta, \qquad \frac{\partial \bar{M}^*}{\partial \Theta} = \Pi(\bar{p}). \tag{5.3}$$

The semigeostrophic potential pseudodensity equation on the β-plane can be written as

$$\frac{\partial \sigma^*}{\partial T} + \frac{\partial}{\partial X}\left(-\frac{\sigma^*}{f(Y)}\frac{\partial M^*}{\partial Y}\right) + \frac{\partial}{\partial Y}\left(\frac{\sigma^*}{f(Y)}\frac{\partial M^*}{\partial X}\right) = 0, \tag{5.4}$$

where the subscript g has been dropped. When linearized about the above basic state (5.4) becomes

$$\left(\frac{\partial}{\partial T} - \frac{1}{f(Y)}\frac{\partial \bar{M}^*}{\partial Y}\frac{\partial}{\partial X}\right)\sigma^{*\prime} + \frac{\partial M^{*\prime}}{\partial X}\frac{\partial}{\partial Y}\left(\frac{\bar{\sigma}^*}{f(Y)}\right) = 0, \tag{5.5}$$

where the prime indicates a deviation from the basic state and

$$\frac{\partial M^{*\prime}}{\partial X} = f(Y)v_g'. \tag{5.6}$$

Let us introduce the northward geostrophic particle displacement η', defined by

$$v_g' = \left(\frac{\partial}{\partial T} - \frac{1}{f(Y)}\frac{\partial \bar{M}^*}{\partial Y}\frac{\partial}{\partial X}\right)\eta'. \tag{5.7}$$

Using (5.6) and (5.7), we can write (5.5) as follows

$$\left(\frac{\partial}{\partial T} - \frac{1}{f(Y)}\frac{\partial \bar{M}^*}{\partial Y}\frac{\partial}{\partial X}\right)\sigma^{*\prime} + f(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\left(\frac{\partial}{\partial T} - \frac{1}{f(Y)}\frac{\partial \bar{M}^*}{\partial Y}\frac{\partial}{\partial X}\right)\eta' = 0, \tag{5.8}$$

which can be integrated to obtain

$$\sigma^{*\prime} + f(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\eta' = 0. \tag{5.9}$$

Multiplying (5.9) by $f(Y)v_g'$ and taking the zonal average at fixed Y, we obtain by using (5.7)

$$f(Y)\overline{v_g'\sigma^{*\prime}} + f^2(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\overline{\eta'\left(\frac{\partial}{\partial T} - \frac{1}{f(Y)}\frac{\partial \bar{M}^*}{\partial Y}\frac{\partial}{\partial X}\right)\eta'} = 0, \tag{5.10}$$

which can be written as

$$\frac{\partial}{\partial T}\left(f^2(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\frac{1}{2}\overline{\eta'^2}\right)+f(Y)\overline{v_g'\sigma^{*\prime}}=0. \tag{5.11}$$

The linearized invertibility relation will allow us to write $f(Y)\overline{v_g'\sigma^{*\prime}}$, which is the geostrophic meridional eddy flux of potential pseudodensity multiplied by the Coriolis parameter, in terms of the divergence of a geostrophic E-P flux. Thus, when we integrate (5.11) over the entire domain the second term vanishes since the boundary fluxes vanish, and for a growing disturbance $\partial\left(f^{-1}(Y)\bar{\sigma}^*\right)/\partial Y$ must have both signs. This proof of the Charney-Stern theorem follows Eliassen's (1983) generalization of it and does not make any assumptions as to the spatial or temporal structure of the disturbance. The definition (5.6) of the Lagrangian field η' makes the description of the wave disturbance simpler in that it allows one to avoid assumptions on its exact nature. This is in agreement with McIntyre (1980) who states: "It has become evident in recent years that the underlying theoretical structure of the subject" (namely wave mean flow interactions) "becomes immeasurably clearer if one describes wave disturbances in terms of *particle displacements* about the mean flow, in place of the more usual eddy velocity fields".

The invertibility relation can be written as

$$\frac{\partial(x,y,p)}{\partial(X,Y,\Theta)}+\sigma^*=0, \tag{5.12}$$

which is in a form slightly different from (2.46), but this is the form that is convenient to linearize. We can separate the dependent variables into basic state ones and ones that represent deviations from the basic state. We have $\bar{x}=X$ since $\bar{v}_g=0$,

$$x'=-\frac{v_g'}{f(Y)} \qquad \text{and} \qquad y'=\frac{u_g'}{f(Y)}. \tag{5.13}$$

Starting the linearization, we obtain

$$\frac{\partial(\bar{y}+y',\bar{p}+p')}{\partial(Y,\Theta)}+\frac{\partial(x',\bar{y},\bar{p})}{\partial(X,Y,\Theta)}+\bar{\sigma}^*+\sigma^{*\prime}=0. \tag{5.14}$$

Noting that

$$\bar{\sigma}^*=-\frac{\partial(\bar{y},\bar{p})}{\partial(Y,\Theta)}, \tag{5.15}$$

allows us to write (5.14) as

$$\sigma^{*\prime} = \bar{\sigma}^* \frac{\partial x'}{\partial X} - \frac{\partial(y', \bar{p})}{\partial(Y, \Theta)} - \frac{\partial(\bar{y}, p')}{\partial(Y, \Theta)}, \tag{5.16}$$

which is the linearized invertibility relation. Let us multiply (5.16) by $f(Y)v_g'$ and take the zonal average using (5.13), to get

$$f(Y)\overline{v_g'\sigma^{*\prime}} = \overline{f(Y)v_g'\left(\frac{\partial \bar{p}}{\partial Y}\frac{\partial y'}{\partial \Theta} - \frac{\partial \bar{p}}{\partial \Theta}\frac{\partial y'}{\partial Y}\right)} + \overline{f(Y)v_g'\left(\frac{\partial \bar{y}}{\partial \Theta}\frac{\partial p'}{\partial Y} - \frac{\partial \bar{y}}{\partial Y}\frac{\partial p'}{\partial \Theta}\right)}. \tag{5.17}$$

Regrouping terms, we can also write (5.17) as

$$\begin{aligned} f(Y)\overline{v_g'\sigma^{*\prime}} = & \frac{\partial}{\partial Y}\left(\frac{\partial \bar{y}}{\partial \Theta} f(Y)\overline{v_g'p'} - \frac{\partial \bar{p}}{\partial \Theta} f(Y)\overline{v_g'y'}\right) + \frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{p}}{\partial Y} f(Y)\overline{v_g'y'} - \frac{\partial \bar{y}}{\partial Y} f(Y)\overline{v_g'p'}\right) \\ & + \overline{y'\frac{\partial}{\partial Y}\left(\frac{\partial \bar{p}}{\partial \Theta} f(Y)v_g'\right)} - \overline{p'\frac{\partial}{\partial Y}\left(\frac{\partial \bar{y}}{\partial \Theta} f(Y)v_g'\right)} + \overline{p'\frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{y}}{\partial Y} f(Y)v_g'\right)} - \overline{y'\frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{p}}{\partial Y} f(Y)v_g'\right)}. \end{aligned} \tag{5.18}$$

Using the zonally averaged thermal wind equation we can show that the four terms in the last line of (5.18) vanish. First, we use (5.6) to write

$$\begin{aligned} \mathcal{G} \equiv & -\overline{y'\frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{p}}{\partial Y} f(Y)v_g'\right)} + \overline{y'\frac{\partial}{\partial Y}\left(\frac{\partial \bar{p}}{\partial \Theta} f(Y)v_g'\right)} \\ & -\overline{p'\frac{\partial}{\partial Y}\left(\frac{\partial \bar{y}}{\partial \Theta} f(Y)v_g'\right)} + \overline{p'\frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{y}}{\partial Y} f(Y)v_g'\right)} \\ = & -\overline{y'\left[\frac{\partial \bar{p}}{\partial Y}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial \Theta}\right) - \frac{\partial \bar{p}}{\partial \Theta}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial Y}\right)\right]} \\ & -\overline{p'\left[\frac{\partial \bar{y}}{\partial \Theta}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial Y}\right) - \frac{\partial \bar{y}}{\partial Y}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial \Theta}\right)\right]} \\ = & -\overline{y'\frac{\partial \bar{p}}{\partial Y}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial \Theta}\right)} - \overline{p'\frac{\partial \bar{y}}{\partial \Theta}\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial Y}\right)}, \end{aligned} \tag{5.19}$$

where in the last line we have used

$$\bar{\Gamma}p' = \kappa\frac{\partial M^{*\prime}}{\partial \Theta}, \tag{5.20}$$

and

$$-\left(f^2(Y) - 2\beta\bar{u}_g\right)y' = \frac{\partial M^{*\prime}}{\partial Y}, \tag{5.21}$$

which are easily derived from the linearized hydrostatic and geostrophic relations. Secondly, the zonally averaged thermal wind equation can be derived by noting that

$$\bar{y} = Y + \frac{\bar{u}_g}{f(Y)}, \tag{5.22}$$

taking the Θ derivative and using (5.3) to get

$$-\left(f^2(Y) - 2\beta\bar{u}_g\right)\frac{\partial \bar{y}}{\partial \Theta} = \frac{\bar{\Gamma}}{\kappa}\frac{\partial \bar{p}}{\partial Y}. \tag{5.23}$$

Using (5.23) in (5.19), we are left with

$$\mathcal{G} = \left(f^2(Y) - 2\beta\bar{u}_g\right)\frac{\kappa}{\bar{\Gamma}}\frac{\partial \bar{y}}{\partial \Theta}\overline{y'\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial \Theta}\right)} - \frac{\partial \bar{y}}{\partial \Theta}\overline{p'\frac{\partial}{\partial X}\left(\frac{\partial M^{*\prime}}{\partial Y}\right)}. \tag{5.24}$$

Using (5.20) and (5.21), we obtain

$$\begin{aligned}\mathcal{G} &= \left(f^2(Y) - 2\beta\bar{u}_g\right)\frac{\partial \bar{y}}{\partial \Theta}\left(\overline{y'\frac{\partial p'}{\partial X} + p'\frac{\partial y'}{\partial X}}\right)\\ &= \left(f^2(Y) - 2\beta\bar{u}_g\right)\frac{\partial \bar{y}}{\partial \Theta}\overline{\frac{\partial}{\partial X}(y'p')} = 0,\end{aligned} \tag{5.25}$$

and (5.18) can be written as

$$f(Y)\overline{v_g'\sigma^{*\prime}} = \frac{\partial}{\partial Y}\left(\frac{\partial \bar{y}}{\partial \Theta}f(Y)\overline{v_g'p'} - \frac{\partial \bar{p}}{\partial \Theta}f(Y)\overline{v_g'y'}\right) + \frac{\partial}{\partial \Theta}\left(\frac{\partial \bar{p}}{\partial Y}f(Y)\overline{v_g'y'} - \frac{\partial \bar{y}}{\partial Y}f(Y)\overline{v_g'p'}\right). \tag{5.26}$$

We can now define a geostrophic E-P flux $\mathbf{F}$ as

$$\mathbf{F} = \left(\frac{\partial \bar{y}}{\partial \Theta}f(Y)\overline{v_g'p'} - \frac{\partial \bar{p}}{\partial \Theta}\overline{v_g'u_g'}\,,\ \frac{\partial \bar{p}}{\partial Y}\overline{v_g'u_g'} - \frac{\partial \bar{y}}{\partial Y}f(Y)\overline{v_g'p'}\right), \tag{5.27}$$

where we have used (5.13) to write y' in terms of u_g'. Combining (5.11) and (5.26) using (5.27), we obtain the adiabatic and frictionless form of the generalized geostrophic Eliassen-Palm relation (5.1),

$$\frac{\partial}{\partial T}\left(f^2(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\frac{1}{2}\overline{\eta'^2}\right) + \nabla\cdot\mathbf{F} = 0, \tag{5.28}$$

where ∇ indicates the del operator in the meridional plane. Let us integrate over the (Y, Θ) plane. As before (chapter 2) we shall assume that the meridional boundary conditions for our β-plane are that v_g' vanishes at both the northern and southern boundaries, thus the meridional boundary fluxes vanish. The boundary flux at the top vanishes since according to (2.48b) both $\partial\bar{p}/\partial Y = 0$ and $p' = 0$ at the top. We shall now apply the lower boundary condition, (2.48c). For the basic state flow we have $\Theta_B\bar{\Pi} - \bar{M} = 0$. Differentiating with respect to Y, we obtain

$$\frac{\partial \bar{M}^*}{\partial Y} - \bar{u}_g\frac{\partial \bar{u}_g}{\partial Y} - \Theta_B\frac{\bar{\Gamma}}{\kappa}\frac{\partial \bar{p}}{\partial Y} = 0. \tag{5.29}$$

