THE FORCING AND MAINTENANCE OF GLOBAL MONSOONAL CIRCULATIONS: AN ISENTROPIC ANALYSIS

by

Donald R. Johnson
Space Science and Engineering Center
and the Department of Meteorology
University of Wisconsin-Madison
Madison, WI 53706

November 1987

published in: Advances in Geophysics (B. Saltzman, ed.) 31,
Academic Press, New York, (1989).

The Forcing and Maintenance of Global Monsoonal Circulations: An Isentropic Analysis

Page

Page

Page

1. Introduction

While the earliest theories for monsoons and the general circulation were based on concepts of thermally driven direct circulations (Halley, 1686; Hadley, 1735), Sir Napier Shaw (1930) was apparently the first to point out some constraints existing in a stratified atmosphere that isolate a relation between differential heating and mass circulations within isentropic coordinates. The constraint that he suggested as well as his arguments provide a perspective that is basic to the thermal forcing of monsoon circulations. His arguments were fundamentally valid, with one exception to be noted shortly.

Sir Napier Shaw divided the atmosphere into an overworld and an underworld with the dividing surface between these two domains being a surface of constant entropy. Within the stratified atmosphere, the overworld consisted of warmer air spanning the entire meridional extent of the atmosphere while the underworld consisted of colder air that was restricted to higher latitudes. A view of the underworld is suggested in Fig. 1 which shows the topography of the 313 K isentropic surface (Nagle, 1979). The colder air with potential temperature less than 313 K is largely confined to polar latitudes. The atmosphere above 313 K would constitute the overworld. Fig. 1

Sir Napier Shaw emphasized that the overworld exchanged properties with the underworld through convection associated with diabatic processes. No exchange between the over- and underworld would be realized with isentropic motion. In this situation the isentropic surfaces become material surfaces, which evolve geometrically through the quasi-horizontal exchange of mass, momentum and energy. He argues, although erroneously, that air within an isentropic layer was free to move laterally without any constraints. While basic theory for instabilities had been developed (Helmholtz, 1868, 1888; Raleigh, 1880; Solberg, 1928, 1930), the theories for baroclinic (Charney, 1947; Eady, 1949) and barotropic (Kuo, 1949) instabilities for planetary scale motions of the atmosphere were not developed prior to the early 40's. In all likelihood Sir Napier Shaw was unaware of the constraints imposed by dynamic stability and the requirements for forced motion in a hydrodynamically stable vortex. He was aware, however, of the stabilizing effects of the increase of potential temperature with height.

In applying his concepts, Sir Napier Shaw noted that heating forced motion upward through isentropic surfaces while cooling forced motion downward, and that such motion constituted upward and downward branches of a mass circulation embedded within a stratified atmosphere. Horizontal branches linking heat sources and sinks were implied by mass continuity. In the overworld the net horizontal mass transport would be from heat source to heat sink, while in the underworld the net mass transport would be from heat sink to heat source. The requirement that horizontal mass transport must occur between heat sources and sinks becomes evident within the time-averaged planetary scale circulation. This requirement exists independently of the requirement for energy transport, although it does not exist independently of the requirement for entropy transport within the atmosphere. In an isentropic analysis of the effects of heating on atmospheric circulation, one must carefully distinguish vertical diabatic transport of mass, energy and entropy through isentropic surfaces from Lagrangian sources of energy and entropy, since each of the governing equations for mass continuity, energy and entropy involve different and unique fundamental principles.

Atmospheric circulation is traditionally classified as monsoonal if the wind direction in a region systematically reverses between winter and summer and its speed is relatively steady with one of the seasons being wet and the other dry (Ramage, 1971; Webster, 1987a and b). Categorization by these criteria excludes baroclinic circulations of extratropical and polar latitudes. The systematic isentropic mass circulations that occur within the atmosphere's stratification are global in extent, seasonal thus monsoonal in nature, and thermally forced by the differential heating of earth and ocean. A perspective that large-scale monsoonal circulations are global in extent is different from the traditional view that monsoons are regional phenomena primarily confined within tropical-subtropical latitudes (Ramage, 1971; Lighthill and Pearce (eds.), 1981; Fein and Stephens (eds.), 1987).

In regard to this perspective, however, Webster's (1987a) description of the fundamental driving mechanisms of planetary scale monsoons is of particular interest. He states that the two foremost are "the differential heating of the land and ocean" and "the swirl introduced to the winds by the rotation of the earth", while a third is attributed to "moist processes". These same mechanisms are basic processes to the forcing of systematic global exchange of energy and entropy in temporally averaged isentropic mass circulations. Hence, in view of 1) the global nature of the thermal forcing by the differential heating of land and ocean, 2) the global response of isentropic mass, energy and entropy transport to this forcing, and 3) the seasonal reversal of the energy and entropy transport between the land and ocean, these systematic isentropic mass circulations are designated global monsoonal circulations. With the Arabic origin of the word monsoon from *mausim*, meaning season, the designation -- global monsoonal circulations -- directs the emphasis towards the seasonal variation of forcing by differential heating and response in the form of mass, energy and entropy transport.

On these matters, the seasonal reversal of energy and entropy transport between oceans and continents constitutes a direct fundamental response to the differential heating, even more so than the reversal of a surface wind direction. The latter is merely a reflection of a kinematic atmospheric structure involved with the seasonal reversal of the energy transport between land and ocean in tropical-subtropical latitudes where the effects of the earth rotation are less important. In extratropical latitudes where rotational influences are more important, the migrating nature of baroclinic waves precludes the steadiness of the wind within seasons and a distinct reversal between seasons. This condition, however, does not preclude the seasonal reversal of the time-averaged thermally forced isentropic mass circulation and the energy and entropy transport between land and oceans.

With regard to energy balance, the quasi-steady nature of the time-averaged circulation demands: 1) that the scale of the mass circulation corresponds with the scale of energy transport, and 2) that energy be transported from heat sources to heat sinks at a rate such that the net energy gained in the regions of heat sources is equal to the energy lost in regions of heat sinks. Diagnostic results have established that the isentropic quasi-horizontal energy transport occurs through the mean mass circulation (Johnson and Townsend, 1981; Johnson, 1985a and b). The mass transport in the overworld transports more energy from heat source to heat sink than is returned by the mass transport from the heat sink to the heat source in the underworld. The imbalances that occur account for the seasonal evolution of the atmosphere as well as differences that occur between seasons of different years, annual means of different years, multi-annual means of different decades, etc. However, no matter how important these differences seem to be, they result from relatively small imbalances of the time-averaged exchange within the thermally forced component of the planetary circulation.

Since the isentropic mass circulations and the transports of energy and entropy stem from separate physical principles, a classic fundamental premise after Halley (1686) will be assumed in the following discussion on the forcing of global monsoonal circulations; namely, a differentially heated atmosphere develops a mass circulation to transport energy from heat source to heat sink in order for the atmosphere's time-averaged response to be quasi-steady. This premise requires the scale and intensity of the mass circulation to be just sufficient to transport energy at a rate such that the net energy gained in the heat source region is equal to the energy lost in the heat sink region. As far as the vertically integrated energy transport is concerned within this quasi-steady behavior, divergence of energy transport must occur in the heat source region and convergence must occur in the heat sink region, while the net vertically integrated divergence of mass transport vanishes. Within the steady time-averaged state the premise does not place any constraint on the momentum structure, its balance, or its mode of response. The only requirement imposed is to demand that modes of atmospheric circulation will develop whereby, either through instabilities or by forcing, systematic isentropic mass circulations occur and in their occurence net energy is transported from heat source to heat sink.

In addition to the common scale of the isentropic mass circulation and energy transport within the steady time-averaged state, a constraint stemming from the atmosphere's entropy balance is imposed. This constraint is implicit from the unique relation between entropy and potential temperature, the use of the isentropic continuity equation and the quasi-steady constraint. Entropy increases in heat source regions and decreases in heat sink regions. Just like the energy balance, the quasi-steady entropy balance of a heat source region is maintained by a net divergence of the horizontal entropy transport in upper isentropic layers which exceeds the net convergence in lower layers[1]. Likewise the entropy balance in a heat sink region is maintained by a net convergence of entropy transport within upper isentropic layers which exceedes the net divergence within lower isentropic layers. Through this structure the upper branch of the isentropic mass circulation transports more entropy from the heat source region to the heat sink region than the lower branch returns. In the heat source region entropy is diabatically fluxed upward through isentropic surfaces, while within the heat sink region entropy is diabatically fluxed downward. However, for the isentropic mass and entropy distributions to be quasi-steady state, the global isentropically area-averaged diabatic diabatic mass transport, $\overline{\rho J_\theta \dot{\theta}}^A$, and entropy advection

---

[1]In the description of temporally or zonally averaged structure, upper and lower denote, respectively, higher-valued and lower-valued isentropic surfaces, layers, branches of the mass circulation, etc. This usage does not uniquely determine relative geometric height. For example, in a zonal average through the east-west thermal contrast of a baroclinic wave, portions of higher-valued isentropic surfaces in extratropical latitudes are actually located at lower elevations than portions of lower-value isentropic surfaces. Also, in a temporal average a higher-valued surface at one time may be located at a lower geometric height than a lower-valued surface at another time.

$\overline{\rho J_\theta \dot{\theta} \partial s/\partial\theta}^A$ must vanish. Such is not the case for the globally area-averaged diabatic energy advection, $\rho J_\theta \overline{\dot{\theta} \partial\psi/\partial\theta}^A$, which will prove to be positive. Later results will show that these conditions in combination with the energy principle require that heating occur at high pressure and cooling occur at low pressure in order for the atmosphere's circulation to be maintained, a result that is reminiscent of Sandström's Theorem (Sandström, 1916; Godske, 1936; Godske et. al, 1957; Dutton, 1976a).

The means to isolate these time-averaged circulations forced by differential heating from the spectrum of atmospheric motion has become feasible for the planetary scale through use of the data sets from the Global Weather Experiment (GWE) and global isentropic analyses of mass and energy transport (Johnson and Townsend, 1981; Johnson et al., 1985). Hence the existence of global monsoonal circulations occurring within the isentropic stratification of the atmosphere which was foretold by Sir Napier Shaw (1930) can now be verified.

The objectives of this summary, a sequel to "Global and Regional Distributions of Atmospheric Heat Sources and Sinks during the GWE" (Johnson, Yanai and Schaack, 1986), are severalfold. Primary objectives are to establish a physical relationship between differential heating and thermally forced mass transport within planetary scale monsoonal circulations and to document that the scale of energy transport by the isentropic mass circulation corresponds with the scale of energy transport. The principles of mass, energy and entropy and their balance requirements, as they are applied within isentropic coordinates, will be used to establish why such simple relations exist and how these simple relations require that monsoonal circulations are determined by the planetary scale of differential heating. No attempt will be made to summarize the results of general circulation research from analyses in isobaric coordinates, although much of the results presented herein have been stimulated by the works of Professors V. P. Starr and E. N. Lorenz (e.g. Starr and Saltzman, 1966; Starr, 1968; Lorenz, 1967, etc.).

The principle results of this summary are organized into four main topics. (See table of contents at beginning of the summary.) In studying this extended summary, readers may wish to consider that the results and discussion in each of the topical areas as well as in Chapter 7 are largely self-contained and may be studied independently.

Following this introduction, Chapter 2 is devoted to diagnostic equations used in analysis of isentropic transport processes. The first main topic, consisting of Chapters 3 through 5, is devoted to the presentation of the empirical evidence for global monsoonal circulations based on analyses in isentropic coordinates. In Chapter 3, a global analyses of the time-averaged fields of heating for wavelengths 10,000 km and greater will be presented for the months of January, April, July and October for the GWE year. In Chapter 4, emphasis will be given to the isentropic monsoonal circulations of January and July in order to show the seemingly complex, but in reality disarmingly simple model of the atmosphere's response to differential heating. In Chapter 5, the analyses will substantiate that the horizontal distributions of the irrotational components of mass and of energy transport are common with each other. The results also show that distributions of the rotational components of mass and energy transport are common with each other.

The second main topic which is covered in Chapter 6 seeks to establish a physical basis for the global monsoonal circulations from the joint application of mass, energy and entropy principles. The results will show that the time-averaged differential heating, isentropic mass circulations and energy and entropy transport must enjoy common scales.

The third main topic, contained in Chapter 7, 8 and 9, sets forth the thermodynamics of mass, energy and entropy balance as they combine to maintain monsoonal circulations. Building on basic principles, in Chapter 7 a theory for an energy phase space of reversible isentropic processes is established, through which underlying theoretical relations between Lorenz's (1955a) available potential energy and Dutton's (1973) entropic energy are established. In Chapters 8 and 9 the relations between the planetary balance of energy and entropy are expressed in isobaric and isentropic coordinates and the global distributions of isentropic and isobaric energy transport processes for January and July are compared.

In the last main topic consisting of Chapter 10, the zonally-averaged isentropic mass circulations and relative angular momentum distributions for January and July are presented in conjunction with an analyses of the forcing of the isentropic mean meridional circulations. The primary purpose of this section is to show that a zonally averaged distribution of angular momentum sources and sinks exists which forces the horizontal branches of the isentropic mass circulation, which in turn transport energy and entropy directly from heat source to heat sink. In Chapter 11, a summary is presented and several unresolved problems concerning diagnosis of the climate of chaotic atmospheric circulation are discussed.

Many details regarding the structure of regional monsoonal circulations and influence of geophysical factors will remain unresolved. The concepts set forth in this summary do, however, provide a physical basis for global monsoonal circulations that will hopefully broaden one's perspective of thermal forcing at the planetary scale and the atmospheric response in the form of global monsoonal circulations.

2. Quasi-horizontal isentropic mass and energy transport

The time-averaged transports of mass (ρ) and *total flow energy* (ν) within an isentropic layer will be represented by transport potential ($\overline{X}_\rho$ and $\overline{X}_\nu$) and stream ($\overline{\Psi}_\rho$ and $\overline{\Psi}_\nu$) functions through the use of Helmholtz's theorem given by

$$\overline{\rho J_\theta \underset{\sim}{U} f} = \nabla_\theta \overline{X}_f + \underset{\sim}{k} \times \nabla_\theta \overline{\Psi}_f \ , \qquad (2.1)$$

where f is the specific property and the overbar represents a time-average. In this summary all averaging operators without a designation of the independent variable over which the averaging is defined will denote a time-average except in Chapter 10. See list of symbols. With f equal to unity, Eq. (2.1) reduces to mass transport. For the representation of mass transport, $\overline{X}$ and $\overline{\Psi}$ will be subscripted ρ. For the representation of the transport of other properties, the functions will be subscripted appropriately, such as $\overline{X}_\nu$ and $\overline{\Psi}_\nu$ for *total flow energy*.

Total energy (e) is defined to be the sum of the internal (ι), geopotential (ϕ) and kinetic (k) energies (Van Mieghem, 1973). The *total flow energy* is defined to be the sum of total energy (e) plus the RT component of pressure work. It will be italicized to minimize confusion with total energy. Changes of total energy occur through the convective transport of total energy and through the nonconvective transport of total energy by boundary pressure work (Johnson, 1980; Johnson and Downey, 1982). The RT component of work, which is also equal to $p\alpha$, is sometimes referred to as "flow energy" (Haltiner and Martin, 1957). Thus, the sum of total and "flow energy" has been designated *total flow energy* in order to recognize exchange processes associated with both quantities. For steady, isentropic motion, *total flow energy* is the effective energy that a parcel enjoys, the isopleths of which are tangent to the streamlines. Within isentropic coordinates, the *total flow energy* could be termed the "Bernoulli Stream Function", since under the condition of isentropic motion the horizontal velocity is given by $(\underset{\sim}{k} \times \nabla_\theta \nu)/(\zeta_\theta + f)$. This function has been called the Bernoulli function (Gill, 1982), total energy and also stagnation enthalpy (Milne-Thompson, 1960).

The irrotational and rotational components of the transport of the property f are defined by

$$(\rho J_\theta \underset{\sim}{U} f)_\chi = \nabla_\theta \chi_f \tag{2.2}$$

and

$$(\rho J_\theta \underset{\sim}{U} f)_\Psi = \underset{\sim}{k} \times \nabla_\theta \Psi_f \ . \tag{2.3}$$

The Poisson equations used to determine X and Ψ in Eq. (2.1) are

$$\nabla_\theta^2 \overline{X}_f = \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} f}) = \overline{\delta}_f \tag{2.4}$$

and

$$\nabla_\theta^2 \overline{\Psi}_f = \underset{\sim}{k} \cdot \nabla_\theta \times (\overline{\rho J_\theta \underset{\sim}{U} f}) = \overline{\zeta}_f \ . \tag{2.5}$$

The divergence and curl of global mass and energy transport were calculated from ECMWF IIb data using finite difference solutions of Eqs. (2.4) and (2.5) on a regular 3.75 degree latitude-longitude grid by successive over-relaxation. For details on the lower boundary condition, numerical methods and tests of convergence used to calculate transport potential, see Zillman (1972), Townsend (1980) and Johnson and Townsend (1981).

Throughout this study a convention after Lorenz (1955a) will be used to specify the atmosphere structure at the earth's surface and "underground". In this convention, it is assumed that all isentropic surfaces form continuous quasi-spherical surfaces which completely circumscribe the earth and that J_θ equal to $\partial z/\partial\theta$ is non-negative everywhere. In regions where the potential temperature of the particular isentropic surface is less than the earth's surface potential temperature, the isentropic surface itself, while termed underground, is assumed to be "contiguous" with the earth's surface. Thus, there are three classes of isentropic surfaces -- one class with $\theta < \theta_{s_o}$ (the minimum value of θ observed at the earth's surface everywhere) where the surface is entirely underground -- one class with $\theta_{s_o} \leq \theta < \theta_{s_m}$ (the maximum value of θ observed at the earth's surface everywhere) where the surface is partially underground -- and one class with $\theta \geq \theta_{s_m}$ everywhere where there is no underground region. In this convention the following state conditions are consistently defined:

$$z(\lambda,\phi,\ \theta \leq \theta_s,\ t) = z(\lambda,\phi,z_s,t) \tag{2.6a}$$

$$p(\lambda,\phi,\ \theta \leq \theta_s,\ t) = p(\lambda,\phi,z_s,t) \tag{2.6b}$$

$$\psi(\lambda,\phi,\ \theta \leq \theta_s,\ t) = \theta\pi(\lambda,\phi,z_s,t) + \phi(\lambda,\phi,z_s,t) \tag{2.6c}$$

$$\underset{\sim}{U}(\lambda,\phi,\ \theta \leq \theta_s,\ t) = \underset{\sim}{U}(\lambda,\theta,z_s,t) = 0 \tag{2.6d}$$

$$\dot{\theta}(\lambda,\phi,\ \theta \leq \theta_s,\ t) = \dot{\theta}(\lambda,\phi,z_s,t) \tag{2.6e}$$

Through this convention, all variables are piecewise continuous at the earth's surface. In the underground region, all vertical derivatives of state variables vanish with the exception of the Montgomery Stream Function and temperature, which are expressed as linear functions of potential temperature by

$$\psi(\lambda,\phi,\ \theta \leq \theta_s,t) = \theta\pi(\lambda,\phi,z_s,t), \tag{2.7a}$$

$$T(\lambda,\phi,\ \theta \leq \theta_s,t) = c_p^{-1}\theta\pi(\lambda,\phi,z_s,t) \tag{2.7b}$$

This convention satisfies the hydrostatic assumption in that within the underground region

$$\rho J_\theta(\lambda,\phi,\theta < \theta_S,t) = -\frac{1}{g}\frac{\partial}{\partial\theta}[p(\lambda,\phi,z_S,t)] = 0, \tag{2.8}$$

and

$$\frac{\partial\psi}{\partial\theta}(\lambda,\phi,\theta < \theta_S,t) = \pi(\lambda,\phi,z_S,t) . \tag{2.9}$$

At the earth's surface ρJ_θ is specified by its atmospheric value. Although this convention introduces a discontinuity in ρJ_θ across the earth-atmosphere interface, with the use of the mass-weighted average fields of transport processes may be temporally or zonally averaged and the interchange of temporal averaging and spatial differentiation may be accomplished without deleterious effect. With regard to lower boundary conditions for convective transport processes, note that with the implicit assumption of a no slip condition, the mass transport, $\rho J_\theta \underset{\sim}{U}$, and the transport of an arbitrary property, $\rho J_\eta \underset{\sim}{U} f$, vanish both at the earth's surface and underground. The diabatic mass transport, $\rho J_\theta \dot{\theta}$ vanishes underground, however, at the earth's surface it remains finite. The finite value for $(\rho J_\theta \dot{\theta})(\lambda,\phi,z_S,t)$ is necessary to account for the condition that a material element at the earth's surface moves relative to the isentropic structure as it receives or losses heat. The vertical integral of the mass continuity equation remains source free, however, since the time dependent boundary term $[(\rho J_\theta)\delta\theta_S/\delta t]_{\theta_S}$ from the vertical integration of the tendency of ρJ_θ cancels $(\rho J_\theta \dot{\theta})_{\theta_S}$ upon equating $\delta\theta_S/\delta t$ equal to $\dot{\theta}(\lambda,\phi,z_S,t)$ under the no slip boundary condition.

From the isentropic mass continuity equation averaged over a sufficient time interval to invoke the steady state assumption, and use of Eq. (2.4), the divergence of the isentropic mass transport is given by

$$\delta = \nabla_\theta^2 \overline{X}_\rho = \nabla_\theta \cdot \overline{(\rho J_\theta \underset{\sim}{U})} = - \frac{\partial}{\partial\theta} \overline{(\rho J_\theta \dot{\theta})} \quad . \tag{2.10}$$

For a simple two-layered model within regions of heating, horizontal mass convergence in lower layers and divergence in upper layers is balanced respectively by increasing upward mass flux in lower layers and decreasing upward mass flux in upper layers. The conditions are reversed within regions of cooling. Thus, this time-averaged relation links the irrotational component of isentropic mass transport directly to differential heating and requires the spatial separation between regions of convergence and divergence within isentropic layers be identical to the distance between planetary scale heat sources and sinks. Since the irrotational component of isentropic mass transport includes geostrophic and ageostrophic components (Johnson and Downey, 1975a), the link between differential heating and the irrotational component of mass transport is direct at all latitudes. Hence, the irrotational component, $\nabla_\theta \overline{X}$, expressed by Eq. (2.10) constitutes the thermally forced component of the mass transport and the isentropic mass continuity relation is fundamental to the isolation of thermodynamically forced global mass circulations.

The implications of this last statement deserve further consideration. The only coordinate system within which vertical mass flux occurs solely through heating is isentropic coordinates. As such, the isentropic vertical mass flux is related to the mean quasi-horizontal mass circulation as it is resolved through analyses within isentropic coordinates. Analysis of thermodynamically forced mass circulation within other coordinate systems requires use of the energy equation. In these coordinate systems, the vertical and horizontal components of both the mass and energy transport are linked implicitly to both adiabatic and diabatic processes. The forcing of the mass circulation by sources and sinks of momentum is also linked jointly to adiabatic and diabatic processes. Thus, there is no means to isolate explicitly a thermally forced component of the mass circulation that is uniquely linked to the planetary scale of monsoonal circulations. If such a mass circulation were isolated by some means, it would need to be identical with the isentropic mass circulation, since its isolation would require the same methods and constraints used in isentropic analysis merely from the uniqueness of the governing equations of atmospheric motion.

As one example of the differences between energy transport in isentropic and other coordinates, it is an established result that the isobaric energy transport by geostrophic motion in other than tropical latitudes is relatively independent of the mean isobaric mass circulation itself. Isobaric convergence of geostrophic energy exchange, which occurs largely by transient components of enthalpy exchange in extratropical latitudes, is considered to modify the diabatic forcing of the time-averaged circulation (Hoskins, 1983). This effective heat source forces the vertical branches of the Ferrel circulation within the zonally-averaged circulation (Kuo, 1956; Lorenz, 1967). In constrast, in isentropic coordinates the time-averaged geostrophic mode of mass and energy transport may be viewed as a direct response to the planetary scale differential heating. One of the consequences is that the scales of time-averaged mass circulations in non-isentropic coordinate systems differs from the larger time and space scales of the monsoonal circulations as revealed by isentropic analyses (Johnson and Downey, 1975;

Johnson and Townsend, 1981). Another consequence is that views of the forcing of these mass circulations also differ (Gallimore and Johnson, 1977a and 1981a; Johnson, 1985a).

A. EMPIRICAL EVIDENCE FOR GLOBAL MONSOONAL CIRCULATIONS

3. The global distribution of heat sources and sinks

While recognizing the validity of Halley's (1686) claim that monsoons are caused by differential heating between land and ocean in conjuction with the seasonal variation of incident solar energy, prior to the Global Weather Experiment (GWE) atmospheric scientists had not been able to study the basic cause of monsoonal circulations strictly from observational evidence of the atmosphere's heat sources and sinks. It is important to establish, insofar as possible, the physical basis for monsoonal circulations from observational evidence of heat sources with the belief that someday the atmosphere will be observed with sufficient accuracy to follow the behavior of its major monsoonal circulations.

The reasons for presenting filtered heating distributions with wavelengths greater than 10,000 km apart from the more detailed distributions shown by Johnson et al. (1986), are: 1) to focus attention on the planetary features of the heating distributions that are basic to summer and winter monsoonal flows, 2) to identify more clearly the impact of continental land masses and, 3) to contrast some key differences between summer and winter monsoonal circulations for land and ocean hemispheres. These results lay the foundation for the following analysis and discussion of isentropic mass and energy transport associated with global monsoonal circulations.

3.1 Global heating distributions for wavelengths greater than 10,000 km

The heating distributions from wavelengths greater than 10,000 km for the four months of January, April, July and October 1979 are presented in Figs. 2a through d. These filtered distributions were prepared through a Fourier

Fig. 2

representation of the heating distribution by truncating wavelengths less than 10,000 km (Wei et al., 1983; Johnson, 1984). Meridional filtering was then applied through the combination of six passes of a low pass (2,3,2) filter and one pass of an inverse (-1,5,-1) filter. The response of selected wavelengths to this combination of filtering is shown in Table 1.

Mass-weighted zonal averaging inherently smooths data in proportion to the meridional distribution of mass, with the smoothing tending to be proportional to the cosine of the latitude. The greatest smoothing occurs in tropical latitudes, while the least occurs in polar latitudes. In recognition of the smoothing by averaging, the meridional filtering that was discussed in the preceeding paragraphy was made proportional to the absolute value of sin ϕ.

Inspection of the heating distributions for January and July confirm that a distinct planetary scale for heat sources and sinks is related to the distributions of continents and oceans. As expected from earlier results (Wei and Johnson, 1981; Wei et al., 1981, 1983; Johnson and Wei, 1985, Johnson et al., 1986), the most prominent feature of the heating distributions for the four months is the region which migrates with the sun from just northeast of Australia in January to southeast Asia in July, after which the path of migration reverses direction. While other distinct centers or belts of heating that migrate with the sun can also be identified over Africa and the Americas, the scale and integrity of this regional feature of Asia, Australia and the Maritime Continent dominates the other features. This primary heat source is best developed in both areal extent and intensity in January and July, the summer and winter seasons. For the planetary scale, the absolute maximum of heating in excess of 3 K day^{-1} is located within this feature east-northeast of Australia in January. In July, this region of heating with a maximum greater 2 K day^{-1} is now centered over southeast Asia. While this maximum rate is less than in January, the areal extent of this feature of heating in July exceeds the areal extent in January.

The contrast of heating associated with the distributions of continents and oceans is critically important to the forcing of monsoonal circulations. A comparison of the heating distributions among all four months verifies that this contrast in the Eastern Hemisphere is maximized during the winter and summer seasons. Note that the region of heating for the planetary scale in the Western Hemisphere is also best developed in January and July. In January, the region of heating covers all of South America with the exception of Chile and Argentina. In July, the area of heating covers all of North America and the northern one-third of South America.

At the planetary scale the heating over Africa primarily consists of a belt which migrates with the sun. The maximum heating rate within this belt occurs in April.

3.2 A comparision of heating distributions for "land" and "ocean" hemispheres

The relationship between global monsoonal circulations and land and ocean hemispheres yields some interesting results. During the Asian summer monsoon, the incident solar energy is maximized over a land hemisphere, while the maximum of net outward radiation occurs over an ocean hemisphere. During the Asian winter monsoon the incident solar radiation is maximized over an ocean hemisphere, while the maximum of net outward radiation occurs over a land hemisphere.

As far as the overall pattern in January with maximum solar energy incident on an ocean hemisphere, five distinct regions of heating constitute the major energy sources for the planetary scale at this time: two over subtropical oceanic regions of the Southern Hemisphere, one over land in the tropical-subtropical belt of the Southern Hemisphere and two over the oceans in extratropical latitudes of the Northern Hemisphere just east of the Asian and North American continents. See Fig. 2a. The dominant region of cooling occurs within a large contiguous area of the Northern Hemisphere including the North American and Asian continents and the Arctic Ocean. Cooling also occurs over the eastern portions of the North and South Pacific subtropical anticyclones and the extratropical latitudes of the Indian Ocean.

In July, with maximum solar energy incident over a land hemisphere only two primary regions of heating occur, the one over Asia which extends southeastward across the equator to the South Pacific convergence zone and the other over North America that also extends southward over the Pacific Ocean near Central America and the northern one-third of South America. See Fig. 2c. A third region elongated zonally is associated with the latent and sensible heating within cyclonic disturbances occurring in the Antarctic Convergence Zone of the southern Indian Ocean. However, the intensity and areal extent of this third region is much less than its counterpart over the Gulf Stream and Kuroshio Current of the Northern Hemisphere in January.

In July, the three distinct regions of cooling occur primarily over the oceans. The dominant region of cooling occurs over the Southern Hemisphere oceans within an extensive area that includes the South Atlantic, southern Africa, Australia, and the southern parts of the Indian and South Pacific Oceans. Distinct regions of cooling occur within the Pacific Oceans and a contigious region over the North Atlantic and Europe.

From a ocean versus land hemispheric perspective for the winter and summer seasons, in the Southern Hemisphere's summertime the number of major heat source regions is five, while in the Northern Hemisphere summertime it is two. In the Southern Hemisphere's wintertime the number of major heat sink regions is three, while in the Northern Hemisphere wintertime it is two. The dominance of and preference for the two major heat sources of the planetary scale to be concentrated over the continents of the Northern Hemisphere in July, in comparison with the four major oceanic heat source regions over the oceans and one over land in January, reflects differences induced by the effects of solar energy incident on a land versus ocean hemisphere. Likewise, the preference for the two major heat sinks of the planetary scale to be localized over continents of the Northern Hemisphere in January similarly reflects influences induced within a land hemisphere.

Net cooling in winter results from the dominance of emission of infrared energy over other diabatic processes within the polar anticyclonic circulations that persist over the continental land masses. The net heating over the Asian and North American continents in summer, as indicated by the mean heating within isobaric layers (Johnson et al., 1986), results from the combination of sensible heating of the low troposphere and latent heating of the mid and upper troposphere associated with summertime convection. These processes are not independent in the sense that the sensible heating of the low troposphere decreases static stability and stimulates the moist convection. Thus, in either season, an important effect of the continents is to concentrate the heating in summer and cooling in winter within their geographical boundaries.

One of the reasons for the differences in the winter season between the Northern and Southern Hemispheres is the strong east-west contrast of wintertime heat sources and sinks found within the Northern but not in the Southern Hemisphere. The one region of wintertime heating along the Antarctic Convergence Zone, while relatively weak, is also elongated zonally in comparison with the strong east-west heating contrast in the Northern Hemisphere. One might suggest that the areal extent and intensity of heat sources over oceans and heat sinks over continents within the Northern Hemisphere are independent of each other, but this is not the case. Climatologically, the oceanic heat sources within the Northern Hemisphere largely result from the movement of polar air masses within the westerlies from continental to oceanic regimes where the atmosphere is heated through the addition of sensible and latent energies. In view of the relatively unlimited heat capacity of oceans, the intensity and areal extent of the wintertime mid-latitude oceanic heat sources is largely determined by the overall intensity and extent of cold air advection from continental to oceanic domains. Sensible heating of the polar air masses within the lower troposphere decreases static stability, which with the baroclinity of the quasi-stationary jet stream of these regions leads to intense cyclogenesis over the Kuroshio Current and Gulf Stream. Substantial latent heating subsequently occurs downstream with the maturation of oceanic extratropical cyclones. Given the constraint that the areally integrated vorticity over a sphere vanishes, the prevelance of cyclonic vorticity that is maintained and destroyed within these oceanic regions must be accompanied by a maintenance and generation of anticyclonic vorticity within the environments of the oceans. In the case of the Northern Hemisphere oceans these environments are the Asian and North American continents. Anticyclogenesis over the continents develops and maintains the quasi-stationary polar anticyclones of these areas, supresses moist convection and cloudiness, and thus enhances radiational cooling. These processes are important for the development of the cold air masses and thus the east-west heating contrast which is induced by cold air advection over warm ocean currents within the westerlies.

The lack of large continental land masses within the Southern Hemisphere precludes the general development of extensive quasi-stationary polar anticyclones within extratropical latitudes and within which extensive extremely cold polar air masses would develop. Thus, wintertime cooling in the ocean hemisphere is less localized than in the land hemisphere.

For the atmosphere as a whole to maintain its quasi-equilibrium state, regions of atmospheric cooling occur in order to balance the regions of heating. Planetary scales of heat sources and sinks are present in both the Northern and Southern Hemispheres regardless of whether a hemisphere consists primarily of land or ocean. The existence of large continents within the Northern Hemisphere simply determines preferred locations and scales of cooling in winter and heating in summer for the land hemisphere. The lack of large continents in the Southern Hemisphere does not preclude planetary scales of heating and cooling which are in excess of 10,000 km. However, the key difference is that the simple inverse pattern of heating in summer and cooling in winter over the continental land masses of the Northern Hemisphere that force the global monsoonal circulations is not a simple inverse pattern in the Southern Hemisphere. The extensive regions of heating within the Southern

Hemisphere during its summertime do not become the same extensive regions of cooling during its wintertime. These differences were captured by the composite FGGE observing system and the ECMWF data analysis and assimilation model (First GARP Global Experiment (FGGE) Advisory Panel, 1985; WMO GARP Reports, 1985a, 1985b and 1986).

4. Global monsoonal circulations

Although the analyses of the heating distributions were filtered to reveal the systematic structure within wavelengths 10,000 km and greater, the results that follow for the transport of mass and energy have not been filtered. The relatively smooth fields result from temporal averaging and the relaxation used to determine potential and stream functions for the time averaged transport. As will become clear, the major planetary features are quite evident.

Monthly averaged rotational and irrotational components of mass transport from an analyses of the ECMWF Level III FGGE data are presented for the months of January, April, July and October. Although isentropic analyses of quasi-horizontal mass and energy transport have been completed for each 10K layer, the global exchange that is linked directly to differential heating is captured primarily by a two-layered representation of the irrotational mass transport. The layers to be presented for each of the four months are a lower isentropic layer, the 300-310 K layer, and an upper layer, the 340-350 K layer (to be called the 300 K and 340 K layers). However, to describe the Asian winter monsoon, the irrotational component of mass transport for four additional 10 K layers between 260 K and 300 K will be presented for January. Similar analyses for NMC's Level III FGGE data have been presented by Johnson and Townsend (1981) and Johnson et al., (1981, 1985).

4.1 The rotational component of mass transport

Global analyses of the rotational component of the isentropic mass transport within the 300 K and 340 K layers for the four months are presented in Figs. 3 and 4, while the corresponding analyses for the irrotational component are presented in Figs. 5 and 6. The analyses of the irrotational component also portray isobars of the isentropically time-averaged pressure in order to provide a frame of reference for the exchange processes within the isentropic layer. The 300 K layer is located in the low troposphere of tropical regimes and the mid and high troposphere of polar latitudes. Since the thermal wind in isentropic coordinates is proportional to the isentropic gradient of pressure, the greatest slopes of the isobaric topography occur within middle latitudes. The 340 K surface is located within the upper troposphere of tropical latitudes and lower stratosphere of polar latitudes.

The dominant features in Figs. 3 and 4 are westerlies in middle latitudes and anticyclonic circulations in subtropical latitudes. At the 300 K layer westerlies persist within extratropical latitudes throughout the year, with the exception of the month of July in the Northern Hemisphere. In July the subtropical anticyclones over the Pacific and the Atlantic oceans of the Northern Hemisphere are particularly well developed. Consequently, westerlies are weak and are shifted poleward over North America, while over Asia no systematic westerlies occur.

An interesting feature of the westerlies in the 340 K layer for the spring, fall and winter seasons is an equatorial displacement of the zone of maximum gradient of the stream functions relative to the zone of maximum gradient at 300 K (Figs. 3 and 4a, b, and d). This displacement shows that the rotational component of mass transport for the subtropical jet stream occurs at higher isentropic layers and lower latitudes than that for the polar jet stream.

During January (Fig. 3a), pronounced anticyclonic circulations within the 300 K layer are located in the Southern Hemisphere over the Indian, Pacific and Atlantic Oceans. Within the Northern Hemisphere subtropical anticyclonic circulations are located over the Sahara, Saudi Arabia, the western Pacific and the Caribbean. The westerlies in both hemispheres are well developed. The waves of the westerlies in the Northern Hemisphere portray relatively large amplitudes in contrast to the nearly zonally directed westerlies of the Southern Hemisphere. Because of relatively low elevation of the 300 K layer within tropical latitudes, easterlies generally prevail. The strongest westward mass transport occurs on the equatorward side of the Southern Hemisphere subtropical anticyclones and also over the Northern Hemisphere subtropical latitudes of the Pacific Ocean.

During April (Fig. 3b), the anticyclonic circulations of the Southern Hemisphere within the 300 K level are relatively weak, while the anticyclonic circulations of the Northern Hemisphere are distinct. Note the intensification of the anticyclonic circulations over the Near East and the Pacific. The westerlies within the Northern Hemisphere are reduced in intensity and oriented more zonally relative to the pronounced wave structure of January. The westerlies in the Southern Hemisphere have intensified considerably relative to January.

During July (Fig. 3c), pronounced anticyclonic circulations within the 300 K layer occur over the north Atlantic and the north Pacific as well as over the subtropical latitudes of the south Pacific. Within the equatorial regions of the low troposphere over the Indian Ocean, the anticyclonic circulation in the 300 K layer reflects the rotational mode of cross equatorial transport of mass and energy by the Somali jet towards the regions of convection over southeast Asia and India. The westerlies during this winter season in the Southern Hemisphere are particularly well developed. The difference between max-min isopleths near 120°W over the Pacific Ocean of the Southern Hemisphere is twelve units in this winter season, while the max-min difference in the winter season near 150°E over the western Pacific in the Northern Hemipshere is only eight units. This difference reflects a relative increase of 50% in the rotational mass transport by westerlies within the Southern Hemisphere over the mass transport in the Northern Hemisphere. As noted earlier, the westerlies of July within the Northern Hemisphere have almost completely disappeared in this layer except over higher latitudes of the Western Hemisphere.

During October (Fig. 3d), anticyclonic circulations within the 300 K layer are well developed in both hemispheres. In the Northern Hemisphere anticyclonic circulations are located over northern Africa-southern Asia, the Pacific Ocean and the Atlantic Ocean. In the Southern Hemisphere an anticyclonic circulation is located over the Pacific Ocean. An elongated anticyclonic circulation extends from South America across south Africa to Australia. The westerlies of the Northern Hemisphere have intensified greatly over their summertime magnitude while the decrease of the westerlies in the Southern Hemisphere over the winter maximum is relatively modest.

In January (Fig. 4a), anticylonic circulations of the 340 K layer are located over or near the major centers of heating in the tropics: the largest one is centered just east of the Philippines, one is over east Africa and one is over western Brazil. In April (Fig. 4b), the anticyclonic circulation of the western Pacific now centered over the Philippines has decreased in intensity. Weak anticyclonic circulations are located over the northern coast of South America and the Gold Coast of Africa. The eastward mass transport of the subtropical jet streams within both hemispheres is zonally oriented and symmetrically located about the equator. During July (Fig. 4c), two anticyclonic circulations straddle the equator within the Eastern Hemisphere, one centered over northern Australia and the other one stretching from the East China Sea to northwestern Africa. The relatively strong gradient of the stream function between these two anticyclonic circulations at the 340 K level is evidence of the Tropical Easterly Jet that has a maximum just south of India. Its entrance region is located near the mid-Pacific, while its exit region lies over the mid-Atlantic. In October (Fig. 4d), two anticyclonic circulations exist in the western Pacific, one near 25°N and the other near 10°S. The westerlies in the Northern Hemisphere meander meridionally, while the more pronounced westerlies of the Southern Hemisphere are zonally oriented.

4.2 The irrotational component of mass transport

In Figs. 5 and 6, the patterns of the mass transport potential functions for the 300 and 340 K layers isolate the time averaged irrotational components of mass transport associated with differential heating and the atmosphere's primary two-layered response. The distance between extrema, which identify the primary regions of convergence and divergence for the planetary scale, reveals monsoonal circulations that are global in nature. The two-layered structure of the mass circulations is indicated by the inverse relation between the potential functions of the lower and upper isentropic layers.

Fig. 7

Fig. 7 presents the irrotational mass transport for the 260 K, 270 K, 280 K and 290 K layers of the Northern Hemisphere and identifies a mass circulation occurring within higher latitudes associated with monsoonal-like flow between the continental land masses and warm surrounding oceans. In the discussion regarding the irrotational components of isentropic mass transport, primary emphasis will be given to the extrema which identify the scale of the monsoonal circulations for the various continents and establish the link of these circulations with differential heating.

In the primary heat source region over the western Pacific and southeast Asia, the mass transport during all months is convergent within the 290, 300, and 310 K layers of the low troposphere and divergent in the layers between 330 and 360 K layers of the upper troposphere. The level of non-divergence is located in the 320 K layer. The vertical mass transport in this region constitutes the diabatically forced upward branch of the monsoonal circulation that exists within the Eastern Hemisphere throughout the annual cycle. In all four months the maximum of the potential functions at the 300K level generally coincides with the regions of maximum vertical mass transport. Compare Figs. 2 with Fig. 5. This center of the ascending branch of this Asian monsoonal circulation shifts seasonally in conjunction with the movement of the region of strong diabatic heating discussed earlier.

Within the general area of the maxima in the 300 K layer potential function to the eastnortheast of Australia in January (Fig. 5a), elongations of maximum curvature to the southeast and to the west are coincident with the bands of heating within the South Pacific Convergence Zone and a region of heating over northeast Australia. A major feature of the January 300 K and

340 K potential functions is the elongated region of low level convergence and upper level divergence associated with the ITCZ over the equatorial region of the Indian Ocean.

In the 340 K layer (Fig. 6a), with the exception of an elongation over the ITCZ of the Indian Ocean, the quasi-circular nature of the potential functions in the upper layers, which span radially 6,000-7,000 km from this region, indicates that the intensity of both meridional and zonal irrotational components of transport in the upper troposphere are equally pronounced. The maximum in the gradient across the Pacific within tropical-subtropical latitudes indicates a relatively strong eastward directed branch across the Pacific within upper layers that is undoubtedly part of the Walker circulation of the region. Meridional mass transport from the Southern Hemisphere to the Northern Hemisphere is also indicated within the 340 K layer in January by the relatively strong poleward directed gradients north of the Indian Ocean ITCZ minimum, the western Pacific minimum and the Brazil minimum.

During January, relative maxima at 300 K and minima at 340 K also exist over Brazil and South Africa. The mass convergence into the lower layer and divergence in the upper layer is associated with the heat sources and vertical branches of the monsoonal mass circulations over these continents. Descending branches from these large scale monsoonal mass circulations are interrelated with the Asian winter monsoon. Primary centers of high level mass convergence and low level mass divergence in the Northern Hemisphere (Figs. 5a and 6a) are also located over the heat sinks of northern Africa and southern Asia. In the Southern Hemisphere a center of low level mass divergence is located just west of Australia. Within the Western Hemisphere, the mass transport in January at

300 K is divergent over the eastern Pacific and the Caribbean. In the upper level the pattern of the potential function indicates upper level convergence over above low level divergence.

In the Asian summer monsoon, the primary region of low level mass convergence at 300 K and upper level mass divergence at 340 K (Figs. 5c and 6c) shifts to the heat source region of southeast Asia, Indonesia and India. The structure for the mass circulation is associated with the deep and extensive convection that constitutes the upward vertical branch of the Asian summer monsoon. The primary downward branch occurs within an extremely large region extending from the Atlantic Ocean across south Africa and the Indian Ocean. This, of course, is due to the extensive region of cooling over the same region that was discussed the previous section of this summary. Within the 300 K layer the strong gradient between Madagascar and southeast Asia reflects the irrotational component of the strong cross equatorial mass transport that supplies energy for the moist convection within the Asian summer monsoon. A reverse gradient at the 340 K layer is evidence of a return branch of the irrotational component of mass transport from the ascending branch over southeast Asia to the descending branch within the Mascarene High.

In the Western Hemisphere, low level convergence and upper level divergence occurs over Central America and the southern United States. The ascending branch of the mass circulation within this heat source region is part of the monsoonal circulation in the Western Hemisphere. An analysis that illustrates the character of the North American monsoon has recently been completed by Tang and Reiter (1984).

The limited areal extent of the North American monsoonal circulation relative to its Asian counterpart must be associated with the limited size of the North American land mass and the Atlantic Ocean relative to the Asian land mass and Pacific Ocean. In comparing the relative intensities of the irrotational components of the isentropic mass circulations, the intensity of the isentropic mass transport between a heat source and sink will depend on both the differences in intensity between heating and cooling and areal extent of land masses. Differences in mass transport due to differential heating are obvious, while differences due to areal extent are less obvious. To illustrate this difference, consider a circular area of heating. The areally integrated diabatic vertical flux out of an isentropic layer is proportional to the square of the radius while the horizontal flux through the lateral boundary is proportional to the radius itself. In order for steady state conditions to prevail in an area with uniform vertical divergence of the diabatic mass flux, the intensity of the lateral mass transport through the boundary of the circular area must increase linearly with increasing radius. Viewed within the context of the Asian and North American monsoonal circulations, it is clear that the differences in intensities of the irrotational component of the isentropic mass transport between the two regions can be largely attributed to the differences in the size of the two continental land masses in the Northern Hemisphere.

The transport potential functions for April (Figs. 5b and 6b) and October (Figs. 5d and 6d) also reveal thermodynamically forced planetary scale mass circulations which are somewhat less intense than their counterparts in winter and summer seasons. Since the continent-ocean surface temperature contrasts associated with differential heating and surface energy balance are minimized during these seasons, the intensity of global monsoonal circulations are reduced.

A principle component of forcing of the global monsoonal mass circulations is moist convection. During the months of April and October low level mass convergence indicated by the mass transport potential functions tend to be oriented along the ITCZ near the equator, while in the months of January and July centers of low level mass divergence are displaced away from the equator toward higher latitudes. This distribution of divergence and convergence leads to the condition that the ageostrophic Hadley circulation tends to be symmetric with respect to the equator during April and October (Townsend and Johnson, 1981).

One interesting difference between the 300 K mass transport potential for April and October (Figs. 5b and d) occurs over eastern Asia and the western Pacific. In April the mass transport is convergent over China and Siberia, while in October the mass transport is divergent. This is likely associated with differences in the sensible heating between the spring and fall seasons. In spring, an upward heat flux from the earth to the atmosphere is associated with the static instability created by more rapid heating of a barren cold land mass than the overlying cold air. In the fall season, the net sensible heat flux is likely neglible, since static stabilization occurs through the more rapid cooling of a warm land mass than the overlying warm air. Hence, the differing responses of a cold land mass and atmosphere to incoming solar energy in early spring in comparsion with a warm land mass and atmosphere in fall impact the planetary scale mass circulation. The net effect of this difference in low tropospheric heating is to induce a difference in phase longitudinally of the low level mass convergence. Note that the region of maximum low level mass convergence associated with heating over eastern Asia in April is shifted downstream just off the east coast of Asia in October.

One important feature of the thermodynamically forced mass circulation of the Northern Hemisphere during winter is a distinct mass circulation within the 260-300 K isentropic layers of higher latitudes (Johnson and Townsend, 1981; Johnson et. al., 1985). This mass circulation is associated with the exchange of polar air between continential and oceanic regions. Within the westerlies, polar air moves equatorward and eastward from anticyclonic circulations over Asia and North America. As it moves over oceanic regions, the polar air is modified by transfer of sensible and latent energies from the ocean to the atmosphere. The mass transport within the 260 and 270 K layers (Figs. 7a and b) is convergent over the Gulf Stream and Kuroshio Currents and divergent over polar latitudes. Within the 280 K and 290 K layers (Figs. 7c and d), the mass transport is divergent over the Gulf Stream and Kuroshio Currents, while it is convergent over polar latitudes. The upward vertical mass flux within the region of low level convergence and upper level divergence is due to the sensible and latent heating over the mid-latitude oceanic regions just east of the Asian and North American continents. These regions are linked with downward vertical mass transport due to infrared cooling in association with upper level mass convergence and low level divergence over polar latitudes of the Asian and North American continents. Within the 280-300 K layer mass transport is from heat source to heat sink, within the 260-280 K layer it is from heat sink to heat source.

While most would not classify these mass circulations to be monsoonal based on momentum considerations and the alternation of the direction of the wind field within baroclinic waves (Ramage, 1971), from thermodynamic considerations it is difficult to draw a distinction between these circulations of higher latitudes and the ones of lower latitudes. The potential functions analyses show that part of the polar air leaving the Asian continent during winter streams eastward over the Pacific Ocean and becomes part of the mid-latitude polar circulation, while part streams southward over the South China Sea towards equatorial latitudes and becomes part of the winter monsoonal mass circulation. The polar origin of these air masses over Asia is common to both trajectories of the flow. The time scale for the alternation of wind direction in higher latitudes is determined by the passage of cyclonic and anticyclonic in disturbances in mid-latitudes, while the time scale for the alternation of wind direction in lower latitudes, which is traditionally classified monsoonal flow, is seasonal. However, viewed within the context that the polar cyclonic and anticyclonic disturbances are simply the means to accomplish the energy transport needed to satisfy the global scale of differential heating, the time scale of this response is identical with the seasonal time scale determined from classification of monsoonal circulations by surface wind direction.

Insight into the nature of planetary scale monsoonal exchange begins with the recognition that distinct physical mass circulations exist within the time-averaged flow that are uniquely coupled with differential heating. The time-averaged isentropic mass transport is determined by the time averaged potential and stream functions. As such, the combination of the time-averaged potential and stream function fields determines the streamlines of time-averaged mass transport. Since the time-averaged potential and stream functions individually determine the vector fields $(\overline{\rho J \underset{\sim}{U}})_\chi$ and $(\overline{\rho J \underset{\sim}{U}})_\psi$, streamlines for the irrotational and rotational components of the time-averaged mass transport are separately and uniquely determined. Streamlines for the rotational motion remain independent of the diabatic mass flux in view of the nondivergent nature of this component. In contrast, the irrotational component of the motion and the flow streamlines defined by this component are directly related to vertical diabatic mass flux relative to isentropic surfaces. Thus, the combination of the irrotational component of motion and diabatic mass flux uniquely determines a three dimensional structure for the thermodynamically coupled mass circulation in isentropic coordinates. This structure is superimposed on the more intense rotational component of mass transport. Within this structure, however, streamlines should not be interpreted as trajectories of the time-averaged flow even though for a steady state field streamlines and trajectories would be coincident. The analyses are averages of non-steady circulation for which the streamlines of each event are not trajectories. Thus the streamlines of mass transport from time-averaged potential and stream functions cannot determine a time-average of Lagrangian trajectories.

Evidence from these and earlier diagnostic analyses verify that global isentropic monsoonal mass circulations exist within the atmosphere which bear a one-to-one correspondence with the planetary scale of differential heating. In actuality, the heat sources and sinks were determined with GWE data from the isentropic mass circulations. However, as pointed out previously (Johnson, 1985a and b), the ability to isolate isentropic mass circulations using the GWE data was only due to the fact that the atmosphere is continually responding to heat sources and sinks in the form of planetary scale isentropic mass circulations which were captured by the GWE information. If the energy balance of heat sources and sinks were of a local nature, such planetary scale mass circulation would not be manifested.

4.3 Mean meridional circulations

With the systematic heating in low latitudes and cooling in high latitudes, the structure of differential heating demands that mean meridional isentropic mass circulations exist which span each hemisphere. An upward branch in tropical latitudes must be associated with net heating while a downward branch in polar latitudes must be associated with net cooling. The quasi-horizontal branches from equatorial to polar latitudes at upper levels and from polar to equatorial latitudes in lower levels are required by mass continuity.

Using diagnostics of the zonally-averaged circulation from GWE data, isentropic Hadley circulations have been isolated which span each hemisphere (Townsend, 1980; Townsend and Johnson, 1981 and 1985; Johnson, 1985a). Isentropic Hadley circulations have been calculated from the isentropic geostrophic mass transport (Henderson, 1971) and from modeling the zonally averaged heating distributions (Johnson and Dutton, 1969; Zillman, 1972; Otto, 1974; Dutton, 1976a; Gallimore and Johnson, 1981b; Otto-Bliesner and Johnson, 1982; and Zillman and Johnson, 1985). In these analyses, the isentropic mass stream function for the zonally averaged mass transport (Townsend, 1980; Townsend and Johnson, 1985) is defined by

$$\frac{\partial \Psi_\theta}{\partial \theta} = -2\pi \ a \cos \phi \ \overline{\rho J_\theta}^{\lambda,t} \ \hat{v}^{\lambda,t} \tag{4.1}$$

and

$$\frac{\partial \Psi_\theta}{\partial \phi} = 2\pi \ a^2 \cos \phi \ \overline{\rho J_\theta}^{\lambda,t} \ \hat{\dot{\theta}}^{\lambda,t} \quad . \tag{4.2}$$

With appropriate boundary conditions the isentropic stream function is defined by

$$\Psi_\theta(\phi,\theta) = -2\pi \ a \cos\phi \int_{\theta_B}^{\theta} \overline{\rho J_\theta}^{\lambda,t} \ \hat{v}^{\lambda,t} \ d\theta \quad . \tag{4.3}$$

This definition is analogous to the definition for isobaric coordinates (Palmen and Vuorela, 1963; Newell et al., 1970; Starr et al., 1970; Oort and Rasmusson, 1970) given by

$$\frac{\partial \Psi_p}{\partial p} = \frac{2\pi \ a \cos\phi}{g} \overline{v}^{\lambda,t} \tag{4.4}$$

and

$$\frac{\partial \Psi_p}{\partial \phi} = \frac{-2\pi \ a^2 \cos\phi}{g} \overline{\omega}^{\lambda,t} \tag{4.5}$$

With appropriate boundary conditions, the stream function is determined by

$$\Psi_p(\phi,p) = \frac{2\pi \ a \cos\phi}{g} \int_{p_s}^{p} \overline{v}^{\lambda,t} \ dp \quad . \tag{4.6}$$

Through these definitions the zonally, temporally averaged meridional mass fluxes, $\overline{\rho J_\theta v}^{\lambda,t}$ and $\overline{v}^{\lambda,t}$, are respectively proportional to the vertical derivative of the isentropic and isobaric stream functions while the zonally, temporally averaged vertical mass fluxes, $\overline{\rho J_\theta \dot{\theta}}^{\lambda,t}$ and $\overline{\omega/g}^{\lambda,t}$, are proportional to their meridional derivatives.

The results for the isentropic mass stream function from ECMWF analyses for January and July 1979 are presented in Figs. 8a and 9a. These results show direct isentropic Hadley circulations with upward mass flux in tropical-subtropical latitudes and downward mass flux in polar latitudes. The meridional branches are hemispheric in extent with poleward mass transport in upper layers and equatorward mass transport in lower layers. The dominance of the Hadley cell in the winter hemisphere, which is associated with the greater differential heating of this season, is quite evident. A summertime Hadley circulation for the Northern Hemisphere is not evident (Fig. 9a). The meridional slope of both branches of the wintertime cell from tropical to polar latitudes is associated with net cooling in upper isentropic layers of the poleward branch and with net heating in lower layers of the equatorward branch. Figs. 8 and 9

In January (Fig. 8a), two secondary isentropic mass circulations are embedded within the Northern Hemispheric cell, one in the subtropics and one in the extratropics. The low latitude cell with downward mass transport in the subtropics is associated with radiational cooling and lack of convective and latent heat release in the trade wind regions of these latitudes. The high latitude cell is associated with upward diabatic mass flux in the extratropics due to sensible and latent heating in baroclinic circulations over the North Atlantic and Pacific Oceans. In the Southern Hemisphere an embedded secondary cell in extratropical latitudes is not evident in July, at least to the extent that the Northern Hemispheric cell stands out in January. While a belt of heating associated with cyclonic disturbances exists over the Southern Ocean surrounding the Antarctic continent, the combined latent and sensible heat sources in these regions is less intense and less developed in vertical extent than the sources over the Kuroshio Current and Gulf Stream. This dissimilarity stems from the difference in the differential heating associated with land and ocean hemispheres and the strength of cold air advection from continents to oceans discussed earlier.

The isobaric results presented in Figs. 8b and 9b for January and July of the GWE are in general agreement with results of other studies of the isobaric zonally averaged circulation. Isobaric Hadley and Ferrel circulations are present in each hemisphere with the Hadley cell of the winter hemisphere being dominant. As in the results for the isentropic mass circulation, the Hadley cell of the Northern Hemisphere is virtually nonexistent in July.

The physical and mathematical relations between the isentropic and isobaric zonally averaged mass circulations will be discussed in a later section. However, in order to gain some insight at this point from the empirical evidence for the zonally-averaged mass circulations, the isentropic circulation for January and July, 1979 is now partitioned into geostrophic and ageostrophic components, given by

$$\overline{\rho J_\theta v}^{\lambda,t} = -g^{-1} \left[\overline{(\partial p/\partial\theta)v_g}^{\lambda,t} + \overline{(\partial p/\partial\theta)v_{ag}}^{\lambda,t} \right], \quad (4.7)$$

and compared with the isobaric circulation. The results presented in Figs. 8 and 9, c and d, show that an ageostrophic mode of isentropic mass transport dominates in low latitudes, while a geostrophic mode dominates in middle latitudes (Townsend and Johnson, 1981, 1985). Both of these modes are a direct response to differential heating since it is through the combination of these two modes whereby the thermally forced Hadley circulation exists. A comparison shows that the meridional scale of the ageostrophic isentropic mass circulations is common with the meridional scale of the isobaric circulations, which through the hydrostatic condition and zonal averaging are constrained to be solely ageostrophic (Townsend and Johnson, 1985).

The meridional scale of the geostrophic component of the isentropic mass transport is determined by the meridional extent of the active baroclinic wave regime. To gain physical insight into this mean meridional geostrophic mass transport, consider the structure of steady and amplifying baroclinic waves in mid-latitudes (Johnson, 1979; Townsend and Johnson, 1985). In the schematic of a steady baroclinic wave (Fig. 10a), the geopotential and potential temperature waves are in phase. In the zonal, vertical cross section extending from a lower isentropic surface, θ_ℓ, to an upper isentropic surface, θ_u, the midvalue isentropic surface θ_m (e.g., θ_m = 315 K) divides the hypothetical atmosphere into two layers. The dotted vertical lines designate trough and ridge locations that separate regions of poleward and equatorward meridional geostrophic motion. With a vertical scale of linear pressure, the relative positions of the isentropes show more mass $(-g^{-1}(\partial p/\partial\theta)d\theta)$ in the trough than in the ridge for the layer between θ_ℓ and θ_m. The mass distribution reverses in the layer above θ_m. Since the potential temperature and geopotential fields are in phase, the meridional mass transport $\rho J_\theta v$, as it is largely determined by its geostrophic component $\rho J_\theta v_g$, is symmetric about the trough line in each layer. Thus, the equatorward geostrophic mass transport to the rear of the trough exactly balances the poleward transport forward of the trough with the result that the mean meridional mass transport within each layer of the steady baroclinic wave vanishes.

In contrast, the systematic structure of an amplifying baroclinic wave (Charney, 1947; Eady, 1949) results in net poleward mass transport by the geostrophic mode in upper isentropic layers and net equatorward mass transport in lower isentropic layers (Fig. 10b). Above 315 K, poleward motion ahead of the trough is located within a deeper layer of mass than the equatorward motion in the rear of the trough. Below 315 K, the structure reverses. More mass moves equatorward in the rear of the trough than ahead of the trough. Consequently, the mean meridional mass transport within this two-layered structure of active baroclinic waves is poleward above 290 K and equatorward below 315 K. Hence, systematic variations of mass and meridional motion within an amplifying baroclinic wave provide for poleward and equatorward branches of a mean isentropic zonally-averaged mass circulation.

To emphasize this unique degree of freedom within the isentropic meridional mass circulations, consider that since $\overline{v_g}^{\lambda}$ is zero, the two components are given by

$$\overline{\rho J_\theta v}^{\lambda} = -g^{-1}\,[\overline{(\partial p/\partial\theta)' v_g'}^{\lambda} + \overline{(\partial p/\partial\theta) v_{ag}}^{\lambda}]\ , \tag{4.8}$$

where the isentropic zonal deviation in this case is defined by

$$f' = f - \overline{f}^{\lambda}\ . \tag{4.9}$$

This result shows that a zonally averaged geostrophic mass transport only occurs through a covariance of the deviations of mass and the meridional geostrophic component. As the schematic for the isentropic structure reveals, a systematic covariance exists between the covariance of ∂p/∂θ and the meridional geostrophic motion within amplifying baroclinic waves. In contrast, in isobaric coordinates the deviation of the hydrostatic mass ρJ_p equal to g^{-1} is identically zero. Thus, isobaric coordinates do not enjoy zonally averaged meridional geostrophic mass transport, at least within isobaric layers which do not intersect the earth's orography (Johnson and Downey, 1975a).

Some suggest that the geostrophic mode of mass transport is an eddy mode since the existence of this mode is associated with zonal deviations of two fields and systematic processes occurring in the atmosphere. However, if this is the only criterion used to define mean and eddy components of physical structure or processes, the criterion rest purely on elementary statistical consideration, such as one uses to prove that a given level of correlation is unlikely to occur by chance. Unless the decomposition into components is based on a well founded physical principle (i.e. some model of a physical process), the result provides little insight in the underlying physics except to support the notion that something systematic is occurring.

It is exactly for these reasons that the mass-weighted average was defined by Johnson and Downey (1975a) to be used in analyses of transport processes. It was introduced into transport theory by Reynolds (1894). In analyses of a physical system, one seeks to determine the systematic structure or processes that occur within a physical system. An average per unit mass that properly defines the mean of a physical property or process in a given region must be defined as the ratio of the integral of the quantity within the given region to the integral of the mass within the same region. This definition is the only one which will uniquely decompose the quadratic property of kinetic energy into the sum of two positive definite quantities.

This feature is essential in seeking equations for the time rate of change of average and deviation components of kinetic energy which are readily related to corresponding degrees of freedom in the exchange of mass and angular momentum.

In isentropic analyses of transport processes, geostrophic and ageostrophic modes are systematic processes which together satisfy the mass balance of the planetary scale. The existence of this systematic mass circulation supports the premise that a differentially heated atmosphere develops motion to transport energy from a heat source to heat sink. For the zonally averaged circulation, the mass circulation must span the hemisphere, and the energy transport by the mass circulation must extend from the tropics to polar regions. While the evidence up to this point has not identified just how the geostrophic mode of the isentropic mass circulation transports energy, the evidence does verify that systematic isentropic mass circulations exist within the baroclinic wave regime which are directly determined by the hemispheric scale of differential heating, not by the scale of the baroclinic waves. This suggests the wave regime is simply a response to the differential heating, within which the isentropic mass circulation occurs by both geostrophic and ageostrophic modes (Townsend and Johnson, 1981, 1985). The same mean mass circulation is also the primary degree of freedom through which the mean meridional energy transport occurs geostrophically in midlatitudes. Later, in a discussion of the forcing of the isentropic Hadley circulation, the geostrophic mass transport will be related to the forcing of the isentropic mass circulation, energy transport and the exchange of angular momentum by pressure torques.

5. Energy balance of global monsoonal circulations

5.1 The rotational component of transport of *total flow energy*

The rotational components of *total flow energy* transport for January and July in the 300 K and 340 K isentropic layers are presented in Fig. 11. Except for the different scaling for the rotational energy transport, the pattern in each of these figures is almost identical with the corresponding pattern for the mass transport (Figs. 3 and 4, a and c). Since the rotational components of either the isentropic mass or *total flow energy* transport are not linked directly to the planetary heat sources and sinks, the close agreement of isentropic mass and energy transport is remarkable, at least in the sense that such agreement does not occur in isobaric analyses.

Fig. 11

The primary purpose of showing results for both the rotational and irrotational components of both mass and *total flow energy* transport is to establish the remarkable similarity of the fields from empirical evidence. This remarkable similarity, which occurs in both high and low latitudes, indirectly suggests that the earth's rotation does not influence the capability of the isentropic mass circulation to transport *total flow energy* from heat sources to heat sinks. From an Eulerian viewpoint energy transport may occur through transient rotational modes which are independent of the mean mass circulation. If such modes exist, however, the remarkable similarity of the transport of mass and energy by mean irrotational components that is evident in the following section would be precluded.

5.2 The irrotational component of transport of *total flow energy*

The irrotational components of *total flow energy* transport for January and July in the 300 K and 340 K isentropic layers are presented in Fig. 12.

Fig. 12

Like the remarkable similarity between the rotational components of mass and energy transport, the pattern of irrotational energy transport in these figures is almost identical to the corresponding pattern for the mass transport. A few exceptions occur in regions of cols. In all cases, regions of mass and energy divergence as well as regions of mass and energy convergence are coincident with each other. This implies that the scales of the time-averaged mass and *total flow energy* transport within the isentropic stratification are identical. Equation (2.10) shows that the irrotational component of the mass transport is linked directly to differential heating. The similarities of the irrotational component of energy transport and the irrotational component of mass transport in individual layers provides empirical evidence that 1) the isentropic energy transport develops as a direct response to differential heating and 2) its exchange occurs through the isentropic mass circulation.

From a global perspective, the summer hemisphere receives an excess of energy while the winter hemisphere experiences a deficit of energy. This leads to strong energy transport from the summer to the winter hemisphere in conjuction with the the thermally forced isentropic mass circulations. For example, the 340 K potential function for energy transport in July (Fig. 12d) shows that the primary region of energy divergence occurs over Asia and the western Pacific in the Northern Hemisphere, while the primary region of energy convergence in the Southern Hemisphere is located over the Southern Atlantic, south Africa, and the Southern Ocean. The maximum in the gradient between these two regions provides evidence of the strong energy transport by the irrotational mode from the Northern to the Southern Hemisphere within the

340 K layer. Likewise, in the 300 K layer (Fig. 12b) the strong gradient between southeastern Asia and the South Atlantic-Africa-Southern Ocean is indicative of the strong energy transport by the irrotational mode from the Southern to the Northern Hemisphere. The net energy transport from the Northern to the Southern Hemisphere occurs because the upper branch of the isentropic mass circulation transports more energy from the heat source to the heat sink than the lower branch returns from the heat sink to the heat source. Through this means, the quasi-steady balance is maintained by a net transport of energy from the Northern Hemisphere to the Southern Hemisphere.

The primary source of energy that forces the upward branch of the mass circulation in the Asian summer and winter monsoons is latent heating through moist convection. Sensible heating also plays an important role both directly and indirectly in the summer and winter monsoons. In the winter monsoon, the sensible heating over the South China Sea increases the equivalent potential temperature of air that eventually becomes part of the upward diabatic mass circulation that develops over the warm waters of the tropical western Pacific and Indonesia (Fig. 12a and c). In the summer monsoon, the sensible heating over the Tibetan Plateau and Asia, combined with the latent energy transport into this region, supplies the energy to force the upward diabatic branch in the summer monsoon (Figs. 12b and d). The energy gained through latent and sensible heating in these heat source regions is transported within upper isentropic levels to heat sink regions. In heat sink regions warming through subsidence offsets the cooling of the atmosphere by infrared emission. In lower latitudes, this warming occurs within descending branches of isentropic Walker circulations. In higher latitudes, it occurs within descending branches of the isentropic Hadley circulations that span the hemisphere.

5.3 Stationary and transient modes of isentropic energy transport in relation to heat sources and sinks

In the discussion regarding the rotational component of *total flow energy* transport it was pointed out that from an Eulerian viewpoint degrees of freedom exist for transient modes in isentropic coordinates to accomplish the energy transport from heat source to sink. In isobaric perspectives of energy balance, the transient mode of energy transport in association with quasi-geostrophic motion is of prime importance in the satisfaction of energy balance (Lau, 1978, 1979a, 1979b; Holopainen, 1983). While the empirical evidence up to this point is in agreement with the physical reasoning that scales of isentropic mass transport correspond with scales of energy transport, explicit determination of transient modes of energy exchange have not been presented.

In order to ascertain empirically whether or not *total flow energy* transport occurs by transient modes that are independent of the mean isentropic mass circulation, the vertically integrated divergence of the isentropic energy transport vector, $\overline{\rho J_\theta \underset{\sim}{U} \nu}$, is divided into stationary and transient components, given by

$$\int_{\theta_{s_o}}^{\theta_T} \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} \nu})\, d\theta = \int_{\theta_{s_o}}^{\theta_T} \nabla_\theta \cdot (\overline{\rho J_\theta} (\hat{\underset{\sim}{U}} \hat{\nu} + \widehat{U^* \nu^*})\, d\theta \; . \qquad (5.1)$$

The operator $\hat{(\;)}$ is a mass weighted area average over time; the an asterisk denotes its time deviation. See list of symbols. The quantity $\overline{\rho J_\theta} \hat{\underset{\sim}{U}} \hat{\nu}$ is the stationary mode of energy transport determined by the product of time-averaged mass transport, $\overline{\rho J_\theta} \hat{\underset{\sim}{U}}$ equal to $\overline{\rho J_\theta \underset{\sim}{U}}$, and the time-averaged *total flow energy* $\hat{\nu}$. This mode of energy transport is determined by the time-averaged rotational and thermally forced irrotational mode of the mass transport in combination with the mean energy distribution. The quantity $\overline{\rho J_{\theta}}\widehat{\underset{\sim}{U}^{*}\nu^{*}}$ is the transient mode of energy transport associated with a covariance between the deviation mass transport vector, $(\rho J_{\theta}\underset{\sim}{U}^{*})$, and the deviation specific energy, ν^{*}. In this partitioning, the usage of "stationary" to define a time-averaged property or transport process may be misleading. Temporal averaging is a mathematical operation that simply defines the mean of the property or process over the interval of time averaging. Hence, the stationary mode of energy transport does not imply that the mean mass circulation is steady or stationary within the time interval of averaging. The averaging operation is merely a means to extract systematic processes from a dynamic system. In this case, the stationary mode of energy transport through the irrotational component of the mean mass circulation isolates systematic energy exchange associated with the thermally forced isentropic mass circulation. The partitioning also isolates systematic energy transport which is not directly linked to the thermal forcing within the degree of freedom defined by the rotational mode of the mean mass transport.

Global distributions of vertically integrated heating and isentropic divergence of the energy transport, including stationary and transient components, are presented in Figs. 13 and 14 (Johnson, 1985b and c). A comparison of the four fields shows an exceptional agreement between the vertical integrated heating, the divergence of the *total flow energy* transport and its stationary component. The scaling in these three fields is one order of magnitude greater than the field for the transient component. Thus, inspection shows that divergence of the *total flow energy* transport and its stationary component are everywhere at least one order of magnitude greater than the divergence of the transient component.

Fig. 13 and 14

The few regions where the ratio of the transient divergence of energy transport to total divergence is as large as five per cent are located in middle latitudes where transient baroclinic disturbances are most pronounced. Whether or not these differences are a reflection of important isentropic transient energy exchange within these regions cannot be ascertained without further study. Such a study would require an increase in resolution of atmospheric structure in time and space in order to determine if ageostrophic energy transport is important in these active baroclinic regions. Later, it will be established that the divergences of the geostrophic mass and energy transport are linked to each other spatially and temporally to the degree that transient modes of isentropic dry static energy transport are restricted to ageostrophic modes.

A key point that deserves emphasis in this analysis is the fact that the distribution of heating was computed from the isentropic equation of mass continuity while the other three fields were computed from the divergence of the energy transport. In the introduction, it was emphasized that heating involves the physical processes of vertical mass flux, heat sources and sinks and changes of entropy, each of which has its roots in three different physical principles. As such, the information estimating heat sources and sinks in Figs. 13a and 14a stems from a different process than the information in Figs. 13b and 14b.

In Figs. 13a and 14a, the estimation of the planetary scale distribution of heating utilized the isentropic mass continuity equation. By an indefinite vertical integration, vertical profiles of diabatic mass flux are determined from the integral of the isentropic mass divergence in a manner analogous to the kinematic estimation of ω in isobaric coordinates. In regions of heating (cooling) the diabatic mass flux is upward (downward), while the isentropic mass transport is convergent (divergent) in lower layers and divergent (convergent) in upper layers. Once profiles of diabatic mass flux are estimated, the horizontal distribution of heating is calculated from vertically averaging the substantial derivative of potential temperature (or the heat addition per unit mass) (Wei et al., 1983; Johnson and Wei, 1985).

In the estimation of heating from the total energy equation, an assumption that the frictional dissipation is small relative to the other terms is usually made. If one assumes, however, that the transfer of energy by viscous stresses is negligible (Johnson, 1980), the component of diabatic heating due to viscous production of internal energy is locally equal to the frictional dissipation of kinetic energy. Consequently, the net of the vertically integrated horizontal divergence of the *total flow energy* transport is the heat addition by all components except that due to viscous processes. The result is

$$\int_{\theta_{S_0}}^{\theta_T} \overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} v)}\, d\theta = \int_{\theta_{S_0}}^{\theta_T} \overline{\rho J_\theta (Q_m)}_{-d}\, d\theta \quad , \tag{5.2}$$

where the subscript (-d) indicates the net heating by all components except the viscous dissipation of kinetic energy.

Since the vertically-integrated divergence of the mass transport tends to zero for quasi-steady states, the information for estimating the vertically-averaged heating from the vertically-integrated energy transport comes from the covariance between the deviations of energy and mass divergence from their vertically-averaged values. Consequently, in regions of heating (cooling) the divergence (convergence) of the energy transport in upper isentropic layers is greater than the convergence (divergence) of energy transport in lower layers. Within this structure for the time-averaged circulation, net energy transport occurs because an upper branch of the isentropic mass circulation transports more energy from source to sink region than is returned by a lower branch. This structure of differential heating and energy divergence is basic to the result that the scale of the irrotational component of the isentropic mass circulation corresponds with the scale of the energy transport (Johnson and Townsend, 1981; Johnson, 1984).

Given these differences in the methods of estimating the distribution of heat sources and sinks from the isentropic mass continuity equation and the energy equation, there seems to be little doubt empirically that the time-averaged energy transport from heat source to heat sink occurs through the isentropic mass circulation. If the energy transport were independent of the isentropic mass circulation, the distribution of the divergence of the transport associated with the stationary component would not correspond with the distribution of heat sources and sinks and the transient component would be first order in importance. With 1) the similarity of the mass and energy transport within isentropic layers, and 2) the similarity between the regional distribution of heat sources and sinks inferred from mass continuity and from vertically integrated stationary component of energy transport, the evidence, while circumstantial, seems irrefutable. Thus, at this point one may note that both parts of the fundamental premise that was assumed in the introduction have been supported.

5.4 Implicit constraints in stationary and transient modes of isentropic energy transport

In the overall perspective of mass and energy transport by rotational and irrotational modes, it is important to recognize that the role of the rotational component is restricted to a transport of mass and energy downstream within their respective stream functions, ψ_ρ and ψ_ν, and that, as the rotational components of $\rho J_\theta \underset{\sim}{U}$ and $\rho J_\theta \underset{\sim}{U}\nu$ are defined, each of these fields are non-divergent everywhere. In a quasi-Lagrangian definition of time-averages determined in a coordinate system where the stream function itself becomes one of the horizontal curvilinear orthogonal coordinates*, the time-averaged mass and energy transport described by rotational components must exist independently of heat sources and sinks. This constraint stems from the non-divergent character of a rotational mode of transport. Since heating involves the vertical exchange of mass and energy and thus vertical convergence or divergence, local horizontal divergence or convergence of both mass and energy must occur by irrotational modes in order for steady conditions to prevail. This perspective from a quasi-Lagrangian viewpoint contrasts with an Eulerian perspective.

Within an Eulerian perspective the movement and fluctuation of the mass transport stream function field permits degrees of freedom for transient components of the transport of a property by the rotational component of the

---

* Such a coordinate system would correspond to a natural coordinate system, where the tangent vector $\underset{\sim}{t}$ is defined by $\underset{\sim}{k} \times \nabla_\eta \Psi / |\underset{\sim}{k} \times \nabla_\eta \Psi|$ and the normal vector $\underset{\sim}{n}$ is defined by $\nabla_\eta \Psi / |\nabla_\eta \Psi|$. The only limitations to this transformation would be the singularities found at maxima and minima of the stream function.

mass transport, i.e., if time-averages are determined within an Eulerian coordinate system. To appreciate this difference, consider an integral of the divergence of the rotational component of mass transport of an arbitrary property f within an incremental channel bounded by ψ_f and $\psi_f + \delta\psi_f$. This integral is expressed by

$$\int_{\delta A} \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} f)_\psi \, dA = \int_{\delta A} \nabla_\theta \cdot (\underset{\sim}{k} \times \nabla_\theta \psi_f) dA \ . \tag{5.3}$$

Note in Fig. 11 that stream function isopleths within the circumpolar vortex for either mass or energy are cyclic closed curves and thus the incremental channel between ψ_f and $\psi_f + \nabla\psi_f$ constitutes a "meandering ring". In this doubly connected region the divergence of the transport of the property f by the rotational component is everywhere zero. Thus, the net divergence within the "ring" as it is expressed through the divergence theorem,

$$\int_{\delta A} \nabla_\theta \cdot (\rho J \underset{\sim}{U} f)_\psi \, dA = \oint \underset{\sim}{n} \cdot (\underset{\sim}{k} \times \nabla_\theta \psi_f) d\ell = 0 \ , \tag{5.4}$$

is also zero. In this and subsequent expressions, the line integral about the area of the "meandering ring" is defined by first traversing the outside border of δA in the positive direction as defined from a Northern Hemisphere perspective. Then, after a complete traverse of the outer boundary of the ring, the path of integration crosses the ring, makes a complete traverse of the inner boundary in the opposite direction and then crosses the ring back to the starting point of the path of integration. Recall that in the evaluation of such integrals, the net contribution to the line integral is solely from the boundaries of the doubly connected region since the portion of the line integral that crosses the ring provides no net contribution.

Equation (5.4) shows that the non-divergence of the rotational component for either mass or *total flow energy* transport within the ring is maintained from the condition that the outer unit normal $\underset{\sim}{n}$, which is perpendicular to the boundary isopleth Ψ_f, is also everywhere perpendicular to $\underset{\sim}{k}x\nabla_\theta\Psi_f$. This insures that the rotational component of transport of any arbitrary property is restricted to a downstream transport within the ring bounded by the isopleths of Ψ_f.

A similar constraint also applies to transport of any property by the rotational component of the mass transport, $(\underset{\sim}{k} \times \nabla_\theta\Psi_\rho)f$, in the sense that the rotational component of the mass transport defined by $\underset{\sim}{k} \times \nabla_\theta\Psi_\rho$ cannot transport the property across the isopleths of the mass streamfunction Ψ_ρ. This constraint is expressed by

$$\int_{\delta A} \nabla_\theta\cdot[(\rho J_\theta \underset{\sim}{U})_\Psi f]dA = \int_{\delta A} \nabla_\theta\cdot(f\underset{\sim}{k}x\nabla_\theta\Psi_\rho)dA = 0 \qquad (5.5a)$$

$$= \oint f(\underset{\sim}{n}\cdot\underset{\sim}{k}x\nabla_\theta\Psi_\rho)d\ell = 0 \ . \qquad (5.6b)$$

However, in this case the divergence of the rotational component of the mass transport of the property f does not vanish everywhere. "Local" divergence is permitted within the "meandering ring" through downstream advection of the property as it is expressed by $(\underset{\sim}{k}x\nabla_\theta\Psi_\rho)\cdot\nabla_\theta f$. Since the divergence operator is invariant with respect to translation, this capability of the rotational component of the mass transport to redistribute properties by downstream advection within the channel of the "ring" applies to both quasi-Lagrangian and Eulerian perspectives.

In order to establish the differences between time-averaging in a quasi-Lagrangian and Eulerian perspectives, consider the time-averaged divergence of the transport of an arbitrary property within the "meandering ring" bounded by the mass transport stream functions Ψ_ρ and $\Psi_\rho + \delta\Psi_\rho$, as it is expressed by

$$\overline{\int_{\delta A} \nabla_\theta \cdot (\rho \underset{\sim}{J}\underset{\sim}{U} f) dA}^{t(qL)} = \overline{\int_{\delta A} \nabla_\theta \cdot [f(\nabla_\theta X_\rho + \underset{\sim}{k} \times \nabla_\theta \Psi_\rho)] dA}^{t(qL)} \qquad (5.6a)$$

$$= \overline{\oint \underset{\sim}{n} \cdot [f(\nabla_\theta X_\rho + \underset{\sim}{k} \times \nabla_\theta \Psi_\rho)] d\ell}^{t(qL)} \qquad (5.6b)$$

$$= \overline{\oint f \underset{\sim}{n} \cdot \nabla_\theta X_\rho d\ell}^{t(qL)} , \qquad (5.6c)$$

where t(qL) indicates that the time-average is determined following the "meandering ring". In this result the net quasi-Lagrangian time-averaged divergence of the transport of an arbitrary property f over the areal extent of the ring is solely determined by the irrotational component of mass transport. However, a property may be redistributed by convergence and divergence within the confines of the "meandering ring" by the rotational component of mass transport advection, as was just emphasized. Stationary and transient modes are implicitly included in this degree of freedom as well as in the degree of freedom for the redistribution of a property by the irrotational mode. Still, as far as the transport across a "meandering ring", the process in quasi-Lagrangian temporal averaging is restricted to the irrotational mode.

Now consider the Eulerian time-averaged divergence of the transport of an arbitrary property, given by

$$\overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} f)} = \nabla_\theta \cdot \{\overline{\rho J_\theta f[(\nabla_\theta X_\rho/\rho J_\theta) + (\underset{\sim}{k} x \nabla_\theta \Psi_\rho)/\rho J_\theta]}\} \quad , \tag{5.7}$$

where the mass transport potential and stream functions have been divided by ρJ_θ in order to define components of velocity needed for partitioning the transport into stationary and transient components. The order of the time-averaging and differentiation have been exchanged in order to show that the time-averaged divergence is determined by the divergence of the time-averaged stationary and transient components of transport. With these steps, the partitioning yields

$$\overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} f)} = \nabla_\theta \cdot \{\overline{\rho J_\theta} \hat{f}[\widehat{(\nabla_\theta X_\rho/\rho J_\theta)} + \widehat{((\underset{\sim}{k} x \nabla_\theta \Psi_\rho)/\rho J_\theta)}]$$
$$+ \overline{\rho J_\theta}[\widehat{f^*(\nabla_\theta X_\rho/\rho J_\theta)^*} + \widehat{f^*((\underset{\sim}{k} x \nabla_\theta \Psi_\rho)/\rho J_\theta)^*}] \ . \tag{5.8}$$

In this result, explicit degrees of freedom are identified for the transport of the property by stationary and transient components of both rotational and irrotational modes of mass transport.

In order to identify the modes of transport that redistribute energy within and across the time-averaged mass stream function channel, the mean modes of transport in Eq. (5.8) through use of mass-weighted averaging are expressed by

$$[\widehat{\nabla_\theta X_\rho/\rho J_\theta} + \widehat{((\underset{\sim}{k} x \nabla_\theta \Psi_\rho)/\rho J_\theta)}] = (\nabla_\theta \overline{X}_\rho + \underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho)/\overline{\rho J_\theta} \quad , \tag{5.9}$$

a result which explicitly identifies the relation between the stationary mode of rotational and irrotational transport and the time-averaged potential and stream functions. With the combination of Eqs. (5.8) and (5.9) and an integration over the "meandering ring" bounded by $\overline{\Psi}$ and $\overline{\Psi} + \overline{\delta\Psi}$, the net divergence of the time averaged transport of an arbitrary property f becomes

$$\int_{\delta A} \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} f})dA = \oint \hat{f} \underset{\sim}{n} \cdot \nabla_\theta \overline{X}_\rho d\ell + \oint \hat{f} \underset{\sim}{n} \cdot (\underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho) d\ell$$

$$+ \oint \underset{\sim}{n} \cdot \overline{\{f^*[(\nabla_\theta X_\rho / \rho J_\theta)^* + ((\underset{\sim}{k} x \nabla_\theta \Psi_\rho)/\rho J_\theta)^*]\}} d\ell \ . \qquad (5.10)$$

In this partitioning and integration over the "meandering ring", the net divergence of the time-averaged transport of the mean property $\hat{f}$ by the mean rotational component of mass transport will vanish because $\underset{\sim}{n}$ is orthogonal to to $\underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho$. Thus, the planetary global transport of energy by the stationary component is restricted to the thermally-forced irrotational mode of mass transport, as it is in quasi-Lagrangian-averaging. It is important to realize that the divergence of the rotational component of the time-averaged mass transport of the mean property $\hat{f}$ does not vanish locally within the ring. Thus the property $\hat{f}$ may be redistributed downstream. In this case, however, the divergence by advective processes is associated with the advection of the mean property by the time-averaged rotational component of mass transport, expressed by

$$\nabla_\theta \cdot (\hat{f} \underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho) = (\underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho) \cdot \nabla_\theta \hat{f} \ . \qquad (5.11)$$

The last integral of Eq. (5.10) defines the transient components of the divergence of $\overline{\rho J_\theta \underset{\sim}{U} f}$. Although not attempted in this study, it would be revealing to resolve each of the transient vector functions into the components along $\underset{\sim}{n}$ and $(\underset{\sim}{k} x \underset{\sim}{n})$, thereby isolating transport across and along the time-averaged stream function $\overline{\Psi}_\rho$.

In regard to the exchange of *total flow energy*, the empirical evidence suggests that isentropic transient energy transport is for the most part negligible relative to the mean mode. Thus, the divergence of the transport of *total flow energy* Eq. (5.10) is approximated by

$$\int_{\delta A} \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} \nu})\, dA \simeq \oint \nu \hat{\underset{\sim}{n}} \cdot \nabla_\theta \overline{X}_\rho \; d\ell \; . \tag{5.12}$$

The redistribution of energy within the channel by the rotational component, if it occurs, is restricted to

$$\nabla_\theta \cdot [\hat{\nu}(\underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho)] \simeq (\underset{\sim}{k} x \nabla_\theta \overline{\Psi}_\rho) \cdot \nabla_\theta \hat{\nu} \quad . \tag{5.13}$$

The nearly identical fields of the mass and energy stream functions suggest that the advection of *total flow energy* by the time-averaged rotational component of the mass circulation is negligible. In such regions, the isopleths of $\hat{\nu}$ and $\overline{\Psi}_\rho$ would be coincident. This suggests a "Bernoulli" like relation, in that the time-averaged rotational component of the isentropic mass transport would be parallel to surfaces of entropy and time-averaged *total flow energy*. Such a result would be consistent with the fact that the thermally-forced isentropic mass circulation, as it is described by the time-averaged irrotational component, is the sole means by which the available potential energy generated by differential heating is related to the generation of kinetic energy (Johnson, 1986). Recall that for steady-isentropic motion, the mass transport $\rho J_\theta \underset{\sim}{U}$ is non-divergent and parallel to surfaces of entropy and *total flow energy*, while potential vorticity is conserved (Milne-Thompson, 1960). In this case, the only component of the time-averaged mass circulation that is related the heating and the generation of kinetic energy is the irrotational component

(Johnson, 1986). Within the time-averaged circulation, the generation of kinetic energy by the irrotational component primarily balances the kinetic energy dissipation through inertial transformations from the irrotational to rotational component of kinetic energy. If the inertial transformations of the irrotational to the rotational component of kinetic energy are independent or only weakly linked with the isentropic mass circulation, it is reasonable for the time-averaged rotational mass transport to be parallel to the surfaces of entropy and time-averaged *total flow energy*, but not for the time-averaged irrotational mass transport.

In the decomposition expressed in Eq. (5.13), it is still essential to recognize that degrees of freedom exist for the exchange of energy by transient components across the isopleths of the Eulerian time-averaged mass stream function, as it is defined in both isobaric and isentropic coordinates. There is little doubt that energy transport by the transient component of the rotational component dominates in middle and higher latitudes when defined in isobaric coordinates (Lau,1978; Holopainen, 1983). The empirical evidence for isentropic coordinates in Section 5.3 suggests that planetary exchange of energy by transient modes is of second order importance. While this may seem contrary to theory and diagnosis in isobaric coordinates, it is worthwhile noting Starr's (1949) reply to Palmén's (1949) remarks which questioned the validity of Rossby and Starr's (1949) interpretation of angular momentum balance. Starr claimed that "we must encourage free experimentation with various hypotheses and proposals in order to see which one leads to the discovery of new observationally verifiable facts, since definitive criteria for acceptance or rejection are lacking." If the atmosphere were to transport its energy between heat sources and sinks by transient modes that are independent of the mean isentropic mass circulation, the data would reflect such a result in the statistics of the transport processes. Both the empirical results and statistical analyses of the modes of transport suggest that this mode of transport in isentropic coordinates is negligible. These results point towards the importance of the thermally-forced irrotational components of mass and energy transport which maintain global monsoonal circulations.

B. THEORETICAL EVIDENCE FOR GLOBAL MONSOONAL CIRCULATIONS

Within the interdependency of the mass, energy and entropy exchange in maintaining global monsoonal circulations, these are two distinct and important factors to be considered. The one factor involves the requirement for net energy transport from regions of heat sources to regions of heat sink. For regional thermodynamic balance the net energy gained in heat source regions must be equal to the energy lost in heat sink regions both of which must be equal to the net energy transported from heat source to heat sink. As such, this factor involves mass and energy principles as they are expressed through transport theory within isentropic coordinates.

The second factor involves the entropy principle with the generation, the maintenance and the dissipation of the atmosphere's kinetic energy. The dissipation of kinetic energy which constitutes an entropy increasing processes requires an entropy decreasing process by differential heating. Hence, the net generation of total energy by the heat sources and sinks other than viscous dissipation must constitute an entropy decreasing process. The heat sources in some overall sense must occur at higher pressure than heat sinks. This latter process is related directly to the production of kinetic energy through the requirement that within an isentropic layer, the divergence of the isentropic mass transport and the dry static energy must be positively correlated.

6. Energy and entropy balance of the global monsoonal circulation

The physical basis that global monsoonal circulation is determined by the time-averaged isentropic mass circulation, which in turn determines the sense and scale of the energy transport, will now be developed from first principles. This physical basis is established from the combined application of mass continuity, energy and entropy principles (Johnson, 1980; Johnson and Townsend, 1981; Johnson, 1984). The empirical evidence has established that heat sources and sinks determine the sense and scale of the isentropic mass circulation and energy transport. The total potential (internal plus geopotential) energy of a hydrostatic atmosphere is determined by the joint distribution of mass and potential temperature. Consequently, with the use of isentropic coordinates, the total potential energy is determined simply by the mass distribution in that coordinate system; changes in the total potential energy are determined by changes in the mass distribution through the isentropic equation of continuity. The mass distribution and its change in isentropic coordinates also determine the entropy distribution and its change.

6.1 The energy balance of the time-averaged state

The time-averaged isentropic transport equation for total energy (e) defined as the sum of kinetic, gravitational potential and internal energies (including latent energy of phase changes) (Johnson, 1980; Johnson and Townsend, 1981)

$$e_\ell = k + \phi + \iota + Lq, \qquad (6.1)$$

is

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta e_\ell})+\nabla_\theta\cdot(\overline{\rho J_\theta \underset{\sim}{U} e_\ell}) + \frac{\partial}{\partial\theta}(\overline{\rho J_\theta \dot{\theta} e_\ell}) = \nabla_\theta\cdot[\overline{J_\theta(\underset{\approx}{\pi}\cdot\underset{\sim}{U}-\underset{\sim}{H})}$$

$$+ \frac{\partial}{\partial\theta}[\overline{(\underset{\sim}{k}-\nabla_\theta z)\cdot(\underset{\approx}{\pi}\cdot\underset{\sim}{U}-\underset{\sim}{H})}] \quad , \tag{6.2}$$

where $\underset{\approx}{\pi}$ is the stress tensor and $\underset{\sim}{H}$ is the radiative, sensible and latent energy flux. In this form the tendency of the time-averaged total energy, the first term, is determined by boundary processes. The second term is the horizontal divergence of the total energy transport, and the third is the vertical derivative of the diabatic flux of total energy through an isentropic surface. The terms on the right constitute the Lagrangian source/sink function, $\rho J_\theta de/dt$. The first term is the horizontal divergence of the nonconvective transfer of energy through lateral boundaries of a volume element by pressure and viscous stresses and by radiative, sensible and latent energy flux. The second is the vertical derivative of the nonconvective flux of total energy through the upper and lower surfaces of the volume element by the same processes. Since changes in total energy are entirely determined by boundary processes, the principle of energy conservation in isentropic coordinates is revealed under the assumption that boundary exchange vanishes.

A more common form for energy balance is to consider that phase changes of water -- vapor to liquid, vapor to ice or liquid to ice -- are a source of internal energy to the atmosphere. In this case the total energy is defined by

$$e = k + \phi + \iota \quad , \tag{6.3}$$

while separate transport equations for total energy and water vapor are expressed by

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta e}) + \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} e}) + \frac{\partial}{\partial \theta}(\overline{\rho J_\theta \dot{\theta} e}) = \nabla_\theta \cdot \{\overline{J_\theta[\underset{\approx}{\pi} \cdot \underset{\sim}{U} - (\underset{\sim}{H}_s + \underset{\sim}{H}_r)]}\}$$

$$+ \frac{\partial}{\partial \theta}\{\overline{(k - \nabla_\theta z) \cdot [\underset{\approx}{\pi} \cdot U - (H_s + H_r)]}\} + \overline{\rho J_\theta \frac{d}{dt}(Lq)}, \quad (6.4)$$

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta Lq}) + \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} Lq}) + \frac{\partial}{\partial \theta}(\overline{\rho J_\theta \dot{\theta} Lq}) = \{\overline{[\nabla_\theta \cdot (J_\theta \underset{\sim}{H}_\ell) + \frac{\partial}{\partial \theta}[(k - \nabla_\theta z) \cdot \underset{\sim}{H}_\ell]}\}$$

$$+ \overline{\rho J_\theta \frac{d}{dt}(Lq)} \ . \quad (6.5)$$

The last term of Eqs. (6.4) and (6.5) is the Lagrangian source of energy from phase changes. Changes of energy are now no longer uniquely associated with boundary processes since the latent energy is realized as direct heat source within a material element.

A still more common form for the energy balance is to divide boundary work into components associated with pressure and viscous stresses and combine boundary pressure work with the divergence of the transport of total energy. Note, in Eq. (6.4) that through integration over an isentropic volume element, $dAd\theta$, boundary work as a line integral along the lateral boundaries is determined by $\overline{\underset{\sim}{n} \cdot (J_\theta \underset{\approx}{\pi}) \cdot \underset{\sim}{U}}$, while on the upper and lower surfaces it is determined by $\overline{(\underset{\sim}{k} - \nabla_\theta z) \cdot \underset{\approx}{\pi} \cdot \underset{\sim}{U}}$. With the definition of the surface stress by

$$\underset{\sim}{n} \cdot \underset{\approx}{\Pi} = \underset{\sim}{T}_n \,, \tag{6.6}$$

and the colinear vector relation (Johnson, 1980),

$$(k - \nabla_\theta z) = \underset{\sim}{n}(dA_I/dA) \,, \tag{6.7}$$

the incremental boundary work on the lateral and upper and lower surfaces are, respectively, $\overline{J_\theta \underset{\sim}{T}_n \cdot \underset{\sim}{U}}$ and $\overline{(dA_I/dA)\underset{\sim}{T}_n \cdot \underset{\sim}{U}}$. In these two expressions, J_θ serves to transform the incremental area on the lateral boundary from $d\ell dz$ to $d\ell d\theta$, while the ratio, dA_I/dA, determines the area of the inclined upper and lower boundary surfaces dA_I in terms of the projected horizontal area, dA (Johnson, 1980). The transfer of total energy between a region and its environment requires a common component of the velocity and the surface stress. As such, boundary work occurs through the displacement of a fluid element per unit time along the surface stress, a force per unit area. If a net transfer of energy occurs by boundary work, one is assured that an imbalance of the transformation between kinetic and internal energy is ocurring not only in the region of interest, but also in its environment (Johnson and Downey, 1982). Unless the individual source functions of kinetic energy ($-\underset{\sim}{U} \cdot \nabla p$) and internal energy ($-p\nabla \cdot \underset{\sim}{U}$) are explicitly determined within the region and also the environment, no conclusions are permissible regarding just what form of total energy is realized in the region through the boundary work involved and just what form of energy is being transferred from the environment.

In order to isolate the components for boundary work associated with the pressure and viscous stresses, the stress tensor is now expressed as a sum of an isentropic part, $-p\underset{\approx}{I}$, and a non-isotropic part, $\underset{\approx}{\gamma}$, (Weatherburn, 1966; Batchelor, 1967) given by

$$\underset{\approx}{\Pi} = -p\underset{\approx}{I} + \underset{\approx}{\gamma} \ , \tag{6.8}$$

where the deviatoric and strain tensors (Navier, 1822; Stokes, 1845) appropriate to the Navier-Stokes equation are respectively defined by

$$\underset{\approx}{\gamma} = 2\mu[\underset{\approx}{\Phi} - 1/3(\nabla\cdot\underset{\sim}{U})\ \underset{\approx}{I}] \ , \tag{6.9}$$

$$\underset{\approx}{\Phi} = 1/2[\nabla\underset{\sim}{U} + (\nabla\underset{\sim}{U})^T] \ . \tag{6.10}$$

With these definitions the boundary pressure work on an incremental volume element in Eq. (6.4) is defined by

$$\overline{J_\theta\nabla\cdot(p\underset{\approx}{I}\cdot\underset{\sim}{U})} = -\{\nabla_\theta\cdot(\overline{J_\theta p\underset{\approx}{I}\cdot\underset{\sim}{U}}) + \frac{\partial}{\partial\theta}[\overline{(\underset{\sim}{k}-\nabla_\theta z)\cdot p\underset{\approx}{I}\cdot\underset{\sim}{U}}]\} \ , \tag{6.11a}$$

$$= -\{\nabla_\theta\cdot(\overline{\rho J_\theta \underset{\sim}{U} RT}) + \frac{\partial}{\partial\theta}(\overline{\rho J_\theta \dot{\theta} RT}) + \frac{\partial}{\partial\theta}(\overline{p\frac{\partial z}{\partial t_\theta}})\} \ . \tag{6.11b}$$

The form of the equation reveals three components in isentropic coordinates. The first two are uniquely related with the exchange of internal energy by convective transport, $\rho J_\theta \underset{\sim}{U}_I$ and $\rho J_\theta \dot{\theta}_I$, while the last term involves the vertical displacement of isentropic coordinates. It should be emphasized at this point that for isentropic motion, the exchange of total energy between a region and its environment through the action of pressure stresses is a reversible process.

With the Lagrangian source of total energy by boundary pressure work now combined with the divergence of the transport of total energy, the result is

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta e}) + \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} \nu}) + \frac{\partial}{\partial \theta}(\overline{\rho J_\theta \dot{\theta} \nu}) + \frac{\partial}{\partial \theta}(\overline{p \frac{\partial z}{\partial t_\theta}})$$

$$= \nabla_\theta \cdot \{\overline{J_\theta [\underset{\approx}{T} \cdot \underset{\sim}{U} - (\underset{\sim}{H}_s + \underset{\sim}{H}_r)]}\} + \frac{\partial}{\partial \theta}\{\overline{(\underset{\sim}{k} - \nabla_\theta z) \cdot [\underset{\approx}{T} \cdot \underset{\sim}{U} - (\underset{\sim}{H}_s + \underset{\sim}{H}_r)]}\}$$

$$- \rho J_\theta \overline{\frac{d}{dt}(Lq)} \, , \qquad (6.12)$$

where the relations of *total flow energy* ν with the total energy and the "RT" component of work and with the dry static ψ and kinetic energies are expressed by

$$\nu = e + RT \, , \qquad (6.13a)$$

$$= \psi + k \, . \qquad (6.13b)$$

Now, the boundary visous work on an incremental volume in Eqs. (6.4) and also (6.12), as it is defined by

$$\overline{J_\theta \nabla \cdot (\underset{\approx}{T} \cdot \underset{\sim}{U})} = \nabla_\theta \cdot (\overline{J_\theta \underset{\approx}{T} \cdot \underset{\sim}{U}}) + \frac{\partial}{\partial \theta}[\overline{\underset{\sim}{k} - \nabla_\theta z) \cdot \underset{\approx}{T} \cdot \underset{\sim}{U}}] \, , \qquad (6.14)$$

is expanded into a Lagrangian source of kinetic and internal energies given by

$$\overline{J_\theta \nabla \cdot (\underset{\approx}{T} \cdot \underset{\sim}{U})} = \overline{\rho J_\theta \underset{\sim}{U} \cdot \underset{\sim}{F}} + \overline{J_\theta \varepsilon^2} \, . \qquad (6.15)$$

The Lagrangian source of kinetic energy associated with frictional dissipation is

$$\overline{\rho J_\theta \underset{\sim}{U} \cdot \underset{\sim}{F}} = \overline{\underset{\sim}{U} \cdot \{\nabla_\theta \cdot (J_\theta \underset{\approx}{T}) + \frac{\partial}{\partial \theta}[k - \nabla_\theta z) \cdot \underset{\approx}{T}]\}} \, , \qquad (6.16a)$$

$$= \overline{J_\theta \underset{\sim}{U} \cdot (\nabla \cdot \underset{\approx}{T})}, \qquad (6.16b)$$

while the Lagrangian source of internal energy known as the dissipation function (Lamb, 1932; Weatherburn, 1966; Batchelor, 1967) is

$$\overline{J_\theta (e)^2} = \overline{J_\theta \underset{\approx}{T}:\nabla_\theta \underset{\sim}{U}} + \overline{(k-\nabla_\theta z)\cdot\underset{\approx}{T}\cdot\frac{\partial \underset{\sim}{U}}{\partial\theta}} , \tag{6.17a}$$

$$= \overline{J_\theta\ \underset{\approx}{T}:\nabla\underset{\sim}{U}} , \tag{6.17b}$$

$$= \overline{J_\theta \underset{\approx}{T}:\underset{\approx}{\Phi}} . \tag{6.17c}$$

Equations (6.16b) and (6.17b and c) are stated merely to emphasize how through classical relations the temporally averaged loss of kinetic energy due to viscosity and the irreversible generation of internal energy are related to the temporally averaged boundary viscous stresses. The sink of kinetic energy through $\overline{\rho J_\theta \underset{\sim}{U}\cdot \underset{\sim}{F}}$ need not equal the positive dissipation function $\overline{J_\theta \varepsilon^2}$, since total energy may be transferred between a region and its environment by boundary viscous stresses. If a net transfer of total energy occurs by boundary viscous stresses, an imbalance must occur between these two processes within the region and also in the environment. Stress as defined in the Navier-Stokes equations always involves spatial derivatives of the velocity either through $\nabla\underset{\sim}{U}$ or $(\nabla\cdot\underset{\sim}{U})\underset{\approx}{I}$. Thus if viscous stresses exist, internal energy is irreversible generated by virtue of the quadratic nature of the spatial derivatives involving either deformation and/or divergence as expressed in the dissipation function ε^2 (see list of Symbols for its definition).

Now, with the relation between the heating within an isentropic volume element and its individual components expressed by

$$\overline{\rho J_\theta Q_m} = J_\theta \overline{[\varepsilon^2 - \nabla \cdot \underset{\sim}{H}_r - \nabla \cdot \underset{\sim}{H}_s - \rho \frac{d(Lq)}{dt}]} \ , \qquad (6.18)$$

the transport equation for total energy from combination of Eqs. (6.12), (6.15) and (6.18) is

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta e}) + \nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} \nu}) + \frac{\partial}{\partial \theta}(\overline{\rho J_\theta \dot{\theta} \nu}) + \frac{\partial}{\partial \theta}(\overline{p \frac{\partial z}{\partial t_o}}) =$$

$$\overline{\rho J_\theta (Q_m + \underset{\sim}{U} \cdot F)} \ . \qquad (6.19)$$

Apart from its expression in isentropic coordinates, the transport equation for total energy in Eq. (6.19) is the most common form used in energy diagnostics. Since the hydrostatic assumption has not been used in deriving this expression for total energy, the balance equation for total energy as expressed in isentropic coordinates is of a general nature. The development of this equation from Eq. (6.2) places in perspective the physical processes involved. The reversible component of boundary work has been combined with the convective transport of internal energy and expressed as a transport of enthalpy in conjunction with the definition of total flow energy. Hence, if a net boundary transport of enthalpy occurs between a region and its environment, it follows that imbalances exist between the source/sink functions of kinetic and internal energies in association with reversible processes both in the region and also its environment.

6.2 The sense of the energy transport in relation to heat sources and sinks

Inspection of Eqs. (6.2) and (6.19) suggests that a vertical integration will yield insight into the relation between the horizontal transport and diabatic and viscous processes. By averaging over an appropriate time interval for which the steady assumption becomes valid and assuming a rigid earth-atmosphere interface with no transfer of energy by pressure or viscous stresses across the earth's surface, a vertical integration of Eq. (6.19) and use of the hydrostatic equation yields

$$\int_{\theta_{S_o}}^{\theta_T} \nabla_\theta \cdot \overline{(\rho J_\theta \underset{\sim}{U} \nu)} d\theta + \overline{E2} = - \int_{\theta_{S_o}}^{\theta_T} \overline{\psi \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})} d\theta = \int_{\theta_{S_o}}^{\theta_T} \overline{\rho J_\theta \dot{\theta} \frac{\partial \psi}{\partial \theta}} d\theta \quad , \qquad (6.20)$$

where the limits of integration extend from the coldest potential temperature observed globally, θ_{S_o}, to the top of the atmosphere, θ_T. The first left term represents the vertically integrated horizontal divergence of transport of total energy and boundary pressure work. $\overline{E2}$ is the vertically integrated time-averaged viscous dissipation function which under the assumption of no boundary viscous work at the earth-atmosphere interface is equal to the vertical integral of mechanical viscous work, $\overline{\rho J_\theta \tilde{U} \cdot \tilde{F}}$. Boundary viscous work $(\overline{\underset{\sim}{n} \cdot \underset{\approx}{\gamma} \cdot \underset{\approx}{U}})$ vanishes everywhere if the earth-atmosphere interface remains rigid, i.e., with $\underset{\sim}{U}(z_S)$ equal to zero there is no displacement of material elements at the interface. While such an assumption is not satisfied at the interface between the atmosphere and ocean, its violation is of little consequence here, since a loss of energy by viscous work from the atmosphere to the ocean would in effect constitute an added sink to $\overline{E2}$. The direct transfer of atmospheric energy to the hydrosphere through excitation of waves by both pressure and viscous stresses is not generally acknowledged for the importance that it deserves.

The integrand of the last expression in Eq. (16.20) is the heat addition, $\overline{\rho J_\theta Q_m}$. The middle expression, from an integration by parts is merely equal to the integral of the heat addition. The fact that the last integrand equals $\overline{\rho J_\theta Q_m}$ shows that heat addition in isentropic coordinates and the vertical advection of dry static energy across isentropic surfaces are equivalent processes. Through the hydrostatic equation, the vertical derivative of the dry static energy in Eq. (6.4) is equal to the Exner function (π), which is likewise proportional to the R/c_p power of hydrostatic pressure (the mass). Thus it is evident that the source of energy, $\overline{\rho J_\theta Q_m}$, is also directly related to the diabatic mass flux and the isentropic mass distribution.

In the energy balance expressed by Eq. (6.20), the scale and intensity of the vertically integrated horizontal transport of total energy is invariant with respect to coordinate systems. Thus, the balance as well as the sense of the energy transport determined by the following analysis is of a general nature.

For illustrative purposes, geographical regimes of diabatic heating for the planetary scale, such as occur in global monsoonal circulations, are assumed to be distinct from regimes of cooling. In each regime, profiles of diabatic mass flux are also assumed to increase or decrease monotonically from zero to an extremum at some intermediate isentropic surface, θ_m, within the atmosphere, above which the magnitude of the mass flux decreases to zero. Thus, only one maximum in the magnitude of the diabatic mass flux occurs in each geographic regime of cooling or heating. Heating profiles which illustrate this characteristic structure are shown in Fig. 15 (from Johnson and Wei, 1985).

Fig. 15

Vertical integration of the derivative of the vertical mass flux from the earth's surface to the top of the atmosphere, in conjunction with the condition that $\rho J_\theta \dot{\theta}$ vanishes at both θ_T and θ_{s_o}, yields

$$\overline{\int_{\theta_{s_o}}^{\theta_T} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta}) d\theta} = 0 \ . \tag{6.21}$$

With a division into a lower and an upper layer separated by the level of maximum absolute value of the vertical mass flux, $\theta_m(t)$, Eq. (6.2) becomes

$$0 = \overline{\int_{\theta_{s_o}}^{\theta_T} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta}) d\theta} = \overline{\int_{\theta_{s_o}}^{\theta_m(t)} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})_\ell d\theta} + \overline{\int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})_u d\theta} \ . \tag{6.22}$$

The right hand integrals are equal in magnitude and opposite in sign as long as the assumption of a single extremum for the diabatic mass flux is satisfied. The level of maximum upward (downward) diabatic mass transport, $\theta_m(t)$, separates the lower layer (ℓ), where the upward (downward) mass transport is increasing (decreasing) with respect to potential temperature, from the upper layer (u) where the upward (downward) mass transport is decreasing (increasing). Thus, in regimes of heating, $\partial(\rho J_\theta \dot{\theta})_\ell/\partial\theta$ is positive and $\partial(\rho J_\theta \dot{\theta})_u/\partial\theta$ is negative, while in regions of cooling, $\partial(\rho J_\theta \dot{\theta})_\ell/\partial\theta$ is negative and $\partial(\rho J_\theta \dot{\theta})_u/\partial\theta$ is positive. The right hand integral of the energy transport equation Eq. (6.4), similarly split into two parts, is given by

$$\int_{\theta_{s_o}}^{\theta_T} \overline{\nabla_\theta \cdot (\rho J_\theta U \underset{\sim}{\nu})} d\theta + \overline{E2} = -[\overline{\int_{\theta_{s_o}}^{\theta_m(t)} \psi \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})_\ell d\theta} + \overline{\int_{\theta_m(t)}^{\theta_T} \psi \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})_u d\theta}] \ . \tag{6.23}$$

Since ψ is always a positive monotonic increasing function with respect to potential temperature in a hydrostric atmosphere and $\partial(\rho J_\theta\dot{\theta})/\partial\theta$ of each integral is either positive or negative, by the mean value theorem Eq. (6.23) becomes

$$\int_{\theta_{s_o}}^{\theta_T} \overline{\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}\nu)}d\theta + \overline{E^2} = \overline{-\psi_\ell\int_{\theta_{s_o}}^{\theta_m(t)} \frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})_\ell\, d\theta} \overline{-\psi_u\int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})_u d\theta} \;. \quad (6.24)$$

The inequality $\psi(\theta_{s_o}) < \psi_\ell < \psi(\theta_m) < \psi_u < \psi(\theta_T)$ is always satisfied within each vertical column.

With this condition and the uniqueness of the sign of each of the right hand integrals, the mean value theorem is used with respect to the averaging integral over time in combination with Eq. (6.23) to determine the result,

$$\int_{\theta_{s_o}}^{\theta_T} \overline{\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}\nu)}d\theta + \overline{E^2} = -(\tilde{\psi}_u - \tilde{\psi}_\ell)\overline{\int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})_u d\theta} \;. \quad (6.25)$$

The tilde over $\tilde{\psi}_u$ and $\tilde{\psi}_\ell$ represents an appropriate mean value over the time interval of averaging. Since ψ_u is always greater than ψ_ℓ in a hydrostatic atmosphere at each point in time, the difference $(\tilde{\psi}_u - \tilde{\psi}_\ell)$ is always positive. Recognizing that the viscous dissipation function, $\overline{E^2}$, is much less than the magnitude of the heating of the atmosphere the quasi-steady energy transport equation Eq. (6.25) states in the mean value sense that the divergence of energy transport is determined primarily by the difference of the dry static energies between upper and lower isentropic layers and the magnitude of the diabatic vertical mass transport at θ_m.

With $\partial(\rho J_\theta\dot{\theta})_u/\partial\theta$ being negative in regions of heating, the vertically integrated horizontal transport of *total flow energy* within a region of heating must be divergent in that the layers above θ_m must export more energy than the lower levels import. Likewise since $\partial(\rho J_\theta \dot{\theta})_u/\partial\theta$ is positive in regions of cooling, the vertically-integrated energy transport must be convergent, with the result that the layers above θ_m import more energy than the lower layers export. Thus, the result in Eq. (6.25) dictates that the time-averaged circulation transports net total flow energy from a source to sink region because the upper branch of the isentropic mass circulation transports more energy from source to sink region than is returned by the lower branch. Consequently, the sense and scale of the isentropic mass transport uniquely determine both the sense and scale of the energy transport, all of which are determined by the planetary distribution of heat sources and sinks. The actual intensity of energy transported is also strongly dependent on the intensity of the mass circulation and the mass distribution. Note, in Eq. (6.25) that the vertically-integrated divergence of the energy transport in the mean value sense depends directly on both the intensity of diabatic mass flux at θ_m and the difference of the dry static energies $\tilde{\psi}_u$ and $\tilde{\psi}_\ell$. The dry static energy is determined from the mass distribution through integration of the hydrostatic equation.

In preceeding sections, both empirical and theoretical support have been provided for the fundamental premise of this study--that global monsoonal circulations occur in the form of time-averaged isentropic mass circulations. However, a key result that remains to be established theoretically is why the isentropic energy transport occur through the mean isentropic mass circulation and not the transient mode, as is the case in isobaric coordinates. The result in Eq. (6.25) allows the modes of energy transport (i.e., transient or steady, rotational or irrotational, geostrophic or ageostrophic, etc.) to be dependent on the mutual adjustment of mass and momentum, boundary conditions, and nonlinear inertial forcing of the isentropic mass circulation.

If systematic exchanges of energy associated with nonlinear inertial or other processes occur which are independent of thermally-forced isentropic mass circulation, modes of mean energy exchange may exist which are independent of the mass circulation. In this case, the robustness of the above premise would be compromised.

Within this transport of total flow energy from heat source to heat sink, the RT component requires imbalances between the generation of kinetic and internal energies in both the heat source and sink region. No attempt will be made in this study to address the relative importance of sources of energy through boundary processes versus the sources of energy internal to the atmosphere, although this matter is important in understanding climate. The focus in this analysis is on the mass circulation and energy transport in isentropic coordinates. It is the heating internal to the atmosphere as it changes the energy and entropy distributions that is of direct interest.

Apart from the effect of latent heat, the global distribution of energy sources and sinks is largely determined by the net energy flux through the upper and lower boundaries of the atmosphere, as revealed by a vertical integration of Eq. (6.2). However, latent heating introduces manifold complications, in that intense Lagrangian sources of internal energy develop from phase changes. This latent heat release, which forces the vertical branches of isentropic mass circulations, is linked with the horizontal branches that supply the latent energy to the region of moist convection.

Consequently, presence of latent heat within an ascending branch of the isentropic mass circulation alters the scale and intensity of the mass circulation that would occur in an atmosphere without latent energy. Complications are particularly pronounced in regions where circulations occur in conjunction with symmetric, convective or baroclinic instabilities and/or weak hydrodynamic stability. These complexities preclude analytic solutions that adequately describe global monsoonal circulations.

6.3 The sense of the entropy transport in relation to heat sources and sinks

The similarity of the relations for vertically integrated total energy and entropy transport is now studied. In so doing, some key points that distinguish the principles of energy and entropy from each other will be identified and also used to establish the necessity of a differentially heated atmosphere. With the definition of entropy,

$$s - s_o = c_p \ln \frac{\theta}{\theta_o} , \qquad (6.26)$$

the transport equation for entropy is expressed by

$$\frac{\partial}{\partial t_\theta}(\rho J_\theta s) + \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} s) + \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta} s) = c_p(\rho J_\theta \dot{\theta})/\theta \quad . \qquad (6.27)$$

The first term is the tendency of entropy, the second is the horizontal divergence of the entropy transport, the third is the vertical derivative of the diabatic flux of entropy. The right hand term is the Lagrangian source function. In contrast to energy, the Lagrangian source function for entropy cannot be expressed as a boundary flux of entropy (Sommerfeld, 1950). With a rearrangement and vertical integration, the quasi-steady time-averaged entropy transport equation is expressed by

$$\overline{\int_{\theta_{S_0}}^{\theta_T} \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} s)\, d\theta} = - \overline{\int_{\theta_{S_0}}^{\theta_T} s \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta})\, d\theta} \quad . \tag{6.28}$$

With the exception of the positive dissipation function in the energy equation, this form of the entropy transport equation, Eq. (6.28), is identical with its energy counterpart, Eq. (6.20). With the monotonic increase of entropy with respect to potential temperature and an analogous development to the mass and energy relations Eqs. (6.20) through (6.25), the vertically integrated entropy balance becomes

$$\overline{\int_{\theta_{S_0}}^{\theta_T} \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} s)\, d\theta} = - (\tilde{s}_u - \tilde{s}_\ell) \overline{\int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta})_u\, d\theta} \quad , \tag{6.29}$$

where $\tilde{s}_u$ and $\tilde{s}_\ell$ are values of entropy that represent appropriate mean values over the time interval of averaging. This result establishes that the entropy transport between source and sink regions is also determined by the difference of entropy between upper and lower isentropic layers and the same distribution of differential heating that determines the mass and energy transport. Since the entropy in the upper layer is always larger than the entropy in the lower layer, just like the energy transport the vertically integrated transport vector in regions of heating $(\partial(\rho J_\theta \dot{\theta})_u / \partial\theta < 0)$ must be divergent. In this situation the atmosphere above the level of maximum upward diabatic mass flux exports more entropy from its source to its sink region than the atmosphere beneath θ_m imports from its sink to its source region. Since $\partial(\rho J_\theta \dot{\theta})_u / \partial\theta$ is positive in regions of cooling, the vertically integrated entropy transport must be convergent, since the atmosphere above the level of maximum downward diabatic mass flux imports more entropy than the lower atmosphere exports.

A comparision of Eqs. (6.25) and (6.29) shows: 1) the sense of the mass, energy and entropy transport correspond uniquely with each other, and 2) the scales of the transport of all three properties, being determined by the planetary distribution of heat sources and sinks, are also identical. The single important difference between the intensity of the quasi-horizontal transport of energy and entropy is the dissipation function. For steady conditions and given heating (cooling), the vertically integrated divergence of energy transport increases (decreases) as the dissipation function decreases. If frictional dissipation were absent, the temporally vertically integrated balance of total energy and entropy would be identically maintained. The fact that the vertically integrated frictional dissipation in the atmosphere is small in comparison to the vertical integrals of the divergence of the energy transport and diabatic heating in Eq. (6.20) is undoubtedly the key to the observational fact that a one to one correspondence tends to exist between measures of entropy and energy, such as exists between $c_p\theta$ and ψ. The former provides a good approximation in an evaluation of the transport of dry static energy (Zillman, 1972; Zillman and Johnson, 1985). This is related to the fact that the balance of entropy can be used to study the energy balance (Lettau, 1954a; Paltridge, 1978).

6.4 The vertical and horizontal exchange of energy and entropy

The vertical structure of the mass, energy and entropy exchange in regions of diabatic heating and cooling is now examined. From Eqs. (6.25) and

(6.29), the thermodynamics equation and an integration by parts, the energy and entropy transport equations become

$$\overline{(\tilde{\psi}_u - \tilde{\psi}_\ell) \int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial\theta}(\rho J_\theta \dot{\theta})_u d\theta} = \overline{\int_{\theta_{s_o}}^{\theta_T} \rho J_\theta \dot{\theta} \frac{\partial\psi}{\partial\theta}\, d\theta} \qquad (6.30a)$$

$$= \overline{\int_{\theta_{s_o}}^{\theta_T} \rho J_\theta \dot{s} T \, d\theta} \qquad (6.30b)$$

$$\overline{(\tilde{s}_u - \tilde{s}_\ell) \int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial\theta}(\rho J_\theta \dot{\theta})_u d\theta} = \overline{\int_{\theta_{s_o}}^{\theta_T} \rho J_\theta \dot{\theta} \frac{\partial s}{\partial\theta}\, d\theta} \qquad (6.31a)$$

$$= \overline{\int_{\theta_{s_o}}^{\theta_T} \rho J_\theta \dot{s} \, d\theta} \quad . \qquad (6.31b)$$

The form of these equations reveal that the source functions of total energy and entropy in Eqs. (6.30b) and (6.31b) are equivalent, respectively, to the integrals of the vertical advection of dry static energy and entropy in Eqs. (6.30a) and (6.31a). This is somewhat suprising since the Lagrangian source of entropy is an exact differential of the state process of heat addition, while the Lagrangian source of total energy by heat addition is not. In the case of entropy, the vertical advection function $\dot{\theta}\partial s/\partial\theta$ is identically equal to the Lagrangian source of entropy ds/dt. In the integral of Eq. (6.30a), the vertical advection function $\dot{\theta}\partial\psi/\partial\theta$ is only one component of the Lagrangian source $d\psi/dt$, as it is expressed by $(\partial\psi/\partial t_\theta + \underset{\sim}{U}\cdot\nabla_\theta\psi + \dot{\theta}\partial\psi/\partial\theta)$. In the integral form of the energy equation as it is determined from

Eq. (6.19) and the steady assumption, the degree of freedom associated with the tendency has been constrained through temporal integration in isentropic coordinates. The other degree of freedom, $\underset{\sim}{U}\cdot\nabla_\theta\psi$, is embedded in the divergence of the transport of *total flow energy*.

The time-averaged relation as derived in Eq. (6.25) supports a Bernoulli-like relation for the time averaged exchange of energy that was discussed in Sec. 5.4. By definition the time-averaged transport of mass and *total flow energy* by their rotational components, $\underset{\sim}{k} \times \overline{\Psi}_\rho$ and $\underset{\sim}{k} \times \overline{\Psi}_\nu$, respectively, are nondivergent. Thus it is clear from Eq. (6.20) and the nondivergent restriction enjoyed by the rotational modes of isentropic mass and energy transport, that as noted in Section 5.4, it is the irrotational mode of mass and energy transport which satisfies the balance requirements imposed by differential heating.

Following the development in Section 5.4, with the arbitrary property f equal to total flow energy ν, the divergence of the transport of total flow energy is expressed by

$$\nabla_\theta\cdot(\overline{\rho J_\theta \underset{\sim}{U}\nu}) = \nabla_\theta\cdot\overline{\{\rho J_\theta\nu[(\nabla_\theta\chi_\rho/\rho J_\theta) + (k\times\nabla_\theta\Psi_\rho/\rho J_\theta)]\}} \quad . \qquad (6.32)$$

By partitioning into stationary and transient components, the divergence of the total flow energy with the use of Eq. (5.9) is expressed by

$$\nabla_\theta\cdot(\overline{\rho J_\theta \underset{\sim}{U}\nu}) = \nabla_\theta\cdot[\hat{\nu}(\nabla_\theta\overline{\chi}_\rho + k \times \nabla_\theta\overline{\Psi})] +$$

$$\nabla_\theta\cdot\{\overline{\rho J_\theta}[\widehat{\nu^*(\nabla_\theta\chi_\rho/\rho J_\theta)^*} + \widehat{\nu^*(k \times \nabla_\theta\Psi_\rho/\rho J_\theta)^*}] \quad . \qquad (6.33)$$

The empirical results of Section 5.3 indicated that the transient component of planetary scale exchange of total flow energy was negligible in relation to the stationary component. Thus the divergence of the transport of total flow energy is approximated by

$$\nabla_\theta \cdot (\overline{\rho J_\theta \underset{\sim}{U} \nu}) \simeq \nabla_\theta \cdot (\hat{\nu} \nabla_\theta \overline{\chi}_\rho) + (k \times \nabla_\theta \overline{\psi}_\rho) \cdot \nabla_\theta \hat{\nu} \quad . \tag{6.34}$$

Substitution of this result into Eq. (6.25) yields

$$\int_{\theta_{s_o}}^{\theta_T} \nabla_\theta \cdot (\hat{\nu} \nabla_\theta \overline{\chi}_\rho) + (\underset{\sim}{k} \times \nabla_\theta \overline{\Psi}_\rho) \cdot \nabla_\theta \hat{\nu} d\theta + \overline{E^2} \simeq - (\tilde{\psi}_u - \tilde{\psi}_\ell) \overline{\int_{\theta_m(t)}^{\theta_T} \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta})_u d\theta} \; . \tag{6.35}$$

The corresponding result for the entropy balance without approximation is

$$\int_{\theta_{s_o}}^{\theta T} \nabla_\theta \cdot (s \nabla_\theta \overline{\chi}_\rho) \; d\theta = - (s_u - s_\ell) \overline{\int_{\theta_m(t)}^{\theta T} \frac{\partial}{\partial \theta} (\rho J \dot{\theta})_u \; d\theta} \tag{6.36}$$

Except for the advection of mean total flow energy by the mean rotational mass transport, the modes of transport for entropy and total flow energy are identical. In Section 5.4, it was established that this mode consisted of a downstream advection of time-averaged *total flow energy* within the "meandering ring" defined by Ψ_ρ and $\Psi_\rho + \Delta\Psi_\rho$. It was also established that upon area integration over the "meandering ring", the net exchange of *total flow energy* with the environment by this mode of the "meandering ring" vanishes. Thus, in the isentropic time-averaged state, it is the irrotational component of mass transport that is linked with sources and sinks of energy that occur through viscous dissipation and the diabatic vertical mass flux in Eq. (6.35), both of which involve changes of entropy. By virtue of the entropy change, these two processes involve the addition or removal of total energy along trajectories in the vertical branches of the isentropic mass circulation. In regions of entropy increase through heating, the trajectories are upward through the isentropic structure as the total energy added is transported away through the divergence of the irrotational component. In regions of entropy decrease through cooling, the trajectories are downward as the total energy removed is imported to the region through convergence of the irrotational component. In the case of entropy, a similar exchange takes place. However, in this situation the relationship between the vertical advection of entropy and the Lagrangian source of entropy along the trajectory is unique, without the need to impose integral constraints through temporal integration.

The requirement that the sources and sinks of energy and entropy enjoy a relation with trajectories through isentropic surfaces complements Sir Napier Shaw's views (1930) on the vertical exchange of mass between the underworld and overworld, in that the diabatic mass transport through vertical advection of entropy and dry static energy determines both the scale and sense of the horizontal and vertical exchange of energy and entropy. The requirements for ascending and descending trajectories through isentropic surfaces in association with heating and cooling and the mean irrotational modes of mass and energy transport allow the satisfaction of balance required by the nonisentropic processes of heating and viscous dissipation and at the same time permit a Bernoulli-like relation between the mean rotational modes of mass and energy transport.

6.5 The link between global differential heating and the vertical exchange of energy and entropy

The delicate nature and important role of global differential heating in the vertical exchange of energy and entropy in maintaining the atmosphere's circulation is now ascertained. Through the combination of Eqs. (6.20) and (6.30a and b) and an area integration over the global domain of the atmosphere, the transport equation for total energy reduces to

$$\overline{E^2}^{tA} = \int_{\theta_{s_0}}^{\theta_T} \overline{\rho J \dot{s} T}^{tA} \, d\theta \qquad (6.37a)$$

$$= \int_{\theta_{s_0}}^{\theta_T} \overline{\rho J}^{tA} \left[\hat{\dot{s}}^{tA} \, \hat{T}^{tA} + \widehat{\dot{s}^{**} T^{**}}^{tA} \right] d\theta \, , \qquad (6.37b)$$

where the operator $(\ \ )^{\hat{}tA}$ and the double asterisks (**) are respectively the time-area mass-weighted average and its corresponding deviation within the global extent of an isentropic layer. With an areal integration of the transport equation for entropy, Eq. (6.27), and use of Eq. (6.26), the following constraints become evident:

$$0 = \overline{\rho J_\theta \dot{\theta}}^{tA} = \overline{c_p \rho J_\theta \dot{\theta}/\theta}^{tA} = \overline{\rho J_\theta \dot{\theta} \frac{\partial s}{\partial \theta}}^{tA} = \overline{\rho J_\theta \dot{s}}^{tA} \qquad (6.38)$$

Thus, vanishing of the quasi-steady area-averaged mass flux in isentropic coordinates demands the following: that the temporally areally averaged entropy source/sink function from diabatic heating, the temporally, areally averaged vertical advection of entropy on each isentropic surface in Eq. (6.31a) and the temporally, areally averaged vertical entropy flux must all vanish. This is not the case for energy. The substitution of the constraint $\hat{\dot{s}}$ is zero into Eq. (6.38) yields

$$\overline{E^2}^{tA} = \int_{\theta_{s_0}}^{\theta_T} \overline{\rho J}^{tA} \widehat{(\dot{s}^{**}T^{**})}^{tA} d\theta \qquad (6.39a)$$

$$= c_p^{-1} \int_{\theta_{s_0}}^{\theta_T} \overline{\rho J}^{tA} \widehat{(\dot{s}^{**}\pi^{**})_\theta}^{tA} d\theta \quad . \qquad (6.39b)$$

Since viscous processes in the atmosphere are always sinks of total energy even though the net entropy production vanishes in an isentropic layer, generation of total energy is required by differential heating. In some isentropic layers, the temperature [pressure (p^κ)] in regions of diabatic heating must be greater than its mass weighted isentropic average, while in regions of diabatic cooling the temperature [pressure (p^κ)] must be less than its mass weighted average. This result in conjunction with an areal integration of Eq. (6.30a) demands that within an isentropic layer the net vertical energy advection function in the time-area averaged domain must be positive. This net vertical advection of energy in the isentropic structure offsets the energy sink by viscous dissipation. These latter two requirements stem from an entropy constraint. Within the isentropic structure heat sinks must occur at higher pressure than the heat sources in order that kinetic energy is generated to offset viscous dissipation.

Since the temporally-areally-averaged vertical energy advection is positive and, since the heat source region must occur at high pressure relative to the heat sink region within an isentropic layer, one can conclude that differential heating increases the dry static energy of the ascending trajectories more than the dry static energy of the descending trajectories is decreased. For given potential temperature (or entropy) change, the difference between the gain and loss of dry static energy along the ascending and descending trajectories is determined by the difference in the temperature (also pressure, p^{κ}) between the ascending and descending trajectories. Specifically, the gain of dry static energy with heating at high pressure through upward diabatic mass flux $\rho J_{\theta}\dot{\theta}$ is greater than the loss of dry static energy with cooling at low pressure. Viewed within classical concepts of a thermodynamic cycle, the addition of heat and expansion are coincident phases as are the removal of heat and compression. The essential requirement in an isentropic layer is that the addition of heat and expansion occur at higher temperature (also pressure) than the removal of heat and compression, which is exactly what occurs from an isentropic perspective. The dry static energy within an isentropic layer that is created by this process ultimately offsets the viscous dissipation of kinetic energy. As emphasized by Dutton and Johnson (1967), the source of total potential energy that maintains the kinetic energy supply of the general circulation against frictional dissipation can only occur through heating at high temperature (also pressure) and cooling at low temperature within an isentropic layer. Any attempt to circumvent this basic result will lead to a violation of one or more of the basic principles of mass, energy, and entropy. This requirement for differential heating, its relation to the isentropic mass circulation energy transport, and entropy transport, and Lagrangian sources of dry static energy are subtle ones that cannot be determined from energy principles alone.

The important role of differential heating has been established in previous works on available potential energy and entropy (Lettau, 1954a; Lorenz, 1955a, 1955b; Dutton and Johnson, 1967; Johnson, 1970; Dutton, 1976a). However, the fact that diabatic heating increases the entropy and energy of a mass circulation as trajectories ascend through atmospheric stratification and decreases entropy and energy as trajectories descend within global monsoonal circulations has not been explicitly elucidated.

The implications of these results are that the addition and extraction of heat within a stratified atmosphere must be viewed as thermodynamical processes through which the vertical branches of the atmosphere's global monsoonal circulations occur and are forced. Through mass continuity the vertical branches link with horizontal branches within which circulation and kinetic energy are realized. The circulation and kinetic energy that is realized within the horizontal branches of monsoons should not be viewed in isolation as occurring through isentropic processes, at least not for the global circulation.

6.6 The generation of kinetic energy

The direct link between the generation of kinetic energy within the thermally forced isentropic mass circulations and the differential heating is readily established by an area integration of Eq. (6.20). With this step the transport equation for total energy simplifies to

$$\overline{E^2}^{tA} = - \int_{\theta_{s_0}}^{\theta_T} \overline{\psi \frac{\partial}{\partial \theta} (\rho J \dot{\theta})}^{tA} \, d\theta \,. \qquad (6.40)$$

A substitution of the isentropic mass continuity equation yields

$$\overline{E^2}^{tA} = \int_{\theta_{s_o}}^{\theta_T} \overline{\psi[\frac{\partial}{\partial t_\theta}(\rho J_\theta) + \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})] \quad d\theta}^{tA} \quad . \tag{6.41}$$

With the hydrostatic assumption, a differentiation by parts and the definition of the Montgomery Stream function, Eq. (6.41) becomes

$$\overline{E^2}^{tA} = - \int_{\theta_{s_o}}^{\theta_T} \overline{\frac{1}{g}[\frac{\partial}{\partial \theta}(\psi \frac{\partial p}{\partial t_\theta}) - \frac{\partial \psi}{\partial \theta}\frac{\partial p}{\partial t_\theta}] - \psi \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}) \quad d\theta}^{tA} \quad , \tag{6.42a}$$

$$= - \frac{1}{g}\int_{\theta_{s_o}}^{\theta_T} \frac{c_p}{(1+k)p_{oo}{}^{\kappa}} \{\frac{\partial}{\partial \theta}[\frac{\partial}{\partial t_\theta}\overline{(\theta p^{1+k}}^{tA})] - \frac{\partial}{\partial t_\theta}\overline{p^{1+k}}^{tA}\}$$

$$+ \frac{\partial}{\partial \theta}(\phi \overline{\frac{\partial p}{\partial t_\theta}}^{tA}) \quad d\theta + \int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta \{\psi [\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})/\rho J_\theta]\}}^{tA} \quad d\theta \quad . \tag{6.42b}$$

In the time-averaged state, an underlying assumption for global monsoonal circulations is that the distributions of mass and energy are steady. Thus, the bracketed part of the integrand of the first integral vanishes. The last term of this integrand integrates to upper and lower boundary conditions. At the upper boundary θ_T, the pressure tendency vanishes. At the earth's surface,

$$g^{-1} \phi \overline{\frac{\partial p}{\partial t_\theta}}^{tA} \Big|_{\theta_{s_o}} = \overline{\phi_s \frac{\partial}{\partial t_\theta}(\overline{p_s}^{t} /g)}^{A} \quad , \tag{6.43}$$

vanishes from the assumption that the time-averaged mass distribution is steady. With these simplifications and a partitioning into mean and eddy components, (Eq. (6.42b) becomes

$$\overline{E^2}^{tA} = \int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta}^{tA} \;\widehat{\{\psi^{**}[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})/\rho J_\theta]^{**}\}}^{tA} \, d\theta \quad , \qquad (6.44)$$

where the double asterisks indicates a deviation from the mass-weighted area average. The mass weighted average of $[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})/\rho J_\theta]$ vanished from areal integration. The integrand of Eq. (6.44) may be expressed by

$$\overline{\rho J_\theta}^{tA} \;\widehat{\{\psi^{**}[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})/\rho J_\theta]^{**}\}}^{tA} = \overline{\rho J_\theta\{\psi^{**}[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})/\rho J_\theta]\}}^{tA} \quad ,$$

$$= \overline{\psi^{**}\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})}^{t,A} \quad . \qquad (6.45)$$

A differentiation identity yields

$$\overline{\psi^{**}\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})}^{tA} = \overline{\nabla_\theta\cdot(\psi^{**}\rho J_\theta \underset{\sim}{U})}^{tA} - \overline{\rho J_\theta \underset{\sim}{U}\cdot\nabla_\theta\psi^{**}}^{tA} \quad . \qquad (6.46)$$

Substitution of this result into Eq. (6.44) yields

$$\overline{E^2}^{tA} = -\int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta \underset{\sim}{U}\cdot\nabla_\theta\psi^{**}}^{tA} \, d\theta \qquad (6.47a)$$

$$= -\int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta \underset{\sim}{U}\cdot\nabla_\theta\psi}^{tA} \, d\theta \quad , \qquad (6.47b)$$

where the condition that $\nabla_\theta \hat{\psi}^{tA}$ is identically zero has been utilized.

Although the partitioning of the integrand of the last integral of Eq. (6.42b) is not essential in deriving Eq. (6.47b), Eqs. (6.39a and b), (6.44) and (6.47) provide insight into the maintenance of the atmosphere's balance of total energy between the entropy increasing processes by kinetic energy dissipation and entropy decreasing processes by differential heating. In regions of heating, the gain of total energy in Eq. (6.39a) through entropy increases in ascending trajectories is accompanied by isentropic mass divergence and positive deviation of dry static energy in the upper branch of an isentropic mass circulation. In regions of cooling, the loss of total energy through entropy decreases in descending trajectories is accompanied by isentropic mass convergence and negative deviation of dry static energy in the lower branch of an isentropic mass circulation. Within the mass circulation, volume expansion and contraction in association with thermodynamic work (Johnson and Downey, 1982) provide for the transformation of total potential to kinetic energy through the combination of boundary work by pressure stresses, transport of internal and geopotential energies and advection of dry static energy. The nonconvective transport of internal energy by boundary pressure work and the convective transport of internal and geopotential energies from a heat source region to heat sink region is implicit in the first right hand term of Eq. (6.46) (which vanishes upon areal integration). The generation of kinetic energy in Eq. (6.47) associated with these processes occurs through the ageostrophic advection of dry static energy.

Within this structure of thermodynamic processes, the generation of total energy by the entropy decreasing process of differential heating, the production of kinetic energy by volume expansion and contraction and the degradation of total energy by the entropy increasing process of viscous dissipation are linked intrinsically with each other in global monsoonal circulations. Basic to these processes are the commonality of scales among differential heating, the isentropic mass circulation and the transport of dry static energy and entropy by the mass circulation.

6.7 On sources and sinks of entropy

The viscous sink of energy on the left hand sides of Eqs. (6.37a) and (6.37b) is an implicit source of entropy and thus also a source of energy on the right hand side. The energy balance expressed by Eqs. (6.37a) and (6.37b) leaves open the possibility that the atmosphere's circulation could be maintained by net heating from viscous dissipation of mechanical energy through $\dot{s}$, without any requirement for differential heating by nonviscous components. However, this condition immediately violates the mass transport and entropy source constraints of Eq. (6.38), since no counterpart sink of entropy by atmospheric cooling exists to offset the source of entropy by viscous dissipation of kinetic energy.

To establish the balance between entropy decreasing and increasing processes, the relation between the heat addition and time-area-averaged entropy source/sink function in Eq. (6.38) is now expressed by

$$0 = \overline{\rho J_\theta \frac{ds}{dt}}^{tA} = \overline{\rho J_\theta [(Q_m)_{-d} + \varepsilon^2/\rho])/T}^{tA} \quad , \tag{6.48}$$

where ε^2/ρ is the specific heating by viscous process. The remaining components of heat addition $(Q_m)_{-d}$ were defined previously in Eq. (5.2). Since viscous dissipation is an ever present source of entropy (Sommerfeld, 1950; Dutton, 1976a) which occurs independently of the temperature distribution, the other components of atmospheric heating must constitute a net sink of entropy. This relation is expressed by

$$\overline{\rho J_\theta (Q_m)_{-d}/T}^{tA} = -\overline{(\rho J_{\theta\varepsilon^2})/T}^{tA} < 0 \quad . \tag{6.49}$$

This subtle condition from the second law of thermodynamics emphasizes that the time-space domain of atmospheric cooling in global monsoonal circulations is equal in importance to the domain of heating. Sources of heating, particularly latent heat release associated with convection, are intense and dramatic, drawing immediate attention to such phenomena and emphasizing the importance of the tropics. In contrast, the sinks of energy by infrared radiation dominate in larger time-space scales of polar latitudes, desert regions, subtropical anticyclones, etc., which, for the most part, receive less attention. Although the importance of tropical sources of energy can be supported by energy considerations alone, the combination of mass, energy and entropy principles requires that equal importance be given to regions of entropy decrease. It is the combination of the regions of entropy decrease and increase through differential heating that force global monsoonal circulations. Studies of planetary scale processes that create energy sinks and sources must receive equal emphasis in order to properly resolve the thermal forcing of global monsoonal circulations and ultimately climate.

6.8 The link between sources of total energy by differential heating and the generation of available potential energy

The requirement for differential heating and thermally-forced planetary scale circulations was ascertained as a natural consequence of the combined application of mass, energy and entropy principles and the steady constraint, without any recourse to available potential energy concepts; in particular, a reference state atmosphere or its energy was never defined (Lorenz, 1955a; Van Mieghem, 1956, Dutton and Johnson, 1967).

The time-averaged energy source integral Eq. (6.39b) from diabatic heating is equivalent to the generation integral in available potential energy theory (Dutton and Johnson, 1967). To establish this equivalence, in accord with the steady constraint, a time-area-averaged distribution of pressure is defined by

$$p_{r_o}(\theta) = \overline{p}_r(\theta,t) = g\int_{\theta}^{\theta_T} \overline{\rho J_\theta}^{tA}\, d\theta \quad . \tag{6.50}$$

With this definition and the constraint on the time-area-averaged entropy flux, Eq. (6.38), the product of the entropy source and the time-averaged reference state pressure defined by an Exner function π_{r_o}, equal to $c_p(p_{r_o}/p_{oo})^\kappa$, is

$$\overline{\rho J_\theta \dot{s}\, \theta (p_{r_o}/p_{oo})^\kappa}^{tA} = \overline{\rho J_\theta \dot{s}}^{tA}\, \theta \pi_{r_o}/c_p = 0 \quad . \tag{6.51}$$

Subtraction of this quantity from the integrand of Eq. (6.39) and use of Poisson's equation yields

$$\overline{E^2}^{tA} = \int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta \dot{s} T[1-(p_{r_o}/p)^\kappa]}^{tA}\, d\theta \quad . \tag{6.52}$$

With the definition of an efficiency factor by

$$\varepsilon_{r_o} = [1 - (p_{r_o}/p)^\kappa] \ , \tag{6.53}$$

Eq. (6.52) becomes

$$\overline{E^2}^{tA} = \int_{\theta_{s_o}}^{\theta_T} \overline{\rho J_\theta \dot{s} T \varepsilon_{r_o}}^{tA} \, d\theta \,, \tag{6.54}$$

a form that is equivalent to the generation of available potential energy (Dutton and Johnson, 1967).

The equivalence stems from the steady state of the time averaged mass distribution, which in turn demands that the time-average of the reference state atmosphere defined in available potential energy theory be steady. The energy of the hydrostatic reference state atmosphere of available potential energy theory depends solely on the mass distribution with respect to entropy. The combined application of mass, energy, and entropy principles led to the identical result derived by Lorenz (1955a)--that the kinetic energy of the atmosphere is maintained by differential heating--without any recourse to the argument that the kinetic energy of the atmosphere is realized by an isentropic redistribution of mass from the actual to the reference state atmosphere. In this context, controversies over the proper definition of a reference state atmosphere and the proper definition of available potential energy are not of primary importance (Van Mieghem, 1956; Pfeffer et al., 1965).

7. Thermodynamics and global monsoonal circulations

With the emphasis on isentropic analysis in this study, it is important to consider exchange processes that govern the balance of entropy itself. However, in a broader sense, a question basic to atmospheric dynamics concerns the role of entropy in atmospheric circulation. As Dutton (1973) noted, little emphasis has been given to the significance of the second law of thermodynamics except to acknowledge that both kinetic energy dissipation and heat diffusion increase entropy from the requirement that the viscosity and heat conductivity coefficients are positive. It is customary to utilize the first law of thermodynamics, emphasize determinism expressed through the equation of motion and frequently disregard the second law in theoretical, numerical or diagnostic studies through assumptions of isentropic motion. As such, the primary emphasis is on mechanical aspects, without enquiring how entropy sources and sinks in the form of differential heating produce motion and how the motion interacts as part of a thermodynamic process. The reason for the failure of meteorologists to embrace entropy is due in part to a completeness of the governing equations of atmosphere motion with inclusion of heat diffusion and viscous dissipation in energy equations (Tolman and Fine, 1948). However, the reason is also in part related to Shaw's (1930) statement; "Meteorologists generally express aversion from any suggestion to introduce the idea of entropy into meteorological reasoning on the ground that is is an incomprehensible entity which suggests no physical reality and therefore confuses the argument".

In atmosphere circulation both the sources that create and the sinks that destroy total energy are thermodynamic processes involving an exact differential relation between heat addition and entropy change. Sources and sinks of entropy occur through the molecular exchange of energy. Thus, a classic problem of atmospheric thermodynamics is to determine from first principles how thermal processes involving entropy sources and sinks at the molecular scale creates the systematic mass, energy and entropy exchanges at the planetary scale. Insight into this problem is fundamental to our understanding of global monsoonal circulations and also climate.

To a large degree the traditional view of planetary circulation developed over the last four decades is based on the theories of baroclinic instability (Charney, 1947; Eady, 1949) and available potential energy (Lorenz, 1955a, 1967). This view holds that the zonal available potential energy generated by the systematic meridional gradient of differential heating is transformed through baroclinic instability to eddy available potential energy and realized as eddy kinetic energy. Part of the eddy kinetic energy is dissipated by friction and part is transformed to zonal kinetic energy of the circumpolar vortex after which it too is dissipated by friction.

In these concepts, the maintenance of the atmosphere's circulation through the mechanisms and scales of kinetic energy generation are considered to be relatively independent of the differential heating. The generation of zonal available potential energy by differential heating is viewed to be a static process which increases the meridional gradient of temperature. As the temperature gradient increases to a critical value, eddy kinetic energy is realized dynamically by isentropic processes during the instability of the zonal current. Eventually the instability is damped as the meridional temperature gradient is decreased below its critical value of instability and the kinetic energy realized through the instability is removed through the increased frictional dissipation. With the decrease in meridional exchange, the generation of zonal available potential energy again increases the temperature gradient above its critical value for instability and the sequence repeats itself.

An alternative and not necessarily contradictory view that emerges from the results of this study based on mass, energy and entropy exchange is that the differential heating in the form of vertical mass flux through isentropic surfaces is a dynamic component of a thermally forced isentropic mass circulation that is continually in motion. This mass circulation with horizontal and vertical branches demanded by the mass continuity principle in turn brings about the vertical and horizontal exchange of energy and entropy in planetary monsoonal circulations, and at the same time in conjunction with boundary pressure work produces the kinetic energy of motion to maintain these circulations. Through these circulations, the atmosphere finds the means to exchange its momentum and thereby force the horizontal branches of the mass circulations which transport sufficient energy and entropy to satisfy concomitantly the balance required by differential heating. The modes of energy transport in the baroclinic instability process are simply one of the means by which the atmosphere continuously carries out the required energy and entropy exchange. The purpose of this section is to add to this perspective by isolating the reversible isentropic component of total energy through the joint application of basic thermodynamic principles. It will be established that this component of total energy is equivalent to the sum of available potential and kinetic energy (Lorenz, 1955a) and that this definition proceeds logically from thermodynamics without requiring a virtual isentropic redistribution of mass to the reference state of minimum total potential energy.

7.1 On total energy, entropy and available potential energy.

In the steady time-averaged circulation discussed in Section 6.5, the kinetic energy dissipation, an irreversible thermodynamic process, was balanced solely by differential heating with the entropy source occurring at high pressure and the entropy sink at low pressure. The area-averaged entropy source from kinetic energy dissipation did not generate energy in a form that may be transformed to kinetic energy through reversible processes. Such a constraint on the distributions of entropy sources and sinks stems from the second law of thermodynamics which is the fundamental principle for the definition of entropy. The question that naturally follows is: did Lorenz (1955a) succeed in establishing the requirement that the atmosphere's circulation is maintained by differential heating solely from energy principles, or did he succeed from a combination of energy and entropy principles, with the latter being introduced through the definition of the energy of the reference state and the means by which the reference state is determined from the actual atmosphere? There is some disagreement on this issue between several investigators. Both prior to and after Lorenz's (1955a) work, investigators have emphasized the entropy principle in determining the efficiency of heating to create motion (Bridgman, 1941; Tolman and Fine, 1948; Wulf and Davis, 1952; Lettau, 1954a; Dutton, 1973; Livezy and Dutton, 1976a; Pearce, 1978).

In his discussion of basic energy forms and conversions, Lorenz (pp. 97-102, 1967) draws attention to Lettau's (1954a) use of the steady state entropy constraint that differential heating maintains the atmosphere's circulation. Lettau (1954a) used the entropy principle to estimate a value of 2.0 watt/m^2 for the frictional dissipation of kinetic energy in the atmosphere with a remarkably simple formulation. With a steady state assumption, the generation of entropy by the irreversible process of mechanical energy dissipation must be balanced by entropy decreasing processes through systematic meridional distribution of differential heating. In accord with Eq. (6.49), Lettau (1954a) used the systematic mean meridional distributions of heating and temperature to estimate the rate of entropy decrease from which kinetic energy dissipation was then calculated. It is through these considerations that the efficiency of a system to maintain its circulation against irreversible processes can be set forth.

Lorenz (pps. 101-103, 1967) suggests that to utilize the entropy constraint effectively for these purposes, "a knowledge of the vertical variations of Q but not the horizontal variations of Q would be needed." He also notes, however, that the methods used to estimate the generation of energy based on Lettau's use of the entropy principle apparently do not "make full use of the known restrictions upon the distribution of Q," a comment which suggests that other principles were needed to gain insight on these matters.

Lorenz proceeds to state in his discussion of available potential energy that "the sum of APE and KE resembles negative entropy in that friction decreases it and internal reversible adiabatic processes leave it unaltered, while heating is needed to increase it. It is nevertheless a separate concept from entropy, since it involves the field of motion, while entropy depends only upon the thermodynamic state." It is on this point that the development in this section will establish that the sum of A and K defines a reversible component of total energy and that its time rate of change may be isolated from mass, energy and entropy principles alone as a natural consequence of the determination of an equilibrium temperature $T_\alpha(\theta, t)$, the total energy of this state, and its time-rate of change. The development will also establish the relationship between the theory of available potential energy and the entropy concepts of Lettau (1954a), Dutton (1973) and Pearce (1978).

In developing these relationships, the concept of a statistically stationary state will be utilized in lieu of temporal averaging the structure of the atmosphere and thereby ascertaining local balance. In this state, the time dependencies of integrals of basic properties for the entire atmosphere are assumed to vanish from the condition that the global integral of the Lagrangian sources and sinks of the same properties vanishes; however, local tendencies of time averaged structure need not vanish. For the atmosphere as a whole there is no net change by convective transport processes, since the atmosphere's upper and lower boundaries are assumed to be closed with respect to mass transport. Within the statistically stationary state for entropy, $\dot{S} = 0$, it will be useful at times to contrast results when this state is realized from the condition that the mass-weighted area-averaged entropy source vanishes ($\hat{\dot{s}}^A = 0$) with results when it does not vanish ($\hat{\dot{s}}^A \neq 0$). The condition $\hat{\dot{s}}^A$ equal zero everywhere is a sufficient condition for $\dot{S}$ to vanish, however, $\dot{S}$ equal zero is not a sufficient condition for $\hat{\dot{s}}^A$ to vanish.

7.2 A global perspective

In studying the efficiency of differential heating in maintaining the atmosphere's circulation using classical entropy principles, a strategy employed either directly or indirectly is to define the time rate of change of the total potential energy of an atmosphere in unconstrained thermodynamic equilibrium with a uniform temperature T_0. In this case, the change of total potential energy given by the product of T_0 and the time rate of change of the atmosphere's entropy (Dutton, 1973; Pearce, 1978) is expressed by

$$\dot{\Pi}_0 = T_0 \int \rho J_\theta \dot{s} dV_\theta \quad , \tag{7.1a}$$

$$= T_0 \dot{S} \quad . \tag{7.1b}$$

Sommerfeld (1950) emphasizes that equality of temperature within a system is a necessary condition for unconstrained thermodynamic equilibrium and that a process which starts spontaneously from a state of unconstrained equilibrium is impossible. He also notes for an isolated system that at the state of equilibrium, the internal energy assumes a minimum value while the entropy is a maximum. Differential heating in an equilibrium atmospheric state does not immediately create energy in a form that is capable of maintaining the atmosphere's kinetic energy supply against viscous dissipation.

The maintenance of the atmosphere's kinetic energy supply in a statistically stationary state from Eq. (6.19) is expressed by

$$D(E) = G(E), \tag{7.2}$$

where

$$D(E) = - \int \rho J_\theta \underset{\sim}{U} \cdot \underset{\sim}{F} \; dV_\theta \;, \tag{7.3a}$$

$$= \int J_\theta \varepsilon^2 dV_\theta \;, \tag{7.3b}$$

and

$$G(E) = \int \rho J_\theta Q_m \; dV_\theta \;, \tag{7.4a}$$

$$= \int \rho J_\theta \dot{s} T \; dV_\theta . \tag{7.4b}$$

G(E) and D(E) are volume integrals of the Lagrangian source of internal energy and sink of kinetic energy through viscous processes, respectively. For simplicity in this section, it is assumed that boundary pressure and viscous work vanish at the earth-atmosphere interface, thus the equality between the frictional sink of kinetic energy in Eq. (7.3a) and the viscous source of internal energy in Eq. (7.3b) (Johnson and Downey, 1982). This assumption also ensures for the atmosphere as a whole that the Lagrangian source of kinetic energy ($- \underset{\sim}{U} \cdot \nabla p$) and Lagrangian sink of internal energy ($p \nabla \cdot \underset{\sim}{U}$) are equal.

The need for a covariance between entropy sources and temperature to maintain atmosphere circulation is most readily established by dividing the domain of the atmosphere in two regions, one where entropy sources are positive and one where entropy sources are negative (Dutton, 1976a). With this division, the energy balance becomes

$$D(E) = \int_{V_{\theta+}} (\rho J_\theta \dot{s})_+ T_+ \; dV_\theta + \int_{V_{\theta-}} (\rho J_\theta \dot{s})_- T_- \; dV_\theta \;, \tag{7.5}$$

where the plus and minus signs identify state variables and processes within entropy source and sink regions, respectively.

Since the entropy sources in the region of heating and also the entropy sinks in the region of cooling are uniformally of one sign, the energy balance with the use of the mean value theorem is expressed by

$$D(E) = \tilde{T}_+ \int_{V_{\theta+}} (\rho J_\theta \dot{s})_+ \, dV_\theta + \tilde{T}_- \int_{V_{\theta-}} (\rho J_\theta \dot{s})_- \, dV_\theta \quad , \tag{7.6}$$

where the tilde denotes appropriate mean values. However, under the condition of a statistically stationary state with $\dot{S} = 0$, the two integrals of opposite sign are equal in magnitude. Thus, the energy balance may be expressed by

$$D(E) = (\tilde{T}_+ - \tilde{T}_-) \int_{V_{\theta+}} (\rho J_\theta \dot{s})_+ \, dV_\theta \tag{7.7a}$$

$$= - (\tilde{T}_+ - \tilde{T}_-) \int_{V_{\theta-}} (\rho J_\theta \dot{s})_- \, dV_\theta \tag{7.7b}$$

Since the dissipation is positive, these two expressions based on global energy and entropy concepts point out two key facets concerning the question of differential heating and the atmosphere's efficiency. The production of mechanical energy through differential heating requires that nonequilibrium thermodynamic conditions exist in the form of temperature differences. The intensity of the production of mechanical energy by differential heating is primarily dependent on the difference between the mean temperatures of the heat source and the heat sink in combination with the difference between the intensities of the entropy source and the entropy sink. In a statistically stationary state, the atmosphere's efficiency in creating mechanical energy by differential heating is independent of the mean of the two temperatures, $\tilde{T}_+$ and $\tilde{T}_-$.

While these considerations are fundamentally important and have been emphasized in other work previously (e.g., Sandström, 1915, 1916; Bjerknes et al., 1937; Dutton, 1976a), the result Eq. (7.7a or b) is of limited value in applications. In order to appreciate the importance of the entropy principle and to derive expressions that are readily applied, a constraint between the time rate of change of total energy, entropy and the reversible component of total energy of a statistically stationary atmospheric system as a whole will now be studied.

7.3 The constraint between energy and entropy sources of a statistically stationary state

The time-rate of change of the atmosphere's total energy (after Eq. 6.19) will be expressed by

$$\dot{E} = (\dot{\Pi} + \dot{K}) \ , \tag{7.8a}$$

$$= G(E) - D(E) \ , \tag{7.8b}$$

where

$$\dot{E} = \int \frac{\partial}{\partial t} (\rho J_\theta e) \, dV_\theta \ . \tag{7.9}$$

In these relations, E, Π and K are volume integrals of the atmosphere's total, total potential and kinetic energies. Use of the Lorenz convention in isentropic coordinates (Dutton and Johnson, 1967) and appropriate boundary conditions as set forth in Section 2 ensures energy conservation for the system.

For the moment, assume that the atmosphere's total energy is statistically stationary in that the dissipation of energy by irreversible processes D(E) is balanced by the generation of total energy G(E). In the derivation of the theory of available potential energy (Lorenz, 1955a), the only explicit irreversible process considered in the energy balance was the loss of kinetic energy through viscous forces. The increase of entropy by heat flux is implicitly included in the specific heat addition Q_m of Eq. (7.4a). In order to develop a requirement for differential heating through the entropy principle which is identical with the generation integral of available potential energy theory, viscous dissipation is regarded to be the only explicit irreversible entropy increasing process involved in the degradation of energy. As pointed out in Section 6.6, this balance expressed could seemingly be satisfied with the sink of mechanical energy constituting the source of thermodynamic energy through viscous processes. Hence, if the atmospheric system were isolated and only the energy principles were involved, the atmosphere's kinetic energy level could be maintained against viscous dissipation by the heat generated through this process. Such a condition however violates the principle of the second law of thermodynamics since heating by viscous dissipation is an irreversible process.

In view of the inability of an atmosphere with uniform temperature distribution to maintain a supply of kinetic energy against viscous dissipation through differential heating, an equilibrium state with an arbitrary temperature distribution is postulated. With an addition and subtraction of the equilibrium temperature T_e in the integral of Eq. (7.4b), the energy balance is expressed by

$$D(E) = \int \rho J_\theta \dot{s}(T-T_e)\, dV_\theta + \int \rho J_\theta \dot{s} T_e\, dV_\theta \quad . \tag{7.10}$$

In this integral $\rho J_\theta \dot{s}$ is the Lagrangian source of entropy, while the distribution of T_e remains to be determined. The integral of the product of $\dot{s}T_e$ by equivalence with $\dot{s}T$ defines the source of energy of the equilibrium state stemming from entropy sources. An alternative form of Eq. (7.10) is

$$D(E) = \int \rho J_\theta Q_m \left(1 - \frac{T_e}{T}\right) dV_\theta + G(E_e) \ , \tag{7.11}$$

where the generation of total energy of the equilibrium state is defined by

$$G(E_e) = \int \rho J_\theta \dot{s} T_e \, dV_\theta \quad . \tag{7.12}$$

In their study of the energy balance associated with the irreversible production of entropy, Tolman and Fine (1948) defined an "efficiency equation" in which the sources of energy through heat addition, boundary work and convective transport processes offset the degradation of energy by the irreversible production of entropy. For a closed atmospheric system without boundary work or boundary transport of energy, the source of energy by differential heating in their "efficiency equation" simply offsets the sink of energy by irreversible processes. Furthermore, their balance of energy between the source of energy by differential heating and sink of energy through the irreversible production of entropy is identical in form to Eq. (7.11) provided that the last integral of this equation vanishes. The condition that must be satisfied is expressed by

$$G(E_e) = \int \overline{\rho J_\theta}^A \, \widehat{\dot{s} T_e}^A \, dV_\theta \ , \tag{7.13a}$$

$$= \int \overline{\rho J_\theta}^A \left(\hat{\dot{s}}^A \, \hat{T}_e^A + \widehat{\dot{s}^{**} T_e^{**}}^A\right) dV_\theta = 0 \ , \tag{7.13b}$$

where with the use of a mass-weighted area-average and its deviation, the integrand of Eq. (7.12) has been partitioned into mean and eddy components. In a statistically stationary state with the area-averaged source of entropy $\hat{\dot{s}}^A$ equal to zero, the necessary condition to be satisfied for the equilibrium state reduces to

$$G(E_e) = \int \widehat{\rho J_\theta}^A \, \widehat{(\dot{s}^{**} T_e^{**})}^A dV_\theta = 0 \quad . \tag{7.14}$$

Since in general the entropy deviation $\dot{s}^{**}$ does not vanish, the condition requires the areal deviation of the equilibrium temperature defined by

$$T_e^{**} = T_e - \hat{T}_e^A \tag{7.15}$$

to vanish everywhere. In this case the equilibrium temperature T_e is restricted to be a function of potential temperature and time, $T_e(\theta,t)$.

In a statistically stationary state with $\dot{S}$ equal to zero but $\hat{\dot{s}}^A$ unequal to zero, the contraint requires the deviation temperature defined by

$$T_e^{***} = T_e - \hat{T}_e^V \tag{7.16}$$

to vanish everywhere. In this case the equilibrium temperature is restricted to be a function of time, $T_e(t)$.

Now, under the definition that the equilibrium temperature T_e is horizontally uniform, with $\hat{\dot{s}}^A$ equal to zero (spatially uniform, with $\hat{\dot{s}}^A$ unequal to zero) the statistically stationary condition for kinetic energy dissipation and generation of total energy becomes

$$D(k) = G(E) = \int \rho J_\theta \dot{s}^{**}(T-T_e)\, dV_\theta \ , \qquad (7.17a)$$

$$= \int \rho J_\theta \dot{s}^{**}(T\varepsilon_e)\, dV_\theta \ , \qquad (7.17b)$$

$$= \int \rho J Q_m \varepsilon_e dV_\theta \ , \qquad (7.17c)$$

where the efficiency ε_e and the atmosphere's deviation temperature $T\varepsilon_e$ from $T_e(\theta,t)$ [$T_e(t)$] are defined by

$$\varepsilon_e = (1 - T_e/T) \ , \qquad (7.18)$$

$$T\varepsilon_e = (T-T_e) \ . \qquad (7.19)$$

By virtue of the volume integration, the notation for area (or volume) averaging has been suppressed for simplicity in Eqs. (7.17a and b) and wherever redundant in the remainder of this Section. This result shows that generation of total energy occurs where an entropy sink occurs in conjunction with a negative deviation temperature [$(T\varepsilon_e) < 0$] and where an entropy source occurs with a positive deviation temperature [$(T\varepsilon_e) < 0$]. Although not yet specified, the equilibrium temperature T_e is presumably bounded by the extremes of temperature either within an isentropic layer if $T_e(\theta,t)$, or within the atmosphere if $T_e(t)$. This result establishes that the generation of total energy which offsets the atmosphere's dissipation of kinetic energy occurs through the combination of differential heating and the relatively warm ($T\varepsilon_e$>0) and cold ($T\varepsilon_e$<0) air. In a statistically stationary atmosphere, the deviation temperature may be defined with the use of either equilibrium temperature, $T_e(\theta,t)$ or $T_e(t)$ [with $\hat{\dot{s}}^A$ unequal to zero, the area deviation $\dot{s}^{**}$ becomes a volume deviation $\dot{s}^{***}$ in Eqs. (7.17a and b).]

Furthermore, in this statistically stationary state the transformation of total potential to kinetic energy, the dissipation of kinetic energy by irreversible processes, and the generation are equal. Since the transformation between kinetic and total potential energy defines the rate at which energy is transformed by reversible processes, the differential heating as expressed in Eqs. (7.17a and b) in combination with the efficiency also determines the generation of energy for reversible processes.

It is only through the combined use of entropy and energy principles that the requirement for maintenance of the atmosphere's circulation by differential heating is established. The form of this result for G(E) in Eq. (7.17) with T_e equal to $T_e(t_o)$ is identical to Tolman and Fine's (1948) result. The form of this result with T_e equal to $T_e(\theta,t)$ is also identical to the expression for the generation of available potential energy (Lorenz, 1955a). The form is also valid if the equilibrium temperature is $T_e(t)$ or $T_e(t_o)$, as is encountered in application of entropy concepts to the atmosphere (Lettau, 1954a; Dutton, 1973 and Pearce, 1978). Since this requirement for differential heating is established without any need to arbitrarily define an isentropic redistribution of mass, the basis of this result must reside in the combination of mass, energy, and entropy principles. The reason for this link between total energy, entropy and available potential energy will be established later after the energetics of the time dependent state are determined.

7.4 The time dependent state

In a statistically stationary atmosphere, two distributions of the equilibrium temperature, either $T_e(\theta,t)$ or $T_e(t)$, will satisfy the balance constraints expressed by Eq. (7.17a). In view of the seasonal nature of atmospheric circulation, the implications for energy and entropy balance are now considered under time dependent conditions.

The difference in the generation of total energies of actual and equilibrium atmospheres is defined by

$$G(\Delta E_e) = G(E) - G(E_e) \tag{7.20a}$$

$$= \int \rho J_\theta \dot{s}(T - T_e) dV_\theta \tag{7.20b}$$

$$= \int \rho J_\theta \dot{s}\ (T\varepsilon_e)\ dV_\theta \tag{7.20c}$$

In this more general condition, neither the total energy nor the total potential and kinetic energies of the atmosphere need be steady, in that the generation of total energy by heating need not equal the viscous dissipation of total energy or the transformation of total potential to kinetic energy. By substituting Eq. (7.20a) into Eq. (7.8b) and rearranging, the relation for the time rate of change of the atmosphere's total energy is expressed by

$$\dot{E} - G(E_e) = G(\Delta E_e) - D(E)\ . \tag{7.21}$$

With an indefinite vertical integration of the atmosphere's isentropic area-averaged hydrostatic equation of continuity, the relation between the tendency of area-averaged pressure and entropy source is given by

$$\overline{\rho J_\theta \dot{s}}^A = c_p g^{-1} \theta^{-1} \frac{\partial \bar{p}^A}{\partial t_\theta}\ , \tag{7.22}$$

where the Lorenz convention is used to satisfy mass conservation. See Eqs. (2.6) and (2.7). With a subsitution of Eq. (7.22) into Eq. (7.12), $G(E_e)$ becomes

$$G(E_e) = \frac{c_p}{g} \int \frac{T_e}{\theta} \frac{\partial}{\partial t_\theta} \overset{A}{p} \, dV_\theta \quad . \tag{7.23}$$

At this point, a thermodynamic structure for the equilibrium state must be postulated. In general the equilibrium state is assumed to be a motionless, hydrostatic state with the same mass M as that of the actual atmosphere. Its thermodynamic structure at most varies vertically and temporally. The relations of the equilibrium temperature with the equilibrium pressure and hydrostatic dry static energy distribution above ground are

$$T_e = \theta (p_e/p_{oo})^{\kappa} \ , \tag{7.24a}$$

$$= \frac{\partial \psi_e}{\partial s} \ . \tag{7.24b}$$

Below ground, the distributions from the Lorenz convention are specified by Eqs. (2.6) and (2.7). With use of Eq. (7.24a), the time rate of change of E_e determined from G(E) in Eq. (7.23) is expressed by

$$\dot{E}_e = G(E_e) \ , \tag{7.25a}$$

$$= \frac{c_p}{g(p_{oo})^{\kappa}} \int p_e^{\kappa} \frac{\partial}{\partial t_\theta} \overset{A}{p} \, dV_\theta \quad . \tag{7.25b}$$

where the volume of integration includes the underground region from θ equal to θ_{s_o}. In this form, the time rate of change of E_e is determined by the equilibrium mass distribution and the tendency of the isentropic mass distribution of the actual atmosphere. By defining the ratio of the isentropic area-averaged mass distribution to the equilibrium state mass distribution,

$$\beta_1 = \overset{A}{p}/p_e \ , \tag{7.26}$$

$\dot{E}_e$ in Eq. (7.25b) is expressed in two alternative forms by

$$\dot{E}_e = \frac{c_p}{g(p_{oo})^{\kappa}} \int \frac{\beta_1}{(1+\kappa)} \frac{\partial}{\partial t_\theta} p_e^{1+\kappa} + p_e^{1+\kappa} \frac{\partial}{\partial t_\theta} \beta_1 \, dV_\theta \quad , \tag{7.27a}$$

$$= \frac{c_p}{g} \int p_{oo}^{-\kappa} \frac{\partial}{\partial t_\theta} (\beta_1 p_e^{1+\kappa}) - \frac{\kappa}{(1+\kappa)} \beta_1 p_e \frac{\partial}{\partial t_\theta} \left(\frac{T_e}{\theta}\right) dV_\theta \quad . \tag{7.27b}$$

Inspection of Eqs. (7.27a) and (7.27b) reveals that the energy of the equilibrium state may be readily defined from state variables. If the isentropic tendency of β_1 is assumed to vanish in Eq. (7.27a), the specification of β_1 in Eq. (7.26) requires the equilibrium pressure and thus the equilibrium state energy to be determined from the tendency of the $(1 + \kappa)$ power of the isentropically area-averaged pressure. If the isentropic tendency of T_e vanishes in Eq. (7.27b), the equilibrium pressure becomes independent of time, and the equilibrium state energy upon subsitution of Eq. (7.26) is determined by the tendency of the area-average pressure, however, weighted by the κ power of the equilibrium pressure. Lorenz (1955a) has emphasized that the total potential energy of a hydrostatic atmosphere is determined by the mass-potential temperature distribution. Thus, the dependency of the equilibrium state energies on the isentropically area-averaged pressure in either definition is consistent with such a condition. With an assumption that the tendency of β_1 vanishes in Eq. (7.27a), the functional dependency of the equilibrium state temperature must be $T_e(\theta, t)$. With an assumption that the tendency of T_e vanishes in Eq. (7.27b), the functional dependency of the equilibrium state temperature is restricted to be $T(\theta, t_o)$ or $T(t_o)$. It is on this point that the developments of Lorenz's (1955a) concept of available potential energy and Lettau's (1954a), Dutton's (1973) and Pearce's (1978) concepts based on entropy differ.

7.5 The reversible component of total energy ΔE_α

In setting forth the theory of available potential energy, Lorenz (1955a) in effect assumed that β_1 of Eqs. (7.26) and (7.27a) was equal to unity by defining the reference state through an isentropic redistribution of mass. With β_1 equal to unity the equilibrium state pressure distribution and thus the equilibrium state temperature T_α is determined from the atmosphere's mass distribution. This relation defined by integration of the hydrostatic equation is

$$p_\alpha(\theta,t) = g\int_\theta^{\theta_T} \overline{\rho J_\theta}^A \, d\theta \ , \qquad (7.28a)$$

$$= g\overline{m}^A(\theta_T,\theta) \ , \qquad (7.28b)$$

$$= \overline{p}^A(\theta,t) \ , \qquad (7.28c)$$

where $\overline{m}^A(\theta_T,\theta)$ is the isentropically area-averaged mass distribution of the the atmosphere and $\overline{p}(\theta_T)$ is assumed to vanish. In order to identify each of the equilibrium states and corresponding thermodynamic properties, variables for the reversible and other components of total energy defined through $T_e(\theta,t)$, $T_e(t_o)$ and $T_e(\theta,t_o)$ are now subscripted alpha (α), naught (o) and alpha-naught (αo), respectively.