Differentiating (5.22) with respect to Y and using it in (5.29), we get

$$\frac{\partial \bar{M}^*}{\partial Y} + f(Y)\bar{u}_g - f(Y)\bar{u}_g\frac{\partial \bar{y}}{\partial Y} - \beta\frac{\bar{u}_g^2}{f(Y)} - \Theta_B\frac{\bar{\Gamma}}{\kappa}\frac{\partial \bar{p}}{\partial Y} = 0,$$

or, by using the geostrophic relation for the basic flow (5.3), we obtain

$$f(Y)\bar{u}_g\frac{\partial \bar{y}}{\partial Y} + \Theta\frac{\bar{\Gamma}}{\kappa}\frac{\partial \bar{p}}{\partial Y} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.30}$$

For the perturbation flow on the lower boundary, we have $M^{*\prime} - \bar{u}_g u_g' - \Theta_B\Pi' = 0$. Multiplying this expression by v_g', taking the zonal average and using (5.6), we obtain

$$\bar{u}_g\overline{v_g' u_g'} + \Theta\frac{\bar{\Gamma}}{\kappa}\overline{v_g' p'} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.31}$$

Combining (5.30) and (5.31) , we conclude that

$$\frac{\partial \bar{p}}{\partial Y}\overline{v_g' u_g'} - \frac{\partial \bar{y}}{\partial Y}f(Y)\overline{v_g' p'} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.32}$$

Thus, the lower boundary flux also vanishes and the result of integrating (5.28) over the entire (Y, Θ) plane is

$$\frac{\partial}{\partial T}\iint f^2(Y)\left(\frac{\partial}{\partial Y}\frac{\bar{\sigma}^*}{f(Y)}\right)\frac{1}{2}\overline{\eta'^2}\,dY\,d\Theta = 0. \tag{5.33}$$

Now, $f^2(Y)$ is always positive and for a growing disturbance $\overline{\eta'^2}$ is positive thus requiring that the meridional derivative of the inverse potential vorticity takes on both signs within the fluid. This last requirement is equivalent to demanding that the meridional derivative of potential vorticity has both signs within the fluid which is exactly the Charney-Stern theorem of necessary conditions for instability.

5.2 The Charney-Stern theorem generalized to semigeostrophic theory on the sphere

The derivation of the spherical case proceeds in a manner identical to the one on the β-plane. Again we consider a zonal current in the absence of friction and diabatic effects. Using overbars to denote the dependent variables for the basic current, we have

$$\frac{\partial \bar{M}}{a\partial \phi} = -\frac{\cos\phi}{\cos\Phi}2\Omega\sin\Phi\,\bar{u}_g, \qquad \frac{\partial \bar{M}}{\partial \theta} = \Pi(\bar{p}), \tag{5.34}$$

or in geostrophic space

$$\frac{\partial \bar{M}^*}{a\partial\Phi} = -2\Omega\sin\Phi\,\bar{u}_g + \frac{\bar{u}_g^2}{a\sin\Phi\cos\Phi}, \qquad \frac{\partial \bar{M}^*}{\partial\Theta} = \Pi(\bar{p}). \tag{5.35}$$

The semigeostrophic potential pseudodensity equation on the sphere can be written as

$$\frac{\partial\sigma^*}{\partial T} + \frac{\partial}{a\partial\Lambda}\left(-\frac{\sigma^*}{2\Omega\sin\Phi}\frac{\partial M^*}{a\partial\sin\Phi}\right) + \frac{\partial}{a\partial\sin\Phi}\left(\frac{\sigma^*}{2\Omega\sin\Phi}\frac{\partial M^*}{a\partial\Lambda}\right) = 0, \tag{5.36}$$

where the subscript g has been dropped. When linearized about the above basic state (5.36) becomes

$$\left(\frac{\partial}{\partial T} - \frac{1}{2\Omega\sin\Phi}\frac{\partial\bar{M}^*}{a\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\right)\sigma^{*\prime} + \frac{\partial M^{*\prime}}{a\partial\Lambda}\frac{\partial}{a\partial\sin\Phi}\left(\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right) = 0, \tag{5.37}$$

where the prime indicates a deviation from the basic state and

$$\frac{\partial M^{*\prime}}{a\cos\Phi\partial\Lambda} = 2\Omega\sin\Phi\, v_g'. \tag{5.38}$$

We introduce the northward geostrophic particle displacement η', defined by

$$v_g' = \left(\frac{\partial}{\partial T} - \frac{1}{2\Omega\sin\Phi}\frac{\partial\bar{M}^*}{a\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\right)\eta'. \tag{5.39}$$

Using (5.38) and (5.39), we can write (5.37) as follows

$$\left(\frac{\partial}{\partial T} - \frac{1}{2\Omega\sin\Phi}\frac{\partial\bar{M}^*}{a\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\right)\sigma^{*\prime}$$

$$+2\Omega\sin\Phi\cos\Phi\left(\frac{\partial}{a\partial\sin\Phi}\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\left(\frac{\partial}{\partial T} - \frac{1}{2\Omega\sin\Phi}\frac{\partial\bar{M}^*}{a\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\right)\eta' = 0, \tag{5.40}$$

which can be integrated to obtain

$$\sigma^{*\prime} + 2\Omega\sin\Phi\cos\Phi\left(\frac{\partial}{a\partial\sin\Phi}\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\eta' = 0. \tag{5.41}$$

In what follows it is convenient to introduce the function $\gamma(\Phi)$, where

$$\gamma(\Phi) \equiv 2\Omega\sin\Phi\,\cos\Phi. \tag{5.42}$$

Multiplying (5.41) by $\gamma(\Phi)v_g'$ and taking the zonal average at fixed Φ, we obtain by using (5.39)

$$\gamma(\Phi)\,\overline{v_g'\sigma^{*\prime}} + \gamma^2(\Phi)\left(\frac{\partial}{a\partial\sin\Phi}\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\overline{\eta'\left(\frac{\partial}{\partial T} - \frac{1}{2\Omega\sin\Phi}\frac{\partial\bar{M}^*}{a\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\right)\eta'} = 0, \tag{5.43}$$

which can be written as

$$\frac{\partial}{\partial T}\left(\gamma^2(\Phi)\left(\frac{\partial}{a\partial\sin\Phi}\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\frac{1}{2}\overline{\eta'^2}\right)+\gamma(\Phi)\,\overline{v_g'\sigma^{*\prime}}=0. \tag{5.44}$$

The linearized invertibility relation will allow us to write $\gamma(\Phi)\,\overline{v_g'\sigma^{*\prime}}$, which is the geostrophic meridional eddy flux of potential pseudodensity multiplied by $\gamma(\Phi)$, in terms of the divergence of a geostrophic Eliassen-Palm flux. Thus, when we integrate (5.44) over the entire domain the second term vanishes since the boundary fluxes vanish, and for a growing disturbance $\partial\left((2\Omega\sin\Phi)^{-1}\bar{\sigma}^*\right)/a\partial\sin\Phi$ must have both signs.

The invertibility relation can be written as follows

$$\frac{\partial(\lambda,\sin\phi,p)}{\partial(\Lambda,\sin\Phi,\Theta)}+\sigma^*=0. \tag{5.45}$$

We can separate the dependent variables into basic state variables and variables that represent deviations from the basic state. We have $\bar{\lambda}=\Lambda$ since $\bar{v}_g=0$,

$$\lambda'=-\frac{v_g'}{2\Omega\sin\Phi\, a\cos\Phi}=-\frac{v_g'}{a\gamma(\Phi)}\qquad\text{and}\qquad \sin\phi'=\frac{u_g'\cos\Phi}{a2\Omega\sin\Phi}. \tag{5.46}$$

Starting the linearization, we obtain

$$\frac{\partial(\overline{\sin\phi}+\sin\phi',\bar{p}+p')}{\partial(\sin\Phi,\Theta)}+\frac{\partial(\lambda',\overline{\sin\phi},\bar{p})}{\partial(\Lambda,\sin\Phi,\Theta)}+\bar{\sigma}^*+\sigma^{*\prime}=0. \tag{5.47}$$

Noting that

$$\bar{\sigma}^*=-\frac{\partial(\overline{\sin\phi},\bar{p})}{\partial(\sin\Phi,\Theta)}, \tag{5.48}$$

allows us to write (5.47) as

$$\sigma^{*\prime}=\bar{\sigma}^*\frac{\partial\lambda'}{\partial\Lambda}-\frac{\partial(\sin\phi',\bar{p})}{\partial(\sin\Phi,\Theta)}-\frac{\partial(\overline{\sin\phi},p')}{\partial(\sin\Phi,\Theta)}, \tag{5.49}$$

which is the linearized invertibility relation. Let us multiply (5.49) by $\gamma(\Phi)v_g'$ and take the zonal average using (5.46), to get

$$\gamma(\Phi)\,\overline{v_g'\sigma^{*\prime}}=\gamma(\Phi)\,\overline{v_g'\left(\frac{\partial\bar{p}}{\partial\sin\Phi}\frac{\partial\sin\phi'}{\partial\Theta}-\frac{\partial\bar{p}}{\partial\Theta}\frac{\partial\sin\phi'}{\partial\sin\Phi}\right)}$$

$$+\gamma(\Phi)\,\overline{v_g'\left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\frac{\partial p'}{\partial\sin\Phi}-\frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\frac{\partial p'}{\partial\Theta}\right)}. \tag{5.50}$$

Regrouping terms, we can also write (5.50) as

$$\gamma(\Phi)\,\overline{v_g'\sigma^{*\prime}} = \frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\gamma(\Phi)\,\overline{v_g'p'} - \frac{\partial\bar{p}}{\partial\Theta}\gamma(\Phi)\,\overline{v_g'\sin\phi'}\right)$$
$$+\frac{\partial}{\partial\Theta}\left(\frac{\partial\bar{p}}{\partial\sin\Phi}\gamma(\Phi)\,\overline{v_g'\sin\phi'} - \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)\,\overline{v_g'p'}\right)$$
$$+\overline{\sin\phi'\frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\bar{p}}{\partial\Theta}\gamma(\Phi)v_g'\right)} - \overline{p'\frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\gamma(\Phi)v_g'\right)}$$
$$+\overline{p'\frac{\partial}{\partial\Theta}\left(\frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)v_g'\right)} - \overline{\sin\phi'\frac{\partial}{\partial\Theta}\left(\frac{\partial\bar{p}}{\partial\sin\Phi}\gamma(\Phi)v_g'\right)}. \tag{5.51}$$

Using the zonally averaged thermal wind equation we can show that the four terms, which we give the symbol $\mathcal{H}$, in the last two lines of (5.51) vanish. First, we use (5.38) to write

$$\mathcal{H} \equiv -\overline{\sin\phi'\frac{\partial}{\partial\Theta}\left(\frac{\partial\bar{p}}{\partial\sin\Phi}\gamma(\Phi)v_g'\right)} + \overline{\sin\phi'\frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\bar{p}}{\partial\Theta}\gamma(\Phi)v_g'\right)}$$
$$-\overline{p'\frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\gamma(\Phi)v_g'\right)} + \overline{p'\frac{\partial}{\partial\Theta}\left(\frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)v_g'\right)}$$
$$= -\overline{\sin\phi'\left[\frac{\partial\bar{p}}{\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\Theta}\right) - \frac{\partial\bar{p}}{\partial\Theta}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\sin\Phi}\right)\right]}$$
$$-\overline{p'\left[\frac{\partial\overline{\sin\phi}}{\partial\Theta}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\sin\Phi}\right) - \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\Theta}\right)\right]}$$
$$= -\overline{\sin\phi'\frac{\partial\bar{p}}{\partial\sin\Phi}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\Theta}\right)} - \overline{p'\frac{\partial\overline{\sin\phi}}{\partial\Theta}\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\sin\Phi}\right)}, \tag{5.52}$$

where in the last line we have used

$$\bar{\Gamma}p' = \kappa\frac{\partial M^{*\prime}}{\partial\Theta}, \tag{5.53}$$

and

$$-\left((a2\Omega\tan\Phi)^2 - 2\frac{a2\Omega}{\cos^3\Phi}\bar{u}_g\right)\sin\phi' = \frac{\partial M^{*\prime}}{\partial\sin\Phi}, \tag{5.54}$$

which are easily derived from the linearized hydrostatic and geostrophic relations. Secondly, the zonally averaged thermal wind equation can be derived by noting that

$$\overline{\sin\phi} = \sin\Phi + \frac{\bar{u}_g\cos\Phi}{a2\Omega\sin\Phi}, \tag{5.55}$$

taking the Θ derivative and using (5.35) to get

$$-\left((a2\Omega\tan\Phi)^2 - 2\frac{a2\Omega}{\cos^3\Phi}\bar{u}_g\right)\frac{\partial\overline{\sin\phi}}{\partial\Theta} = \frac{\bar{\Gamma}}{\kappa}\frac{\partial\bar{p}}{\partial\sin\Phi}. \tag{5.56}$$

Using (5.56) in (5.52), we are left with

$$\mathcal{H} = \left((a2\Omega\tan\Phi)^2 - 2\frac{a2\Omega}{\cos^3\Phi}\bar{u}_g\right)\frac{\kappa}{\bar{\Gamma}}\frac{\partial\overline{\sin\phi}}{\partial\Theta}\overline{\sin\phi'\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\Theta}\right)}$$

$$-\frac{\partial\overline{\sin\phi}}{\partial\Theta}\overline{p'\frac{\partial}{a\partial\Lambda}\left(\frac{\partial M^{*\prime}}{\partial\sin\Phi}\right)}. \tag{5.57}$$

Using (5.53) and (5.54), we obtain

$$\mathcal{H} = \left((a2\Omega\tan\Phi)^2 - 2\frac{a2\Omega}{\cos^3\Phi}\bar{u}_g\right)\frac{\partial\overline{\sin\phi}}{\partial\Theta}\left(\overline{\sin\phi'\frac{\partial p'}{a\partial\Lambda} + p'\frac{\partial\sin\phi'}{a\partial\Lambda}}\right)$$

$$= \left((a2\Omega\tan\Phi)^2 - 2\frac{a2\Omega}{\cos^3\Phi}\bar{u}_g\right)\frac{\partial\overline{\sin\phi}}{\partial\Theta}\overline{\frac{\partial}{a\partial\Lambda}(\sin\phi' p')} = 0, \tag{5.58}$$

and (5.51) can be written as

$$\gamma(\Phi)\overline{v_g'\sigma^{*\prime}} = \frac{\partial}{\partial\sin\Phi}\left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\gamma(\Phi)\overline{v_g'p'} - \frac{\partial\bar{p}}{\partial\Theta}\gamma(\Phi)\overline{v_g'\sin\phi'}\right)$$

$$+\frac{\partial}{\partial\Theta}\left(\frac{\partial\bar{p}}{\partial\sin\Phi}\gamma(\Phi)\overline{v_g'\sin\phi'} - \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)\overline{v_g'p'}\right). \tag{5.59}$$

We define a geostrophic E-P flux on the sphere $\mathbf{G}$ as follows

$$\mathbf{G} = \left(\frac{\partial\overline{\sin\phi}}{\partial\Theta}\gamma(\Phi)\overline{v_g'p'} - \frac{\partial\bar{p}}{\partial\Theta}\frac{\cos^2\Phi}{a}\overline{v_g'u_g'},\ \frac{\partial\bar{p}}{\partial\sin\Phi}\frac{\cos^2\Phi}{a}\overline{v_g'u_g'} - \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)\overline{v_g'p'}\right), \tag{5.60}$$

where we have used (5.46) to write $\sin\phi'$ in terms of u_g'. Combining (5.44) and (5.59) using (5.60) gives us the conservative generalized geostrophic Eliassen-Palm relation on the sphere,

$$\frac{\partial}{\partial T}\left(\gamma^2(\Phi)\frac{\partial}{a\partial\sin\Phi}\left(\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\frac{1}{2}\overline{\eta'^2}\right) + \nabla\cdot\mathbf{G} = 0, \tag{5.61}$$

where ∇ indicates the del operator in the meridional plane. We now integrate (5.61) over the meridional plane, from pole to pole. The boundary flux at the top vanishes since according to (3.54b) both $\partial\bar{p}/\partial\sin\Phi = 0$ and $p' = 0$ at the top. We shall now apply the lower boundary condition, (3.54c). For the basic state flow we have $\Theta_B\bar{\Pi} - \bar{M} = 0$. Differentiating with respect to $\sin\Phi$, we obtain

$$\frac{\partial\bar{M}^*}{\partial\sin\Phi} - \bar{u}_g\frac{\partial\bar{u}_g}{\partial\sin\Phi} - \Theta_B\frac{\bar{\Gamma}}{\kappa}\frac{\partial\bar{p}}{\partial\sin\Phi} = 0. \tag{5.62}$$

Differentiating (5.55) with respect to $\sin\Phi$ and using it in (5.62), we get

$$\frac{\partial \bar{M}^*}{\partial \sin\Phi} + a2\Omega \tan\Phi\, \bar{u}_g - a2\Omega\tan\Phi\, \bar{u}_g \frac{\partial \overline{\sin\phi}}{\partial \sin\Phi} - \frac{\bar{u}_g^2}{\sin\Phi\cos^2\Phi} - \Theta_B \frac{\bar{\Gamma}}{\kappa}\frac{\partial \bar{p}}{\partial \sin\Phi} = 0,$$

or, by using the geostrophic relation for the basic flow (5.35), we obtain

$$a2\Omega\tan\Phi\,\bar{u}_g \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi} + \Theta\frac{\bar{\Gamma}}{\kappa}\frac{\partial\bar{p}}{\partial\sin\Phi} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.63}$$

For the perturbation flow on the lower boundary, we have $M^{*\prime} - \bar{u}_g u_g' - \Theta_B \Pi' = 0$. Multiplying this by v_g', taking the zonal average and using (5.38), we obtain

$$\bar{u}_g \overline{v_g' u_g'} + \Theta\frac{\bar{\Gamma}}{\kappa}\overline{v_g' p'} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.64}$$

Combining (5.63) and (5.64), we conclude that

$$\frac{\partial\bar{p}}{\partial\sin\Phi}\frac{\cos^2\Phi}{a}\overline{v_g' u_g'} - \frac{\partial\overline{\sin\phi}}{\partial\sin\Phi}\gamma(\Phi)\overline{v_g' p'} = 0 \qquad \text{at} \qquad \Theta = \Theta_B. \tag{5.65}$$

Thus, the lower boundary flux also vanishes and the result of integrating (5.61) over the entire $(\sin\Phi, \Theta)$ plane is

$$\frac{\partial}{\partial T}\iint \gamma^2(\Phi)\frac{\partial}{a\partial\sin\Phi}\left(\frac{\bar{\sigma}^*}{2\Omega\sin\Phi}\right)\frac{1}{2}\overline{\eta'^2}\cos\Phi\, d\Phi d\Theta = 0. \tag{5.66}$$

In order for the disturbance to grow, $\overline{\eta'^2}$ must be positive and $\partial\left((2\Omega\sin\Phi)^{-1}\bar{\sigma}^*\right)/\partial\sin\Phi$ must have both signs, i.e., the meridional derivative of the inverse potential vorticity must have both signs which proves the Charney-Stern theorem.

Chapter 6

TOWARDS A GLOBALLY VALID BALANCED THEORY

As was discussed in chapter 2, section 2.2.4 for the β-plane case and in chapter 3, section 3.2.4 for the spherical case, semigeostrophic theory breaks down on the equator. The equator represents a singular point in the meridional structure equation; a point where v_g becomes infinite. When considering the two dimensional zonally symmetric case, v_g is identically zero and one gets around the problem of the singular point. However, there can be no flow across the equator. Thus it is of very limited interest to study flows close to the equator and semigeostrophic theory should be viewed as basically a midlatitude theory.

In this chapter zonally symmetric balanced flow in the equatorial region will be studied. Specifically we are interested in studying how the distribution of potential vorticity on isentropic surfaces changes with time as a tropical heat source is allowed to act. Under undisturbed conditions the potential vorticity increases monotonically northwards, from negative values south of the equator to positive values north of it and isolines of constant potential vorticity are straight and vertical. Deep convection induces a positive potential vorticity anomaly at low levels and a negative anomaly aloft. The following question arises. Are the requirements of the Charney-Stern theorem ever fulfilled, i.e., does a reversal of the potential vorticity gradient on isentropic surfaces ever develop? This would set the stage for instability and the breakup of the ITCZ into tropical waves. Indeed, McBride and Holland (1989) observed a fairly regular breakup of the ITCZ into individual weather systems during the 1987 monsoon season in Australia. Perhaps this regular breakup of the ITCZ can in part be attributed to its own powers, i.e., the ITCZ carries the seeds of its own destruction.

Again, potential temperature is the vertical coordinate. As a horizontal coordinate we use potential latitude, which was introduced by Hack et al. (1989). The potential latitude can be thought of as an angular momentum coordinate and it allows for cross equatorial flow while having the advantage of being a vortex coordinate. In section 6.1 the advantage of the coordinate transformation will be explored. The development is very similar to the study by Schubert and Alworth (1987) of the tropical cyclone where they use potential temperature and potential radius, which is also an angular momentum coordinate. Similar to the semigeostrophic derivations in chapters 2 and 3 everything reduces to one prognostic equation in σ^*, the potential pseudodensity, and an invertibility principle which allows one to diagnose the field of M^*, the Bernoulli function, and thus derive the wind- and massfields. In section 6.2 friction is neglected and a simple heating function thought to represent the ITCZ is introduced which allows one to solve the prognostic equation analytically. The potential pseudodensity equation was also solved analytically in the tropical cyclone study of Schubert and Alworth (1987) and the semigeostrophic study of a squall line on an f–plane by Schubert et al. (1989). Our simple study indicates that indeed the latent heat release thought to represent the ITCZ produces potential vorticity gradient reversals on the time scale of 2–3 days both at low levels, on the poleward side of the ITCZ, and at upper levels, on the equatorward side of the ITCZ. Thus, the ITCZ can be thought of as a very dynamic phenomenon — one which sets the stage for its own destruction.

Section 6.3 offers some speculation on the direction and framework for future work in deriving a globally valid three dimensional balanced theory. This theory would not be restricted by the limitations of semigeostrophic theory which arise from the geostrophic momentum approximation and its neglect of curvature vorticity compared to shear vorticity and it would be valid anywhere on the earth. Specifically the zonally symmetric balanced theory would be a two dimensional special case of the full three dimensional theory.

6.1 Zonally symmetric theory and the potential latitude/potential temperature coordinate

The equations for zonally symmetric, balanced flow can be written

$$\frac{Du}{Dt} - \left(2\Omega \sin\phi + \frac{u\tan\phi}{a}\right) v = F, \tag{6.1}$$

$$\left(2\Omega \sin\phi + \frac{u\tan\phi}{a}\right) u + \frac{\partial M}{a\partial\phi} = 0, \tag{6.2}$$

$$\frac{\partial M}{\partial\theta} = \Pi, \tag{6.3}$$

$$\frac{D\sigma}{Dt} + \sigma\left(\frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi} + \frac{\partial\dot{\theta}}{\partial\theta}\right) = 0, \tag{6.4}$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + v\frac{\partial}{a\partial\phi} + \dot{\theta}\frac{\partial}{\partial\theta} \tag{6.5}$$

is the total derivative, F represents the effects of friction and all other definitions are the same as in chapter 3. Specifically (6.1)–(6.4) can be obtained from (3.1)–(3.4) using the assumptions that all fields are independent of longitude and that the zonal flow is balanced.

6.1.1 Conservation relations

The zonal momentum equation (6.1) can also be written in the angular momentum form

$$\frac{D}{Dt}(\Omega a\cos^2\phi + u\cos\phi) = F\cos\phi. \tag{6.6}$$

In the absence of friction the absolute angular momentum is conserved. The ITCZ is highly nonconservative, however as we shall see in the next section, transforming to a type of angular momentum coordinate simplifies the dynamics considerably.

The kinetic energy equation can be derived by combining (6.1) and (6.2) to get

$$\frac{DK}{Dt} + v\frac{\partial M}{a\partial\phi} = uF, \tag{6.7}$$

where $K \equiv \frac{1}{2}u^2$ is the kinetic energy per unit mass. Combining (6.7) with the continuity equation (6.4) gives

$$\frac{\partial(\sigma K)}{\partial t} + \frac{\partial(\sigma v\cos\phi K)}{a\cos\phi\partial\phi} + \frac{\partial(\sigma\dot{\theta}K)}{\partial\theta} + \sigma v\frac{\partial M}{a\partial\phi} = \sigma uF. \tag{6.8}$$

Using the continuity and hydrostatic equations the kinetic energy equation can be written

$$\frac{\partial(\sigma K)}{\partial t}+\frac{\partial(\sigma v\cos\phi(K+gz))}{a\cos\phi\partial\phi}+\frac{\partial\left(\sigma\dot{\theta}(K+gz)\right)}{\partial\theta}-\frac{\partial}{\partial\theta}\left(gz\frac{\partial p}{\partial t}\right)=\sigma\left(uF-\alpha\omega\right), \quad (6.9)$$

where $\omega = Dp/Dt$.