With β_1 equal to unity in Eq. (7.27a), the time rate of change of total energy for the equilibrium state becomes

$$\dot{E}_\alpha = \frac{c_p\,(1+\kappa)^{-1}}{g(p_{oo})^\kappa} \int \frac{\partial}{\partial t_\theta}\, p_\alpha{}^{(1+\kappa)} \; dV_\theta \ . \qquad (7.29)$$

Through use of the hydrostatic relation expressed by Eq. (7.28a) and a differentiation, the equilibrum state pressure is expressed by

$$p_\alpha = \frac{\partial}{\partial\theta}(\theta p_\alpha^{1+\kappa}) + g(1+\kappa)\,\theta p_\alpha^{\kappa}\,\overline{\rho J_\theta}^{A} \quad . \tag{7.30}$$

A substitution of this relation into Eq. (7.29), interchange of temporal and vertical differentiation, and use of Poisson's equation as expressed by Eq. (7.24a) yields

$$\dot{E}_\alpha = \frac{c_p}{g\, p_{oo}^{\kappa}} \int \frac{\partial}{\partial\theta}\left(\theta\, p_\alpha^{\kappa} \frac{\partial p_\alpha}{\partial t_\theta}\right) dV_\theta + \int \frac{\partial}{\partial t_\theta}(\rho J_\theta\, c_p\, T_\alpha)\, dV_\theta \quad . \tag{7.31}$$

With integration to upper and lower boundaries, the first integral of Eq. (7.31) vanishes from the conditions that p_α (equal to $\overline{p}^{A}$) vanishes at θ_T, while both θ and $\partial p_\alpha/\partial t_\theta$ (equal to $\partial \overline{p}^{A}/\partial t_\theta$) vanish at θ equal to zero from the hydrostatic structure of the equilibrium state and mass conservation for the atmosphere. Now with an indefinite integration with respect to time, the equilibrium state energy from Eqs. (7.29) and (7.31) is expressed by

$$E_\alpha = \frac{c_p\,(1+\kappa)^{-1}}{g(p_{oo})^{\kappa}} \int_A \int_0^{\theta_T} p_\alpha^{1+\kappa}\, d\theta\, dA \quad , \tag{7.32a}$$

$$= \int \rho J_\theta c_p T_\alpha\, dV_\theta = M c_p\, \hat{T}_\alpha^{V} \quad , \tag{7.32b}$$

$$= \int \rho J_\theta (c_v T_\alpha + \phi_\alpha) dV_\theta = M(c_v \hat{T}_\alpha^{V} + \hat{\phi}_\alpha^{V}) \quad . \tag{7.32c}$$

The equilibrium state energy E_α is identical with Lorenz's (1955a) total potential energy for the reference state of available potential energy theory. The equivalence of Eqs. (7.32b) and (7.32c) follows from the hydrostatic structure of the reference state and the equality of the geopotential energy with the RT component of enthalphy for integrals extending throughout the vertical extent of the atmosphere. The equivalence of Eqs. (7.32a) and (7.32b) involves the transformation identity, Eq. (7.30), and inclusion of the boundary term by use of the Lorenz convention through which the underground isentropic layer extending from θ equal 0 to θ_{s_0} is included in the vertical integration. See Dutton and Johnson (1967) for details.

With the substitution of Eq. (7.25a) into Eq. (7.21), and the definition of a dissipation function $D(\Delta E_\alpha)$ equal to $D(E)$, the time rate of change of the reversible component of total energy $\Delta \dot{E}_\alpha$ and its relation to available potential and kinetic energies are defined by

$$\Delta\dot{E}_\alpha = \dot{E}-\dot{E}_\alpha = G(\Delta E_\alpha) - D(\Delta E_\alpha) \ , \tag{7.33a}$$

$$= (\dot{\Pi} - \dot{\Pi}_r) + \dot{K} \ , \tag{7.33b}$$

$$= \dot{A} + \dot{K} \ . \tag{7.33c}$$

Here the equality of $\dot{E}_\alpha$ with $\dot{\Pi}_r$ is evident. In the atmosphere, available potential and kinetic energies (Lorenz, 1955a) engage in energy transformations through reversible isentropic processes. In view of the exact differential forms for E_α in Eqs. (7.29) and (7.31), Eq. (7.33) is integrated to define the reversible component of total energy by

$$\Delta E_\alpha = E - E_\alpha \ , \tag{7.34a}$$

$$= A + K \ . \tag{7.34b}$$

This step enables the sum of available potential and kinetic energies to be appropriately identified as the reversible component of total energy. Although the indefinite integration of Eq. (7.29) [and also Eq. (7.31)] with respect to time yields a constant of integration, in noting the equivalence of the expressions for E_α in Eqs. (7.32b and c) with the total potential energy of the atmosphere as it is commonly expressed through enthalpy, the constant has been set to zero. The reversible component of total energy ΔE_α would, however, be independent of this constant, since any constants from the indefinite integration of $\dot{E}$ and $\dot{E}_\alpha$ in Eq. (7.33a) would for consistency simply cancel.

In this development based solely on application of mass, energy and entropy principles, only thermodynamic concepts were involved, in the sense that an equilibrium temperature for each incremental isentropic layer was defined. In a statistically stationary state, this step led to the definition of the generation of a component of total energy by entropy decreasing processes through differential heating. This generation offsets the viscous dissipation of total energy by irreversible processes of entropy production. Then with dependency of the equilibrium state temperature being restricted to sources and sinks of entropy in the atmosphere and an equilibrium temperature determined by the area-averaged pressure, the time dependency of total energy of an equilibrium state was isolated through the isentropic mass continuity equation. It was not necessary to require that total potential energy of a reference state be defined through a virtual isentropic redistribution of mass.

7.6 An energy phase space and reversible isentropic processes

The concept of an energy phase space for reversible isentropic processes in the atmosphere is now developed following Dutton's (1976a) arguments for entropic energy. A schematic of the energy phase space, in which an admissible region of energy levels is defined for reversible isentropic processes, is presented in Fig. 16. The abscissa is the kinetic energy of the atmosphere, while the ordinate is the total potential energy. The family of sloping lines represents isopleths of total energy. Under conditions of isentropic motion, a trajectory representing the exchange between kinetic and total potential energy would be constrained to move along an isopleth of total energy $E(t_0)$. If only the energy principle were involved, one would presume that any state represented by this isopleth would be admissible. However, with utilization of the entropy principle, the energy exchange between kinetic and total potential energy is restricted to the upper left hand segment of each sloping isopleth. The lower boundary of the admissible region of reversible thermodynamic processes is determined by the intersection of E_α (equal to π_r) with the given isopleth of total energy. At the coordinate of (K_{max}, $E_\alpha = \pi_r$) all of the reversible component of total energy would be transformed to kinetic energy, while at the coordinate (0, π_{max}), all of the kinetic energy would be transformed to total potential energy. With the ever continuous differential heating and viscous dissipation such minima and maxima of kinetic energy in the atmosphere are never realized. If all the reversible component of total energy were realized as kinetic energy, its generation would vanish since T would be equal to $T_\alpha(\theta,t)$ (Lorenz, 1955b, 1960). Likewise, if no kinetic energy were present, the effects of differential heating would immediately introduce accelerations and thus produce kinetic energy (Dutton, 1976a).

A trajectory in the energy phase space to higher or lower isopleths of total energy for the atmosphere as a whole requires a nonconvective flux of energy across the system's external boundary and/or release of latent energy within the atmosphere. Both of these constitute an entropy source or sink to the system. Increases of total energy due to an imbalance between G(E) and D(E) in Eq. (7.8b) force a trajectory towards higher isopleths of E, while decreases force the trajectory towards lower isopleths. However, imbalances between G(E) and D(E), if considered apart from the deviation temperature $T\varepsilon_{\alpha}$, do not determine whether or not the admissible region of energy exchange by reversible thermodynamic processes is expanding or contracting. The actual contraction or expansion of the region of reversible isentropic processes is determined by the imbalance between $G(\Delta E_{\alpha})$ and $D(\Delta E_{\alpha})$ as expressed by Eq. (7.33). Since $D(\Delta E_{\alpha})$ always contracts the region of reversible isentropic processes, any expansion must occur through $G(\Delta E_{\alpha})>0$. The region can also be contracted by $G(\Delta E_{\alpha})<0$. With steady conditions the expansion by $G(\Delta E_{\alpha})$ equals the contraction by $D(\Delta E_{\alpha})$.

Tolman and Fine (1948) emphasized the importance of considering the second law and the production of entropy by irreversible processes. In their discussion, Tolman and Fine point out that when the irreversible entropy production is explicitly defined through the definition of viscous stresses "the information provided by the second law equality is already available to us on other grounds". In principle, with employment of viscous dissipation and heat diffusion in atmospheric studies, the second law is implicitly included.

The relevance of the second law, however, is seldom considered. The view that differential heating maintains the atmosphere's circulation against frictional dissipation is based on Lorenz's (1955a) application of the theory of available potential energy to the general circulation. While this view seemingly stems from energy concepts and appears to be independent of the entropy balance, the physical basis for the generation of available potential energy by differential heating stems from thermodynamic relationships between mass, energy and entropy. The relation between the second law and the contraction or expansion of the region of reversible isentropic processes by entropy sources is evident from the following postulate by Carathéodory (1909, 1925) (from Sommerfeld, 1950). "In the neighborhood of every state which can be reached reversibly there exist states which cannot be reached along a reversible adiabatic path, or, in other words, which cannot be reached irreversibly or which cannot be reached at all". Sommerfeld (1950) points out that this postulate suffices to provide a proof of the existence of the second law and entropy. For a more extended development and discussion, see Chandrasekhar (1939) and Margenau and Murphy (1943).

The introduction of the entropy distribution enters through Lorenz's definition of a reference state atmosphere, whose structure in time in combination with the observed atmospheric heating distribution determines a generation of reference state total potential energy. This generation is identical with the generation of total energy by one of a family of possible equilibrium states that may be postulated. The time dependent structure of this equilibrium state, as it is defined by requiring that the total energy of the equilibrium state to be determined by the mass-potential temperature distribution (i.e. mass-entropy distribution), is governed by the isentropic equation of continuity. Thus, although scaled in terms of energy, the time rate of change of the equilibrium state atmosphere is determined by Lagrangian entropy sources in the atmosphere which include both the entropy increase by irreversible processes and decrease through differential heating. Changes in the entropy distribution, when combined with the temperatures of actual and equilibrium states, determine the generation of total energies for actual and equilibrium state atmospheres. In the quasi-equilibrium conditions of atmospheric circulation, the generation of total energy by entropy decreasing processes through differential heating equals its degradation by irreversible entropy production.

7.7 Reversible isentropic processes and the march of seasons.

From the definition of $G(\Delta E_\alpha)$ in Eq. (7.20a), the reversible component $\Delta \dot{E}_\alpha$ of Eq. (7.33a) is also expressed by

$$\Delta \dot{E}_\alpha = [G(E) - G(E_\alpha)] - D(E) \quad . \tag{7.35}$$

Within the energy phase space, net heating [G(E)>0] expands the admissible region of reversible thermodynamic trajectories. A decrease in the equilibrium state energy [$G(E_\alpha)<0$] through an entropy sink also expands the region, while viscous dissipation [D(k)>0] contracts the region. Apart from latent heat release, G(E) is determined by the net flux of energy through the upper and lower boundaries of the atmosphere. If latent energy were included as part of the internal energy, such as defined in the transport equation for total energy in Eq. (6.2), G(E) would be uniquely determined by boundary flux of energy. The gravitational potential and internal energies so defined would correspond with the potential plus internal energies of the actual atmosphere that Lorenz (1978) used in defining moist available potential energy. The reference states however are different. With such a definition, the reversible component of total energy ΔE_α would be conserved by the combination of isentropic motion and zero net boundary flux of latent, sensible and radiative energies. In contrast, $G(E_\alpha)$ is determined by the distribution of entropy sources and sinks internal to the atmosphere, a characteristic of the second law of thermodynamics in that changes associated with the entropy of a system are not solely determined by boundary integrals. Entropy conservation does not occur through isolation as long as heat diffusion and kinetic energy dissipation are present.

With regard to the expansion or contraction of the region of reversible processes by $G(E_\alpha)$, it is informative to express the change in equilibrium state energy by

$$G(E_\alpha) = \int \rho J_\theta \dot{s}\, (\hat{T}^\theta_\alpha + T_\alpha^*) dV_\theta \quad , \tag{7.36a}$$

$$= \hat{T}^{\theta}_\alpha \dot{S} + \int \rho J_\theta \dot{s}\; \hat{T}^A_\alpha (1 - (\hat{T}^\theta_\alpha / T_\alpha)\, dV_\theta \quad , \tag{7.36b}$$

$$= \hat{T}^{\theta}_\alpha \dot{S} + \int \rho J_\theta \dot{s}\, (T_\alpha \varepsilon_\theta) dV_\theta \quad . \tag{7.36c}$$

The mean vertically integrated equilibrium temperature of the atmosphere $\hat{T}^\theta_\alpha(t)$, its deviation $T^*_\alpha(\theta,t)$ and an efficiency $\varepsilon_\theta(\theta,t)$ are defined by

$$\hat{T}^\theta_\alpha(t) = \overline{\int \rho J_\theta \hat{T}^A_\alpha d\theta} / \overline{\int \rho J_\theta d\theta}^{A} \quad , \tag{7.37a}$$

$$= E_\alpha / c_p M \quad , \tag{7.37b}$$

$$T^*_\alpha(\theta,t) = T_\alpha - \hat{T}^\theta_\alpha = T_\alpha \varepsilon_\theta \quad , \tag{7.38}$$

$$\varepsilon_\theta = 1 - (\hat{T}^\theta_\alpha / T_\alpha) \quad , \tag{7.39}$$

where M is the mass of the atmosphere. The functional equivalence of the mass-weighted vertically averaged temperature $\hat{T}^\theta_\alpha$ to the total energy of the equilibrium atmosphere follows from the definition of the equilibrium atmosphere's average enthalpy, given by the product of the atmosphere's mass

M and its mean specific enthalpy, $c_p\hat{T}_\alpha^\theta$. The generation of the equilibrium state energy in Eq. (7.36) has two components. The one component is determined directly by the product of the average equilibrium temperature of the atmosphere $\hat{T}_\alpha^\theta$ and the time-rate change of the atmosphere's entropy $\dot{S}$. The second involves the covariance of the area-averaged entropy source and the deviation equilibrium temperature defined by Eq. (7.38).

Now consider the admissible regions of reversible isentropic processes for the winter and summer circulations of the two hemispheres and the two distinct thermally forced mean isentropic Hadley mass in each hemisphere. Except for the unlikely possibility of an eddy mode of total energy transport, the one-to-one correspondence between scales of differential heating and the transport of energy and entropy by mean isentropic mass circulations within each Hadley circulation requires the net exchanges of mass, energy and entropy between the two zonally averaged isentropic Hadley circulations to vanish. If this is the case, consider that in an energy phase space for each hemisphere the envelope for the admissible region that bounds the exchange between kinetic and available potential energy by reversible processes is greater in the winter than in the summer hemisphere. The expansion of the envelope for the winter hemisphere does not occur through the generation of total energy [G(E)>0] during the fall season. On the contrary this expansion occurs from the negative generation of equilibrium energy [$G(E_\alpha)$<0] through the sink of entropy by radiative processes. During the transition from summer to winter, $G(E_\alpha)$ in Eq. (7.35) is negative and greater in magnitude than G(E), which also is negative. Through this expansion of the envelope for reversible isentropic processes by the sink of entropy, the efficiency of the hemisphere to generate the reversible component of total energy increases, the intensity of the atmosphere's circulation increases and viscous dissipation increases. In addition, the ability of the wintertime circulation to experience larger oscillations by reversible energy transformations between A and K is enhanced.

The expansion of the envelope for reversible processes is due to both components in Eq. (7.36c) being negative, the one associated with a hemispheric decrease of entropy and the other associated with net cooling in isentropic layers where $T_\alpha \varepsilon_\theta$ is positive. Evidence for the former is provided by the decrease in a hemisphere's entropy through the formation of extensive polar air masses during the onset of winter, while evidence for the latter is provided by the increase in static stability of the wintertime airmasses over their summer counterparts. Such an increase in the ability of the atmosphere to maintain a greater intensity for its circulation in the wintertime is consistent with the increased efficiency of a thermodynamic system to produce mechanical energy by extracting heat at a colder temperature.

In contrast, in the transition from winter to summer the generation of equilibrium energy by entropy sources is greater than the generation of E; consequently contraction of the envelope occurs by this process in addition to its contraction by kinetic energy dissipation. The contraction of the envelope of reversible processes by heating is due to both components in Eq. (7.36c) being positive. Evidence for the greater generation of E_α over E by the first component is provided by the increase in a hemisphere's entropy through the formation of extensive subtropical air masses over much of the hemisphere during the onset of summer. The decrease by the second component is manifested through the decrease in static stability of the summer airmasses over their winter counterparts due to net heating in isentropic layers where

$T_\alpha \varepsilon_\alpha$ is positive. As a result of this contraction, the ability of the atmosphere to generate the reversible component of total energy by differential heating as well as the intensity of the differential heating decreases, therefore the intensity of the atmosphere's circulation decreases. Consequently, with a contraction of the envelope that bounds trajectories for energy exchange by reversible isentropic processes, the ability of summertime circulation to undergo planetary scale oscillations is suppressed.

7.8 A component model of total energy $\Delta E_{\alpha o}$

Now a component model of total energy with an equilibrium state determined by $T_{\alpha o} = T(\theta, t_o)$ will be developed. This derivation, which differs from the one for the reversible component E_α, reveals in a direct way the role of isentropic mass exchange in the energetics of the equilibrium state.

The atmosphere's total energy balance from Eq. (7.21) is expressed by

$$\dot{E} - G(E_{\alpha o}) = G(\Delta E_{\alpha o}) - D(E) \ , \qquad (7.40)$$

where the generation integral $G(\Delta E_{\alpha o})$ and the efficiency $\varepsilon_{\alpha o}$, after Eqs. (7.20c) and (7.18), are defined by

$$G(\Delta E_{\alpha o}) = \int \rho J_\theta \dot{s} \ (T \varepsilon_{\alpha o}) dV_\theta \ , \qquad (7.41a)$$

$$\varepsilon_{\alpha o} = (1 - T_{\alpha o}/T) \ . \qquad (7.41b)$$

The relations of the temperature $T_{\alpha o}$ with the hydrostatic distribution of the dry static energy and pressure of the equilibrium state are

$$T_{\alpha o} = \frac{\partial \psi_{\alpha o}}{\partial s} \ , \qquad (7.42a)$$

$$= \theta \left(\frac{p_{\alpha o}}{p_{oo}}\right)^\kappa \ . \qquad (7.42b)$$

The generation of the equilibrium state energy $G(E_{\alpha_o})$ defined from the combination of Eqs. (7.12) and (7.42a) along with a differentiation is expressed by

$$G(E_{\alpha o}) = \int \rho J_\theta \dot{\theta} \frac{\partial \psi_{\alpha o}}{\partial \theta} dV_\theta \ , \tag{7.43a}$$

$$= \int \frac{\partial}{\partial \theta}[\rho J_\theta \dot{\theta} \psi_{\alpha o}) - \psi_{\alpha o} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta}) \ dV_\theta \ . \tag{7.43b}$$

With the use of the isentropic mass continuity equation, and a vertical integration the time rate of change of $E_{\alpha o}$ as it is defined by $G(E_{\alpha o})$ becomes

$$\dot{E}_{\alpha o} = G(E_{\alpha o}) = \int \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta} \psi_{\alpha o}) dV_\theta + \int \psi_{\alpha o} [\frac{\partial}{\partial t_\theta}(\rho J_\theta) + \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})] dV_\theta , \tag{7.44a}$$

$$= - \int_A (\rho J_\theta \dot{\theta} \psi_{\alpha o})_{\theta_s} dA_\theta + \int_A \int_{\theta_s}^{\theta_T} \{[\frac{\partial}{\partial t_\theta}(\rho J_\theta \psi_{\alpha o}) + \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} \psi_{\alpha o})]$$

$$- (\rho J_\theta \frac{\partial \psi_{\alpha o}}{\partial t_\theta} + \rho J_\theta \underset{\sim}{U} \cdot \nabla_\theta \psi_{\alpha o})\} \ d\theta dA \ , \tag{7.44b}$$

where the vertical transport of energy vanishes from the boundary condition that $\rho J_\theta \dot{\theta}$ at θ_T is zero.

In Eq. (7.44b), the tendency and advection of $\psi_{\alpha o}$ are zero from the condition that $\psi_{\alpha o}$ is a function of θ only. With the use of the Leibnitz' rule for differentiation of integrals, Eq. (7.44b) is expressed by

$$\dot{E}_{\alpha o} = \int [\rho J_\theta \psi_{\alpha o} (\frac{\partial \theta_s}{\partial t_\theta} + \underset{\sim}{U} \cdot \nabla \theta_s) - \dot{\theta}]_{\theta_s} + \frac{\partial}{\partial t} \int_A \int_{\theta_s}^{\theta_T} \rho J_\theta \psi_{\alpha o} \ d\theta dA$$

$$+ \int_A \nabla_\theta \cdot \int_{\theta_s}^{\theta_T} \rho J_\theta \underset{\sim}{U} \psi_{\alpha o} d\theta) dA \ . \tag{7.45}$$

The last term vanishes from area integration over the entire atmosphere. The lower boundary term also vanishes from the condition that the tendency plus advection of $\theta_s(\lambda,\phi,t)$ equal $d\theta_s/dt$, which in turn is equated to $\dot{\theta}(\theta_s)$, thereby ensuring mass conservation. With a no-slip boundary condition, the advection vanishes (Johnson and Downey, 1975a). With these simplifications, an indefinite integration with respect to time yields,

$$E_{\alpha o} = \int \rho J_\theta \psi_{\alpha o}\, dV_\theta + C_{\alpha o} , \tag{7.46}$$

where a constant of integration is retained. With the hydrostatic assumption, use of Eq. (7.42a), and a differentiation, the equilibrium state energy is expressed by

$$E_{\alpha o} = -\frac{1}{g}\int \frac{\partial}{\partial\theta}(\overline{p}\psi_{\alpha o})dV_\theta + \overline{p}\frac{\partial\psi_{\alpha o}}{\partial\theta}\, dV_\theta + C_{\alpha o} , \tag{7.47a}$$

$$= \frac{1}{g}\int_A (\overline{p}\psi_{\alpha o})_{\theta_{s_o}}\, dA + \frac{c_p}{g}\int_A\int_{\theta_{s_o}}^{\theta_T} (\overline{p}\, T_{\alpha o}/\theta)dV_\theta + C_{\alpha o} . \tag{7.47b}$$

Through use of the Lorenz convention for the equilibrium state as defined by Eqs. (2.6) and (2.7) in conjunction with the assumption that the earth's surface is uniform at mean sea level, the area integral of Eq. (7.47b) is expressed by

$$\frac{1}{g}\int_A(\overline{p}\psi_{\alpha o})_{\theta_{s_o}}\, dA = \frac{1}{g}\int_A\int_0^{\theta_{s_o}} \overline{p}\frac{\partial\psi_{\alpha o}}{\partial\theta}\, d\theta dA . \tag{7.48}$$

With this result and use of Eqs. (7.42a) and (2.7b), an alternative form for the equilibrium state energy is given by

$$E_{\alpha o} = \frac{c_p}{g}\int_A\int_0^{\theta_T} \overline{p}\, T_{\alpha o}/\theta\, d\theta dA + C_{\alpha o} . \tag{7.49}$$

Eq. (7.49) shows that the equilibrium state energy $E_{\alpha o}$ is determined solely by the mass-potential temperature (entropy) distribution, provided that the distribution of $T_{\alpha o}$ and the constant $C_{\alpha o}$ are specified. Hence, this dependency is consistent with hydrostatic energetics in that the total potential energy is determined by the hydrostatic mass-potential temperature distribution, while entropy sources in conjunction with the temporally invariant equilibrium temperature uniquely determine its time rate of change. With this result, the component of total energy $\Delta E_{\alpha o}$ is given by

$$\Delta E_{\alpha o} = E - E_{\alpha o} \, . \tag{7.50}$$

If the atmosphere were actually in this hydrostatic equilibrium state, the mass distribution would be given by $\rho_{\alpha o} \, J_{\theta\alpha o}$ and its total energy from Eqs. (7.46) and (7.49) would be defined by

$$(E_{\alpha o})_e = \int \rho_{\alpha o} \, J_{\theta\alpha o} \, (c_v T_{\alpha o} + \phi_{\alpha o}) \, dV_\theta \ , \tag{7.51a}$$

$$= \frac{c_p}{g} \int_A \int_o^{\theta_T} p_{\alpha o} \, T_{\alpha o}/\theta \, d\theta dA \ + C_{\alpha o} \ , \tag{7.51b}$$

where the equilibrium mass distribution, the geopotential energy and the constant of integration would be given by

$$\rho_{\alpha o} \, J_{\theta_{\alpha o}} = - \frac{p_{oo}}{g} \frac{\partial}{\partial\theta} \left(\frac{T_{\alpha o}}{\theta}\right)^{1/\kappa} \ , \tag{7.52a}$$

$$\phi_{\alpha o} = - \int_{\theta_{so}}^{\theta} \theta \frac{\partial}{\partial\theta} \pi_{\alpha o} \, d\theta \ , \tag{7.52b}$$

$$C_{\alpha o} = - \kappa (E_{\alpha o})_e \, . \tag{7.52c}$$

In Eq. (7.52b), the boundary condition that $\phi_{\alpha o}(\theta_{s_o})$ equals zero at mean sea level has been used. With these definitions, the energy $(E_{\alpha o})$ for the atmosphere in its equilibrium state from Eqs. (7.46) and (7.49) is expressed by

$$(E_{\alpha o})_e = \frac{1}{(1+\kappa)} \int \rho_{\alpha o} J_{\theta_{\alpha o}} \psi_{\alpha o} \, dV_\theta \quad , \tag{7.53a}$$

$$= \frac{c_p}{g(1+\kappa)} \int_A \int_0^{\theta_T} p_{\alpha o} T_{\alpha o}/\theta \, d\theta dA \quad . \tag{7.53b}$$

One means for specifying the equilibrium temperature $T_{\alpha o}$ is to derive its functional dependency from the isentropic time-area-averaged distribution of pressure. Using this result for the steady time-averaged balance from Eq. (6.50), the equilibrium temperature and pressure distributions are defined in the free atmosphere by

$$T_{\alpha o} = \theta \, [\overline{p}^{At}(\theta)/p_{oo}]^\kappa \quad , \tag{7.54}$$

$$p_{\alpha o} = \overline{p}^{At}(\theta) \; . \tag{7.55}$$

In the underground region of integration, the temperature by Eq. (2.7b) remains a linear function of potential temperature under the requirement of piecewise continuity at θ_{s_o}. A substitution of Eq. (7.54) and (7.55) into Eq. (7.53b) yields

$$E_{\alpha o} = (E_{\alpha o})_e = \frac{c_p(1+\kappa)^{-1}}{g(p_{oo})^\kappa} \int_A \int_0^{\theta_T} (\overline{p}^{At})^{1+\kappa} \, d\theta dA \quad , \tag{7.56}$$

where in this case $E_{\alpha o}$ is equal to $(E_{\alpha o})_e$, since the time-area-averaged pressure in Eq. (7.55) through designation becomes the equilibrium state pressure $p_{\alpha o}$.

A comparison of Eq. (7.56) with Eq. (7.32a) reveals that the equilibrium state energy $E_{\alpha o}$ so defined is also equal to the equilibrium state energy E_α that was derived for reversible isentropic processes, provided of course that the isentropically area-averaged pressure is independent of time. This condition is satisfied if the isentropically area average entropy source $\hat{\dot{s}}^A$ vanishes such as was assumed in the analysis of global monsoonal circulations. In general, however, the energy $E_{\alpha o}$ with its mass specified by $\overline{p}(\theta,t)$ does not equal the energy of the equilibrium atmosphere $(E_{\alpha o})_e$ with its mass specified by the pressure distribution $p_{\alpha o}$. $E_{\alpha o}$ varies in time from the condition that the area-averaged mass in the integrand of Eq. (7.46) and the area averaged pressure in Eq. (7.49) have a time dependence from the vertical mass flux induced by the area-averaged entropy source $\hat{\dot{s}}^A$. By definition, the energy $(E_{\alpha o})_e$ of the equilibrium state corresponding to the energy $E_{\alpha o}$ is independent of time. Any occasion, however, when $\hat{\dot{s}}^A$ vanishes, or if $p_\alpha = p_{\alpha o}$, the generations $G(\Delta E_\alpha)$ and $G(\Delta E_{\alpha o})$ are identical. The equality when $\hat{\dot{s}}^A$ is zero stems from the condition that $G(E_\alpha)$ and $G(E_{\alpha o})$ both vanish under steady conditions for the pressure distributions, $\overline{p}^A(\theta,t)$ and $\overline{p}^{At}(\theta,t)$. In view of the invariance of the time-area-averaged entropy source $\hat{\dot{s}}^A$ over extended time periods, the component model ΔE_{α_o} is most appropriate for climate diagnostics or modeling.

Finally, if an energy phase space is to be defined for $\Delta E_{\alpha o}$, the critical issue is to insure that the difference of energies $(E - E_{\alpha o})$ is positive. This matter will be resolved later after the entropy component model of total energy ΔE_o based on the equilibrium temperature $T_o = T(t_o)$ is derived.

7.9 An entropy component model of total energy ΔE_o

With the same methods used previously, a component model of total energy utilizing the global balance of entropy is now developed based on the definition of an equilibrium temperature $T_o = T(t_o)$. The time rate of change of the atmosphere's total energy from Eq. (7.21) is expressed by

$$\dot{E} - G(E_0) = G(\Delta E_0) - D(E) \ , \tag{7.57}$$

where the generation integral and efficiency for the entropy component of total energy from Eqs. (7.20c) and (7.18) are given by

$$G(\Delta E_0) = \int \rho J_\theta \dot{s}\ T\varepsilon_0\ dV_\theta \ , \tag{7.58a}$$

$$\varepsilon_0 = (1 - T_0/T) \ . \tag{7.58b}$$

The relations of the uniform temperature T_0 with the hydrostatic distribution of dry static energy and pressure of the equilibrium state are

$$T_0(t) = \frac{\partial\psi_0}{\partial s} \ , \tag{7.59a}$$

$$= \theta(p_0/p_{00})^\kappa \ . \tag{7.59b}$$

With the use of Eq. (7.59a), the generation of total energy of the equilibrium state from Eq. (7.12) is expressed by

$$G(E_0) = T_0 S \ , \tag{7.60a}$$

$$= \int \rho J_\theta \dot{s} T_0\ dV_\theta \ , \tag{7.60b}$$

$$= \int \rho J_\theta \dot{\theta} \frac{\partial\psi_0}{\partial\theta}\ dV\theta \ , \tag{7.60c}$$

where the link with the global balance of entropy is evident. By virtue of its dependency on the global entropy balance, this model is called the entropy component model of total energy. Above ground the thermodynamic variables of the equilibrium atmospheric structure with uniform temperature T_0 are defined after Dutton (1973) by

$$p_0 = p_0(\phi_{00})\ e^{-\phi_0/RT_0} \ , \tag{7.61a}$$

$$\rho_0 = \rho_0(\phi_{00})\ e^{-\phi_0/RT_0} \ , \tag{7.61b}$$

$$p_0 = \rho_0 R T_0 \quad , \tag{7.61c}$$

where ϕ_{00} is the reference geopotential that is assigned a value of zero at the lower boundary of the equilibrium atmosphere. Through the hydrostatic relation, the geopotential height, the Montgomery Stream Function and the isentropic mass distribution are expressed by

$$\phi_0 = \phi_{00} + c_p T_0 \, \ln \, [\theta/\theta(\phi_{00})] \quad , \tag{7.62a}$$

$$\psi_0 = \psi_0(\phi_{00}) + c_p T_0 \ln[\theta/\theta(\phi_{00})] \; , \tag{7.62b}$$

$$\rho_0 J_{\theta_0} = p_0(\phi_{00}) \; (\kappa\theta)^{-1} \; e^{-\phi_0/RT_0} \quad . \tag{7.62c}$$

Below ground, the structure will be defined in accord with the Lorenz convention set forth in Eqs. (2.6) and (2.7). Consistency between Poisson's equation and the distribution of dry static energy in Eq. (7.62b) suggests that T_0 be equal to the product of the coldest observed potential temperature $\theta_{so}(\phi_{00})$ and mean sea level pressure $[p_s(\phi_{00})/p_{00}]^\kappa$ to ensure piecewise continuity at the boundary between the underground and above ground portions of the equilibrium state. In contrast to the time dependency of the equilibrium temperature T_α, such that the entire thermodynamic structure for E_α evolves, the equilibrium temperature T_0, the pressure p_0 and the dry static energy ψ_0 specified for the entropy component model in Eqs. (7.61) and (8.62) do not change with time.

Inspection of Eqs. (7.60c) and (7.43a) reveals identical forms of the integral expressions. Thus, the expressions for E_0 may be stated directly from Eqs. (7.46) and (7.49) without repeating the steps involved. The total energy E_0 and the constant of integration are expressed by

$$E_0 = \frac{c_p}{g} \int_A \int_0^{\theta_T} \overline{p}\, T_0/\theta \, d\theta dA + C_0 \, , \tag{7.63a}$$

$$E_0 = \int \rho J_\theta \psi_0 dV_\theta + C_0 \, , \tag{7.63b}$$

$$= \int \rho J_\theta (c_v T_0 + \phi_0) dV_\theta \, , \tag{7.63c}$$

$$= M(c_v T_0 + \hat{\phi}_0^V) \, , \tag{7.63d}$$

$$C_0 = - \kappa (E_0)_e \, , \tag{7.63e}$$

where the equilibrium thermodynamic structure is defined by Eqs. (7.61) and (7.62). Since the equilibrium temperature T_0 is invarient in time and space, its mass weighted average $\hat{T}_0^V$ is simply equal to T_0. The constant of integration in Eq. (7.63e) readily combines with the enthalpy component of Eq. (7.63b), thus the final expression in Eq. (7.63d). The energy E_0 changes with time in accordance with the intensity of entropy sources and sinks weighted by the equilibrium temperature. This change in energy is uniquely linked with time dependency of the isentropic hydrostatic mass distribution, $\overline{p}(\theta,t)$, as represented by Eq. (7.23).

With the substitution of the following relation in Eq. (7.63a),

$$(\overline{p}/\theta) = g\rho J_\theta + \frac{1}{\theta} \frac{\partial}{\partial \theta} (\theta \overline{p}) \, , \tag{7.64}$$

the equilibrium state energy E_0 may be expressed by

$$E_0 = c_p T_0 \int \rho J_\theta dV_\theta + c_p \int_A \int_0^{\theta_T} \frac{T_0}{\theta} \frac{\partial}{\partial \theta} (\theta \overline{p}/g) d\theta dA - \kappa (E_0)_e \, . \tag{7.65}$$

The first integral on the right is equal to the total potential energy of the equilibrium state $(E_0)_e$, since the mass of the actual and equilibrium states are equal. Thus, Eq. (7.65) may be expressed by

$$E_0 = (1-\kappa)(E_0)_e + c_p \int_A \int_0^{\theta_T} \frac{T_0}{\theta} \frac{\partial}{\partial\theta} (\theta \overline{p}^A/g)\, d\theta dA . \tag{7.66}$$

A substraction of Eqs. (7.66) and (7.63c) yields

$$\int \rho J_\theta \phi_0 dV_\theta = \frac{c_p}{g} \int_A \int_0^{\theta_T} \frac{T_0}{\theta} \frac{\partial}{\partial\theta} (\theta\overline{p}) d\theta dA . \tag{7.67}$$

In general, the total energy E_0 as it is determined by Eq. (7.66) from the actual atmospheric mass distribution and the temperature T_0 does not equal to its equilbrium state energy $(E_0)_e$. However, by virtue of a uniform time independent equilibrium temperature T_0, the internal energies I_0 and $(I_0)_e$ (and also the enthalpies) are equal. Thus the time dependency of the total energy is associated solely with the time dependency of the geopotential energy as it is determined by Eq. (7.67). If $\overline{p}(\theta,t)$ in this integral equals the equilibrium state pressure $p_0(\theta)$, then Eq. (7.67) becomes equal as it should to $\kappa(E_0)_e$, i.e. the RT component of equilibrium state enthalpy.

With these results the entropy component model of total energy and its time rate of change from Eq. (7.57) are expressed by

$$\Delta\dot{E}_0 = \dot{E} - \dot{E}_0 = \dot{K} + (\dot{\pi} - \dot{E}_0) = G(\Delta E_0) - D(\Delta E_0) , \tag{7.68a}$$

$$\Delta E_0 = E - E_0 = K + (\pi - E_0) . \tag{7.68b}$$

Whether the difference $(\pi-E_0)$ is positive or negative in this component model depends on the selection of the equilibrium temperature T_0. Ideally, it should remain positive under all states attainable by reversible isentropic processes. This requires that the equilibrium state energy E_0 be less than the reference state energy E_α. If as suggested T_0 is determined by the coldest potential temperature, $(\pi-E_0)$ is positive. This matter is addressed in the next Subsection.

7.10 The relation between the reversible (ΔE_α) and entropy (ΔE_0) component models of total energy

With addition and subtraction of the equilibrium temperature T_α, the generation integral $G(\Delta E_0)$ from Eq. (7.60b) and an efficiency after Eq. (7.18) may be expressed by

$$G(\Delta E_0) = G(\Delta E_\alpha) + \int \rho J_\theta \dot{s}(T_\alpha - T_0) dV_\theta \ , \qquad (7.69a)$$

$$= G(\Delta E_\alpha) + \int \rho J_\theta \dot{s} \ (T_\alpha \varepsilon_{0\alpha}) \ dV_\theta \ , \qquad (7.69b)$$

$$= G(\Delta E_\alpha) + G(\Delta(E_\alpha, E_0)) \ , \qquad (7.69c)$$

$$\varepsilon_{0\alpha}(\theta,t) = (1 - T_0/T_\alpha) \ . \qquad (7.69d)$$

The last term of Eq. (7.69c) defines a generation component associated with the vertical variation of $(T_\alpha - T_0)$ and the area-averaged entropy source $\hat{\dot{s}}^A$. Thus, $G(\Delta E_0)$ is the sum of components associated with horizontal and vertical variations in differential heating and temperature.

In Eq. (7.69c), $G(\Delta E_\alpha)$ generates the reversible component of total energy and maintains the atmosphere's circulation against viscous dissipation, while $G(\Delta E_\alpha, E_0)$ generates a component of energy associated with vertical variation of the equilibrium state temperature T_α. If the isentropic mass distribution is statistically stationary with $\hat{\dot{s}}^A$ equal to zero, the second integral vanishes, the two generation integrals $G(\Delta E_0)$ and $G(\Delta E_\alpha)$ become equivalent, and the increase or decrease of the ranges of ΔE_α and ΔE_0 are identical. As noted earlier, Lorenz (1967) suggested that the concept of maintenance of the atmosphere's circulation based on available potential energy theory, which emphasized horizontal variation of differential heating and temperature, embodied different principles from concepts based on entropy, which outwardly appeared to emphasize vertical variations of differential heating and temperature. However, this analysis shows that the generation of an entropy component of total energy includes as a component the generation of available potential energy. For the most part, the estimates of $G(E_\alpha)$ and $G(E_0)$ will be quite similar, since the global area averaged entropy source is minimal relative to the intensity of the differential entropy sources in an isentropic layer. Lorenz's (1967) suggestion was made apart from an analysis in isentropic coordinates.

With the relation between two generation integrals established, the addition and subtraction of $\dot{E}_\alpha$ yields

$$\Delta\dot{E}_0 = \Delta\dot{E}_\alpha + (\dot{E}_\alpha - \dot{E}_0) \ , \tag{7.70a}$$

$$= G(\Delta E_\alpha) + G[\Delta(E_\alpha,E_0)] - D(\Delta E_\alpha) \tag{7.70b}$$

Now $\Delta\dot{E}_0$ is also composed of two components, $\Delta\dot{E}_\alpha$ associated with the reversible component of total energy, and $\dot{E}_\alpha - \dot{E}_0$ associated with an irreversible component of total energy. With T_0 being time independent, the time rates of change of the two components following the form of Eq. (7.23) and the results for $\Delta\dot{E}_\alpha$ and $\Delta\dot{E}_0$ are expressed by

$$\Delta\dot{E}_\alpha = (\dot{E} - \dot{E}_\alpha) \ , \tag{7.71a}$$

$$= \frac{c_p}{g(1+\kappa)} \int_A \int_0^{\theta_T} \frac{\partial}{\partial t_\theta} [(T - T_\alpha)\ p/\theta] dA d\theta \quad , \tag{7.71b}$$

$$\Delta(\dot{E}_\alpha,\dot{E}_o) = (\dot{E}_\alpha - \dot{E}_o) \ , \tag{7.71c}$$

$$= \frac{c_p}{g(1+\kappa)} \int_A \int_0^{\theta T} \frac{\partial}{\partial t_\theta} [(T_\alpha - (1+\kappa)T_o)p/\theta] \, dA d\theta \ . \tag{7.71d}$$

Furthermore, with the earlier results of integration, Eq. (7.70a) and the two components expressed in Eqs. (7.71a and b) are given by

$$\Delta E_o = \Delta E_\alpha + \Delta(E_\alpha,E_o) = (A+K) + \Delta(E_\alpha,E_o) \ , \tag{7.72a}$$

$$\Delta E_\alpha = (E - E_\alpha) = \frac{c_p}{g(1+\kappa)} \int_A \int_0^{\theta T} [(T-T_\alpha)\, p/\theta] d\theta dA \ , \tag{7.72b}$$

$$= \int \rho J_\theta [c_v(T-T_\alpha) + (\phi-\phi_\alpha)] dV_\theta \ , \tag{7.72c}$$

$$= M[c_v(\hat{T}^V-\hat{T}^V_\alpha) + (\hat{\phi}^V-\hat{\phi}^V_\alpha)] \ . \tag{7.72d}$$

$$\Delta(E_\alpha,E_o) = (E_\alpha - E_o) = \frac{c_p}{g(1+\kappa)} \int_A \int_0^{\theta T} \{[T_\alpha - (1+\kappa)T_o]p/\theta]\} d\theta dA$$

$$+ \kappa(E_o)_e \ , \tag{7.72e}$$

$$= \int \rho J_\theta [c_v(T_\alpha-T_o) + (\phi_\alpha-\phi_o)] dV_\theta \ , \tag{7.72f}$$

$$= M[c_v[\hat{T}^V_\alpha-T_o] + (\hat{\phi}^V_\alpha-\hat{\phi}^V_o)] \ , \tag{7.72g}$$

$$= M[c_p\hat{T}^V_\alpha - (c_vT_o+\hat{\phi}^V_o)] \ . \tag{7.72h}$$

where $\hat{T}^V_\alpha$ and $\hat{\phi}^V_o$ are the mass-weighted volume averages of the equilibrium temperature T_α and the geopotential energy ϕ_o. As noted earlier, the difference component $\Delta(E_\alpha, E_o)$ should be positive. This component will be positive as long as the equilibrium temperature T_o from Eq. (7.72h) satisfies the following inequality,

$$T_o < (c_p \hat{\check{T}}_\alpha - \hat{\check{\phi}}_o)/c_v \quad . \tag{7.73}$$

This decomposition of ΔE_o in Eq. (7.72a) is like that of Pearce's (1978) model. Through a variance analysis, he partitioned the temperature variance of the atmosphere into horizontal and vertical components. He identified the variance of temperature about its isobaric area-averaged temperature with the atmosphere's available potential energy A and the variance of the isobaric area-averaged temperature about the mean global temperature with an unavailable component. His vertical component corresponds with the energy component $\Delta(E_\alpha, E_o)$.

Now, if the equilibrium temperature T_o were defined by

$$T_o = (c_p \hat{\check{T}}_\alpha - \hat{\check{\phi}}_o)/c_v \quad , \tag{7.74a}$$

$$= T_o(\hat{\check{T}}_\alpha, \hat{\check{\phi}}_o) \quad , \tag{7.74b}$$

the energy component $\Delta(E_\alpha, E_o)$ vanishes and Eq. (7.72a) reduces to

$$\Delta E_o = \Delta E_\alpha = (A+K) \quad . \tag{7.75}$$

However, for the equality of ΔE_o and $(A+K)$ to be time independent, $G(\Delta E_\alpha, E_o)$ necessarily must be zero. Under the statistically stationary condition with $\hat{\dot{s}}^A$ equal to zero, $G(\Delta E_\alpha, E_o)$ does vanish and Eq. (7.70b) reduces to

$$\Delta\dot{E}_0 = \Delta\dot{E}_\alpha = (\dot{A}+\dot{K}) \qquad [\hat{\dot{s}}^A = 0,\ T_0 = T_0(\hat{T}^V_\alpha, \hat{\phi}^V_0)] \tag{7.76a}$$

$$= G(E_\alpha) - D(\Delta E_\alpha) \tag{7.76b}$$

Thus, with the same condition that has been used to isolate the thermal forcing of the global monsoonal circulations, i.e. $\hat{\dot{s}}^A$ equal to zero, an equality of the ranges for the reversible isentropic and the entropy component model of total potential energy, ΔE_α and ΔE_0, can be determined from classical entropy concepts provided of course that the equilibrium temperature is suitably determined and time independent. In this context, there can be no doubt that the concepts of the maintenance of atmosphere's circulation based on available potential energy and entropy considerations have a common foundation in thermodynamics stemming from the combination of mass, energy and entropy principles.

In using entropy concepts to study atmospheric energetics and estimate a global value of kinetic energy dissipation, Lettau (1954a) assumed a constant T_0. Thus his approach, as it would be related to ΔE_0, has an equivalence with the results discussed earlier for ΔE_α, the admissible region of reversible isentropic processes. As noted, previously, with $\hat{\dot{s}}^A$ equal to zero, the time dependency of the ranges for ΔE_0 and ΔE_α become identical.

In his development of entropic energy, Dutton (1973) defined an equilibrium temperature determined by the total energy of the atmosphere, and thus its evolution was determined by the rate of change of the atmosphere's total energy. He then defined an energy phase space as the product of T_0 and the difference between the entropies of equilibrium and actual states of the atmosphere. This definition introduces a degree of complexity that precludes an explicit and unique dependency on the hydrostatic mass distribution in isentropic coordinates. At the same time, this definition includes a time dependency for T_O determined by changes in total energy.

Pearce (1978) followed a suggestion by Lorenz (1967) and defined T_O to be the reciprocal of the mass-weighted average of T^{-1} over the entire atmosphere. In this case, the expression in isentropic coordinates for T_O is

$$T_O = (\hat{T^{-1}}^{V})^{-1} = [\int \rho J_\theta T^{-1}\, dV_\theta / \int \rho J_\theta\, dV_\theta]^{-1} , \qquad (7.77)$$

while its time rate of change, a rather complicated expression, is given by

$$\frac{\partial}{\partial t_\theta} T_O = T_O^2 \int_{V_\theta} [\frac{\kappa \rho J_\theta}{Tp} \frac{\partial p}{\partial t_\theta} - (T^{-1})^{***} \frac{\partial}{\partial t_\theta}(\rho J_\theta)] dV_\theta / (\int \rho J_\theta dV_\theta) . \qquad (7.78)$$

Since the temperature T is determined by the pressure distribution in isentropic coordinates and hydrostatic pressure is determined by the mass distribution, Eq. (7.78) does verify that the time rate of change of T_O for Pearce's entropy model is determined by changes in the isentropic mass distribution. While Eq. (7.78) makes possible direct estimation of T_O for use in the entropy component model, complexity is introduced in defining an equilibrium temperature T_O based on the average of the reciprocal of the atmosphere's temperature. Furthermore, the retention of a time dependency for T_O precludes in a strict sense the temporal integration of $\Delta\dot{E}_O$ to define uniquely a component of total energy as a function of state variables provided the energy of the equilibrium state and its time rate of change are required to be specified solely by the joint distribution of mass and potential temperature and its time rate of change.

7.11 Total entropic energy

In his development of the theory of entropic energy, Dutton (1973) defined an equilibrium state $S_{n(t)}$ of maximum entropy. A state of maximum entropy is one of uniform temperature $T_{n(t)}$. In justifying this particular choice of an equilibrium state, he noted that an isolated atmosphere under the constraints of conservation of mass and total energy would naturally evolve to the state of maximum entropy through the entropy increasing processes of viscous dissipation and heat diffusion. With the condition of isolation imposed, he defined the total energy of the actual and equilibrium states to be identical under the integral constraint of energy conservation. However, he noted that the entropy of the equilibrium state would be greater than that of the actual state, since the processes of heat diffusion and viscous dissipation internal to the atmosphere are irreversible, entropy increasing processes. He used these conditions to define total entropic energy as the difference of the entropies of the equilibrium and actual states multiplied by the temperature of the equilibrium atmosphere.

He also emphasized that the theories of available potential energy and entropic energy "have the fundamental difference that the S_n (his S_o) reference state is the equilibrium one towards which an atmosphere in isolation will naturally tend; in available energy theory there is no assurance that the reference state has any physical significance other than representing a portion of the total potential energy that could never become kinetic energy in an isolated atmosphere. He also noted that the idea for the kinetic energy being limited in an isolated system was a dynamical consequence of Gibb's stability concept. While available potential and entropic energies are seemingly quite different concepts, the conditions under which a commonality of these concepts emerges will become evident.

7.11.1 The definition of total entropic energy

Under the constraints of conservation of mass and total energy, and maximization of total entropy, the relations between the temperature and total energy of the equilibrium state $E_n(t)$ defined by Dutton (1973) and the total energy of the actual atmosphere are

$$T_n(t) = E_n(t)/c_pM \quad , \tag{7.79a}$$

$$= E(t)/c_pM \quad . \tag{7.79b}$$

In order to relate the concept of total entropic energy to the concepts developed for the reversible isentropic component ΔE_α and the entropy component ΔE_0 developed herein, it is helpful to distinguish between Dutton's equilibrium state and the ones defined previously; thus, the equilibrium state and associated state variables defined by Dutton (1973) will be subscripted n, a subscript which corresponds to Dutton's symbol for total entropic energy, N.

In defining entropic energy, Dutton established the relation that if two natural states of the atmosphere have the same mass M and total energies E_1 and E_2, then the entropies of their associated equilibrium states with uniform temperatures are expressed by

$$(S_2 - S_1) = c_pM\, \ell n(T_2/T_1) \quad , \tag{7.80a}$$

$$= c_pM\, \ell n(E_2/E_1) \quad . \tag{7.80b}$$

Hence, the relation between entropy and the total energy of equilibrium states is a unique monotonic increasing function. His definition of total entropic energy N(t) (his Eq. 4.10) as the product of the equilibrium temperature T_n and the difference of the entropy of the equilibrium state S_n minus the entropy of the actual state S is expressed by

$$N(t) = T_n(t)[S_n(t) - S(t)] \ , \qquad (7.81a)$$

$$= K(t) + T_n(t)\,\Sigma(t) \ , \qquad (7.81b)$$

where $T_n(t)\,\Sigma$ is the static component of entropic energy. Since the entropy of the actual atmosphere is less than the equilibrium state, the total entropic energy N is always positive. Dutton established that the static entropic energy $T_n\Sigma(t)$ was positive definite. It is through multiplication of the entropies of the equilibrium and actual states by T_n that the difference in entropy of two states becomes a derived property in the two dimensional phase space of kinetic and total potential energies.

7.11.2 The time rate of change of total entropic energy

The time rate of change of total entropic energy (Dutton, 1973; Livezey and Dutton, 1976a) is

$$\dot{N} = T_n(\dot{S}_n - \dot{S}) + \dot{T}_n(S_n - S) \ , \qquad (7.82a)$$

$$= \dot{K} + T_n\dot{\Sigma} + \dot{T}_n\Sigma \ . \qquad (7.82b)$$

With a designation in Eq. (7.80) that the equilibrium state variables E_n, S_n and T_n correspond with E_2, S_2 and T_2 while E_1, S_1 and T_1 are global reference values, a time differentiation yields

$$\dot{S}_n = c_pM\,\dot{T}_n/T_n = \dot{E}_n/T_n = \dot{E}/T_n \ . \qquad (7.83)$$

Hence, the time rates of change of both temperature and entropy of the equilibrium state are determined from the time rate of change of total energy.

A substitution of Eq. (7.83) into Eq. (7.82a) and Eq. (7.81b) into Eq. (7.82b) with rearrangement yields

$$\dot{N} = (\dot{E}_n - T_n\dot{S}) + \dot{T}_n(S_n - S) \ , \tag{7.84a}$$

$$= \dot{K} + T_n(\dot{\Sigma} - K\dot{T}_n/T_n^2) + \dot{T}_n(S_n - S) \ . \tag{7.84b}$$

Total entropic energy increases from a gain in total energy through boundary flux ($\dot{E}_n>0$, $\dot{T}_n>0$ and $\dot{S}_n>0$), but decreases through the entropy increasing processes in the atmosphere ($\dot{S}<0$).

If the net flux of energy through the boundary of the atmosphere vanishes, both $\dot{T}_n$ and $\dot{S}_n$ vanish and the time rate of change of total entropic energy is determined by internal changes of the atmosphere's entropy,

$$\dot{N} = -T_n\dot{S} \ , \qquad [\dot{T}_n = \dot{S}_n = 0] \tag{7.85a}$$

$$= K + T_n\dot{\Sigma} \ . \tag{7.85b}$$

This result emphasizes Dutton's (1973) global perspective that with isolation of the atmosphere, entropy increasing processes in the atmosphere ($\dot{S} > 0$) will inexorably decrease the total entropic energy N to zero. While this result is certainly not to be faulted, in considering the maintenance of the atmosphere's circulation it is necessary to distinguish between the condition of complete isolation of an atmosphere in which the boundary flux of energy vanishes everywhere and the condition in which only the net boundary flux vanishes. In both situations $\dot{E}(=\dot{E}_n)$ is zero; however, in the first $\dot{S}$ would be positive, while in the second $\dot{S}$ need not be positive. Thermally forced global monsoonal circulation may still be maintained in a statistically stationary state with both $\dot{E}_n$ and $\dot{S}$ equal to zero by the combination of a systematic meridional variation of temperature and differential heating. In global monsoonal circulations with $\dot{\hat{s}}^A$ equal to zero (thus $\dot{S}$ equal to zero) the entropy decrease through differential heating within the atmosphere balances the entropy increase from viscous dissipation and sensible heat flux. These conditions are now studied using the strategy that was employed earlier to develop concepts of reversible isentropic and entropy component models of total energy.

7.11.3 The generation of entropic energy by differential heating

With equality of the total energies of actual and equilibrium atmospheres, the time rate of change of Dutton's equilibrium total energy is determined by the flux of energy through the upper and lower boundaries of the atmosphere. Thus, with an integration of Eq. (6.2) over the entire atmosphere and using the conditions that boundary transport and work vanish, the time rate of change of the equilibrium state total energy based on Dutton's definition of internal energy as a dry gas is

$$\dot{E}_n = T_n\dot{S}_n \tag{7.86a}$$

$$= -\int \underset{\sim}{n}\cdot\underset{\sim}{H}d\sigma \quad . \tag{7.86b}$$

However, the corresponding expression defined by Eq. (6.2) used throughout this study is

$$\dot{E}_n = \int \rho J_\theta \dot{s} T\, dV_\theta + \int \rho J_\theta \underset{\sim}{U}\cdot\underset{\sim}{F}\, dV_\theta \quad , \tag{7.87a}$$

$$= G(E) - D(E) \quad . \tag{7.87b}$$

A substitution of Eq. (7.87b) into Eq. (7.84a) yields

$$\dot{N} = [G(E) - D(E)] - T_n\dot{S} + \dot{T}(S_n - S) \quad , \tag{7.88}$$

The quantity, $G(E) - T_n\dot{S}$, in Eq. (7.88) corresponds to the difference in the generation of total energies of the actual and equilibrium atmospheres defined in general by Eq. (7.20) and for the entropy component of total energy by Eq. (7.58a). Thus, the difference in generation, $G(E) - T_n\dot{S}$, is defined by

$$G(\Delta E_n) = G(E) - T_n\dot{S} \ , \tag{7.89a}$$

$$= \int \rho J_\theta \dot{s} T\varepsilon_n \, dV_\theta \ , \tag{7.89b}$$

where the efficiency ε_n and a deviation temperature for entropic energy are defined by

$$\varepsilon_n = (1 - T_n/T) \ , \tag{7.90}$$

$$T\varepsilon_n = (T - T_n) \ . \tag{7.91}$$

With the definition for $G(\Delta E_n)$, the time-rate of change for entropic energy from Eqs. (7.88) and (7.89a) is expressed by

$$\dot{N} = G(\Delta E_n) - D(E) + \dot{T}_n(S_n - S) \ , \tag{7.92}$$

With the equality of the entropy increasing processes by viscous dissipation $D(\Delta E_o)$ in Eq. (7.68a) and $D(E)$ in Eq. (7.92), the subtraction of Eq. (7.68a) from Eqs. (7.92) and (7.84b) respectively along with a transposition of $\dot{T}_n(S_n - S)$ yields

$$[\dot{N} - \dot{T}_n(S_n - S)] - \Delta\dot{E}_o = G(\Delta E_n) - G(\Delta E_o) \ , \tag{7.93a}$$

$$= T_n(\dot{\Sigma} - K\dot{T}_n^2/T_n) - (\dot{\pi} - \dot{E}_o). \tag{7.93b}$$

Equations (7.93a) and (7.93b) constitute the differences in energy generation by differential heating for Dutton's entropic energies and the entropy component model of total energy ΔE_o. For the moment let the equilibrium temperature T_o be determined by T_n, a definition that leads to the equality of $G(\Delta E_n)$ and $G(\Delta E_o)$ in Eq. (7.93a). With the equality of the generation integrals,

$$\Delta\dot{E}_o = \dot{N} - \dot{T}_n (S_n - S) \,, \qquad [T_o = T_n] \qquad (7.94)$$

and

$$\dot{\Pi} - \dot{E}_o = T_n (\dot{\Sigma} - K\dot{T}_n/T_n^2) \,. \qquad [T_o = T_n] \qquad (7.95)$$

Furthermore, if the net boundary flux of energy to the atmosphere vanishes ($\dot{E}_n = \dot{T}_n = 0$), then

$$\Delta\dot{E}_o = \dot{N} \,, \qquad [T_o = T_n \text{ and } \dot{E}_n = 0] \qquad (7.96a)$$

$$= \dot{K} + T_n\dot{\Sigma} \,, \qquad (7.96b)$$

$$= \dot{K} + \dot{A} + (\dot{E}_\alpha - \dot{E}_o) \,, \qquad (7.96c)$$

and

$$\dot{\Pi} - \dot{E}_o = T_n\dot{\Sigma} \,, \qquad [T_o = T_n \text{ and } \dot{E}_n = 0] \qquad (7.97a)$$

$$= \dot{A} + (\dot{E}_\alpha - \dot{E}_o) \,. \qquad (7.97b)$$

Thus, when a statistically stationary condition exists for the total energy of the atmosphere, i.e., $\dot{E}_n = \dot{T}_n = 0$, the time rate of change of the entropy component model of total energy with T_o equal to T_n is identical with the Dutton's time rate of change of entropic energy.

Earlier it was established that the time rates of change of the component models of energy -- $\Delta\dot{E}_\alpha$, $\Delta\dot{E}_{\alpha o}$ and $\Delta\dot{E}_o$ -- were equal in the statistically stationary state with $\hat{\dot{s}}^A$ equal to zero regardless of the distribution of equilibrium temperatures. With the conditions that the net boundary energy flux vanishes ($\dot{E}_n = \dot{T}_n = 0$) and the area-averaged entropy flux $\hat{\dot{s}}^A$ equal to zero, the time rates of change of the reversible component model, the entropy component model and Dutton's total entropic energy are all equal regardless of the distributions of equilibrium temperature, T_α, $T_{\alpha o}$, T_o and T_n.

$$\Delta\dot{E}_o = \dot{N} \ , \qquad [\hat{\dot{s}}^A = 0 \text{ and } \dot{E}_n = 0] \tag{7.98a}$$

$$= \Delta\dot{E}_\alpha \ . \tag{7.98b}$$

Furthermore, the time rates of change the static component of entropic energy $T_n \dot{\Sigma}$, and A are also equal

$$(\dot{\Pi} - \dot{E}_o) = T_n\dot{\Sigma} \ , \qquad [\hat{\dot{s}}^A = 0 \text{ and } \dot{E}_n = 0] \tag{7.99a}$$

$$= \dot{A} \ . \tag{7.99b}$$

These results verify that both the theories of available potential and entropic energies have common underlying principles--the first and second laws of thermodynamics. Thus, the requirement from the entropy principle that atmospheric circulations must be maintained by differential heating is common to the concepts of Lettau (1954a), Lorenz (1955a, 1967), Dutton and Johnson (1967), Dutton (1973) and Pearce (1978). Difference that emerge among the various concepts are due to the temperature structure used to define equilibrium states and the time dependence of the temperature structure.

7.11.4 The components of generation of entropic energy

In Eq. (7.92), entropic energy increases through the covariance of entropy sources and the deviation temperatures $T\varepsilon_n$ and through the boundary flux of energy, while decreases occur through entropy increasing processes of kinetic energy dissipation (and also heat diffusion). The increase from the product of the time rate of change of the equilibrium temperature and the entropy difference of the two states is the process that introduces the nonlinear relation between changes of entropic energy and total energy. Apart from this consideration the component of entropic energy generation by $G(\Delta E_n)$ is identical with results from the entropy model of total energy derived earlier.

With use of Dutton's (1973) mean temperature T_n of approximately 245 K, contributions to the generation $G(\Delta E_n)$ in Eq. (7.92a) will be positive for heating at temperatures greater than 245 K and for cooling at tempertures less than 245 K. This generation $G(\Delta E_n)$, however, divides into three components with the definition of the time rate of change of entropy as the sum of

$$\dot{s} = \hat{\dot{s}}^{V} + \hat{\dot{s}}^{A^*} + \dot{s}^{**} \quad , \tag{7.100}$$

where $\hat{\dot{s}}^V$ is the mass-weighted average entropy source in the atmosphere, $\hat{\dot{s}}^{A^*}$ is the deviation of the mass-weighted area-averaged entropy source from $\hat{\dot{s}}^V$ and $\dot{s}^{**}$ is the deviation of the entropy source from $\hat{\dot{s}}^A$. A substitution of Eq. (7.100) into Eq. (7.89b) yields

$$G(\Delta E_n) = \hat{\dot{s}}^V \int \rho J_\theta T\varepsilon_n dV_\theta + \int \rho J_\theta \, \hat{\dot{s}}^{A^*} T\varepsilon_n dV_\theta + \int \rho J_\theta \dot{s}^{**} T\varepsilon_n \, dV_\theta \; . \tag{7.101}$$

The first term represents generation that its associated with the product of the atmosphere's mean entropy source and the volume integral of $T\varepsilon_n$. From the definitions of $T\varepsilon_n$ in Eq. (7.91) and total potential energy for the actual and equilibrium states, the integral of the first term of Eq. (7.101) is

$$\int \rho J_\theta T\varepsilon_n dV_\theta = (\pi - E_n)/c_p \tag{7.102a}$$

$$= - K/c_p. \tag{7.102b}$$

Thus, the first component of the generation $G(\Delta E_n)$ is proportional to the product of the atmosphere's kinetic energy and the time rate of change of the atmosphere's entropy.

Generation by the second component involves a covariance between the deviations of the mass-weighted area-averaged entropy source and the temperature $T\varepsilon_n$ from vertically averaged values. Generation by the third component involves a positive covariance between the areal deviations of the entropy source and the temperature $T\varepsilon_n$. The second and third components of the generation $G(\Delta E_n)$ correspond with the two components defined by Pearce (1978) and the two components of $G(\Delta E_0)$. The third component of $G(\Delta E_n)$ is is identical with the generation of the reversible component of total energy $G(\Delta E_\alpha)$ and available potential energy $G(A)$. As noted earlier with $\dot{\hat{s}}^A$ and $\dot{E}_n$ equal to zero, the generations by the reversible component $G(\Delta E_\alpha)$, the entropic component $G(\Delta E_0)$, Dutton's entropic energy $G(\Delta E_n)$, and available potential energy $G(A)$ are all equal. In this case the thermodynamics of the maintenance of the atmosphere's circulation by differential heating whether viewed through Lorenz's concept of available potential energy or through Dutton's concept of entropic energy yield identical results.

7.11.5 On the similarity of available potential and entropic energies

The similar nature of Lorenz's concept of available potential energy and Dutton's concept of entropic energy is evident insofar as differential heating maintains the atmosphere's circulation. As emphasized, a primary difference lies in the definition of an equilibrium state temperature and the nature of its time dependency. In the theory of available potential energy, the equilibrium temperature defined by Lorenz's reference state is $T_\alpha(\theta,t)$ with a time dependency determined by changes in the isentropic mass distribution. In the entropy component model of total energy, the equilibrium temperature $T_O(t_O)$ is invariant with time. In the case of Dutton's entropic energy, the equilibrium temperature is $T_n(t)$ with a time dependency determined by the changes of total energy for the system. As previously noted, Dutton's definition of time dependent equilibrium temperature T_n introduces a non-linear relation between $\Delta \dot{E}_O$ and $\dot{N}$ that in general precludes a simple correspondence between ΔE_O and N.

The fact that Dutton (1973) defined an equilibrium state for entropic energy which is different than the equilibrium state for the reversible component of total energy (the reference state of available potential energy theory) should not be confused with the fact that the processes that increase and decrease ΔE_O and N are quite similar. Since $\dot{E}_n$ equals $\dot{E}$, changes in the energy of Dutton's equilibrium state correspond with changes of total energy in the equations for the reversible and entropy component models of total energy, while the changes of entropic energy identified with $T_n\dot{S}$ in Eq. (7.82a) correspond with changes of $\dot{E}_O$ in Eq. (7.68a). Note $\dot{E}_O$ equals $G(E_O)$ by definition, while $G(E_O)$ equals $T_O\dot{S}$ by Eq. (7.60a). With T_O equal to T_n the latter two changes are equal to each other. With $\dot{T}_n$ equal to zero and T_n equal to T_O, $G(\Delta E_O)$ and $\dot{N}$ are equal. The relation among all three energies will now be examined.

Under the conditions of $\dot{T}_n$ equal to zero the relation between the time rate of change of energy and entropy (after Dutton, 1973) is

$$\dot{E} = d(T_n S)/dt \,. \qquad [\dot{T}_n = 0] \qquad (7.103)$$

An integration yields

$$E = T_n S + C_n \,, \qquad [\dot{T}_n = 0] \qquad (7.104)$$

where C_n is the constant of integration in this linear relation between the energy and entropy. With the definition of entropic energy from Eq. (7.81a), use of the Eq. (7.104) to express the total energy E_n by $T_n S_n$ and also the entropy component E_o by $T_n S$ in conjunction with an addition and subtraction of π and E_α, the relation between entropic energy and the entropy component of total energy is

$$N = E_n - E_o = \Delta E_o \,, \qquad [\dot{T}_n = 0 \text{ and } T_n = T_o] \qquad (7.105a)$$

$$= (E_n - \pi) + (\pi - E_\alpha) + [E_\alpha - E_o) \,, \qquad (7.105b)$$

$$= (K + A) + [(E_\alpha - E_o) \,. \qquad (7.105c)$$

In these expressions entropic energy is expressed as a linear combination of the reversible component of total energy and the difference of total energy $(E_\alpha - E_o)$. The constants of integration in Eq. (7.104) from the relation of E_n with $T_n S_n$ and also E_o with $T_n S$ cancel. From Eq. (7.81b), N is also expressed by

$$N = K + T_n \Sigma \quad . \qquad (7.106)$$

Noting the equality of ΔE_α and A + K derived earlier, the relation between N and ΔE_α is expressed by

$$N = \Delta E_\alpha + (E_\alpha - E_o) \,, \qquad [\dot{T}_n = 0 \text{ and } T_n = T_o] \qquad (7.107)$$

while the difference of Eq. (7.106) and (7.107) yields

$$T_n\Sigma = A + (E_\alpha - E_o) \, . \qquad [\dot{T}_n = 0 \text{ and } T_n = T_o] \qquad (7.108)$$

The question to be resolved is whether or not the range of N is larger or smaller than the range of ΔE_α. If it is larger, then $(E_\alpha - E_o)$ is positive, if it is smaller then $(E_\alpha - E_o)$ is negative. This question is now examined deductively.

With an assumption of isentropic motion ($\dot{E}_\alpha = 0$ and $\dot{S}_n = \dot{S} = 0$), the time rate of change of the energies from Eqs. (7.105c) and (7.106) are

$$0 = \dot{K} + \dot{A} \, , \qquad [\dot{E} = \dot{E}_\alpha = 0 \text{ and } \dot{S}_n = \dot{S} = 0] \qquad (7.109a)$$

$$0 = \dot{K} + T_n\dot{\Sigma} \, . \qquad (7.109b)$$

Eq. (109b) states that the maximum kinetic energy attainable from entropic energy upon depletion of its static component through reversible isentropic processes with $T_n \, \Sigma$ equal to zero is

$$K_{max}(T_n\Sigma) = K(t_1) + T_n\Sigma \, (t_1) \; , \quad [T_n\Sigma = 0] \qquad (7.110)$$

while Eq. (7.109a) states that the maximum attainable kinetic energy upon depletion of the available potential energy through like processes is

$$K_{max}(A) = (A(t_1) + K(t_1) \; , \qquad [A = 0] \qquad (7.111)$$

where t_1 is an arbitrary initial time. If the range of $K + T_n\Sigma$ were greater than A + K, then the excess kinetic energy of $K_{max}(T_n\Sigma)$ over and above $K_{max}(A)$ would require that A become negative by the amount of this difference. Since (A + K) constitutes the range of all reversible thermodynamic processes within the range of total energy, such a result is contradictory to the thermodynamics involved unless one presumes that the difference K_{max} $(T_n\Sigma)$ minus $K_{max}(A)$ is a component of static entropic energy that cannot be transformed to kinetic energy.

On the contrary, if the range of (A + K) were greater than the range of $K + T_N\Sigma$, total entropic energy conservation under reversible transformations restricts the maximum of K attainable from A unless $T_n\Sigma$ becomes negative. However, this is also contradictory to the embodied thermodynamics since the theory of entropic energy does not place any restrictions on the range of reversible processes and $T_n\Sigma$ is always a positive definite quantity (page 101, Dutton, 1973).

Thus, the combination of the two conditions require for physical consistency that

$$K + T_n\Sigma \geqslant K + A \ , \qquad (7.112a)$$

$$T_n\Sigma \geqslant A \qquad (7.112b)$$

i.e., the energy associated with reversible isentropic processes must always be less than total entropic energy.

Since the range of $K + T_n\Sigma$ must exceed the range of (K+A), one may partition entropic energy into the component ΔE_α equal to (A + K) involving reversible isentropic processes and another component defined by $(T_n\Sigma - A)$ involving irreversible processes of a nonnegative quantity. Dutton's (1973) computational results (his Table (6.1), p. 104) reveal that A is on the order of 10^{-2} $T_n\Sigma$, thus substantiating that the range of ΔE_α is less than the range of N. In Fig. 16, an energy level T_nS is indicated on the axis of total potential energy at a level less than the energy level E_α in accord with Dutton's results. Entropy sources at high temperatures and sinks at low temperatures within an isentropic layer maintain ΔE_α, while entropy sources in lower isentropic layers and sinks in the upper isentropic layers maintain $(E_\alpha - E_0)$ under the restricted condition that $\dot{E}_n$ equals zero. With $\dot{E}_n$ unequal to zero, noninearity is introduced in the relation between ΔE_0 and N. Finally, as Dutton emphasized, if the atmosphere were isolated, then the entropy increasing processes of heat diffusion and viscous dissipation would force N to zero; however as indicated by the figure and the results herein, such a condition would also force ΔE_α to zero.

In these concepts, much research has been devoted to the reversible component ΔE_α [equal to (A+K)] of total energy; however, little attention has been devoted to the component [N $-\Delta E_\alpha$), or its counterpart $(E_\alpha - E_0)$. This is due to a neglect of the role of heat diffusion in increasing the atmosphere's entropy and thereby degrading energy in its ability to produce circulation. On this matter, the contrast of tropospheric and stratospheric energy balance provokes an interesting consideration.

The reversible component of total energy in the stratosphere is maintained by the vertical propagation of energy from the troposphere (Charney and Drazin, 1961). With its relatively uniform vertical distribution of temperature, the stratosphere constitutes a region of minimum $(E_\alpha - E_0)$, while the tropsophere with a marked lapse rate constitutes a region of large $(E_\alpha - E_0)$. Baroclinic stability is suppressed in the stratosphere by its high static stability (Holton, 1972), a state of maximal entropy in a relative sense (Dutton, 1973). In contrast, baroclinic stability in the troposphere is enhanced by low static stability, equilvalently a state of minimal entropy in a relative sense. With minimal thermal forcing of global monsoonal circulations in the stratosphere and thus no requirement for energy or entropy transport by amplifying baroclinic waves, the relative uniform temperature of this region likely involves a delicate atmospheric energy balance that is constrained by the irreversible entropy increasing processes of both heat diffusion and viscous dissipation. In contrast, with a sustained systematic thermal forcing of global monsoonal circulations in the troposphere, and thus a requirement for energy and entropy transport, the baroclinic structure involves an atmosphere energy balance that is much less constrained by irreversible entropy increasing processes. A contrast of the troposphere's and stratosphere's mass, energy and entropy exchange should provide insight on the role of entropy sources and sinks in modulating the intensity and mode of global monsoonal circulations and in the vertical propagation of energy from the troposphere to the stratosphere.

7.12 Reversible and entropy component models of total energy for open systems

The theory of available potential energy (Lorenz, 1955a; Dutton and Johnson, 1967) is based on the concept that the reference state atmosphere involves a global isentropic redistribution of mass. Johnson (1970) extended the theory to open systems with lateral boundaries extending vertically from the earth's surface to the top of the atmosphere. However, some question the results for open systems due to the need to determine a reference state by the isentropic redistribution of mass within the limited region. This constraint, however, is no more arbitrary than the constraint of an isentropic mass redistribution within an atmosphere that is continuously undergoing differential heating. The problem of open systems is always encountered in atmospheric applications of the theory of available potential energy, since data is never available throughout the entire vertical extent of the atmopshere. It will become apparent that the modes for vertical exchange of available potential energy are equally if not more complicated than the ones for horizontal exchange. For progress in understanding the interaction between a region and its environment, the importance of evaluation of boundary processes must be recognized whether the region and its environment be the troposphere and stratosphere or a quasi-Lagrangian phenomenon and its environment (Johnson and Downey, 1975a).

In their development of energy balance for a steady process, Tolman and Fine (1948) developed an energy equation which, in effect, included all the processes set forth by Johnson (1970) in the available potential energy theory for open systems. Thus, the equations developed earlier by Johnson (1970) and those to be developed herein for the reversible component of total energy of open systems have a basis in the early work of Tolman and Fine (1948). From a theoretical view point these relations are well founded. Numerical results from isentropic modeling (Anthes, 1971; also discussed in Johnson, 1970; Gall, 1972) have also shown that the balance of kinetic-available potential energy for open systems is accurately determined from calculations of generation, conversion, boundary work and transport.

In view of the development of integral equations for the reversible component of total energy for the atmosphere as a whole without requiring a virtual isentropic redistribution of mass to a horizontally uniform stably stratified state, it seems timely to consider the development of like equations for open systems. From the postulate of a state of thermodynamic equilibrium and joint application of mass, energy and entropy principles for open systems, the reversible component of total energy and its time rate of change will be defined for isentropic layers that are limited in both vertical and horizontal extent. Thus, both the vertical and horizontal exchange of the reversible component of total energy can be defined and examined. The development will also reveal how internal modes of energy exchange are impacted by the different choices made in the definition of the temperature of equlibrium states for both the available potential energy concept of Lorenz (1955a) and the entropy concepts of Lettau (1954a), Dutton (1973) and Pearce (1978).

The reasons for setting forth these equations for open systems at this time are two fold. The first is to establish a more general theoretical basis employing transport theory and thermodynamic principles for the reversible and entropic components of total energy and thereby be able to discuss how exchange processes through both lateral and vertical boundaries impact their time rates of change. The second is a practical one stemming from the need to study the thermodynamics of regional circulations and thereby focus on physical processes within monsoons. It is only through the employment of kinetic-available potential energy theory that sufficient accuracy is realized in diagnostic studies of the maintenance of the atmosphere's energy cycle, either globally or regionally. This development also provides a theoretical basis for the discussion of the march of the seasons and the circulation in summer and winter hemispheres that was offered in Section 7.7.

7.12.1 Constraints on the equilibrium state imposed by requirements for reversible transformations

The development here is based on the premise that an equilibrium state can be defined for an open system by intensive thermodynamic variables. With specification of the dependence of the total energy of the equilibrium state on the isentropic mass distribution within a limited domain, definitions for both the reversible and the entropic components of total energy for open systems can be stated and sources through differential heating, boundary transport and boundary work can be determined. However, before determining these relations, internal constraints on the equilibrium state imposed by requirements for reversible isentropic processes will be examined.

The Lagrangian source of kinetic energy is expressed in generalized coordinates as the sum of reversible and irreversible processes by

$$\rho J_\eta \frac{dk}{dt} = \rho J_\eta [s_r(k) + s_{ir}(k)] \quad , \tag{7.113}$$

where the reversible and irreversible source/sink functions are

$$s_r\,(k) = -\,\underset{\sim}{U} \cdot (\alpha \nabla p + \nabla \phi) \quad , \tag{7.114}$$

$$s_{ir}(k) = \underset{\sim}{U} \cdot \underset{\sim}{F} \quad . \tag{7.115}$$

For discussion purposes here, the kinetic energy sink, $\underset{\sim}{U} \cdot \underset{\sim}{F}$, is assumed locally equal to the positive dissipation function, ε^2/ρ. In applying entropy concepts for an open system, the equilibrium state will be specified by intensive thermodynamic variables. Variables and processes within this state when appropriately multiplied by mass will be integrable, obey a continuity principle and determine the time rate of change of the equilibrium state. However, in the energy-entropy relations, physical processes that must remain invariant are the transformation between kinetic and total potential energy by reversible processes Eq. (7.114) and the viscous dissipation of kinetic energy by irreversible processes Eq. (7.115).

Following the earlier general development, T_e is the equilibrium temperature and ψ_e is the dry static energy within the domain of the system. The domain of the system may be limited in both vertical and horizontal extent.

Let the Lagrangian source of kinetic energy by reversible processes from Eq. (7.114) expressed by

$$s_r(k) = - \underset{\sim}{U} \cdot (\nabla\psi - T\nabla s) \ , \tag{7.116}$$

$$= - \underset{\sim}{U} \cdot [\nabla(\psi-\psi_e) - (T-T_e)\nabla s] - \underset{\sim}{U} \cdot (\nabla\psi_e - T_e\nabla s)] \ , \tag{7.117}$$

where the thermodynamic relation, (Milne-Thomson, 1960; Batchelor, 1967)

$$\alpha\nabla p + \nabla\phi = \nabla\psi - T\nabla s \ , \tag{7.118}$$

has been used and an equivalent function of Eq. (7.118) for the equilibrium state,

$$\underset{\sim}{G}_e = \nabla\psi_e - T_e\nabla s \ , \tag{7.119}$$

has been added and subtracted in Eq. (7.117). This vector function is the sum of the equilibrium state's pressure gradient and gravitational accelerations. In Eq. (7.116) the source of kinetic energy by reversible processes involves the advection of dry static energy and entropy. While Eq. (7.117) is mathematically valid regardless of whether or not the last term on the right hand vanishes, for invariance of Lagrangian source/sink functions it is essential to determine an equilibrium state that remains independent of the reversible isentropic process involved in kinetic energy generation. Thus, the vector $(\nabla\psi_e - T_e\nabla s)$ must either vanish or always be orthogonal to to the velocity $\underset{\sim}{U}$. With a transformation to generalized coordinates, Eq. (7.119) is expressed by

$$\nabla\psi_e - T_e\nabla s = (\nabla_\eta\psi_e - T_e\nabla_\eta s) + (\underset{\sim}{k} - \nabla_\eta z)\left(\frac{\partial z}{\partial s}\right)^{-1}\left(\frac{\partial\psi_e}{\partial s} - T_e\right) \quad . \tag{7.120}$$

Sufficient conditions for $(\nabla\psi_e - T_e\nabla s)$ to vanish are

$$\eta = \theta \quad , \tag{7.121}$$

$$\psi_e = \psi_e(\theta, t) \quad , \tag{7.122}$$

$$T_e = \partial\psi_e/\partial s \quad . \tag{7.123}$$

With

$$T_e = \theta(p_e/p_{oo})^\kappa \quad , \tag{7.124}$$

Eq. (7.123) is a statement of hydrostatic balance for the equilibrium state. Through use of Eqs. (7.121) through (7.124), the source of kinetic energy by reversible processes is now expressed by

$$\rho J_\theta s_r(k) = -\rho J_\theta \underset{\sim}{U}\cdot[\nabla_\theta(\psi-\psi_e) + (\underset{\sim}{k}-\nabla_\theta z)\left(\frac{\partial z}{\partial s}\right)^{-1}\left(\frac{\partial\psi}{\partial s} - T\right)] \quad . \tag{7.125}$$

With the definition of the hydrostatic defect (Dutton and Johnson, 1967)

$$\chi = \left(\alpha\frac{\partial p}{\partial z} + \frac{\partial\psi}{\partial z}\right) = \left(\frac{\partial z}{\partial s}\right)^{-1}\left(\frac{\partial\psi}{\partial s} - T\right) \quad , \tag{7.126}$$

Eq. (7.125) simplifies to

$$\rho J_\theta s_r(k) = -\rho J_\theta \underset{\sim}{U}\cdot[\nabla_\theta(\psi-\psi_e) + (\underset{\sim}{k} - \nabla_\theta z)\chi] \quad , \tag{7.127}$$

a result which is identical with a multiplication of Eq. (7.116) by ρJ_θ. The first right term of Eq. (7.127) constitutes the source of horizontal kinetic energy, while the last constitutes a source of both horizontal and vertical kinetic energies, which vanishes in a hydrostatic atmosphere. For simplicity, the atmosphere is assumed to be in hydrostatic balance, consequently Eq. (7.127) reduces to a source of horizontal kinetic energy given by

$$\rho J_\theta s_r(k_h) = -\rho J_\theta \underset{\sim}{U}\cdot\nabla_\theta(\psi-\psi_e) \quad . \tag{7.128}$$

These results in conjunction with Eqs. (7.121) through (7.124) verify that an equilibrium state can be defined by intensive thermodynamic properties within the open system, which when multiplied by mass becomes extensive and is integrable. The invariance needed for the reversible transformation between kinetic and total potential energies within the domain of the system is also ensured. Fortunately, the conditions needed to ensure this invariance are identical with the integral conditions used earlier to derive reversible and entropic component models of total energy for the atmosphere as a whole.

7.12.2 An energy continuity equation for open systems

In the preceding section it has just been established that the states determined by the equilibrium temperatures $T_e(\theta,t)$, $T_e(\theta,t_o)$ and $T_e(t_o)$ correspond with the most general conditions that can be defined under the constraint that the Lagrangian source of kinetic energy by reversible processes must remain invariant. In order to study thermodynamical exchange processes and the generation of kinetic energy within an open system under these conditions, the development begins by defining the energy sources for the actual and equilbrium states by differential heating within the system. Then the isentropic mass continuity equation is used to determine a general continuity equation, after which the continuity equations appropriate to the reversible and entropic components of total energy will be determined.