For deriving the thermodynamic energy equation, multiply (6.4) by c_pT to obtain

$$\frac{\partial}{\partial t}\left(\sigma c_pT\right)+\frac{\partial\left(\sigma v\cos\phi c_pT\right)}{a\cos\phi\partial\phi}+\frac{\partial}{\partial\theta}\left(\sigma\dot{\theta}c_pT\right)=\sigma\left(Q+\alpha\omega\right), \quad (6.10)$$

where $Q = \Pi\dot{\theta}$. Adding (6.9) and (6.10) we obtain the total energy equation

$$\frac{\partial}{\partial t}\left(\sigma(K+c_pT)\right)+\frac{\partial}{a\cos\phi\partial\phi}\left(\sigma v\cos\phi(K+M)\right)$$
$$+\frac{\partial}{\partial\theta}\left(\sigma\dot{\theta}(K+M)-\phi\frac{\partial p}{\partial t}\right)=\sigma(uF+Q). \quad (6.11a)$$

The lower boundary will be regarded in the same way as before. The bottom isentropic surface θ_B is the largest value of θ which remains everywhere below the earth's surface. Assuming the top boundary θ_T is both an isentropic and isobaric surface and assuming no topography and a vanishing $\dot{\theta}$ at the top and bottom, we can integrate (6.11a) over the entire atmosphere to obtain

$$\frac{\partial}{\partial t}\iint(K+c_pT)\sigma a\cos\phi d\phi d\theta=\iint(uF+Q)\sigma a\cos\phi d\phi d\theta. \quad (6.11b)$$

The equation for the absolute vorticity

$$\zeta=2\Omega\sin\phi-\frac{\partial(u\cos\phi)}{a\cos\phi\partial\phi} \quad (6.12)$$

is derived from (6.1) and can be written as either

$$\frac{D\zeta}{Dt}+\zeta\frac{\partial(v\cos\phi)}{a\cos\phi\partial\phi}-\frac{\partial u}{\partial\theta}\frac{\partial\dot{\theta}}{a\partial\phi}+\frac{\partial(F\cos\phi)}{a\cos\phi\partial\phi}=0, \quad (6.13)$$

or, in the alternative flux form,

$$\frac{\partial(\sigma P)}{\partial t}+\frac{\partial}{a\cos\phi\partial\phi}\left[\left(v\sigma P-\dot{\theta}\frac{\partial u}{\partial\theta}+F\right)\cos\phi\right]=0, \quad (6.14)$$

where $P = \zeta/\sigma$ is the potential vorticity. Earlier remarks on the Haynes-McIntyre theorem (1987) apply, i.e. the potential vorticity flux is exactly isentropic. Eliminating the horizontal divergence between (6.4) and (6.13) gives

$$\sigma\frac{DP}{Dt}=\frac{\partial u}{\partial\theta}\frac{\partial\dot{\theta}}{a\partial\phi}+\zeta\frac{\partial\dot{\theta}}{\partial\theta}-\frac{\partial(F\cos\phi)}{a\cos\phi\partial\phi}, \quad (6.15)$$

which is the usual form of the potential vorticity equation.

6.1.2 Coordinate transformation

The potential latitude coordinate Φ was first defined by Hack et al. (1989) as

$$\sin\Phi = \sin\phi - \frac{u\cos\phi}{a\Omega(\sin\Phi + \sin\phi)}, \tag{6.16a}$$

which is equation (3.8a) in their paper. Comparing (6.16a) to (3.21) shows that the coordinate used in three dimensional semigeostrophic theory on the sphere (chapter 3) can be thought of as an approximate potential latitude. We obtain (3.21) from (6.16a) if u is replaced by u_g and if ϕ in the second term on the right hand side is replaced by Φ. We can rewrite (6.16a) as

$$\Omega a\cos^2\Phi = \Omega a\cos^2\phi + u\cos\phi, \tag{6.16b}$$

which makes apparent the connection of potential latitude to the total angular momentum: the potential latitude can be interpreted as the latitude to which an air parcel must be moved (conserving absolute angular momentum) in order for its zonal wind component to vanish. The advantages of using an angular momentum coordinate in balanced zonal models have been discussed by Shutts (1980). Since the argument of the inverse cosine function must not exceed unity, we limit our attention to flows for which $u\cos\phi \le \Omega a\sin^2\phi$. Note that this excludes westerly flows at the equator, but that frictionless flows which develop by thermal forcing from an initial state of rest are never westerly at the equator.

Transforming from (ϕ,θ,t) space to (Φ,Θ,T) space, where $\Theta = \theta$ and $T = t$, gives the following relations between derivatives

$$\frac{\partial}{\partial t} = \frac{\partial\Phi}{\partial t}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial T}, \tag{6.17}$$

$$\frac{\partial}{\partial\phi} = \frac{\partial\Phi}{\partial\phi}\frac{\partial}{\partial\Phi}, \tag{6.18}$$

$$\frac{\partial}{\partial\theta} = \frac{\partial\Phi}{\partial\theta}\frac{\partial}{\partial\Phi} + \frac{\partial}{\partial\Theta}. \tag{6.19}$$

Another way of writing (6.18) is

$$\frac{\partial}{\cos\phi\partial\phi} = \left(\frac{\zeta}{2\Omega\sin\Phi}\right)\frac{\partial}{\cos\Phi\partial\Phi}, \tag{6.20}$$

since the vorticity can be expressed as

$$\zeta = 2\Omega \sin\Phi \frac{\partial \sin\Phi}{\partial \sin\phi}. \tag{6.21}$$

We limit our attention to flows in which Φ is a monotonically increasing function of ϕ so that $\zeta/(2\Omega\sin\Phi) > 0$. In regions where $\zeta > 2\Omega\sin\Phi$ the Φ coordinate provides a natural stretching which is analogous to the stretching provided in similar situations by the generalized geostrophic coordinate in semigeostrophic theory (chapter 3). From (6.17)–(6.19) we can easily show that (6.5) can also be written as

$$\frac{D}{Dt} = \frac{\partial}{\partial T} + \dot{\Phi}\frac{\partial}{\partial \Phi} + \dot{\theta}\frac{\partial}{\partial \Theta}, \tag{6.22}$$

where (6.16b) in (6.6) allows us to write

$$a\dot{\Phi}\cos\Phi 2\Omega\sin\Phi = -F\cos\phi. \tag{6.23}$$

The advantage of (6.22) over (6.5) is the elimination of the divergent wind component v.

Introducing the potential pseudodensity σ^*, which as before is defined such that $P\sigma^* = 2\Omega\sin\Phi$, the Bernoulli function M^*, and a function of the zonal flow u^*, we have,

$$\sigma^* = \left(\frac{2\Omega\sin\Phi}{\zeta}\right)\sigma, \tag{6.24}$$

$$u^* = \left(\frac{\cos\phi}{\cos\Phi}\right)u, \tag{6.25}$$

$$M^* = M + \frac{1}{2}u^2. \tag{6.26}$$

The potential pseudodensity is the pseudodensity a parcel would acquire if ζ were changed to $2\Omega\sin\Phi$ under conservation of potential vorticity. The new dependent variable u^* allows us to write transformation relations in more compact form such as the form we get when (6.20) is applied to $u\cos\phi$:

$$\frac{2\Omega\sin\phi - \frac{\partial(u\cos\phi)}{a\cos\phi\partial\phi}}{2\Omega\sin\phi} = \frac{2\Omega\sin\Phi}{2\Omega\sin\Phi + \frac{\partial(u^*\cos\Phi)}{a\cos\Phi\partial\Phi}}. \tag{6.27}$$

Now, (6.16b) and (6.25) can be combined to yield

$$\frac{\Omega + \frac{u}{a\cos\phi}}{\Omega} = \frac{\Omega}{\Omega - \frac{u^*}{a\cos\Phi}}. \tag{6.28}$$

Thus, as $-\partial(u^* \cos\Phi)/(a\cos\Phi\partial\Phi)$ approaches $2\Omega\sin\Phi$ the absolute vorticity becomes much larger than the local Coriolis parameter, and as $u^*/(a\cos\Phi)$ approaches Ω the absolute circulation per unit area becomes infinite.

To ensure a one to one correspondence between Φ and ϕ we require that

$$\frac{\partial \sin\Phi}{\partial \sin\phi} > 0. \tag{6.29}$$

Writing (6.16b) in terms of sines, we get

$$\sin\Phi = \pm\left(\sin^2\phi - \frac{u\cos\phi}{\Omega a}\right)^{1/2}. \tag{6.30}$$

We can find a condition on how to choose the sign in (6.30) by differentiating (6.30) and using (6.29). We obtain,

$$\frac{\partial \sin\Phi}{\partial \sin\phi} = \pm\left\{\frac{2\Omega\sin\phi - \frac{\partial u\cos\phi}{a\cos\phi\partial\phi}}{2\Omega\sin\phi}\right\}\frac{\sin^2\phi}{\sin^2\Phi}\left(1 - \frac{u\cos\phi}{\Omega a\sin^2\phi}\right)^{1/2} > 0. \tag{6.31}$$

Thus, when the term in brackets (the dimensionless vorticity) is negative we choose the negative sign. Conversely, when the term in brackets is positive we choose the positive sign.

With the new variables u^* and M^*, the balance equation (6.2) and the hydrostatic equation (6.3) transform to

$$\left(\frac{2\Omega\sin\Phi}{1 - \frac{u^*}{\Omega a\cos\Phi}}\right)u^* + \frac{\partial M^*}{a\partial\Phi} = 0, \tag{6.32}$$

$$\frac{\partial M^*}{\partial\Theta} = \Pi. \tag{6.33}$$

Formally, (6.33) is identical to (6.3) while (6.32) is simpler than (6.2) in that (6.32) allows only one u^* for a given $\partial M^*/\partial\Phi$.

We now want to derive the potential pseudodensity equation which will be the fundamental prognostic equation of the model. The σ^* equation is derived from the potential vorticity equation (6.15). We first note that (6.18) and (6.19) can be combined to yield

$$\frac{\partial u}{\partial\theta}\frac{\partial}{a\partial\phi} + \zeta\frac{\partial}{\partial\theta} = \zeta\frac{\partial}{\partial\Theta}, \tag{6.34}$$

which makes apparent the nomenclature vortex coordinates. Using (6.20), (6.23) and (6.34) we can rewrite the right-hand side of (6.15) to obtain the potential pseudodensity equation

$$\frac{D\sigma^*}{Dt} + \sigma^* \left(\frac{\partial(\dot{\Phi}\cos\Phi)}{\cos\Phi\partial\Phi} + \frac{\partial\dot{\theta}}{\partial\Theta} \right) = 0. \tag{6.35}$$

In the absence of heating and friction σ^* is conserved. However, the ITCZ is highly nonconservative. As we shall see, a midtropospheric maximum in $\dot{\theta}$ plays the crucial role of a sink of σ^* in the lower troposphere and a source of σ^* in the upper troposphere. The flux form of (6.35) will prove to be particularly useful. With D/Dt given by (6.22), the flux form of (6.35) becomes

$$\frac{\partial\sigma^*}{\partial T} + \frac{\partial(\sigma^*\dot{\Phi}\cos\Phi)}{\cos\Phi\partial\Phi} + \frac{\partial(\sigma^*\dot{\theta})}{\partial\Theta} = 0. \tag{6.36}$$

The advantage of (6.36) is that, if the source terms $\dot{\Phi}$ and $\dot{\theta}$ are known functions of (Φ, Θ, T), then the problem of solving for the time evolution of σ^* is separate from the rest of the dynamics. If $\dot{\Phi}$ and $\dot{\theta}$ are simple enough, (6.36) can even be solved analytically, as was discussed by Schubert and Alworth (1987) and Schubert et al. (1989). Such analytic solutions of (6.36) will be further discussed in section 6.2.1.

6.1.3 Invertibility principle

The potential pseudodensity σ^* is a combination of the mass field σ and the balanced wind field u expressed in terms of the dimensionless vorticity $\zeta/2\Omega\sin\Phi$. Since σ is related to M^* through hydrostatic balance (6.33) and ζ is related to M^* through gradient balance (6.32) the complete flow field can be obtained from σ^* by inverting it to get M^*. To derive the invertibility principle we use the transformation relations (6.19) and (6.20) in the definition (6.24) to obtain the Jacobian form

$$\frac{\partial(\sin\phi, p)}{\partial(\sin\Phi, \Theta)} + \sigma^* = 0, \tag{6.37}$$

which can be written as

$$\frac{\partial(\sin\phi, \Pi)}{\partial(\sin\Phi, \Theta)} + \Gamma\sigma^* = 0, \tag{6.38}$$

when noting that $\Gamma(p) = d\Pi/dp$. To derive the Jacobian in (6.38) it is easiest to express the balance condition (6.32) in terms of ϕ, Φ and M^* as follows

$$2\Omega \sin\Phi\, \Omega a \left(\frac{\cos^2\Phi}{\cos^2\phi} - 1 \right) + \frac{\partial M^*}{a \cos\Phi \partial\Phi} = 0 \tag{6.38}$$

and then take the derivatives. Using the hydrostatic and balance conditions we get

$$\left[4\Omega^2 \sin^2\Phi - \sin\Phi \cos^3\Phi \frac{\partial}{a\partial\Phi} \left(\frac{1}{\sin\Phi \cos^3\Phi} \frac{\partial M^*}{a\partial\Phi} \right) \right] \frac{\partial^2 M^*}{\partial\Theta^2} + \left(\frac{\partial^2 M^*}{a\partial\Phi\partial\Theta} \right)^2$$

$$+\Gamma\sigma^* \frac{\sin\phi}{\sin\Phi} \left(2\Omega \sin\Phi - \frac{\partial M^*}{\Omega a^2 \cos\Phi \partial\Phi} \right)^2 = 0, \tag{6.39a}$$

which is the desired relation between M^* and σ^*. If the upper isentropic surface $\Theta = \Theta_T$ is also an isobaric surface with Exner function Π_T, the upper boundary condition for (6.39a) is

$$\frac{\partial M^*}{\partial\Theta} = \Pi_T \quad \text{at} \quad \Theta = \Theta_T. \tag{6.39b}$$

We now approximate the lower boundary condition by assuming that the geopotential vanishes on the lower isentropic surface $\Theta = \Theta_B$, so that $M = \Theta\Pi$ there. Using (6.26) to express M in terms of M^* and u, (6.25) to express u in terms of u^*, (6.32) to express u^* in terms of M^*, and (6.33) to express Π in terms of M^*, we can write the lower boundary condition as

$$2\Omega \sin\Phi \left(2\Omega \sin\Phi - \frac{\partial M^*}{\Omega a^2 \cos\Phi \partial\Phi} \right) \left(\Theta \frac{\partial M^*}{\partial\Theta} - M^* \right) + \frac{1}{2} \left(\frac{\partial M^*}{a\partial\Phi} \right)^2 = 0 \quad \text{at} \quad \Theta = \Theta_B. \tag{6.39c}$$

For the boundary conditions at the poles, symmetry requires

$$\frac{\partial M^*}{\partial\Phi} = 0 \quad \text{at} \quad \Phi = \pm\frac{\pi}{2}. \tag{6.39d}$$

We can now summarize the results of our analysis as follows. If the time evolution of the σ^* field can be determined from (6.36), we can then solve the diagnostic problem (6.39) for M^*, after which the wind field u^* and the mass field Π can be determined from (6.32) and (6.33). This is all accomplished in (Φ, Θ) space. The transformation to other representations, e.g., $u(\phi, \theta)$ or $u(\phi, p)$, is straightforward.

The diagnostic problem (6.39) involves nonlinearities in both the partial differential equation (6.39a) and the lower boundary condition (6.39c). Because of the nonlinearities, problem (6.39) must be solved using an iterative technique. Note that the coefficients Γ and $\sin\phi$ in (6.39a) are functions of M^*, and must therefore be included in the iterative procedure.

6.2 The breakdown of the ITCZ

We now want to use the idealized model developed in the previous section to study the response of the zonally symmetric atmosphere to a deep tropical heat source. To simplify the solution of (6.36) friction is ignored ($\dot{\Phi} = 0$). A heating function separable in potential latitude and potential temperature is assumed. Its latitudinal part is of Gaussian structure, the same as the one used by Hack et al. (1989). The vertical structure is a simple sine function with a midtropospheric maximum and it is zero on both boundaries. Given this heating profile the prognostic equation can be solved analytically. Then equation (6.39) need only be inverted when output is desired and thus efficiency is not a primary concern.

6.2.1 An analytical solution of the σ^* equation

In the spirit of Schubert and Alworth (1987) and Schubert et al. (1989) we will use the method of characteristics to solve (6.36). Let us consider the case where the heating is a sine function in the vertical

$$\dot{\theta} = Q(\Phi)\sin(\pi Z), \tag{6.40}$$

where $Z = (\Theta - \Theta_B)/(\Theta_T - \Theta_B)$, and $Q(\Phi)$ is the Gaussian latitude distribution of the specified heating and will be considered later. Multiplying (6.36) by $\dot{\theta}$ and using (6.40) we obtain

$$\frac{\partial}{\partial \mathcal{T}}(\dot{\theta}\sigma^*) + \sin(\pi Z)\frac{\partial}{\partial Z}(\dot{\theta}\sigma^*) = 0, \tag{6.41}$$

where $\mathcal{T}(\Phi) = Q(\Phi)T/(\Theta_T - \Theta_B)$. According to (6.41) the quantity $\dot{\theta}\sigma^*$ is constant along each characteristic curve determined from $dZ/\sin(\pi Z) = d\mathcal{T}$. By integration of this equation we can show that the characteristic through the point $(Z, \mathcal{T})$ intersects the $\mathcal{T} = 0$ axis at

$$Z_0(Z, \mathcal{T}) = \frac{2}{\pi} \tan^{-1}\left[e^{-\mathcal{T}} \tan\left(\frac{\pi Z}{2}\right)\right]. \tag{6.42}$$

Since $\dot{\theta}\sigma^*$ is constant along each characteristic, then

$$\dot{\theta}(Z, \mathcal{T})\sigma^*(Z, \mathcal{T}) = \dot{\theta}(Z_0(Z, \mathcal{T}), 0)\, \sigma^*(Z_0(Z, \mathcal{T}), 0). \tag{6.43}$$

Assuming that the initial value of σ^* is the constant σ_0, we can use (6.40) and (6.42) to write (6.43) as

$$\sigma^*(Z, \mathcal{T}) = \sigma_0 \frac{\sin\left\{2 \tan^{-1}\left[e^{-\mathcal{T}} \tan\left(\frac{\pi Z}{2}\right)\right]\right\}}{\sin(\pi Z)}. \tag{6.44a}$$

Although (6.44a) is indeterminate at the boundaries $Z = 0$ and $Z = 1$, use of l'Hopital's rule yields

$$\sigma^*(Z, \mathcal{T}) = \sigma_0 \begin{cases} e^{\mathcal{T}} & Z = 1 \\ e^{-\mathcal{T}} & Z = 0. \end{cases} \tag{6.44b}$$

Equations (6.44) constitute the analytic solution of the frictionless version of the potential pseudodensity equation when the diabatic source has the form (6.40). The complete solution $\sigma^*(\Phi, \Theta, \mathcal{T})$ can be plotted once $Q(\Phi)$, and hence $\mathcal{T}(\Phi)$, is specified. Since the $\mathcal{T}$ clock runs faster where $Q(\Phi)$ is large, the largest anomalies in the σ^* field will occur in the ITCZ.

For the latitudinal distribution of the heating we choose the form

$$Q(\Phi) = Q_0 4\alpha\pi^{-\frac{1}{2}}\left\{\mathrm{erf}[\alpha(1 + \sin\Phi_c)] + \mathrm{erf}[\alpha(1 - \sin\Phi_c)]\right\}^{-1} \exp[-\alpha^2(\sin\Phi - \sin\Phi_c)^2]. \tag{6.48}$$

By varying Φ_c and α we can consider simulated ITCZ's centered at different latitudes and with different widths. Through integration of (6.48) it can be shown that

$$\frac{1}{2}\int_{-\pi/2}^{\pi/2} Q(\Phi) \cos\Phi \, d\Phi = Q_0, \tag{6.49}$$

so that different values of Φ_c and α all result in the same area averaged heating Q_0. In particular we shall set $\alpha = 15$, $0 \leq \Phi_c \leq 30$ degrees, $\Theta_T = 360$ K and $\Theta_B = 300$ K. The latitude e-folding width of heating corresponding to $\alpha = 15$ is 8 degrees which is in agreement with Marshall Islands rainfall data (Yanai et al., 1973, Fig. 12) and has been discussed by Hack et al. (1989). Choosing $Q_0 = 0.30$ K/day results in a peak heating $Q(\Phi_c) \approx 5.1$ K/day.

6.2.2 Results

The fields of σ^*/σ_0 computed from (6.44) at $T = 2,3,4$ days are shown in figures 6.1a–6.3a. The corresponding fields of potential vorticity P, normalized by $2\Omega/\sigma_0$ are shown in the lower part of the figures (6.1b–6.3b). In the ITCZ, a region of low potential pseudodensity develops at lower levels and a region of high potential pseudodensity at upper levels. Due to vertical advection, the upper tropospheric maximum in σ^* begins to get pinched off as a small σ^* spreads throughout the troposphere. The convective modification of the P field occurs within a background state which has a northward increase of P. As convection continues the gradient of P becomes locally reversed in the lower troposphere poleward of the ITCZ and in the upper troposphere equatorward of the ITCZ. These features are consistent with observations made by Burpee (1972) in his study of the origins of easterly waves in the lower troposphere of the north African region. According to the Charney-Stern theorem, such zonal flows (i.e., those with a reversal in the meridional gradient of the potential vorticity) are unstable. Thus, it would appear that ITCZ convection alone can lead to the generation of unstable zonal flows. This may be the cause of periodic breakdowns of the ITCZ.

6.3 Framework for future work

The task at hand is to extend the theory so as to fill the top right-hand box in figure 1.1, i.e., to derive a fully consistent globally valid three dimensional balanced theory.

The most general system of balanced equations is that of Charney (1962). However, in its present form the equations are highly implicit in the dependent variables and the dynamics are far from being reducible to the simple mathematical form of semigeostrophic theory. Additionally, even though the balance equations include a full representation of curvature effects, they are formally not valid in regions of large vorticity. A transformation to vortex coordinates should solve that. Because of the complexity of the nonlinear balance equations the coordinate transformations might be expected to take the form of differential relations. However, similar to semigeostrophic theory, the prediction and diagnosis of the balanced mass and wind fields would be performed in transformed space so

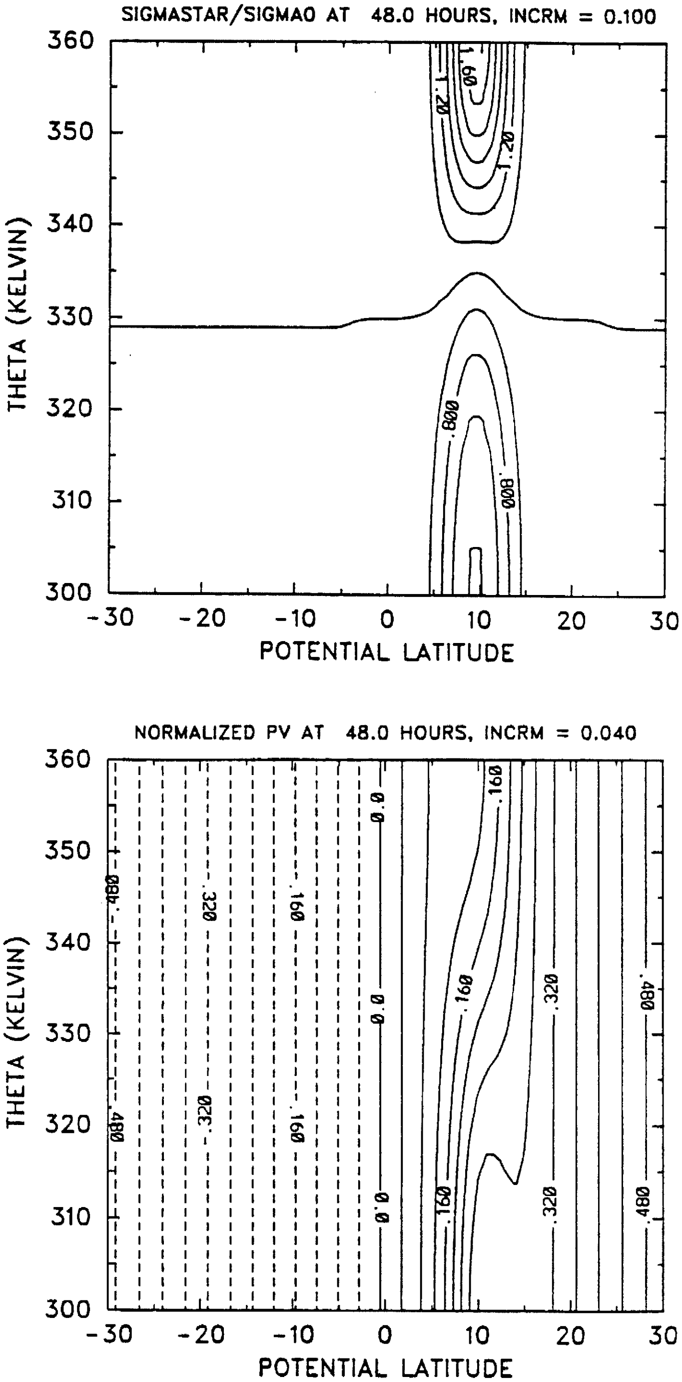

Figure 6.1: Results for $T = 2$ days. (a) Isolines of σ^*/σ_0 (i.e., potential pseudodensity measured in units of σ_0) in (Φ, Θ) space. Note that the convection of the ITCZ generates a lower tropospheric layer of high potential vorticity and an upper tropospheric layer of low potential vorticity. (b) Isolines of $\sigma_0 P/(2\Omega)$ (i.e., potential vorticity measured in units of $2\Omega/\sigma_0$).