With use of the equilibrium temperature T_e, the difference in the generation of total energy $g(\Delta e_e)$ between the actual and equilibrium atmospheres is defined by

$$\rho J_\theta g(\Delta e_e) = \rho J_\theta \dot{s}\ (T-T_e)\ , \tag{7.129a}$$

$$= \rho J_\theta \dot{\theta} \frac{\partial}{\partial \theta} (\psi-\psi_e)\ . \tag{7.129b}$$

Through differentiation and use of the isentropic continuity equation, the generation is expressed by

$$\rho J_\theta g(\Delta e_e) = \frac{\partial}{\partial \theta} [\rho J_\theta \dot{\theta}(\psi-\psi_e)] + (\psi-\psi_e)[\frac{\partial}{\partial t_\theta} (\rho J_\theta) + \nabla_\theta \cdot (\rho J \underset{\sim}{U})]\ , \tag{7.130a}$$

$$= \{\nabla_\theta \cdot [\rho J_\theta \underset{\sim}{U}(\psi-\psi_e)] + \frac{\partial}{\partial \theta} [\rho J_\theta \dot{\theta}(\psi-\psi_e)]\}$$

$$- \rho J \underset{\sim}{U} \cdot \nabla_\theta(\psi-\psi_e) + (\psi-\psi_e) \frac{\partial}{\partial t_\theta} (\rho J_\theta)\ . \tag{7.130b}$$

The transport equation for kinetic energy,

$$\frac{\partial}{\partial t_\theta} (\rho J_\theta k) + \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} k) + \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta} k) = - \rho J_\theta \underset{\sim}{U} \cdot \nabla_\theta (\psi-\psi_e) + \rho J_\theta d(e_e), \tag{7.131}$$

in combination with Eq. (7.130b) and area-averaging yields

$$\overline{\rho J_\theta [g(\Delta e_e) - d(\Delta e_e)]} = \overline{\{\nabla_\theta \cdot [\rho J_\theta \underset{\sim}{U}(k+\psi-\psi_e)]\}} + \overline{\frac{\partial}{\partial \theta} [\rho J_\theta \dot{\theta}(k+\psi-\psi_e)}$$

$$+ \frac{\partial}{\partial t_\theta} \overline{(\rho J_\theta k)} + \overline{(\psi-\psi_e) \frac{\partial}{\partial t_\theta} (\rho J_\theta)}\ . \tag{7.132}$$

Through a partitioning of ψ and ψ_e into internal and gravitational energies and an RT work component, and a differentiation by parts, the mass tendency component of Eq. (7.132) is expressed by

$$\overline{(\psi-\psi_e)\frac{\partial}{\partial t_\theta}(\rho J_\theta)} = \frac{\partial}{\partial t_\theta}\overline{\{\rho J_\theta[c_v(T-T_e) + (\phi-\phi_e)]\}}$$

$$+ [\overline{RT\frac{\partial}{\partial t_\theta}(\rho J_\theta)} - \overline{\rho J_\theta\frac{\partial}{\partial t_\theta}(c_v T)} - \overline{\rho J_\theta\frac{\partial \phi}{\partial t_\theta}}]$$

$$- [RT_e\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta}) - \overline{\rho J_\theta}\frac{\partial}{\partial t_\theta}(c_v T_e) - \overline{\rho J_\theta}\frac{\partial \phi_e}{\partial t_\theta}] \; . \quad (7.133)$$

With use of the hydrostatic relations for the actual and equilibrium states,

$$\overline{\rho J_\theta} = -\frac{1}{g}\frac{\partial \overline{p}}{\partial \theta} \; , \quad (7.134)$$

$$\rho_e J_{\theta_e} = -\frac{1}{g}\frac{\partial p_e}{\partial \theta} \; , \quad (7.135)$$

the ratio of the mass distributions is defined by,

$$\beta_2 = \overline{\rho J_\theta}/\rho_e J_{\theta_e} = \frac{\partial \overline{p}}{\partial \theta}/\frac{\partial p_e}{\partial \theta} = \frac{\partial \overline{p}}{\partial p_e} \; , \quad (7.136)$$

while the relation between β_1 defined in Eq. (7.26) and β_2 is

$$\beta_2 = \beta_1 + p_e\frac{\partial \beta_1}{\partial \theta}/\frac{\partial p_e}{\partial \theta} = \beta_1 + p_e\frac{\partial \beta_1}{\partial p_e} \; . \quad (7.137)$$

A subsitution into Eq. (7.133) and rearrangement yields

$$\overline{(\psi-\psi_e)\frac{\partial}{\partial t_\theta}(\rho J_\theta)} = \overline{\frac{\partial}{\partial t_\theta}\{\rho J_\theta[c_v(T-T_e) + (\phi-\phi_e)]\}}$$

$$+ [\overline{RT\frac{\partial}{\partial t_\theta}(\rho J_\theta)} - \overline{\rho J_\theta \frac{\partial}{\partial t_\theta}(c_v T)} - \overline{\rho J_\theta \frac{\partial \phi}{\partial t_\theta}}]$$

$$- \beta_2 [RT_e \frac{\partial}{\partial t_\theta}(\rho_e J_{\theta_e}) - \rho_e J_{\theta_e} \frac{\partial}{\partial t_\theta}(c_v T_e) - \rho_e J_{\theta_e} \frac{\partial \phi_e}{\partial t_\theta}]$$

$$- \rho_e J_{\theta_e} RT_e \frac{\partial \beta_2}{\partial t_\theta} \quad . \qquad (7.138)$$

In Eq. (7.138), β_2 is a scaling parameter which in effect both normalizes processes in the equilibrium state to specific units through division by $\rho_e J_{\theta_e}$, and assigns proper weight within any isentropic layer through multiplication by the area-averaged mass $\overline{\rho J_\theta}$. Now, the height tendency terms of Eq. (7.138) are expressed by

$$\overline{\rho J_\theta \frac{\partial \phi}{\partial t_\theta}} = - \frac{\partial}{\partial \theta}(p \overline{\frac{\partial z}{\partial t_\theta}}) + \overline{RT \frac{\partial}{\partial t_\theta}(\rho J_\theta)} - \overline{RTJ_\theta \frac{\partial \rho}{\partial t_\theta}} \quad , \qquad (7.139)$$

$$\rho_e J_{\theta_e} \frac{\partial \phi_e}{\partial t_\theta} = - \frac{\partial}{\partial \theta}(p_e \frac{\partial z_e}{\partial t_\theta}) + RT_e \frac{\partial}{\partial t_\theta}(\rho_e J_{\theta_e}) - RT_e J_{\theta_e} \frac{\partial}{\partial t_\theta} \rho_e \quad . \qquad (7.140)$$

With the combined substitution of Eqs. (7.136), (7.139) and (7.140) and simplification, Eq. (7.138) becomes

$$\overline{\rho J_\theta[g(\Delta e_e) - d(\Delta e_e)]} = \frac{\partial}{\partial t_\theta}\overline{\{\rho J_\theta\ [k + c_v(T-T_e) + (\phi-\phi_e)]\}}$$

$$+ \nabla_\theta \cdot \overline{\{\rho J_\theta \underset{\sim}{U}[k + c_p(T-T_e) + (\phi-\phi_e)]\}}$$

$$+ \frac{\partial}{\partial \theta}\overline{\{\rho J_\theta \dot{\theta}[k + c_p(T-T_e) + (\phi-\phi_e)]\}}$$

$$- \frac{\partial}{\partial \theta}\overline{(p \frac{\partial z}{\partial t_\theta})} - \beta_2 \frac{\partial}{\partial \theta}(p_e \frac{\partial z_e}{\partial t_\theta}) - \rho_e J_{\theta_e} RT_e \frac{\partial}{\partial t_\theta}\beta_2 \quad . \qquad (7.141)$$

An energy principle in this case requires conservation for isentropic motion in the absence of convective transport and boundary work. The form of Eq. (7.141) shows that the sum of kinetic and deviation internal and geopotential energies would be conserved under these conditions provided that the last two terms vanish or are altered to be determined by exact differentials involving boundary conditions.

7.12.2.1 The reversible component model of total energy for open systems

With the requirement just noted for the last two terms, one condition for energy conservation under isentropic motion is for β_2 to be independent of potential temperature and time. In this case, the last term vanishes and the second to last term is an exact differential. From Eq. (7.136), the relation between the tendencies of the mass distributions in the actual and equilibrium states is given by

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta}) = \beta_2 \frac{\partial}{\partial t_\theta}(\rho_\alpha J_{\theta_\alpha}) \quad . \qquad (7.142)$$

Although the selection of a value for β_2 seems arbitrary, consistency between the time dependency of mass and energy balance in the actual and equilibrium states for an open system requires that β_2 be assigned the value of unity. In the event that the structures of the actual and equilibrium atmospheres are equal within an open system, with β_2 equal to unity the next to last term in Eq. (7.141) would, as it should, just cancel the preceding term. Furthermore, with β_2 equal to unity, Eqs. (7.136) and (7.137) require the hydrostatic equilibrium state pressure p_α to be equal to $\overline{p}$, β_1 to be unity and $\partial\beta_1/\partial p_e$ to be zero.

The dependency of the hydrostatic equilibrium state pressure p_α on the mass distribution in the open system is given by

$$p_\alpha(\theta) = \overline{p_\alpha}(\theta_u) + g \int_\theta^{\theta_u} \overline{\rho J_\theta}\, d\theta \, . \tag{7.143}$$

The overbar defines an area average over the domain of the limited region. The domain of the open region extends vertically from an arbitrary surface θ to an upper boundary surface θ_u and horizontally over the area used in the definition of the area-average. Through temporal differentiation and use of the isentropic mass continuity equation, the time dependency of the equilibrium state pressure is

$$\frac{\partial}{\partial t_\theta} p_\alpha(\theta) = \frac{\partial}{\partial t_\theta} \overline{p}(\theta_u) - g \int_\theta^{\theta_u} \overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})} + \frac{\partial}{\partial\theta} \overline{(\rho J_\theta \dot{\theta})}\, d\theta \, . \tag{7.144}$$

By Gauss's Theorem the integral of Eq. (7.144) is determined by lateral and vertical boundary mass transport.

The relation between the tendency of the equilibrium state temperature and pressure from Eq. (7.124) is expressed by

$$\frac{\partial T_\alpha}{\partial t_\theta} = \frac{\kappa T_\alpha}{p_\alpha}\frac{\partial p_\alpha}{\partial t_\theta} \quad . \tag{7.145}$$

With the time dependency of T_α and p_α specified through the isentropic mass continuity relation, the tendencies of all of the thermodynamic variables of the equilibrium atmosphere are determined.

Now with the definitions of reversible components of total energy and *total flow energy* as intensive variables

$$\Delta e_\alpha = [k + c_v(T-T_\alpha) + (\phi - \phi_\alpha)] \ , \tag{7.146}$$

$$\Delta \nu_\alpha = e_\alpha + R(T-T_\alpha) \ , \tag{7.147}$$

in combination with Eq. (7.141) and β_2 equal to unity, the continuity equation for the reversible component of total energy is expressed by

$$\overline{\rho J_\theta[g(\Delta e_\alpha) - d(\Delta e_\alpha)]} = \frac{\partial}{\partial t_\theta}\overline{(\rho J_\theta \Delta e_\alpha)} + [\nabla_\theta\cdot\overline{(\rho J_\theta \underset{\sim}{U}\Delta \nu_\alpha)}$$

$$+ \frac{\partial}{\partial \theta}\overline{(\rho J_\theta \dot{\theta}\Delta \nu_\alpha)} + g^{-1}\frac{\partial}{\partial \theta}(\overline{p\frac{\partial \phi}{\partial t_\theta}} - \overline{p_\alpha\frac{\partial \phi_\alpha}{\partial t_\theta}})] \quad . \tag{7.148}$$

A vertical integration from θ equal zero to the upper isentropic surface, θ_u, yields

$$\dot{\Delta E_\alpha} = G(\Delta E_\alpha) + D(\Delta E_\alpha) + B(\Delta E_\alpha) \ , \tag{7.149}$$

where

$$\Delta \dot{E}_\alpha = \iint_{A0}^{\theta_u} \frac{\partial}{\partial t_\theta} (\rho J_\theta \Delta e_\alpha)\, d\theta dA \quad , \tag{7.150a}$$

$$= \frac{\partial}{\partial t_\theta} \int_A \int_0^{\theta_u} \rho J_\theta \Delta e_\alpha \, d\theta dA \;, \tag{7.150b}$$

$$G(\Delta E_\alpha) = \iint_{A0}^{\theta_u} \rho J_\theta s T \dot{\varepsilon}_\alpha \, d\theta dA \;, \tag{7.151}$$

$$D(\Delta E_\alpha) = \iint_{A0}^{\theta_u} \rho J_\theta \underset{\sim}{U} \cdot \underset{\sim}{F} \, d\theta dA \quad , \tag{7.152}$$

$$B(\Delta E_\alpha) = - \iint_{A0}^{\theta_u} [\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} \Delta \nu_\alpha) + \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta} \Delta \nu_\alpha) + g^{-1} \frac{\partial}{\partial \theta} (p \frac{\partial \phi}{\partial t_\theta} - p_\alpha \frac{\partial \phi_\alpha}{\partial t_\theta})] d\theta dA \;. \tag{7.153a}$$

$$= - \oint \underset{\sim}{n} \cdot (\int_0^{\theta_u} \rho J_\theta \underset{\sim}{U} \Delta \nu_\alpha d\theta) d\ell - (\int_A \rho J_\theta \dot{\theta} \Delta \nu_\alpha + g^{-1} (p \frac{\partial \phi}{\partial t_\theta} - p_\alpha \frac{\partial \phi_\alpha}{\partial t_\theta}) dA)_{\theta_u} \;, \tag{7.153b}$$

where for simplicity the Lorenz convention is employed. In this convention discussed earlier in Sections 2 and 7.8 the region of integration includes an underground portion. In this balance equation, $G(\Delta E_\alpha)$ and $D(\Delta E_\alpha)$ constitute the Lagrangian source/sink functions by differential heating and viscous dissipation, while $B(\Delta E_\alpha)$ includes all of the processes that integrate to boundary exchange of total energy--the divergence of the horizontal transport, the derivative of the vertical transport, and the component of boundary pressure work associated with the vertical displacement of the isentropic surfaces within the actual and equilibrium states. Eq. (8.153b) reflects the assumption utilized throughout this Section that boundary work at the earth atmosphere interface vanishes. If boundary pressure work were to be retained at the earth's interface in association with, for example, a transfer of energy for the atmosphere to the ocean, a surface boundary term involving the difference of the products of pressure and the height tendency from the integration of Eq. (7.153a) would need to be retained in Eq. (7.153b).

The domain of the open system ideally does not need to extend vertically to the earth's surface provided that ϕ_α is defined at some lower isentropic surface. However, since ϕ_α is determined by a vertical integration of the equilibrium state's hydrostatic mass distribution which requires use of a lower boundary value at the earth's surface, the practical solution is to divide the atmosphere vertically into several layers of interest extending from the earth's surface upwards.

There are several important results from this derivation that merit attention. Since Eq. (7.148) has been derived without vertically including the entire atmosphere, changes of the reversible component of total energy due to convective transport and boundary work correspond one-to-one with changes by the same processes in the total energy equation. The fundamental difference from other theoretical results for available potential energy is that an equilibrium state based on the entropy principle and use of intensive thermodynamic properties is first defined for open systems. This step isolates the role of boundary processes and Lagrangian entropy sources and sinks in expanding or contracting the admissible region of the reversible component of total energy within the phase space defined earlier. Both vertical and horizontal exchange of energy within the domain of the open system are also defined.

The vertical exchange of the reversible component includes degrees of freedom for both diabatic and adiabatic processes. The diabatic flux of *total flow energy* includes the vertical transport of internal and gravitational energies and an $R(T-T_\alpha)$ component of work. In the absence of diabatic processes, the vertical exchange through boundary work is restricted to the difference in changes by $\overline{p\partial\phi/\partial t_\theta}$ and $p_\alpha \partial\phi_\alpha/\partial t_\theta$. Thus, the balance of the reversible component of total energy is determined by physical processes within the domain of the open system, boundary conditions, and boundary transport. As such, changes in the environment only directly impact the reversible component of total energy through boundary processes. If the system is closed with respect to these processes and isentropic motion prevails, the reversible component of total energy is conserved. However, in contrast to total energy which is conserved simply by closure with respect to boundary processes, nonisentropic processes internal to the open domain preclude conservation of the reversible component of total energy. Upon integration over the entire atmosphere, these balance equations for the open system reduce to the global balance equations.

Just as for the global system, differences in the balance of total energy and of its reversible component that stem from the introduction of the entropy principle have both theoretical and pragamatic consequences. Through differencing of the energies for the actual and equilibrium states, an envelope of reversible processes is defined which bounds trajectories within an energy phase space for open systems. This envelope separates admissible regions that can occur within the open system through isentropic motion and no net boundary exchange from inadmissible regions. The envelope of the admissible region can be expanded by boundary transport and work as well as by generation. The envelope can be contracted by all of these processes in addition to the viscous dissipation of kinetic energy.

If the system is energetically closed, the entropy increasing processes of heat diffusion and viscous dissipation will completely collapse the admissible region as a state of uniform temperature and maximum entropy is realized. As such, these relations embody Dutton's (1973) concept of a predictive dimension regarding entropic energy and natural processes.

The evolution of the envelope also provides insight on the degree to which an atmospheric system may engage in reversible isentropic processes internal to the system, interact energetically with its environment and generate energy within the system. For example, in a system with a nominal reversible component of total energy and in the absence of boundary processes and generation by differential heating, oscillations associated with kinetic/available potential energy exchange are constrained to be minimal in contrast with systems with large amounts of reversible total energy. The differences of the wintertime and summertime circulations within each hemisphere illustrate this behavior. Through the larger amounts of reversible total energy within the circulation of the winter hemisphere relative to its summer counterpart, the generation by differential heating within wintertime is large relative to summertime. The increased generation of the winter hemisphere stems from the combination of larger negative and positive efficiencies acting in concert with the increased contrast between entropy sources and sinks.

The degree to which atmospheric predictability in a limited domain is impacted by boundary transport versus Lagrangian sources of the reversible component also follows naturally from these concepts. The future evolution of a state with minimal ΔE_α within a limited atmospheric domain will largely be constrained by boundary processes and less determined by its initial state. The future evolution of a state with relatively large ΔE_α in the same domain is more determined by its initial state and less constrained by boundary processes. As the domain of prediction is increased, the importance of the internal structure relative to boundary transport must also increase. This increase stems in part from the greater rate of increase of volume than surface area as a domain expands in size. For example, the volume of the spherical domain increases with the cube of the radius while the surface area increases with the square.

There are also pragmatic considerations in applying the reversible component of total energy studies of open systems. If not already ascertained, diagnostic studies in the future will establish that reliable comprehensive estimates of the energy cycle of open systems can only be determined from employment of reversible total energy concepts. Estimation of the global energy cycle only became feasible with Lorenz's (1955a) development of available potential energy theory. Increased accuracy results from differencing the energies or processes of actual and equilibrium atmospheres as defined by intensive properties within isentropic coordinates and then multiplication of the intensive properties or processes for both states by the isentropic mass distribution of the actual state of the atmosphere. This latter step largely eliminates errors that stem from an inability to diagnose atmospheric mass and its balance with sufficient accuracy to determine mass tendency. The differencing of the thermodynamic properties and processes between actual and associated equilibrium states systematically cancels errors that are inherently present in empirical data and for that matter even exist in numerical simulations of atmosphere phenomena due to lack of conservation of fundamental properties. The current disparity between isobaric and isentropic estimates of the balance of available potential energy (Dutton and Johnson, 1967; Min and Horn, 1982) are associated with the quadratic and uniform static stability approximations introduced in isobaric studies. These discrepancies are not to be confused with sampling and analysis errors.

7.12.2.2 The entropy component model of total energy for open systems

Sufficient conditions for the Lagrangian source of kinetic energy by reversible processes to remain invariant in Eqs. (7.116) and (7.117) are satisfied by the thermodynamic relations Eqs. (7.121) through (7.124). Within an equilibrium state satisfying these conditions, the entropy component of total energy with uniform temperature $T(t_0)$ is defined as an intensive thermodynamic variable by

$$\Delta e_0 = k + c_v(T-T_0) + (\phi-\phi_0) \quad , \tag{7.154}$$

while the corresponding *total flow energy* becomes

$$\Delta\nu_0 = \Delta e_0 + R(T-T_0) \quad . \tag{7.155}$$

Using these definitions in Eq. (7.141), the continuity equation for the entropy component model of total energy is given by

$$\rho J_\theta[g(\Delta e_0) - d(\Delta e_0)] = \frac{\partial}{\partial t_\theta}(\rho J_\theta \Delta e_0) + \nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}\Delta\nu_0)$$

$$+ \frac{\partial}{\partial\theta}(\rho J_\theta \dot{\theta}\Delta\nu_0) + \frac{\partial}{\partial\theta}\, p\frac{\partial z}{\partial t_\theta} - \rho_0 J_{\theta_0} RT_0 \frac{\partial \beta_2}{\partial t_\theta} \quad , \tag{7.156}$$

where it has been noted that the boundary work component involving $\partial z_0/\partial t_\theta$ is zero by definition. The left term constitutes the Lagrangian source/sink function by the combination of differential heating and viscous dissipation. The right terms are the tendency, the divergence of the horizontal transport, the derivative of the vertical transport and a component of boundary pressure work. In the entropy model, there is no means for the mass of the equilibrium state to be related directly to the area-averaged mass distribution $\overline{\rho J_\theta}$, since the equilibrium temperature T_o is invariant in time and space. The thermodynamic state is determined by the surface pressure $p(\phi_{oo})$, and the temperature T_o [see Eqs. (7.61) through (7.62)]. Thus with the need for the relation between $\overline{\rho J_\theta}$ and $\rho_o J_{\theta_o}$ to be arbitrary, the tendency of β_2 cannot vanish.

Through a differentiation in combination with Eq. (7.136), the last term of Eq. (7.156) is given by

$$\rho_o J_{\theta_o} RT_o \frac{\partial \beta_2}{\partial t_o} = \frac{\partial}{\partial t_\theta} (\overline{\rho J_\theta} RT_o) - \beta_2 \frac{\partial}{\partial t_\theta} (\rho_o J_{\theta_o} RT_o) \; . \tag{7.157}$$

From the condition that the equilibrium state temperature T_o and its hydrostatic surface pressure $p(\phi_{oo})$ are independent of time, the last term of Eq. (7.157) vanishes. With this simplification, use of the hydrostatic relation and an exchange of the order of differentiation, Eq. (7.157) is expressed by

$$\rho_o J_{\theta_o} RT_o \frac{\partial \beta_2}{\partial t_o} = \frac{\partial}{\partial \theta} [RT_o \frac{\partial}{\partial t_o} (-\overline{p}/g)] \; . \tag{7.158}$$

A substitution into Eq. (7.156) yields

$$\rho J_\theta[g(\Delta e_o) - d(\Delta e_o)] = \frac{\partial}{\partial t_\theta} (\rho J_\theta \Delta e_o) + [\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} \Delta \nu_o) + \frac{\partial}{\partial \theta} (\rho J_\theta \dot{\theta} \Delta \nu_o)]$$

$$+ \frac{\partial}{\partial \theta} [p \frac{\partial z}{\partial t_\theta} - RT_o \frac{\partial}{\partial t_\theta} (-\overline{p}/g)] \; . \tag{7.159}$$

With a time independent equilibrium temperature T_0, there is no degree of freedom for the RT_0 component of work to be associated with the quasi-Lagrangian nature of isentropic surfaces. Hence, the RT_0 component of the flow energy is replaced by the vertical derivative of the product of RT_0 and the tendency of the area-averaged mass distribution, which integrates to a boundary condition. See the last term of Eq. (7.159). Thus, all of the physical processes associated with changes of the equilibrium atmosphere's energy on the right hand side are determined by the changes in the isentropic mass distribution and are related to boundary processes.

The form of the transport equation, Eq. (7.159), demonstrates that with isentropic motion and energetically closed boundary conditions, the entropy component of total energy is conserved. Except for the differing structures for the equilibrium states and the vertical component of RT boundary work term, the transport equations in Eqs. (7.148) and (7.159) and integrals of the transport equations for the reversible and entropy components of total energy are formally identical. [See Eqs. (7.149) through (7.153) for integrals of transport equation for reversible component. Integrals of the entropy component are not stated due to identical nature of the two sets.] If the domain of the open system includes the entire atmosphere, the integration of Eq. (7.159) yields identical results with global results expressed by Eqs. (7.68a and b). In the expansion or contraction of the region for the entropy component, just as for the reversible component of total energy, it is the difference of the time rates of change of actual and equlibrium state energies that determine the time dependency of an envelope in the energy phase space. The expansion or contraction of the envelope of the reversible and entropy components of total energy for open systems will be equal with each other provided the area-averaged mass distribution is steady. Such conditions will prevail when the area-averaged mass circulation is steady. This requires that the area-integrated vertical derivative of the diabatic mass flux exactly balances the area-integrated divergence of the isentropic mass transport, which by Gauss' Theorem equals the line integral of the normal component of the mass transport. With closed boundary conditions, just as for the reversible component, entropy increasing processes will degrade this component of energy to a state of maximum entropy with uniform temperature that is determined by the total energy of the system (Dutton, 1973).

7.12.3 The equilibrium state and its functional dependency.

In discussing thermodynamics, Sommerfeld (1950) notes that the science of thermodynamics is an axiomatic science that introduces the new concept of temperature. Through postulating states of thermal equilibria, a communality between the results of Lettau (1954), Lorenz (1955), Dutton (1973) and Pearce (1978) has been established.

Classically, in thermodynamic systems the determination of the efficiency of heating in producing mechanical energy under steady equilibrium conditions involves the temperature T_1 at which heat is added and a general temperature of an environment T_2 at which heat is removed (Tolman and Fine, 1948). In atmospheric applications intensive variables locally are assumed to be close enough to thermal equilibrium so that the temperature is well-defined and unambiguous and the definition of $(-\nabla \cdot H)/\rho$ equal to $\dot{s}T$ does not create difficulty in the application of classical thermodynamics. Thus, the determination of energy sources/sinks by the product of temperature and entropy sources/sinks in G(E) is well determined. The logical difficulty, both in theory and in applications for the atmosphere however, is how to define the "global" temperature T_0 of the "surroundings" that is needed for understanding the thermodynamic efficiency of the atmosphere and how to determine its time rate of change. As noted earlier, it is primarily on this point that the available potential energy model of Lorenz (1955a, 1955b, and 1960) and the entropy models of Lettau (1954a), Dutton (1973) and Pearce (1978) differ.

From the time rate of change of entropy, a single value of temperature T_0 when multiplied by the time rate of change of the atmosphere's entropy as a whole ($\dot{S}$) determines the time rate of change of the energy of an equilibrium state for the atmosphere as a whole. With the mass and temperature being invariant, changes in the energy E_0 are restricted to changes in geopotential energy. The entropy models of Lettau (1954a), Dutton (1973) and Pearce (1978) are based (directly or indirectly) on this approach, where in accord with classical thermodynamics uniformity of temperature defines a state of unconstrained thermal equilibrium (Sommerfeld, 1950). The equilibrium temperatures that were (explicitly or implicitly) defined by each were somewhat different.

In his discussion of thermodynamic equilibrium, Sommerfeld (1950) contrasts the state of a system in unconstrained thermodynamic equilibrium with states which are not in equilibria. From nonequilibrium states, he defines a state of constrained equilibrium such that "the entropy of the system may be taken to be the sum of the entropy of all volume elements" within which the state of a volume element is one of thermodynamic equilibrium. As such the temperature distributions T_α and $T_{\alpha 0}$ for the component models of energy ΔE_α and $\Delta E_{\alpha 0}$ are characterized by states of constrained equilibrium. With mass variation within isentropic layers due to entropy sources and sinks, both the internal and geopotential energies of E_α and $E_{\alpha 0}$ will vary.

From the local viewpoint of thermodynamics in an atmosphere where the temperature of the "surroundings" is not uniform, it is reasonable heuristically to regard the mass in each layer of equal entropy as a "system" and define an equilibrium temperature T_α or $T_{\alpha o}$. In deriving the reference state for available potential energy, Lorenz assumed a horizontally invariant hydrostatic state attained by an isentropic redistribution of mass. As such, his reference state corresponds to a constrained equilibrium state since the temperature T_α is uniform within each isentropic layer. The temperature structure $T_\alpha(\theta,t)$ must remain a function of entropy by virtue of the isentropic assumption in attaining the reference state.

Entropy change enjoys an exact differential relation with heat addition divided by temperature. The global mass balance of an isentropic layer is precisely determined by the diabatic vertical mass flux associated with entropy sources and sinks. In this layer differential heating generates a reversible component of total energy, provided the "local" temperature differs systematically from the layer's equilibrium temperature. For the system as a whole then, the generation of the reversible component of total energy is simply the sum of the reversible component energy generated in each of the isentropic layers, while the net vertical exchange of mass, total energy and entropy between layers by diabatic mass flux and boundary work vanishes with vertical integration over all layers that define the system. Thus, from recognition of the systematic isentropic stratification of the atmosphere and use of isentropic coordinates, questions regarding the inhomogeneity of thermodynamic variables and the atmosphere's efficiency in maintaining circulation as energy is gained and lost seems to be resolved. Lorenz's (1955a) theory of available potential energy, although developed from different concepts, is consistent with this heuristic explanation.

A retention of functional dependence for the equilibrium temperature T_α on entropy and time can also be justified pragmatically. If the differential heating occurs in an isentropic layer in the troposphere as opposed to a layer in the stratosphere, each of these regions may enjoy circulations which thermodynamically are independent of each other. Thus, an equilibrium temperature for each of these isentropic layers needs to be defined in order to determine the efficiency of differential heating in the generation of a reversible component of total energy.

It is important to realize however, that the equilibrium state temperature $T_\alpha(\overline{p}(\theta,t))$ just defined is only one option. As long as the system is in a statistically stationary state in which the mass-weighted area-averaged vertical flux of entropy $\hat{s}^A$ vanishes, Eq. (7.17a) shows that the generation of the reversible component of total energy is independent of the functional form utilized for $T_\alpha(\theta,t)$. Hence, estimates of the generation $G(\Delta E_0)$ by differential heating from the entropy component model corresponding with Lettau's (1954a), Dutton's (1973) and Pearce's (1978) estimates and the generation $G(\Delta E_\alpha)$ by the available potential energy model of Lorenz (1955a) are identical.

With the lack of global statistical equilibrium during the annual cycle, however, Lorenz's (1955a) definition of the reference state in available potential energy theory (the equilibrium state T_α) ensures that the reversible component is isolated and accurately bounded when through entropy sources and sinks the total energies of the actual and equilibrium states increase or decrease. Furthermore, of all of the candidates for the equilibrium temperature distribution, namely $T_\alpha(\theta,t)$, $T(\theta,t_0)$ and $T(t_0)$, $T_\alpha(\theta,t)$ is the one that determines the sharpest delineation of the region of reversible thermodynamic trajectories, in the sense that with expansion or contraction of the envelope no component of total energy is included which is unable to engage in reversible exchange between kinetic and total potential energies. This conclusion stems from our understanding in that the hydrostatic reference atmosphere is the state of minimum total potential energy that can be attained by an arbitrary isentropic mass redistribution, and that on attainment of this state the kinetic energy is a maximum (Lorenz, 1955a; Dutton and Johnson, 1967). This attribute stems from Lorenz's (1955a) choice for an equilibrium temperature T_α that is dependent on the isentropic area-averaged hydrostatic pressure distribution.

The condition that the reversible component of energy $E - E_\alpha$ defines the maximum kinetic energy attainable by isentropic processes is the counterpart of the theorem of maximum work (Sommerfeld, 1950 p. 52-55). In this theorem the change in free energy, ΔF, which is described as the maximum work $p\nabla\cdot \underset{\sim}{U}$ that a system can perform when it passes reversibly from an initial state of given temperature T_0 to another state of equal temperature, corresponds with available potential energy. The internal energy minus the free energy, which is called the "unavailable energy", is equal to the product of temperature and entropy, T_0S. Although the thermodynamics processes involved in determining the maximum work are more limiting that those permitted in the component energy models, it is interesting to note that the "unavailable energy" corresponds with the energy of the equilibrium state for the entropy component model E_0 as it is determined from Eq. (7.104) by T_0S, and also with the energy of the reversible component model provided that one recognizes that E_α is the energy of a state in constrained equilibrium.

7.13 Kinetic energy generation in global monsoonal circulations

With the isolation of the reversible component of total energy ΔE_α, the means by which kinetic energy is realized in global monsoonal circulations is now examined in more detail. One aim is to point out how, through thermal forcing at the planetary scale, the isentropic mass circulation directly maintains the atmosphere's supply of kinetic energy against viscous dissipation. Another aim is to illustrate the reversibility of energy transformations by identifying geographical regions where the generation of kinetic energy is systematically positive and other regions where it is systematically negative. The existence of regions with systematic transformation of kinetic to available potential energy lends credence to the importance of reversible isentropic processes. Although the empirical evidence to be presented is only preliminary, the results illustrate: 1) that direct generation of kinetic energy by the irrotational component occurs in thermally forced mass circulations and 2) that the redistribution of kinetic energy in the circumpolar vortex by the rotational component involves reversible isentropic energy exchange at the planetary scale.

7.13.1 Irrotational and rotational components of kinetic energy generation

With the hydrostatic assumption and use of generalized coordinates, the generation of horizontal kinetic energy by reversible processes (after Eq. (7.127) is given by

$$\rho J_\eta s(k_h) = - \rho J_\eta \underset{\sim}{U} \cdot [\nabla_\eta(\psi - \psi_\alpha) - (T\varepsilon)\nabla_\eta s] \ , \qquad (7.160)$$

where the subscript r is suppressed for simplicity. The source/sink function of kinetic energy is by mechanical work (Johnson and Downey, 1982) through displacement of a fluid element in the presence of a body force. However, in Eq. (7.160) the form of the kinetic energy source also shows that this mechanical work involves two processes, advection of dry static energy and entropy.

Through use of Helmholtz's theorem, the generation of kinetic energy divided into irrotational and rotational components is expressed by

$$\rho J_\eta s(k_h) = \rho J_\eta [s_\chi(k) + s_\psi(k)] \ , \tag{7.161}$$

where

$$\rho J_\eta s_\chi(k_h) = - \nabla_\eta \chi \cdot [\nabla_\eta(\psi - \psi_\alpha) - (T\varepsilon)\nabla_\eta s] \ , \tag{7.162}$$

$$\rho J_\eta s_\psi(k_h) = - (\underset{\sim}{k} \times \nabla_\eta \psi) \cdot [\nabla_\eta(\psi - \psi_\alpha) - (T\varepsilon)\nabla_\eta s] \ . \tag{7.163}$$

With a differentiation, Eqs. (7.162) and (7.163) are expressed by

$$\rho J_\eta s_\chi(k_h) = - \{\nabla_\eta \cdot [(\psi - \psi_\alpha)\nabla_\eta \chi] - (\psi - \psi_\alpha)\nabla_\eta^2 \chi$$

$$- T\varepsilon[\nabla_\eta \cdot (s\nabla_\eta \chi) - s\nabla_\eta^2 \chi)\} \ , \tag{7.164}$$

$$\rho J_\eta s_\psi(k_h) = - \{\nabla_\eta \cdot [(\psi - \psi_\alpha)\underset{\sim}{k} \times \nabla_\eta \psi] - (T\varepsilon)\nabla_\eta \cdot (s\underset{\sim}{k} \times \nabla_\eta \psi)\} \ . \tag{7.165}$$

The first terms of Eq. (7.164) and (7.165) are determined by the divergence of the transport of dry static energy by the irrotational and rotational components of motion. These terms include boundary transport of internal and gravitational energies in addition to the RT component of boundary pressure work. The second term of Eq. (7.164) associated with the horizontal divergence of the mass transport involves changes internal to a system. The other terms of Eqs. (7.164) and (7.165) involve entropy exchange that is also internal to a system.

From fundamental principles, the generation of kinetic energy by mechanical work is equal to the sum of boundary pressure and thermodynamic work (Johnson and Downey, 1982). Classically, boundary pressure work is determined by the surface integral of the scalar product of the pressure stress and velocity, while thermodynamic work involves the product of the pressure and three dimensional velocity divergence. With expression of the pressure gradient force as a function of the gradients of dry static energy and entropy, an explicit identification of these two fundamental processes is not apparent in generalized coordinates.

Note in the differentiation that the terms involving the product of the deviation temperature and the divergence of the entropy transport [the third term of Eq. (7.166) and the second term of Eq. (7.167)] do not integrate to boundary integrals; thus one must infer that boundary pressure work is in general linked with boundary transport of energy and with mass and entropy exchange within a region. This introduces an inherent complexity in considering the generation of kinetic energy, since the mass, dry static energy and entropy balance are interrelated through boundary transport processes and internal source/sink functions.

With the choice of isentropic coordinates, a fundamental simplification occurs. Since entropy is a unique function of potential temperature, the sources of kinetic energy through the rotational and irrotational components of mass transport reduce to

$$\rho J_\theta s_\chi(k_h) = -\nabla_\theta \chi \cdot \nabla_\theta(\psi - \psi_\alpha) , \tag{7.166a}$$

$$= -\nabla_\theta \cdot [(\psi-\psi_\alpha)\nabla_\theta \chi] + (\psi - \psi_\alpha)\nabla_\theta^2 \chi , \tag{7.166b}$$

$$\rho J_\theta s_\Psi(k_h) = -(\underset{\sim}{k} \times \nabla_\theta \Psi) \cdot \nabla_\theta(\psi - \psi_\alpha) , \tag{7.167a}$$

$$= -\nabla_\theta \cdot [(\psi - \psi_\alpha)\underset{\sim}{k} \times \nabla_\theta \Psi] . \tag{7.167b}$$

The generation of kinetic energy is now uniquely related to the advection of dry static energy in Eqs. (7.166a) and (7.167a), while in Eqs. (7.166b) and (7.167b) generation is through the divergence of the isentropic dry static energy transport and also in Eq. (7.166b) through the product of the divergence of the isentropic mass transport and deviation dry static energy. As such, the exchange of the reversible component of the dry static energy becomes uniquely related with kinetic energy generation. The divergence of the dry static energy transport in Eqs. (7.166b) and (7.167b) constitutes a source of kinetic energy through boundary transport of internal and gravitational energies and through boundary pressure work. In the last term of Eq. (7.166b), thermodynamic work is a source of horizontal kinetic energy through expansion and contraction of fluid elements along horizontal components of motion.

In an isentropic global perspective, the thermodynamic work of Eq. (7.166b) is the ultimate source for maintenance of kinetic energy against viscous dissipation. With areal integration, the kinetic energy generation by the rotational component as well as the divergence of the energy transport of the irrotational component vanishes. Furthermore, the net source by the rotational component in an isentropic layer also vanishes if one integrates over the areal extent of a channel bounded by ψ and $\psi + \Delta\Psi$. The rotational component can in effect only redistribute kinetic energy within the channel; it does not transport energy normal to the channel. Positive kinetic energy generation by the rotational component in one segment of the channel must be balanced by negative generation in another segment of the channel.

Insight into the link between the thermally forced mass circulations and the generation of kinetic energy is gained with an areal integration over the entire isentropic surface and use of the mass-weighted area-average and its deviation. In this case, the kinetic energy source in Eqs. (7.166b) and (7.167b) reduces to

$$\overline{\rho J_\theta s \chi (k_h)}^{A} = \overline{\rho J}^{A} \widehat{(\psi-\psi_\alpha)^{**}(\nabla^2_\theta \chi/\rho J_\theta)^{**}}^{A} \quad . \tag{7.168}$$

The mean component of Eq. (7.166b) vanishes from application of the divergence theorem to the mass-weighted area-average of $(\nabla^2_\theta \chi/\rho J_\theta)$. Global generation of kinetic energy occurs through a positive covariance of the deviation dry static energy and the divergence of the isentropic mass transport. In the overworld, the combination of mass divergence with positive deviation dry static energy in entropy source regions and mass convergence with negative deviation energy in entropy sink regions constitutes the primary structure for the kinetic energy generation of global monsoonal circulations. Although likely less important at the global scale, kinetic energy generation in the underworld occurs through the combination of mass divergence with positive deviation of dry static energy in anticyclonic circulations and mass convergence with negative deviation of dry static energy in cyclonic circulations.

7.13.2 Stationary and transient components of kinetic energy generation

With the definition of a mass-weighted time-average and its deviation, the rotational and irrotational components of kinetic energy generation partitioned into stationary and transient components are expressed by

$$\overline{\rho J_\theta s_\chi}^{t}(k_h) = \overline{\rho J_\theta}^{t} [\widehat{(\nabla_\theta \chi/\rho J_\theta)}^{t} \cdot \widehat{\nabla_\theta \psi}^{t} + \widehat{(\nabla_\theta \chi/\rho J_\theta)^{*} \cdot (\nabla_\theta \psi)^{*}}^{t}] \quad , \tag{7.169}$$

$$\overline{\rho J_\theta s_\Psi}^{t}(k_h) = \overline{\rho J_\theta}^{t} \{ [\widehat{(\underset{\sim}{k} \times \nabla_\theta \Psi)/\rho J_\theta}^{t}] \cdot \widehat{\nabla_\theta \psi}^{t}$$

$$+ \widehat{[(\underset{\sim}{k} \times \nabla_\theta \Psi)/\rho J_\theta]^{*} \cdot [\nabla_\theta \psi]^{*}}^{t} \} \quad . \tag{7.170}$$

The first term of Eq. (7.169) which is determined by the scalar product of the mean mass transport and the mean pressure force, is directly linked with the thermally forced mean mass circulation. The second term of Eq. (7.169) as well as both terms in Eq. (7.170) are not directly linked with the thermally forced mean mass circulation.

In this partitioning, the question to be explored is whether or not the generation of kinetic energy is linked directly or indirectly with the diabatic mass transport in the isentropic mass circulation. One may suggest that if kinetic energy generation in middle latitudes were realized primarily through baroclinic instability processes that occurred independently of the differential heating in global monsoonal circulations, the time-space scale of these processes would differ from the time-space scale of the entropy sources and sinks of isentropic mass circulations. However, if the pattern of kinetic energy generation is primarily determined by the time-averaged irrotational component of the isentropic mass circulation, the time-space scale of these processes will be consistent with each other. Such evidence would support the premise that even the realization of kinetic energy through the process of baroclinic instability is an integral component of the thermally forced mass circulation. In this case the first term of Eq. (7.169) would be larger than the second term; however, it need not be larger than either generation term in Eq. (7.170), since globally the area integral of the rotational component vanishes. As emphasized earlier, the generation by the rotational component merely redistributes the kinetic energy in the global domain.

7.13.3 A preliminary result

The pattern of the 340-350 K time-averaged Montgomery Stream Function Fig. 17 for January 1979 in Fig. 17A shows low values of dry static energy in polar latitudes and high values of dry static energy in tropical latitudes. In middle latitudes of the Northern Hemisphere, the pattern shows the major planetary waves that exist, with troughing over the eastern portions of Asia and North America and ridging over the eastern Pacific and Atlantic Oceans downstream of the Asian and North American geostrophic wind maximums. The pattern of the Southern Hemisphere indicates greater symmetry of the circumpolar circulation compared to its Northern Hemisphere counterpart.

Without partitioning into irrotational and rotational components the pattern of kinetic energy generation in Fig. 17b reveals a somewhat disorganized pattern. In the middle and subtropical latitudes of the Northern Hemisphere a systematic variation in sign occurs in the westerlies. Positive regions tend to occur in confluent regions of the Asiatic and North American geostrophic wind maxima evident in Fig. 17a, while negative regions occur in the diffluent regions. Such a pattern is consistent with the concept of direct and indirect secondary circulations attending wind maxima. While a pattern of negative and positive regions is indicated in the Southern Hemisphere, the patterns in this hemisphere lack the systematic organization evident in the Northern Hemisphere.

Figures 17c and d present the preliminary results for the kinetic energy generation by the standing components of the irrotational and rotational modes of mass transport (Johnson, 1986). Only the standing component of the two modes will be discussed, since the earlier results (Johnson, 1986) established that the standing component dominated the transient component, at least in this layer. The distribution of generation by the standing component of the irrotational mode (Fig. 17c) is explained by comparing the 340-350 K isentropic mass transport potential function of Fig. 6a and the Montgomery Stream Function of Fig. 17a. In regions where common components of the irrotational mass transport and the mean pressure gradient force ($-\nabla_\theta\psi$--not shown, but indicated by the time-averaged Montgomery Stream Function $\overline{\psi}$) occur, the kinetic energy generation is positive (Fig. 17c). These regions generally extend along the equator in the subtropics and meridionally from pole to pole in the eastern Asia-Pacific-Australia region. Maxima of generation are located just north of the ITCZ over the Indian Ocean and over the western central Pacific along the axis of the east Asian subtropical jet maximum. A similar maximum occurs just south of Australia, where the strong meridional component of the irrotational mass transport and the strong meridional pressure gradient of the Southern Hemisphere westerlies have common components. The negative generation at this level in the North American wind maximum is associated with the isentropic mass convergence over the Caribbean and western Africa. A positive region of kinetic energy generation occurs over northern South America and western Africa in association with the divergent mass transport from tropical convection occurring in monsoon circulations over South America and Africa. These patterns of mass transport and kinetic energy generation in the 340-350 K layer are indicative of kinetic energy generation in the subtropical jet stream by the thermally forced Hadley circulation in conjunction with the poleward transport of absolute angular momentum. Hence, the patterns are consistent with Krishnamurti's (1961) results for the subtropical jet stream.

A comparison of the kinetic energy generation through the standing component of the rotational mass transport (Fig. 17d) with the total generation (Fig. 17b) shows that the zonal alteration of sources and sinks in the total field stems from the redistribution of kinetic energy by the rotational component. This alteration is imposed by the constraint of zero net generation by this component discussed earlier. Locally, magnitudes of the kinetic energy generation and destruction by the rotational component are somewhat larger than the magnitudes for the irrotational component. Note particularly the alteration of sources/sinks in the westerlies, where in the confluent regions of the Asian and North American wind maxima kinetic energy generation is positive, but in the diffluent regions it is negative. Patterns of positive and negative regions in the westerlies support the concept that reversible isentropic exchange of kinetic and available potential energy is important. At least in the overworld, the magnitude of this exchange locally is larger than the net transformation of available potential energy to kinetic energy.

These preliminary results have not resolved the question of whether or not thermally forced isentropic mass circulations play a key role in higher latitudes or in lower isentropic layers. However, the pattern displayed in the 340-350 K layer for the stationary irrotational mode also shows substantial kinetic energy generation in extratropical latitudes (see Fig. 17c).

Until a comprehensive analysis is completed, final conclusions concerning the links between the kinetic energy generation and the thermally forced isentropic mass circulation within a perspective of the reversible component of total energy must be held in abeyance. However, these preliminary results point out that:

(1) Simplifications of the exchange of dry static energy and entropy in the analyses of kinetic energy generation are achieved through use of isentropic coordinates,

(2) Kinetic energy generation occurs primarily through the time-averaged structure of the isentropic mass circulation, and

(3) Reversible isentropic processes are important in the downstream advection of kinetic and available potential energy in the westerlies and in the determination of the structure of the time-averaged planetary scale waves.

Overall, the results verify that the net source of kinetic energy, occurs directly through the thermodynamically forced component of the isentropic mass circulation.

7.13.4 On the necessity of reversible isentropic processes

The empirical results of the previous section have isolated the existence of direct and indirect circulations in the westerlies that are associated with the advection of dry static energy by the rotational component of the isentropic mass transport. Such circulations are traditionally considered to be inertially forced secondary circulations associated with propagating baroclinic waves. Since the rotational component of the isentropic mass transport is independent of the forcing by differential heating, nor is it a net global source of kinetic energy, these results are not contradictory with the concepts of thermally forced global monsoonal circulations. On the contrary, the isolation of the kinetic energy generation by the irrotational mass transport and the constrained exchange between the dry static and kinetic energies through reversible processes dictate that direct and indirect circulations are a means by which the rotational component of the mass transport zonally redistributes energy in the westerly wind regime. The role of the rotational component of the mass transport in the zonal redistribution of the reversible component of total energy and the transformation between available potential and kinetic energies will now be considered.

The hydrostatic equation of motion,

$$\frac{\partial \underset{\sim}{U}}{\partial t_\theta} + \underset{\sim}{U}\cdot\nabla_\theta \underset{\sim}{U} + \dot{\theta}\,\frac{\partial \underset{\sim}{U}}{\partial \theta} = -\nabla_\theta \psi - f\underset{\sim}{k}\times\underset{\sim}{U} + \underset{\sim}{F}\,, \qquad (7.171)$$

in combination with the vector identity,

$$\underset{\sim}{U}\cdot\nabla_\theta \underset{\sim}{U} = \zeta_\theta k \times \underset{\sim}{U} + \nabla_\theta k \ , \tag{7.172}$$

yields

$$\frac{\partial \underset{\sim}{U}}{\partial t_\theta} + \zeta_{\theta_a} \underset{\sim}{k} \times \underset{\sim}{U} + \dot{\theta}\frac{\partial \underset{\sim}{U}}{\partial \theta} = -\nabla_\theta(\psi + k) + \underset{\sim}{F} \ . \tag{7.173}$$

The corresponding kinetic energy equation from a scalar multiplication by $\underset{\sim}{U}$ is

$$\frac{\partial k}{\partial t_\theta} + \dot{\theta}\frac{\partial k}{\partial \theta} = -\underset{\sim}{U}\cdot\nabla_\theta(\psi + k) + \underset{\sim}{U}\cdot\underset{\sim}{F} \ . \tag{7.174}$$

With use of the isentropic equation of continuity and definitions for the reversible component of *total flow energy* from Eq. (7.147) and viscous dissipation after Eq. (7.152), the time-averaged kinetic energy transport equation is

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta k}^t) + \overline{k\,\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U})}^t + \frac{\partial}{\partial \theta}(\overline{\rho J_\theta \dot{\theta} k}^t) = -\overline{\rho J_\theta \underset{\sim}{U}\cdot\nabla_\theta \nu_\alpha}^t - \overline{\rho J_\theta d(\Delta e_\alpha)}^t \tag{7.175}$$

The local balance of kinetic energy expressed by Eq. (7.175) involves several physical processes. However, from a global perspective, the important terms involve the viscous dissipation of kinetic energy and the advection of the reversible component of *total flow energy*, $\Delta\nu_\alpha$. This latter process includes the Lagrangian source of kinetic energy by baroclinic processes and the advection of kinetic energy. Presumably, in the time-averaged state of the global circulation, the tendency of kinetic energy, the covariance of the kinetic energy with the mass divergence, and the vertical derivative of the diabatic transport of kinetic energy are of lesser importance.

The empirical results of Section 5 established that both the time-averaged potential and stream functions for the mass transport were nearly identical with corresponding functions for the transport of *total flow energy*. Since the contribution by the transient energy transport is negligible, these results require the isopleths of the *total flow energy* $\hat{\nu}$ and the time-averaged stream function to be coincidental. Earlier, it was suggested: 1) that a Bernouilli-like relation would exist, with the time-averaged rotational component of the isentropic mass transport being parallel to surfaces of entropy and time-averaged *total flow energy* and 2) that such a result is not inconsistent, since the time-averaged irrotational component is the sole degree of freedom by which the generation of kinetic energy is linked to the thermally-forced isentropic mass circulation. The thermally forced mass circulation generates the kinetic energy of the irrotational component. Although not explicitly determined in this study, the net source of the rotational component of kinetic energy comes from the transformation of the irrotational component of kinetic energy to the rotational component by Coriolis and inertial accelerations as the quasi-gradient balance of the rotational motion is maintained against dissipation.

With the definitions from Eqs. (7.146) and (7.147) and the results in Section 5, the divergence of the transport of the reversible component of *total flow energy* by the stationary component is now expressed by

$$\nabla_\theta \cdot (\overline{\rho J_\theta}^{t} \hat{\underset{\sim}{U}}^{t} \Delta\hat{\nu}_\alpha^{t}) = \nabla_\theta \cdot [\Delta\hat{\nu}_\alpha^{t}(\nabla_\theta \overline{X}^{t} + k \times \nabla_\theta \overline{\Psi}^{t})] \qquad (7.176)$$

The integral constraint of no net transport of time-averaged *total flow energy* across the isopleths of the time-averaged stream function requires that the area-averaged advection of the reversible component of *total flow energy* by the time-averaged rotational component vanish within the stream function channel. This result is expressed by

$$\overline{\nabla_\theta \cdot (\overline{\rho J_\theta \hat{U}_\psi}^t \hat{\Delta\nu_\alpha}^t)}^{\delta A} = \overline{\nabla_\theta \cdot (\hat{\Delta\nu_\alpha}^t \underset{\sim}{k} \times \nabla_\theta \overline{\Psi}^t)}^{\delta A} = 0 \quad , \qquad (7.177)$$

$$= \overline{(\underset{\sim}{k} \times \nabla_\theta \overline{\Psi}^t) \cdot \nabla_\theta \hat{\nu}_\alpha^t}^{\delta A} = 0 \quad , \qquad (7.178)$$

where the overbar with the superscript δA indicates area-averaging over the stream function channel bounded by $\overline{\Psi}^t$ and $(\overline{\Psi}^t + \overline{\Delta\Psi}^t)$. With the definition that the reversible component of the *total flow energy* $\Delta\nu_\alpha$ is the sum of k plus $(\psi-\psi_\alpha)$, this constraint is expressed by

$$\overline{(k \times \nabla_\theta \overline{\Psi}^t) \cdot \nabla_\theta \hat{k}^t}^{\delta A} = - \overline{(k \times \nabla_\theta \overline{\Psi}^t) \cdot \nabla_\theta \widehat{(\psi-\psi_\alpha)}^t}^{\delta A} \quad . \qquad (7.179)$$

Within the nondivergent area-averaged transport of *total flow energy* as expressed in Eq. (7.179), the rotational component of kinetic energy generation is negative (positive) in regions where kinetic energy advection is positive (negative). Thus, a local increase (decrease) of kinetic energy by advection simply balances the Lagrangian sink (source) of kinetic energy associated with indirect (direct) circulation and its transformation deviation of kinetic energy to (from) dry static energy. Note from the distribution of the Montgomery stream function in Fig. 17a and the mass stream function in Fig. 3a that positive kinetic energy advection and negative generation of kinetic energy are indicated in the diffluent region located between between the trough and ridge downstream of the wind maximum. Opposite patterns are found in the confluent region located between the trough and upstream ridge.

In models of finite length jet streaks energy transformations between kinetic and available potential energy are attributed to direct circulation in the confluent region and to indirect circulation in the diffluent region. Although such transformations are internally consistent with the concept of an inertially forced ageostrophic component of the isentropic mass transport (Uccellini and Johnson, 1979), it is interesting to find that such energy transformations are required from constraints on the transport of total energy by the rotational component of isentropic mass transport. Since the planetary waves are relatively stationary, these circulations must in part determine the location of thermal ridge and trough axes of planetary waves through the longitudinal redistribution of dry static and kinetic energies. Hence, these processes should be recognized for there role in determining the position of the time-averaged planetary scale waves through transverse secondary circulations of the Asiatic and North American wind maxima, but not for a primary role in maintaining the atmosphere's kinetic energy supply against viscous dissipation. It is the thermally forced isentropic mass circulation which ultimately generates the component of kinetic energy that maintains the rotational component of kinetic energy against viscous dissipation.

7.14 A background summary

The results of this section were addressed to the determination of the role of reversible and irreversible thermodynamic processes in the maintenance of global monsoonal circulations. At the same time, the different approaches in studying the maintenance of the atmosphere's circulation that have been set forth by Lettau (1954a), Lorenz (1955a, 1967), Dutton (1973) and Pearce (1978) have been discussed. All of the approaches utilized principles of mass, energy and entropy. The differences, which outwardly appear substantial, are explained in terms of differences in postulated equilibrium temperatures and their time dependency. The derivation of reversible and entropy component models of total energy from the premise of underlying thermal equilibria reveals a common basis for these models involving the second law of thermodynamics.

Before concluding this summary, however, brief attention is now directed to the classical works of scientists prior to World War II who provided contributions to the thermodynamics discussed in this section that are basic to global monsoonal circulations, namely, Stokes (1845, 1846), Reynolds (1894), Margules (1903), Sandström (1916) and V. Bjerknes (1916).

Margules (1903) established the concept of available potential energy by showing that the kinetic energy of storms could be realized from the total potential energy stored within the baroclinic structure of the atmosphere. Margules (1903) credited Blasius (1875) and Bigelow (1903) for the idea that the energy of cyclones is associated with existing temperature gradients in storms and the overturning of air masses of unequal temperature. Margules' emphasis that a baroclinic atmosphere constitutes a store of available potential energy and that the kinetic energy of storms need not be realized immediately from differential heating constitutes a fundamental contribution to thermodynamics of reversible processes.

Two other classical contributions to thermodynamics involve the dissipation function and the transformation between kinetic and internal energies. Reynolds (1894) credited Stokes (1845, 1846) with obtaining energy equations that contained a positive definite function for arbitrary viscous stresses from a transformation of the equations of motion. This function represents the difference between the work done on the fluid by the environment through boundary stresses and the rate of increase of kinetic energy. He concluded that the positive definite function constitutes the transformation of kinetic energy into heat. Reynolds credited Lord Rayleigh (1880) for naming this, the dissipation function.

In developing his criterion for the onset of turbulence, Reynolds (1894) emphasized that the difference between the rate of increase of kinetic energy for a fluid element and the rate at which work is done by boundary stresses equals the rate at which heat is converted into energy of motion. Thus the concept of a transformation between kinetic and internal energies within an open system that provided for the exchange of energy and momentum through the action of stresses was also introduced early in the development of thermodynamics A detailed treatment of energy transformations based upon the classical work of Stokes, Rayleigh and Reynolds for both open and closed systems is set forth by Van Mieghem (1973). This approach also formed the basis for Starr's (1948a; 1951b) claim for hydrostatic fluids that regions of horizontal velocity divergence are the primary source of kinetic energy.

A fourth contribution involves the generation of kinetic energy by differential heating. Roots for the emphasis regarding the relative positions of heat sources and sinks in the atmosphere may be traced to Sandström's (1915, 1916) and V. Bjerknes (1916) works. V. Bjerknes et al., (1933, p. 186; Godske et al., 1957, p. 262) acknowledged Sandström (1915, 1916) for extending the Bjerknes' circulation theorem to include a thermodynamic cycle, although V. Bjerknes (1916) parallel contribution seems to be of equal importance. Sandström's theorem [after Godske et al., (1957)] states, "If a circulating motion is to increase (or be maintained against frictional forces) in the atmosphere, the heat must be supplied at low and extracted at high levels". In a study of idealized thermally produced circulation within closed tubes, V. Bjerknes stated [after Godske et al., (1957) pp. 280-281]: "The normal circulation in a tube in which the warm and cold sources be in different branches is characterized by ascending motion through the warm source and descending motion through the cold source. The circulation is accelerated, transforming heat into kinetic energy, when the cold source is situated above the warm source and retarded in the opposite case". Both of these results are emphasized in this study of the forcing of global monsoonal circulations.

Jeffreys (1925, 1936) critized both Sandström's and Bjerknes' results concerning the statement that the heat source and heat sink must be at different heights. In his emphasis on the importance of temperature gradients on level (geopotential) surfaces in creating motion, Jeffreys noted that the heat source and heat sink need not occur at different heights. The dilema over these issues at the time of this criticism is readily apparent in the discussion that follows Jeffreys' (1925) paper. In an exchange of correspondence, Godske (1936) and Jeffreys (1936) partially resolve the source of the differences. However, no clear perspective on just how differential heating is related to the maintenance of baroclinic structure and how kinetic energy is realized through baroclinic processes is gained until results emerge from extensive general circulation research after World War II (see Starr and Saltzman, 1966).

Up to the point in Sandström's and Bjerknes' development, where it is concluded that mechanical energy can only be realized in circulations when the work of expansion occurs under greater pressure than the contraction, Jeffreys (1925) agreed with Sandström's and Bjerknes' results. The critical issue is to resolve which part of the energy gained within regions of heating is used to increase the kinetic energy of motion and which part is delivered to the regions of cooling to offset the energy lost. Available potential energy theory and the results in this study clarify these matters and Jeffreys' criticism. These results state that, at least within some isentropic layers, net generation of available potential energy requires heating to occur at higher pressure than the cooling, while net production of kinetic energy requires the horizontal divergence of the isentropic mass transport to occur at higher values of dry static energy than the convergence. Since the reference levels against which these processes must be referred are isentropic surfaces and since the dry static energy is determined by the combination of pressure and height, it seems that neither Sandström or Jeffreys were totally correct, at least in this general application to global monsoonal circulations. These conclusions are only valid insofar as systems are closed with regard to boundary exchange of mass, momentum and energy by convective transport processes and boundary work.

In the portion of Bjerknes' (1916) results with which Jeffreys agreed, i.e., the requirement for work of expansion to occur at higher pressure than the work of contraction, Bjerknes stated that the increase of kinetic energy in the steady state system was equal to the product of an efficiency and the work of expansion. His result for kinetic energy balance (see pp. 186-192, Bjerknes et al., 1933 and/or pp. 279-281, Godske et al., 1957) within closed tubes is analogous to Eq. (7.7a) for the global circulation expressed in the form of

$$D(E) = \tilde{T}_+ \varepsilon_+ \int_{V_\theta +} (\rho J_\theta \dot{s})_+ \, dV_\theta \quad , \tag{7.180}$$

where the efficiency that Bjerknes employed from classical thermodynamics (Sommerfeld, 1950) is

$$\varepsilon_+ = (1 - \tilde{T}_- / \tilde{T}_+) \; . \tag{7.181}$$

Equivalent expressions for Bjerknes definition of work of expansion and energy received in the region of heating ($V_{\theta+}$) and work of contraction and energy lost in the region of cooling ($V_{\theta-}$) [W_W and W_C in Eqs. (8.22(3) and (4), p. 281, Godske et al. 1957] are, respectively, given by

$$W_W = W_+ = \tilde{T}_+ \int_{V_{\theta+}} (\rho J_\theta \dot{s})_+ \, dV_\theta \ , \tag{7.182}$$

$$W_C = W_- = \tilde{T}_- \int_{V_{\theta+}} (\rho J_\theta \dot{s})_+ \, dV_\theta \ . \tag{7.183}$$

The net energy involved in the kinetic energy production is given by the difference, $(W_+ - W_-)$. In these expressions it is assumed as in Section 7.2 that the magnitudes of the time rate of change of entropy gained in regions of heating and lost in regions of cooling are equal under the condition of $\dot{S}$ equal zero. Bjerknes et al., (1957) note that the efficiency ε_+ is equal to that of a Carnot cycle working between the temperatures $\tilde{T}_+$ and $\tilde{T}_-$. The efficiency increases with decreasing $\tilde{T}_-$ and increasing $\tilde{T}_+$, however it only approaches unity as $\tilde{T}_-$ decreases to absolute zero.

Sandström's (1916) extension of Bjerknes' circulation theorem and Bjerknes' (1916) definition of efficiency are fundamental to a theory of the reversible component of total energy. The efficiency defined by Bjerknes and the efficiency of available potential energy theory appear different in that the temperature $\tilde{T}_-$ in Eq. (7.181), which corresponds with the equilibrium state temperature T_α of available potential energy theory, is the mean temperature of the heat sink region. With the definition for an equilibrium temperature for the efficiency in Eq. (7.181) given by

$$T_e = (\tilde{T}_+ + \tilde{T}_-)/2 \tag{7.184a}$$

$$= \tilde{T} \ , \tag{7.184b}$$

and substitution of the efficiency into Eq. (7.20c), the difference in the generation of total energies of actual and equilibrium atmospheres, $G(\Delta E_e)$, under steady conditions is given by

$$G(\Delta E_e) = \int \rho J_\theta \dot{s} \; (\tilde{T}e_+) \, dV_\theta \quad . \tag{7.185}$$

This expression for Bjerknes' result is valid provided that $\tilde{T}$ is given by the mean value theorem in accord with the results of Section 7.2. This expression is also equivalent with the results of Lettau (1954a), Lorenz (1955a), Dutton and Johnson (1967), Dutton (1973) and Pearce (1978), thus revealing the relevance of Sandström's and Bjerknes' results.

The results of Eqs. (7.7a and b) and now Eq. (7.180), however, are of limited value in applications. Besides an inability to determine readily the the temperatures $\tilde{T}_+$ and $\tilde{T}_-$ from data, the thermodynamics of the atmosphere is not determined by a Carnot cycle with phases of isothermal expansion and contraction during heat addition and extraction that are distinct from phases of isentropic expansion and contraction.

Of all the candidates for diagnostic analysis of the atmosphere's static energy, i.e. $\Pi\text{-}E_\alpha$, $\Pi\text{-}E_{\alpha o}$ and $\Pi\text{-}E_o$, estimates of the available potential energy $\Pi\text{-}E_\alpha$ will be the most accurate. Available potential energy is primarily determined by the integrand $p^{1+\kappa}(\lambda,\phi,\theta,t) - p_e^{1+\kappa}(\theta,t)$. Since the dominant component of this integral is second moment of the pressure about p_e, the effect of bias errors from measurements and analysis of pressure will be minimized in estimation if the equilibrium pressure is estimated by the average pressure $\overline{p}(\theta,t)$. As such, the bias component from any systematic pressure errors in estimation of its quadratic component, are eliminated.

In extending Margules (1903) work, one of the most important of Lorenz's (1955a) contribution was his emphasis on differential heating and setting forth a means to determine precisely the generation of available potential energy that maintains the kinetic energy against dissipation. Lorenz pointed out vividly that the combination of heat sinks and sources creates available potential energy. Through the analysis developed herein, relations between Lorenz's results, Carathéodory's statement of the second law (Sommerfeld, 1950) and entropy exchange have been established. The entropy decreasing process of differential heating expands the admissible region of reversible isentropic processes within an energy phase space, while the entropy increasing process of viscous dissipation contracts the admissible region. At the same time within the "modus operandi" of global monsoonal circulations, differential heating forces global monsoonal circulations, the transport of energy and entropy from sources to sinks and the generation of kinetic energy.

The motivation to integrate Lorenz's concept in a perspective of reversible and irreversible processes of atmosphere circulation stems from Dutton's emphasis on the role of entropy increasing processes within the atmospheric system and a desire to resolve diverse results through a perspective of thermally forced global monsoonal circulations. As noted previously (Johnson and Downey, 1982), if the atmosphere's total energy had not been partitioned into kinetic and total potential energy, the question of maintaining the atmosphere's kinetic energy supply against viscous dissipation would not arise. Instead, the question that would need to be resolved is how the atmosphere's mass, energy and entropy balance is maintained in the presence of entropy increasing processes that occur in any natural system. Through analysis employing the postulate of an underlying equilibrium temperature for reversible and entropy components of total energy, these questions have been addressed and the means to reconcile the emphases of Lorenz (1955a), Dutton (1973) and Pearce (1978) has been provided.

C. A COMPARISON OF ISOBARIC AND ISENTROPIC RESULTS FOR GLOBAL MONSOONAL CIRCULATIONS

8. Planetary scale differential heating, mass circulation and energy transport in isobaric and isentropic coordinates

The Global Weather Experiment provided opportune data sets for isentropic analyses of planetary transport processes. Prior to the GWE, the tropical and Southern Hemispheric circulations were inadequately observed and assimilation methods used in global models were not advanced to the degree that the planetary circulation was resolved with the accuracy needed to estimate the atmosphere's heat sources and sinks. Through the GWE, however, all of the necessary resources in the form of a global observational system, the acquisition of a global data base and the development of assimilation models were realized. The results were global analyses produced with the accuracy needed to estimate the global distribution of heat source and sinks. These estimates in combination with analyses of the atmosphere's balance of mass, momentum and energy have provided insight on global monsoonal circulations. In addition, the availability of several GWE Level III data sets to the scientific community has fostered numerous studies and comparison of results from many different investigations, including the comparison of results from isentropic and isobaric analyses (First GARP Global Experiment (FGGE) Advisory Panel, 1985; WMO GARP Report, 1985a, 1985b, 1986).

Nearly all atmospheric research and numerical simulations of the past three decades have been conducted using either isobaric or sigma coordinates. Consequently, views of atmospheric circulation and the forcing and maintenance of the atmosphere's circulation have been founded on isobaric results, both from empirical and theoretical analyses (e.g., Smagorinsky, 1953; Starr and

Saltzman, 1966). With the exception of the isobaric zonally averaged circulations, the only results presented heretofore in this study have been developed using isentropic coordinates. Johnson and Downey (1975a and b) and Johnson (1980) pointed out how viewpoints become coordinate dependent in the study of a particular pheonomena through use of transport equations that relate the Eulerian tendency and divergence transport of a property to a Lagrangian source or sink. Given the established nature and general acceptance of isobaric views of the planetary circulation, in contrast with the relative infancy of isentropic views, attention will now be focused on establishing the relation between isentropic and isobaric transport processes for mass, energy and entropy. The purpose of the next three sections is to isolate the source of some of the differences between isobaric and isentropic results and to foster an appreciation of the isentropic viewpoint from a comparison of some empirical results from isobaric and isentropic analyses.

8.1 A contrast of isentropic and isobaric mass circulations for a simple system

At an International Conference on the Scientific Results of the Monsoon Experiment at Bali, a question was asked which dealt with why atmospheric heating can be determined with relative ease from analysis in isentropic coordinates and why the matter is so difficult in isobaric coordinates. For an understanding of the nature of the problem, an analysis of the mass circulation attending convection in a simple system is first discussed, after which the more general problem of contrasting mass and energy transport in isentropic and isobaric coordinates is pursued.

Based on methodology used in convective parameterization (e.g., Arakawa and Schubert, 1974) the system will consist of a cumulus cloud with upward vertical motion ($\omega_1 < 0$) within its areal extent (A_1) and downward vertical motion ($\omega_2 > 0$) within the areal extent (A_2) of the immediate environment. See schematic in Fig. 18. The areally-averaged large scale vertical motion $\overline{\omega}^A$ for the total area A equal to A_1 plus A_2 is given by

$$\overline{\omega}^A = (A_1\omega_1 + A_2\omega_2)/A \ . \tag{8.1}$$

The isobaric vertical motion is divided into adiabatic and diabatic components

$$\omega = \omega_a + \omega_d \ , \tag{8.2}$$

which through the thermodynamic equation are defined by

$$\omega_a = -\left(\frac{\partial\theta}{\partial t_p} + \underset{\sim}{U}\cdot\nabla_p\theta\right)\Big/\frac{\partial\theta}{\partial p} \text{ and } \omega_d = \dot{\theta}\Big/\frac{\partial\theta}{\partial p} \ . \tag{8.3}$$

A substitution in Eq. (8.1) with the diabatic component of the environment equal to zero and a rearrangement yields

$$A_1\omega_{d_1} = A\overline{\omega}^A - [A_1\omega_{a_1} + A_2\omega_{a_2}] \ . \tag{8.4}$$

Given the nature of atmospheric compensation as manifested in isobaric coordinates, the latent heat released within clouds does not directly warm the environment. Instead, the environment is warmed through subsidence and compression, while the latent heat release offsets cooling associated with upward motion and expansion within the cumulus. As a first approximation, the upward mass flux $A_1\omega_1/g$ within the cloud tends to balance the downward mass flux $A_2\omega_2/g$ in the environment. In such a case, the area-averaged vertical mass flux vanishes and Eq. (8.4) reduces to

$$\omega_{d_1} = -[\omega_{a_1} + (A_2/A_1)\omega_{a_2}] \ . \tag{8.5}$$

In this result, an isobaric estimate of heating is determined from the difference between the adiabatic vertical motion in the environment ($\omega_{\alpha_2} > 0$) and in the region of the convection ($\omega_{\alpha_1} < 0$) in combination with an accurate determination of the ratio of the areas of ascent to descent. The intensities of the adiabatic component in both the environment and the convective region would be calculated from Eq. (8.3). This result points out the fine state of balance that exists between diabatic and adiabatic components of isobaric vertical motion in regions of cumulus convection.

This fine state of balance and compensation is characteristic of an atmosphere in which the large scale motion tends to be non-divergent (Johnson and Downey, 1975a). Integration of the isobaric equation of continuity over the total area A and use of the divergence theorem yields

$$\int_A \frac{\partial \omega}{\partial p}\, dA = \frac{\partial}{\partial p} \int_A \omega\, dA = 0 \tag{8.6a}$$

$$= - \int \underset{\sim}{n} \cdot \underset{\sim}{U}\, d\ell = 0\ , \tag{8.6b}$$

provided that exact compensation occurs at all levels. The constraint Eq. (8.6b) simply emphasizes the result that the areally-averaged horizontal divergence of the larger scale in isobaric coordinates need not be uniquely linked with diabatic processes, thus allowing the larger scale flow away from the convective region to be non-divergent even in the presence of convection. Consequently, the isobarically analyzed mass circulation attending convection may be of a local nature that is not systematically linked with a large scale mass circulation which transports energy to a heat sink. The requirement for energy transport from the local region to its large scale environment stems from the quasi-steady state assumption as applied to the thermodynamic equation. In geostrophic regimes this energy transport will occur primarily through the isobaric eddy transport of enthalpy. Consequently, mean isobaric mass circulations do not yield direct insight on the planetary scale distribution of differential heating and energy transport in the atmosphere.

Now the consequences of the heating are explored for the same example within isentropic coordinates. For this example, quasi-steady conditions are assumed for the hydrostatic mass ($g^{-1}\partial p/\partial\theta$) distribution, an inverse measure of the thermal stratification. The areally averaged vertical mass flux in isentropic coordinates defined by

$$\overline{\rho J_\theta \dot{\theta}}^A = \frac{1}{A}\int \rho J_\theta \dot{\theta}\, dA \tag{8.7}$$

is given by

$$\overline{\rho J_\theta \dot{\theta}}^A = (A_1/A)\, \rho J_\theta \dot{\theta} > 0 \,. \tag{8.8}$$

With the condition that the vertical mass flux through the top of the atmosphere vanishes and use of Eq. (8.8), a vertical integration of the areally averaged isentropic equation of continuity from the level of maximum heating θ_m to the top of the atmosphere θ_T, yields

$$\frac{1}{A}\int_A \int_{\theta_m}^{\theta_T} \frac{\partial}{\partial\theta} \rho J_\theta \dot{\theta}\, d\theta dA = -\, (A_1/A) \rho J_\theta \dot{\theta}\Big|^{\theta_m} \tag{8.9a}$$

$$= -\,\frac{1}{A}\int_{\theta_m}^{\theta_T}\oint \underset{\sim}{n} \cdot \rho J_\theta \underset{\sim}{U}\, d\ell d\theta < 0 \,. \tag{8.9b}$$

This establishes that the mass circulation above the level of maximum heating must be divergent, while below this level the mass circulation must be convergent. Since these conditions result in systematic horizontal mass transport through the lateral boundary, the mass circulation in the region of convection is linked with the mass circulation of the environment regardless of the characteristic state of dynamic balance of the motion field, provided that the mass distribution remains quasi-steady.

The contrast of the isobaric and isentropic results for this simple example points out an important consideration. It is the quasi-steady approximation in the isentropic mass continuity equation which requires a link between heating and transport processes within a system and its environment. In isobaric coordinates the link is established indirectly through the quasi-steady state nature of the temperature distribution as this state is maintained by nonlinear advection within the thermodynamic equation. This difference stems from the thermodynamic principle that entropy change is related to heating through an exact differential which uniquely determines the vertical mass flux through isentropic surfaces. This is not so in isobaric coordinates. In isobaric coordinates knowledge of volume expansion (thermodynamic work--an inexact differential) is needed in order to be able to estimate heating from changes in internal energy (or enthalpy). The fact that consistent heating distributions and isentropic mass circulations have been determined is evidence that the atmosphere responds directly to the heating at the planetary scale. In the GWE data, the information for inference of the heating actually comes from observations of the systematic isentropic mass transport and the isentropic distribution of mass at two different time periods.

Another important consideration is evident in the schematic in Fig. 19. The length scale of the horizontal branches of the isentropic mass circulation is determined by the distance between heat source and sink. Analysis of the isentropic mass circulation thus yields insight directly on the planetary scale differential heating. Analysis of the isobaric mass circulation does not.

The thermodynamics of this simple system up to this point have not involved explicitly any requirements for isentropic energy transport. However, the requirement for energy transport from the region is implicit in the quasi-steady state applied to the isentropic mass continuity equation. As emphasized earlier, the hydrostatic mass distribution in isentropic coordinates determines the total potential energy distribution. Consequently, the quasi-steady state assumption applied to the mass continuity equation and the requirement of a systematic mass circulation with convergence in lower isentropic layers, divergence in upper isentropic layers, and upward diabatic mass flux provides for systematic export of energy from the domain. As established earlier, in a hydrostatic atmosphere the dry static energy increases monotonically upwards. Thus in the upper layers of the mass circulation more dry static energy is transported away from the region of convection than is returned in the lower layers thus, insuring a net divergence of energy from the simple system.

8.2 A contrast of energy exchange processes in isobaric and isentropic coordinates

With regard to the premise stated in the introduction and the perspective of the forcing of global monsoonal circulations developed in isentropic coordinates, a key question remains. How can the emphasis placed on transient modes of isobaric energy transport within baroclinic circulations of middle latitudes be reconciled with the prime importance of mean energy transport by the isentropic mass circulation? Johnson and Downey (1975a) called attention to the fact that where geostrophy prevails, mean modes of mass transport in isentropic coordinates exist which when averaged differ from corresponding averages in isobaric coordinates.

Up to this point one can conclude that the mass, energy and entropy transport are convergent in lower isentropic layers and divergent in upper isentropic layers in heat source regions, while the opposite distribution of divergence occurs in heat sink regions. The condition that the regions of heat sources and heat sinks are distinct requires that irrotational modes of energy and entropy transport must be of the same scale as the irrotational mode of mass transport. This constraint exists regardless of the nature of the dynamical balance of the wind field. In view of the constancy of entropy within an isentropic layer, the transport of entropy is uniquely linked with the isentropic mass circulation. Even though entropy and dry static energy are closely related in the sense that $c_p\theta$ approximates ψ, this result does not reveal why energy transport occurs through the isentropic mean mass circulation in the time-averaged structure.

If a link can be established between energy transport associated with transient circulations in isobaric coordinates and mean mass circulations in isentropic coordinates, the view that the transient baroclinic waves in middle latitudes are simply a mode of response to transport energy from heat source to heat sink can be established. Thus, the fundamental premise that a differentially heated atmosphere will develop a mass circulation which transports energy from heat source to heat sink can then be defended. In this section the relations between the isentropic and isobaric divergence of mass, energy and entropy are studied in order to gain insight into the link between mean modes of transport in isentropic and mean and transient modes in isobaric coordinates.

The reasons for similarities and differences between energy transport in isobaric and isentropic coordinates will be established by expressing the energy balance in generalized coordinates (Johnson, 1980). This balance from Eq. (6.19) is expressed by

$$\frac{\partial}{\partial t_\eta}(\rho J_\eta e)+\nabla_\eta\cdot(\rho J_\eta \underset{\sim}{U}\nu) + \frac{\partial}{\partial \eta}(\rho J_\eta \dot{\eta}\nu) + \frac{\partial}{\partial \eta}\left(\frac{\partial z}{\partial \eta}J_\eta^{-1} p \frac{\partial z}{\partial t_\eta}\right) = \rho J_\eta Q_m + \rho J_\eta \underset{\sim}{U}\cdot\underset{\sim}{F} \quad .(8.10)$$

The factor $(\partial z/\partial \eta)J_\eta^{-1}$ in the fourth term equals unity for a monotonic increasing relation between η and z, as in isentropic coordinates, and minus unity for a monotonic decreasing relation between η and z, as in isobaric and sigma coordinates.

With Poisson's equation and a rearrangement, the energy transport is given by

$$\frac{\partial}{\partial t_\eta}(\rho J_\eta e)+\nabla_\eta\cdot(\rho J_\eta \underset{\sim}{U}\nu) + \frac{\partial}{\partial \eta}\left[\rho J_\eta\left(\dot{\eta} - \frac{\partial \eta}{\partial \theta}\dot{\theta}\right)\psi\right] + \frac{\partial}{\partial \eta}(\rho J_\eta \dot{\eta} k) + \frac{\partial}{\partial \eta}\left(\frac{\partial z}{\partial \eta}J_\eta^{-1} p \frac{\partial z}{\partial t_\eta}\right)$$

$$= -\frac{\partial \theta}{\partial \eta}\psi\frac{\partial}{\partial \theta}\left(\frac{\partial z}{\partial \eta}J_\eta^{-1}\rho J_\theta \dot{\theta}\right) + \rho J_\eta \underset{\sim}{U}\cdot\underset{\sim}{F} \quad . \qquad (8.11)$$

Traditionally (Panofsky, 1946), the vertical motion ($\dot{\eta}$) is considered to be the sum of adiabatic ($\dot{\eta}_a$) and diabatic ($\dot{\eta}_d$) components where, from the thermodynamic equation, the two components are defined by

$$\dot{\eta}_a = -\frac{\partial \eta}{\partial \theta}\left(\frac{\partial \theta}{\partial t_\eta} + \underset{\sim}{U}\cdot\nabla_\eta \theta\right) \quad \text{and} \quad \dot{\eta}_d = \frac{\partial \eta}{\partial \theta}\dot{\theta} \quad . \qquad (8.12)$$

A substitution yields

$$\frac{\partial}{\partial t_\eta}(\rho J_\eta e) + \nabla_\eta \cdot (\rho J_\eta \underset{\sim}{U} \nu) + \frac{\partial}{\partial \eta}(\rho J_\eta \dot{\eta}_a \psi) + \frac{\partial}{\partial \eta}(\rho J_\eta \dot{\eta} k) + \frac{\partial}{\partial \eta}\left(\frac{\partial z}{\partial \eta} J_\eta^{-1} p \frac{\partial z}{\partial t_\eta}\right)$$

$$= -\left(\frac{\partial z}{\partial \eta} J_\eta^{-1}\right)\frac{\partial \theta}{\partial \eta}\, \psi \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta}) + \rho J_\eta \underset{\sim}{U} \cdot \underset{\sim}{F} \; . \tag{8.13}$$

In the first right hand term of Eq. (8.13), the diabatic component of vertical energy flux from the third left hand term of Eq. (8.10) has been combined with the heat addition of the right hand side to yield a net term involving the product of a static stability measure, the dry static energy and a component from the isentropic mass continuity equation. This key term, when expressed in isentropic coordinates and vertically integrated, was used to determine the sense of the energy and entropy transport in relation to the mass circulation. The general result shows that the vertical exchange of dry static energy in all coordinate systems is linked with both adiabatic and diabatic components of vertical motion in thermally forced systems, the exception is isentropic coordinates where the adiabatic component vanishes.

A contrast of Eq. (8.13) expressed in isobaric and isentropic coordinates will help to resolve the physical relation between the time-averaged energy transport processes. With Eq. (8.13) the quasi-steady time-averaged isobaric energy equation becomes

$$g^{-1}\left[\overline{\nabla_p \cdot (\underset{\sim}{U}\nu)} + \frac{\partial}{\partial p}\overline{(\omega_a \psi)} + \frac{\partial}{\partial p}\overline{(\omega k)}\right] - \frac{\partial}{\partial p}\overline{\left(p \frac{\partial z}{\partial t_p}\right)} = \overline{\psi \frac{\partial \theta}{\partial p} \frac{\partial}{\partial \theta}(\rho J_\theta \dot{\theta})} + g^{-1}\, \overline{\underset{\sim}{U} \cdot \underset{\sim}{F}} \; . \tag{8.14}$$

In this form the boundary flux of energy in isobaric coordinates on the left hand side is related to the physical processes of heating and viscous dissipation on the right hand side. The source of energy by heating involves a product of the dry static energy, a static stability measure and the vertical derivative of isentropic mass transport with respect to potential temperature. Although the left hand side of the equation is determined by boundary processes, an ideal form for insuring energy conservation, the product of ($\psi\ \partial\theta/\partial p$) in the heating term is not necessarily a monotonic function of pressure. This precludes the development of a unique relation between the isobaric mass circulation and vertically integrated energy transport such as derived earlier for isentropic coordinates.

Through expression of Eq. (8.13) in isobaric coordinates, multiplication by the Jacobian of transformation from isobaric to isentropic coordinates ($-\partial p/\partial\theta$), and subsequent isentropic time-averaging the isobaric energy equation becomes

$$-\{g^{-1}\left[\overline{\frac{\partial p}{\partial\theta}\frac{\partial}{\partial t_p}(e)} + \overline{\frac{\partial p}{\partial\theta}\nabla_p\cdot(\underset{\sim}{U}\nu)} + \overline{\frac{\partial p}{\partial\theta}\frac{\partial}{\partial p}(\omega a\psi)} + \overline{\frac{\partial p}{\partial\theta}\frac{\partial}{\partial p}(\omega k)}\right] - \overline{\frac{\partial p}{\partial\theta}\frac{\partial}{\partial p}\left(p\frac{\partial z}{\partial t_p}\right)}\}]/\overline{\rho J_\theta}$$

$$= \left[-\psi\overline{\frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})} + g^{-1}\,\overline{\rho J_\theta \underset{\sim}{U}\cdot\underset{\sim}{F}}\right]/\overline{\rho J_\theta}\ . \tag{8.15}$$

The quasi-steady time-averaged isentropic energy balance determined directly from Eq. (8.13) is

$$\left[\overline{\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}\nu)} + \overline{\frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta}k)} + \overline{\frac{\partial}{\partial\theta}\left(p\frac{\partial z}{\partial t_\theta}\right)}\right]/\overline{\rho J_\theta}$$

$$= \left[-\psi\overline{\frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})} + \overline{\rho J_\theta \underset{\sim}{U}\cdot\underset{\sim}{F}}\right]/\overline{\rho J_\theta}\ . \tag{8.16}$$

The right hand sides of Eqs. (8.15) and (8.16) are identical, an equality which requires that the sums of all left hand terms of the two equations are also equal. In a comparison of the left hand side terms, the terms involving the correlation between inverse static stability and the isobaric tendency of total energy and also the adiabatic vertical exchange of dry static energy in isobaric coordinates have no explicit counterpart in isentropic coordinates. The left hand terms of Eq. (8.16) individually integrate to boundary processes and are directly related to the isentropic mass circulation and frictional dissipation. As noted earlier, with a vertical integration over the entire atmosphere, Eq. (8.16) becomes the expression used to determine the sense of the energy transport in Eq. (6.25) by the isentropic mass circulation. In the case of Eq. (8.15), the left hand terms are likewise related to the same processes, but the individual terms will only integrate to boundary conditions with the simplifying assumption of an invariant Jacobian of transformation $-\partial p/\partial\theta$. With a barotropic atmosphere, as the case tends to be in the tropics, $\partial p/\partial\theta$ is invariant in time and space and ω equals ω_d as ω_a vanishes. Thus, Eqs. (8.15) and (8.16) become identical and energy transport by the isobaric and the isentropic mass circulations will correspond with each other. Furthermore, in this region the results derived previously for the close relation between isentropic mass, energy and entropy transport heat sources and sinks become directly applicable in isobaric coordinates.

In regions other than the tropics, the isobaric energy balance $(c_pT+\phi)$ is complicated by the vertical flux of energy associated with ω_a and the related fact that large scale exchange of all properties occurs through geostrophic motion. The constraint that the isobaric geostrophic advection of dry static energy in extratropical latitudes is restricted to the advection of enthalpy requires a vertical exchange between geopotential and internal energies in the maintenance of hydrostatic balance (Johnson, 1970). Thus, in the second and third terms of Eq. (8.15) the isobaric horizontal divergence of the energy transport and the vertical derivative of the adiabatic component of energy transport tend to be large and usually of opposite sign. This exchange is compounded by static stability which is quite variable within the baroclinic structure of extratropical latitudes. Consequently, isobaric energy balance within baroclinic waves in the atmosphere is subtle and difficult to resolve diagnostically, particularly if one attempts to relate the vertical and horizontal structure of the energy balance to the scale of the heat sources and sinks that ultimately force the atmosphere's circulation.

8.3 On geostrophic modes of mass, energy and entropy transport in isobaric and isentropic coordinates

Diagnostic results comparing energy transport at the planetary scale during the GWE have confirmed empirically for the same data sample that the energy transport in isentropic coordinates occurs by the mean mode, while the corresponding energy transport in isobaric coordinates occurs by the combination of mean and transient modes (Johnson, 1984). Reasons for the link between the mass circulation and geostrophic energy transport in isentropic coordinates and for the differences between isentropic and isobaric coordinates are now identified.

In this analysis generalized coordinates will be used in order to insure that differences in results do not stem from mathematical definitions. Modes of *total flow energy* transport are now determined through an expansion of the divergence of the horizontal energy transport given by

$$\overline{\nabla_\eta \cdot (\rho J_\eta \underset{\sim}{U} \nu)} = \overline{\nu \, \nabla_\eta \cdot (\rho J_\eta \underset{\sim}{U})} + \overline{\rho J_\eta \underset{\sim}{U} \cdot \nabla_\eta \nu} \quad . \tag{8.17}$$

Assuming that the transport in latitudes other than the tropics tends to be by geostrophic motion and the condition that $(c_pT+\phi)>>k$, the divergences of the *total flow energy* transport in hydrostatic isobaric and isentropic coordinates become, respectively,

$$\overline{\nabla_p \cdot (\rho J_p \underset{\sim}{U}_g \psi)} = \overline{(1/g) \underset{\sim}{U}_g \cdot \nabla_p c_p T} \tag{8.18}$$

and

$$\overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}_g \psi)} = \overline{\psi \nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}_g)} \quad . \tag{8.19}$$

The isobaric horizontal divergence of the geostrophic energy transport occurs through geostrophic enthalpy advection. The isentropic equivalent occurs through the divergence of the geostrophic mass transport, a mode of transport that is directly linked to the distribution of diabatic heating via the isentropic equation of continuity. With use of the geostrophic relation

$$f(\underset{\sim}{k} \times \underset{\sim}{U}_g) = - \nabla_p \phi = - \nabla_\theta \psi \ , \tag{8.20}$$

Eqs. (8.17) and (8.18) are expressed by

$$\overline{\nabla_p \cdot (\rho J_p \underset{\sim}{U}_g \psi)} = (g^{-1}) \overline{(\underset{\sim}{k} \times \nabla_p \phi) \cdot \nabla_p (\psi/f)} \tag{8.21}$$

and

$$\overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}_g \psi)} = \overline{\psi (\underset{\sim}{k} \times \nabla_\theta \psi) \cdot \nabla_\theta (\rho J_\theta / f)} \tag{8.22a}$$

$$= \overline{[\underset{\sim}{k} \times \nabla_\theta (\psi^2/2)] \cdot \nabla_\theta (\rho J_\theta / f)} \quad . \tag{8.22b}$$

The divergences of the geostrophic entropy transport are

$$\overline{\nabla_p \cdot (\rho J_p \underset{\sim}{U}_g s)} = (g^{-1}) \ \overline{(\underset{\sim}{k} \times \nabla_p \phi) \cdot \nabla_p (s/f)} \tag{8.23}$$

and

$$\overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}_g s)} = s\ \overline{\nabla_\theta \cdot (\rho J \underset{\sim}{U}_g)}$$

$$= s\overline{(\underset{\sim}{k} x \nabla_\theta \psi) \cdot \nabla_\theta (\rho J_\theta / f)} \qquad (8.24)$$

Since the entropy is uniform on an isentropic surface, the field of entropy divergence is directly proportional to the mass divergence. The corresponding divergences of the geostrophic mass transport are

$$\overline{\nabla_p \cdot (\rho J_p \underset{\sim}{U}_g)} = (g^{-1})\ \overline{(\underset{\sim}{k} x \nabla_p \phi) \cdot \nabla_p (1/f)} \qquad (8.25)$$

and

$$\text{and } \overline{\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}_g)} = \overline{(\underset{\sim}{k} x \nabla_\theta \psi) \cdot \nabla_\theta (\rho J_\theta / f)}\ . \qquad (8.26)$$

A comparison of the mass, energy and entropy divergence reveals a commonality in isentropic coordinates that is not present within isobaric coordinates. The isentropic advecting velocities for all three are colinear, while the gradient of the property being advected $\nabla_\theta(\rho J_\theta/f)$ is identical. The isobaric advecting velocities for all three are identical (a special case of colinearity), but, none of the properties being advected, (ψ/f), (s/f) and $(1/f)$, are uniquely related with each other. Thus there is no constraining relation between the divergence of the geostrophic transport mass, energy and entropy in isobaric coordinates, nor is one expected since the gradients of the three properties in isobaric coordinates are different. In contrast, the isentropic relations <u>do not permit</u> a degree of freedom at any time in the space domain for geostrophic motion to generate a transient component of energy or entropy transport that is independent of the mass circulation. If transient modes develop that transport dry static energy from heat source to heat sink, the modes must occur through the ageostrophic component of motion.

With regard to the distributions of isentropic divergence of geostrophic transport of mass, energy and entropy, a comparison of Eqs. (8.22), (8.24) and (8.26) shows that the zero isopleths of all three fields will be coincident. Thus, regions of the divergence and convergence of mass, energy and entropy will be coincident. Such a result is essential for the scale of the irrotational mode of transport of all properties to be the same. The results of Eqs. (6.25) and (6.29) for vertically integrated transport shows that locations of maxima and minima of heating (disregarding E^2 as being small) will also correspond respectively with the locations of maxima and minima of energy and entropy divergence. Since the rotational mode of transport for each of these properties is nondivergent, the only mode that is involved directly with the satisfaction of energy balance requirements for each of these three properties is the irrotational mode.

It is important to realize, however, that actual mass, energy and entropy transport occur by both rotational and irrotational modes at all times. At the planetary scale, the divergence of horizontal energy and entropy transport associated with convective heating which initially forces irrotational modes that subsequently, through adjustment, shift to rotational modes. The commonality of the divergence of the isentropic geostrophic mass, energy and entropy transport permits the irrotational modes of each of these properties to be common with each other. As a consequence, the energy and entropy transport by irrotational motion remains coupled with the irrotational mode of the mass circulation.

With the lack of commonality of the divergences of the geostrophic mass, energy and entropy in isobaric coordinates and the close relationship between geostrophic and rotational modes of atmospheric motion, the distribution of the rotational and irrotational modes among these properties cannot be common with each other. Consequently, as the simple example illustrated, the length scales identified with mass circulations in isobaric coordinates differ from their counterpart in isentropic coordinates and cannot be readily related to planetary scales of differential heating.

8.4 Relations between isentropic and isobaric modes of mass, energy and entropy transport

In the past decade, Lagrangian perspectives of mean meridional circulations have been set forth by Kida (1977) and Andrews and McIntyre (1978a, 1978b). The Lagrangian mean meridional circulations are remarkably similar to isentropic Hadley circulations (Henderson, 1971; Zillman, 1972; Otto, 1974; Dutton, 1976a; Gallimore and Johnson, 1977a, 1981a; Townsend, 1980, Townsend and Johnson, 1985). These analyses were preceded by Johnson and Dutton's (1969) results of isentropic Hadley circulations presented at the London Conference on the Global Circulation of the Atmosphere.

Prior to the GWE, analyses of the isentropic Hadley circulations were based on climatotlogical modeling of the temporally, zonally averaged heating distribution with one exception. Henderson's (1971) analysis of the Northern Hemisphere Hadley circulation was based on the National Meteorological Center's gridded height information and use of the geostrophic approximation. Kida's results were determined from an analysis of the Lagrangian mean meridional circulation that was numerically simulated in isobaric coordintes.

Andrews and McIntyre results were based directly on a theoretical analysis of the role of nonlinear baroclinic waves in forcing Lagrangian circulations. Through use of quasi-geostrophic theory, Edmon et al. (1980) approximated the Lagrangian mean circulations by defining a transformed meridional circulation from the sum of the Eulerian isobaric zonally-averaged meridional velocity and an eddy entropy function. The zonally averaged mass transport for the transformed equation of continuity is called the residual meridional circulation. Recently, the Southern Hemisphere circulation has been studied through an equivalent transformation of the isobaric Eulerian time-averaged circulation to time-averaged residual meridional circulation (Trenberth, 1986a and b).

The key to determining a residual meridional circulation from isobaric analysis is to redefine eddy entropy transport as a "component" of the meridional velocity. Townsend and Johnson (1985) employing transformations developed in this section discussed the relations between zonally-averaged isobaric residual and zonally-averaged isentropic mass circulations. The following development sets forth the transformations from isentropic to isobaric coordinates within the time-averaged domain.

With the vector identity

$$\nabla_\theta \cdot (\ ) = \nabla_p \cdot (\ ) - \nabla_p \theta \cdot \frac{\partial}{\partial \theta} (\ ) \quad , \qquad (8.27)$$

and the hydrostatic assumption, the divergences of the isentropic mass, energy and entropy transport are, respectively,

$$[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})]/\rho J_\theta = \nabla_p \cdot \underset{\sim}{U} - \frac{\partial}{\partial p}[(\underset{\sim}{U} \cdot \nabla_p \theta)/\frac{\partial \theta}{\partial p}] \ , \tag{8.28}$$

$$[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}\psi)]/\rho J_\theta = \nabla_p \cdot (\underset{\sim}{U}\psi) - \frac{\partial}{\partial p}[(\psi \underset{\sim}{U} \cdot \nabla_p \theta)/\frac{\partial \theta}{\partial p}] \ , \tag{8.29}$$

$$[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} s)]/\rho J_\theta = \nabla_p \cdot (\underset{\sim}{U} s) - \frac{\partial}{\partial p}[(s \underset{\sim}{U} \cdot \nabla_p \theta)/\frac{\partial \theta}{\partial p}] \ . \tag{8.30}$$

In tropical latitudes, where the atmosphere tends to be barotropic and temperature advection is minimal, the following time-averaged approximate relations hold

$$\overline{[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U})]/\rho J_\theta} \simeq \nabla_\theta \cdot \overline{\underset{\sim}{U}}_{ag} \simeq \nabla_p \cdot \overline{\underset{\sim}{U}}_{ag} \ , \tag{8.31}$$

$$\overline{[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U}\psi)]/\rho J_\theta} \simeq \nabla_\theta \cdot (\overline{\underset{\sim}{U}_{ag}\phi}) \simeq \nabla_p \cdot (\overline{\underset{\sim}{U}_{ag}\phi}) \ , \tag{8.32}$$

$$\overline{[\nabla_\theta \cdot (\rho J_\theta \underset{\sim}{U} s)/\rho J} = s \ \overline{[\nabla_\theta \cdot (\rho J \underset{\sim}{U})/\rho J]} \simeq s \nabla_\theta \cdot \overline{\underset{\sim}{U}}_{ag} \simeq s \nabla_p \cdot \overline{\underset{\sim}{U}}_{ag} \ . \tag{8.33}$$

The fact that the isentropic and isobaric divergences of various properties tend to be equal in tropical regions accounts for the similarities of the mass circulation, energy transport and entropy transport in these latitudes. Thus, as noted earlier, the isobaric mass circulation is directly linked with heat sources and sinks. Furthermore, in view of the coincidence of isentropic and isobaric surfaces in a barotropic atmosphere, there will be a direct one-to-one correspondence, respectively, between the mean and eddy modes in isentropic coordinates with mean and eddy modes in isobaric coordinates. Inspection of Eqs. (8.13) and (8.14) will verify that the adiabatic component of ω also vanishes in barotropic regions.

In extratropical latitudes where the atmosphere is baroclinic, the motion is primarily geostrophic and in turn is linked with adiabatic components of isobaric vertical motion. It is in these regimes that differences emerge with regard to the perspective of transport processess. If the motion is geostrophic, the divergences of the mass, energy and entropy transport become

$$[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}_g)]/\rho J_\theta = \nabla_p\cdot\underset{\sim}{U}_g - \frac{\partial}{\partial p}\Big[(\underset{\sim}{U}_g\cdot\nabla_p\theta)/\frac{\partial\theta}{\partial p}\Big] = -\Big[\frac{\partial}{\partial\theta}(\rho J_\theta\dot{\theta})\Big]/\rho J_\theta , \quad (8.34)$$

$$[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}_g\psi)]/\rho J_\theta = \nabla_p\cdot(\underset{\sim}{U}_g c_p T) - \frac{\partial}{\partial p}\Big[(\psi\underset{\sim}{U}_g\cdot\nabla_p\theta)/\frac{\partial\theta}{\partial p}\Big] , \quad (8.35)$$

$$[\nabla_\theta\cdot(\rho J_\theta \underset{\sim}{U}_g s)]/\rho J_\theta = \nabla_p\cdot(\underset{\sim}{U}_g s) - \frac{\partial}{\partial p}\Big[(s\underset{\sim}{U}_g\cdot\nabla_p\theta)/\frac{\partial\theta}{\partial p}\Big] . \quad (8.36)$$

In extratropical latitudes, isentropic and isobaric surfaces in baroclinic systems are far from being coincident. Therefore, temporal averages in one coordinate system will be different from averages in the other coordinate system. In order to relate time-averaged transports in the two coordinates systems, Eqs. (8.34), (8.35) and (8.36) will be temporally averaged isentropically and then partitioned into stationary and transient components.

With the use of the mass-weighted time-average and its deviation in isentropic coordinates, Eqs. (8.34), (8.35) and (8.36) are expressed by

$$[\nabla_\theta\cdot(\overline{\rho J_\theta}\hat{\underset{\sim}{U}}_g)]/\overline{\rho J_\theta} = \widehat{\nabla_p\cdot\underset{\sim}{U}_g} - \Big\{\frac{\partial}{\partial\theta}\{\overline{\rho J_\theta}[\hat{\underset{\sim}{U}}_g\cdot\widehat{\nabla_p\theta} + \widehat{\underset{\sim}{U}_g^{*}\cdot(\nabla_p\theta)^{*}}]\}\Big\}/\overline{\rho J_\theta} , \quad (8.37a)$$

$$= -\Big[\frac{\partial}{\partial\theta}(\overline{\rho J_\theta}\hat{\dot{\theta}})\Big]/\overline{\rho J_\theta} , \quad (8.37b)$$

$$\{\nabla_\theta \cdot [\overline{\rho J_\theta}(\hat{U}_g \hat{\psi} + \widehat{U_g^{*} \psi^{*}})]\}/\overline{\rho J_\theta} = c_p[\underline{\hat{T}\widehat{\nabla_p \cdot U_g}} + \widehat{T^{*}(\nabla_p \cdot U_g)^{*}} + \hat{U}_g \cdot \widehat{\nabla_p T} + \widehat{U_g^{*} \cdot (\nabla_p T)^{*}}]$$

$$\underline{- \{\frac{\partial}{\partial\theta}\{\overline{\rho J_\theta}[\hat{\psi}(\hat{U}_g \cdot \widehat{\nabla_p \theta} + \widehat{U_g^{*} \cdot \nabla_p \theta^{*}})]\}\}/\overline{\rho J_\theta}}$$

$$- \{\frac{\partial}{\partial\theta}\{\overline{\rho J_\theta}[\widehat{\psi^{*}(U_g \cdot \nabla_p \theta)^{*}}]\}\}/\overline{\rho J_\theta} \ , \tag{8.38}$$

$$[\nabla_\theta \cdot (\overline{\rho J_\theta}\hat{U}_g s)]/\overline{\rho J_\theta} = \{\underline{s\widehat{\nabla_p \cdot U_g}} + [\hat{U}_g \cdot \widehat{\nabla_p s} + \widehat{U_g^{*} \cdot (\nabla_p s)^{*}}]\}$$

$$\underline{- \{\frac{\partial}{\partial\theta}\{s\overline{\rho J_\theta}[\hat{U}_g \cdot \widehat{\nabla_p \theta} + \widehat{U_g^{*} \cdot (\nabla_p \theta)^{*}}]\}\}/\overline{\rho J_\theta}}$$

$$= - s[\frac{\partial}{\partial\theta}(\overline{\rho J_\theta}\hat{\dot{\theta}})]/\overline{\rho J_\theta} \ . \tag{8.39}$$

The relation in Eq. (8.37) for the time-averaged isentropic mass divergence expressed in isobaric coordinates, while complicated, shows that both stationary and transient modes of isobaric geostrophic advection of potential temperature (entropy) are components of the isentropic mean mass circulation and hence are directly linked to the thermally forced isentropic mass circulation.

The only link between mean isentropic and mean isobaric geostrophic mass circulations is through the divergence of mass associated with $v\partial f/\partial y$, since this is the only degree of freedom for the isobaric divergence of the geostrophic mass transport. This degree of freedom in the three relations, however, is not the primary link with isentropic divergence of the mass transport, nor in turn with the vertical derivative of the diabatic mass flux.

If it were, the isobaric and isentropic mass circulations would be similar everywhere.

The primary degree of freedom for the link with the divergence of the isentropic mass transport in middle latitudes is the vertical derivative of the isobaric geostrophic advection of potential temperature multiplied by the hydrostatic mass $\overline{\rho J_\theta}$. This quantity may be inverted to become $-g\ \partial\theta/\partial\overline{p}$, in which case the last term of Eq. (8.37a) may be expressed by $\partial\widehat{[\underset{\sim}{U}_g^{*}\cdot(\nabla_p\theta)^{*}/(\partial\theta/\partial\overline{p})]}/\partial\overline{p}$, where $\overline{p}$ is the isentropic time-averaged pressure. With an assumption that $(\partial\overline{\theta}/\partial p)^{-1}$ in isobaric coordinates approximates $\partial\overline{p}/\partial\theta$ in isentropic coordinates, a non-divergent quasi-geostrophic approximation ($f = f_o$) and an indefinite integration meridionally, the resulting expression is equivalent in form to the eddy quantity $\overline{v'\theta'}/(\partial\overline{\theta}/\partial p)$ used in defining residual meridional motion (Eq. 2.4b, Edmon et al., 1980). This quantity is the primary difference between the Eulerian isobaric and isentropic mass circulations. This relation which transforms the Eulerian isobaric mass circulation in Eq. (8.37a), as it is determined by the isobaric mass divergence to the isentropic mass circulation, deals directly with potential temperature (entropy) advection as a physical process. Its link with energy (or enthalpy) occurs because the isobaric potential temperature (entropy) advection and enthalpy advection are uniquely related to each other. In Eq. (8.37a), it is the mass exchange in isentropic coordinates that is being determined from exchange processes in isobaric coordinates, and as such, the advection of potential temperature (entropy) is basic to the transformation.

In the isobaric zonally-averaged circulation, the divergence of the eddy energy transport is the primary mechanism forcing the vertical branch of the

Ferrel cell (Lorenz, 1967). In quasi-geostrophic theory for baroclinic disturbances, the rate of decrease with height of the enthalpy advection is associated with increases of geopotential energy, while enthalpy advection itself is linked with ascending motion (Holton, 1972). When viewed from an isobaric perspective, the transient potential temperature advection, which is uniquely identified with the transient modes of enthalpy exchange, plays a fundamental role in the forcing of middle latitude isobaric circulations including the Ferrel cell. However, in the perspective of the time-averaged isentropic circulation, the vertical derivative of transient geostrophic potential temperature advection within baroclinic waves is simply a component of the thermodynamically forced mode of response of the isentropic mass circulation to the planetary scale differential heating.

As emphasized earlier, the zonally-averaged mass circulation in isentropic coordinates spans each hemisphere. By partitioning into geostrophic and ageostrophic components, meridional ageostrophic circulations were isolated that included a Hadley circulation restricted to tropical-subtropical latitudes and a Ferrel circulation in middle latitudes, while a meridional geostrophic circulation was evident in middle latitudes. Taken together, the sum of the geostrophic and ageostrophic modes provide for a mass circulation that spans the hemisphere.

With regard to the isentropic and isobaric energy exchange in Eq. (8.38), the result shows mathematically that stationary and transient components are present in both. Diagnostic results summarized earlier for January and July 1979 have verified for the left hand side of Eq. (8.38) that the divergence of the transient isentropic energy transport is negligible in comparison with the divergence of the stationary energy transport (Johnson, 1985a and b). The theoretical results of Eqs. (8.17) through (8.26) that the isentropic geostrophic divergences of mass, energy and entropy are uniquely linked explain why transient isentropic energy transport is negligible in middle latitudes.

The isobaric form of the right hand side in Eq. (8.38) reveals a complicated expression with several stationary and transient components. The terms have been expanded and averaged in a form such that the isobaric degrees of freedom involving the isentropic mass transport of Eq. (8.37a) are identified by the underlined terms. The first four terms of the right hand side involve stationary and transient components of the isobaric divergence of energy transport. Of these four terms, the two involving stationary and transient components of enthalpy advection are of prime importance in middle latitudes. The last two terms involve the vertical derivative of the product of dry static energy and isobaric potential temperature advection. These are primary degrees of freedom that would account for transport of energy isentropically if one attempted to ascertain the energy transport within each branch of the isentropic Hadley circulation from information expressed in isobaric coordinates. When vertically averaged, however, differences between the isentropic and isobaric transport processes stemming from this mode vanish.

With regard to the relation between the isentropic divergence of mass and entropy transport in Eqs. (8.37a) and (8.39), a one-to-one correspondence exists between the two transport processes in isentropic coordinates. However, the expression of the isentropic divergence of entropy transport in isobaric components requires a combination of stationary and transient components of isobaric entropy advection. Again, the terms have been expanded and averaged so that the isobaric degrees of freedom involving the isentropic mass transport of Eq. (8.37a) can be identified (underlined terms). A basic difference stems from the condition that the isentropic advection of entropy is zero, while the isobaric geostrophic advection is nonzero. Although no diagnostic results will be presented, one can infer from knowledge of geostrophic temperature advection and transient modes of energy exchange in isobaric coordinates that transient modes of isobaric entropy transport are of first order importance in middle latitudes.

With regard to the overall differences of Eulerian isentropic and isobaric transport processes in middle latitudes, several points are now summarized. These are: 1) the isobaric transport of most properties will be primarily determined by modes other than by the mean mass circulation since the isobaric mass circulation of middle latitudes is not linked uniquely with the thermally forced isentropic mean mass circulation; 2) with an isentropic mean mass circulation that is linked to planetary scale heat sources and sinks, the isentropic mean mode of transport of all properties will be important; 3) in isobaric coordinates the divergence of the geostrophic transport of dry static energy is linked with stationary and transient components of enthalpy advection, while in isentropic coordinates the divergence of the transport of dry static energy is directly linked with the divergence of the geostrophic mass transport and involves exchanges of both enthalpy and geopotential energy through the thermally forced mass circulation; and 4) the divergence of the isobaric geostrophic transport of entropy includes degrees of freedom for both stationary and transient modes of isobaric exchange, while in isentropic coordinates the divergence of the entropy transport is uniquely linked to the mass transport.

9. A comparison of the global distribution of isentropic and isobaric energy transport processes

In this section, diagnostic results of planetary transport processes for both isentropic and isobaric analyses will be presented and compared for January and July, 1979. One objective is to discuss isobaric analyses of the planetary energy exchange during January and July of the GWE. Another is to show from empirical evidence that the energy transport of global monsoonal circulations is realized in isobaric coordinates through a combination of stationary and transient transport processes. For this comparison the distributions of isentropic and isobaric horizontal energy exchange are vertically integrated. In this way, the focus is placed on modes and scales of the quasi-horizontal energy transport, since differences stemming from the lack of coincidence of isentropic and isobaric layers are removed.

The vertical integration of divergence of *total flow energy* transport in generalized coordinates with a division into stationary and transient components is given by

$$\int_{\eta_{S_o}}^{\eta_T} \nabla_\eta \cdot \overline{(\rho J_\eta \underset{\sim}{U} \nu)}\, d\eta = \int_{\eta_{S_o}}^{\eta_T} \nabla_\eta \cdot [\overline{\rho J}_\eta (\underset{\sim}{\hat{U}} \hat{\nu} + \widehat{\underset{\sim}{U}^* \nu^*})] d\eta \quad , \tag{9.1}$$

where the mass-weighted temporal average isolates the mean component of energy transport by the mass circulation. In isentropic coordinates the form of the divergence of *total flow energy* transport is identical to the integrand in Eq. (5.1). However, since ρJ_p is equal to g^{-1} in hydrostatic isobaric coordinates, the divergence of energy reduces to

$$\nabla_p \cdot (\overline{\rho J_p} \widehat{\underset{\sim}{U} \nu}) = g^{-1} \nabla_p \cdot (\overline{\underset{\sim}{U}\nu} + \overline{\underset{\sim}{U}' \nu'}) \quad . \tag{9.2}$$

9.1 Comparison of vertically integrated divergence of isentropic and isobaric stationary and transient modes of energy transport

The vertically integrated fields of isentropic and isobaric divergence of *total flow energy* transport and respective potential functions are presented in Figs. 20 through 23. In each figure, panels a and b present isentropic and isobaric results for the left hand sides of Eqs. (9.1) and (9.2), respectively, while the isobaric results for stationary and transient components of energy transport are presented in panels c and d. The potential functions have been computed from defining equations which are analogous to Eqs. (2.2) and (2.4).

In the earlier results for the isentropic mean and transient components in Section 5.3, the divergence of the energy transport by the transient component was generally two orders of magnitude less than the stationary component (Johnson, 1985a and b). Consequently, the results that the energy divergences by the total and stationary components were nearly identical verified that the isentropic exchange of energy between heat sources and sinks occurs through the isentropic mass circulation. Johnson and Townsend (1981) had determined this result earlier in their analyses of NMC IIIa data.

The comparison of the vertically integrated isentropic and isobaric results for *total flow energy* transport for January and July (Figs. 20 through 23, a and b) shows the expected agreement between the two methods of analyses. All of the major extrema in one are also present in the other. The relatively small differences that occur are associated in part with the different upper boundaries that were used in the isobaric and isentropic analyses, 10 mb and 650 K respectively, and in part with the vertical interpolation used in preparing the isentropic data set from the GWE Level III data.

In accord with Eq. (5.2) and the results of Section 5.3, vertical integrated isobaric energy divergence occurs over heat source regions and energy convergence over heat sink regions. Note the primary region of isobaric energy divergence is located in the heat source region of the western Pacific. In January (Figs. 20b and 22b), this center is northeast of Australia in exactly the same location of the heating maximum of 2.5 K day^{-1} (Johnson et al., 1986). In July (Figs. 21b and 23b), this center has migrated northwestward to be located over southeast Asia with a heating maximum of 3 K day^{-1} over the southern Tibetan Plateau. In January, the isobaric analyses show energy divergence over Brazil, equatorial Africa and along the ITCZ of the Indian Ocean. Centers of energy divergence are also evident just off the east coasts of Asia and North America. Divergence in these regions is associated with the strong sensible and latent heating occurring in cyclogenetic areas over the relatively warm ocean currents. Energy convergence generally occurs over Asia in both analyses (Figs. 20a and b), although some dissimilarity is evident in western Siberia. Over Canada, the isentropic analysis indicates energy convergence, while the isobaric analysis suggests weak divergence. With the cooling rates averaging 1.5 K day^{-1} in both ECMWF and NMC analyses (Wei et al., 1983, Johnson et al., 1985a and b), the isentropic results appear more realistic. A comparison of the corresponding transport potential for these fields in Figs. 22a and b also shows the reasonable agreement between the isentropic and isobaric analyses of the planetary scale of the vertically integrated energy transport.

The partitioning of the total isobaric energy divergence shows that the stationary mode dominates in tropical latitudes; however, the transient and stationary modes are similar in magnitude in extratropical latitudes (Figs. 20 through 23, c and d). The distributions of the potential function in January also show that the importance of transient energy transport is largely confined to extratropical latitudes (Fig. 22d), while the stationary component is of a more global nature (Fig. 22c).

In the Northern Hemisphere circulation, transient energy convergence tends to occur in the major troughs of the mass streamfunction field (Figs. 3a and c), while transient energy divergence occurs over the ridges (Johnson, 1985a). Poleward of 50°N within the wave structure of the Northern Hemisphere, the pattern of stationary energy convergence tends to be out of phase with the transient energy convergence.

In the Southern Hemisphere circulation the transient energy transport is generally convergent within a meandering belt between 45° and 70°S, while the stationary mode of transport is divergent. Equatorward of this transient energy convergence is another nearly continuous belt of transient energy divergence and also stationary energy convergence. This belt stretches from South America through the South Atlantic across the Indian Ocean, Australia, and the South Pacific. Through this systematic latitudinal distribution of energy exchange within the baroclinic regime of the Southern Hemisphere, the transient component transports energy poleward, while the stationary component transports energy equatorward. In the Northern Hemisphere between 30° and 60°N, the poleward and equatorward meridional transports by the two components have the same sense; however, substantial variation of the patterns of energy exchange occurs zonally in association with the long wave structure within the wintertime circulation of this hemisphere. See Fig. 3a. In general the patterns of the transient and stationary isobaric energy divergence within the baroclinic waves of middle latitudes tend to be inverse to each other.

The tendency that the patterns of the transient and stationary isobaric energy divergence are inverse to each other stems from the fact that isobaric divergence of *total flow energy* transport by geostrophic motion is primarily restricted to the advection of enthalpy. See Section 8.3. Consequently, in order to maintain the classical ratio of geopotential to internal energy in hydrostatic atmospheres in middle latitudes, an increase of temperature in the lower troposphere of baroclinic systems through convergence of enthalpy transport requires a corresponding increase of geopotential energy within and above the level of transient enthalpy convergence (Johnson, 1970). In turn, the adjustment to hydrostatic balance forces a divergent component of geopotential energy transport from the region. This divergence of energy transport will be reflected as a stationary component of energy transport since it occurs through the isobaric mass circulation. In regions of transient energy divergence in isobaric layers of the lower troposphere, the stationary component of geopotential energy transport at upper levels will be convergent. Although this simple model is not appropriate in all cases, such as maxima of transient energy transport near the tropopause, it is consistent with quasi-geostrophic theory of extratropical cyclones and provides a simple explanation of the inverse nature of the patterns of transient and stationary energy divergence in extratropical latitudes.

This model is also basic to the isobaric theory of the forcing of the vertical branches of the Ferrel circulation of extratropical latitudes (Kuo, 1956; Lorenz, 1967). The convergence of transient energy transport in higher latitudes and divergence in lower latitudes are in effect a heat source and sink to the isobaric zonally-averaged circulation, which respectively force upward isobaric mass flux at higher latitudes and downward mass flux in subtropical latitudes. This pattern of energy divergence is associated with poleward transport of enthalpy by the transient component in the low troposphere, while in turn net geopotential energy is transported equatorward by the stationary component. The net geopotential energy transport towards the subtropics by the Ferrel circulation occurs by a combination of the equatorward mass transport in upper isobaric levels and poleward mass transport in lower isobaric levels. The patterns of energy divergence and convergence in the two meandering belts of the nearly symmetric circumpolar circulation of the Southern Hemisphere during July provides evidence of these processes.

In July (Fig. 21b), the main regions of isobaric energy divergence are located over southeast Asia and the western Pacific, Central America and the southern United States, and along a belt in the Antarctic circumpolar vortex extending from 10°W to 110°E. Isobaric energy divergence still exists in middle latitudes over the Kuroshio Current and Gulf Stream. Centers of energy convergence are located primarily over oceanic anticyclonic circulations. Note in Fig. 21b that regions of convergence occur over the Pacific and Atlantic Oceans in the Nothern Hemisphere where the Bermuda/Azores and Pacific anticyclones are well developed. In the Southern Hemisphere, regions of energy convergence are located over the eastern portion of the Pacific subtropical anticyclone in the subsiding branch of the Walker circulation, and a region east-southeast of New Zealand. According to Hill (1981), transient anticyclonic circulations dominate in the New Zealand region.

Trenberth (1985a) has noted the blocking that was prevalent in this region during the month of July, 1979. A large region of energy convergence, extending from the southern Atlantic eastward across southern Africa, the Indian Ocean and Australia, offsets a heat sink within the same region (Fig. 2c). The heat sink and associated downward diabatic mass transport within this region that includes the Mascarene anticyclone are linked through the cross equatorial flow in the low troposphere with the heat source and upward diabatic mass transport within the Asian summer monsoon. All of these isobaric features correspond well with patterns of differential heating estimated from the ECMWF and NMC data (Wei et al., 1983; Johnson and Wei, 1985; and Johnson et al., 1986).

The partitioning of the isobaric energy transport for July (Figs. 21 and 23, c and d) shows that in the Northern Hemisphere transient isobaric energy convergence is confined largely to polar latitudes, while centers of divergence are located over Asia, North America and the Aleutian Islands. Within the tropics, the stationary isobaric energy divergence is quite pronounced over major heat sources in southeast Asia and Central America. Patterns of transient isobaric energy convergence and divergence within the Southern Hemisphere circumpolar vortex are quite pronounced. Note in particular that looking downstream the transient component is divergent to the left and convergent to the right of the South Pacific and South Atlantic convergence zones. Such a pattern reflects a "down-gradient" transport of enthalpy from warmer to colder regions by transient extratropical disturbances as they propagate southeastward along these zones. Again, the potential function (Fig. 21d) shows that the importance of this component of energy transport is largely confined to extratropical latitudes, while the potential function (Fig. 21c) for the stationary component is global.

9.2 Implicit constraints in stationary and transient modes of isobaric energy exchange

Additional insight into the isobaric energy transport by stationary and transient components is gained by an analysis of energy transport by rotational and irrotational modes of the isobaric mass circulation. This analysis follows the development for isentropic coordinates in Eqs. (5.10) through (5.13). The net divergence of the isobaric transport of enthalpy and geopotential energy ψ within a "meandering ring" bounded by isopleths of the temporally averaged mass stream function $\overline{\Psi}_\rho$ and $\overline{\Psi}_\rho + \overline{\delta\Psi}_\rho$ is given by

$$\int_{\delta A} \nabla_p \cdot (\overline{\underset{\sim}{U}\nu})dA = \oint \overline{\psi}\underset{\sim}{n} \cdot \nabla_p \overline{X}_\rho d\ell + \oint \overline{\psi}\underset{\sim}{n} \cdot [\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho]d\ell$$

$$+ \oint \underset{\sim}{n} \cdot \{\overline{\psi' [\nabla_p X_\rho)' + (\underset{\sim}{k} \times \nabla_p \Psi_\rho)']}\} d\ell \quad , \tag{9.3}$$

where in accord with Eq. (2.1), the isobaric mass transport is defined by

$$\rho J_p \underset{\sim}{U} = g^{-1} \underset{\sim}{U} = \nabla_p X_\rho + \underset{\sim}{k} \times \nabla_p \Psi_\rho \quad . \tag{9.4}$$

In deriving this expression, $\overline{\rho J_p \underset{\sim}{U} f}$ has been partitioned into stationary and transient components. The divergence of the isobarically averaged transport of *total flow energy* within the "meandering ring" is associated with four components, two by the stationary component and two by the transient component. However, the net divergence of the second left hand term vanishes over the area of the stream function channel. Thus, the net divergence of energy transport is through the irrotational mode of the mean mass circulation and the transient components associated with both rotational and irrotational modes. The rotational mode of the mean isobaric mass circulation has the degree of freedom to redistribute energy within the "meandering ring" through advection of mean energy by the time-averaged rotational component of mass transport. As in the isentropic results, it is expressed by

$$\nabla_p \cdot [\overline{\psi}(\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho)] = (\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho) \cdot \nabla_p \overline{\psi} \quad . \tag{9.5}$$

With regard to the isobaric global energy balance in tropical latitudes, it is reasonable to suggest from the barotropic nature of these regions that the gradients of $\overline{\psi}$ and ψ^* tend to zero. Thus, the divergence of the energy transport in the tropics is approximated by

$$\int_{\delta A} \nabla_p \cdot (\overline{\underset{\sim}{U}\nu}) dA \simeq \oint \overline{\psi} \underset{\sim}{n} \cdot \nabla_p \overline{X}_\rho d\ell \quad , \tag{9.6}$$

a result which shows that the divergence of the transport of energy occurs primarily through the time-averaged irrotational mode of mass transport. In view of the correspondence between isobaric and isentropic modes of mass and energy transport established in Sections 8.2 and 8.4, one can conclude that isobaric mass and energy transports are directly linked with thermally forced isentropic mass and energy transports in these regions.

All four components are important to the isobaric global energy balance in middle latitudes. The quasi-geostrophic nature of extratropical circulations and the scale limitations imposed by ageostrophic mass circulations dictate that transient modes become essential for achieving global energy balance. Since geostrophic motion, as it is represented by the rotational component, can only advect enthalpy in the second and fourth right hand terms, ψ is expressed by c_pT. In this case the net divergence from the "meandering ring" is approximated by

$$\int_{\delta A} \nabla_p \cdot (\overline{U\nu}) \simeq \oint \overline{\psi}\underset{\sim}{n} \cdot \nabla_p \overline{X}_\rho d\ell + \oint c_p \overline{T}\underset{\sim}{n} \cdot \{\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho\} d\ell$$

$$+ \oint \underset{\sim}{n} \cdot [\overline{\psi'(\nabla_p X_\rho)'} + c_p \overline{T'(\underset{\sim}{k} \times \nabla_p \Psi_\rho)'}] d\ell \quad , \tag{9.7}$$

Although the net divergence of the stationary component of the rotational mode of energy transport (the second right hand integral in Eq. (9.7)) vanishes, the redistribution of energy within the isobaric "meandering ring" is given by

$$\nabla_p \cdot [\overline{\psi}\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho] \simeq (\underset{\sim}{k} \times \nabla_p \overline{\Psi}_\rho) \cdot \nabla_p c_p \overline{T} \quad . \tag{9.8}$$

This mode of energy exchange is primarily associated with geostrophic energy transport by the long stationary waves of wintertime circulations. Furthermore, this is the mode which generates the wavy structure of the stationary components in middle latitudes. Both meridional and zonal energy exchange may occur by this mode, with divergence near troughs and convergence near ridges. In Eq. (9.8) the pattern of divergence and convergence is associated with cold air advection into the base of the trough and warm air advection into the axis of the ridge, a characteristic pattern of isobaric temperature in both the stationary and transient waves. The importance of energy transport by the long waves in the Northern Hemisphere is evident in the results of many investigators. Wiin-Nielsen et al. (1963) found that the maximum poleward enthalpy transport in the Northern Hemisphere in January occurred in middle latitudes at low altitudes and that 50% of the zonally-averaged meridional transport was accomplished by wave numbers one through four. The pronounced transport by the long waves of the Northern

Hemisphere is primarily associated with distribution of continents and oceans and the global monsoonal circulations associated with these features.

The transient mode of isobaric energy exchange, which transports heat zonally and meridionally, primarily tends to occur at higher wave numbers. In Eq. (9.7), this mode is the primary means by which energy is transported poleward across the isopleths of the temporally-averaged mass stream function. As such, the patterns of the divergence of the transient component indicate a net transport of energy from warmer to colder regions, cross-stream of the large scale temperature structure of the longer stationary waves, i.e., a downgradient transport of heat. Such patterns are particularly pronounced along the baroclinic zones of the jet stream maxima found east of Asia, North America and along the South Pacific and South Atlantic Convergence Zones; regions with high frequency of occurrence of transient baroclinic circulations.

The importance of transient isobaric energy exchange in extratropical latitudes and stationary energy exchange globally is well established from empirical results over the past three decades (Hoskins, 1983; Holopainen, 1983; Oort and Peixóto, 1983). The isobaric viewpoint is that transient energy exchange in middle latitudes is forced through the baroclinic instability. Baroclinic instability is attributed primarily to the vertical shear of the geostrophic zonal motion which is created by the meridional gradient of the differential heating of the atmosphere. According to Lau and Oort (1981, 1982), this downgradient transient energy transport extracts energy from the time-mean zonally-averaged flow and the time-mean stationary waves. Likewise, Youngblut and Sasamori (1980) concluded from their analysis that the net effect of the transient eddies on the stationary waves was dissipative. It must be emphasized that these are isobaric Eulerian concepts of the maintenance of the planetary scale circulations.

The isobaric viewpoint, however, contrasts markedly with the result that the time-averaged energy balance of the planetary scale circulation is maintained through transport by the isentropic mass circulation. In view of the negligible transient isentropic energy transport and in view of the link between isobaric transient enthalphy advection and the isentropic mass circulation, one may conclude: 1) that this baroclinic mode of isobaric energy exchange is embedded within the thermally forced isentropic mean mass circulation and, 2) that the isobaric stationary and transient modes are merely a reflection of the energy transport by the isentropic mass circulation. Given the constraint that the rotational mode of the mean isobaric mass circulation cannot advect energy across its isopleths and the fact that the isobaric mass circulation is primarily ageostrophic, the primary degree of freedom for cross stream transport of energy from the heat sources of warmer latitudes to the heat sinks of higher latitudes must be the transient modes of enthalpy exchange. This process then becomes a key mode in isobaric coordinates through which the energy transport by the thermally forced isentropic mass circulation can be realized in middle latitudes and by which the energy balance of global monsoonal circulations is satisfied.

9.3 Comparison of meridional distributions of vertically, zonally integrated energy transport

Meridional profiles of vertically, zonally integrated transport of *total flow energy* and its components--mean meridional, stationary and transient

--from the ECMWF analyses for both isentropic and isobaric coordinates are presented for January and July in Fig. 24. The meridional profiles of isobaric and isentropic *total flow energy* transport are in good agreement for both January and July.

In January, meridional energy transport is divergent between 10°N and 30°S with the dividing latitude between poleward transport occurring at 10°S (Figs. 24a and c). In the stream function shown in Fig. 8a, the Hadley mass circulation for the winter hemisphere also extends to 10°S. Hence, the meridional extent of the mean mass circulation determines the dividing latitude of the poleward energy transport. Between 10°S and 30°N the energy transport is convergent in a region of net zonally averaged cooling in subtropical latitudes (Johnson et al., 1986). From 30°N to 60°N, divergence occurs in the region of the oceanic heat sources, while energy convergence occurs poleward of 60°N in association with the high latitude cooling. See Fig. 2a. In the Southern Hemisphere, the poleward transport of *total flow energy* is convergent between 10°S and 30°S and between 50°S and 60°S. It is divergent between 30°S and 50°S, and between 60°S and the pole. While a net energy source associated with baroclinic circulations over the Southern Ocean between 50°S and 60°S is implied from the meridional transport, its magnitude is much less pronounced than in the Northern Hemisphere. Compare Figs. 2a and c.

In July, during the Asian summer monsoon, the transport of *total flow energy* towards the South Pole begins at 30°N and reaches a maximum near 10°S (Figs. 24b and d). As such, the meridional extent of the heat transport from the summer to the winter hemisphere begins at higher latitudes of the summer hemisphere in July than in January. This difference is likely associated with incident solar radiation being a maximum over the Asian and North American continents of the Northern Hemisphere in July, while in January the maximum incident solar radiation is over an ocean hemisphere. Based on an analyses of the net radiation at the top of the atmosphere, Campbell and Vonder Haar (1980) determined that the oceans gain more energy than land regions, thus requiring a net annual transport of energy from the oceans to the lands. They also determined that the transport of energy should be from oceans to lands during the winter and from land to oceans during the summer. Johnson's et al. (1986) global distributions of mean heating of the surface-800 mb layer show definitely more heating over land areas in July than in January, with maxima over Asia and North America. Since there is much less ocean surface to absorb summertime incident solar energy in the Northern than in the Southern Hemisphere, less incident solar energy must be absorbed globally in the oceans in July than in January. Consequently, if one assumes that, globally the atmosphere attempts to maintain uniform energy balance throughout the annual cycle, more energy must be given to the atmosphere from the ocean in January than in July. It follows that the poleward energy transport within the oceans should be greater in January than July. Conversely, since less energy is absorbed in the oceans during July than in January, the poleward energy transport in the atmosphere of the winter hemisphere should be greater in July than in January while the poleward energy transport within the oceans should be less. Oort and Vonder Haar (1976) and Carissimo et al. (1985) have established that a substantial fraction of the poleward energy transport occurs in the oceans and that this poleward transport is maximized in the winter hemisphere. In an analysis of meridional heat transport in the world oceans, Hastenrath (1982, 1985) concludes that the annually averaged meridional transport at the equator is towards the Northern Hemisphere.

Carissimo et al. (1985) found that poleward heat transport is substantially stronger in the Northern than in the Southern Hemisphere during wintertime. These results support the rational offered. Townsend's (1980) analyses of seasonally zonally averaged dry static energy flux using NMC's Level III data shows substantially greater atmospheric poleward energy flux in the winter hemisphere during June-July-August than during December-Janauary-February, also supporting the rational offered. The analysis using ECMWF Level III data indicates that the poleward energy flux in the winter hemisphere is greater in January than in July, thus discounting the rational offered. In view of uncertainties in calculations of oceanic heat transport within the Indian and Pacific Oceans (Anderson, 1983; Bryan, 1983) as well as uncertainities in calculation of atmospheric heat transport within the Southern Hemisphere, analyses with GWE data sets can not resolve this issue.

The isentropic mean mode energy transport by the Hadley cell dominates all other components in the isentropic framework. The stationary and transient isentropic components of energy transport are negligible and lie within the level of uncertainty of the analyses (Figs. 24a and b). Within a hemisphere, the upper branch of the isentropic Hadley cell simply transports more energy poleward than the lower branch returns. While the importance of the poleward energy transport by the isentropic Hadley circulation had been previously determined by Zillman (1972), Townsend (1980) established from an analysis of NMC's Level IIIa data that the zonally averaged isentropic transient energy transport was negligible.

The isobaric results (Figs. 24c and d) are in marked contrast with the isentropic results. The mean mode transport in isobaric coordinates is dominant in lower latitudes, while all three components become important in middle latitudes. Equatorward of the maximum of isobaric transient energy transport, the transient component is divergent, while poleward it is convergent. At the same time, the mean meridional component tends to be convergent equatorward and divergent poleward of the maximum in the transient meridional flux. As noted earlier, this distribution of zonally averaged convergence and divergence of transient poleward enthalpy transport through geostrophy forces the vertical branches of the indirect ageostrophic Ferrel circulations. The forced indirect Ferrel circulations in turn, transports geopotential energy towards lower latitudes (Lorenz, 1967). This pattern of isobaric meridional energy exchange was evident in the January and July distributions of the transient and stationary components for the Southern Hemisphere (Figs. 13 and 14, c and d).

For all practical purposes the time-averaged isentropic mass circulation is the only mode by which the isentropic energy exchange occurs. Since fields of the vertically integrated isentropic and isobaric transport are identical, within the accuracy of these calculations for the GWE one must conclude: 1) that the isobaric components of energy transport are in themselves components of the transport of energy by the thermally forced isentropic mass circulation, and 2) that standing eddy and transient components of isobaric energy transport associated with baroclinic circulations are not independent, but are constrained by the mean isentropic energy transport. Thus, the baroclinic instability process becomes simply a means by which the atmosphere transports energy at the planetary scale in middle latitudes, where the stability of the atmosphere due to rotation precludes zonally symmetric meridional circulations. Within the constraint of energy transport by the isentropic mass circulation, the relative importance of the isobaric components is a reflection of adjustment processes through which the atmosphere finds the means to transport energy from heat source to heat sink.

As such, this thermally forced mass isentropic circulation must play a fundamental role in the isobaric energy exchange in the global circulation.

In view of the importance of isobaric transient enthalpy transport in extratropical latitudes and the restricted length scale of the isobaric mass circulations associated with transient cyclones and anticyclones, it is not surprising that the isobaric energy exchange in middle latitudes has not been directly linked to global monsoonal circulations. This apparent lack of an explicit link to differential heating illustrates a confounding inherent in isobaric statistics of transport processes and in physical inference based on isobaric theory. This confounding stems from the combination of geostrophy, the selection of a coordinate system in which the geostrophic exchange of mass, dry static energy and entropy bear little relation with each other. Compare Eqs. (8.21) through (8.26) and the results discussed in Section 8.3.

D. ZONALLY AVERAGED CIRCULATIONS

10. The structure and forcing of isentropic mean meridional circulations for January and July

10.1 Background

The primary focus of this summary up to this point has been on the thermal forcing of mass, energy and entropy transport in global monsoonal circulations with little regard to the relation of these processes to momentum exchange. It was implicit in the initial premise that the atmosphere would develop modes of momentum exchange through which isentropic mass circulations would arise and satisfy energy balance requirements in relation to differential heating. The diagnostic results presented heretofore provide evidence that isentropic mass circulations exist which transport energy and entropy from heat source to heat sink. The occurence of an isentropic mass circulation and energy and entropy transport with scales common to the scale of differential heating dictate that a corresponding scale of momentum exchange also exists. The momentum exchange forces the quasi-horizontal branches of the isentropic mass circulation.

At this point, this supposition cannot be elucidated from analyses of the three-dimensional momentum balance in isentropic coordinates, since a comprehensive study of the isentropic momentum exchange has not been completed as of this date. Studies of the relationship between diabatic heating and absolute angular momentum exchange in forcing the isentropic zonally averaged mass circulation (Townsend, 1980; Gallimore and Johnson, 1981a and b; Johnson, 1983a and b, 1985a) and an analyses of the time-averaged global angular momentum exchange have been completed (Schaack, 1982; Johnson et al., 1982). Recently, an analysis of the planetary exchange of isentropic dynamic vorticity defined as $\underset{\sim}{k} \cdot (\nabla_\theta \times \rho J_\theta \underset{\sim}{U})$ was completed by Hoerling (1987). Preliminary results of this analysis were presented by Hoerling and Johnson (1986).

All of these results are consistent with the premise that the atmosphere develops modes of momentum exchange through which isentropic mass circulations exist and transport sufficient energy and entropy to satisfy balance requirements in relation to differential heating.

The purpose of this section is to provide a perspective for the forcing of isentropic Hadley circulations (Gallimore and Johnson, 1981a) using diagnostic results for January and July of the GWE year (Townsend, 1980; Townsend and Johnson, 1981; Johnson 1983a and b, 1985a). The results will illustrate how torque about the earth's axis of rotation forces quasi-horizontal branches of the isentropic Hadley circulation. The results will also show how the three components of torque combine to force momentum exchange on a scale of hemispheric forcing that is equal to the hemispheric scale of differential heating. Thus, the scales of the differential heating, the isentropic Hadley mass circulation, the mean mode of meridional energy transport and the zonally averaged angular momentum torque are all hemispheric in meridional extent, hence enjoying a common scale.

10.2 The balance of the zonally averaged transport of mass and angular momentum

For the steady time-averaged isentropic circulation, the zonally averaged transport equations for mass and angular momentum (Gallimore and Johnson, 1981a) are expressed by

$$\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi_\theta} \overline{(\overline{\rho J_\theta} \hat{v} \cos \phi)}^t + \frac{\partial}{\partial \theta} \overline{(\overline{\rho J_\theta} \hat{\dot{\theta}})}^t = 0 \quad , \qquad (10.1)$$

$$\frac{1}{a \cos \phi} \frac{\partial}{\partial \phi_\theta} \overline{(\overline{\rho J_\theta} \hat{v} \cos \phi)}^t + \frac{\partial}{\partial \theta} \overline{[\overline{\rho J_\theta} \hat{\dot{\theta}} \hat{g}_a]}^t = \overline{[\overline{\rho J_\theta} \Sigma \hat{T}_i]}^t \quad , \qquad (10.2)$$

where in this chapter the overbar and caret without an accompanying superscript indicate zonal averaging. The overbar with the superscript t denotes a time average.

The zonal average of the component of the absolute angular momentum about the earth axis of rotation is given by

$$\hat{g}_a = (\underset{\sim}{\Omega}/|\underset{\sim}{\Omega}|)\cdot\underset{\sim}{r} \times \hat{\underset{\sim}{U}}_a \quad , \tag{10.3a}$$

$$= g_e + g_r \quad , \tag{10.3b}$$

$$= \Omega a^2 \cos^2\phi + \hat{U}\, a \cos\phi \quad , \tag{10.3c}$$

where the definitions of the earth (g_e) and relative (g_r) components of angular momentum are evident. The sum of the torques is expressed by

$$\Sigma \hat{T}_i = \hat{T}_p + \hat{T}_F + \hat{T}_I \quad , \tag{10.4}$$

where the pressure, friction and inertial torques are defined by

$$\hat{T}_p = - \overline{\rho J_\theta \partial\psi/\partial\lambda_\theta}/\overline{\rho J_\theta} \quad , \tag{10.5a}$$

$$= - a \cos\phi \; f \; \hat{v}_g \quad , \tag{10.5b}$$

$$\hat{T}_F = a \cos\phi \; \overline{\rho J_\theta F_\lambda}/\overline{\rho J_\theta} \quad , \tag{10.6}$$

$$\hat{T}_I = - (\overline{\rho J_\theta})^{-1} \left[\frac{1}{a \cos\phi}\frac{\partial}{\partial\phi_\theta}(\overline{\rho J_\theta}\,\widehat{u^* v^*}\,a \cos^2\phi) + \frac{\partial}{\partial\theta}(\overline{\rho J_\theta}\,\widehat{u^* \dot{\theta}^*}\,a \cos\phi)\right] . \tag{10.7}$$

In the zonally averaged equations, the convergence of the horizontal and vertical eddy momentum transport becomes in effect a source/sink function of angular momentum. For brevity, this process is called an inertial torque.

Through an integration of the mass continuity equation over the meridional domain from a latitude ϕ_1 to the pole of the Northern Hemisphere and elimination of the zonally averaged mass continuity equation from the transport equation for absolute angular momentum, the mean meridional transport from Eqs. (10.1) and (10.2) is expressed by

$$\overline{\rho J_\theta v}^{t\lambda}(\phi_1) = (a \cos \phi_1)^{-1} \int_{\phi_1}^{\pi/2} \frac{\partial}{\partial \theta} (\overline{\rho J_\theta \dot{\theta}}^{t\lambda}) a^2 \cos \phi \, d\phi \,, \tag{10.8}$$

$$\overline{\rho J_\theta v}^{t\lambda}(\phi_1) = - \overline{[\overline{\rho J_\theta}(\Sigma \hat{T}_i - \hat{\dot{\theta}} \, \partial \hat{g}_a/\partial \theta)/(a \cos \phi \; Z_a)]}^{t} (\phi_1) \quad . \tag{10.9}$$

The relation between the meridional gradient of absolute angular momentum and absolute vorticity of the mean zonal motion is

$$\frac{1}{a}\frac{\partial \hat{g}_a}{\partial \phi_\theta} = \frac{1}{a}\frac{\partial}{\partial \phi_\theta} [\Omega \; a^2 \cos^2 \phi + \hat{u} \; a \cos \phi] \,, \tag{10.10a}$$

$$= - a \cos \phi \; Z_{a_\theta} \quad , \tag{10.10b}$$

where

$$Z_{a_\theta} = f - \frac{1}{a^2 \cos \phi} \frac{\partial}{\partial \phi_\theta} (\hat{u} \; a \cos \phi) \; . \tag{10.11}$$

With the potential vorticity of the zonally averaged motion expressed by

$$P_\theta = Z_{a_\theta} / \overline{\rho J_\theta} \quad , \tag{10.12}$$

the temporally zonally averaged mass transport is expressed by

$$\overline{\rho J_\theta v}^{t\lambda} \; (\phi_1), = - \overline{[(\Sigma \hat{T}_i - \hat{\dot{\theta}} \frac{\partial \hat{g}_a}{\partial \theta})/(a \cos \phi \; P_\theta)]}^{t}(\phi_1) \; . \tag{10.13}$$

In a hydrodynamically stable vortex with positive potential vorticity, the isentropic meridional mass transport in the balance expressed by Eqs. (10.8) and (10.9) requires differential heating and torques. Vertical branches of the mass circulation occur through entropy sources in tropical latitudes and entropy sinks in polar latitudes. Horizontal branches of the mass circulation occur through torques. A negative (positive) torque in effect makes the zonally averaged motion subgradient (supergradient), thereby requiring poleward (equatorward) advection of absolute angular momentum [equivalently a poleward (equatorward) transport of absolute vorticity in Eq. (10.13)]. If the absolute vorticity vanishes, the joint requirements for differential heating and torques are decoupled.

10.3 The forcing of the isentropic Hadley mass circulation

Following Eliassen's (1951, 1960) and Kuo's (1956, 1957) work on the forcing of mean meridional circulations, Gallimore and Johnson (1977a, 1981a) established a theory for the forcing of zonally-averaged isentropic mass circulation. In the zonally averaged circulation, mass transport is forced towards (from) an axis of rotation by negative (positive) torques, while forcing of the upward (downward) branch of the mass circulation is by heating (cooling). The zonally averaged torques in isentropic coordinates are pressure and frictional torques in combination with the inertial torque associated with convergence of eddy angular momentum transport. There is no degree of freedom for the forcing of the vertical branches by isentropic eddy energy convergence.

In isobaric coordinates, the frictional torque, diabatic heating, and the convergences of eddy angular momentum and eddy sensible heat transport constitute the forcing of the mean mass circulation (Kuo, 1956; Lorenz, 1967). Because the mean zonal pressure gradient force vanishes, there is no forcing of the meridional mass transport by pressure torques in the Eulerian isobaric perspective (Johnson and Downey, 1975a and b). Thus, there can be no direct link with the isentropic geostrophic mode of the mass circulation.

10.4 The isentropic relative angular momentum distribution

The isentropic time-averaged zonally integrated relative angular momentum distributions for January and July are presented in Figs. 25b and 26b. For ready reference in this and the following sections, the temporally, zonally averaged mass circulations for January and July, which were discussed in detail in Section 4.3, are presented in Figs. 25a and 26a.

A comparison of the mean zonally integrated relative angular momentum distributions between the hemispheres shows marked differences. In January, the westerlies in Fig. 25b are well developed in both hemispheres, the wintertime maximum of the zonally integrated relative angular momentum for the Northern Hemisphere being one and one-half times the summertime maximum for the Southern Hemisphere. In July, the westerlies in Fig. 26b are well developed in the Southern Hemisphere, but not in the Northern Hemisphere. The wintertime maximum of angular momentum for the Southern Hemisphere is three times the summertime maximum of the Northern Hemisphere. In conjunction with the mimima of relative angular momentum of the Northern Hemisphere in July, a thermally forced isentropic Hadley cell in the Northern Hemisphere is not readily apparent at this time. As such, the primary circulation which maintains the westerlies of the Northern Hemisphere against frictional torques is suppressed.

Another difference between the two months is the limited depth of the tropical-subtropical easterlies during January in contrast to the unlimited depth during July. The easterlies in the upper troposphere during July reflect the existence of the Tropical Easterly Jet of the Asian-African sector that is associated with the Asian summer monsoon.

10.5 The forcing of the isentropic Hadley mass circulation by torque

The combined forcing of the horizontal branches of the mean meridional circulation by pressure and frictional torques and the convergence of the horizontal and vertical eddy angular momentum transport are presented in Figs. 25c and 26c. The poleward branch of an isentropic Hadley circulation in the upper layers is forced by negative torque, while an equatorward branch is forced by positive torque. The one-to-one correspondence between the scale and the sense of the horizontal branches of the mean mass circulation and the distribution of torques is evident. Regions of negative and positive torques are respectively coincident with poleward and equatorward motion of the Hadley circulation. This is remarkable in view of the fact that: 1) torques and the mass transport are estimated from atmospheric processes that are distinctly different from each other and, 2) the zonally integrated frictional torque is estimated by residual methods (Schaack, 1982). Such estimates traditionally involve a large degree of uncertainity in the determination of momentum balance.

A comparison of the winter and summer hemispheric Hadley circulations and torques in January (Figs. 25a and c) show a difference in intensity that is consistent with the greater differential heating in a winter hemisphere. The region of positive torque within the lower layers of the Northern Hemisphere has two maxima, which are consistent with the forcing required to maintain the two embedded isentropic circulations.

A comparison of the winter and summer Hadley circulations and torques in July (Figs. 26a and c) show that the mass circulation and forcing in the winter Southern Hemisphere are intense, while circulation and forcing in the summer Northern Hemisphere are lacking. The mean mass circulation that exists in the Northern Hemisphere is from the extension of the winter Hadley circulation into the summer hemisphere. At the equator, the distribution of torques reverses in sign in each of the horizontal branches of the Hadley circulation, thus consistently accounting for the forcing of the extension of the Southern Hemisphere Hadley circulation into the Northern Hemisphere. Between 0° and 30°N, negative torques in lower isentropic layers force motion towards the latitudinal belt of the ascending branches of the summer monsoonal circulations over the Asian and North American continents, while positive torques in upper isentropic layers, although weak, force motion towards the equator. The apparent inconsistency of the positive torque in the lowest layers between 0° and 30°N and the overlying negative torque above is accounted for by the differences between the isentropic layers of poleward mass transport over the Indian Ocean and equatorward mass transport over the Atlantic and Pacific Oceans. In the lowest layers below 300 K, the mass transport within the trades of the Atlantic and Pacific Oceans is forced equatorward by negative torque, while between 300 and 315 K positive torque forces mass transport in the subtropical regions of the Indian Ocean towards the Asiatic continent during its summer monsoon. The separation of the positive and negative torques into two distinct isentropic layers is due primarily to the warmer subtropical mixed layer over the Indian Ocean relative to the trades over the Atlantic and Pacific Oceans.

From hydrodynamic stability considerations, Eliassen pointed out that for a given forcing the response of the circulation will be strong in regions of weak hydrodynamic stability and weak in regions of strong stability. Thus the relatively weak response of the meridional circulation to atmospheric forcing in extratropical latitudes with strong stability and the relatively strong response in low latitudes with weak stability are also in agreement with Eliassen's conclusions. See Figs. 25 and 26, a and c.

The relative importance of the pressure torque, the frictional torque and the inertial torque associated with the horizontal and vertical eddy convergence of angular momentum transport are evident from comparison of the results in panels d, e and f of Figs. 25 and 26. The distributions of pressure torque in Figs. 25e and 26e indicate pronounced forcing within the active baroclinic wave regime of the winter hemispheres. The distributions in the summer hemispheres show the pressure torque to be well organized in the Southern Hemisphere, but virtually nonexistent in the Northern Hemisphere. In extratropical latitudes, the pressure torque dominates other torques and forces the poleward branch in upper isentropic layers and the equatorward branch in lower isentropic layers.

Johnson (1979, 1980) and Gallimore and Johnson (1981a) emphasize that over level terrain, isentropic pressure torques transfer angular momentum from the poleward branch to the equatorward branch of amplifying baroclinic waves, while the vertically integrated pressure torque vanishes. The geostrophic mode of mass transport when multiplied by the Coriolis parameter and a cos ϕ becomes the pressure torque, which is proportional to the vertical derivative of the non-convective flux of angular momentum by pressure stresses (Johnson and Downey 1975b; Johnson, 1980). Thus, in the idealized baroclinic wave in Fig. 10, angular momentum is transferred from the upper to the lower isentropic layer by the pressure stress across the isentropic surface θ_m that divides the upper and the lower branches of the isentropic Hadley circulation. This angular momentum sink in the upper isentropic layer forces the net poleward mass transport, while the source in the lower layers forces net equatorward mass transport.

In the schematic of the amplifying baroclinic wave (Fig. 10b), the transfer of absolute angular momentum from the poleward to the equatorward branch constitutes a longitudinal transfer of angular momentum from the warmer subtropical air in the forward part of the wave to the colder polar air in the rear part of the wave. This combination also forces net meridional transport of energy poleward. The poleward branch of the geostrophic mass transport with greater enthalpy is warm relative to the equatorward branch, while the meridional geostrophic geopotential energy transport of the poleward and equatorward branches tends to be equal.

In this manner, the isentropic geostrophic mass circulation, the forcing of the mass circulation by pressure torques and the poleward energy transport within amplifying baroclinic waves become directly coupled to the meridional distribution of differential heating. This direct link of the mass, momentum and energy exchange with the differential heating constitutes the basis for claiming that the isentropic Hadley circulations are thermally forced and constitute a direct response to differential heating. The supposition that the atmosphere develops systematic sources and sinks of momentum to force the horizontal branches of the temporally averaged mass circulation with a scale that is common to the thermal forcing seems well supported. If such mean mass circulations exist, the means to force these circulations against dissipative processes must also exist.

In the equatorward branch of the Hadley circulation in tropical latitudes, the frictional torques are negative and dominate other torques in forcing the mass transport towards the tropical latitudes. The negative frictional torque in the core of the maximum westerlies enhances the poleward mass transport of upper isentropic layers, but opposes the forcing by the pressure torque in the equatorward mass transport of lower isentropic layers.

As noted previously, the frictional torque (Figs. 25d and 26d) was estimated by residual methods from a consistent analyses of both the absolute and relative angular momentum balance (Schaack, 1982). Such methods of analyses involve uncertainity. In the meridional distributions in Figs. 25d and 26d, the frictional torque does oppose both the westerlies and easterlies in the zonally-averaged circumpolar circulation. For the most part, larger frictional torque occurs in lower isentropic layers in the planetary boundary layer. However, in the extratropical latitudes of the Northern Hemisphere during winter, the negative frictional torque extends throughout the vertical extent of the westerlies. By virtue of the large zonal temperature contrast ranging from 240 to 290 K in the boundary layers over the extremely cold continents and relatively warm oceans in extratropical latitudes of the Northern Hemisphere, the isentropic zonally averaged frictional torque extends vertically. With the more uniform temperatures of the oceanic boundary layer of the Southern Hemisphere, the frictional torque there remains more vertically confined.

Overall, these features are in accord with expectations and suggest a consistency in estimating angular momentum balance that does not seem possible prior to the GWE. The suggestion that these results may simply reveal the dynamics of the assimilation model cannot be ruled out. However, the large and consistent differences that are evident between the January and July results and the consistency between the mass circulation and the distribution of forcing by total torque indicate that, at least qualitatively, the results are valid.

A comparison of the magnitude and distribution of the inertial torques indicates that this component is generally of lesser importance than pressure and frictional torques in the forcing of the isentropic Hadley mass circulation. An exception occurs in the upper isentropic layers of the subtropics, where the negative torque that occurs equatorward of the relative angular momentum maximum constitutes the primary forcing of the poleward mass transport. The convergence poleward of the relative angular momentum maxima forces the isentropic mass transport equatorward but is secondary in importance to the pressure torque. Through the distribution of inertial torque, relative angular momentum is transported poleward across the maximum of relative angular momentum. In the vertically zonally integrated balance, the westerlies that lie poleward of this relative angular momentum maximum are maintained by this process.

The result that the divergence of the eddy angular momentum transport is of secondary importance in the forcing of the quasi-horizontal branches of the isentropic zonally-averaged circulation is in marked contrast to its relative importance in isobaric coordinates. In isobaric coordinates, this torque being positive in the upper branch of the Ferrel circulation is primary to the forcing of the equatorward mass transport. It is also primary to the forcing of the poleward branch of the Hadley circulation, being negative in subtropical latitudes. With the maximum of relative angular momentum transport occuring near 30°, the poleward extent of the isobaric Hadley circulation is restricted to subtropical latitudes.

The contradiction between the relative importance of eddy angular momentum transport in isentropic and isobaric coordinates stems from: 1) differences in the geometric position of a zonal tube within which the transport processes are zonally integrated and, 2) the allowable modes of zonally averaged mass transport and pressure torque, the net of the latter being zero in isobaric coordinates (Johnson and Downey, 1975a). As the earlier results in Figs. 8 and 9 revealed, a zonal-average of the mass transport in isentropic coordinates includes a systematic component of meridional exchange through the geostrophic modes which, in effect, determines the first order importance of the net pressure torques.

A similarity of the scale of isentropic and isobaric ageostrophic motion is suggested from the results for zonally averaged circulations. As noted earlier, the results in Figs. 8 and 9 establish an equivalent meridional scale for the isentropic and isobaric ageostrophic mass circulations. The basis of this equivalence stems from the coordinate independence of the definition of an ageostrophic component of meridional velocity and the lack of a systematic zonal covariance between ρJ_θ and v_{ag}. Consequently, the isentropic and isobaric zonally averaged ageostrophic circulations may be similar, provided that both coordinate representations enjoy a common scale for the forcing of v_{ag}. The scale of the isentropic ageostrophic mass circulation (Figs. 8d and 9d) is determined by the scale of the imbalance between the thermally forced advection of zonally averaged relative angular momentum, pressure torques and the divergence of eddy angular momentum transport in the upper troposphere. In isobaric coordinates the scale of the ageostrophic mass circulation is determined primarily by the eddy exchange of angular momentum and sensible heat.

10.6 On the relation between the isentropic Hadley mass circulation and sloped convection in a rotating fluid

The key to the existence of the hemispheric scale of the torque distribution is the geostrophic component of mass transport and the net nonconvective zonal transfer of angular momentum from the poleward branch of the isentropic Hadley circulation to its equatorward branch in amplifying baroclinic waves (Johnson and Downey, 1975a, 1975b). This transfer of angular momentum not only forces the poleward and equatorward branches of the isentropic mass circulation, but also forces the net meridional energy transport to be poleward. The net poleward energy transport occurs because the dry static energy of the poleward branch excedes the energy of the equatorward branch. The resulting systematic divergence of energy in tropical latitudes and convergence in low latitudes fulfills the zonally averaged energy balance required by differential heating.

An analogy between the geostrophic mass circulation and forcing of the isentropic Hadley circulation by pressure torques with concepts of sloping convection in a rotating fluid (Hide and Mason, 1975) is evident. Such a perspective is also implicit in Ludlam's (1966) views of meridional isentropic exchange in baroclinic waves and in Palmen and Newton's (1969) schematic of the mass exchange in extratropical cyclones. In the perspective of isentropic exchange in middle latitudes, the equatorward mass transport of the lower isentropic layers in the rear of a pressure trough occurs through the motion of polar air masses from high to low latitudes, while the poleward mass transport forward of the pressure trough occurs through motion of subtropical air masses from low to high latitudes. By virtue of the contiguous nature of air masses, the polar outbreaks constitute a sloped convective plume within a shallow atmosphere extending from the heat sink region of polar latitudes to the heat source region of subtropical latitudes, while the subtropical air constitutes a plume extending from subtropical to polar latitudes.

In the topography of the 313 K isentropic surface (Fig. 1), the equatorward moving plumes of polar air lie beneath the protrouding ridges that extend towards lower latitudes, while the poleward moving plumes of subtropical air lie within the valleys that extend towards the pole. As the polar air is forced towards subtropical latitudes by positive torques, it gains energy at the earth's surface and moves to higher isentropic layers. There it eventually becomes subtropical air, part of which moves towards the ITCZ and eventually returns to higher latitudes within the upper branch of the Hadley circulation and part of which quickly returns towards higher latitudes within the warm sectors of middle latitude cyclones. The subtropical air, which is forced towards polar latitudes by negative pressure torques, moves to lower isentropic layers through cooling and eventually becomes polar air. Through this combination of processes, an isentropic Hadley circulation is realized which spans the hemisphere.

Empirical evidence for the structure of the poleward and equatorward branches of isentropic geostrophic mass transport in middle latitude baroclinic waves is provided by three cross sections. One (Fig. 27) is an east-west vertical cross section from Tuscon to Bermuda through a large amplitude wave studied extensively by Newton and Palmén (1963). The other two (Figs. 28a and b) are zonal-vertical cross sections of the isentropic and isobaric time-averaged meridional velocities along 50°N for the winter season (Townsend, 1980).

In the cross section through the large amplitude wave, Newton and Palmén's analysis has been modified by rescaling the vertical coordinate to be linear in pressure and the meridional component of the wind has been calculated and analyzed. Since the vertical distance between isentropes is proportional to mass, the incremental area between adjacent isopleths of the meridional wind and the potential temperature when multiplied by the intensity of the meridional wind estimates the isentropic meridional mass transport. The layer extending from the earth's surface to 315 K contains more mass in regions of equatorward motion than in regions of poleward motion. Thus, strong equatorward mass transport of cold polar air occurs below 315 K. In the 320 K to 335 K layer, more mass is located in the cross-hatched region of poleward motion (which is also more intense) than in the dotted region of equatorward motion, thus reflecting a mean poleward transport of subtropical air. The slope of the zero isotach indicates a westward tilt of the pressure wave with height in conjunction with the colder air to the rear of the surface pressure trough. While this is the configuration of an intensifying baroclinic wave with a geostrophic mode of mass and energy transport, this long wave is in an advanced stage of development and the covariance of mass ρJ_θ and the meridional motion is not particularly pronounced. A wave structure which is much more similar to the structure in the schematic of the amplifying wave in Fig. 10b is shown in Danielsen's cross section through a baroclinic wave over the western United States for 1200 GMT, 13 April 1972 (see Danielsen's Figs. 1 and 2, 1974).

This structure for systematic isentropic meridional mass transport is also evident in Townsend's (1980) time-averaged cross section for the winter season of the FGGE year (Figs. 28a and b). The isobaric time-averaged meridional wind $\overline{v}$ and potential temperature distributions are presented in panel a, while the isentropic time-averaged meridional wind $\hat{v}$ and pressure distributions are presented in panel b.

A comparison of these two cross sections reveals some interesting differences. The primary regions of isobaric meridional exchange occur between 500 and 100 mb. In the isentropic cross section, the larger values of meridional mass transport occur between the time-averaged isobars of 800 and 400 mb. This displacement downward of the isentropic meridional mass transport relative to the isobaric meridional mass transport is associated with the positive correlation of ρJ_θ and v in the atmosphere. In isobaric coordinates, this correlation exists in the form of low static stability associated with poleward mass transport and high static stability associated with equatorward mass transport. Such a correlation is particularly evident in all three cross sections. The systematic westward tilt of the zero isopleth of meridional mass transport is also a reflection of the systematic exchange of polar and subtropical air masses associated with the systematic correlation of colder and more stable air in equatorward motion and warmer and less stable air in poleward motion.

While the structure in the isobaric cross section does not permit zonally averaged geostrophic meridional motion such as displayed in the isentropic cross section, the structure does permit zonally averaged poleward transport of energy by the time-averaged stationary waves (Wiin-Nielsen et al., 1963). Global diagnostic results reveal that the transient and stationary waves share equally in the energy flux in the Northern Hemisphere, but transient waves completely dominate the energy flux in the Southern Hemisphere (See Fig. 1.5 of James (1983)). In contrast, the net isentropic meridional transport of energy occurs through the mean mass circulation, which is dominated by the geostrophic component at these latitudes. It should be emphasized again that the geostrophic mass and energy transport is forced by the transfer of absolute angular momentum by pressure stresses in amplifying baroclinic waves; thereby polar air masses gain angular momentum at the expense of subtropical air masses. In the isentropic Hadley circulation, forcing through angular momentum exchange is simply a manifestation of the atmospheric response needed for hemisphere scales of mass, energy and entropy transport that are common with the scale of the zonally averaged differential heating.

In regard to global monsoonal circulations, note that the equatorward mass transport in the longitudinal structure of Figs. 28a and b, primarily occurs in the sectors from 90E to 160E and 120W to 60W. These are the sectors of equatorward monsoonal flow of cold polar air that occur on the average over the Asian and North American continents as well as the oceanic regions immediately to the east of these continents. The poleward transport of subtropical air occurs in the longitudinal sectors primarily over the eastern halves of the North Pacific and North Atlantic Oceans. Such systematic mass transport is consistent with the relatively cold polar climates of the east coasts of North America and Eurasia and the relatively warm maritime climates of the west coasts of North America and Europe.

In contrasting these results from isobaric and isentropic cross sections and interperting differences, it is important to realize that the fields in Fig. 28 represent an average structure of transport processes by a spectrum of baroclinic features propagating within the westerlies. As baroclinic features propagate, the equatorward mass transport of polar air on the average is more intense in the sectors over the eastern halves of continents and adjacent oceans, while the poleward transport of subtropical air is more intense over the eastern halves of the Pacific and Atlantic Oceans. As episodes of baroclinic amplification occur, time-averaging in isentropic coordinates includes the increased intensity of the mass and energy transport within the stationary component of the circulation, while within isobaric coordinates the geostrophy of the circulations precludes the complete isolation of the enhanced energy transport in the stationary component. Thus the zonally averaged isobaric meridional energy transport occurs through a combination of stationary and transient components in the Northern Hemisphere.

On the matter of isentropic energy transport by the stationary component, it was pointed out earlier in Section 9.2 that the zonally averaged meridional transport of energy by the stationary rotational component of the mass transport must occur through divergence of energy transport within the troughs and convergence within the ridges of the time-averaged mass stream function. However, there can be no net divergence of energy transport from an incremental channel bounded by $\overline{\Psi}_\rho$ and $\overline{\Psi}_\rho + \Delta\overline{\Psi}_\rho$ through the time-averaged rotational component of mass transport (see Section 5.4). Thus the meridional extent of the net poleward energy transport by the stationary rotational mode within a given incremental channel of the time-averaged mass stream function is restricted to the latitudinal band in which the channel is located. Zonally there is no restriction.

10.7 On the origin of the easterlies in subtropical latitudes

In his monograph on the general circulation, Lorenz (1967) drew attention to the lack of a complete understanding of the causes of the trade winds by the quote from Hadley (1735), "I think the causes of the General Trade-Winds have not been fully explained by any of those who wrote on that Subject...".

In attempting to resolve this problem using the isobaric results for the zonally averaged circulation, one notes that the systematic downward mass transport of the isobaric Hadley mass circulation must induce westerlies by transporting the eastward relative angular momentum of the upper troposphere to the lower troposphere. Hence, for the maintenance of easterlies against this systematic downward transport of eastward momentum by the isobaric Hadley mass circulation, the sink of angular momentum at these latitudes through the isobaric divergence of eddy angular momentum transport must exceed the source by downward transport. Consequently, the maintenance of the easterlies against friction is directly dependent on the poleward eddy angular momentum transport by baroclinic waves in the westerlies.

Without a means to directly link the eddy angular momentum exchange to systematic processes in the zonally averaged circulation, the secret for the maintenance of the easterlies appears to reside within the eddies themselves. In his concluding chapter, Lorenz stated "We regard the problem of explaining the pattern of the transport of angular momentum by the eddies as the most important problem in general circulation theory among those for which we now lack a fairly adequate qualitative explanation". It may be argued by some that partitioning of atmosphere transport processes into isobaric temporal averages and eddy components in lieu of zonal averages and eddy components helps to resolve this problem. While additional insight is gained on the global distribution of baroclinic processes, the problem still remains in that an isobaric zonal average of the time-averaged stationary and transient eddy transport processes will leave one with the same problem that Lorenz (1967) discussed. Now the question of the maintenance of the easterlies will be addressed after which the maintenance of the westerlies will be discussed in the following subsection.

In the lower isentropic layers of extratropical latitudes, frictional torque opposes the westerlies and forces mass circulation poleward; however, pressure torque in the lower layers dominates the frictional torque, with the result that the mass circulation is equatorward in this region. In the subtropical latitudes, where the net pressure torque of baroclinic waves in the lower isentropic layers is still positive but weak (Fig. 25e), the frictional torque becomes positive and dominant (Fig. 25c). Since both torques are now positive they force the lower branch of the Hadley circulation equatorward. However, since both torques oppose the easterlies, by themselves these forces do not create easterlies by direct means through acceleration of the flow. Thus, the easterlies must be realized through the transformation of the relative to the earth component of absolute angular momentum during equatorward flow, i.e., the Coriolis torque. It is important, however, to recognize that before the frictional torque is able to force the lower branch of the mass circulation equatorward and thereby create easterlies from the Coriolis torque, some mechanism must exist to establish easterlies in the equatorward branch of the isentropic mass circulation. At the latitude of zero relative angular momentum in the lower branch of the Hadley circulation the frictional torque vanishes; therefore this torque does not force the equatorward mass transport at this latitude.

Inspection of the total torque distribution in Figs. 25 and 26 reveals that in the isentropic layers immediately above the earth's surface the total torque remains positive at all latitudes and forces equatorward mass transport, even at latitude near 30° where the relative angular momentum vanishes. At this latitude, the positive torque stems from the equatorward geostrophic mode of mass transport and the transfer of angular momentum by pressure stresses from the upper to lower isentropic layers through amplifying baroclinic waves. Also, within this latitudinal belt that separates ageostrophic Hadley and Ferrel circulations, the ageostrophic component of the isentropic mass circulation vanishes. Consequently, without ageostrophic mass transport the equality of the actual and geostrophic mode of mass transport at this latitude is directly and uniquely linked to the thermally forced isentropic mass circulation, which through the pressure torque becomes the sole means to force an equatorward transport of absolute angular momentum from the extratropics. This balance is readily noted in Eqs. (10.9) and (10.10). Under the conditions of negligible frictional torque, inertial torque and mean vertical advection of mean angular momentum, the zonally averaged angular momentum balance as expressed by Eq. (10.10) simplifies to

$$\overline{\rho J_\theta v}^{t\lambda} = - \overline{\overline{\rho J_\theta}\, \hat{T}_p / a \cos\phi\, f}^{t} , \tag{10.16a}$$

$$= \overline{\rho J v_g}^{t\lambda} \quad . \tag{10.16b}$$

The absolute vorticity has been set equal to the Coriolis parameter, since the vorticity of the mean zonal motion is minimal in comparison to f in the lower branch of the Hadley circulation. This angular momentum balance stems from the mutually consistent conditions in the latitudinal belt separating the ageostrophic Hadley and Ferrel circulations that: 1) the mean meridional ageostrophic mass transport vanishes and 2) the frictional torque, inertial torque and mean diabatic advection of zonally averaged angular momentum are negligible.

The latitudinal belt separating the ageostropic Hadley and Ferrel circulations also separates the easterlies and westerlies. The lack of mean relative angular momentum in this latitudinal belt requires the equatorward transport of angular momentum by the thermally forced mass circulation to be restricted primarily to the earth component, as expressed in Eq. (10.9). In conjunction with the meridional forcing by the positive pressure torque associated with the thermally forced geostrophic mass transport, easterlies develop through the westward relative acceleration from a sink of relative angular momentum through the Coriolis torque. This sink is a concomitant source of earth angular momentum. With the development of easterlies in the equatorward branch, friction eventually becomes the dominant torque in subtropical latitudes forcing equatorward mass transport since the pressure torque diminishes towards lower latitudes.

Thus, the combination of the thermally forced isentropic Hadley mass circulation, transport of absolute angular momentum by the mean mass circulation and the forcing of the mass circulation by the torques provides an explanation for the cause of the trade winds. As Lorenz suggested, the secret of the easterlies involves angular momentum exchange in isobaric coordinates. However, it is in isentropic coordinates that the direct coupling of these processes to the thermally forced mass circulation may be isolated.

10.8 On the maintenance of the westerlies

The same source of angular momentum by frictional torque that forces the mass transport equatorward in the subtropics is also the ultimate source of absolute angular momentum for the westerlies of middle latitudes (Jeffreys, 1926; Starr, 1948b; Widger, 1949; Lorenz, 1967). Westerlies are realized in the upper branch through the poleward transport of absolute angular momentum from the tropical-subtropical latitudes to extratropical-polar latitudes by the thermally forced Hadley circulation. In the upper branch of the Hadley circulation a transformation of earth to relative angular momentum occurs through the Coriolis torque. While the poleward eddy flux of relative angular momentum augments the poleward transport of mean angular momentum by the mass circulation, it is not as important as the Coriolis torque in determining the sense of the zonal motion in either the upper or lower isentropic layers. However, upon vertical integration, the net poleward transport of angular momentum occurs through the eddy flux of relative angular momentum, since the vertically integrated Coriolis torque vanishes in a steady zonally averaged mass circulation. The nature of this problem is now outlined by a discussion of the zonally and vertically integrated balance of angular momentum, after which a heuristic explanation for the eddy angular momentum exchange will be offered.

10.8.1 The temporally, vertically, zonally averaged balance of angular momentum

The balance of angular momentum that follows is determined from vertically averaging the zonally averaged transport equation for absolute angular momentum expressed in generalized coordinates, after which temporal averaging is applied. The zonally averaged transport equation for absolute and relative angular momentum in generalized coordinates (Johnson, 1980) is expressed by

$$\frac{\partial}{\partial t_\eta}(\overline{\rho J_\eta}\hat{g}_a) + \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}(\overline{\rho J_\eta}\hat{v}\hat{g}_a) + \frac{\partial}{\partial\eta}(\overline{\rho J_\eta}\hat{\dot{\eta}}\hat{g}_a) =$$

$$= \overline{\rho J_\eta}(\hat{T}_p + \hat{T}_F + \hat{T}_I) \ , \qquad (10.17)$$

$$\frac{\partial}{\partial t_\eta}(\overline{\rho J_\eta}\hat{g}_r) + \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}(\overline{\rho J_\eta}\hat{v}\hat{g}_r) + \frac{\partial}{\partial\eta}(\overline{\rho J_\eta}\hat{\dot{\eta}}\hat{g}_r) = \overline{\rho J_\eta}(\hat{T}_C + \hat{T}_p + \hat{T}_F + \hat{T}_I), \qquad (10.18)$$

where the Coriolis, pressure, friction and inertial torques are expressed by

$$\hat{T}_C = f\,\overline{\rho J_\eta v}/\overline{\rho J_\eta}\ , \tag{10.19}$$

$$\hat{T}_p = a\cos\phi\,\overline{\rho J_\eta P_\lambda}/\overline{\rho J_\eta}\ , \tag{10.20}$$

$$\hat{T}_F = a\cos\phi\,\overline{\rho J_\eta F_\lambda}/\overline{\rho J_\eta}\ , \tag{10.21}$$

$$\hat{T}_I = -\left[\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}(\overline{\rho J_\eta}\,\widehat{v^* g_a^*}\cos\phi) + \frac{\partial}{\partial\eta}(\overline{\rho J_\eta}\,\widehat{\dot{\eta}^* g_a^*})\right]\ , \tag{10.22a}$$

$$= -\left[\frac{1}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}(\overline{\rho J_\eta}\,\widehat{v^* g_r^*}\cos\phi) + \frac{\partial}{\partial\eta}(\overline{\rho J_\eta}\,\widehat{\dot{\eta}^* g_r^*})\right]\ . \tag{10.22b}$$

The equality of the two expressions for the inertial torques, while easily verified, stems from the condition that in the zonally averaged equations, the net eddy transport of earth angular momentum vanishes.

With an integration over the vertical extent of the atmosphere, the temporally, vertically, zonally averaged transport equations for absolute and relative angular momentum are expressed by

$$\frac{\partial}{\partial t_\eta}\overline{(\overline{\rho J_\eta}^{\lambda\eta}\,\hat{g}_a^{\lambda\eta})}^t + \frac{1}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}\overline{[\overline{\rho J_\eta}^{\lambda\eta}(\hat{v}^{\lambda\eta}\,\hat{g}_a^{\lambda\eta} + \widehat{\hat{v}^{*\lambda}\,\hat{g}_a^{*\lambda}}^{\eta})\cos\phi]}^t$$

$$= \overline{\overline{\rho J_\eta}^{\lambda\eta}(\hat{T}_p^{\lambda\eta} + \hat{T}_F^{\lambda\eta} + \hat{T}_I^{\lambda\eta})}^t\ , \tag{10.23}$$

$$\frac{\partial}{\partial t_\eta}\overline{(\overline{\rho J_\eta}^{\lambda\eta}\,\hat{g}_r^{\lambda\eta})}^{t} + \frac{1}{a\cos\phi}\frac{\partial}{\partial \phi_\eta}\overline{[\overline{\rho J_\eta}^{\lambda\eta}\,(\hat{v}^{\lambda\eta}\,\hat{g}_r^{\lambda\eta} + \widehat{\hat{v}^{\lambda *}\,\hat{g}_a^{\lambda *}}^{\eta})\cos\phi]}^{t}$$

$$= \overline{\overline{\rho J_\eta}^{\lambda\eta}\,(\hat{T}_C^{\lambda\eta} + \hat{T}_P^{\lambda\eta} + \hat{T}_F^{\lambda\eta} + \hat{T}_I^{\lambda\eta})}^{t}\ , \qquad (10.24)$$

where the vertically, zonally mass-weighted average and the deviation of the zonal average from this average are defined by

$$\hat{f}^{\lambda\eta} = \overline{\rho J_\eta f}^{\lambda\eta} / \overline{\rho J_\eta}^{\lambda\eta}\ , \qquad (10.25)$$

$$\hat{f}^{\lambda *} = \hat{f}^{\lambda} - \hat{f}^{\lambda\eta}\ . \qquad (10.26)$$

It is easily verified that mass-weighted averaging operations permute, in that

$$\hat{f}^{\lambda\eta} = \hat{\hat{f}^{\lambda}}^{\eta} = \hat{\hat{f}^{\eta}}^{\lambda}\ , \qquad (10.27)$$

$$\overline{\rho J_\eta}^{\lambda\eta} = \overline{\overline{\rho J_\eta}^{\lambda}}^{\eta} = \overline{\overline{\rho J_\eta}^{\eta}}^{\lambda}\ . \qquad (10.28)$$

Likewise,

$$\hat{f}^{\lambda\eta t} = \hat{\hat{f}^{\lambda\eta}}^{t} = \hat{\hat{f}^{t}}^{\lambda\eta}\ , \qquad (10.29)$$

$$\overline{\rho J_\eta}^{\lambda\eta t} = \overline{\overline{\rho J_\eta}^{\lambda\eta}}^{t} = \overline{\overline{\rho J_\eta}^{t}}^{\lambda\eta}\ . \qquad (10.30)$$

Both absolute and relative angular momentum equations are included to draw attention to differences in the degrees of freedom for exchange in the zonally averaged equations and the reasons why these differences vanish upon vertical integration.