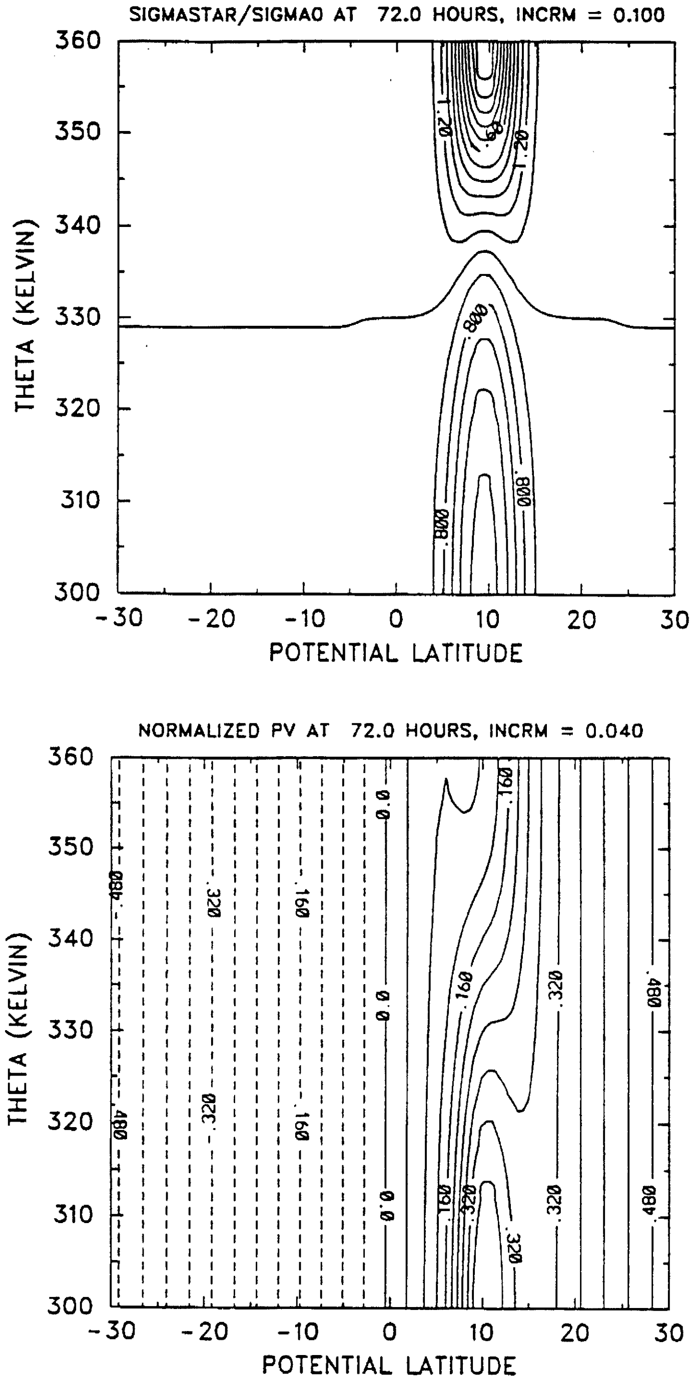

Figure 6.2: Results for $T = 3$ days. (a) Isolines of σ^*/σ_0 (i.e., potential pseudodensity measured in units of σ_0) in (Φ, Θ) space. (b) Isolines of $\sigma_0 P/(2\Omega)$ (i.e., potential vorticity measured in units of $2\Omega/\sigma_0$).

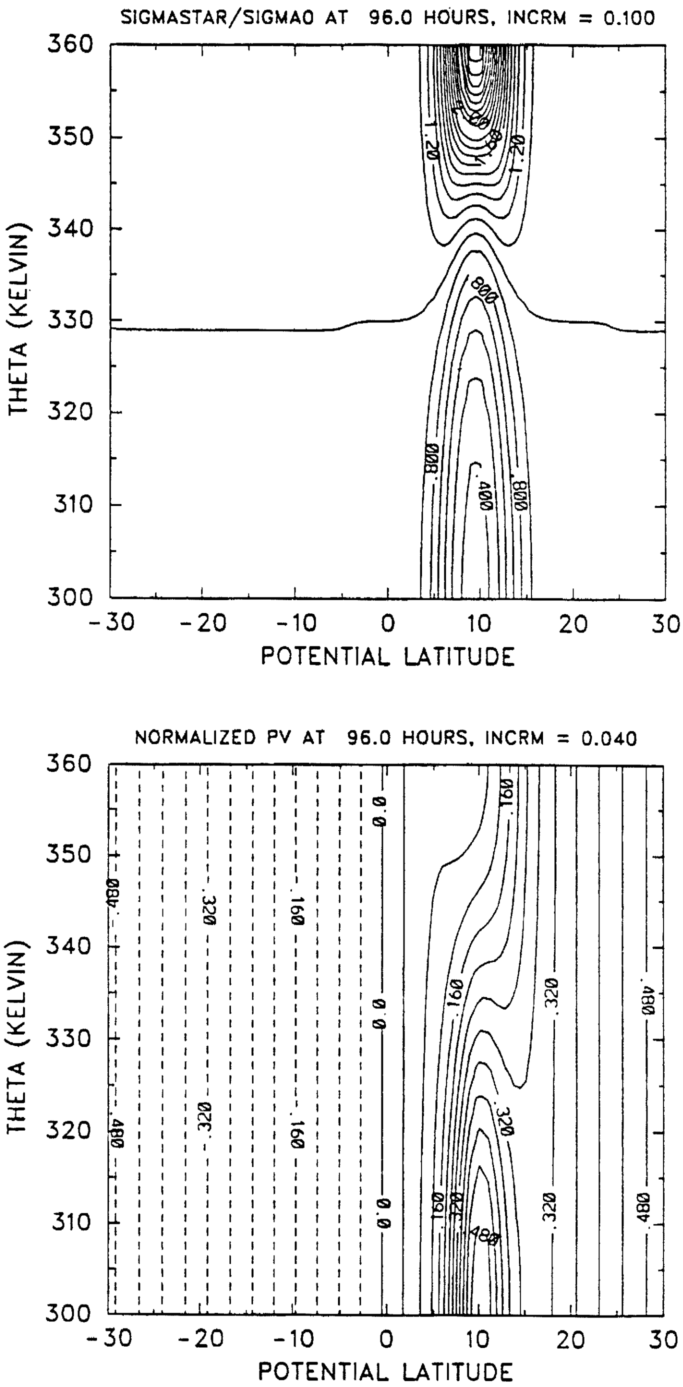

Figure 6.3: Results for $T = 4$ days. (a) Isolines of σ^*/σ_0 (i.e., potential pseudodensity measured in units of σ_0) in (Φ, Θ) space. (b) Isolines of $\sigma_0 P/(2\Omega)$ (i.e., potential vorticity measured in units of $2\Omega/\sigma_0$).

that a transformation to physical space, and thus a solution of the coordinate differential equations, would only be required at times when output was desired. Within such a theory the nonlinear balance system could be expected to blossom out in the same way that the geostrophic momentum approximation system blossomed out within semigeostrophic theory.

From the discussion in chapter 4, one expects Hamilton's principle to be the ideal tool with which to perform the task at hand. The problem is, how to formulate Hamilton's principle, i.e., how to define the Lagrangian of the system so as to arrive at the desired result. Seliger and Whitham (1968) addressed a similar problem within the Eulerian description of fluid dynamics. They were concerned with finding the simplest variational principle that gave exactly the equations of motion of a certain system and no others. They concluded that the Clebsch representation of the wind would lead to the simplest principle, and the Lagrangian was simply the pressure. The Clebsch representation may be written

$$\mathbf{v} = \nabla\chi + S\nabla\eta + \alpha\nabla\beta, \tag{6.50}$$

where S is entropy, χ, η, α and β are scalar potentials. The vortex coordinates are equal to α and β. Independent variations of χ, S, η, α and β lead to the equations of motion. For our purposes the Clebsch representation seems especially promising, since we need to approximate the wind of the system with a balanced wind. This is certainly worth further study.

Chapter 7

SUMMARY AND CONCLUSIONS

Balanced models offer an alternative to the full primitive equations for studying the rotational modes of atmospheric motions. The fact that these models filter out gravity waves makes it easier to interpret the dynamical processes that are involved in many weather phenomena; thus balanced models hold the promise of providing valuable insight into the workings of our atmosphere. A requirement one strives to fulfill when developing balanced models is that of keeping the fundamental conservation principles intact and thus assuring that the models are consistent. However, the models will always be limited by the severity of the approximations made to the momentum equations, the assumptions of the underlying balance, and by how the earth's rotation is represented. Again, the importance of the transformation to vortex coordinates should be emphasized. Then the equations can be expressed in closed form without any horizontal ageostrophic advection. Additionally, these coordinates provide the natural stretching of regions of low vorticity and shrinking of regions of high vorticity resulting in more symmetric length scales. Combining vortex coordinates with the isentropic coordinate makes the divergent part of the wind entirely implicit and the whole dynamics reduce to two equations, a prognostic equation for the potential pseudodensity and an invertibility relation from which both the mass and the wind fields may be diagnosed. This work is concerned with generalizing semigeostrophic theory to take account of the variability of the Coriolis parameter. The assumption of geostrophic balance in semigeostrophic theory makes it impossible to generalize it to the entire sphere, but a hemispheric model was developed in chapter 3. A β-plane version of the theory was derived in chapter 2. Chapter 4 described a powerful technique for deriving approximate dynamical models. This technique involves applying Hamilton's principle to an approximate Lagrangian while preserving the symmetries of the original Lagrangian and thus guaranteeing the existence of the important conservation principles. In particular it was shown how the theory in chapter 3 can be derived using this approach. In fact a whole ensemble of balanced models was derived; each of those models had a particular balance condition and a corresponding coordinate transformation, but all of them satisfied the canonical momentum equations. Our choice of the models in chapters 2 and 3 was motivated by the appearance of the approximate momentum equations in physical coordinates and their resemblance to the f-plane momentum equations with the geostrophic momentum approximation. Chapter 5 addressed the Charney-Stern theorem which was then used in the following chapter to examine the periodic breakdown of the ITCZ in a zonally symmetric atmosphere with a gradient wind balance. This work is but a stroll on the road to a globally valid balanced theory, the two dimensional form of which is this last model. Note that in the meridional direction we have a natural balance, so to speak, guided by the angular momentum principle, and a corresponding vortex coordinate. We are not so fortunate in the zonal direction.

Perhaps the primary virtue of the semigeostrophic β-plane model developed in chapter 2 over and above that of Salmon (1985) is that it is fully three dimensional in the elegant and concise version of isentropic and vortex coordinates. The linearized version of this model leads to a generalized Charney-Stern theorem for barotropic-baroclinic instability and to Rossby wave solutions with a meridional structure different from that in quasi-geostrophic theory. This model would seem to be ideal for studying the occlusion process in synoptic scale baroclinic waves. Similarly, the semigeostrophic hemispheric model developed in chapter 3 takes advantage of combining vortex and isentropic coordinates, again reducing the dynamics to just two fundamental equations. Once more, the linearized version leads directly to the Charney-Stern theorem. This theory would seem ideal for studying large scale processes such as stratospheric dynamics. The only other semigeostrophic spherical model known to us is that of Shutts (1988) which suffers from a rather severe approximation to its kinetic energy and therefore is not as general. Admittedly, the computational problem poses somewhat of a challenge. The fundamental diagnostic equation of both models is a nonlinear second order problem with a nonlinear lower boundary condition. It is of primary importance to have an efficient solver since this equation must be solved at each timestep. Fulton (1989) used multigrid methods (Fulton et al., 1986) to solve the invertibility for a two dimensional, f-plane problem. This numerical technique holds great promise for the future because it is fast, efficient and should be able to handle the discontinuity introduced at the lower boundary. The prognostic equation has to predict a discontinuous positive field, but as was mentioned earlier, workable schemes do exist.

A globally valid three dimensional balanced theory would be extremely valuable since it would combine all the balanced flow processes of the atmosphere under one hat, independent of location, scale or curvature of the system. In that way it would unify midlatitude and tropical filtered theories into one framework. For example, this theory could be used to study the transformation of an easterly wave into a hurricane. It may be argued that the equatorial atmosphere is inherently unbalanced and that Kelvin waves play a major part in its dynamics. From that point of view, a filtered three dimensional theory is not beneficial and one may resort to the long wave approximation theory which was derived in the last section of chapter 4. In fact, the whole approach of potential vorticity thinking is incomplete in the equatorial atmosphere, since Kelvin waves are devoid of potential vorticity.

REFERENCES

Abramowitz, M., and I. Stegun, 1964: Handbook of Mathematical Functions. Dover edition, 1965, Dover, New York, 1046pp.

Andrews, D. G., 1983: A finite-amplitude Eliassen-Palm theorem in isentropic coordinates. *J. Atmos. Sci.*, **40**, 1877–1883.

Andrews, D. G., and M. E. McIntyre, 1976: Planetary waves in horizontal and vertical shear: The generalized Eliassen-Palm relation and the mean zonal acceleration. *J. Atmos. Sci.*, **33**, 2031 –2048.

Andrews, D. G., and M. E. McIntyre, 1978: Generalized Eliassen-Palm and Charney-Drazin theorems for waves on axisymmetric flows in compressible atmospheres. *J. Atmos. Sci.*, **35**, 175 –185.

Bretherton, F. P., 1966: Critical layer instability in baroclinic flows. *Quart. J. R. Met. Soc.*, **92**, 325–334.

Burpee, R.W., 1972: The origin and structure of easterly waves in the lower troposphere of North Africa. *J. Atmos. Sci.*, **29**, 77–90.

Buzzi, A., A. Trevisan and G. Salustri, 1981: Internal frontogenesis: A two-dimensional model in isentropic, semi-geostrophic coordinates. *Mon. Wea. Rev.*, **109**, 1053–1060.

Charney, J. G., 1948: On the scale of atmospheric motions. *Geofys. Publ.*, **17**, No. 2, 1–17.

Charney, J. G., 1962: Integration of the primitive and balance equations. *Proc. Intern. Symp. Numerical Weather Prediction*, Meteor. Soc. Japan, Tokyo, 131–152.

Charney, J. G. and M. E. Stern, 1962: On the stability of internal baroclinic jets in a rotating atmosphere. *J. Atmos. Sci.*, **19**, 159–172.

Courant, R., and D. Hilbert, 1962: *Methods of Mathematical Physics*, volume II. Wiley-Interscience, 830 pp.

Eckart, C., 1960: Variation principles of hydrodynamics. *Phys. Fluids*, **3**, 421–427.

Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, 1980: Eliassen-Palm cross sections for the troposphere. *J. Atmos. Sci.*, **37**, 2600–2616.

Eliassen, A., 1948: The quasi-static equations of motion. *Geofys. Publ.*, **17**(3).

Eliassen, A., 1952: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norv.*, **5**, 60 pp.

Eliassen, A., 1983: The Charney–Stern theorem on barotropic–baroclinic instability. *Pure Appl. Geophys.*, **121**, 563—572.

Eliassen, A., and E. Raustein, 1968: A numerical integration experiment with a model atmosphere based on isentropic surfaces. *Meteorologiske Annaler*, **5**, No. 2, 45–63.

Eliassen, A., and E. Raustein, 1970: A numerical integration experiment with a six-level atmospheric model with isentropic information surfaces. *Meteorologiske Annaler*, **5**, No. 11, 429–449.

Ertel, H., 1942: A new hydrodynamical vorticity equation. English translation by W. H. Schubert, personal communication. Ein neuer hydrodynamischer Wirbelsätz. *Meteorologische Zeitschrift*, **59**, 22–30.

Fulton, S. R., 1989: Multigrid solution of the semigeostrophic invertibility relation. *Mon. Wea. Rev.*, **117**, in press.

Fulton, S. R., P. E. Ciesielski and W. H. Schubert, 1986: Multigrid methods for elliptic problems: A review. *Mon. Wea. Rev.*, **114**, 943–959.

Gill, A. E.,1980: Some simple solutions for heat-induced tropical circulation. *Quart. J. R. Met. Soc.*, **106**, 447–462.

Gill, A. E., 1981: Homogeneous intrusions in a rotating stratified fluid. *J. Fluid. Mech.*, **103**, 275–295.

Gill, A. E. and P. J. Phlips, 1986: Nonlinear effects on heat induced circulation of the tropical atmosphere. *Quart. J. R. Met. Soc.*, **112**, 69–91.

Goldstein, H., 1980: Classical Mechanics. Second Edition. Addison-Wesley Publishing Company, 672 pp.

Hack, J.J., W.H. Schubert, D.E. Stevens, and H.-C. Kuo, 1988: Response of the Hadley circulation to convective forcing in the ITCZ. Submitted to *J. Atmos. Sci.*

Haynes, P. H., and M. E. McIntyre, 1987: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.*, **44**, 828–841.

Heckley, W. A., and A. E. Gill, 1984: Some simple analytical solutions to the problem of forced equatorial long waves. *Quart. J. R. Met. Soc.*, **110**, 203–217.

Heckley, W. A., and B. J. Hoskins, 1982: Baroclinic waves and frontogenesis in a non-uniform potential vorticity semi-geostrophic model. *J. Atmos. Sci.*, **39**, 1999–2016.

Hoskins, B.J., 1971: Atmospheric frontogenesis: some solutions. *Quart. J. 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., 1976: Baroclinic waves and frontogenesis. Part I: Introduction and Eady waves. *Quart. J. R. Met. Soc.*, **102**, 103–122.

Hoskins, B. J., and F. P. Bretherton, 1972: Atmospheric frontogenesis models: Mathematical formulation and solution. *J. Atmos. Sci.*, **29**, 11–37.

Hoskins, B. J., and I. Draghici, 1977: The forcing of ageostrophic motion according to the semigeostrophic equations and in an isentropic coordinate model. *J. Atmos. Sci.*, **34**, 1859–1867.

Hoskins, B. J., M. E. McIntyre and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. R. Met. Soc.*, **111**, 877–946.

Hoskins, B. J., A. J. Simmons and D. G. Andrews, 1977: Energy dispersion in a barotropic atmosphere. *Quart. J. R. Met. Soc.*, **103**, 553–567.

Lamb, H., 1932: Hydrodynamics. Sixth edition. Dover Publications, 738 pp.

Lanczos, C., 1970: The Variational Principles of Mechanics. Fourth Edition. Dover Publications, 418 pp.

Lin, C. C., 1963: Hydrodynamics of helium II. *Proc. Int. Sch Phys. XXI.* pp 93–146. New York: academic.

Lindzen, R. S., 1967: Planetary waves on beta-planes. *Mon. Wea. Rev.*, **95**, 441–451.

Longuet-Higgins, M.S., 1964: Planetary waves on a rotating sphere I. *Proc. R. Soc.*, **A 279**, 446–473.

Lorenz, E. N., 1955: Available potential energy and the maintenance of the general circulation. *Tellus*, **7**, 157–167.

Lorenz, E.N., 1960: Energy and numerical weather prediction. *Tellus*, **12**, 17–26.

Magnus, W., F. Oberhettinger, R. P. Soni, 1966: Formulas and theorems for the special functions of mathematical physics. Third edition. Springer Verlag. 508 pp.

Matsuno, T., 1970: Vertical propagation of stationary planetary waves in the winter hemisphere. *J. Atmos. Sci.*, **27**, 871–883.

Matsuno, T., 1971: A dynamic model of the stratospheric sudden warming. *J. Atmos. Sci.*, **28**, 1479–1494.

McBride, J. L., and G. J. Holland, 1989: The Australian monsoon experiment (AMEX): Early results. *Aust. Met. Mag.*, **37**, 23–35.

McIntyre, M. E., 1980: An introduction to the generalized Lagrangian-mean description of wave, mean-flow interaction. *Pure Appl. Geophys.*, **118**, 152–176.

McIntyre, M. E., 1987: Dynamics and tracer transport in the middle atmosphere: an overview of some recent developments. *Transport Processes in the Middle Atmosphere.* Proceeding of the NATO Workshop held in November 1986 at Erice, Sicily. Edited by G. Visconti and R. Garcia. Dordrecht, Reidel.

McWilliams, J. C., and P. R. Gent, 1980: Intermediate models of planetary circulations in the atmosphere and ocean. *J. Atmos. Sci.*, **37**, 1657–1678.

Merzbacher, E., 1970: Quantum Mechanics. Second edition. John Wiley & Sons Inc. 609 pp.

Pedlosky, J., 1964: The stability of currents in the atmosphere and the ocean. Part I. *J. Atmos. Sci.*, **21**, 159–177.

Phillips, N. A., 1966: The equations of motion for a shallow rotating atmosphere and the "traditional approximation." *J. Atmos. Sci.*, **23**, 626–628.

Rayleigh, Lord, 1880: On the stability or instability of certain fluid motions. *Proc. Lond. Math. Soc.*, **11**, 57–70.

Richardson, L. F., 1922: Weather prediction by numerical process. Cambridge University Press. 236 pp.

Saarloos, W. van, 1981: A canonical transformation relating the Lagrangian and Eulerian description of ideal hydrodynamics. *Physica*, **108A**, 557–566.

Salmon, R., 1982: Hamilton's principle and Ertel's theorem. *Am. Inst. Phys. Proc.*, **88**, 127–135.

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.

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

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

Schubert, W. H., and B. T. Alworth, 1987: Evolution of potential vorticity in tropical cyclones. *Quart. J. R. Met. Soc.*, **113**, 147–162.

Schubert, W. H., S. R. Fulton, and R. F. A. Hertenstein, 1989: Balanced atmospheric response to squall lines. *J. Atmos. Sci.*, **46**, 2478– 2483.

Schubert, W. H., and J. J. Hack, 1983: Transformed Eliassen balanced vortex model. *J. Atmos. Sci.*, **40**, 1571–1583.

Seliger, R. L., and G. B. Whitham, 1968: Variational principles in continuum mechanics. *Proc. R. Soc. Lond.*, **A305**, 1–25.

Shutts, G. J., 1980: Angular momentum coordinates and their use in zonal, geostrophic motion on a hemisphere. *J. Atmos. Sci.*, **37**, 1126–1132.

Shutts, G. J., 1988: Planetary semigeostrophic equations derived from Hamilton's principle. *Met. O. 11 Scientific Note no. 9*, 63pp.

Shutts, G. J., and M. J. P. Cullen, 1987: Parcel stability and its relation to semigeostrophic theory. *J. Atmos. Sci.*, 44, 1318–1330.

Stevens, D. E., H. C. Kuo, W. H. Schubert and P.E. Ciesielsky, 1989: Quasi-balanced dynamics in the tropics. Submitted to *J. Atmos. Sci.*.

Taylor, G. I., 1915: Eddy motion in the atmsphere. *Philos. Trans. R. Soc. Lond.*, **A 115**, 1–26.

Yanai, M., S. Esbensen, and J.-H. Chu, 1973: Determination of bulk properties of tropical cloud clusters from large-scale heat and moisture budgets. *J. Atmos. Sci.*, **30**, 611–627.

Appendix A

ROSSBY WAVES ON THE β-PLANE

Here we consider solutions of

$$\frac{\partial}{\partial T}\left\{\frac{\partial^2\mathcal{M}}{\partial X^2}+f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial\mathcal{M}}{\partial Y}\right)+\frac{f^2(Y)}{\Gamma_0\sigma_0}\frac{\partial^2\mathcal{M}}{\partial\Theta^2}\right\}+\beta\frac{\partial\mathcal{M}}{\partial X}=0, \tag{A.1a}$$

$$\frac{\partial\mathcal{M}}{\partial\Theta}=0 \qquad \text{at} \qquad \Theta=\Theta_T, \tag{A.1b}$$

$$\Theta\frac{\partial\mathcal{M}}{\partial\Theta}-\mathcal{M}=0 \qquad \text{at} \qquad \Theta=\Theta_B, \tag{A.1c}$$

$$\mathcal{M}=0 \qquad \text{when} \qquad Y=Y_0, \tag{A.1d}$$

$$\mathcal{M}=0 \qquad \text{as} \qquad Y\to\infty, \tag{A.1e}$$

where $Y_0 > -f_0/\beta$, which simply means that the β-plane does not cross the equator.