The following discussion of the vertically, zonally integrated transport equations for mass and angular momentum will assume for simplicity that the zonally integrated mass distribution is steady. With this assumption, both the vertically zonally averaged mass transport and Coriolis torque vanish, a result expressed by

$$\hat{v}^{\lambda\eta} = 0 \; . \tag{10.31}$$

With this constraint and a division by $\overline{\rho J_\eta}^{\lambda\eta t}$, Eqs. (10.23) and (10.24) simplify to

$$\frac{\partial}{\partial t_\eta}(\hat{g}_a^{\lambda\eta t}) + \frac{(\overline{\rho J_\eta}^{\lambda\eta t})^{-1}}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}\overline{(\overline{\rho J_\eta}^{\lambda\eta}\,\widehat{\hat{v}^{*\lambda}\,\hat{g}_a^{*\lambda}}^{\eta}\cos\phi)}^{t}$$

$$= (\overline{\rho J_\eta}^{\lambda\eta t})^{-1}[\overline{T_P}^{\lambda t}(\eta_0) + \overline{T_F}^{\lambda t}(\eta_0)] + \hat{T}_I^{\lambda\eta t} , \tag{10.32}$$

$$\frac{\partial}{\partial t_\eta}(\hat{g}_r^{\lambda\eta t}) + \frac{(\overline{\rho J_\eta}^{\lambda\eta t})^{-1}}{a\cos\phi}\frac{\partial}{\partial\phi_\eta}\overline{(\overline{\rho J_\eta}^{\lambda\eta}\,\widehat{\hat{v}^{*\lambda}\,\hat{g}_r^{*\lambda}}^{\eta}\cos\phi)}^{t}$$

$$= (\overline{\rho J_\eta}^{\lambda\eta t})^{-1}[\overline{T_P}^{\lambda t}(\eta_0) + \overline{T_F}^{\lambda t}(\eta_0)] + \hat{T}_I^{\lambda\eta t} \; . \tag{10.33}$$

The left hand terms are, respectively, the tendencies of angular momentum and the meridional divergence of the net transport of angular momentum by the mean mass circulation. The right hand terms are, respectively, the temporally, vertically, zonally averaged pressure and frictional torques that integrate to temporally, zonally averaged surface pressure and frictional stresses, and the temporally, vertically, zonally averaged inertial torque. From the assumption that the vertically, zonally integrated mass transport vanishes and the condition that the earth component of angular momentum is invariant with elevation, in that $\hat{g}_e^{\lambda *}$ vanishes, corresponding terms in the transport equations for absolute and relative angular momentum become equal in a given coordinate system. Among coordinate systems, the temporally, vertically, zonally integrated pressure and frictional torques are invariant; however, the mean meridional transport of angular momentum and the inertial torque in one coordinate system are not respectively equal to the mean meridional transport and inertial torque in another coordinate system. From inspection of Eqs. (10.24) and (10.25) it is clear, however, that the sum of the meridional transport of angular momentum by mean and eddy modes must be invariant from one coordinate system to another, since all the remaining terms individually are invariant among coordinate systems.

10.8.2 On the relative importance of eddy angular momentum transport.

For simplicity in discussing the role of the meridional exchange of angular momentum, it is now assumed that the sum of the mean pressure and viscous stresses between the earth and the atmosphere is positive in the mean easterlies of subtropical latitudes and negative in the mean westerlies of extratropical latitudes. In a steady state for the vertically, zonally averaged angular momentum, Eqs. (10.32) and (10.33) reveal only two degrees of freedom for the maintenance of the westerlies against the sink of angular momentum by pressure and viscous stresses at the earth-atmosphere interface. These are the transport of zonally averaged angular momentum by the mean mass circulation, as expressed by the covariance of $\hat{v}^{\lambda *}$ and $\hat{g}_a^{\lambda *}$ (or $\hat{g}_r^{\lambda *}$) and the meridional exchange through the eddy angular momentum transport in the inertial torque. Of these two degrees of freedom, Townsend's (1980) seasonally averaged results for the GWE show that in both isobaric and isentropic coordinates eddy transport is the primary mode, a result that is in agreement with the longstanding results from isobaric analysis (Lorenz, 1967; Oort and Peixóto, 1983). Thus, in the vertically, zonally averaged angular momentum balance of both coordinate systems, the negative inertial torque offsets the positive stress torque in subtropical latitudes, while the positive inertial torque offsets the negative stress torque in extratropical latitudes.

In the isentropic zonally averaged circulation, this balance is manifested primarily through three conditions. The first is that the negative pressure torque of the overworld offsets the positive pressure torque in the underworld, the net difference of the two being equal to the transfer of angular momentum by pressure stresses across the earth-atmosphere interface. The second is that in a steady zonally averaged mass circulation the poleward advection of earth angular momentum in the overworld is exactly offset by the equatorward advection of earth angular momentum in the underworld. Thus, no net meridional exchange of the earth angular momentum component occurs in the averaged circulation. The third condition is that the average transport of mean relative angular momentum by the mean mass circulation is small relative to the intensity of the vertically, zonally averaged eddy angular momentum transport. This is true although the isentropic Hadley circulation transports some angular momentum poleward (Townsend, 1980) in association with the positive vertical wind shear of the easterlies and westerlies.

A fine state of balance also exists in the isobaric zonally averaged circulation, although for somewhat different reasons. In the ageostrophic isobaric Ferrel circulation, the mean mode of angular momentum transport is equatorward, while the eddy mode is poleward. Thus, even though the mean modes of transport in isentropic and isobaric coordinates differ in the sense of direction in middle latitudes, these differences are secondary to the poleward transport of eddy angular momentum. Consequently, the question of why the vertically, zonally averaged angular momentum balance is maintained through a poleward eddy mode of transport (Lorenz, 1967) is common to both isentropic and isobaric coordinates.

This question in part involves understanding why a symmetric Hadley circulation spanning the hemisphere does not exist. Lorenz discussed this question in considerable detail and concluded that while a symmetric Hadley circulation is consistent with the governing physical laws it does not occur because it is baroclinically unstable. Besides being baroclinically unstable, physical arguments suggest that a steady zonally symmetric circulation is in all probability not an attainable state. With the same rotation rate and meridional distribution of mean heat sources and sinks, the generation of available energy would need to be substantially greater in an axially symmetric atmospheric in order to offset the greater viscous dissipation than would occur from the increased kinetic energy of that state. Absolute angular momentum would be conserved in the upper branch of the symmetric isentropic mass circulation provided that vertical exchange of angular momentum between the upper and lower branches by convection and/or viscous stresses was not a factor. Consequently, relative zonal motion of the symmetric state would be greater than the mean relative zonal motion of observed wintertime westerlies by more than one order of magnitude. In turn, the relative kinetic energy of the symmetric state would be greater by more than two orders of magnitude than the present state.

Presumably, with the mean meridional distribution of differential heating unchanged, the time required for the mean isentropic meridional mass circulation to complete one circuit would be unchanged. Within that circuit the kinetic energy would increase from zero to its maximum value and then decrease to zero. Zero kinetic energy would occur in the lower branch of the isentropic Hadley circulation between the westerly and easterly regimes. The kinetic energy maximum would occur in the upper branch within polar latitudes. In order for the kinetic energy to increase from zero to its maximum value through ageostrophic motion and decrease back to zero through viscous dissipation in one circuit, the kinetic energy generation must be two orders of magnitude or more greater than present rates. Such a rate, excedes the rate of incoming energy to the atmosphere. Thus this generation rate is not feasible even apart from considering that the maximum rate for kinetic energy generation involves differences between temperatures of the polar and tropical latitudes in conjunction with the intensity of the differential heating. In a study of these factors, Lorenz (1955b, 1960) concluded that the atmosphere at present is operating near its maximum efficiency in generating kinetic energy. Thus, even apart from other considerations such as inhomogenous boundary conditions of continents and oceans and differential heating that give rise to global monsoonal circulations, the issue of symmetric circulation is academic.

These conjectural discussions concerning why the atmosphere's circulation is unable to develop symmetrically, however interesting, fail to resolve why the zonally averaged angular momentum balance is not satisifed through the transport of zonally averaged angular momentum by thermally forced isentropic mass circulations which span the hemisphere.

In their successful numerical simulation of the isentropic zonally averaged circulation, Gallimore and Johnson (1981b) parameterized the geostrophic component of meridional mass transport as a component of the thermally forced mass circulation and neglected eddy angular momentum transport. This implies that eddy angular momentum exchange is not as important as other processes forcing the mean meridional mass circulation; a conclusion that is consistent with the empirical results for the forcing of the January and July zonally averaged circulations presented in Section 10. Based on physical considerations, neglect of the forcing by pressure torques in conjunction with the geostrophic mode of mass transport in deference to the eddy angular momentum transport would not lead to a successful simulation of the mean Hadley circulation. Still, in the vertically, zonally integrated absolute angular momentum balance, the eddy angular momentum transport determines the meridional distribution of the westerlies in the mean Hadley circulation (Townsend, 1980). Although seemingly so, Gallimore and Johnson's and Townsend's results are not necessarily contradictory.

In view of the successful numerical simulation without inclusion of the eddy mode, it would appear that this mode of transport develops to satisfy vertically integrated angular momentum balance as a natural consequence of other more essential exchange processes. Inspection of Townsend's (1980) results shows that the poleward transport of eddy angular momentum transport occurs primarily in the overworld of isentropic coordinates and the upper troposphere of isobaric coordinates. While these results suggest a mechanism involving barotropic processes in the overworld of isentropic coordinates, it is also feasible that the mechanism involves baroclinic processes that are common to both the overworld and the underworld. These issues are now addressed.

10.8.3 A plausible mechanism for the poleward transport of eddy angular momentum

The origin of the explanation that follows is to some extent linked with Rossby's (1949a) statement that "the kinetic energy gained from local baroclinic circulations must ultimately through dispersion followed by mixing, accumulate in the jet stream zone". This vague statement seems to be related to his earlier argument (Rossby, 1947, 1949b) that the meridional distribution of westerlies could be attributed to a mixing of vorticity by eddies in extratropical polar latitudes. In this concept, the latitude of strongest westerlies occurs at the juncture of the tropical-subtropical regime of angular momentum conservation and the polar-extratropical regime of uniformly mixed mean vorticity. The exchange of angular momentum through mixing of vorticity maintains the westerlies of extratropical and polar latitudes against friction. Lorenz (1967) justifiably criticizes such concepts in noting that "there is no physical basis for applying classical turbulence theory to eddies of cyclone scale". He also argues from relative angular momentum concepts that such a "theory would yield incorrect results, since throughout most of the tropics and subtropics angular momentum is transported towards latitudes of stronger westerlies."

The mechanism for poleward eddy momentum transport, to be proposed heuristically, stems from a discussion by Eliassen (1951) concerning the mixing of entropy and absolute angular momentum. In discussing forced and free convection, Eliassen cautioned that to conclude that a state of neutral static or inertial stability was indicative of a prior episode of instability could be erroneous. He pointed out that a state of neutral stability may stem from the respective mixing of either entropy or angular momentum through physical processes associated with forcing of the circulation. In this latter context, it should be noted that processes which reduce static and/or inertial stability will more readily permit the growth of selective instabilities (Van Mieghem, 1951) such as barotropic (Kuo, 1949) or baroclinic instability (Charney, 1947 and Eady, 1949) and/or lead to intensification of the mass circulation being forced in hydrodynamically stable atmospheric states (Eliassen, 1951).

In an illustrative example Eliassen (1951) noted that mechanically induced turbulence in a statically stable planetary boundary layer over rough terrain creates a neutral layer by decreasing static stability. Within such a regime, Lagrangian elements moving upward with lower potential temperature become embedded in an environment of higher potential temperature, while elements moving downward with higher potential temperatures become embedded in an environment of lower potential temperature. The local diffusion of entropy that ensues throughout the vertical extent of the mixing tends to create uniform potential temperature and thus reduced static stability.

Eliassen also noted that the meridional mixing of absolute angular momentum in an isentropic layer would likewise reduce the inertial stability of a regime. The basis for the following heuristic explanation of why the eddy angular momentum transport is poleward in the earth's atmosphere is founded on this latter example. In this explanation, both the meridional and zonal exchange of absolute angular momentum in the latitudinal regime of sloped baroclinic convection will be considered together.

In westerlies of the overworld, a baroclinic wave regime that has been characterized as geostrophic turbulence (Charney, 1971), Lagrangian elements moving poleward with higher absolute angular momentum will become embedded in a latitudinal regime of lower absolute angular momentum, while Lagrangian elements with lower angular momentum moving equatorward will become embedded in a latitudinal regime of higher absolute angular momentum. In seeking an equilibrium with the pressure distribution, the local mass-momentum imbalance from these latitudinal displacements requires a longitudinal transfer of angular momentum through the action of pressure stresses in order to maintain a quasi-geostrophic state. In the maintenance of a quasi-geostrophic state a poleward moving Lagrangian element loses angular momentum through a zonal transfer to its environment by pressure stresses; likewise an equatorward moving Lagrangian element gains angular momentum from its environment through a zonal transfer by pressure stresses. With the extratropical-polar latitudes becoming the local terminus of the poleward moving Lagrangian element and with the subtropical-tropical latitudes becoming the local terminus of the equatorward moving element, a net poleward transport of angular momentum results from this process.

The mechanism for this longitudinal transfer of absolute angular momentum by pressure stresses is evident from the flux form of the pressure torque, expressed in isentropic coordinates by

$$T_p = - \left[\frac{\partial}{\partial \lambda_\theta}\left(p\frac{\partial z}{\partial \theta}\right) + \frac{\partial}{\partial \theta}\left(p\frac{\partial z}{\partial \lambda_\theta}\right)\right] \quad . \qquad (10.34)$$

In Eq. (10.34), the zonal derivative of the pressure stress $p\,\partial z/\partial\theta$ acting on the surface area $\Delta\theta\Delta\phi$ involves the zonal transfer of angular momentum horizontally within an isentropic tube from regions of poleward motion with higher values of angular momentum to regions of equatorward motion with lower values of angular momentum. The net effect of the longitudinal transfer of angular momentum by this term is to make the zonal angular momentum distribution more uniform within the zonal tube. The vertical derivative of $p\,\partial z/\partial\lambda_\theta$ acting on the surface area $\Delta\lambda\Delta\phi$ involves the zonal transfer of angular momentum horizontally across bounding inclined upper and lower isentropic surfaces of zonal tubes. The net effect of this is to also make the vertical distribution of the zonally averaged angular momentum with respect to potential temperature more uniform. Upon zonal integration the first term integrates to zero, while the second reduces to an alternative form of the isentropic pressure torque,

$$\hat{T}_p = - \frac{\partial}{\partial \theta}\left(p\,\overline{\frac{\partial z}{\partial \lambda_\theta}}\right)/\overline{\rho J_\theta} \quad . \qquad (10.35)$$

The net effect of the zonal transfer of angular momentum within an isentropic layer in conjunction with the meridional exchange of angular momentum is to make the meridional distribution of zonally averaged angular momentum more uniform throughout the region of sloped baroclinic convection. The end result of the mixing of angular momentum through a combination of convective and nonconvective exchange entails a net transport of absolute angular momentum from lower to higher latitudes by eddy processes.

A reason for the failure to relate the longitudinal transfer of angular momentum by pressure torque to the systematic structure of eddy angular momentum transport is the apparent differences of the physical processes involved and the condition that upon zonal averaging the net of the zonal transfer by pressure stresses acting on the crosssectional area $\Delta\theta\Delta\phi$ within a zonal tube vanishes. The consequence of the vanishing zonal average is that an important systematic dynamical process involving angular momentum transfer by pressure stresses within baroclinic waves is relegated to an eddy degree of freedom and thus considered to be an inertial process. Hence, this statistic involving eddy angular momentum transport does not adequately reflect the underlying physical processes involved.

Although never emphasized, the existence of geostrophic regimes within baroclinic wave activity requires a zonal transfer of absolute angular momentum in order for this balance to be established. By virtue of Newton's Third Law, for each negative (positive) component of pressure torque, $\partial(p\,\partial z/\partial\theta)/\partial\lambda_\theta$, acting locally in an isentropic zonal tube, a net positive (negative) counterpart of equal magnitude must also exist. Likewise, for each negative (positive) component, $\partial(p\partial z/\partial\lambda_\theta)/\partial\theta$, acting locally in a tube a net positive (negative) counterpart must exist elsewhere in either another isentropic layer or a combination of layers. Thus, counterpart torques are located partly within the same isentropic tube and partly in adjacent isentropic layers. Through this action absolute angular momentum is exchanged zonally at a rate such that to the first approximation, the meridional motion of both currents remains geostrophic. The heuristic explanation offered herein is simply an application of physical principles under the condition of geostrophy.

The explanation offered here in isentropic coordinates also explains the poleward eddy angular momentum transport in isobaric coordinates and thus the meridional distribution of the isobaric inertial torque. There are, however, subtle but important differences between isentropic and isobaric viewpoints. In the Eulerian isentropic zonally averaged circulation, the thermally forced mass transport in amplifying baroclinic waves constitutes the horizontal branches of the Hadley mass circulation. As such, a direct relation exists between the scales of the mean mass circulation, energy transport, forcing by the zonally averaged torques and differential heating.

In the Eulerian zonally averaged circulation of isobaric coordinates, such links cannot be established directly. The forcing of the ageostrophic isobaric Ferrel and Hadley circulations are primarily determined by the distributions of eddy angular momentum and eddy energy transport. No direct relation between the forcing and response of isobaric mean circulations and the thermally forced mass circulation embedded in the baroclinic structure is evident from the isobaric zonally averaged statistics. Thus, physical inference on these matters from use of isobaric statistics alone is limited.

Besides maintaining the vertically integrated angular momentum balance, the net effect of the eddy angular momentum transport and its associated divergence in subtropical latitudes and convergence in extratropical latitudes, is to shift the latitude of the maximum westerlies poleward. As noted earlier, Gallimore and Johnson (1977b, 1981b) successfully simulated a mean isentropic mass circulation and mean zonal relative angular momentum distribution that compared favorably with the observed distribution without including eddy angular momentum transport. They noted from their results, however, that inclusion of eddy angular momentum as a superposition of angular momentum exchange would have displaced the maximum of the westerlies poleward and thus would have made the meridional distribution of relative angular momentum somewhat more realistic.

Whether the mechanism of poleward eddy angular momentum transport is a barotropic process occurring in the overworld or a baroclinic process that is common to both the overworld and underworld is an important question. The mechanism for the poleward eddy angular momentum transport, which occurs primarily through geostrophic motion (Starr, 1948b), requires a covariance of the meridional and zonal components of the thermal wind field. Such a covariance could develop in the baroclinic temperature contrast between the overworld and underworld, in association with the inclination of isentropic surfaces and frontal deformation occurring in amplifying baroclinic waves. In a pilot study of the thermal wind contribution to the poleward eddy angular momentum transport, Townsend (1980) found limited evidence in the form of episodic contributions to eddy angular momentum transport by a covariance of thermal wind components. However, the contribution from this mechanism appeared to be insufficient to account for its intensity in the temporally zonally averaged circulation. Limitations in the National Meteorological Center Level III global data set at that time may have precluded resolving this issue.

10.8.4 On angular momentum exchange within "zonally averaged" isentropic-angular momentum coordinates.

The heuristic explanation just offered, which has its roots in Dove's explanation of the general circulation (Lorenz, 1967), may appear overly simplistic. In his review of theories of the general circulation Lorenz (1967) credits Exner (1925) with identifying the role of east-west pressure gradients as the means by which absolute angular momentum was exchanged zonally. This exchange precludes the development of excessive upper-level winds that would be realized with absolute angular momentum conservation in a symmetric circulation. Lorenz (1967) discusses the complexity associated with the nonlinearity of this problem. He suggests that equilibirum may never be reached between the zonal flow and the time dependent eddy transport of angular momentum and energy. He places particular emphasis on understanding the relationship between the structure of the baroclinically amplifying disturbances, eddy angular momentum and energy transport and states that "the mathematical work required to find the form of the most rapidly amplifying disturbance, when the zonal flow varies both horizontally and vertically, is extremely involved. The investigator who has solved this problem may still gain little physical insight as to why the eddies prove to have one particular shape rather than another".

The explanation just offered for the underlying physical processes involved in eddy angular momentum transport provides no insight on the temporal and spatial structure of the finite-amplitude features involved in the meridional and zonal exchange of the angular momentum. There is substantiating evidence, however, that the heuristic explanation offered here is correct, at least in regard to isolating the fundamental process that gives rise to the poleward eddy angular momentum transport. If one defines the "zonally averaged" meridional circulation and "mean zonal" motion in isentropic-absolute angular momentum coordinates by averaging within meandering tubes bounded meridionally by absolute angular momentum surfaces and vertically by isentropic surfaces (i.e., tubes with meridional crosssectional area, $\Delta g_a \Delta\theta$), all degrees of freedom for the meridional eddy exchange of absolute angular momentum vanish. The mean mass circulation would be determined by the vertical mass flux $\overline{\rho J(\lambda,\phi,r/\lambda,g_a,\theta)\dot{\theta}}$, associated with heat sources and sinks while the meridional mass flux would be determined by $\overline{\rho J(\lambda,\phi,r)/\lambda,g_a,\theta)\dot{g}_a}$ associated with pressure and viscous torques. In this mean meridional circulation, the meridional exchange of absolute angular momentum would be restricted solely to a transport by the mean mass circulation, which for attainment of a quasi-steady state would necessarily involve a zonal transfer of angular momentum by pressure and viscous stresses. The "zonally averaged" transfer of absolute angular momentum across inclined isentropic surfaces would occur by the mean pressure stress $\overline{\tilde{i}\; a \cos\phi \cdot p\tilde{n}d\sigma_I(\theta)}$, while the component of mean torque associated with this process would be determined by the vertical derivative $(\partial(\;)\partial\theta)$ of this mean stress. The "zonally averaged" meridional transfer of absolute angular momentum across inclined absolute angular momentum surfaces would occur by the mean pressure stress $\overline{\tilde{i}\; a \cos\phi \cdot p\tilde{n}d\sigma_I(g_a)}$, while the component of mean torque associated with this process would be determined by the meridional derivative $(\partial(\;)/\partial g_a)$ of this mean stress. The zonally averaged frictional torque involving transfer by mean viscous stresses across the inclined surfaces would be determined by like degrees of freedom, i.e., mean vicous stresses defined by $\overline{\tilde{i}\; a \cos\phi \cdot (\tau_{\tilde{n}})_F \; d\sigma_I(\theta)}$ and $\overline{\tilde{i}\; a \cos\phi \cdot (\tau_{\tilde{n}})_F \; d\sigma_I(g_a)}$. In these expressions, $d\sigma_I(\theta)$ is an inclined incremental area $\Delta\lambda\Delta g_a$ on the vertically bounding isentropic surfaces of the zonal tube that may be inclined with respect to latitude and longitude

(Johnson, 1980), while $(\underset{\sim}{T}_n)_F$ is the frictional component of the surface stress. Similarly, $d\sigma_I(g_a)$ is an inclined incremental area $\Delta\lambda\Delta\theta$ on the meridionally bounding absolute angular momentum surfaces that may be inclined with respect to latitude and height.

The heuristic explanation for poleward eddy angular momentum transport offered earlier and the result for the mean pressure torque in isentropic angular momentum coordinates just noted have the common physical process of angular momentum transfer by pressure stresses. Since zonal averaging in isentropic-absolute angular momentum coordinates completely removes the degree of freedom associated with the meridional eddy angular momentum transport, the underlying physical process which gives rise to poleward eddy angular momentum transport must stem from a longitudinal transfer of angular momentum by pressure stresses across the meridional meanders of absolute angular momentum surfaces.

The primary difference between the results of conventional isentropic and isentropic-angular momentum analyses stems from the definition of averaging and the strategy used to isolate systematic structure and/or processes. In isentropic-angular momentum coordinates, the form of averaging leads to an explicit analysis and identification of the systematic angular momentum transfer by pressure stresses that is intrinsic to amplifying baroclinic waves within thermally forced mass circulations. In conventional isentropic coordinates, a component of pressure torques systematically transfers angular momentum longitudinally across isentropic surfaces as a mean process, while the remaining part which transfers angular momentum within an isentropic layer is only implicit in an eddy statistic that appears devoid of a physical basis. Consequently, poleward eddy angular momentum transport has been considered primarily to be a nonlinear barotropic process, without any direct relation to the thermally forced hemispheric mass circulation. In isobaric coordinates, neither component of the longitudinal transfer of absolute angular momentum by pressure stesses is isolated in the zonally averaged balance. The contrast of these results and the inferences drawn points out that diagnostic strategy, traditionally termed the "design of an experiment", is an important consideration for isolating systematic processes occurring in the atmosphere.

Whether analysis is conducted in isentropic or isobaric coordinates, the puzzle regarding the eddy mode of angular momentum exchange has its roots in the use of relative angular momentum and the partitioning of atmospheric structure and transport processes into mean and eddy components. Without the use of relative angular momentum as a property and partitioning into mean and eddy processes, the maintenance of the westerlies in the isentropic zonally averaged circulation would simply be explained by the fact that the zonal average of the isentropic poleward transport of absolute angular momentum in the overworld is greater than the equatorward transport in the underworld. At the same time, it is also important to recognize that in isentropic coordinates, the partitioning into mean and eddy components has isolated a common scale for the mean mass circulation, the energy transport and the forcing of the mean mass circulation by torques. Such a commonality cannot be established in isobaric zonally averaged circulation. A global analysis of the exchange in isentropic-absolute angular momentum coordinates should provide insight on relations between global monsoonal circulations, the convective transport of angular momentum, energy and entropy and the nonconvective transfer of angular momentum by pressure and viscous stresses.

10.9 Some additional comments regarding zonally averaged circulations

In the baroclinic wave regime of the overworld of the isentropic zonally averaged circulation, both the mean poleward advection of the absolute angular momentum and the mean pressure torque dominate the inertial torque in the overworld. These degrees of freedom are directly linked to the thermally forced mass circulation in that they occur through the geostrophic mass transport in amplifying baroclinic waves that span much of the meridional extent of the winter hemisphere. Hence, sloped baroclinic convection constitutes a thermally forced Lagrangian mechanism by which all atmospheric properties including absolute angular momentum are exchanged meridionally. In addition to this meridional exchange of angular momentum, the zonal exchange of angular momentum also occurs through pressure stresses. It is this zonal exchange of absolute angular momentum by pressure stresses that permits the atmosphere to develop geostrophic meridional motion in amplifying baroclinic waves.

In angular momentum exchange by pressure stresses, the pressure torque in a layer is intrinsically linked to exchange processes in other layers of the stratified atmosphere. As emphasized repeatedly for the isentropic zonally averaged circulation, negative pressure torque with poleward geostrophic mass transport in the overworld and positive pressure torque with equatorward geostrophic mass transport in the underworld requires a zonally averaged transfer of angular momentum from higher to lower isentropic layers. The net horizontal transfer of angular momentum from the warm sector located in the overworld of an amplifying baroclinic wave to the cold sector located in the underworld of a baroclinic wave occurs by the zonally averaged pressure stresses. See Figs. 10, 25 and 26.

This process is linked to the poleward geostrophic eddy energy transport in the transformation of zonal to eddy available potential energy in the Lorenz atmospheric energy cycle (Oort, 1964; Lorenz, 1967). This process is also linked to geostrophic eddy component of mass transport in the transformed equations of Edmon et al., (1980) and the Eliassen-Palm flux in Andrew's (1983) isentropic, finite-amplitude version of the Charney-Drazin non-acceleration theorem (Eliassen and Palm, 1961; Charney and Drazin, 1961). The underlying physical process in the transformed equations is a zonal transfer of angular momentum by pressure stresses, which is distinct from an eddy convective transport of energy. It is through the quasi-geostrophic approximation that the relation with concomitant geostrophic energy transport is established.

It is also extremely important to understand that the nonconvective flux of absolute angular momentum by pressure stresses in a baroclinic wave regime is essential for the forcing of the equatorward and poleward branches of the Hadley mass circulations that span the hemisphere. By both its physical nature and its mathematical definition in the isentropic zonally averaged equations, the pressure torque is a mean Lagrangian source. As has been established (Johnson and Downey, 1975b, Johnson, 1980), the net pressure torque within a zonal tube is determined by the surface integral of the pressure stresses acting on the tube, which in turn transfers angular momentum between source and sink regions through the zonally averaged pressure stresses. The requirement for angular momentum transfer by both pressure and viscous stresses in conjunction with pressure and frictional torques stems from conservation principles embodied in Newton's Third Law. In both global monsoonal and zonally averaged circulations, the challenge in understanding the balance of momentum lies in isolating systematic exchange of momentum by nonconvective processes in association with the convective processes.

In theoretical developments (Gallimore and Johnson 1977a, 1981a; Johnson, 1980) and diagnostic applications of the isentropic zonally averaged circulation, particular attention has been given to the understanding of the transport of mass, momentum and energy by both convective and nonconvective exchange processes and the relation of these exchange processes to boundary processes. As a measure of the robust nature and accuracy of the results, meridional distributions of both monthly and zonally averaged diabatic heating and frictional torques have been estimated which are surprisingly realistic.

In view of the similarity of isentropic Hadley, the Lagrangian mean (Andrews and McIntyre, 1976, 1978a; Kida, 1977) and residual meridional circulations (Edmon et al., 1980), the basic question should not be over the use of Lagrangian versus Eulerian methods, but over the ability to: 1) resolve a hemisphere scale of mass circulation, 2) insure that integral constraints are satisfied, and 3) reflect accurately transport processes in relation to the forcing imposed. The determination of the isentropic zonally averaged mass circulation based on an Eulerian representation accomplishes all three. For the zonally averaged circulation, the basic forcing is differential heating at the hemispheric scale. Thus the challenge is to determine the relation between differential heating and the response of a rotating baroclinic fluid in the form of a thermally forced mass circulation and its transport of fundamental properties. The results summarized herein substantiate that the fundamental underlying reason for global monsoonal circulations is the differential heating of atmosphere and that the link between forcing and response can be isolated through analysis employing the isentropically averaged governing equations.

Johnson and Dutton (1969) emphasized at the 1969 London Conference on the General Circulation that one's viewpoint of the forcing of an Eulerian zonally averaged circulation depends upon the choice of coordinate system in which averaging is defined. They contrasted the isobaric and isentropic viewpoints by examining the role of diabatic heating in forcing of mean isentropic meridional circulations. This work initiated a series of preliminary isentropic investigations into the nature and structure of atmospheric circulations in isentropic coordinates (Downey, 1971; Henderson, 1971; Downey, 1972; Zillman, 1972; and Otto, 1974). Johnson and Downey (1975a and b) carefully showed how scales of mean and eddy modes of transport and their relations to thermal forcing were critically dependent on averaging.

The isentropic results relegating eddy transport processes to secondary importance contrast with isobaric results which consider the distributions of eddy fluxes of heat and momentum in the baroclinic waves to be the primary means to force the mass circulation. In the isobaric perspective, the unique relations between the mean geostrophic mode of isentropic mass transport, the exchange of angular momentum by pressure torques and the geostrophic transport of energy are not readily apparent. Furthermore, the determination of a hemisphere scale for all of these processes that is common with the planetary scale of differential heating has not received the attention that it merits, since the focus is directed towards eddy exchange processes in the baroclinic wave regime.

The successful isolation of systematic exchange and forcing crucially depends on being able to relate convective and nonconvective exchange processes to differential heating and boundary processes. The choice of a coordinate system and the averaging operation are critical steps for the success of this effort. Since Lagrangian source-sink functions of various physical properties, as they are expressed by substantial derivatives, are invariant, and since these functions constitute the only forcing functions in an absolute sense, the natural choice would seem to be to utilize the Lagrangian mean as envisaged by Andrews and McIntyre (1978b). (The variance of Lagrangian derivatives of vector properties between various non-inertial coordinate systems need not be considered in this discussion since all of the analyses have been developed through governing equations expressed in earth relative coordinates.) However, there are at least two limitations regarding the use of the Lagrangian mean for averaging. Andrews and McIntyre (1978b) call attention to one of these by noting that the definition of the Lagrangian mean at finite amplitudes is itself a non-trivial part of the problem. The other limitation deals with the problem that by their vary nature atmospheric phenomena are not Lagrangian.

Andrews and McIntryre, 1978b defined the Lagrangian mean by integrating or averaging properties in "materials tubes". In a vertically sheared flow, these tubes must move relative both to each other and to the phenomena of interest. Eventually, however, the tubes pass through the phenomenon and are no longer representative of the phenomenon or its forcing.

Concerning global monsoonal circulations, an application of Lagrangian mean operations on true steady-state and stationary phenomena would be well defined due to the coincidence of trajectories and streamlines and the ease with which Lagrangian sources of fundamental properties could be related to the structure of the phenomenon as a whole. In transient atmospheric circulation, however, attempts to apply a Lagrangian mean operator in diagnostic studies of global monsoonal circulations and the processes forcing such circulations leads to insurmountable problems. The spatial structure of trajectories in the global circulation is extremely complex and in view of the transient character of the atmosphere it is particularly difficult to define and determine the average of a "material tube" (Andrews and McIntyre, 1978b) that systematically moves in space and evolves with time. The transient nature and the intrinsic shear and deformation of atmospheric flow precludes determination of the time-averaged Lagrangian trajectories from streamlines of mass transport. In view of this ergodic nature one might ask, if time-averaged Lagrangian trajectories in a global domain cannot be determined, how is it possible to estimate a Lagrangian mean source or sink occuring along trajectories which would be truly representative of the Lagrangian behavior of phenomena in global monsoonal circulations? A hint of the complexity of the problem is evident in Welander's (1955) results from a barotropic model where horizontally sheared flow leads to a distortion of elements of fluid that quickly become unrecognizable within the local confines of a single wave pattern [Fig. 13, p. 34, (Rossby, 1957); also Fig. 5.9 pp. 160, 161 (Dutton, 1976a)].

A strategy employed in diagnostic studies studies of secondary atmospheric circulations has been to define an Eulerian coordinate system that translates with the phenomena (commonly called a quasi-Lagrangian system) and isolate systematic processes within the phenomena by judicious integration (Johnson and Downey, 1975a and b, 1976). When determining the thermal forcing associated with the earth-fixed features--continents and the oceans--and the atmosphere's response, the application of Eulerian isentropic coordinates is simply a special case of the more general quasi-Lagrangian diagnostic methods, where the focus is directed towards determining systematic processes within the global monsoonal circulations.

11. Isentropic analysis, chaos, monsoons, and climate: a summary reveiw

11.1 Isentropic analysis of thermally forced circulations

A primary objective of this study has been to provide a physical basis for global monsoonal circulations from inference based on empirical and theoretical evidence. A perspective based on Shaw's (1930) vision of an overworld and an underworld has been set forth which emphasizes the direct role of differential heating in the planetary exchange of mass, energy and entropy in global monsoonal circulations. The importance of the isentropic mass circulation emerged from the earlier work on the theory of available potential energy (Dutton and Johnson, 1967) and the isentropic zonally averaged circulation (Johnson and Dutton, 1969; Henderson, 1971; Gallimore, 1973; Dutton, 1976a). The results for global monsoonal circulations have been based primarily on isentropic analysis of GWE data (Townsend, 1980; Johnson and Townsend, 1981; Schaack, 1982; Wei et al., 1983; Johnson, 1984, 1985a and b). The Level III data sets from the GWE provided an opportunity for a direct diagnosis of the isentropic mass and energy transport from which the planetary distribution of differential heating and the global nature of monsoonal circulations have been established. Prior to the GWE, studies of the thermally forced circulation were based primarily on climatological estimates of differential heating (Zillman, 1972; Otto, 1974; Otto-Bliesner and Johnson, 1982; Zillman and Johnson, 1985).

The perspective of global monsoonal circulations draws heavily on classic concepts that differential heating creates atmospheric circulation (Halley, 1686; Hadley, 1735; Sandström, 1916; Shaw, 1930). It provides for the mutual existence of global scale Hadley and Walker circulations. It also draws heavily on the concepts of forcing of vortices (Eliassen, 1951; Kuo, 1956) and of exchange processes within extratropical latitudes (Jeffreys, 1926; Starr, 1951a, 1951b, 1968; Lorenz, 1955, 1967).

With the primary aim of providing a perspective for global monsoonal circulations based on isentropic analysis, and the difficulty of relating these results with isobaric analysis, no attempt has been made here to include recent isobaric results from other investigations of moonsoons (e.g., Boyle and Chen, 1987; Chang et al., 1979; Chang and Lau, 1980 and 1982; Krishnamurti and Ramanthan, 1982; Krishnamurti, 1985; Krishnamurti and Surgi, 1987; Luo and Yanai, 1983 and 1984; Nitta, 1972) or the the planetary circulation (e.g., Hoskins and Pearce (eds.), 1983; Oort and Peixoto, 1983; WMO GARP Reports, 1985a and b). Neither was an attempt made to summarize other research employing isentropic analysis. The recent works of Hoskins et al. (1985), Haynes and McIntyre (1987) and others in the application of potential vorticity to planetary circulation have rekindled interests in isentropic theory and analysis that prospered initially in the 30's and early 40's with the work of Rossby (1940), Namias (1940) and others. Danielsen (1961, 1968) and Reiter (1972) employed isentropic analysis in their research on baroclinic circulations and stratospheric-tropospheric exchange during the 1960's. See Reiter (1972) for an extensive list of references. Shapiro (1976 and 1980) has extensively studied the mesoscale structure of jet stream-frontal systems of the upper levels through isentropic analsysis. In a recent paper presented at a symposium to honor Namias, Eliassen (1986) has summarized developments that have stemmed directly or indirectly from isentropic theory and analysis.

In this concluding section, isentropic perspectives of global monsoonal and zonally averaged circulations are first summarized; after which several issues relating to nonlinearity and the statistical partitioning of transport processes are addressed. This discussion of transport processes draws from results for the general circulation that were thoughtfully and eloquently summarized by Lorenz (1967). Then some results concerning isentropic analysis and monsoons are interwoven with thoughts concerning the diagnosis of climate of a chaotic atmospheric circulation. One objective is to emphasize the importance of experimental design and strategy in analysis of transport processes. Another is to suggest by example that a determination of the thermally forced component of atmospheric mass circulation would advance understanding of the nonlinear relation between exchange processes and Lagrangian sources in an analysis of climate. The section concludes with an epilogue.

11.1.1 An overview of global monsoonal circulations

Halley's (1686) original finding that the primary cause of monsoons stems from differential heating of the atmosphere over continents and oceans proved to be conclusive (Kutzbach, G.,1987). Differential heating creates density gradients whereby atmospheric circulation is accelerated. The systematic meridional distribution of differential heating is induced through the maximum solar energy received within tropical-subtropical latitudes and relatively uniform emission of infrared energy over the atmosphere's spherical domain. The systematic longitudinal distribution of heat sources and sinks stems from the differing responses of the surface energy balance of continents and oceans during the annual cycle. The global monsoonal circulations that develop transport energy and entropy from regions of heating to regions of cooling. While many of the regional features of monsoonal circulations have been previously studied in detail, the conduct of the Global Weather Experiment (GWE) and acquisition of global data sets have provided the information and impetus to extend such studies to a global domain (WMO GARP Reports, 1985a and b). The results presented herein have established that global monsoonal circulations exist after a manner evisaged by Shaw (1930).

In monsoonal circulations based on Shaw's (1930) concepts, mean mass circulations occur by upward vertical mass transport from the under- to the overworld in the heat source region and downward mass transport from the over- to the underworld in the heat sink region. These vertical branches are linked by mass continuity with quasi-horizontal mass transport from heat source to heat sink in the overworld and from heat sink to heat source in the underworld. A net energy transport results from the condition that the branch of the mass circulation in the overworld transports more *total flow energy* from the region of the heat source to the region of the heat sink than the branch in the underworld returns (Johnson and Townsend, 1981; Johnson et al., 1985). The net transport of entropy from heat source to heat sink occurs in a manner like the energy transport. Through this net energy and entropy transport, the atmosphere's time-averaged thermal structure maintains its quasi-equilibrium with respect to the systematic distribution of heat sources and sinks occurring in monsoonal circulations. At the same both available potential energy is generated and kinetic energy is produced to offset the dissipation of the atmosphere's kinetic energy by friction. The generation of available potential energy through differential heating requires heat sources to occur at higher pressure than heat sinks, at least within some isentropic layers. The production of kinetic energy in an isentropic layer requires the divergence of the mass transport to occur at high values of dry static energy and convergence at low values, the structure of which involves net boundary work of the heat source region on the heat sink region.

Implicit in this classical concept of monsoonal circulations is the premise that with the requirements for the isentropic mass, energy and entropy transport, the atmosphere will develop sources and sinks of momentum to force the quasi-horizontal branches of the mass circulation to accomplish this transport. Globally, the scale of the isentropic energy transport occurring in the thermally forced mass circulation is not constrained by the latitudinal effects of rotation. In quasi-barotropic atmospheric structure of tropical-subtropical latitudes, this response occurs through ageostrophic mass transport with the upper branch being located at higher elevations than the lower branch. Within this structure, the upper branch transports more geopotential energy from heat source to heat sink than the lower branch returns, while at the same time the lower branch transports more enthalpy from heat sink to heat source than the upper branch returns. Still, the net horizontal energy exchange associated with the geopotential energy component of the dry static energy dominates the enthalpy component. In the tropical regions, this direct response to thermal forcing is common to both isobaric and isentropic transport processes.

In extratropical latitudes, the response of isentropic mass circulation to thermal forcing occurs primarily through the geostrophic mode of mass transport. In this case, branches of the isentropic mass circulation are embedded within amplifying baroclinic waves. Within these branches, the net dry static energy exchange from heat source to heat sink by the geostrophic mode occurs largely through the transport of enthalpy in a seemingly more complicated, but still in actuality quite a simple response. The net poleward transport of enthalpy in an amplifying baroclinic wave occurs through poleward transport of subtropical air in the overworld in association with warm air advection between the trough and downstream ridge and equatorward transport of polar air in the underworld in association with cold air advection between the trough and upstream ridge. Since in an amplifying baroclinic wave the geostrophic warm air advection in the overworld and the cold air advection in the underworld tend to be located at the same geopotential level, the net meridional geostrophic transport of geopotential energy is relatively less important. The conceptual and dynamical similarity between the isentropic geostrophic mass circulation, its forcing by pressure torques and net poleward energy transport embedded in baroclinic waves with the meridional exchange of mass, momentum and energy through sloping convection in shallow rotating atmospheres (Hide and Mason, 1975) is evident.

In isobaric coordinates, the primary mode of energy transport within the geostrophic structure of amplifying baroclinic waves is through transient modes that are independent of the isobaric mean mass circulation. However, with the recognition that the transient geostrophic enthalpy transport in isobaric coordinates is simply a component of the thermally forced isentropic mass circulation, the premise that the atmosphere develops monsoonal circulations which extend throughout tropical, subtropical and extratropical latitudes can be defended.

At this time, our knowledge of the global momentum balance in isentropic coordinates from empirical studies is less complete than in isobaric coordinates. The modes of momentum exchange that force mass and energy transport in the isentropic zonally averaged circulation have been identified

(Johnson, 1980; Townsend, 1980; Gallimore and Johnson, 1981; Johnson, 1984; Townsend and Johnson, 1985). The global exchange of angular momentum and vorticity in the isentropic time-averaged circulation has also been studied (Johnson et. al. 1982; Schaack, 1982; Hoerling and Johnson, 1986; Hoerling, 1987). The results of these studies are in accord with the premises that the atmosphere develops thermally forced mass circulations which systematically transport energy and entropy between heat sources and sinks and that a corresponding scale of momentum exchange develops to force isentropic mass circulation. From the empirical evidence that thermally forced mass circulations exist globally, the premise of corresponding momentum exchange is assured. However, as in the case of the forcing of the isentropic zonally averaged mass circulation, the total forcing by momentum exchange will occur through a combination of physical processes involving convective exchange by mean and eddy transport processes and nonconvective exchange by pressure and viscous stresses (Townsend, 1980; Johnson et al., 1982). This exchange appears more complicated than the exchange of energy and entropy through the mean mass circulation.

Many factors affect the instantaneous energy and momentum exchange -- the earth's rotation, static stability, transfer of momentum and energy across the the land- and ocean-atmosphere interface, nonlinear inertial accelerations, etc. However, to a large degree the surface characteristics of land and ocean, orography and the seasonal migration of the latitude of maximum incoming solar energy determine the seasonal evolution of the monsoonal circulations. The basic cause for the reversal of global monsoonal circulations between summer and winter is the atmosphere's response to the incoming solar energy in terms of a seasonal variation of the surface energy balance that is linked with the differing thermal heat capacities of land and ocean.

Ideally, the diabatic mass transport of global monsoonal circulations is upward over the continental land masses of the summer hemisphere and downward over the continental land masses of the winter hemisphere. The upward diabatic mass transport over continents in summer is due to the dominance of latent and sensible heating over the cooling by infrared emission. In winter, cooling by infrared emission dominates. Over Asia and North America, the diabatic mass transport is in accord with these expectations and the ensuing winter-and summertime monsoonal circulations dominate the planetary circulation of the Northern Hemisphere. This idealization is complicated by energy sources from the ocean to the atmosphere over the Kuroshio Current and Gulf Stream of the Northern Hemisphere. Also, in tropical-subtropical latitudes of the western Pacific, a sizable fraction of the upward diabatic mass transport from moist convection occurs over oceanic regions and not over land masses, even during the peak of the Asian summer monsoon.

In the Southern Hemisphere, regional monsoonal circulations occur over Africa, Australia, and South America. However, the ideal model for monsoons based on a land hemisphere does not fully explain the observed diabatic mass transport. The size of the Southern Hemisphere continents is too limited for planetary scale quasi-stationary anticyclonic circulations to develop in wintertime to concentrate the net cooling of the atmosphere over the same land masses. Thus, the wintertime loss of energy over land masses in the Southern Hemisphere is insufficient to offset the summertime gain of energy in the

Northern Hemisphere. Consequently, in the Southern Hemisphere's wintertime the descending branch of the global monsoonal circulation through infrared emission is located over an extremely large contiguous oceanic region that includes the South Atlantic, the Indian Ocean and the South Pacific, while ascending branches are located over Asia and North America. In the Southern Hemisphere summer, the dominant ascending branches of the isentropic mass circulation occur over the Maritime Continent, South America and to a lesser extend Southern Africa, while descending branches occur over Asia and North America.

In contrasting the nature of global monsoonal circulations over land and ocean hemispheres, a critical factor is the ratio of land to ocean surface area in combination with orography. The large ratio of land to ocean area combined with the orography of the Northern Hemisphere have a marked impact on the circulation through the distribution of heat sources and sinks (Smagorinsky, 1953, 1960; Hahn and Manabe, 1975; Murakami, 1987a and b). The small ratio of land to ocean and the orography of the Southern Hemisphere have much less impact on the overall global circulation. Besides these factors the distribution of the sea surface temperature and the circulations of the atmosphere and ocean as they determine systematic heat sources and sinks through the hydrological cycle and through surface sensible heat transfer are also key elements.

11.1.2 An overview of isentropic Hadley circulations

In agreement with Halley's (1686) results for the forcing of monsoons by differential heating, Hadley's (1735) explanation that the principle cause of the trade winds can be attributed to thermally forced hemispheric mass circulation is correct, even though the structure of the circulation is more complicated than originally envisaged. The systematic meridional transport of absolute angular momentum by the isentropic Hadley circulation is a primary process in the maintenance of angular momentum balance of the overworld. The supply of angular momentum for the westerlies of the overworld occurs through upward diabatic transport within tropical latitudes and poleward transport within the upper branch of the isentropic mass circulation; where in middle latitudes, the pressure torque through the geostrophic mode of mass transport transfers angular momentum to the underworld. The loss of angular momentum by pressure torque in the overworld constitutes the primary forcing of the poleward branch of the isentropic mass circulation in middle latitudes. In polar latitudes the angular momentum is also removed from the overworld by downward diabatic mass transport; however, this loss to the underworld at high latitudes is minimal compared to the gain by upward transport at low latitudes. Within this angular momentum balance of the overworld, the meridional transport by eddy modes is secondary to the poleward transport of the earth component of absolute angular momentum by the thermally forced mass circulation. However, for the vertically, zonally integrated balance the eddy mode is primary in importance.

Within the underworld, the equatorward branch of the thermally forced mass circulation gains angular momentum by the pressure torque in middle latitudes, while losing a smaller amount through the frictional torque. The loss of angular momentum to the earth by the frictional torque is directly linked to a transfer of angular momentum to the earth by viscous stresses; a still lesser amount is transferred to the earth by pressure stresses in association with the "mountain torque." In the equatorward branch of the mass circulation, easterlies are eventually realized through a transformation of relative to earth angular momentum through the Coriolis torque. Absolute angular momentum is gained in subtropical latitudes of the underworld through the transfer of angular momentum from the earth to the atmosphere by viscous stresses within the trade winds. The primary forcing of the equatorward branch of the isentropic mass circulation in the underworld is through sources of angular momentum by pressure torque in middle latitudes and by frictional torque in subtropical latitudes.

The perspective set forth in this summary for the forcing of the time-averaged circulation is applicable to monthly, seasonal, annual and climate time scales. The explanations for global monsoonal and mean Hadley circulations just offered do not address the question of just how is the distribution of individual components of differential heating determined. Differential heating associated with the hydrological cycle in the atmosphere, sensible heating, the impact of clouds, water vapor and other constituents on radiation, etc. pose many interesting problems for the future.

Insight into the relation between differential heating and mass, momentum and energy exchange is of value in understanding the circulations within which the daily weather evolves. At this time scale, the nonlinear interaction between differential heating and exchange processes is pronounced. By virtue of the transport of water vapor, the momentum exchange itself directly impacts both latent heat release through vertical motion and thru the distribution of differential heating. Latent and sensible heating in many circumstances markedly alter the evolution of atmospheric phenomena, e.g. tropical and extratropical circulations, squall lines, etc. Insight into these processes and the factors determining their systematic distribution at different times will prove to be important to the understanding of the temporal variations at longer time scales, e.g., seasonal anomalies associated with the Southern Oscillation/El Niño (Rasmusson and Carpenter, 1982), the interannual variation of monsoons (Shukla, 1987) and climate evolution (Saltzmann, 1983).

11.2 Chaos, nonlinearity and climate

In the initial union lecture entitled, "Chaos and self organization in the lithosphere and earthquake prediction" by Keilis-Borok at the 1987 conference of the International Union of Geodesy and Geophysics, attention was directed to the chaotic nature of individual earthquakes when compared with each other. Integral statistics of physical processes involved in earthquakes, however, were remarkably similar. This similarity was described as self organization, the consistency of which formed the basis for successful prediction of earthquakes. The underlying reason for the chaotic nature of earthquakes stems from the nonlinearity of the governing equations, while the self organization stems from the integral constraints on exchange processes within the same governing equations.

The state of affairs for earthquake and climate prediction are remarkably similar. This remarkable similarity is not surprising, since the earth system as a whole is governed by the same principles. Viewed within the framework of general circulation theory, global atmospheric structure dominated by reversible isentropic processes appears chaotic in nature. However, the statistics from the apparent chaos reveal satisfaction of the balance requirements for mass, momentum and energy. Hence, the integral constraints of systematic exchange processes for the basic properties in global monsoonal circulations ensures self organization within the chaotic state. This self organization constitutes an underlying fundamental basis for climate prediction.

The condition that climate is determined by integral properties and processes in lieu of the initial state of a chaotic circulation was emphasized at the 1955 Princeton Conference on the Application of Numerical Integration Techniques to the Problem of the General Circulation. Von Neumann (1960) discussed numerical prediction of the atmosphere in terms of three different categories, flows determined by extrapolation of tendencies from an initial state, flows determined by factors other than the initial state, and flows where certain traits of the initial conditions influence the outcome of the prediction. He correctly perceived that the easiest problem would be the numerical prediction of weather from initial state information, the next problem in difficulty would be to predict climate, while the most difficult would be the intermediate range prediction. Robinson (1967, 1977) discusses limitations encountered in weather and climate prediction from initial conditions.

At the same conference, results from the first successful simulation of the atmosphere's general circulation were discussed by Phillips (1956, 1960). This similation must be viewed as the forerunner of later similations of the atmosphere's climate even though the experiment did not include orography or land and oceans. The importance of the zonal asymmetries of orography and heating in accounting for the stationary disturbances within the westerlies of the Northern Hemisphere was emphasized by Smagorinsky (1960) in a summary of his early findings (Smagorinsky, 1953).

With regard to atmospheric predictability (Holloway and West (eds.), 1984; Shukla, 1985), now over two decades later, remarkable advances have been made in the numerical simulation of monsoons (e.g. Krishnamurti, 1987; Krishnamurti et al. 1987) and climate, particularily the different climatic regimes of glacial and interglacial eras (e.g., Gates, 1976; Manabe and Hahn, 1977; Kutzbach and Otto-Bliesner, 1982; Kutzbach, J., 1987). The successful modeling of the planetary scale of differential heating is a critical factor in these experiments. In recent monsoon simulations from initial conditions, Krishnamurti's (1987) successful modeling is linked with recognition that the irrotational component of the mass transport must be included in the initial state structure in order that the thermally forced component of atmospheric circulation is consisently linked with the inertial forcing by the rotational component. In the climate simulations, success is linked with the recognition that variations of solar energy associated with the earth's orbit about the sun (Milanovitch, 1941; Imbrie and Imbrie, 1979) and the earth's distribution of continents, oceans and orography combine to produce glacial and interglacial eras through systematic variations of differential heating. While the underlying solar forcing is external to the earth-atmospheric system, the resulting climatic distribution of differential heating internal to the atmosphere involves nonlinear processes and feedback.

Lorenz (1967) stated "that one gets the impression that numerical experiments will ultimately duplicate the atmospheric circulation to any desired degree of accuracy". However, he also cautioned that an explanation of the circulation from such a prediction must be as complicated as an explanation of the circulation of the actual atmosphere. Perfect simulation by itself is not an explanation. Analyses of atmospheric phenomena are essential for advancing understanding and prediction. An important advantage of numerical simulation of climate is that the sensitivity of the circulation to differences in internal forcing and boundary conditions can be analyzed and quantitative information regarding unmeasureable physical and nonlinear dynamical processes can be ascertained. Still, 1) with the need to include a spectrum of physical processes ranging from molecular to planetary scales in modeling global monsoonal circulations, 2) with unanswered questions regarding the nonlinear nature of the atmosphere's response and 3) with the practical limitations that subgrid scale processes must be prescribed by parameterization, some uncertainty remains intrinsic to results from numerical experiments. A key point of Robinson's (1977) discussion is that predictability from initial value information is limited by the time scale of diffusion of momentum and energy. This limiation is offset, however, by information gained from the systematic nature of the thermal forcing that is realized through successful modeling of physical processes.

In his discussion of these issues, Lorenz (1967) raised the question of whether or not the climatic averaged response of the atmosphere's circulation to forcing in terms of ergodic theory was unique. Lorenz calls attention to results from experiments with thermally forced circulations in rotating dishpans which indicate that the governing equations are intransitive by virtue of inherent nonlinearities. In these experiments, more than one statistically steady and stationary state may be realized from the external forcing imposed through use of identical boundary conditions.

One indication of degrees of freedom for intransitivity within the isentropic zonally averaged equations may be inferred from Eliassen's (1951) theory for meridional circulations. In this theory, the response of the mean meridional mass transport to given forcing by torques is governed by the potential vorticity of the mean zonal state. Specifically, for a given intensity of a statistically steady meridional mass circulation forced by differential heating, the angular momentum balance may be satisfied by different and/or aperodic wave regimes in middle latitudes as long as the ratio of the zonally averaged torque to the absolute vorticity of the mean zonal flow remains uniquely related to the mean meridional mass transport. See the discussion in Sec. 10.2 and Eqs. (10.8) and (10.9). A similar argument can be made for the relationship between time-averaged distribution of heating and the means by which the momentum balance of the time-averaged state is satisfied.

The problem of intransitivity raised by Lorenz (1967) is also related to the dynamics of predictability and recurrence (Dutton, 1976b; Vickroy and Dutton, 1979), multi equilibria (Charney and DeVore, 1979; Moritz and Sutera, 1981), attractor sets (Lorenz, 1980), bifurcation (Mitchell and Dutton, 1981; Curry and Winsand, 1986), equiblibration (Pedlosky, 1970, Mak, 1985; Pedlosky and Paloani, 1986), climatic turbulence (Saltzman, 1983) etc., all of which owe their existence to nonlinearity. These topics are receiving considerable attention in dynamical studies of sets of simplified nonlinear equations (Holloway and West (eds.), 1984); however, the issue of whether or not the atmosphere's climate is intransitive remains unresolved.

The resolution of whether the actual climate of the atmosphere is transitive or intransitive will not be a simple matter for several reasons. One primary reason for difficulties stems from the need to determine whether or not different systematic states of circulation occur in response to given thermal forcing in an atmosphere that is intrinsically aperiodic. Part of the aperiodicity stems from nonlinearity of the momentum and energy exchange, while part stems from the chaotic nature of the thermal forcing itself that becomes linked directly with momentum exchange through vertical motion and latent heat release.

Lorenz states that ergodic theory has not provided a general rule for determining whether or not a given system is transitive or intransitive. He also notes that even if it is assumed that the atmosphere is transitive, "we must decide which statistics ought to be determined". Since the source of the intransitivity stems from nonlinearity of the governing equations, the answer will require insight into how thermally forced global mass circulations are related to transport and dissipative processes within the atmosphere-ocean system.

Presumably if two different regimes can exist with a common thermally forced mass circulation, the statistics on the momentum exchange within the atmosphere and at the earth's surface through viscous and pressure stresses should reflect these differences. Thus, all systematic exchange and dissipative processes should be isolated in climate diagnostic studies in order to determine how the processes of differential heating and momentum exchange relate to each other within the particular state realized.

In the determination of systematic transport processes employing partitioning into mean and eddy components for diagnostic analysis and into basic and perturbation states for theoretical analysis, a deeper appreciation must be established of just how the selection of a coordinate system and transport equations prejudices the determination of nonlinear atmospheric response to Lagrangian sources of momentum and energy. The works of Andrews and McIntyre (1976, 1978a, 1978b), Andrews (1983), McWilliams and Gent (1980), Danielson (1981), Plum (1983), Bleck (1984) Tung (1984, 1986) and others including the author (Johnson and Downey, 1975a and b, Johnson, 1980) have advanced these matters.

Transport equations relate coordinate dependent partial differential functions through an Eulerian expansion with the Lagrangian substantial derivative. As such, the selection of a coordinate system determines the structure of the partial differentiation within the Eulerian expansion and thus involves the interrelation between tendency, the divergence of horizontal and vertical transport and the Lagrangian source. As has been pointed out by Johnson (1980), the individual terms of the Eulerian expansion are not invariant when expressed in different coordinate systems. The only invariant quantity is the Lagrangian source of a property as it is expressed by the fundamental principles underlying the governing equations. Since these quantities relate nonlinearly with each other, the selection of a coordinate system, partitioning into mean and eddy quantities and simplifying assumptions, such as the quasi-geostrophic, balanced approximations, etc., impact the form of nonlinear exchange processes and thus the probability of successful isolation of systematic exchange processes in a chaotic atmosphere.

Involved in these considerations are whether or not averaging in time and/or space will isolate systematic exchange processes by either the mean or eddy components. One of the main purposes is to reduce the degrees of freedom involved in the nonlinear relations, thereby reducing complexity and increasing the probability of accurate inference on the relation between forcing in a chaotic atmosphere and the response of climate. Here, it is essential that an understanding is developed of the purpose that the nonlinear exchange processes fulfill in a particular coordinate system and that consideration is given as to whether or not the particular expression for the nonlinear exchange processes is compatible with the scientific objectives of the particular analysis. These matters, however, are complex.

11.2.1 On mean and eddy processes within planetary circulations

Through partitioning, diagnostic and theoretical results for the general circulation since the 40's have consistently assigned the primary dynamics of the atmosphere within middle latitudes to isobaric eddy processes, while the evolution of the mean state is determined primarily through secondary circulations that are forced by eddy processes (Holton, 1972). The strategy to partition atmospheric quantities into mean and eddy components ultimately led to the demise of symmetric theories of the general circulation by establishing the importance of nonlinear quasi-horizontal exchange of momentum and energy within the general circulation (Starr, 1948a, 1948b, 1951a). The description of the general circulation described by Lorenz (1967) in his monograph, "The Nature and Theory of the General Circulation", was set forth from the results of this work (e.g., Starr and Saltzman, 1966). There can be little doubt that this strategy achieved its objective. Implicit in these results is the recognition of a chaotic atmospheric circulation in the form of geostrophic turbulence (Charney, 1971), the dominant role of the eddies in exchange of fundamental properties, and self organization from the consistency of the temporally, zonally averaged state of the atmosphere and the processes which maintain this state.

In the concluding chapter of his monograph, Lorenz (1967) discusses the problem of nonlinear interaction between the zonally averaged flow and eddies. The following comments, while directed primarily to the zonally averaged flow and the corresponding eddies, are also appropriate to the temporally averaged flow and corresponding eddies and thus the climate problem. With good reasons he stated that the "properties of the eddies cannot be represented in terms of the current zonal flow" for purposes of describing their effect on the zonally averaged circulation. He pointed out that the amplification of the eddies modifies the zonal flow through the transport of angular momentum and energy by the eddies. Since the eddies alter the zonal flow, and in turn the zonal flow alters the eddies, equilibrium may never be reached. Such interdependency which involves the nonlinear advection of momentum and energy of the zonally averaged flow by the eddies as well as the inherent nonlinear dependency of these circulation features with the static stability and differential heating is evident in relatively simplified equations for temporally, zonally averaged circulation (Kuo, 1956, 1960). It is also evident from the empirical evidence regarding index cycles, intra- and inter-annual differences of circulation, etc. Hence, Lorenz's conclusions must be accepted without question.

In the problem of climate, another dimension of these complications involves the systematic nature of global monsoonal circulations. In the governing equations for the zonally averaged circulation, the net component of divergence of zonal energy transport associated with east-west exchange vanishes from the condition of cyclic continuity. The results from the GWE have revealed a marked longitudinal dependency of the transports of mass, momentum and energy within global monsoonal circulations that is forced by both zonal and meridional gradients of heating. The systematic zonal exchange of energy by isentropic Walker type circulations even extends into middle latitudes (Zillman, 1972; Zillman and Johnson, 1985; Johnson, 1985a and b), within which the longitudinal exchange of energy between continents and oceans of the Northern Hemisphere is nearly as intense as the meridional exchange. Since, the divergence of the zonal energy transport impacts the regional energy balance, the meridional exchange becomes dependent longitudinally on the zonal exchange, being different in the Northern and Southern Hemispheres.

In the Northern Hemisphere, substantial wintertime atmospheric heat sources in mid-latitudes result from latent and sensible latent heat fluxes from the ocean to the atmosphere along the Kurishio Current and Gulf Stream. The maintenance of a quasi-equilibrium state for the oceans requires poleward heat transport within the oceans. Thus, for the quasi-equilibrium balance of the earth-atmosphere system as a whole the eddy poleward energy transport in the atmosphere and the ocean in effect become coupled with each other. Such dependency of the heat transport on the distribution of continents and oceans precludes accurate parametric description of mean meridional eddy momentum and energy exchange by simple hypothesis. Thus, efforts to parameterize these nonlinear processes in modeling of zonally averaged circulation and climate are fraught with difficulties. This complexity bears out the wisdom of Lorenz's (1967) conclusions concerning the difficulty of prescribing properties of eddies from properties of the zonally averaged circulation, although for a somewhat more involved reason.

Another dimension of this issue is sharpened from consideration of the following question which was posed to the author some years ago. "I understand eddies in isobaric coordinates, could you please explain what eddies are in isentropic coordinates?" In this matter, it is worthwhile suggesting that the best that can be achieved is an understanding of what purpose eddies fulfill within the chaotic state. If eddies in the chaotic state were physical circulations that were governed by unique dynamical relations, it would be possible to determine the origin of their existence, and an understanding of their structure and dynamics in a general sense. However, in view of the finite amplitude state of the atmosphere, its nonlinear character, the nature of global monsoonal circulations and the fact that an eddy is a mathematical entity with regional dependencies, such a goal does not seem realistic. Eddies as mathematical entities in the chaotic state only assume a statistical meaning within the coordinate system in which they are defined, the governing equations in which they are applied, and the problem towards which they are directed.

On this point, it is important to emphasize that the representation of the chaotic state through the definition of mean and eddy quantities leads to ambiguities provided that certain relations between mathematical definitions and physical processes are not recognized. Very simply, additional degrees of freedom for exchange processes are introduced through mathematical definitions without introducing corresponding governing relations based on physical principles. However, ambiguities in interpretation and uncertainities over the role that eddies serve among different coordinate systems need not stem from these factors. Through the expression of the governing equations in generalized coordinates (Johnson, 1980) and use of the mass-weighted average (Johnson and Downey, 1975a), corresponding terms in the transport equations when partitioned into mean and eddy processes have common forms and physical interpertations among the various coordinate representations.

Johnson and Downey (1975a) followed the definition of a mass-weighted average that was introduced originally by Reynolds (1894). In this definition, the mean property or process was determined by the ratio of the integral of the quantity over a region of interest to the integral of the total mass within the same region. This is the only definition that defines an average specific property which when multiplied by the mass of the region returns the integral of the quantity for that region. It is also the only definition which 1) uniquely divides a quadratic quantity into the two positive definite components, 2) resolves the convective flux of a property into a component associated with the transport of the mean of a property by the mean mass circulation and the eddy component which is statistically independent of the mean mass circulation, and 3) provides for a consistent physical perspective of mean and eddy processes. Lagrangian sources or sinks of a property are determined by direct integration of the process over the region of interest. In the determination of systematic Lagrangian sources and sinks, the mass weighted average is the only operator by which temporally, spatially averaged sources of momentum by pressure-viscouses forces may be directly related to temporally averaged surface stresses. Likewise, it is the only operator through which the sum of temporally, spatially averaged sources of kinetic and internal energies through pressure and viscous forces may be directly related with the surface integral of boundary work (Johnson, 1980; Johnson and Downey, 1982).

Differences in results and physical interpretation for mean and eddy processes among coordinate systems stem primarily from the integration over different domains in the determination of systematic processes in baroclinic regimes. With a different mean being determined by integration over different domains, the structure of eddy deviations must also be different and as such transport through the eddy structure will be different. For example, the results of Section 9 establish that isentropic energy transport occurs through thermally forced mass circulations, while isobaric energy transport occurs through the combination of stationary and transient eddy circulations with the latter being independent of the mean isobaric mass circulation. Hence, the "modus operandi" of the linear and nonlinear degrees of freedom involved in the interaction between mean and eddy states as revealed by empirical evidence is different between the two coordinate systems. This in turn will impact inferences drawn from analysis and theory regarding nonlinear exchange processes within the atmosphere. Here, the need for understanding of the purpose that eddies fulfill in the framework in which they are being analyzed is most evident.

With regard to difficulties in understanding linear and nonlinear processes, attention is directed to Lorenz's (1967) suggestion that deeper physical insight may sometimes be afforded by a simple qualitative description, than for example, a theoretical solution. It is true that a theoretical solution nearly always based on linear dynamics enjoys mathematical rigor and uniquely relates cause and effect provided, of course, that the underlying assumptions and simplifications are valid and the analysis is correct. However, in gaining tractability for theoretical solutions, the sacrifice of nonlinearity and/or interactions between fundamental physical processes and dynamical degrees of freedom enjoyed by the atmosphere, precludes solutions that adequately describe global monsoonal circulations. In view of the atmosphere's nonlinear behavior within a finite amplitude state, it is evident that some analytic solutions and conclusions based on linearization of the chaotic state of the atmosphere limit our insight and understanding of monsoonal circulations.

To a certain degree atmospheric scientists have succumbed to the temptation to accept eddies as physical entities, a dangerous pitfall that Lettau (1954b) cautioned against. This is intrinsic in the development of linearized theory and analysis of planetary circulations by virtue of the assumption that the basic state satisfies the governing equations independently of the equations for the perturbation state. Since ideally the perturbation state does not interact directly with the basic state, criteria for meridional propagation of mass, momentum and energy seldom incorporate adequate degrees of freedom to reflect the structure of the atmosphere, e.g., the finite amplitude of a zonally varying basic state with both horizontal and vertical shear, complex boundary conditions involving mountain ranges, land-ocean contrasts, the biosphere, etc. Clearly, all of these features are intimately involved with nonlinear exchange processes within the governing equations.

11.2.2 On inference of systematic exchange and the forcing and response of climate

Although there are diverse reasons for the introduction of mean and eddy quantities within governing equations, one key reason in diagnostic studies is to determine the systematic behavior of a chaotic state. A multi-dimensional statistical distribution is generated from which various moments of the distribution such as means, variances and covariances within the ensemble are estimated to extract information concerning systematic atmospheric structure and/or processes. Once the set of statistics are determined through employing a set of sufficient estimators (Mood, 1950; Wilks, 1962), ideally the joint residual distribution of data contains no additional information on systematic structure or exchange.

In the case of temporal averaging of the Eulerian isobaric equations to study the planetary circulation, Starr's essay (1948b) reveals that his aim was to establish the importance of quasi-horizontal exchange processes and thereby advance our understanding of the general circulation. This occurred at a time when the primary emphasis was still directed to the importance of symmetric meridional circulations. He and his colleagues accomplished this aim (Starr and Saltzman, 1966).

In the case of temporal averaging in isentropic coordinates, the aim has been to determine the global nature of thermally forced mass circulations and to establish that these systematic circulations are fundamental in the exchange of atmospheric properties. The isentropic statistics presented herein testify to the primary importance of the transport of fundamental properties by global monsoonal mass circulations. In regard to this objective, the isentropic results also testify that the set of statistics from isobaric analysis is not determined from a set of sufficient estimators. Additional information concerning systematic processes occurring within the atmosphere's circulation is provided by the isentropic analysis, and this additional information can not be determined without approximation from transformations of mean isobaric to mean isentropic statistics. It is also true for some purposes that the set of statistics from an isentropic analysis layer are not determined from a sufficient set of estimators. For example, difficulties are encountered in isolating systematic structure and processes occurring in the planetary boundary layer through isentropic analysis.

Given the seemingly diverse nature of strategies which emphasize eddy versus mean statistics, isobaric versus isentropic analysis or a combination of both, it seems timely to discuss diagnostic strategies to study the atmosphere's climate as they relate to objectives, design of experiments and inference.

A most important factor is always a clear statement of the problem to be addressed. From such a statement of objectives and/or hypothesis, the design of the experiment becomes important in determining whether or not the information can be resolved from the data in a form that either fulfills objectives and/or resolves the hypothesis through support or rejection as the case may be.

In the design of an experiment, the selection of a coordinate system for analysis is important. For example, in the study of the forcing and maintenance of a symmetric vortex, the problem is simplfied by selection of a cylindrical (or spherical) coordinate system to analyze the phenonmena; it remains complex otherwise.

The successful isolation of representative mean values through averaging in a chaotic atmosphere depends crucially on being able to isolate the region of systematic structures and/or processes in time or space. Here, both selection of a coordinate system and the choice of averaging operations are critical steps for the success of this effort. If the region over which the mass-weighted average is determined coincides with a systematic processes, the systematic behavior will be reflected in the average. If the two do not coincide, the systematic process will in all probability be reflected in a combination of nonlinear statistics of mean and eddy exchange within which complexity remains due to confounding. In this case, efficient and unbiased estimators (Mood, 1950; Wilks, 1962) of the systematic structure or processes have not been defined.

In analysis of the climatic response of the atmosphere, the isolation of the planetary distribution of heating is of primary importance. Equally important is gaining insight into the relation between differential heating and the nonlinear response of the atmosphere through transport processes. The combination of these two factors suggest that in the determination of sufficient and efficient estimators for diagnosis of climatic response to planetary scale differential heating, the selection of isentropic coordinates for analysis is essential. The discussion that follows serves to illustrate how these choices are important for these two objectives.

For this discussion, a statistical model of an N-dimensional sample space of observations is utilized. Within this sample space, it is assumed that the atmosphere's climate over some period is defined by the spatial distribution of time mass-weighted averages, while the eddy state is defined by deviations of observations from this state. As such, from a statistical perspective for each grid location, one degree of freedom (dimension) of an N-dimensional sample space is assigned to the estimate of the mean of a property or process, while (N-1) degrees of freedom are associated with the distribution of deviations about the mean. Apart from the deviations conforming to some parametric distribution function for each property, all (N-1) degrees of freedom must be retained to describe completely the information contained within the observations.

In a simple example, the systematic information from a normal distribution, such as the annual distribution of temperature on a tropical island, is described by the mean and variance for the ensemble. The average from the sample is an estimate of the mean, while the estimated variance provides information on the uncertainty of the average with regard to the true but unknown mean value. Estimators for the mean and variance define a sufficient set of statistics for a normal distribution of random independent variables. Once these parameters are estimated, estimation theory states that knowledge of individual deviations provide no additional information.

Within a physical context, it is possible that the deviation state vanishes as it would in a steady and stationary state or that the processes associated with the deviation state in an average sense are unimportant relative to the processes associated with the mean state. In this latter case, however, caution must be exercised against inferring that the chaotic nature of the system serves no purpose, since a mean statistic may simply isolate systematic processes occurring within the chaotic state. The isentropic time-averaged mass and energy transports by the geostrophic mode are examples of such systematic processes that occur within chaotic amplifying baroclinic waves. It must not be interperted as a process occurring within a steady stationary state. Hence, the problem becomes one of a correct inference from knowledge of the purposes that partitioning of transport processes fulfill.

A key objective in statistical inference then is to understand the dynamics of a system by studying the systematic structure and behavior of a system. Presume that all the information has been extracted from the data through estimators of a minimum number of sufficient statistics, P. In disciplines without governing dynamical equations, the aim would be to describe the system through the development of a model with P degrees of freedom. In this application, assume that the atmosphere's evolution is described by the governing transport equations of basic physical properties and that its systematic behavior is estimated from temporal averages of its structure and the various transport processes. Furthermore, it is assumed that the averages of all processes are well defined and that within the multi-dimensional sample space, the systematic behavior of the system as described by the temporally averaged dynamical processes may be determined from a subspace of averaged properties and processes.

With the use of the flux form of the equations to represent the dynamical processes, such assumptions are consistent. Through the interchange of temporal averaging with spatial differentiation, time-averaged tendencies within each of the governing equations are related with the divergence of the time-averaged convective (mean motion) and nonconvective (molecular) flux of atmospheric properties (Van Mieghem, 1973; Johnson, 1980). Thus, consistent spatial relations will be determined between mean Lagrangian sources and sinks of properties and the response of the atmosphere in the form of time-averaged convective transport of properties within the atmosphere. Consistent spatial relations will also be established between source/sink functions of momentum and energy associated with the time-averaged nonconvective fluxes and systematic forcing through boundary conditions.

As discussed previously, however, in partitioning that aims through analysis to reduce the degrees of freedom to a minimum, the use of the mass-weighted average is essential. With its use, time-averaged convective fluxes of momentum, energy, and entropy expressed as $\widehat{\underset{\sim}{U}f}$ partition into a mean component, $\hat{\underset{\sim}{U}}\hat{f}$, representing the systematic transport by the time-averaged mass circulation and an eddy component, $\widehat{\underset{\sim}{U}^*f^*}$, representing transport that is independent of the mean mass circulation. If another average is used to determine a mean velocity component, which is not explicitly related with the mean mass circulation through the time-averaged mass continuity equation, the interpretation of mean and eddy modes of transport has only a statistical basis.

Provided the overall strategy is successful, the partitioning into mean and eddy quantities in the governing equations ideally reduces the number of degrees of freedom needed to describe the systematic nature of the atmosphere's structure and the transport processes. Consider for the moment that N_1 is the the number of grid points at which observations are required to specify the global state of the atmosphere for a particular variable, N_2 is the number of different variables that must be observed and N_3 is the number of occurrences that are observed. In this case, the number of dimensions of the sample space to describe the complete temporal and spatial structure for variables within the domain is the product, $N_1N_2N_3$. This number becomes exceedingly large in studies of climate. Clearly, if the climate and the dynamical processes were determined by mass-weighted time-averages $N_1N_2(N_3-1)$ degrees of freedom are removed, and the system is determined by N_1N_2 degrees of freedom. From a physical perspective, the systematic exchange of properties is described by the geographical distribution of the mean structure and processes. In this case, no information is provided by the time-averaged eddy covariance statistics nor is their need for knowledge of the eddy distribution functions as a function of climate state for the time-averaged balance of properties. From both a physical and dynamical perspective, the difficulties inherent in nonlinearity are lessened. From a statistical perspective, the likelihood of accurate inference concerning response of a circulation and its forcing by the geographical distribution of time-averaged Lagrangian sources and sinks of momentum and energy is increased.

A complete analysis of these issues requires that all the governing equations be considered. Here, only the balance of total energy exchange is discussed to illustrate the potential of isentropic analysis to aid in the inference of the response of climate to systematic forcing through simplification of nonlinear exchange processes. A partitioning of the generalized transport equation for total flow energy, Eq. (8.10), into mean and eddy components yields

$$\frac{\partial}{\partial t_\eta}(\overline{\rho J_\eta}\,\hat{e}) + \nabla_\eta\cdot[\overline{\rho J_\eta}(\hat{\underset{\sim}{U}}\hat{\nu} + \widehat{\underset{\sim}{U}^{*}\nu^{*}})] + \frac{\partial}{\partial\eta}[\overline{\rho J_\eta}(\hat{\dot{\eta}}\hat{\nu} + \widehat{\dot{\eta}^{*}\nu^{*}})] + \frac{\partial}{\partial\eta}\{R\overline{\rho J_\eta}[(\widehat{\frac{\partial z}{\partial t_\eta}/J_\eta})\hat{T} + \widehat{(\frac{\partial z}{\partial t_\eta}/J_\eta)^{*}T^{*}}\}$$

$$= \overline{\rho J_\eta}\,[\hat{Q}_m + (\hat{\underset{\sim}{U}}\cdot\hat{\underset{\sim}{F}} + \widehat{\underset{\sim}{U}^{*}\cdot\underset{\sim}{F}^{*}})] \quad . \qquad (12.1)$$

The fourth left hand term is part of the RT component of boundary work as it is related to the local vertical displacement of an η-surface. Through use of the equation of state and use of the mass-weighted time-average, the vertical exchange of energy by this degree of freedom is determined by mean and eddy components. The other components of RT boundary work involve the exchange of total energy in the second and third left hand terms through horizontal and vertical motion.

Based on the results from isentropic analysis summarized in Sections 5 and 6, the time-averaged balance of total energy in isentropic coordinates is expressed by

$$\frac{\partial}{\partial t_\theta}(\overline{\rho J_\theta}\,\hat{e}) + \nabla_\theta\cdot(\overline{\rho J_\theta}\,\hat{\underset{\sim}{U}}\hat{\nu}) + \frac{\partial}{\partial\theta}(\overline{\rho J_\theta}\hat{\dot{\theta}}\hat{\nu}) + \frac{\partial}{\partial\theta}[R\overline{\rho J_\theta}\widehat{(\frac{\partial z}{\partial t_\theta}/J_\theta)T}]$$

$$\simeq \overline{\rho J_\theta}(\hat{Q}_m + \hat{\underset{\sim}{U}}\cdot\hat{\underset{\sim}{F}} + \widehat{\underset{\sim}{U}^*\cdot\underset{\sim}{F}^*}) \; . \qquad (12.2)$$

The advective form of this balance is

$$\frac{\partial}{\partial t_\theta}\hat{e} + \hat{\underset{\sim}{U}}\cdot\nabla_\theta\hat{\nu} + \hat{\dot{\theta}}\frac{\partial}{\partial\theta}\hat{\nu} + \{\frac{\partial}{\partial\theta}[R\overline{\rho J_\theta}\,\widehat{(\frac{\partial z}{\partial t_\theta}/J_\theta)\,T}]\}/\overline{\rho J_\theta}$$

$$\simeq \hat{Q}_m + \hat{\underset{\sim}{U}}\cdot\hat{\underset{\sim}{F}} + \widehat{\underset{\sim}{U}^*\cdot\underset{\sim}{F}^*} \; . \qquad (12.3)$$

A comparison of the equations between generalized and isentropic coordinates reveals simplifications that impact the nonlinear exchange of total energy. In (12.1), the balance of total energy involves mean and eddy convective exchange as it would be determined in isobaric coordinates. In Eq. (12.2) the convective exchange is determined from the transport of the mean total flow energy by the mean isentropic mass circulation. In Eq. (12.3), the balance between the sources and sinks occurs through the advection of the mean total flow energy by the mean motion. These simplifications in isentropic coordinates were supported by theoretical and empirical evidence that: 1) the scale of the mass circulation, energy exchange and differential heating are common and 2) the dynamical constraint that geostrophic advection of mass, dry static energy and entropy are uniquely related with each other, thereby precluding the existence of a transient component of geostrophic eddy energy transport. These simplifications are also supported by the remarkable similarity of the distributions of the time-averaged irrotational and rotational components of the mass transport with like components of energy transport. No results have been presented that the eddy diabatic transport and the eddy RT component of work are negligible in the third and fourth left hand terms. These have been presumed negligible by inference from the condition that the mean mass circulation is the primary degree of freedom determining energy exchange.

In summary, the subspace of mass weighted averages specifing the energy exchange for the climate state consists of N_1N_2 dimensions within the sample space of the $N_1N_2N_3$ dimensions : N_1 degrees of freedom from the geographical distribution of grid points and N_2 from the variables relevant to systematic exchange processes. Furthermore, within the governing equations, the colinear nature of the horizontal divergence of the isentropic geostrophic transport of mass, energy and entropy in Eqs. (8.22a), (8.24) and (8.26), 1) further constrains the degrees of freedom for transport processes among these three properties, 2) simplifies physical understanding of nonlinear exchange processes in the determination of climate and 3) increases the likelihood of accurate inference regarding a system's physical behavior. It is exactly through these simplifications that a perspective of global monsoonal circulations has been established and a means is set forth for analysis of the climatic response of the atmosphere to differential heating.

The simplifications achieved by use of isentropic coordinates may be traced to the condition that the transient exchange of entropy both vertically and horizontally is precluded by the very condition that Lagrangian sources and sinks of entropy uniquely determine the vertical mass flux associated with differential heating. The same Lagrangian sources and sinks of entropy when weighted by temperature determine energy sources and sinks by heating and also the vertical advection of dry static energy. It is interesting to note a traditional view that the use of isentropic coordinates only enjoy certain advantages if the motion is isentropic, otherwise not so. This is equivalent to stating that the use of isobaric coordinates only enjoys advantages if the motion is isobaric. Advantages accrue to the use of isentropic coordinates in the presence of differential heating and frictional processes from the fundamental principle that entropy change in the atmosphere is an exact differential, and thus through use of isentropic coordinates, the vertical mass flux becomes independent of isentropic exchange processes.

The simplifications achieved by time-averaging of transport processes in isentropic coordinates stand in marked contrast with a complexity that remains in time-averaged transport processes of mass, energy and entropy in isobaric coordinates. With the distributions of isobaric transport processes being dependent on both the mean and eddy components of larger scale flow except in the tropics, the number of statistics and physical processes needed to determine the time-averaged state of total energy and the relation of its exchange to differential heating is not reduced. The failure to reduce the degrees of freedom is due to the differing functional dependence of the isobaric divergence of the geostrophic transport of mass, energy and entropy in Eqs. (8.21), (8.23) and (8.25). Consequently, the time-averaged planetary distributions for the isobaric stationary and transient modes of transport for total energy and entropy will differ from each other, and also from the distribution of the time-averaged mass transport.

In view of the direct link between mass, energy and entropy transport established in isentropic coordinates and the lack of a direct link in isobaric coordinates, it must be recognized that the mean and eddy modes of isobaric transport of these properties have a confounded relation to the planetary distribution of differential heating. The confounding of these isobaric statistics precludes identification of a sufficient and efficient set of estimators to prescribe the relation between differential heating and the thermally forced nonlinear exchange of mass, total energy and entropy in a chaotic atmosphere. Whether or not the set of estimators prescribing the isentropic time-averaged exchange of the same three properties constitutes a sufficient set for these two objectives can only be determined from further analysis. Relative to the isobaric set, the isentropic set of estimators more nearly satisfies the criteria of sufficiency and efficiency than the isobaric set.

These consideration suggest that additional insight in climate, the processes that determine the climate and intransitivity can be gained through isentropic diagnostic analysis. Presumably, an analysis of the planetary distribution of heating and exchange processes within the chaotic structure of the atmosphere would resolve whether of not different climatic states and/or systematic distributions of momentum exchange occur in association with a particular planetary distribution of mean heating. The extent to which climate is determined by external forcing, internal forcing, nonlinearity and other processes is surely one of the most challenging problems of the coming century. Only future studies will reveal the utility of isentropic analysis is addressing this challenge.

11.3 An epilogue

There can be little doubt that the Eulerian viewpoint of the isentropic and isobaric time- and zonal-averaged circulations contrast markedly. At the same time, the results are mutually consistent provided recognition is given to the dynamical factors that determine their respective scales. Shaw's (1930) vision of an overworld and underworld with vertical exchange processes occuring through differential heating provides a basis for the definition of global monsoonal circulations. The isentropic analyses presented herein and the application of Eliassen's (1951) results to the isentropic zonally averaged circulation provide a dynamical framework for isentropic Hadley circulations. The primary physical processes that force the zonally averaged mass circulation and also the energy transport are differential heating and pressure and frictional torques (Gallimore and Johnson, 1971a). In a Lagrangian sense, differential heating and pressure and frictional forces are the only physical processes that force circulation within a hydrostatic atmosphere.

On these matters it is interesting to call attention to an exchange between Palmén (1949) and Starr (1949) regarding Rossby and Starr's (1949) emphasis on the importance of quasi-horizontal exchange processes in the maintenance of the angular momentum balance of the atmosphere. In an effort to defend classical concepts, Palmén (1949) argued that meridional circulations were necessary in some form to account for the maintenance of the zonal circulation. He called attention to the fact that a mean meridional circulation of only a few centimeters per second could maintain the westerlies. In this exchange, Palmén stated:

> "It is not clear whether the authors intended to question the necessity of meridional circulations for the maintenance of the kinetic energy of the atmosphere. Such an intention would mean a complete change of the whole foundation of dynamic meteorology, and I doubt strongly that the genius of the authors, recognized by all meteorologists, will be sufficient for that goal."