To eliminate the vertical structure in (A.1a) we define the vertical inner product

$$\langle u,v\rangle=\int_{\Theta_B}^{\Theta_T}u(\Theta)v(\Theta)d\Theta. \tag{A.2}$$

For any functions u and v, we seek a vertical transform of the form

$$\mathcal{V}\left[u(\Theta)\right]=u_l=\langle u,\Psi_l\rangle, \tag{A.3}$$

where the kernel $\Psi_l(\Theta)$ of the transform is to be chosen so that

$$\mathcal{V}\left[\frac{1}{\Gamma_0\sigma_0}\frac{\partial^2\mathcal{M}}{\partial\Theta^2}\right]+\frac{1}{c_l^2}\mathcal{M}_l=0, \tag{A.4}$$

with c_l a constant. Using (A.2) and (A.3) we can integrate the left-hand side of (A.4) by parts twice to obtain

$$\mathcal{V}\left[\frac{\partial^2\mathcal{M}}{\partial\Theta^2}\right]=\left[\Psi_l\frac{\partial\mathcal{M}}{\partial\Theta}-\frac{d\Psi_l}{d\Theta}\mathcal{M}\right]_{\Theta_B}^{\Theta_T}+\int_{\Theta_B}^{\Theta_T}\frac{d^2\Psi_l}{d\Theta^2}\mathcal{M}d\Theta. \tag{A.5}$$

The boundary conditions (A.1b)–(A.1c) then imply that the desired property (A.4) will hold provided that we choose $\Psi_l(\Theta)$ and c_l as solutions of the Sturm-Liouville eigenproblem

$$\frac{1}{\Gamma_0\sigma_0}\frac{d^2\Psi_l}{d\Theta^2}+\frac{1}{c_l^2}\Psi_l=0, \tag{A.6a}$$

$$\frac{d\Psi_l}{d\Theta}=0 \qquad \text{at} \qquad \Theta=\Theta_{\mathrm{T}}, \tag{A.6b}$$

$$\Theta\frac{d\Psi_l}{d\Theta}-\Psi_l=0 \qquad \text{at} \qquad \Theta=\Theta_{\mathrm{B}}. \tag{A.6c}$$

According to the Sturm-Liouville theory (A.6) has a countably infinite set of solutions $\{c_l,\Psi_l\}_{l=0}^{\infty}$, the eigenfunctions Ψ_l are orthogonal in the inner product (A.2), may be chosen to be real and they form a complete set. Thus any well-behaved function $u(\Theta)$ may be expanded in terms of the eigenfunctions and we may write

$$u(\Theta)=\sum_{l=0}^{\infty}u_l\Psi_l(\Theta). \tag{A.7}$$

The coefficients in the expansion are given by $u_l=\langle u,\Psi_l\rangle$ where the Ψ_l have been normalized. Therefore the vertical transform pair is given by (A.3) and (A.7). Now (A.6) can easily be solved and we get

$$\Psi_l(\Theta)=\cos\left[\frac{(\Gamma_0\sigma_0)^{1/2}}{c_l}(\Theta_T-\Theta)\right], \tag{A.8}$$

where the c_l are the solutions of

$$\frac{(\Gamma_0\sigma_0)^{1/2}}{c_l}\Theta_B\tan\left[\frac{(\Gamma_0\sigma_0)^{1/2}}{c_l}(\Theta_T-\Theta_B)\right]=1. \tag{A.9}$$

The c_l can be interpreted as the gravity wave speeds corresponding to the different equivalent depths h_l satisfying $c_l=(gh_l)^{1/2}$ where g is the acceleration due to gravity. However, this is a balanced model and we have no propagating gravity waves.

Taking the vertical transform of (A.1a) we get

$$\frac{\partial}{\partial T}\left\{\frac{\partial^2\mathcal{M}_l}{\partial X^2}+f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial\mathcal{M}_l}{\partial Y}\right)-\frac{f^2(Y)}{c_l^2}\mathcal{M}_l\right\}+\beta\frac{\partial\mathcal{M}_l}{\partial X}=0. \tag{A.10}$$

The β-plane circles the earth and the resulting cyclic boundary conditions and simple X dependence allow us to define the zonal transform as the usual Fourier transform. Let us define $\mathcal{M}_{lm}(Y,T)$ as the zonal Fourier transform of $\mathcal{M}_l(X,Y,T)$. The transform pair is

$$\left.\begin{aligned} \mathcal{M}_{lm}(Y,T) &= \frac{1}{2\pi a}\int_0^{2\pi a} \mathcal{M}_l(X,Y,T)e^{-imX}\,dX, \\ \mathcal{M}_l(X,Y,T) &= \sum_{m=-\infty}^{\infty} \mathcal{M}_{lm}e^{imX} \end{aligned}\right\}, \tag{A.11}$$

where a is the distance from the earth's rotation axis. Transforming (A.10) zonally we obtain

$$\frac{\partial}{\partial T}\left\{f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial \mathcal{M}_{lm}}{\partial Y}\right)+\left[-\frac{f^2(Y)}{c_l^2}-m^2\right]\mathcal{M}_{lm}\right\}+im\beta\mathcal{M}_{lm}=0. \tag{A.12}$$

We will now design a transform to eliminate the meridional structure. Let us define the meridional inner product

$$(p|q)=\int_{Y_0}^{\infty} p(Y)q(Y)f^{-2}(Y)dY, \tag{A.13}$$

for any functions $p(Y)$ and $q(Y)$ on $[Y_0,\infty)$. We seek a meridional transform of the form

$$\mathcal{T}\,[p(Y)] = p_n = (p|\mathcal{K}_n)\,, \tag{A.14}$$

where the kernel $\mathcal{K}_n(Y)$ of the transform is to be chosen such that the meridional transform of (A.12), which completes the transformation to spectral space, has the following form

$$\frac{d\mathcal{M}_{lmn}(T)}{dT} - i\nu_{lmn}\mathcal{M}_{lmn}(T) = 0. \tag{A.15}$$

Taking the meridional transform of (A.12) we get

$$\int_{Y_0}^{\infty}\left\{f^2(Y)\frac{\partial}{\partial Y}\left(\frac{1}{f^2(Y)}\frac{\partial \hat{\mathcal{M}}_T}{\partial Y}\right)+\left[-\frac{f^2(Y)}{c_l^2}-m^2\right]\hat{\mathcal{M}}_T+im\beta\hat{\mathcal{M}}\right\}\mathcal{K}_n f^{-2}(Y)dY=0, \tag{A.16}$$

where the subscript T indicates a partial derivative with respect to T and where we have substituted the subscripts lm, indicating that a vertical and a zonal transform have been taken, with the hat over $\mathcal{M}$. Integrating by parts twice and using the boundary conditions (A.1d) and (A.1e) we obtain

$$\begin{aligned}\int_{Y_0}^{\infty}\left\{\left[\frac{d}{dY}\left(\frac{1}{f^2(Y)}\frac{d\mathcal{K}_n}{dY}\right)+\left(-\frac{1}{c_l^2}-\frac{m^2}{f^2(Y)}\right)\mathcal{K}_n\right]\hat{\mathcal{M}}_T+i\frac{m\beta}{f^2(Y)}\hat{\mathcal{M}}\mathcal{K}_n\right\}dY \\ +\left[\frac{\partial\hat{\mathcal{M}}_T}{\partial Y}\frac{1}{f^2(Y)}\mathcal{K}_n\right]_{Y_0}^{\infty}=0.\end{aligned} \tag{A.17}$$

Therefore the required property holds provided we choose the kernel $\mathcal{K}_n$ such that

$$\frac{d}{dY}\left[\frac{1}{f^2(Y)}\frac{d\mathcal{K}_n}{dY}\right]+\left[-\frac{1}{c_l^2}-\frac{m^2}{f^2(Y)}+\frac{m\beta}{\nu f^2(Y)}\right]\mathcal{K}_n=0, \tag{A.18a}$$

with the boundary conditions

$$\frac{1}{f^2(Y)}\mathcal{K}_n=0 \qquad \text{when} \qquad Y=Y_0, \tag{A.18b}$$

$$\frac{1}{f^2(Y)}\mathcal{K}_n=0 \qquad \text{when} \qquad Y\to\infty. \tag{A.18c}$$

The above is a Sturm-Liouville problem, thus the eigenfunctions $\mathcal{K}_n$ are orthogonal in the inner product (A.13), may be chosen to be real and they form a complete set. We now want to determine the eigenfunctions and eigenvalues. It is convenient to make a change of variable as follows

$$\mathcal{Y}=f(Y)=f_0+\beta Y, \tag{A.19}$$

then (A.18) becomes

$$\beta^2\frac{d}{d\mathcal{Y}}\left[\frac{1}{\mathcal{Y}^2}\frac{d\mathcal{K}_n}{d\mathcal{Y}}\right]+\left[-\frac{1}{c_l^2}-\frac{m^2}{\mathcal{Y}^2}+\frac{m\beta}{\nu\mathcal{Y}^2}\right]\mathcal{K}_n=0, \tag{A.20a}$$

$$\frac{1}{\mathcal{Y}^2}\mathcal{K}_n=0 \qquad \text{when} \qquad \mathcal{Y}=f_0+\beta Y_0, \tag{A.20b}$$

$$\frac{1}{\mathcal{Y}^2}\mathcal{K}_n=0 \qquad \text{when} \qquad \mathcal{Y}\to\infty. \tag{A.25c}$$

Let us make (A.20) dimensionless by defining the units of time and length as follows

$$[\text{time}]=(1/c_l\beta)^{1/2}, \qquad [\text{length}]=(c_l/\beta)^{1/2}. \tag{A.21}$$

Then (A.20a) becomes

$$\frac{d^2\mathcal{K}_n}{d\mathcal{Y}^2}-\frac{2}{\mathcal{Y}}\frac{d\mathcal{K}_n}{d\mathcal{Y}}+\left[-m^2-\mathcal{Y}^2+\frac{m}{\nu}\right]\mathcal{K}_n=0, \tag{A.22a}$$

Defining

$$\mathcal{K}_n(\mathcal{Y})=\mathcal{Y}\,\mathcal{F}_n(\mathcal{Y}) \tag{A.23}$$

and writing (A.22) in terms of $\mathcal{F}_n$ we get

$$\frac{d^2\mathcal{F}_n}{d\mathcal{Y}^2}+\left[-m^2+\frac{m}{\nu}-\mathcal{Y}^2-\frac{2}{\mathcal{Y}^2}\right]\mathcal{F}_n=0 \tag{A.24a}$$

$$\frac{1}{\mathcal{Y}}\mathcal{F}_n = 0 \qquad \text{when} \qquad \mathcal{Y} = (c_l\beta)^{-1/2}(f_0 + \beta Y_0), \tag{A.24b}$$

$$\frac{1}{\mathcal{Y}}\mathcal{F}_n = 0 \qquad \text{when} \qquad \mathcal{Y} \to \infty. \tag{A.24c}$$

This differential equation has the solutions (see Abramowitz and Stegun, 1965)

$$\mathcal{F}_n(\mathcal{Y}) = e^{-\mathcal{Y}^2/2}\,\mathcal{Y}^{\alpha+1/2}\,L_n^{(\alpha)}(\mathcal{Y}^2), \tag{A.25}$$

where $\alpha = \pm 3/2$, $L_n^{(\alpha)}$ denotes the generalized Laguerre polynomials, and

$$\nu_{mn} = \frac{m}{m^2 + 4n + 2\alpha + 2}. \tag{A.26}$$

The boundary condition at infinity has been satisfied. Let us write

$$\mathcal{K}_n(\mathcal{Y}) = e^{-\mathcal{Y}^2/2}\left[A\mathcal{Y}^3 L_n^{(3/2)}(\mathcal{Y}^2) + BL_n^{(-3/2)}(\mathcal{Y}^2)\right]. \tag{A.27}$$

The other boundary condition will determine B in terms of A. When the $\mathcal{K}_n$ are normalized in the inner product (A.13), the value of A is derived. Choosing the interval for $\mathcal{Y}$ to be $(0, \infty)$ is particularly convenient since then

$$\mathcal{K}_n(\mathcal{Y}) = Ae^{-\mathcal{Y}^2/2}\mathcal{Y}^3 L_n^{(3/2)}(\mathcal{Y}^2), \tag{A.28}$$

and $L_n^{(3/2)}$ are orthogonal polynomials so that A can easily be determined. Note that this corresponds to the β-plane extending infinitly close to the equator, since then $f_0 = -\beta Y_0$ or $f(Y_0) = f_0 + \beta Y_0 = 0$.

To summarize, the full solution of (A.1) is of the form

$$\mathcal{M} \sim e^{-\mathcal{Y}^2/2}\mathcal{Y}^{\alpha+3/2}L_n^{(\alpha)}(\mathcal{Y}^2)\cos\left[c_l^{-1}(\Gamma_0\sigma_0)^{1/2}(\theta_T - \theta)\right]e^{i(mX+\nu T)},$$

with the dispersion relation

$$\nu_{mn} = \frac{\beta m}{m^2 + \frac{\beta}{c_l}(4n + 2\alpha + 2)}.$$