Starr (1949) responded to Palmen's challenge by first calling attention to the empirical evidence of Widger (1949) which revealed a distribution of poleward eddy angular momentum transport, the magnitude of which was sufficient to maintain the westerlies. With this evidence in mind, he pointed out that there was no requirement for a systematic transport by mean meridional circulations and in a concluding paragraph, then noted that "apparently Palmén suspects me of highest heresy". The intensity of the debate concerning the maintenance of the circumpolar circulation is evident in Starr's reply:

> "Dr. Palmén speaks of "the whole foundation of dynamic meteorology." What does he mean by it? Certainly he does not mean the collection of sundry differential equations such as the hydrostatic equation, the continuity principle, etc., which is found in textbooks. Does he mean some rational solution of these equations which purports to give the essential mechanism of the general circulation? To the best of my information the problem is so difficult that we have no such solution and cannot hope for one in the directly forseeable future. Any insistence that at the present time the fundamental facts concerning the mechanism of the general circulation have been established is unfounded and misleading. The history of science and all of our general experience in meteorological research points to the inherent danger of the premature acceptance of superficially plausible hypotheses as fact. I therefore maintain that under present circumstances we must encourage free experimentation with various hypotheses and proposals in order to see which ones lead to the *discovery of new observationally verifiable facts*, since definitive criteria for acceptance or rejection are lacking."

Now, several decades later, it is interesting to consider that within the isentropic framework, thermally forced systematic hemispheric meridional circulations have been isolated.

On the overall question of the relative importance of mean meridional circulation versus quasi-horizontal exchange processes, there is little doubt that both Palmén and Starr were partially correct in their views; however, it is also true that both views were understandably incomplete. Tempered by later results, the views of both Palmén (Palmén and Newton, 1969) and Starr (1968) allowed for meridional circulations in tropical-subtropical latitudes and quasi-horizontal exchange in extratropical latitudes. Both of these views contribute to an isentropic perspective of the maintenance of the angular momentum balance within the zonally averaged circulation. The thermally forced mean isentropic mass circulation is fundamental to the meridional exchange of both angular momentum and energy. However, quasi-horizontal exchange processes within baroclinic waves are equally fundamental in accounting for the poleward meridional transport of angular momentum and energy within isentropic Hadley circulations. In his book, Starr (1968) called attention to a need to understand the problem of the eddy transport of angular momentum against the gradient of relative angular momentum and pointed out the universality of this problem in phenomena where rotation is involved.

In the early forties, isentropic data were included with the radiosonde coded reports and routinely transmitted with regular weather information. Dr. Jerome Namias once noted in a informal conversation with the author that the deletion of the coded isentropic radiosonde data was a decision made by a communications official in order to relieve overburdened weather communications sometime during World War II. With this deletion of isentropic data, interest in the use of isentropic analysis in weather forecasting waned.

It is interesting to ponder how views of the atmosphere's general circulation may have evolved if along with his isobaric analysis, Professor Starr had undertaken isentropic analysis of the general circulation building on the results that he set forth in his work, "A quasi-Lagrangian system of hydrodynamical equations" (Starr, 1945). In the thirties and forties research work was being conducted in isentropic coordinates and an extensive set of governing dynamical equations remained to be set forth in isobaric coordinates (Eliassen, 1949). It is likely that if isobaric and isentropic analyses had been applied concurrently to analysis of the general circulation, an understanding of Lagrangian concepts as embodied through isentropic coordinates would have advanced more rapidly and that our views of the general circulation would have evolved differently.

While this author was not privileged to know Professor Starr personally, other than spending an hour with him being introduced to the importance of the hydrological cycle during a visit to MIT in 1966, one may presume that he retained a quiescent interest in the use of isentropic coordinates in combination with the exchange of mass and angular momentum. Just prior to his death, he was one of the very few scientists who contacted the author regarding initial results on isentropic analysis concerning the dependency of the partitioning of mass and angular momentum exchange on coordinate systems (Johnson and Downey, 1975a and b, 1976).

In a footnote to his discussion of "the physical basis for the general circulation", Starr (1951a) noted the great diversity of views expressed on individual topics in atmospheric science; views that often were in direct conflict. He noted that meteorology was in a stage of development where like "alchemists" we are "still in the process of grouping for a common denominator of unifying principles. He concludes by stating, "We must not relax but rather continue to stumble and grope with more determination. The writer (Starr) has the immutable belief that a proper foundation does exist and will be found."

In considering the challenges of the next half century regarding the planetary circulation and climate, Starr's pointed remarks noted here and in his reply to Palmén are still applicable. Remarkable advances in theory, analysis and prediction have been made. However, in weighing our limited knowledge of how the atmosphere interacts both internally and externally with its geophysical and biophysical environment through nonlinear processes, the challenges of the next half century are equally exciting; most of all to those who carry out their scientific inquiry and inference in the spirit of Professor Starr's unbounded and unfettered quest for new discoveries, with insight alike Professor Lorenz's and with vision akin to Sir Napier Shaw's.

Acknowledgements

I express my sincere appreciation to colleagues who have contributed to these efforts either directly or indirectly during the past two decades. In view of the extended time during which this work has been conducted, I will only name those who have contributed directly to the results for global monsoonal circulations: Professors Lyle Horn and John Dutton, Drs. Robert Gallimore, Martin Hoerling, Bette Otto-Bliesner, Ronald Townsend, Ming-Ying Wei, John Zillman and Mr. Todd Schaack. Other colleagues who contributed indirectly through isentropic diagnostic or numerical studies are evident in the references. Dr. Allen Lenzen and Messrs. Thomas Whittaker and Rich Selin conducted many of the diagnostic calculations. Mr. Eric Nelson assisted in the preparation of the manuscript. Mrs. Judy Mohr and Ms. Vicki Allaback deserve special recognition for their patience and cheerfullness in typing the almost endless revisions of the manuscript.

The permission of Mr. Fredrick Nagel and the American Meteorological Society to include Fig. 1 and assistance of Dr. Chester Newton in providing information for the preparation of Fig. 27 are gratefully acknowledged. I also wish to acknowledge the invaluable work of the scientific centers who prepared GWE Level III data sets and many pleasant and useful scientific exchanges with fellow scientists engaged in GWE research. This research has been supported by grants from the National Science Foundation, National Aeronautics and Space Administration, and the National Oceanographic and Atmospheric Administration. Current grant support is provided by NSF Grants ATM-8602860 and ATM-8517273 and NASA Grant NAG5-81.

LIST OF SYMBOLS

a	Mean radius of the earth
c_p	Specific heat at constant pressure
c_v	Specific heat at constant volume
d	Specific frictional sink of kinetic energy
e	Specific total energy equal to the sum of kinetic, geopotential and internal energies; also includes latent energy in Eqs. (6.1) and (6.2)
Δe	Component of specific total energy
f	Specific property; arbitrary function; Coriolis parameter
g	Acceleration of gravity
g_a	Component of absolute angular momentum along earth's axis of rotation
k	Specific kinetic energy
ℓ	Distance along path of integration defined in $\int()d\ell$
p	Pressure
p_{oo}	Pressure of 1000 mb
q	Specific humidity
s	Specific entropy, see Eq. (6.26)
t	Time
u	Zonal wind component
v	Meridional wind component
x,y,z	Position in Cartesian coordinates
A	Area; available potential energy

LIST OF SYMBOLS (Cont.)

J_η	Jacobian for transformation of vertical coordinate to generalized coordinate, $\left\|\frac{\partial z}{\partial \eta}\right\|$ (notation includes absolute value)
J_θ	Jacobian for transformation of vertical coordinate to potential temperature, $\left\|\frac{\partial z}{\partial \theta}\right\|$ (notation includes absolute value)
J_p	Jacobian for transformation of vertical coordinate to pressure, $\left\|\frac{\partial z}{\partial p}\right\|$ (notation includes absolute value)
L	Latent heat energy
P_θ	Potential vorticity of zonally averaged circulation, see Eq. (10.12)
Q_m	Specific heat addition
R	Gas constant for dry air
T	Temperature; time period of averaging; Torque
ΣT_i	Sum of pressure, friction and inertial torques
Z_a	Absolute vorticity of zonally averaged circulation, see Eq. (10.11)
α	Specific volume
β_1	Ratio of isentropic area-averaged pressure to equilibrium state pressure, see Eq. (7.26)
β_2	Derivative of isentropic area-averaged pressure with respect to equilibrium state pressure, see Eq. (7.136)
δ	Horizontal divergence of the transport of an arbitrary property, f, see Eq. (2.4); incremental value
ε	Efficiency for the generation of a component of total energy

LIST OF SYMBOLS (Cont.)

ε^2 Heating by viscous dissipation equal to

$$(\varepsilon)^2 = \mu\{-2/3(\nabla\cdot U)^2 + 2[(\frac{\partial u}{\partial x})^2 + (\frac{\partial v}{\partial y})^2 + (\frac{\partial w}{\partial z})^2] + [(\frac{\partial u}{\partial y} + \frac{\partial v}{\partial x})^2 + (\frac{\partial w}{\partial y} + \frac{\partial v}{\partial z})^2 + (\frac{\partial u}{\partial z} + \frac{\partial w}{\partial x})^2]\}$$

ζ Vertical component of the curl of the transport of an arbitrary property f, see Eq. (2.5); vertical component of vorticity

η Generalized vertical coordinate

θ Potential temperature

ι Specific internal energy

κ R/c_p

λ Longitude

ϕ Latitude; specific geopotential energy

μ Dynamic coefficient of viscosity

ν Specific *total flow energy* equal to sum of total energy (e) and RT component of work

$\Delta\nu$ Component of specific *total flow energy*

π Exner function $c_p(p/p_{oo})^{\kappa}$

ρ Density

χ Hydrostatic defect, see Eq. (7.126)

LIST OF SYMBOLS (Cont.)

X	Potential function for transport of property, see Eq. (2.2)
ψ	Montgomery stream function equal to dry static energy, $c_pT + \phi$
Ψ	Stream function for transport of property, see Eq. (2.3)
Ψ_θ	Stream function for isentropic temporally, zonally averaged mass circulation, see Eq. (4.3)
Ψ_p	Stream function for isobaric temporally, zonally averaged mass circulation, see Eq. (4.6)
ω	Vertical velocity in isobaric coordinates, dp/dt

<u>Integral quantities</u>

$m(\theta_T,\theta)$	Mass of atmosphere between arbitrary isentropic surface θ and upper surface θ_T, see Eq. (7.28b)
A	Available potential energy
B	Boundary transport plus work, see Eq. (7.153)
C	Constant of integration, see Eqs. (7.52c), (7.63e), (7.104)
D	Frictional sink of kinetic energy, see Eq. (7.3); also frictional sink of total energy and of components of total energy
E	Total energy equal to the sum of kinetic, geopotential and internal energies
E^2	Vertically integrated heating by viscous dissipation
ΔE	Component of total energy, see Eqs. (7.34a), (7.50) and (7.68b)
$\Delta(E_\alpha,E_o)$	Component of total energy $(E_\alpha - E_o)$, see Eqs. (7.72a and e)
G	Generation of energy by heating; also generation of components of total energy by heating

LIST OF SYMBOLS (Cont.)

K	Kinetic energy
M	Mass of the atmosphere
N	Total entropic energy, see Eqs. (7.81a and b)
S	Entropy
S_n	Entropy of equilibrium atmosphere, see Eq. (7.81a)
$T_n\Sigma$	Static entropic energy, see Eqs. (7.81b)
W	Work of expansion and contraction, see Eqs. (7.182) and (7.183)
π	Total potential energy

Subscripts or superscripts

a	Adiabatic component of ω; denotes inertial frame of reference
ag	Ageostrophic
d	Diabatic component of ω
-d	Net heating by all components except viscous dissipation of kinetic energy, see Eq. (5.2)
e	General equilibrium state variable; earth component of g_a; equilibrium energy for constants of integration in Eqs. (7.52c) and (7.63e)
g	Geostrophic
ir	Irreversible process
m	Middle surface
n	Equilibrium state variable for entropic energy
o	Equilibrium state variable for entropy component model of total energy; reference value for definition of entropy, see Eq. (6.26)

LIST OF SYMBOLS (Cont.)

αo	Equilibrium state variable for $T_{\alpha o} = T(\theta, t_o)$ component model of total energy
$o\alpha$	Efficiency for generation, $G((\Delta E_\alpha, \Delta E_o)$ in Eq. (7.69d)
p	Isobaric coordinate system; pressure surface; pressure torque
r	Reversible process; reference state for available potential energy; relative component of g_a; radiative component of energy flux
r_o	Reference state for available potential energy determined by isentropic time-area-averaged pressure distribution, see Eq. (6.50)
s	Earth's surface, sensible heat component of energy flux
s_o	Minimum value of observed potential temperature at earth's surface
t	Time
t(qL)	Quasi-Lagrangian time-average of area integral bounded by Ψ and $\Psi + \delta\Psi$, see Eq. (5.6)
u	Upper surface; upper layer
ℓ	Lower surface; lower layer; total energy including latent energy of phase changes, see Eq. (6.1)
A	Area
F	Frictional torque
I	Inertial torque; denotes inclined surface area on isentropic and absolute angular momentum surfaces in lieu of projected area
T	Top of the atmosphere; top level; transpose of tensor
α	Equilibrium state variable for reversible component of total energy

LIST OF SYMBOLS (cont.)

η	Generalized coodinate system; η surface
θ	Isentropic coordinate system; isentropic surface
λ	Longitude
ν	Total flow energy transport potential or stream function
ρ	Mass transport potential or stream function
X	Irrotational component
Ψ	Rotational component
+	State variable within and integral for domain of atmospheric heating, see Eq. (7.5); also denotes efficiency defined by Bjerknes, see Eq. (7.181)
-	State variable within and integral for domain of atmospheric cooling, see Eq. (7.5)

Vectors

$\underset{\sim}{k}$	Unit vector for vertical coordinate
$\underset{\sim}{n}$	Outer unit normal
$\underset{\sim}{r}$	Radius vector
$\underset{\sim}{F}$	Frictional acceleration
$\underset{\sim}{G}$	Sum of pressure gradient and gravitational accelerations, see Eq. (7.119)
$\underset{\sim}{H}$	Radiative, sensible and latent energy flux, see Eq. (6.2)
$\underset{\sim}{U}$	Velocity
$\underset{\sim}{\Omega}$	Earth's angular velocity
$\underset{\sim}{T}_n$	Stress vector equal to $\underset{\sim}{n} \cdot \underset{\approx}{\pi}$

LIST OF SYMBOLS (Cont.)

Tensors

$\underset{\approx}{\Pi}$	Stress tensor equal to $(-p\underset{\approx}{I} + \underset{\approx}{\gamma})$
$\underset{\approx}{\gamma}$	Deviatoric tensor equal to $2u[\underset{\approx}{\Phi} - 1/3(\nabla\cdot\underset{\sim}{U})\underset{\approx}{I}]$
$\underset{\approx}{\Phi}$	Strain tensor equal to $1/2[\nabla\underset{\sim}{U} + (\nabla\underset{\sim}{U})^T]$
$\underset{\approx}{I}$	Unit tensor

Operators, differential and differentiation

$\overline{f}^{t}$	Time average, $\frac{1}{T}\int_0^T f\,dt$, superscript suppressed in all Sections except 7 and 10
$\overline{f}^{\lambda}$	Zonal average, $2\pi^{-1}\int_0^{2\pi} f d\lambda$, superscript suppressed in Section 10
f'	Isobaric time deviation, $f - \overline{f}^{t}$
f'	Isentropic time deviation, $f - \overline{f}^{t}$
$\overline{f}^{A}$	Area-average $A^{-1}\int_A f\,dA$, superscript suppressed in Section 7
$\overline{f}^{\lambda t}$	Zonal-time-average, $(2\pi T)^{-1}\int_0^T\int_0^{2\pi} f\,d\lambda dt$
$\overline{f}^{tA}$	Time-area-average, $(AT)^{-1}\int_0^T\int_A f\,dAdt$

LIST OF SYMBOLS (Cont.)

$\hat{f}^{t}$	Mass-weighted time-average, $\overline{\rho J_\theta f}^{t} / \overline{\rho J_\theta}^{t}$, superscript suppressed in all Sections except 10
$\hat{f}^{\lambda}$	Mass-weighted zonal-average $\overline{\rho J_\theta f}^{\lambda} / \overline{\rho J_\theta}^{\lambda}$; superscript suppressed in Section 10
$\hat{f}^{A}$	Mass-weighted area-average $\overline{\rho J_\theta f}^{A} / \overline{\rho J_\theta}^{A}$
$\hat{f}^{V}$	Mass-weighted volume-average, $\int_{V_\theta} \rho J_\theta f dV_\theta / \int_{V_\theta} \rho J_\theta dV_\theta$
$\hat{f}^{\theta}$	Mass-weighted vertical-average, $\int_{\theta_s}^{\theta_T} \rho J_\theta f d\theta / \int_{\theta_s}^{\theta_T} \rho J_\theta d\theta$
$\hat{f}^{t,A}$	Mass-weighted time-area-average $\overline{\rho J_\theta f}^{tA} / \overline{\rho J_\theta}^{tA}$
$\hat{f}^{\lambda\eta t}$	Mass-weighted zonal-vertical-time-average $\int_{\eta_s}^{\eta_T} \overline{\rho J_\eta f}^{\lambda t} d\eta / \int_{\eta_s}^{\eta_T} \overline{\rho J_\eta}^{\lambda t} d\eta$
f^*	Time deviation, $f - \hat{f}$, except in Chapter 10; zonal deviation in Section 10; also deviation of area average from vertical average in Section 7 and of zonal average from vertical-zonal average in Section 10.
f^{**}	Time-area deviation, $f - \hat{f}^{tA}$, in Section 6; area deviation, $f - \hat{f}^{A}$, in Section 7
f^{***}	Volume deviation, $f - \hat{f}^{V}$, in Section 7
~	Appropriate value of a function in the mean value sense in region of heating and cooling, see Eqs. (6.25) and (7.184b)
•	Substantial derivative

LIST OF SYMBOLS (Cont.)

: double-dot product defined for multiplication of tensors, e.g.,

$$\underset{\approx}{\pi} : \nabla \underset{\sim}{U} = \sum_{i=1}^{3} \sum_{j=1}^{3} \pi_{ij} (\nabla U)_{ij}$$

where π_{ij} and $(\nabla U)_{ij}$ are the scalar values associated with nine dyad elements $\underset{\sim\sim}{ii}$, $\underset{\sim\sim}{ij}$, $\underset{\sim\sim}{ik}$, $\underset{\sim\sim}{ji}$, . . ., $\underset{\sim\sim}{kk}$, of the respective tensors $\underset{\approx}{\pi}$ and $\nabla \underset{\sim}{U}$

$\frac{\partial}{\partial t_\theta}$ Local time derivative in isentropic coordinates

$\frac{\partial}{\partial t_p}$ Local time derivative in isobaric coordinates

$\frac{\partial}{\partial t_\eta}$ Local time derivative in generalized coordinates

dV_θ Incremental volume $dAd\theta$

dV_η Incremental volume $dAd\eta$

$d\sigma_I(\theta)$ Incremental area on isentropic surface, subscript denotes arbitrary inclination of surface

$d\sigma_I(g_a)$ Incremental area of absolute angular momentum surface, subscript denotes arbitrary inclination of surface

∇ Three-dimensional del operator

∇_θ Isentropic del operator, $\underset{\sim}{i}\partial(\ )/\partial x_\theta + \underset{\sim}{j}\partial(\ )/\partial y_\theta$

∇_p Isobaric del operator, $\underset{\sim}{i}\partial(\ )/\partial x_p + \underset{\sim}{j}\partial(\ )/\partial y_p$

∇^2_θ Isentropic Laplacian operator, $\partial^2(\ )/\partial x^2_\theta + \partial^2(\ )\partial y^2_\theta$

LIST OF SYMBOLS (Cont.)

∇_η General quasi-horizontal del operator, $\underset{\sim}{i}\partial(\ )/\partial x_\eta + \underset{\sim}{j}\partial(\ )/\partial y_\eta$

$\int_{\delta A}(\ )dA$ Area integral over incremental channel bounded by stream function Ψ_f and $\Psi_f + \delta\Psi_f$, see Eq. (5.3)

$\int(\ )d\ell$ Line integral about incremental channel bounded by stream function Ψ_f and $\Psi_f + \delta\Psi_f$, see Eq. (5.4)

Vector-tensor transformation (Johnson, 1980)

$$J_\theta\nabla(\ ) = \nabla_\theta[J_\theta(\ )] + \frac{\partial}{\partial\theta}[(\underset{\sim}{k}-\nabla_\theta z)(\ )]$$

$$J_\theta\nabla_z(\ ) = \nabla_\theta[J(\ )] - \frac{\partial}{\partial\theta}[(\nabla_\theta z)(\ )]$$

$$J_\theta\nabla\cdot(\ ) = \nabla_\theta\cdot[J_\theta(\ )] + \frac{\partial}{\partial\theta}[\underset{\sim}{k}-\nabla_\theta z)\cdot(\ )]$$

$$J_\theta\nabla_z\cdot(\ ) = \nabla_\theta\cdot[J_\theta(\ )] - \frac{\partial}{\partial\theta}[(\nabla_\theta z)\cdot(\ )]$$

REFERENCES

Anderson, D. T., 1983: The oceanic general circulation and its interaction with the atmosphere. In: "Large-scale Dynamical Processes in the Atmosphere". B. Hoskins and R. Pearce (Eds.), Academic Press, New York, 305-336.

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

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

Andrews, D. G., and McIntyre, M. E. (1978b). An exact theory of nonlinear waves on a Lagrangian-mean flow. J. Fluid Mech., 89, 609-646.

Anthes, R. A. (1971). A numerical model of the slowly varying tropical cyclone in isentropic coordinates. Mon. Weather Rev., 99, 617-635.

Arakawa, A., and Schubert, W. H. (1974). Interaction of a cumulus cloud ensemble with the large-scale environment, Part I. J. Atmos. Sci., 31, 674-701.

Batchelor, G. K. (1967). "An Introduction to Fluid Dynamics". Cambridge University Press, 615 pp.

Bigelow, F. (1903). The mechanism of countercurrent of different temperatures in cyclones and anticyclones. Mon. Weather Rev., 31, 72-84.

Bjerknes, V. (1916). Über thermodynamische Maschinen die unter Mitwirkung der Schwerkraft arbeiten. Abh. K. Sächs. Gesellsch. Wissensch., Leipzig, math.-phys. Kl, 30, 37-65.

Bjerknes, V., Bjerknes, J., Solberg, H., and Bergeron, T. (1933). "Physikalische Hydrodynamik mit Anwendung auf die Dynamische Meteorologie". Verlag Von Julius Springer, Berlin, 797 pp.

Bjerknes, V., Bjerknes, J., Solberg, H., and Bergeron, T. (1937). "Hydrodynamique Physique Avec Applications à la Météorologie Dynamique." Paris: Les Presses Universitaires, Paris 864 pp.

Blasius, W. (1975). "Storms, their Nature, Classification and Laws." Porter and Coates, Philadelphia, Pennsylvania, 342 pp.

Bleck, R. (1984). On the conversion between mean and eddy components of potential and kinetic energy in isentropic and isopynic coordinates. Atmos. Oceans Dyn., 9, 17-37.

Boyle, J. S. and Chen, T.-J. (1987). Synoptic aspects of wintertime East Asian monsoon. In: "Monsoon Meteorology", (C.-P. Chang and T. N. Krishnamurti, Eds.), Oxford University Press, 125-160.

Bridgman, P. W. (1961). "The Nature of Thermodynamics," Harper and Brothers, New York, pp. 239.

Bryan, K., 1983: Poleward heat transport by the oceans. Rev. of Geophys. and Space Phys., 21, 1131-1137.

Campbell, G. G., and Vonder Haar, T. H. (1980). "Climatology of Radiation Budget Measurements from Satellites," Atmos. Sci. Pap. No. 323., Dept. Atmos. Sci., Colorado State University, Fort Collins.

Carathéodory, C. (1909). Untersuchungen über die Grundlagen der Thermodynamik. Mathematische Annalen, 67, 355-386.

Carathéodory, C. (1925). Über die Bestimmung der Energie und der absoluten Temperatur mit Hilfe von reversiblen Prozessen. Preussiche Akademie der Wissenschaften zu Berlin Sitzungsberichte, 39-47.

Carissimo, B. C., Oort, A. H., and Vonder Harr, T. H. (1985). Estimating the meridional energy transports in the atmosphere and ocean. J. Phys. Oceanogr., 15, 82-91.

Chandrasekhar, S. (1939). "An Introduction to the Study of Stellar Structure". University of Chicago Press, Chicago, 509 pp. Reprinted by Dover Publications, 1957.

Chang, C.-P., Erickson, J. E., and Lau, K. M. (1979). Northeasterly cold surges and near-equatorial disturbances over the Winter MONEX area during December 1974. Part I: Synoptic aspects. Mon. Weather Rev., 107, 812-829.

Chang, C.-P., and Lau, K. M. (1980). Northeasterly cold surges and near equatorial disturbances over the Winter MONEX area during December 1974. Part II: Planetary-scale aspects. Mon. Weather Rev., 108, 298-312.

Chang, C.-P., and Lau, K. M. (1982). Short-term planetary-scale interactions over the tropics and midlatitudes during northern winter. Part I: Contrast between active and inactive periods. Mon. Weather Rev., 110, 933-946.

Charney, J. G. (1947). The dynamics of long waves in a baroclinic westerly current. J. Meteorol., 4, 135-162.

Charney, J. G. (1971). Geostrophic Turbulence. J. Atmos. Sci., 28, 1087-1095.

Charney, J. G. and DeVore, J. G. (1979). Multiple flow equilibria in the atmosphere and blocking. J. Atmos. Sci., 36, 1205-1216.

Charney, J. G., and Drazin, P. G. (1961). Propagation of planetary waves in the atmosphere. J. Geophys. Res., 66, 83-109.

Curry, J. H. and Winsand, D. (1986). Low-order intermediate models: bifurcation, recurrence and solvability. J. Atmos. Sci., 43, 2360-2373.

Danielson, E. F. (1974). The relationship between severe weather, major dust storms and rapid large-scale cyclogenesis (II). In: "Subsynoptic Extratropical Weather Systems: Observations, Analysis, Modeling, and Prediction, Volume II, Colloquium notes of the Advanced Studies Program and Small-Scale Analysis and Prediction Project, National Center for Atmospheric Research, Boulder, CO [NTIS PB-247286], 226-241.

Danielsen, E. F. (1961). Trajectories, isobaric, isentropic and actual. J. Meteorol. 19, 479-486.

Danielsen, E. F. (1968). Stratospheric-tropospheric exchange based on radioactivity, ozone and potential vorticity, J. Atmos. Sci., 25, 502-518.

Danielsen, E. F. (1981). An objective method for determining the generalized transport tensor for two-dimensional Eulerian models. J. Atmos. Sci., 38, 1319-1339.

Downey, W. K. (1971). The angular momentum budget of a mid-latitude cyclone using a quasi-Lagrangian coordinate system. M.S. Thesis, Dept. of Meteorol., University of Wisconsin, Madison, 53 pp.

Downey, W. K. (1972). The dynamics of the extratropical cyclone--An angular momentum approach. Ph.D. thesis, Dept. of Meteorol., University of Wisconsin, Madison, 254 pp.

Dutton, J. A. (1973). The global thermodynamics of atmospheric motion. Tellus, 25, 89-110.

Dutton, J. A. (1976a). "The Ceaseless Wind". McGraw-Hill, New York, 579 pp.

Dutton, J. A. (1976b). The nonlinear quasi-geostrophic equation. Part II: Predictability, recurrence, and limit properties of thermally-forced and unforced flows. J. Atmos. Sci., 33, 1431-1453.

Dutton, J. A., and Johnson, D. R. (1967). The theory of available potential energy and a variational approach to atmospheric energetics. Adv. Geophys., 12, Academic Press, New York, 333-436.

Eady, E. T. (1949). Long waves and cyclone waves. Tellus, 1, 3, 35-52.

Edmon, H. J., Hoskins, B. J. and McIntyre, M. E. (1980). Eliassen-Palm cross sections for the troposphere. J. Atmos. Sci., 37, 2600-2615.

Eliassen, A. (1949). The quasi-static equations of motion with pressure as the independent variables. Geofys. Publ. 17, 5-44 pp.

Eliassen, A. (1951). Slow thermally or frictionally controlled meridional circulation in a circular vortex. Astrophys. Norv., 5, 19-60.

Eliassen, A. (1960). Remarks on the problem of long-range weather prediction. In: "Dynamics of Climate" (R. L. Pfeffer, ed.) Pergamon Press, New York, 33-37.

Eliassen, A., and Palm, E. (1961). On the transfer of energy in stationary mountain waves, Geofys. Publ., 22, 1-23.

Eliassen, A. (1986). A method pioneered by Jerome Namias: Isentropic analysis and its aftergrowth. In: Namias Symposium (J. O. Roods, ed.) Scripps Institution of Oceanography Reference Series 86-17, 70-81.

Exner, F. M. (1925). "Dynamische Meteorologie", J. Springer, Vienna, 421 pp.

Fein, J. S. and Stephens, P. L. (eds.) (1987). "Monsoons". John Wiley and Sons, New York, 632 pp.

First GARP Global Experiment (FGGE) Advisory Panel (1985). "Proceedings of the First National Workshop in the Global Weather Experiment: Current Achievements and Future Directions", Volumes I & II., Nat. Acad. of Sci., 75 pp. and 807 pp.

Gall, R. L. (1972). Prediction of a quasi-steady propogating jet core with an isentropic numerical model. Ph.D. thesis, Dept. of Meteorol., University of Wisconsin, Madison, 121 pp.

Gallimore, R. G. (1973). A diagnostic model of the zonally averaged circulation in isentropic coordinates. Ph. D. thesis, Dept. of Meteorol., University of Wisconsin, Madison, 294 pp.

Gallimore, R. G., and Johnson, D. R. (1977a). The forcing of the meridional circulation of the isentropic zonally averaged circumpolar vortex. In: "Isentropic Numerical Models: Results on Model Development for Zonally Averaged and Secondary Circulations", Scientific Report to the National Science Foundation, Dept. of Meteorol. and Space Science and Engineering Center, University of Wisconsin, Madison, [NTIS PB283480/AS], 89-126.

Gallimore, R. G. and Johnson, D. R. (1977b). A numerical diagnostic model of the zonally averaged circulation in isentropic coordinates. In: "Isentropic Numerical Models: Results on Model Development for Zonally Averaged and Secondary Circulations", Scientific Report to the National Science Foundation, Dept. of Meteorol. and Space Science and Engineering Center, University of Wisconsin, Madison, [NTIS PB283480/AS], 127-178.

Gallimore, R. G., and Johnson, D. R. (1981a). The forcing of the meridional circulation of the isentropic zonally averaged circumpolar vortex. J. Atmos. Sci., 38, 583-599.

Gallimore, R. G., and Johnson, D. R. (1981b). A numerical diagnostic model of the zonally averaged circulation in isentropic coordinates. J. Atmos. Sci., 38, 1870-1890.

Gates, W. L. (1976). The numerical simulation of ice-age climate with a global general circulation model. J. Atmos. Sci., 33, 1844-1873.

Gill, A. E. (1982). "Atmosphere-Ocean Dynamics". Academic Press, New York, 662 pp.

Godske, C. L., (1936). Note on an apparent discrepancy between Bjerknes-Sandström and H. Jeffreys. Q. J. R. Meteorol. Soc., 62, 446-449.

Godske, C. L., Bergeron, J., Bjerknes, J., and Bundgaard, R. C. (1957). "Dynamic Meteorology and Weather Forecasting". Amer. Meteorol. Soc., Boston, and Carnegie Institution of Washington, Washington, D.C., 800 pp.

Hadley, G. (1735). Concerning the cause of the general trade-winds. Phil. Trans., 29, 58-62.

Hahn, D. G. and Manabe, S. (1975). The role of mountains in the South Asian monsoon circulation. J. Atmos. Sci., 32, 1515-1541.

Halley, E. (1686). An historical account of the trade winds and monsoon observable in the seas between and near the tropics with an attempt to assign a physical cause of the said winds. Phil. Trans. Roy. Soc. London, 26, 153-68.

Haltiner, G. J. and Martin, T. L. (1957). "Dynamical and Physical Meteorology". McGraw Hill. New York, 470 pp.

Hastenrath, S. (1982). On meridional heat transports in the World Ocean. J. Phys. Oceanogr., 12, 992-927.

Hastenrath, S. (1985). "Climate and Circulation of the Tropics". Reidel Publishing Co., Dordrecht, Holland, 455 pp.

Haynes, P. H. and McIntyre, M. E. (1987). On the evolution of vorticity and potential vorticity in the presence of diabatic heating and fictional or other forces. J. Atmos. Sci., 44, 828-844.

Helmholtz, H. v. (1868). Über discontinuirliche Flüssigkeitsbewegungen. Monatsber. Kön. Akad. Wiss., Berlin, 215-228. (Transl. by Abbe, C. (1893). "The Mechanics of the Earth's Atmosphere". Smithsonian Inst., 58-66).

Helmholtz, H. v. (1888). Über atmosphärische Bewegungen. Meteorologische Zeitung, 5, 329-340. (Transl. by Abbe, C. (1893). The Mechanics of the Earth's Atmosphere", Smithsonian Inst., Washington, D.C., 78-93.)

Henderson, H. W. (1971). The atmospheric energy budget in isentropic coordinates. M.S. thesis, Dept. of Meteorol., The Pennsylvania State University, University Park, 78 pp.

Hide, R., and Mason, P. J. (1975). Sloping convection in a rotating fluid. Adv. Phys., 24, 47-100.

Hill, H. W. (1981). A synoptic case study using the FGGE buoy data of a typical low index regime in the Western South Pacific Ocean. ICSU/WMO GARP Conv. Proc. of Int. Conf. on Preliminary FGGE Data Analysis and Results (23-27 June 1980, Bergen, Norway), World Meteorol. Organ., Geneva, 450-464.

Hoerling, M. P. (1987). The global balance of dynamic vorticity in relation to heat sources and sinks. Ph.D. Thesis, Dept. of Meteorol., University of Wisconsin, Madison, 246 pp.

Hoerling, M. P., and Johnson, D. R. (1986). On the maintenance of the semi-permanent centers of action during January 1979. National Conference on the Scientific Results of the First GARP Global Experiment, (January 14-17, 1986, Miami, Florida). American Meteorol. Soc., Boston, 91-94.

Holloway, G. and West, B. J. (eds.), (1984). "Predictability of Fluid Motions (La Jolla Institute - 1983)". Amer. Inst. Phys., New York, 612 pp.

Holopainen, E. O. (1983). Transient eddies in mid-latitudes: Observations and interpretation. In: "Large Scale Dynamical Processes in the Atmosphere", (B. J. Hoskins and R. P. Pearce, Eds.), Academic Press, London, 201-233.

Holton, J. R. (1972). "An Introduction to Dynamic Meteorology". Academic Press, New York, 219 pp.

Hoskins, B. J. (1983). Modeling of the transient eddies and their feedback on the mean flow. In: "Large-Scale Dynamical Processes in the Atmosphere", (B. J. Hoskins and R. P. Pearce, Eds.) , Academic Press, London, 169-199.

Hoskins, B. and Pearce R. P. (eds.) (1983). "Large-Scale Dynamical Processes in the Atmosphere". Academic Press, New York, 397 pp.

Hoskins, B. J., McIntyre, M. D. and Robertson, A. W. (1985). On the use and significance of isentropic potential vorticity maps. Q. J. R. Meteorol. Soc., 111, 877-848.

Imbrie, J., and Imbrie K. P. (1979). "Ice Ages: Solving the Mystery". Enslow Publishers, 224 pp.

James, I. N. (1983). Some aspects of the global circulation of the atmosphere in January and July 1980. In: "Large-Scale Dynamical Processes in the Atmosphere", (B. J. Hoskins and R. P. Pearce, Eds.), Academic Press, London, 5-25.

Jeffreys, H. (1925). On fluid motions produced by differences of temperature and humidity. Q. J. R. Meteorol. Soc., 51, 347-356.

Jeffreys, H. (1926). On the dynamics of geostrophic winds. Q. J. R. Meteorol. Soc., 52, 85-104.

Jeffreys, H. (1936). Note on Dr. Godske's paper, Q. J. R. Meteorol. Soc., 62, 449-450.

Johnson, D. R. (1970). The available potential energy of storms. J. Atmos. Sci., 27, 727-741.

Johnson, D. R. (1979). Systematic stratospheric-tropospheric exchange through quasi- horizontal transport processes within active baroclinic waves. WMO Symp. on the Long-Range Transport of Pollutants and Its Relation to General Circulation Including Stratospheric/Tropospheric Exchange Processes, (October 1-5, 1979, Sofia, Bulgaria), WMO No. 538, World Meteorol. Organ., Geneva, 401-408.

Johnson, D. R. (1980). A generalized transport equation for use with meteorological coordinate systems. Mon. Weather Rev., 108, 733-745.

Johnson, D. R. (1983a). The seasonal variation of the angular momentum balance of the zonally-averaged circulation of the Southern Hemisphere. Preprint Volume for First Conf. on Southern Hemisphere Meteorology (31 July - 6 August, 1983, San Jose Dos Campos, Brazil), Amer. Meteorol. Soc., Boston, 19-22.

Johnson, D. R. (1983b). On the forcing and maintenance of the isentropic zonally averaged circumpolar vortex. IAMAP-WMO Symposium on Maintenance of the Quasi-Stationary Components of the Flow in the Atmosphere and in Atmospheric Models (30 August - 6 September, 1983, Paris, France), World Meteorol. Organ., Geneva, 19-22.

Johnson, D. R. (1984). Global monsoonal circulations forced by the planetary scale of differential heating. Proceedings of the Symposium of the I.U.G. Working Group on Tropical Climatology and Human Settlements (August 21-25, 1984, Dijon, France), Centre de Recherches de Climatologies, Faculte des Sciences Humaines, University de Dijon, Dijon, France, 5-26.

Johnson, D. R. (1985a). The global circulation during the FGGE year: On the balance of mass, energy, and angular momentum within isentropic and isobaric coordinates as revealed by different FGGE data sets. ICSU/WMO Report of the Seminar on Progress in Diagnostic Studies of the Global Atmosphere Circulation as a Result of the Global Weather Experiment (August 28-31, 1984, Helsinki, Finland), GARP Special Report No. 42, World Meteorol. Organ., Geneva, I 1-15.

Johnson, D. R. (1985b). On the global distribution of heat sources and sinks and their relation to mass and energy transport. ICSU/WMO Report of the Seminar on Progress in Tropical Meteorology as a Result of the Global Weather Experiment (October 8-12, 1984, Tallahassee, Florida), GARP Special Report No. 44, World Meteorol. Organ., Geneva, III 1-13.

Johnson, D. R. (1985c). On the planetary diabatic differential heating and energy exchange inferred from GWE level IIIb analyses. Preprint Volume for Seventh Conference on Numerical Weather Prediction (June 17-20, 1985, Montreal, Canada), Amer. Meteorol. Soc., Boston, 504-510.

Johnson, D. R. (1986). Thermally forced energy transport and kinetic energy generation in relation to planetary monsoonal circulations. Preprint Volume for National Conference on the Scientific Results of the First GARP Global Experiment, (January 14-17, 1986, Miami, Florida), Amer. Meteorol. Soc., Boston, 155-158.

Johnson, D. R., and Dutton, J. A. (1969). Atmospheric energetics and the general circulation viewed from isentropic coordinates. London Conference on The Global Circulation of the Atmosphere, (August 25-29, 1969, London). Bull. Amer. Meteorol. Soc., 50, 385.

Johnson, D. R., and Downey, W. K. (1975a). Azimuthally averaged transport and budget equations for storms: Quasi-Lagrangian diagnostics 1. Mon. Weather Rev., 103, 967-979.

Johnson, D. R., and Downey, W. K. (1975b). The absolute angular momentum of storms: Quasi-Lagrangian diagnostics 2. Mon. Weather Rev., 103, 1063-1076.

Johnson, D. R., and Downey, W. K. (1976). The absolute angular momentum budget of an extratropical cyclone: Quasi-Lagrangian diagnostics 3. Mon. Weather Rev., 104, 3-14.

Johnson, D. R., and Townsend, R. D. (1981). Diagnostics of the heat sources and sinks of the Asiatic monsoon and the thermally-forced planetary scale response. ICSU/WMO GARP Conference Proceedings, Proceedings of International Conference on Preliminary FGGE Data Analysis and Results, Data Analysis and Results, (23-27 June, 1980, Bergen, Norway), A. Grammeltvedt, Ed.), World Meteorol. Organ., Geneva, 523-533.

Johnson, D. R., Townsend, R. D., Wei, M.-Y. and Selin, R. D. (1981). The global structure of energy transport within thermally-forced planetary scale mass circulations. ICSU/WMO GARP Condensed Papers and Meeting Report of Aspects of its Monsoon Experiments, (12-17 January, 1981, Tallahassee, Florida), World Meteorol. Organ., Geneva, (3-27)-(3-34).

Johnson, D. R., and Downey, W. K. (1982). On the energetics of open systems. Tellus, 34, 458-470.

Johnson, D. R., Wei, M.-Y. and Schaack, T. (1982). Planetary-scale mass, momentum and energy exchange. ICSU/WMO GARP Extended Abstracts and Panel Session of the Int. Conf. on Scientific Results of the Monsoon Experiment (October 26-30, 1981, Bali, Indonesia), World Meteorol. Organ., Geneva, 2-40 - 2-43.

Johnson, D. R., Townsend, R. D. and Wei, M.-Y. (1985). The thermally coupled response of the planetary scale circulation to the global distribution of heat sources and sinks. Tellus, 37A, 106-125.

Johnson, D.R., and Wei, M.-Y. (1985). The planetary distribution of heat sources and sinks during FGGE. FGGE Workshop, Woods Hole, Massachusetts, July 9-20, 1984. Proceedings of the First National Workshop on the Global Weather Experiment: Current Achievements and Future Directions, Natl. Acad. Sci., Vol. II, Part I, 299-316.

Johnson, D. R., Yanai, M. and Schaack, T. K. (1986). Global and regional distributions of atmospheric heat sources and sinks during the GWE. In: "Monsoon Meteorology", (C.-P. Chang and T. N. Krishnamurti, Eds.), Oxford University Press, 271-297.

Kida, H. (1977). A numerical investigation of the atmospheric general circulation and stratospheric-tropospheric mass exchange: II. Lagrangian motion of the atmosphere. J. Meteorol. Soc. Japan, 55, 71-88.

Krishnamurti, T.N. (1961). The subtropical jet stream of winter. J. Meteorol., 18, 172-191.

Krishnamurti, T. N. and Ramanthan, Y. (1982). Sensitivity of the monsoon onset to differential heating. J. Atmos. Sci., 39, 1290-1306.

Krishnamurti, T. N. (1985). Summer monsoon experiment - a review. Mon. Weather Rev., 113, 1590-1626.

Krishnamurti, T. N. (1987). Monsoon models. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 467-522.

Krishnamurti, T. N. and Surgi, N. (1987). Observational aspects of summer monsoon. In: "Monsoon Meteorology", (C.-P. Chang and T. N. Krishnamurti, Eds.), Oxford University Press, 3-25.

Krishnamurti, T. N. Low-Nam, S., Kumar, A., Sheng, J. and Sugi, M. (1987). Numerical weather prediction of monsoons. In: "Monsoon Meteorology", (C.-P. Chang and T. N. Krishnamurti, Eds.), Oxford University Press, 501-544.

Kuo, H.-L. (1949). Dynamic instability of two-dimensional nondivergent flow in a barotropic atmosphere. J. Meteorol., 6, 105-122.

Kuo, H.-L. (1956). Forced and free meridional circulations in the atmosphere. J. Meteorol., 13, 561-568.

Kuo, H.-L. (1957). Further studies of thermally driven motions in a rotating fluid, J. Meteorol., 14, 553-558.

Kuo, H.-L. (1960). Theoretical findings concerning the effects of heating and rotation on the mechanism of energy release in rotating fluid systems. In: "Dynamics of Climate" (R. L. Pfeffer, ed.) Pergamon Press, New York, 72-78.

Kutzbach, G. (1987). Concepts of monsoon physics in historical perspective: the Indian Monsoon (seventeenth to early twentieth century). In "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 159-210.

Kutzbach, J. E. (1987). The changing pulse of the monsoon. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 247-268.

Kutzbach, J. E. and Otto-Bliesner, B. L. (1982). The sensitivity of the African-Asian monsoonal climate to orbital parameter changes for 9000 Years B. P. in a low-resolution general circulation model. J. Atmos. Sci., 39, 1177 - 1188.

Lamb, H. (1932). "Hydrodynamics". Cambridge University Press. pp. 738.

Lau, N.-C. (1978). On the three-dimensional structure of the observed transient eddy statistics of the Northern Hemisphere wintertime circulation. J. Atmos. Sci., 35, 1900-1923.

Lau, N.-C. (1979a). The structure and energetics of transient disturbances in the Northern Hemisphere wintertime circulation. J. Atmos. Sci., 36, 982-995.

Lau, N.-C. (1979b). The observed structure of tropospheric stationary waves and the local balances of vorticity and heat. J. Atmos. Sci., 36, 999-1016.

Lau, N.-C., and Oort, A. H. (1981). A comparative study of observed Northern Hemisphere circulation statistics based on GFDL and NMC analyses. Part I: The time-mean fields. Mon. Weather Rev., 109, 1380-1403.

Lau, N.-C., and Oort, A. H. (1982). A comparative study of observed Northern Hemisphere circulation statistics based on GFDL and NMC analyses. Part II: Transient eddy statistics and the energy cycle. Mon. Weather Rev., 109, 889-906.

Lettau, H. (1954a). A study of the mass, momentum and energy budget of the atmosphere. Archiv. Meteor. Geophys. Bioklima., A, 7, 133-157.

Lettau, H. (1954b). Notes on the transformation of mechanical energy from eddying motion. J. Meteorol., 11, 196-201.

Lighthill, J. and Pearce, R. P. (eds.) (1981). "Monsoon Dynamics". Cambridge University Press, London, 735 pp.

Livezey, R. E., and Dutton, J. A. (1976). The entropic energy of geophysical fluid systems. Tellus, 28, 138-158.

Lorenz, E. N. (1955a). Available potential energy and the maintenance of the general circulation. Tellus 7, pp. 157-167.

Lorenz, E. N. (1955b). Generation of available potential energy and the intensity of the general circulation. In: "Large Scale Synoptic Processes" (J. Bjerknes, Proj. Director). Final Report, 1957, Contract AF19(604)-1286, Dept. Meteorol. University of California, at Los Angeles, Los Angeles, California. 36 pp.

Lorenz, E. N. (1960). Generation of available potential energy and the intensity of the general circulation. In: "The Dynamics of Climate" (R. L. Pfeffer, Ed.), Pergamon Press, New York, 86-92.

Lorenz, E. N. (1967). "The Nature and Theory of the General Circulation of the Atmosphere". WMO, Geneva, No. 218, TP115, 161 pp.

Lorenz, E. N. (1978). Available energy and the maintenance of a moist circulation. Tellus, 30. 15-31.

Lorenz, E. N. (1980). Attractor sets and quasi-geostrophic equilibrium. J. Atmos. Sci., 1685-1699.

Ludlam, F. H. (1966). "The Cyclone Problem: A History of Models of the Cyclonic Storm", Inaugural Lecture, Imperial College of Science and Technology, London. 49 pp.

Luo, H. and Yanai, M. (1983). The large-scale circulation and heat sources over the Tibetan Plateau and surrounding areas during the early summer of 1979. Part I: Precipitation and kinematic analysis. Mon. Weather Rev., 111, 922-944.

Luo, H. and Yanai, M. (1984). The large-scale circulation and heat sources over the Tibetan Plateau and surrounding areas during the early summer of 1979. Part II: Heat and moisture budgets. Mon. Weather Rev., 112, 966-989.

Mak, M. (1985). Equilibration of nonlinear baroclinic instability. J. Atmos. Sci., 42, 2764-2782.

Manabe, S. and Hahn, D. G. (1977). Simulation of the tropical climate of an Ice Age. J. Geophys. Res., 82, 3889-3911.

Margenau, H. and Murphy, G. M. (1943). "The Mathematics of Physics and Chemistry". D. Van Nostrand Company, Inc., New York, 581 pp.

Margules, M. (1903). Uber die Energie der Stüme. Jahrb. Zentralanst. Meteorol., 40, (Transl. by Abbe, C., (1910). "The Mechanics of the Earth's Atmosphere". 3rd Coll., Smithsonian Inst., Washington, D.C., 533-595.)

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

Milankovitch, M. (1941). "Canon of insolation and the ice-age problem". Royal Serb. Acad. Spec. Publ. Vol. 132, 484 pp. [Translated by Israel Program for Scientific Translations, Jerusalem, 1969. Available from the U.S. Dept. of Commerce, Clearinghouse for Federal Scientific and Technical Information, Springfield, VA 22151].

Milne-Thomson, L. M. (1960). "Theoretical Hydrodynamics". The Macmillan Company, New York, 660 pp.

Min, K. D. and Horn, L. H. (1982). Available potential energy in the Northern Hemisphere during the FGGE year. Tellus, 34, 526-539.

Mitchell, K. E. and Dutton, J. A. (1980). Bifurcations from stationary to periodic solutions in a low-order model of forced, dissipative barotropic flow. J. Atmos. Sci., 38, 690-716.

Mood, A. M. (1950). "Introduction to the Theory of Statistics". McGraw-Hill, New York, 433 pp.

Moritz, R. E. and Sutera, A. (1981). The predictability problem: Effects of stochastic perturbations in multiequilibrium systems. Adv. Geophys., 23, Academic Press, New York, 345-383.

Murakami, T. (1987a). Orography and monsoons. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 331-364.

Murakami, T. (1987b). Effects of the Tibetan plateau. In: "Monsoon Meteorology", (C.-P. Chang and T. N. Krishnamurti, eds.), Oxford University Press, 235-270.

Nagle, F. W. (1979). Visualizing a cold air dome. Bull. Amer. Meteorol. Soc. 60, 128.

Namias, J. (1940). "An Introduction to the Study of Air Mass and Isentropic Analysis". Amer. Meteorol. Soc., 232 pp.

Navier, M. (1822). Mémoire sur les lois du movement des fluides. Mem. de l'Acad. des Sciences, 6, 389-440.

Newell, R. E., Vincent, D. G., Dopplick, T. G., Ferruza, D., and Kidson, J. W. (1970). The energy balance of the global atmosphere. In: The Global Circulation of the Atmosphere, (G. A. Corby, ed.), Roy. Meteor. Soc., London, 42-90.

Newton, C. W., and Palmén, E. (1963). Kinematic and thermal properties of a large amplitude wave in the westerlies. Tellus, 15, 99-119.

Nitta, T. (1972). Observational study of heat sources over the eastern Tibetan Plateau during the summer monsoons. J. Meteorol. Soc. Japan, 61, 590-605.

Oort, A. H. (1964). On estimates of the atmospheric energy cycle. Mon. Weather Rev., 92, 483-493.

Oort, A. H., and Vonder Haar, T. H. (1976). On the observed annual cycle in the ocean-atmosphere heat balance over the Northern Hemisphere. J. Phys. Oceanogr. 6, 781-800.

Oort, A. H. and Peixóto, J. P. (1983). Global angular momentum and energy balance requirements from observations. Adv. Geophys., 25, 355-490.

Otto, B.L. (1974). Isentropically time-averaged mass circulations in the Northern Hemisphere. M.S. thesis, University of Wisconsin, Madison, 94 pp.

Otto-Bliesner, B. L., and Johnson, D. R. (1982). Thermally-forced mean mass circulation in the Northern Hemisphere. Mon. Weather Rev., 110, 916-932.

Palmén, E. (1949). Meridional circulations and the transfer of angular momentum in the atmosphere. J. Meteorol., 6, 429.

Palmén, E., and Vuorela, L. A. (1963). On the mean meridional circulations in the Northern Hemisphere during the winter season. Q. J. R. Meteorol. Soc., 89, 131-138.

Palmén, E. and Newton, C. W. (1969). "Atmospheric Circulation Systems". Academic Press, New York, 603 pp.

Paltridge, G. W. (1978). The steady-state format of global climate. Q. J. Meteorol. Soc., 104, 927-945.

Pearce, R. P. (1978). On the concept of available potential energy. Q. J. R. Meteorol., Soc., 104, 737-755.

Panofsky, H. A. (1946). Methods of computing vertical motion in the atmosphere. J. Meteorol., 3, 45-49.

Pedlosky, J. (1970). Finite amplitude baroclinic waves. J. Atmos. Sci., 27, 15-30.

Pedlosky, J. and Polvani, L. M. (1986). Wave-wave interaction of unstable baroclinic waves. J. Atmos. Sci., 44, 631-647.

Pfeffer, R. L., Mardon, D. Sterbenz, P. and Fowlis, W. (1965). A new concept of available potential energy. In: "Dynamics of Large-Scale Atmospheric Processes", Proceedings of the International Symposium on the Dynamics of Large-Scale Processes in the Atmosphere, (June 23-30, Moscow), World Meteorol. Organ., 179-189.

Pfeffer, R. L. (1985). Review of research on wave mean flow interactions using FGGE data. Proceedings of the First National Workshop on the Global Weather Experiment: Current Achievements and Future Directions, Natl. Acad. Sci., Vol. II, Part II, 422-437.

Phillips, N. A. (1956). The general circulation of the atmosphere: A numerical experiment. Q. J. R. Meteorol. Soc., 82, 123-163.

Phillips, N. A. (1960). The general circulation of the atmosphere: numerical experiment. In: "Dynamics of Climate" (R. L. Pfeffer, ed.) Pergamon Press, New York, 13-18.

Plum, R. A. (1983). A new look at the energy cycle. J. Atmos. Sci., 40, 1669-1688.

Ramage, C. (1971). "Monsoon Meteorology". Academic Press, New York, 196 pp.

Rasmusson, E. M. and Carpenter, T. H. (1982). Variations in tropical sea surface temperature and surface wind fields associated with the Southern Oscillation/El Niño. Mon. Weather Rev., 110, 354-384.

Rayleigh, Lord (1880). On the stability, or instability, of certain fluid motions. Proc. London Math. Soc., 9, 57-70. [Also in "Scientific Papers," Cambridge Univ. Press, (1900) 2, 200-207.]

Reiter, E. R. (1972). "Atmospheric Transport Processes Part 3: Hydrodynamic Tracers". U. S. Atomic Energy Commission. 212 pp.

Reynolds, O. (1894). On the dynamical theory of incompresible viscous fluids and the determination of the criterion. London Philos. Trans., 186A, 123-164.

Robinson, G. D. (1967). Some current projects for global meteorological observation and experiment. Q. J. R. Meteorol. Soc., 93, 409-418.

Robinson, G. D. (1977). Weather and climate forecasting as problems in hydrodynamics. Mon. Weather Rev., 106, 448-457.

Rossby, C. G. (1940). Planetary flow patterns in the atmosphere. Q. J. R. Meteorol. Soc., 66 (Suppl.) 68-87.

Rossby, C.-G. (1947). On the distribution of angular velocity in gaseous envelopes under the influence of large-scale horizontal mixing processes. Bull. Amer. Meteorol. Soc., 28, 53-68.

Rossby, C.-G. (1949a). On a mechanism for the release of potential energy in the atmosphere. J. Meteorol., 6, 163-180.

Rossby, C.-G. (1949b). On the nature of the general circulation of the lower atmosphere. "The Atmospheres of the Earth and Planets" (Kuiper, Ed.), Chicago, 16-48.

Rossby, C.-G. (1957). Aktuella Meteorologiska problem. Svensk Naturventenskap 1956, 15-80. (English translation, 1959: Current problems in meteorology. In. "The Atmosphere and Sea in Motion", (B. Bolin, Ed.) Rockefeller Inst. Press, New York, 9-50).

Rossby, C.-G., and Starr, V. P. (1949). Interpretations of the angular-momentum principle as applied to the general circulation of the atmosphere. J. Meteorol., 6, 288.

Saltzman, B. (1983). Climate system analysis. Adv. Geophys. 25, Academic Press, New York, 173-233.

Sandström, J. W. (1915). The origin of the wind. Mon. Weather Rev., 43, 161-163.

Sandström, J. W., (1916). Meteorologische Studien in schevedischen Hochgebirge. Göteborgs K. Vetensnaps-och Vitterhetssamhälles Handl., sec. 4, 22, No. 2, 48 pp.

Schaack, T. (1982). The global angular momentum balance for January 1979, M. S. thesis, University of Wisconsin, Madison, 104 pp.

Shapiro, M. A. (1976). The role of turbulent heat flux in the generation of potential vorticity in the vicinity of upper-level jet stream systems. Mon. Weather Rev., 104, 892-906.

Shapiro, M. A. (1980). Turbulent mixing within tropopause folds as a mechanism for the exchange of chemical constituents between the stratosphere and troposphere. J. Atmos. Sci., 37, 994-1004.

Shaw, Sir Napier (1930). "Manual of Meteorology, The Physical Processes of Weather, Vol. III". Cambridge University Press, 445 p.

Shukla, J. (1985). Predictability. Adv. Geophys., 28B, Academic Press, New York, 87-122.

Shukla, J. (1987). Interannual variability of monsoons. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.), John Wiley and Sons, New York, 399-414.

Smagorinsky, J. (1953). The dynamical influence of large-scale heat sources and sinks on the quasi-stationary mean motions of the atmosphere. Q. J. R. Meteorol. Soc., 79, 342-366.

Smagorinsky, J. (1960). A synopsis of research on quasi-stationary perturbations of the mean zonal circulation caused by topography and heating. In: "Dynamics of Climate" (R. L. Pfeffer, ed.) Pergamon Press, New York, 38-44.

Solberg, H. (1928). Integrationen der atmosphärischen Störungsgleichungen. Geofys. Publ., 5, 120 pp.

Solberg, H. (1930). Das Zyklonenproblem. Proc. 3d Internat. Cong. Appl. Mech. (Stockholm), 121-131 pp.

Sommerfeld, A. (1950). "Thermodynamics and Statistical Mechanics: Lectures on Theoretical Physics", 5, Academic Press, New York, 401 pp.

Starr, V. P. (1945). A quasi-Lagrangian system of hydrodynamical equations. J. Meteorol., 2, 227-237.

Starr, V. P. (1948a). On the production of kinetic energy in the atmosphere. J. Meteorol., 5, 193-196.

Starr, V. P. (1948b). An essay on the general circulation of the earth's atmosphere. J. Meteorol., 5, 39-43.

Starr, V. P. (1949). Reply. J. Meteorol., 6, 430.

Starr, V. P. (1951a). The physical basis of the general circulation. In: "Compendium of Meteorology" (T. Malone, ed.) Am. Meteorol. Soc., Boston, 541-550.

Starr, V. P. (1951b). Applications of energy principles to the general circulation. Compendium of Meteorology, (T. Malone, ed.). Amer. Meteorol. Soc., Boston, 568-576.

Starr, V. P. (1968). "Physics of Negative Viscosity Phenomena". McGraw-Hill, New York, 256 pp.

Starr, V. P. and Saltzman, B. (1966). "Observational Studies of the Atmospheric General Circulation". Planetary Circulations Project, Scientific Report No. 2, Contract No. AF19(628)-5816, Dept. of Meteorol., Massachussetts Institute of Technology, Boston, 700 pp.

Starr, V. P., Peixoto, J. P., and Gaut, N. E. (1970). Momentum and zonal kinetic energy balance of the atmosphere from five years of hemispheric data. Tellus, 22, 251-274.

Stokes, G. G. (1845). On the theories of the internal friction of fluids in motion and of the equilibrium and motion of elastic solids. Cambridge Philos. Soc. Trans. VIII, 287-319.

Stokes, G. G. (1846). On the theories of the internal friction of fluids in motion, and of the equilibrium and motion of elastic solids. Cambridge Philos. Soc. Proc., 1, 16-18.

Tang, M., and Reiter, E. R. (1984). Plateau monsoons of the Northern Hemisphere: A comparison between North America and Tibet. Mon. Weather Rev., 112, 617-637.

Tolman, R. C. and Fine, P. C. (1948). On the irreversible production of entropy. Rev. Mod. Phys., 20, 51-77.

Townsend, R. D. (1980). A diagnostic study of the zonally-averaged global circulation in isentropic coordinates. Ph.D. thesis, University of Wisconsin, Madison, 221 pp.

Townsend, R. D., and Johnson, D. R. (1981). The mass and angular momentum balance of the zonally-averaged global circulation. ICSU/WMO GARP Conf. Proc. of Int. Conf. on Preliminary FGGE Data Analysis and Results, (23-27 June 1980, Bergen, Norway), (A. Grammelwedt, Ed.), World Meteorol. Organ., Geneva, 542-552.

Townsend, R.C., and Johnson, D. R. (1985). A diagnostic study of the isentropic zonally averaged mass circulation during the First GARP Global Experiment. J. Atmos. Sci., 42, 1565-1579.

Trenberth, K. E. (1985). Transient eddy activity in relation to the mean flow over the Southern Hemisphere during FGGE. ICSU/WMO Report of the Seminar on Progress in Diagnostic Studies of the Global Atmosphere Circulation as a Result of the Global Weather Experiment, (August 28-31, 1984, Helsinki, Finland), GARP Special Report No. 42, World Meteorol. Organ., Geneva, I, 1-15.

Trenberth, K. E. (1986a). The signature of a blocking espisode on the general circulation in the Southern Hemisphere. J. Atmos. Sci., 19, 2061-1069.

Trenberth, K. E. (1986b). An assessment of the impact of transient eddies on the zonal flow during a blocking episode using localized Eliassen-Palm flux diagnostics. J. Atmos. Sci., 19, 2070-2069.

Tung, K. K. (1984). Modeling of tracer transport in the middle atmosphere. In: "Dynamics of the Middle Atmosphere". (J. R. Holton and T. Matsumo, eds.). Reidel Publishing Company, Tokyo, 417-444.

Tung, K. K. (1986). Nongeostrophic theory of zonally averaged circulation. Part I: Formulation. J. Atmos. Sci., 43, 2600-2618.

Uccellini, L. W. and Johnson, D. R. (1979). The coupling of upper and lower tropospheric jet streaks and implications for the development of severe convective storms. Mon. Weather Rev., 107, 682-703.

Van Mieghem, J. (1951). Hydrodynamic instability. In: "Compendium of Meteorology", (T. Malone, ed.), Am. Meteorol. Soc., Boston, 435-453.

Van Mieghem, J. (1956). The energy available in the atmosphere for conversion into kinetic energy. Beitr. Phys. Atmosphäre, 29, 4-17.

Van Meighem, J. (1973). "Atmospheric Energetics". Clarendon Press, Oxford, 306 pp.

Vickroy, J., and J. Dutton, 1979: Properties of solutions to a spectral model of forced, dissipative, nonlinear quasi-geostrophic flow. J. Atmos. Sci., 36, 42-52.

Vonder Haar, T. H. and Oort, A. H. (1973). New estimates of annual poleward energy transport by Northern Hemosphere Oceans. J. Phys. Oceanogr., 2, 169-172.

Von Neumann, J. (1960). Some remarks on the problem of forecasting climatic fluctuations. In: "Dynamics of Climate" (R. L. Pfeffer, ed.) Pergamon Press, New York, 9-11.

Weatherburn, C. E., (1966). "Advanced Vector Analysis". G. Bell and Sons, 222 pp.

Webster, P. J. (1987a). The elementary monsoon. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 3-32.

Webster, P. J. (1987b). The variable and interactive monsoon. In: "Monsoons" (J. S. Fein and P. L. Stephens, eds.) John Wiley and Sons, New York, 269-330.

Wei, M.-Y. and Johnson, D. R. (1981). The spatial and temporal variation of diabatic heating during the summer monsoon. ICSU/WMO GARP International Conference on the Results of the Monsoon Experiments, (26-30 October, 1981, Denpasar, Bali, Indonesia), World Meteorol. Organ., Geneva, (1-3)-(1-6).

Wei, M.-Y., Johnson, D. R. and Townsend, R. D. (1981). The structure of the planetary scale diabatic processes for the atmosphere during FGGE. ICSU/WMO GARP Condensed Papers and Meeting Report of the International Conference on Early Results of FGGE and Large-scale Aspects of its Monsoon Experiments, (12-17 January, 1981, Tallahassee, Florida), World Meteorol. Organ., Geneva, (3-35)-(3-42).

Wei, M.-Y., Johnson, D. R., and Townsend, R. D. (1983). Seasonal distributions of diabatic heating during the First GARP Global Experiment. Tellus, 35A, 241-255.

Welander, Pierre. (1955). Studies on the general development of motion in a two-dimensional, ideal fluid. Tellus, 7, 141-156.

Widger, W. K. (1949). A study of the flow of angular momentum in the atmosphere, J. Meteorol., 6, 291-299.

Wiin-Nielsen, A., Brown, J. A., and Drake, M. (1963). On atmospheric energy conversion between the zonal flow and the eddies. Tellus, 15, 261-279.

Wilks, S. S. (1962). "Mathematical Statistics". John Wiley and Sons, Inc., New York, 644 pp.

WMO GARP Report (1985a). ICSU/WMO Report of the Seminar on Progress in Diagnostic Studies of the Global Atmosphere Circulation on a Result of the Global Weather Experiment (August 28-31, 1984, Helsinki, Finland), GARP Special Report No. 42, World Meteorol. Organ., Geneva, 238 pp.

WMO GARP Report (1985b). ICSU/WMO Report of the Seminar on Progress in Tropical Meteorology as a Result of the Global Weather Experiment. (October 8-12, 1984, Tallahassee, Florida), GARP Special Report. No. 44, World Meteorol. Organ., Geneva, 464 pp.

WMO GARP Report. (1986). International Conference on the Results of the Global Weather Experiment and Their Implications for the World Weather Watch. GARP Publ. Ser. No. 26, Vol. I & II, World Meteorol. Organ., Geneva, 291 and 670 pp.

Wulf, O. R., and Davis, L. Jr. (1952). On the efficiency of the engine driving the atmosphere circulation. J. Meteorol. 9, 79-82.

Youngblut, C. E., and Sasamori, T. (1980). The nonlinear effects of transient and stationary eddies on the mean winter circulation. Part I: The diagnostic analysis. J. Atmos. Sci., 37, 1944-1957.

Zillman, J. W. (1972). Isentropically time-averaged mass circulations in the Southern Hemisphere. Ph.D. thesis, University of Wisconsin-Madison, 205 pp.

Zillman, J. W., and Johnson, R. D. (1985). Thermally-forced mean circulations in the Southern Hemisphere. Tellus, 37A, 56-76.

Table 1.
Frequency Response of Meridional Filter

Wavelength (km)	% of original amplitude retained
12,000	94
10,000	91
8,000	86
6,000	76
4,000	55
3,000	29
2,500	7

List of Figures

Fig. 7 Same as Fig. 5 except for 260-270 K (a), 270-280 K (b), 280-290 K (c) and 290-300 K (d) isentropic layers for January 1979. Dashed lines represent pressure of bottom level of the isentropic layer (10^2 mb).

Fig. 8 Mass stream functions (10^{10} kg s^{-1}) for mean meridional circulations computed from ECMWF Level III analyses for January 1979, isentropic (a), isobaric (b), isentropic geostrophic component (c) and isentropic ageostrophic component (d).

Fig. 9 Same as 8 except for July 1979.

Fig. 10 Schematics of horizontal (λ,ϕ) and vertical (λ,p) distributions of geopotential and potential temperature within a steady (a) and amplifying (b) wave. In the zonal, vertical cross sections, the dashed lines designate trough and ridge positions which separate regions of poleward and equatorward geostrophic motion within two isentropic layers. The two layers are bounded by θ_ℓ, a lower isentrope, θ_m, an intermediate isentrope, and θ_u, an upper isentrope (figure after Johnson, 1979).

Fig. 11 Energy transport stream functions (10^{14}J $K^{-1}s^{-1}$) from ECMWF Level III analyses for the 300-310 K and 340-350 K isentropic layers of January (a and c, respectively) and July (b and d, respectively) 1979. Rotational component of isentropic energy transport $(\rho J \underset{\sim}{U} \nu)_\Psi$ is parallel to the contours with lower values to the left.

Fig. 12 Energy transport potential functions (10^{13}J $K^{-1}s^{-1}$) from the ECMWF Level III analyses for the 300-310 K and 340-350 K isentropic layers of January (a and c, respectively) and July (b and d, respectively) 1979. Irrotational component of isentropic energy transport $(\rho J \underset{\sim}{U} \nu)_\chi$ is perpendicular to the contours from lower to higher values.

Fig. 13 Vertically integrated global distributions for January 1979 of heat addition (a), and total (b), standing (c), and transient (d) divergence of total energy flux (units, a, b and c (10^2 watts m^{-2}), d (10^1 watts m^{-2})).

Fig. 14 Same as Fig. 13 except for July 1979.

Fig. 15 Area-averaged vertical profiles of horizontal mass divergence (10^{-3} kg m^{-2} $10K^{-1}$ s^{-1}), vertical mass flux (10^{-3} kg m^{-2} s^{-1}) and diabatic heating (K day^{-1}) for January 1979. Potential temperature is used as ordinate in a and b, and pressure is used in c and d. The area averaged in a and c is from 0N to 15S and from 165E to 175W, roughly east of New Guinea. The area averaged in b and d is from 0N to 15S and from 110W to 90W, roughly over eastern equatorial Pacific.

Fig. 16 An energy phase space diagram illustrating the relations between total, total potential, kinetic and available potential energies. The abscissa is kinetic energy, the ordinate is total potential energy, while the family of sloping lines designates isopleths of total energy. The region of admissible energy levels under reversible isentropic energy transformation between kinetic and available potential energies lies above the isopleth of Π_r.

Fig. 17 Time-averaged Montgomery Stream Function (a), kinetic energy generation (b) and standing irrotational (c) and rotational (d) modes of kinetic energy generation for the 340-350 K isentropic layer of January 1979 (Units, a (10^3 m^2 s^{-2}); b, c; and d (10^{-1} kg s^{-3} K^{-1}).

Fig. 18 Schematic of a simple convective system that contrasts vertical mass flux in isobaric and isentropic coordinates.

Fig. 19 Schematic to illustrate the determination of length scale of mass transport in isentropic coordinates.

Fig. 20 Global distributions of vertically integrated divergence of energy transport for January 1979 from ECMWF Level III analyses. (a) Integration in isentropic coordinates, (b), (c) and (d) show integration in isobaric coordinates for total divergence, standing component and transient component (10^2 watts m^{-2}), respectively.

Fig. 21 Same as Fig. 20 except for July 1979.

Fig. 22 Potential functions determined for January 1979 from the corresponding global distribution of isentropic and isobaric energy divergence fields shown in Fig. 20 (units, a, b and c (10^{14} J s^{-1}), d (10^{13} J s^{-1})).

Fig. 23 Potential functions for July 1979 determined from the corresponding global distribution of isentropic and isobaric energy divergence fields shown in Fig. 21 (units a, b and c (10^{14} J s^{-1}), d (10^{13} J s^{-1})).

Fig. 24 Profiles of vertically and zonally integrated meridional energy transport (10^{15}J s^{-1}) from ECMWF Level III analyses for January (a and c) and July (b and d). Results in a and b, are from isentropic coordinates, results in c and d are from isobaric coordinates. Panels show total (solid), mean meridional (short dash), stationary (long dash) and transient components (mixed dash).

Fig. 25 Meridional cross sections of: (a) mean meridional circulation (10^{10} kg s^{-1}), (b) zonally integrated relative angular momentum (10^{16} kg m s^{-1} K^{-1}) and zonally integrated (c) total torques, (d) frictional torque, (e) pressure torque, and (f) forcing by horizontal and vertical divergence of eddy angular momentum transport (10^{11} kg m s^{-2} K^{-1}) for January of the FGGE year from ECMWF Level III analyses.

Fig. 26 Same as Fig. 25 except for July 1979.

Fig. 27 Vertical cross section from Tucson to Bermuda through a large amplitude wave studied by Newton and Palmén (1963). Dotted lines are isentropes, solid lines are isotachs of the meridional wind while heavy lines are frontal boundaries and tropopause. Within the 320-335 K isentropic layer the cross hatched area denotes poleward mass transport while the dotted area denotes equatorward mass transport. The larger area of cross-hatching indicates the positive covariance of mass and meridional motion embedded in wave structure.

Fig. 28 Zonal cross sections along 50°N of (a) Time-averaged potential temperature, θ^t (units, 10 K), and meridional velocity v^t (units, m s^{-1}) in isobaric coordinates; and (b) Time-averaged pressure, $\overline{p}^t$ (units, 10^2 mb), and meridional mass transport $\rho J_\theta v^t$ (units, 10^2 kg m^{-1} K^{-1} s^{-1}), in isentropic coordinates for December, January and February of the GWE year.

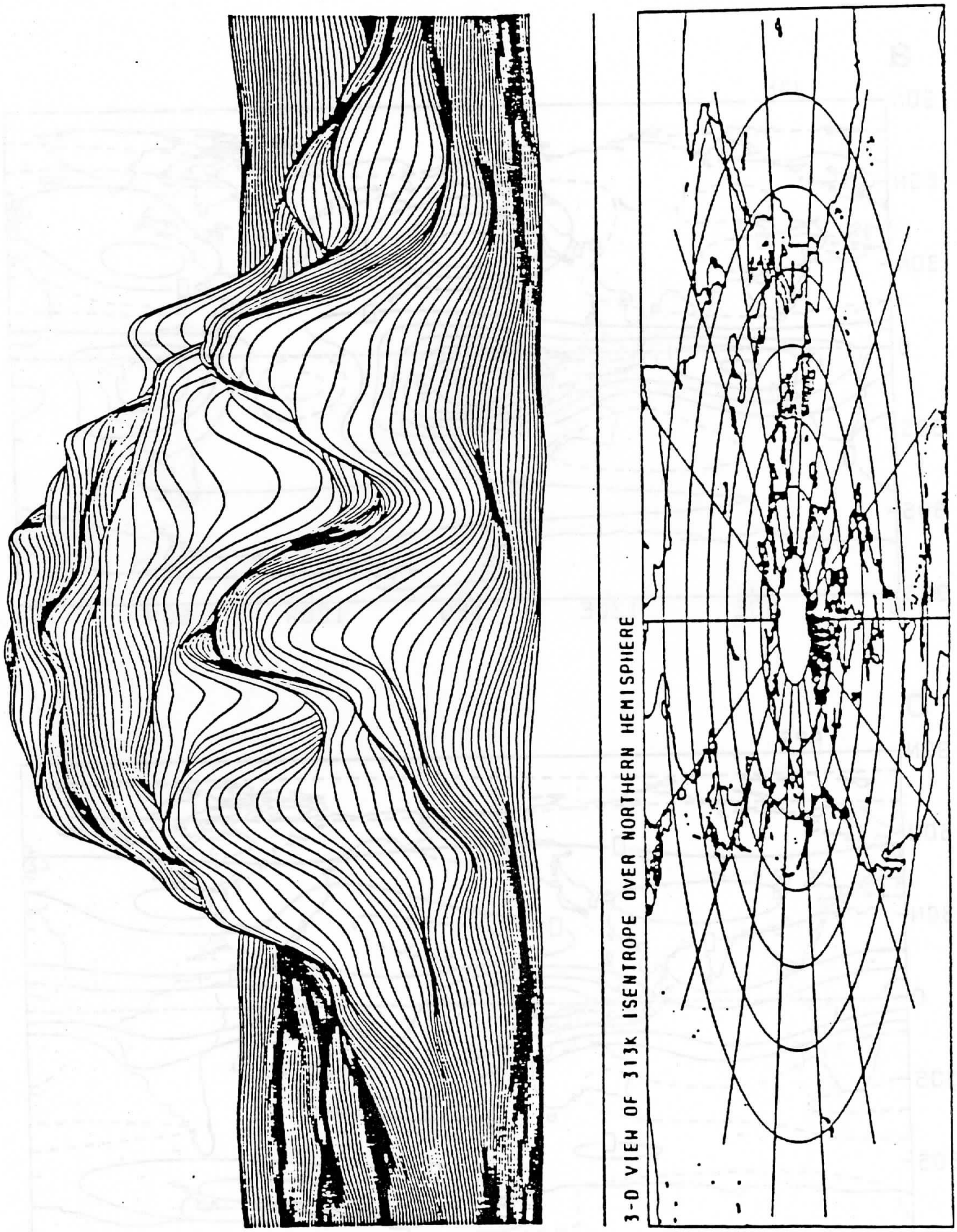


Fig. 1 Three-dimensional pattern that illustrates the topography of the 313 K isentropic surface over the Northern Hemisphere as viewed from 80°W longitude, 1200 GMT, 12 November 1977 (Nagle, 1979).

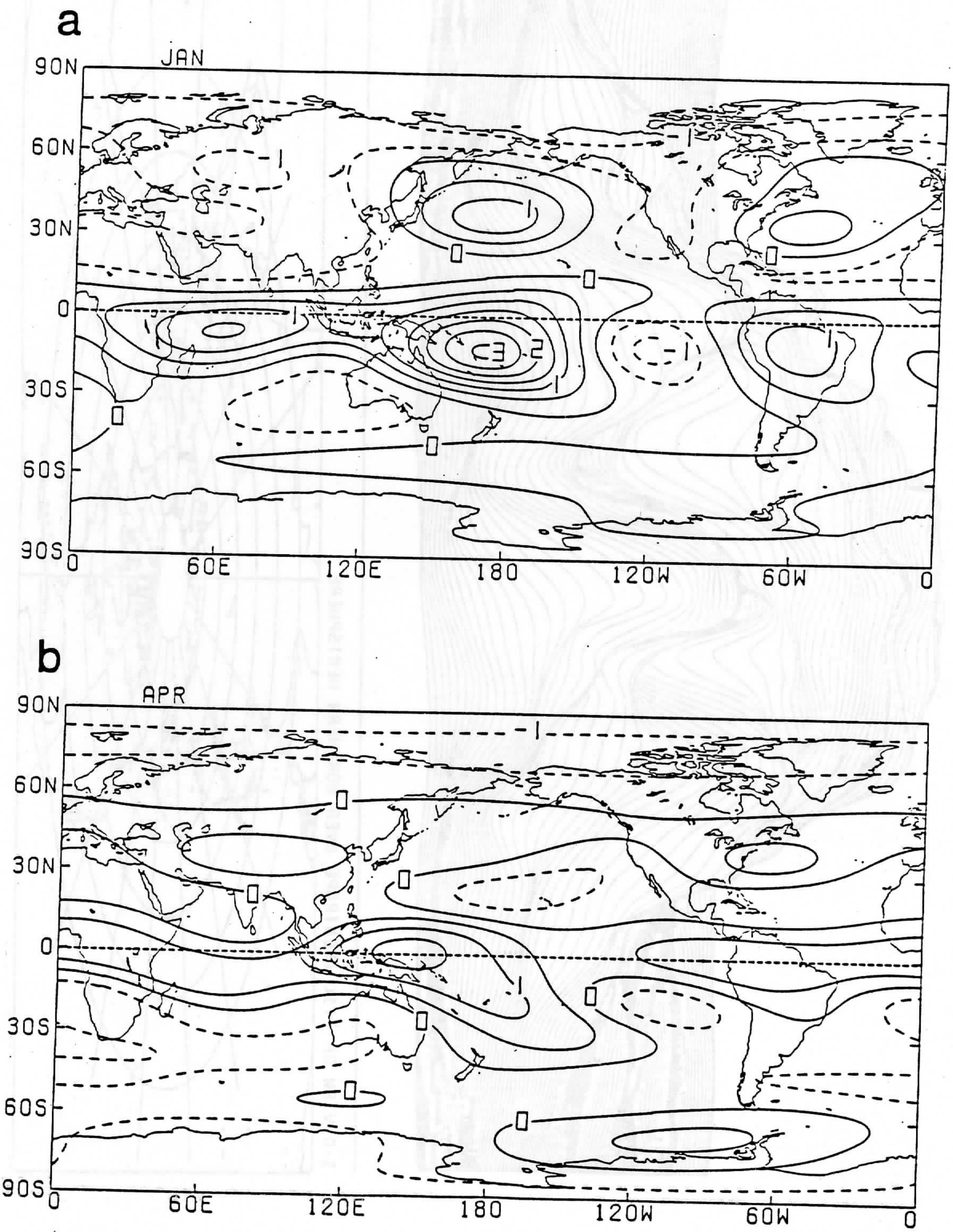
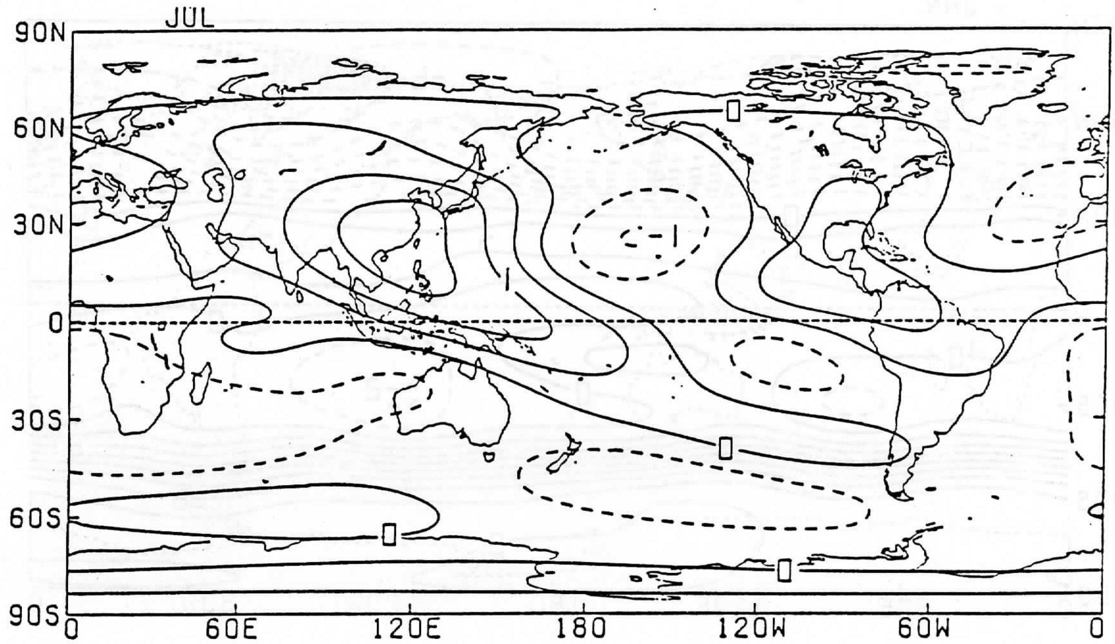


Fig. 2 Mass-weighted vertically averaged heating rates (K day^{-1}) computed from ECMWF Level III analyses for (a) January, (b) April, (c) July, and (d) October 1979. Heating distributions are filtered to emphasize heating from wavelengths greater than 10,000 km.

c



d

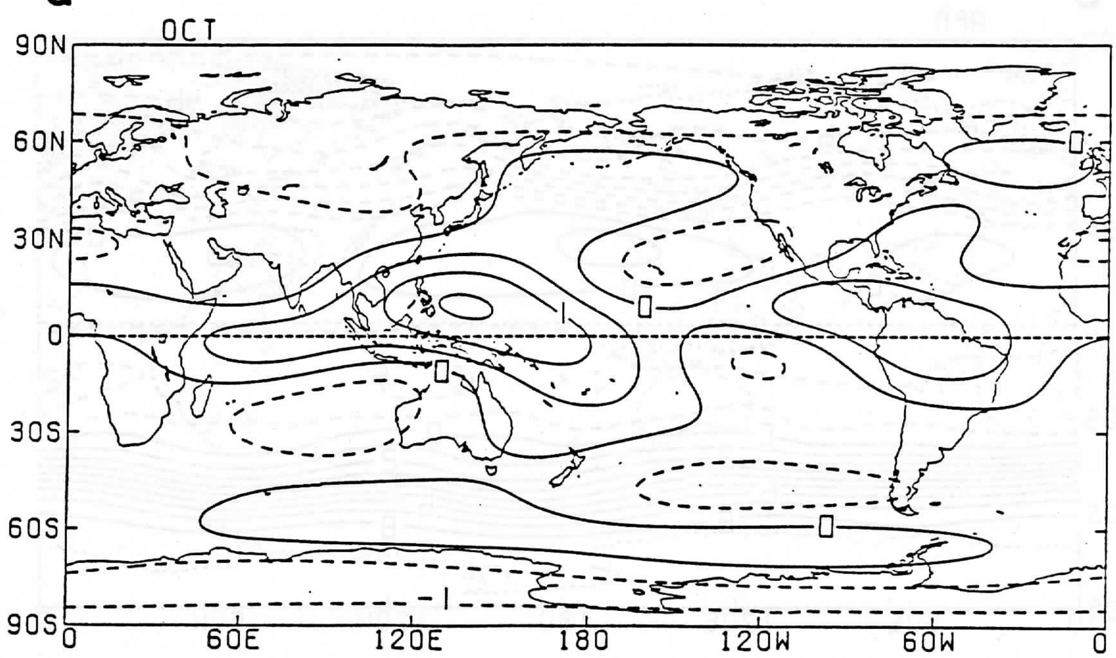


Fig. 2 (Continued)

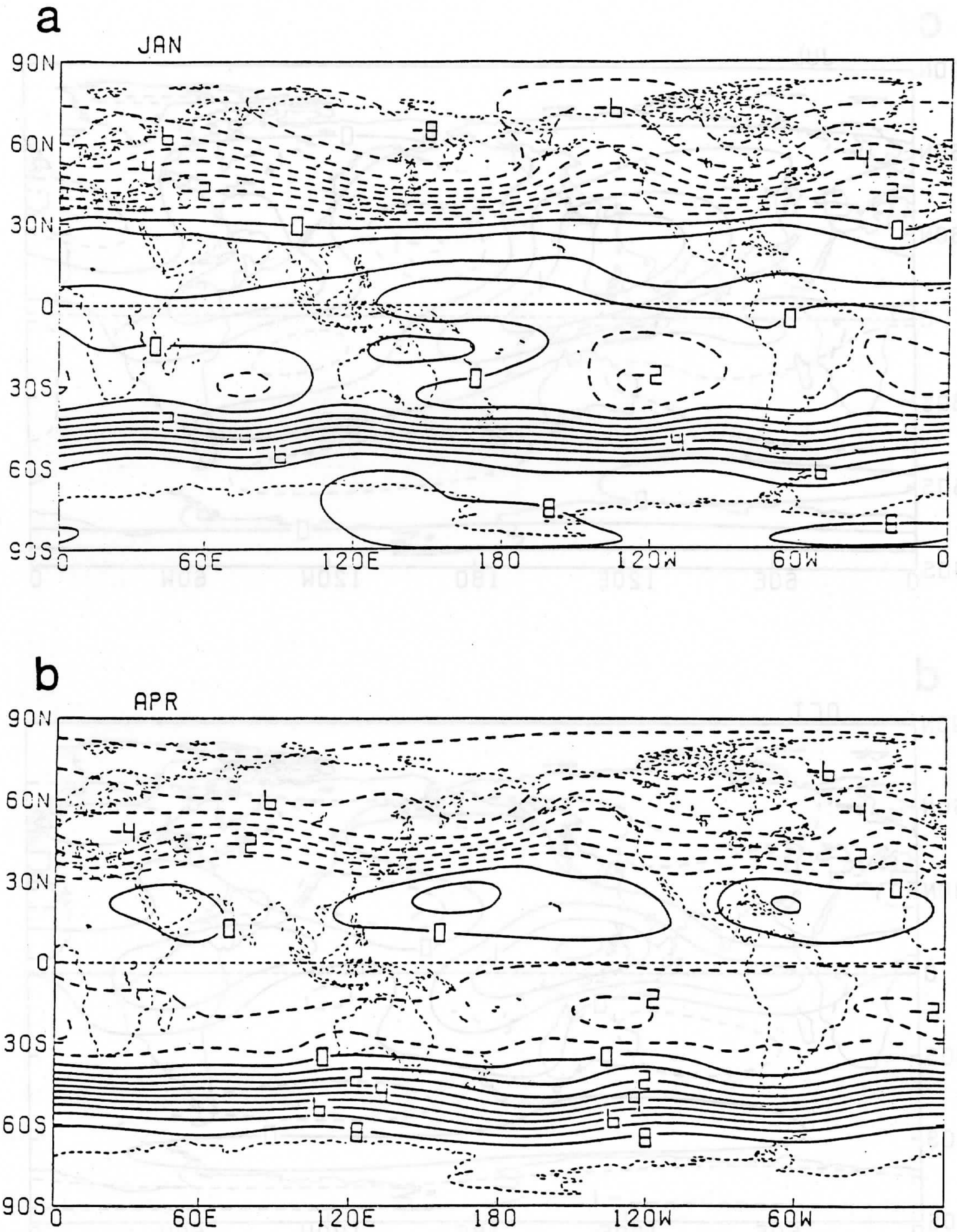


Fig. 3 Mass transport stream functions (10^9 kg $K^{-1}s^{-1}$) from ECMWF Level III analyses for the 300-310 K isentropic layer of January (a), April (b), July (c) and October (d) 1979. Rotational component of isentropic mass transport $(\rho J\underset{\sim}{U})_\psi$ is parallel to the contours with lower values to the left.

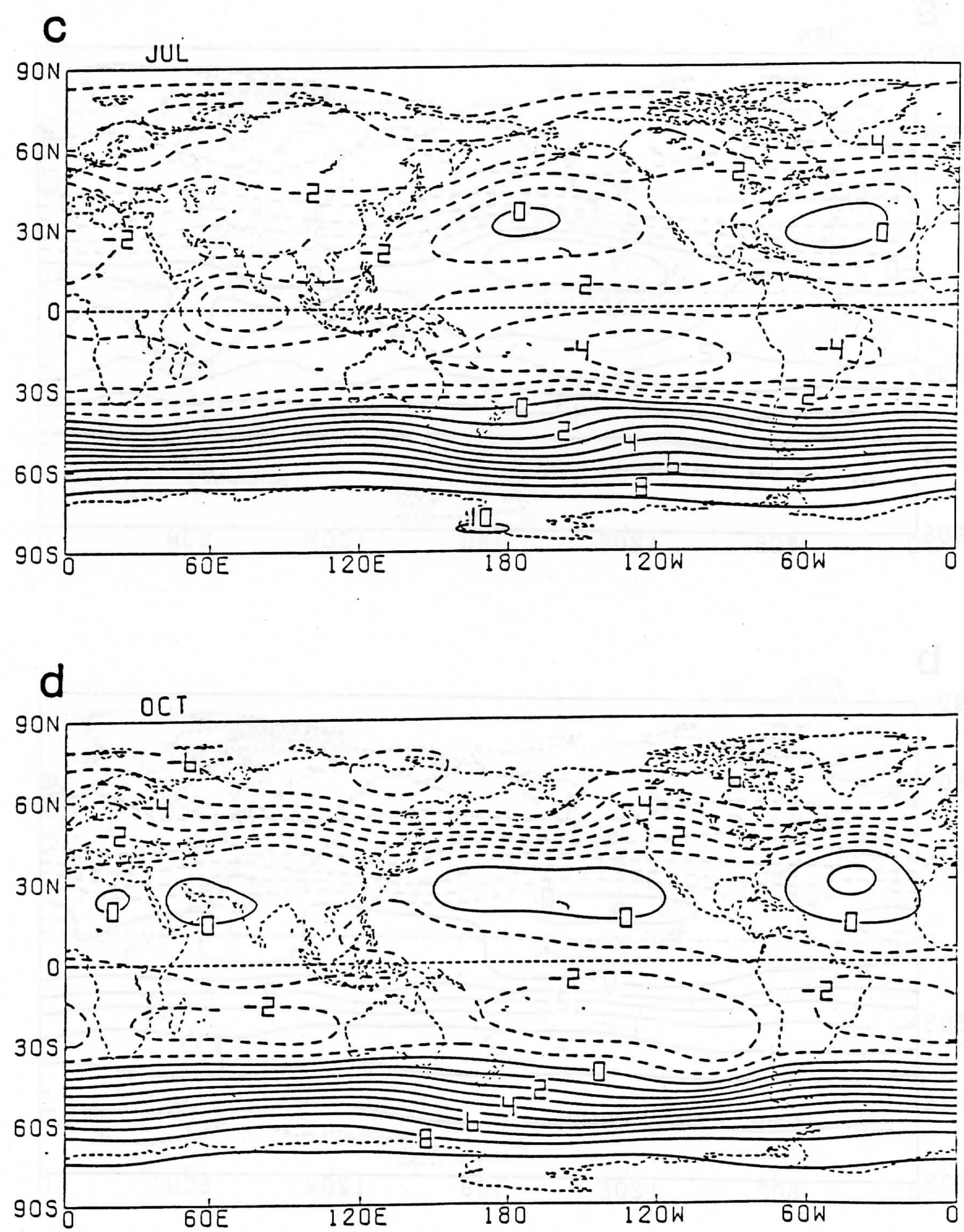


Fig. 3 (Continued)

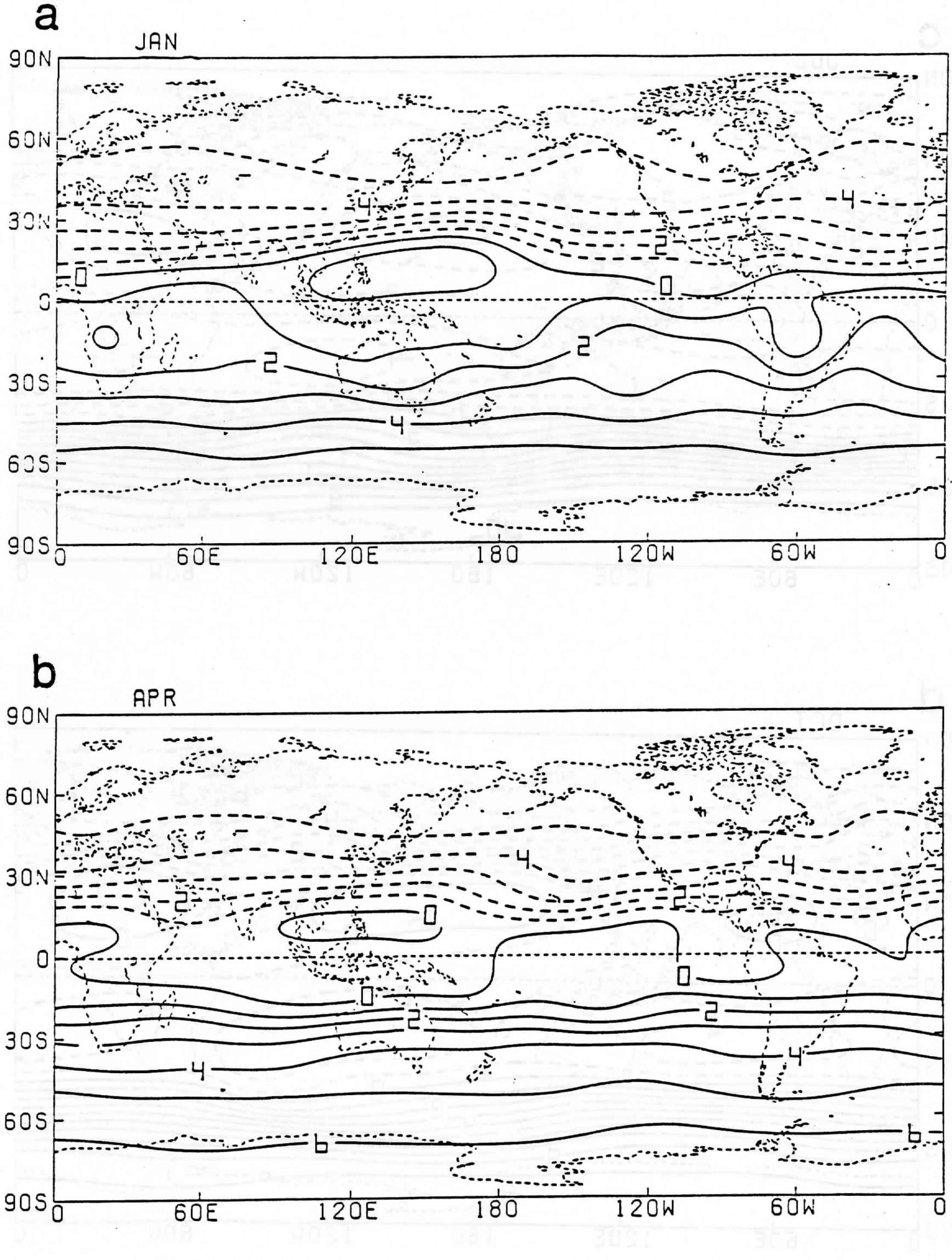
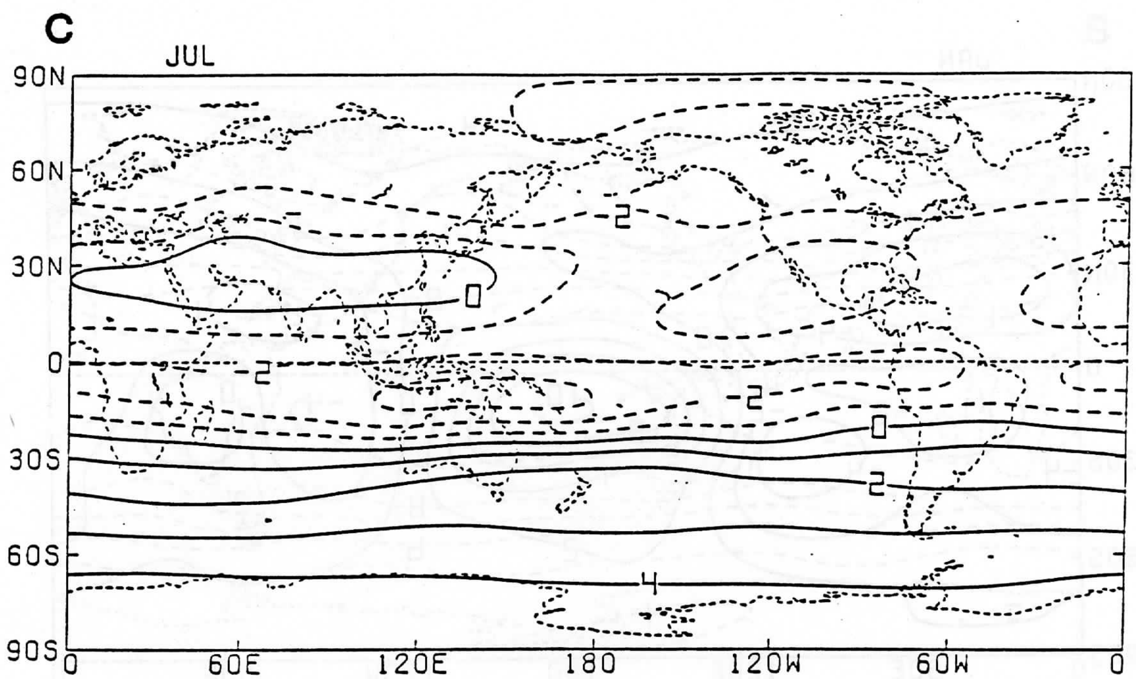


Fig. 4 Same as Fig. 3 except for 340-350 K isentropic layer.



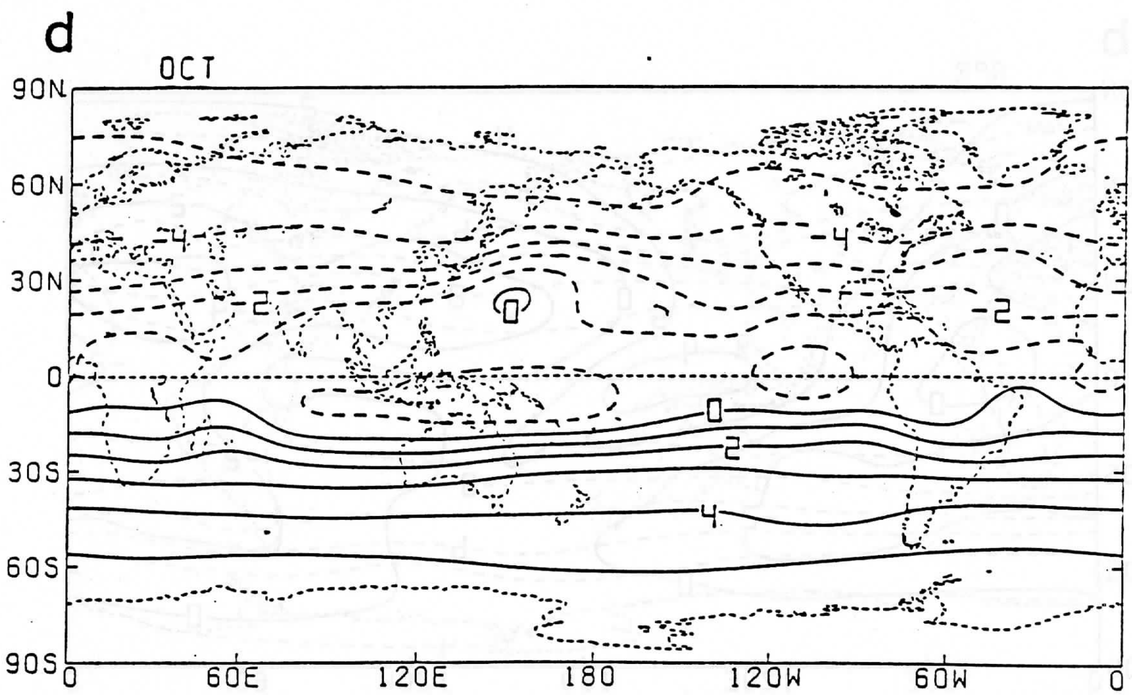


Fig. 4 (Continued)

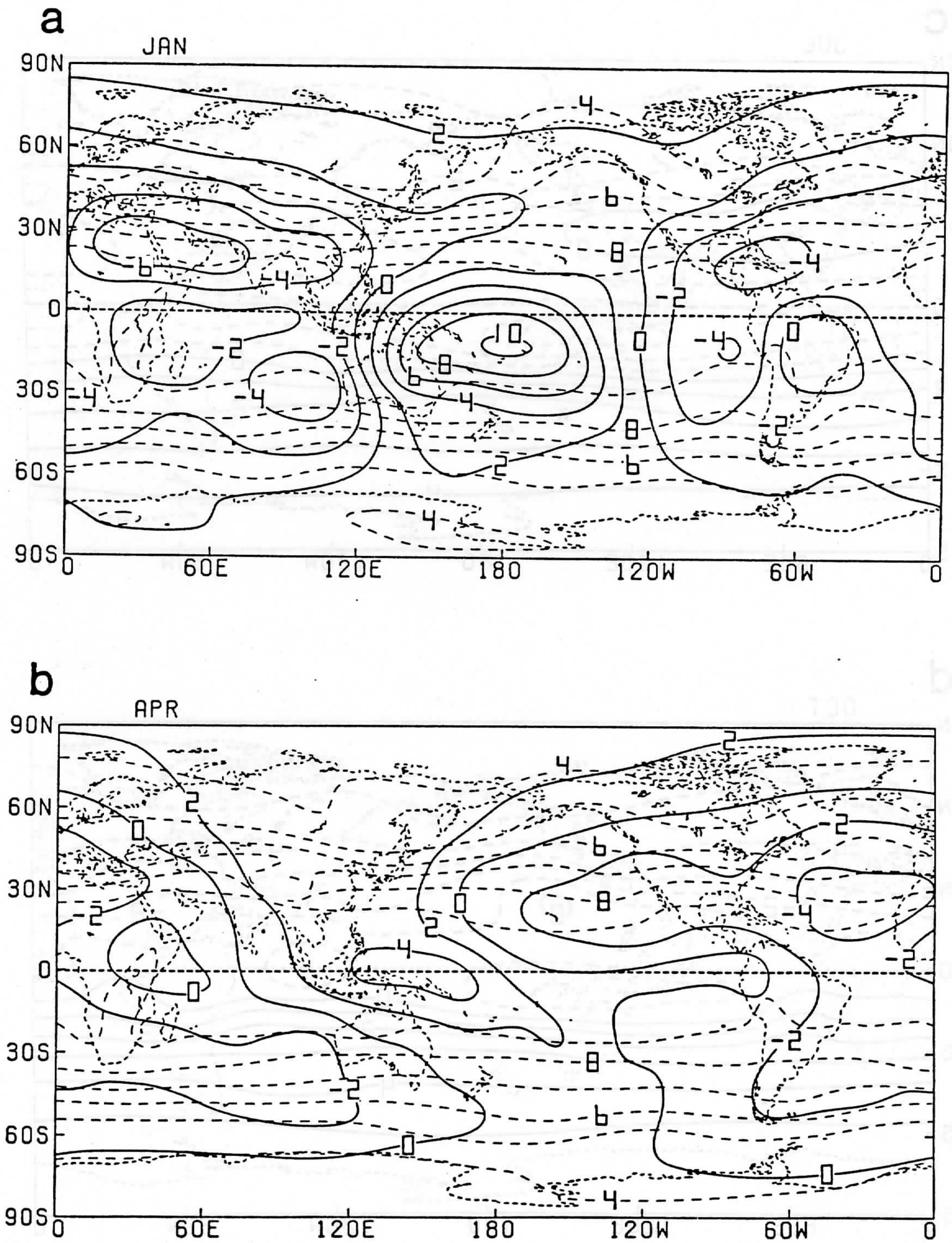
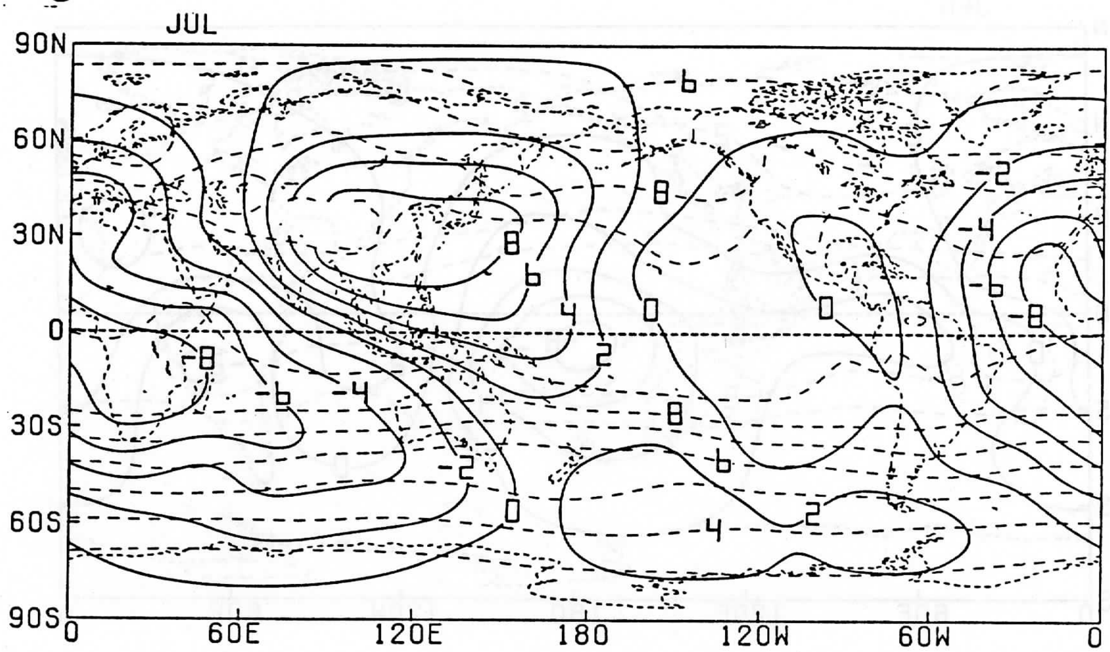


Fig. 5 Mass transport potential functions (10^8 kg $K^{-1}s^{-1}$) from ECMWF Level III analyses for the 300-310 K isentropic layer of January (a), April (b), July (c) and October (d) 1979. Irrotational component of mass transport $(\rho JU)_\chi$ is perpendicular to contours from lower to higher values. Dash lines represent the time-averaged pressure (10^2 mb) of the lower surface of the isentropic layer.

c



d

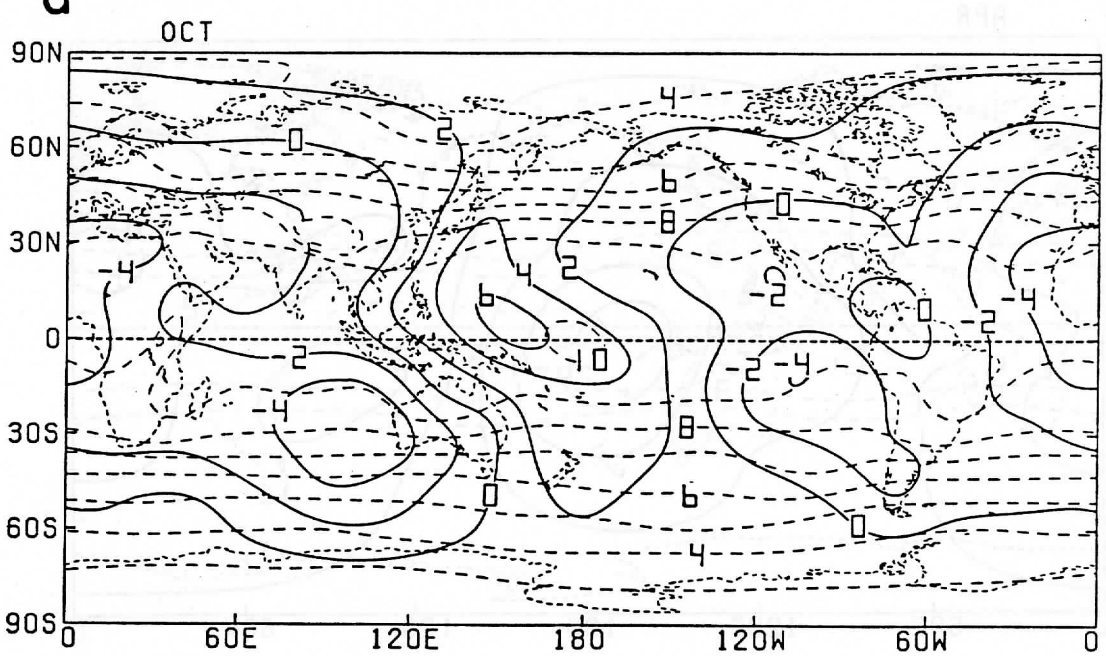


Fig. 5 (Continued)

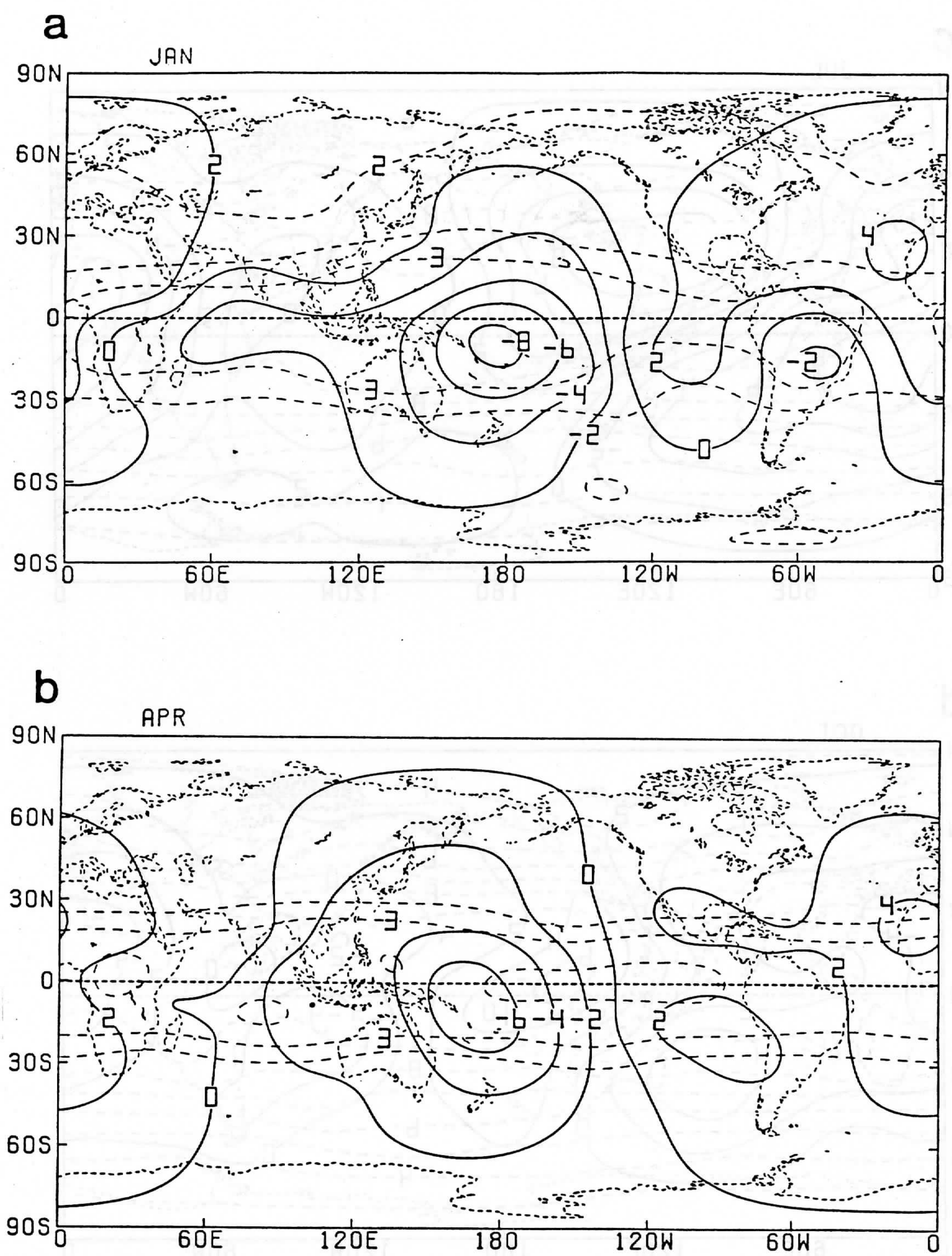
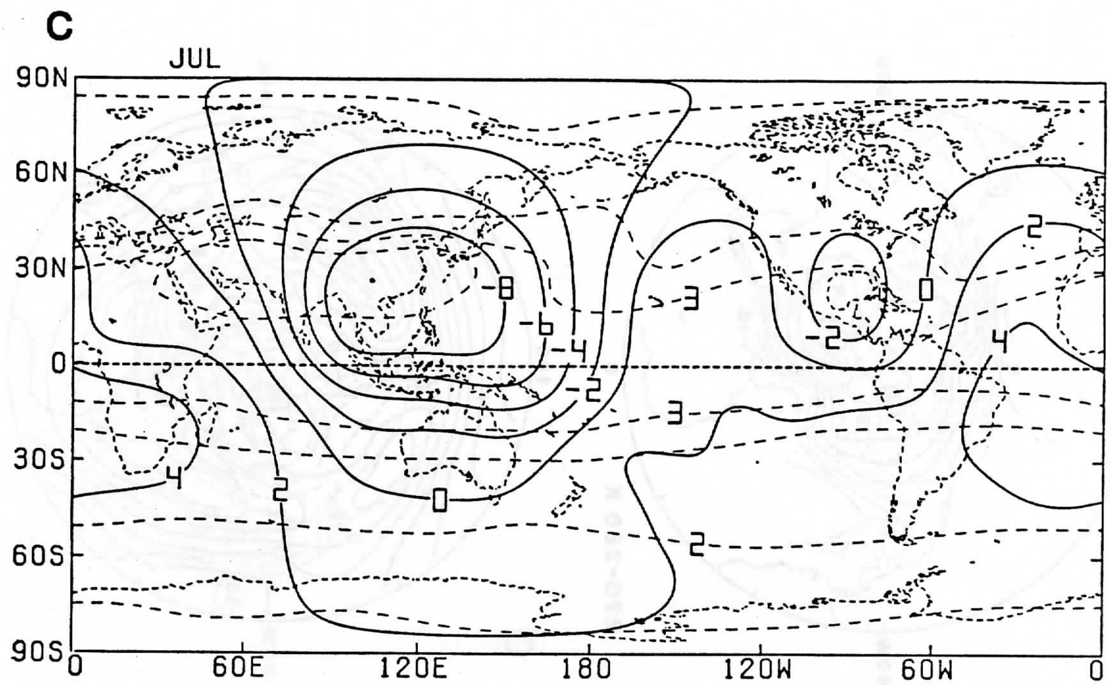


Fig. 6 Same as Fig. 5 except for the 340-350 K isentropic layer.



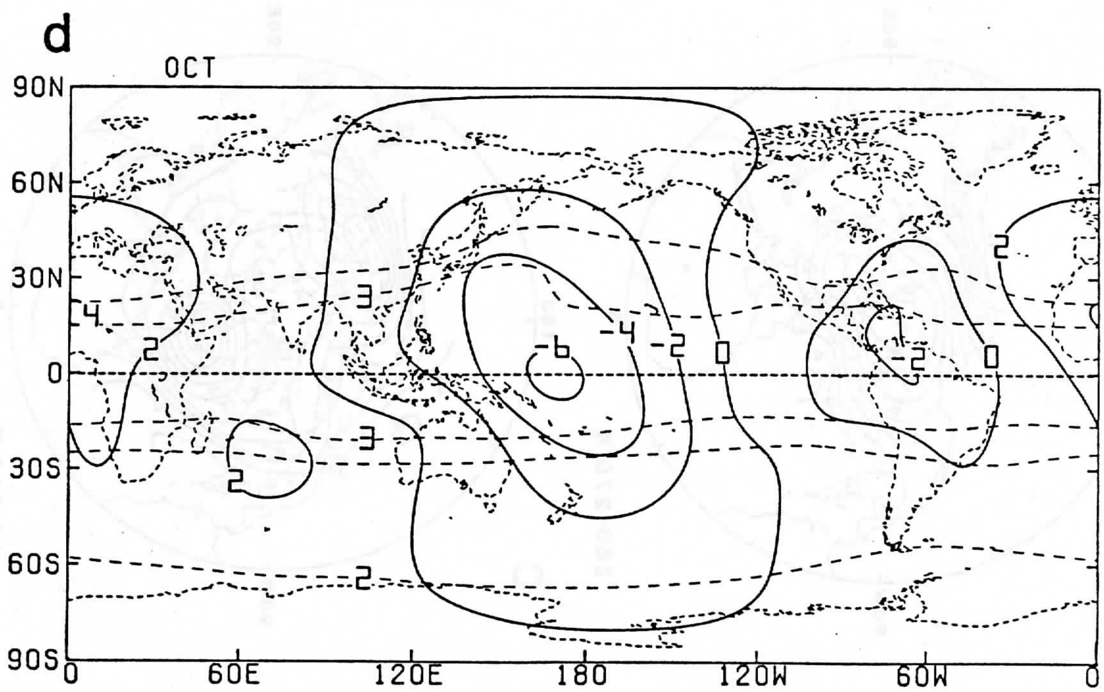


Fig. 6 (Continued)

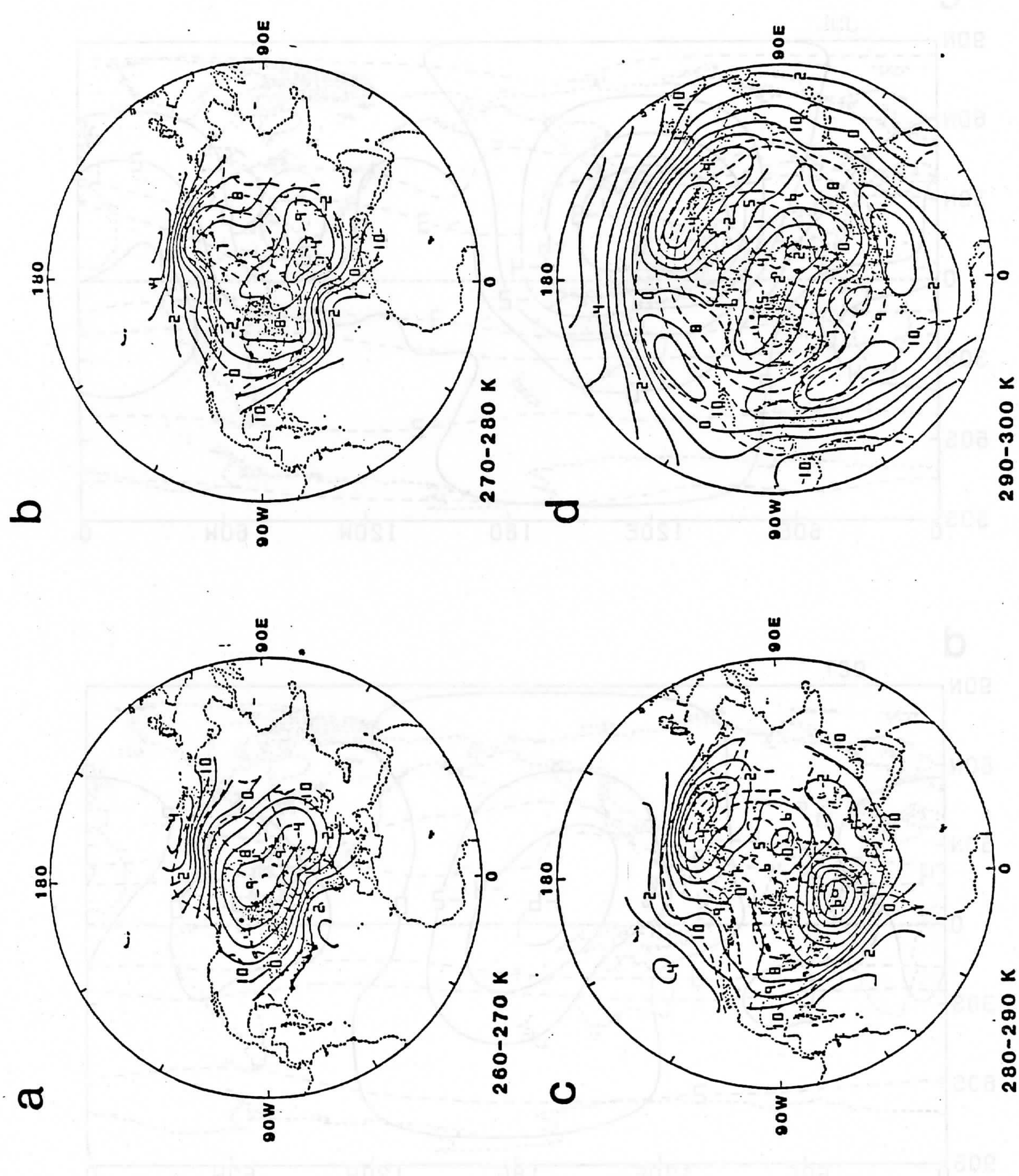


Fig. 7 Same as Fig. 5 except for 260-270 K (a), 270-280 K (b), 280-290 K (c) and 290-300 K (d) isentropic layers for January 1979. Dashed lines represent pressure of bottom level of the isentropic layer (10^2 mb).

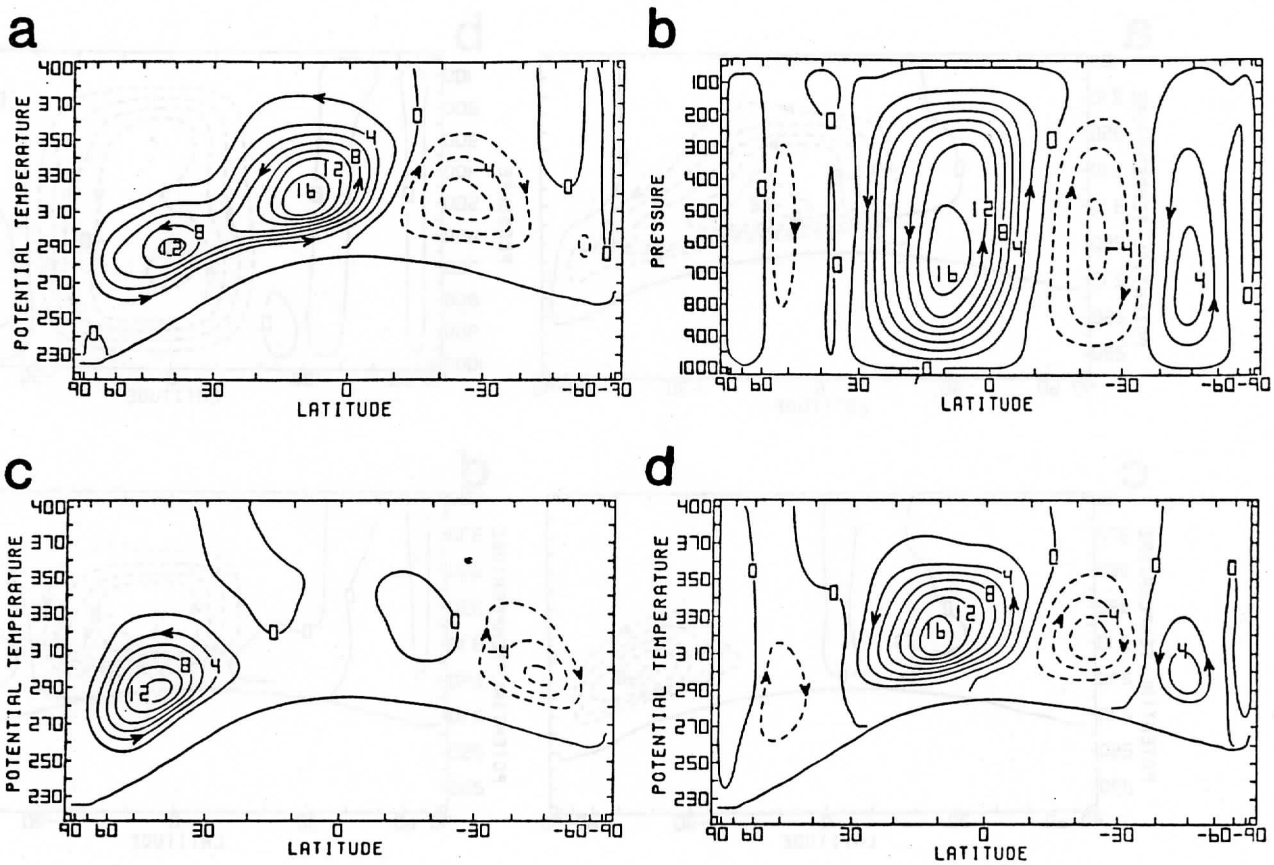


Fig. 8 Mass stream functions (10^{10} kg s^{-1}) for mean meridional circulations computed from ECMWF Level III analyses for January 1979, isentropic (a), isobaric (b), isentropic geostrophic component (c) and isentropic ageostrophic component (d).

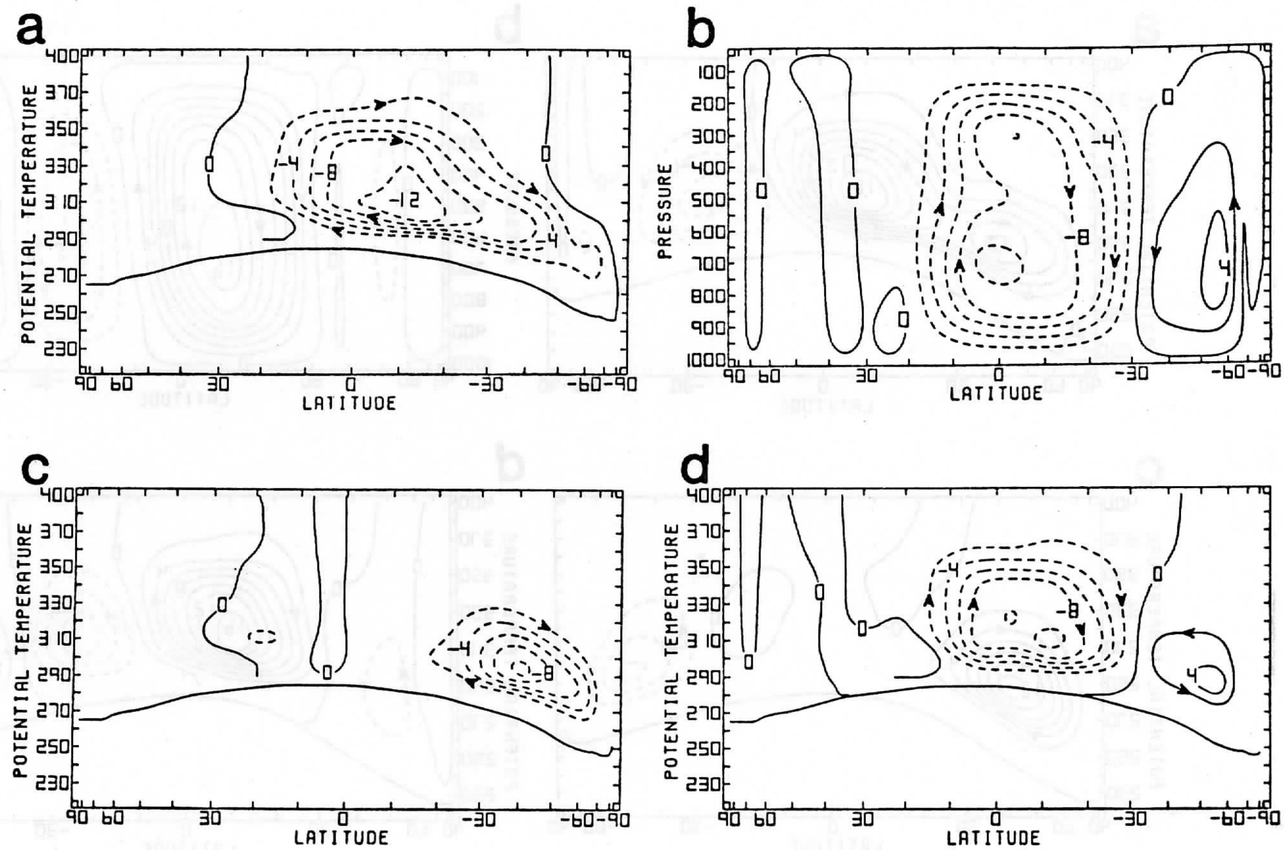


Fig. 9   Same as 8 except for July 1979.

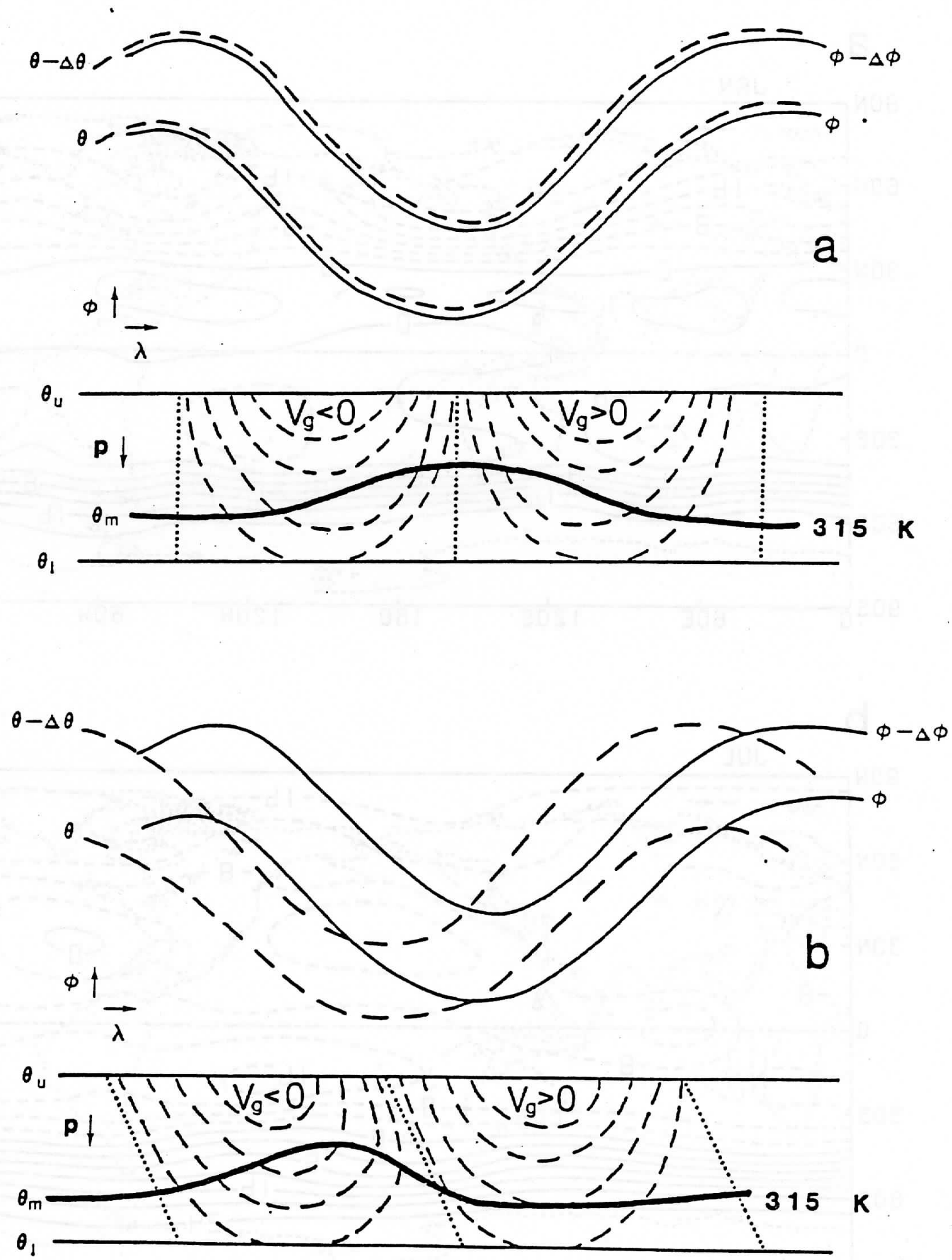


Fig. 10 Schematics of horizontal (λ,ϕ) and vertical (λ,p) distributions of geopotential and potential temperature within a steady (a) and amplifying (b) wave. In the zonal, vertical cross sections, the dashed lines designate trough and ridge positions which separate regions of poleward and equatorward geostrophic motion within two isentropic layers. The two layers are bounded by θ_ℓ, a lower isentrope, θ_m, an intermediate isentrope, and θ_u, an upper isentrope (figure after Johnson, 1979).

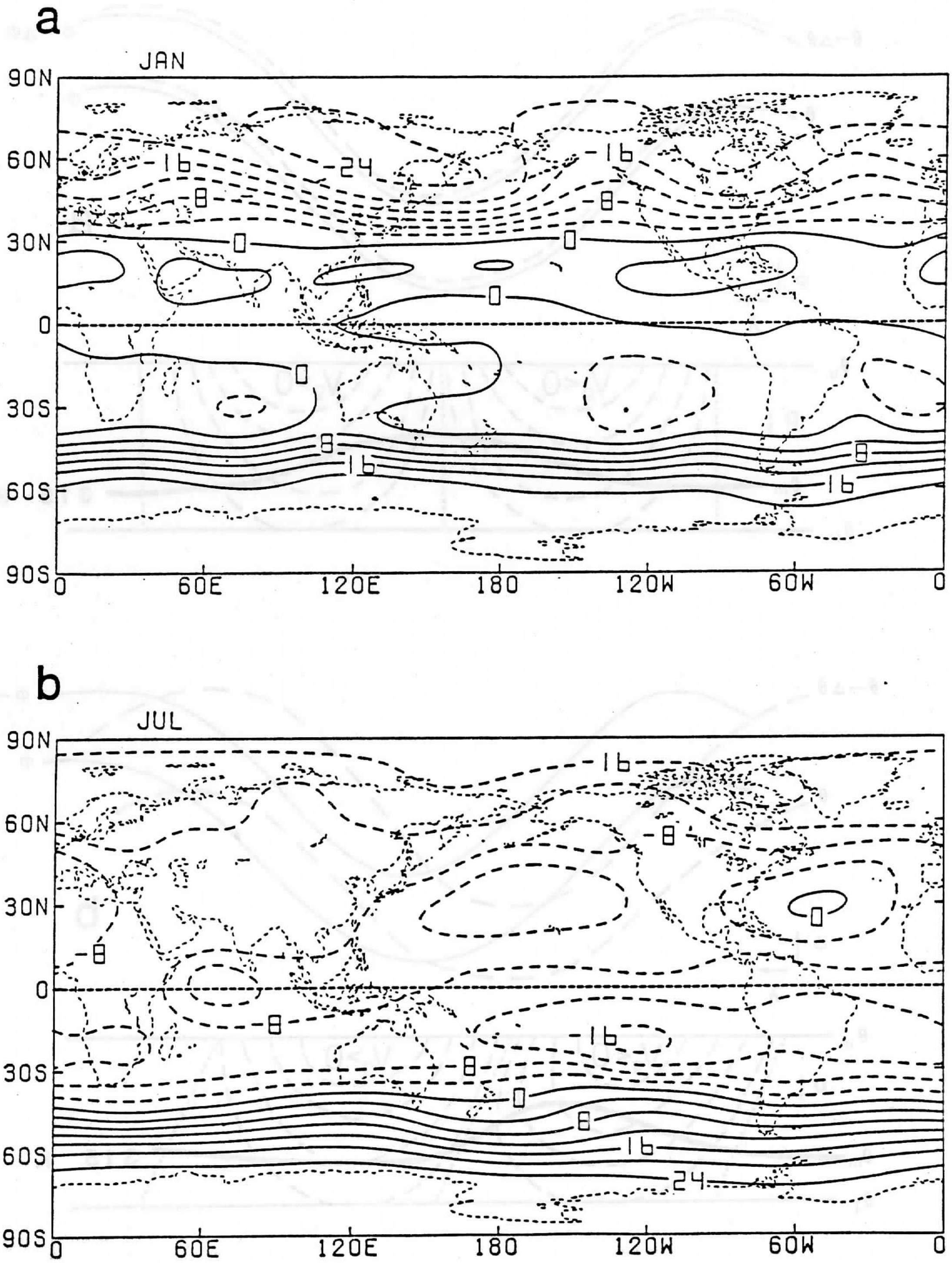


Fig. 11 Energy transport stream functions (10^{14} J $K^{-1}s^{-1}$) from ECMWF Level III analyses for the 300-310 K and 340-350 K isentropic layers of January (a and c, respectively) and July (b and d, respectively) 1979. Rotational component of isentropic energy transport $(\rho J U v)_{\psi}$ is parallel to the contours with lower values to the left.

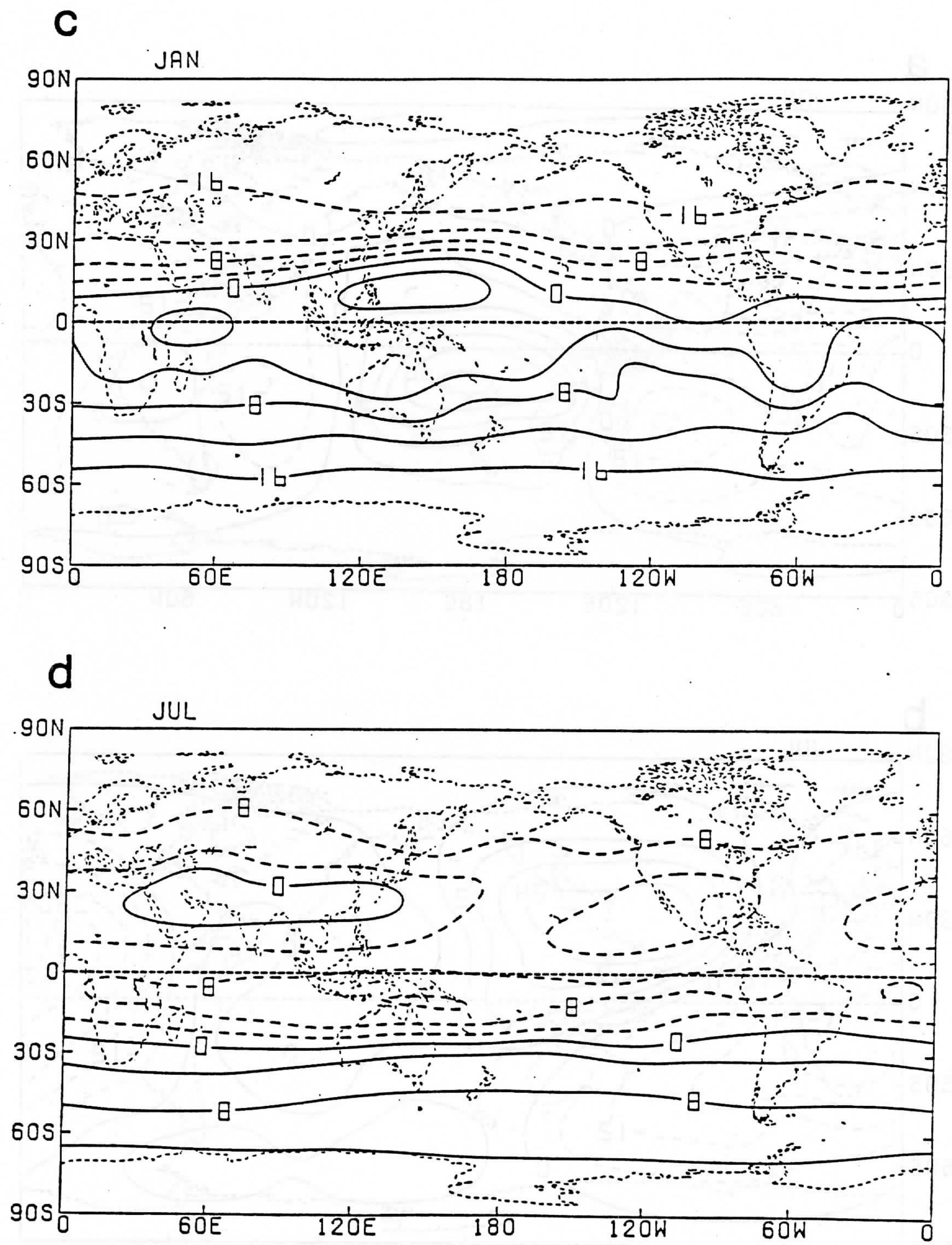
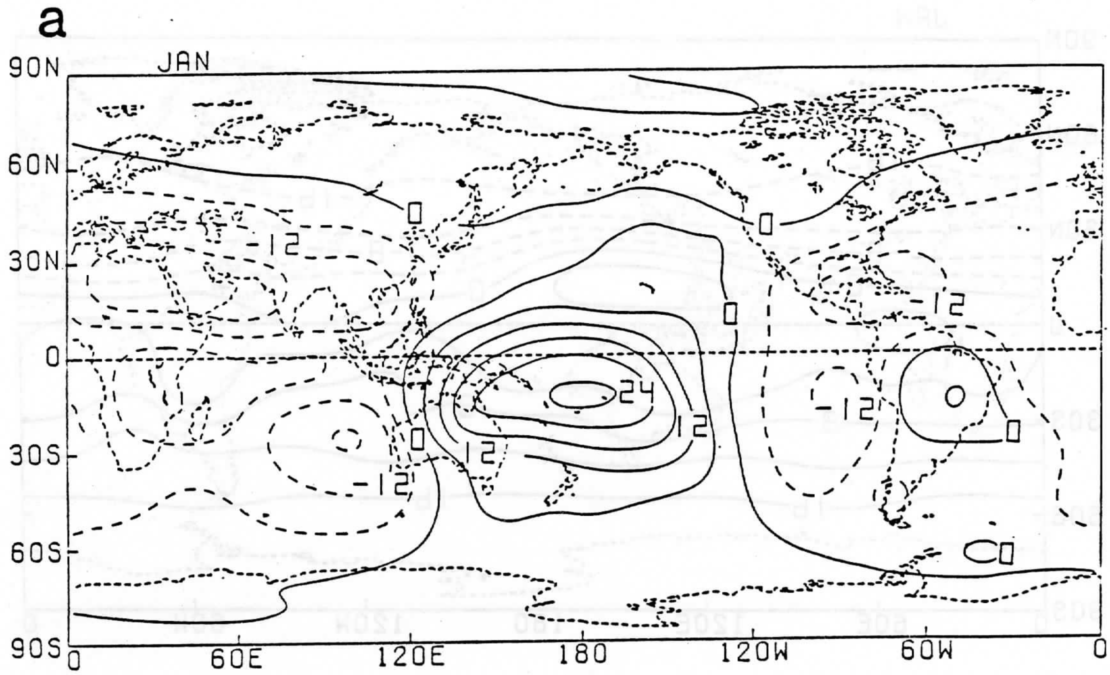


Fig. 11 (Continued)



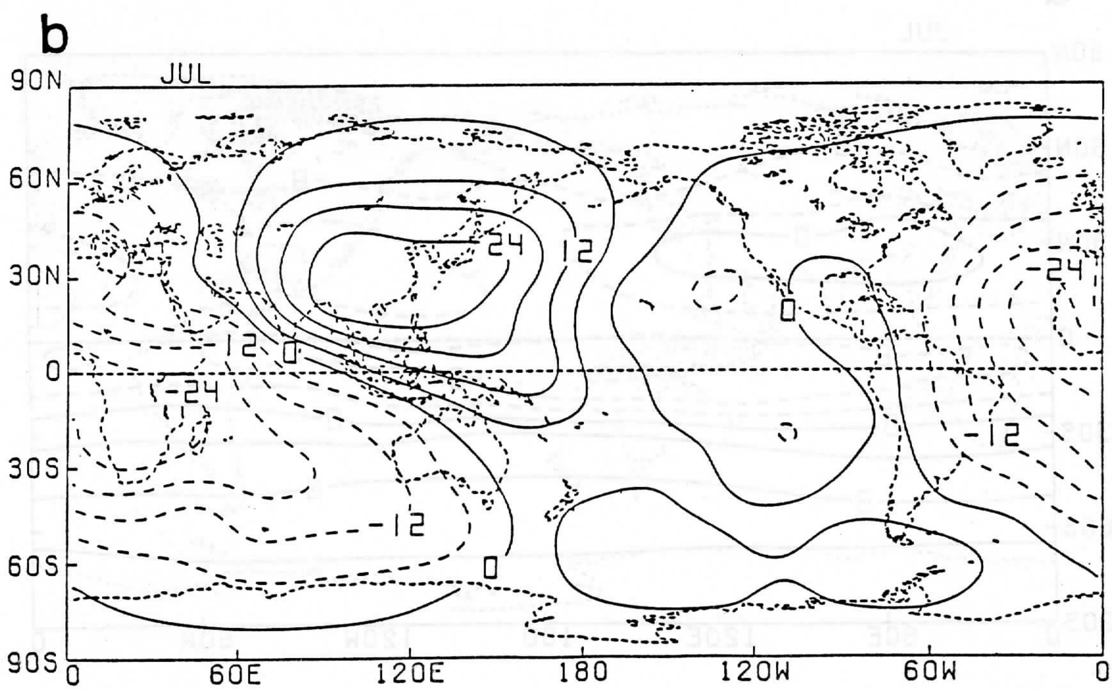
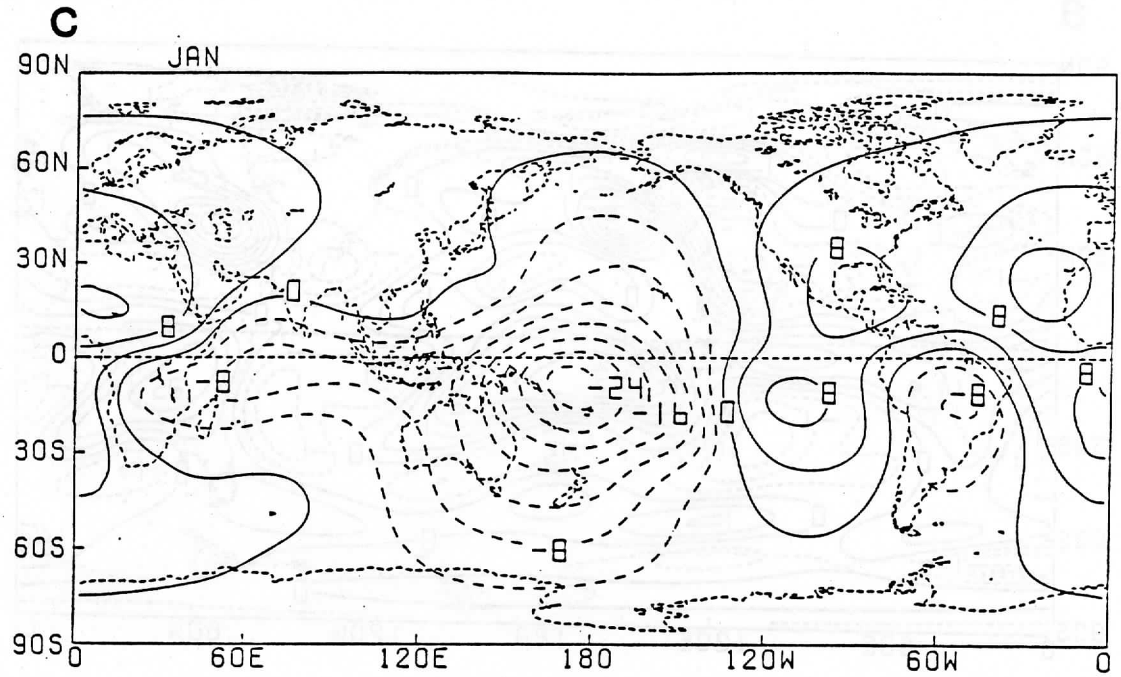


Fig. 12 Energy transport potential functions (10^{13}J $K^{-1}s^{-1}$) from the ECMWF Level III analyses for the 300-310 K and 340-350 K isentropic layers of January (a and c, respectively) and July (b and d, respectively) 1979. Irrotational component of isentropic energy transport $(\rho J U_\nu)_\chi$ is perpendicular to the contours from lower to higher values.



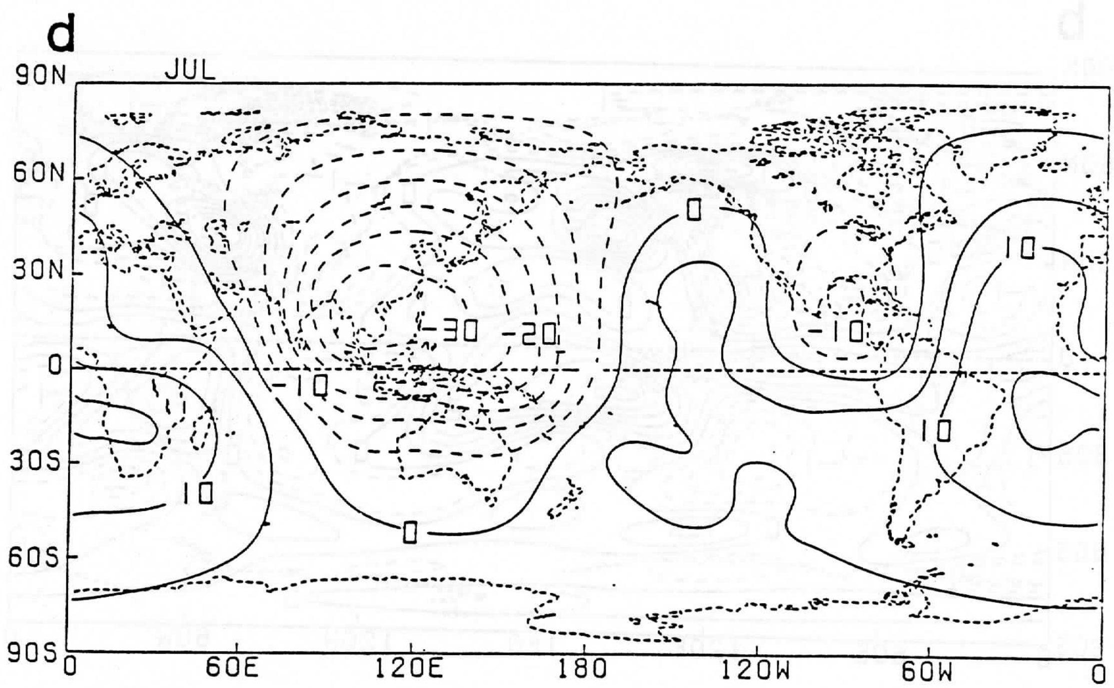


Fig. 12 (Continued)

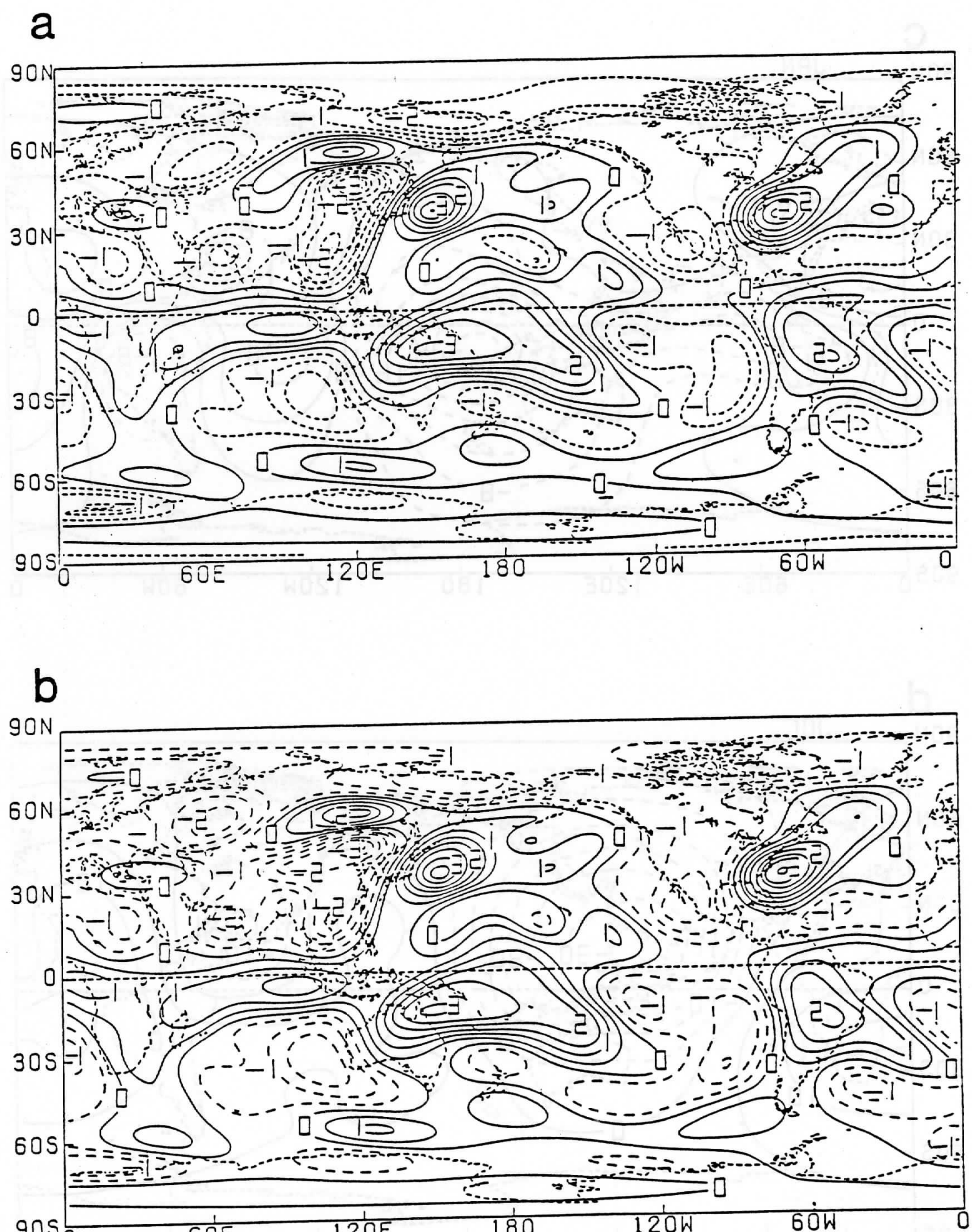


Fig. 13 Vertically integrated global distributions for January 1979 of heat addition (a), and total (b), standing (c), and transient (d) divergence of total energy flux (units, a, b and c (10^2 watts m^{-2}), d (10^1 watts m^{-2})).

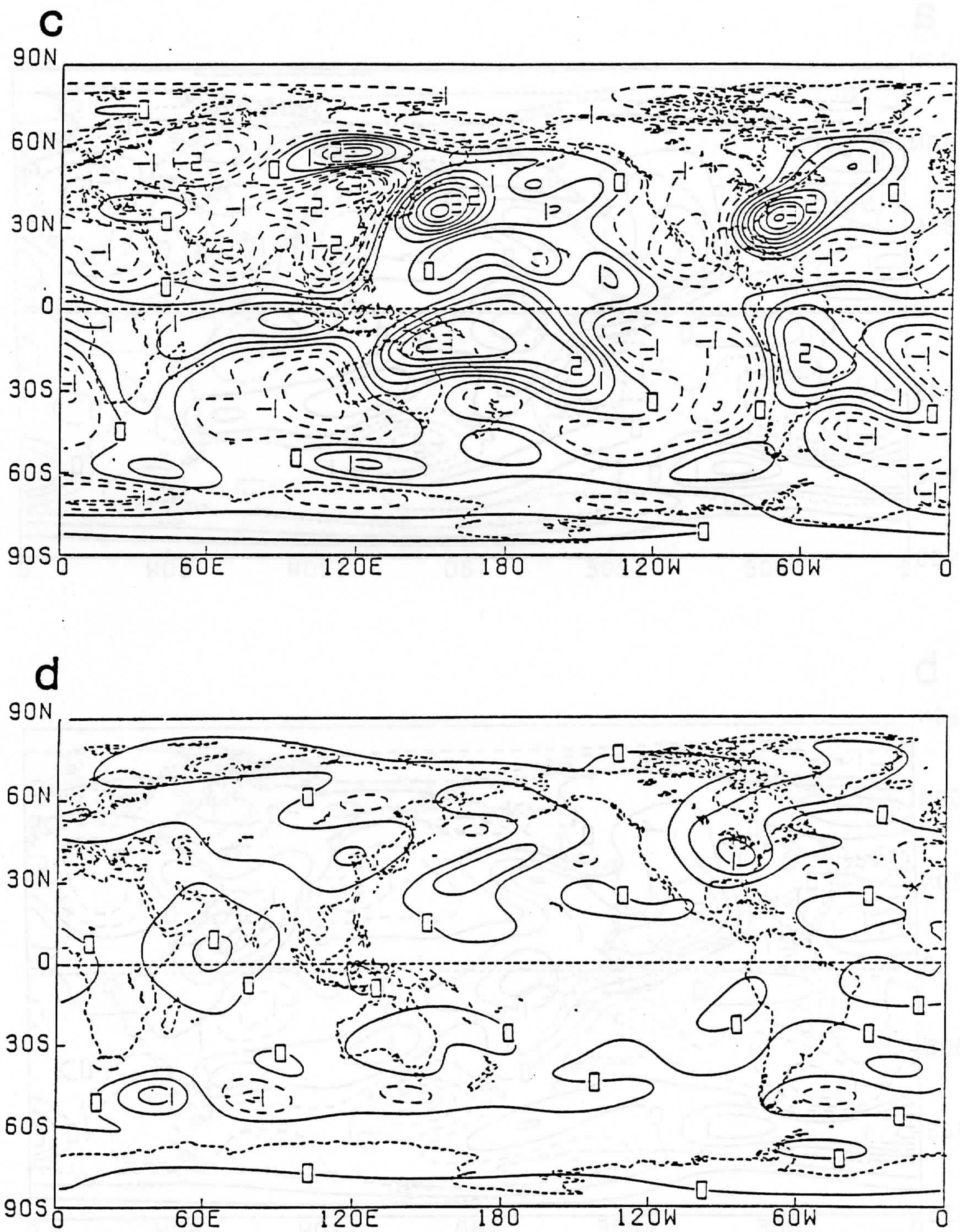


Fig. 13 (Continued)

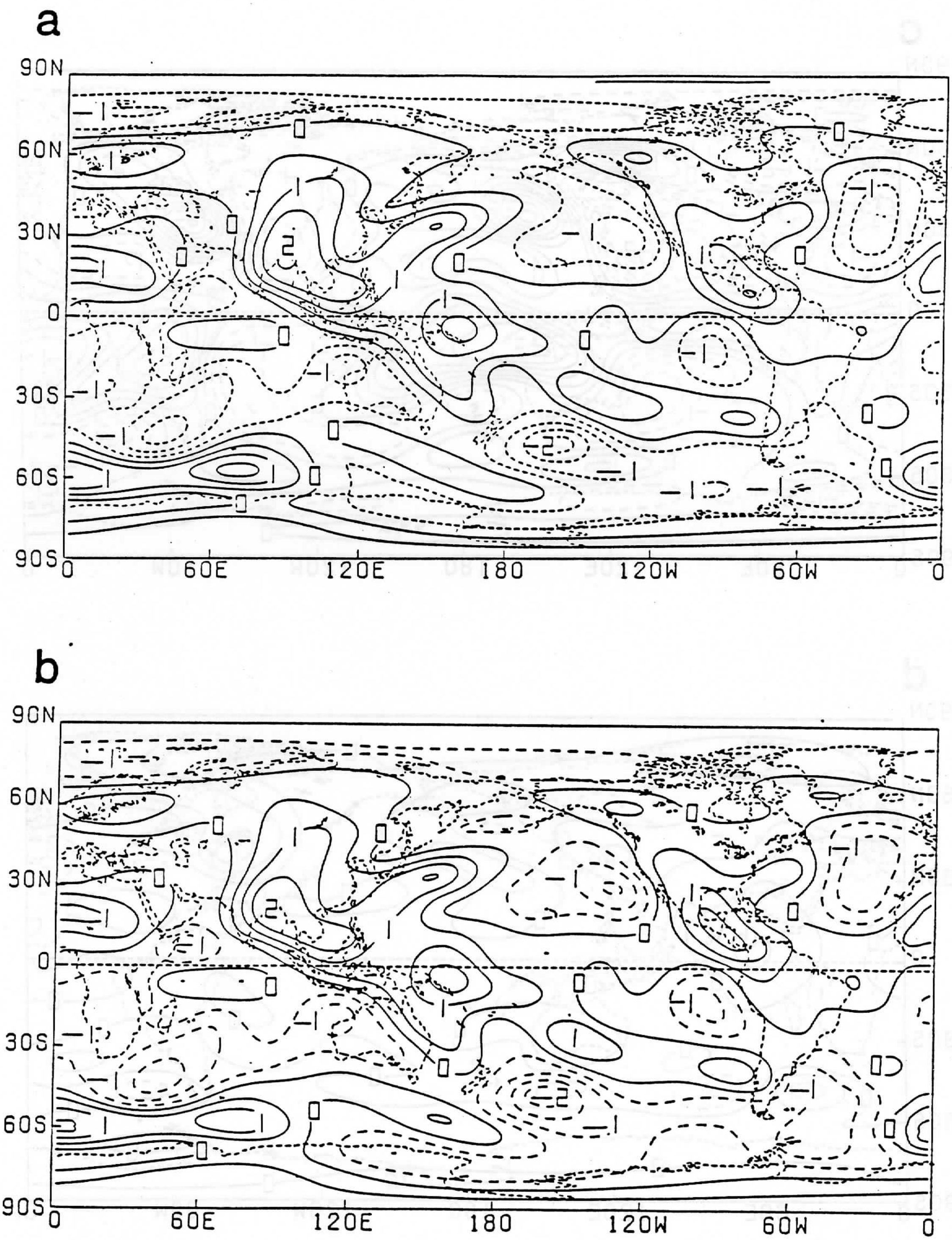
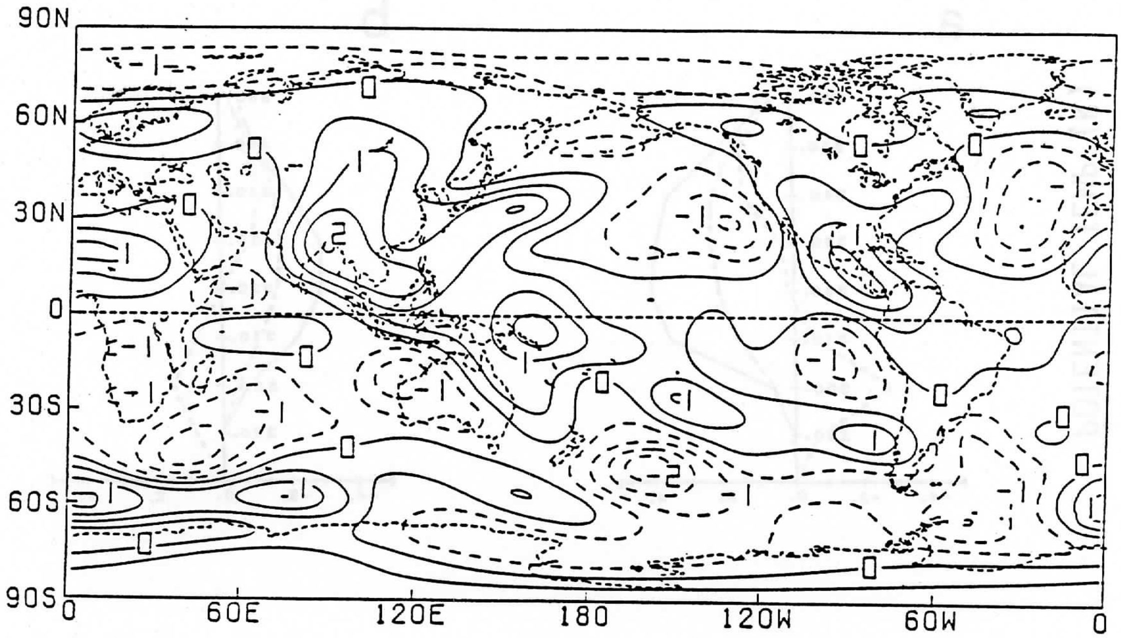


Fig. 14 Same as Fig. 13 except for July 1979.

c



d

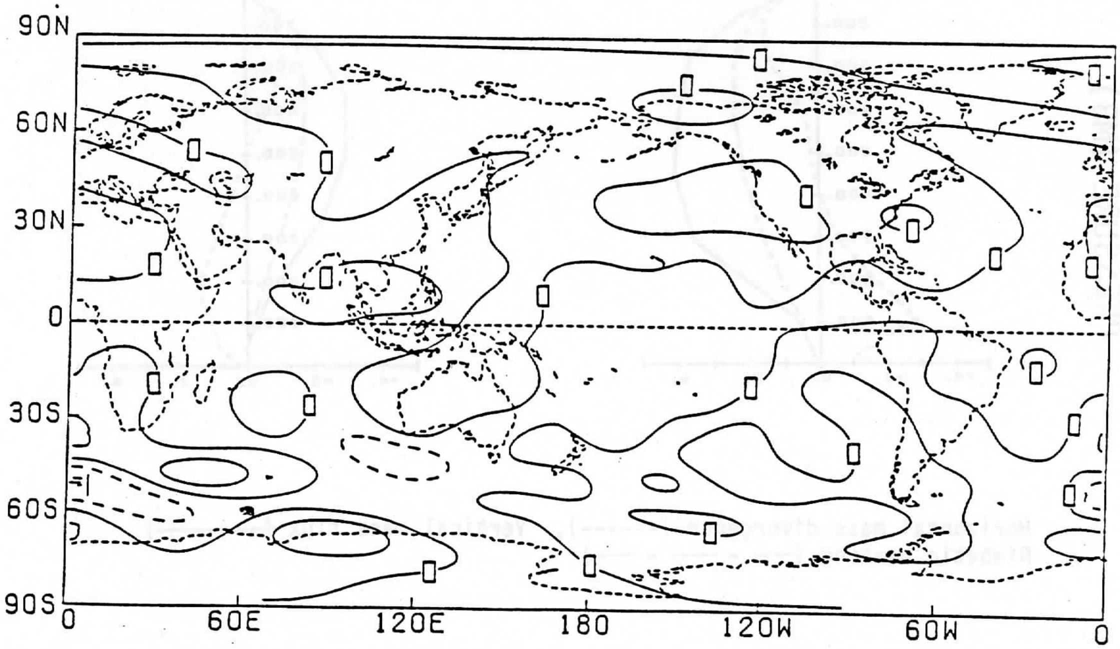


Fig. 14 (Continued)

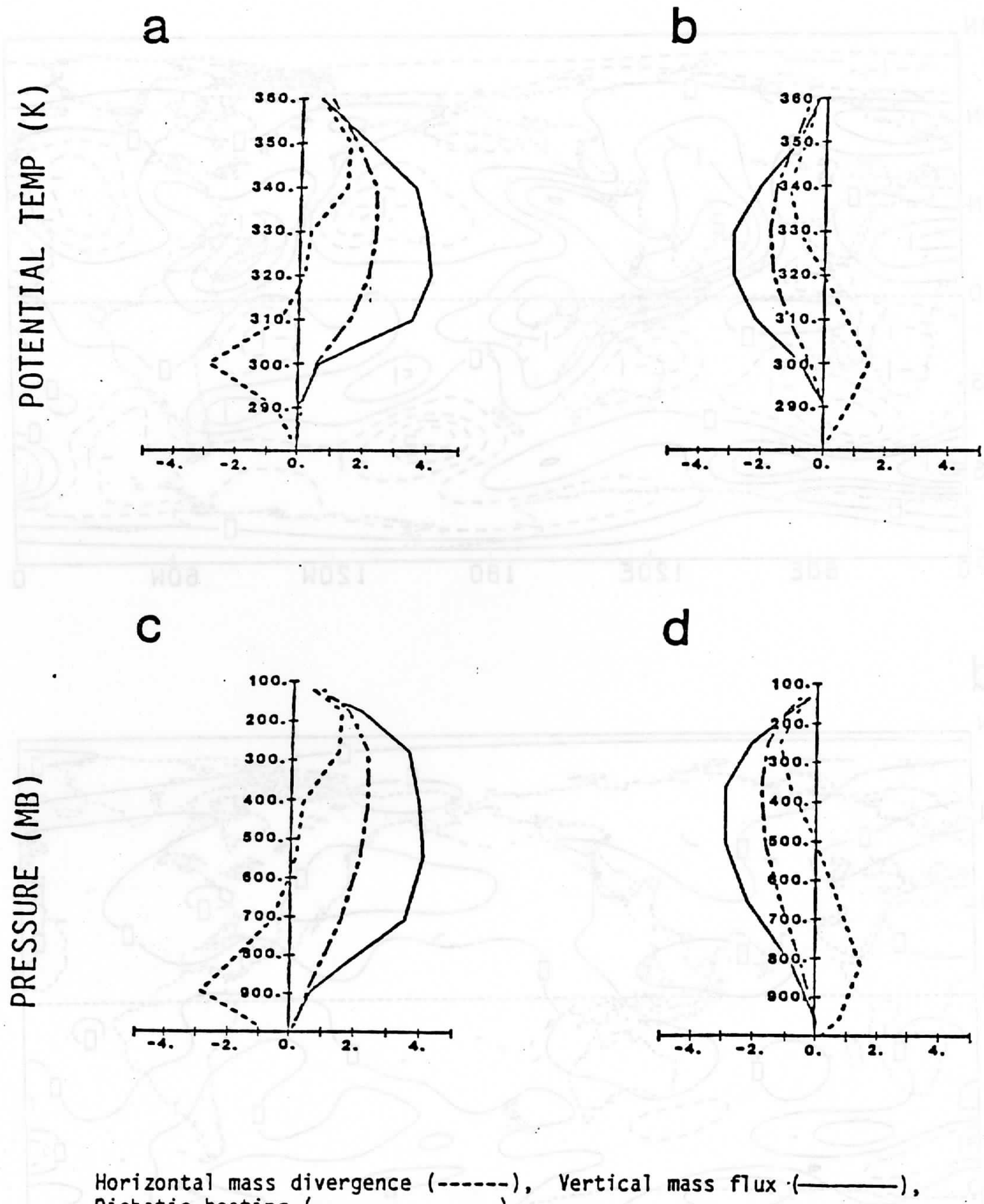


Fig. 15 Area-averaged vertical profiles of horizontal mass divergence (10^{-3} kg m^{-2} 10K^{-1} s^{-1}), vertical mass flux (10^{-3} kg m^{-2} s^{-1}) and diabatic heating (K day^{-1}) for January 1979. Potential temperature is used as ordinate in a and b, and pressure is used in c and d. The area averaged in a and c is from 0N to 15S and from 165E to 175W, roughly east of New Guinea. The area averaged in b and d is from 0N to 15S and from 110W to 90W, roughly over eastern equatorial Pacific.

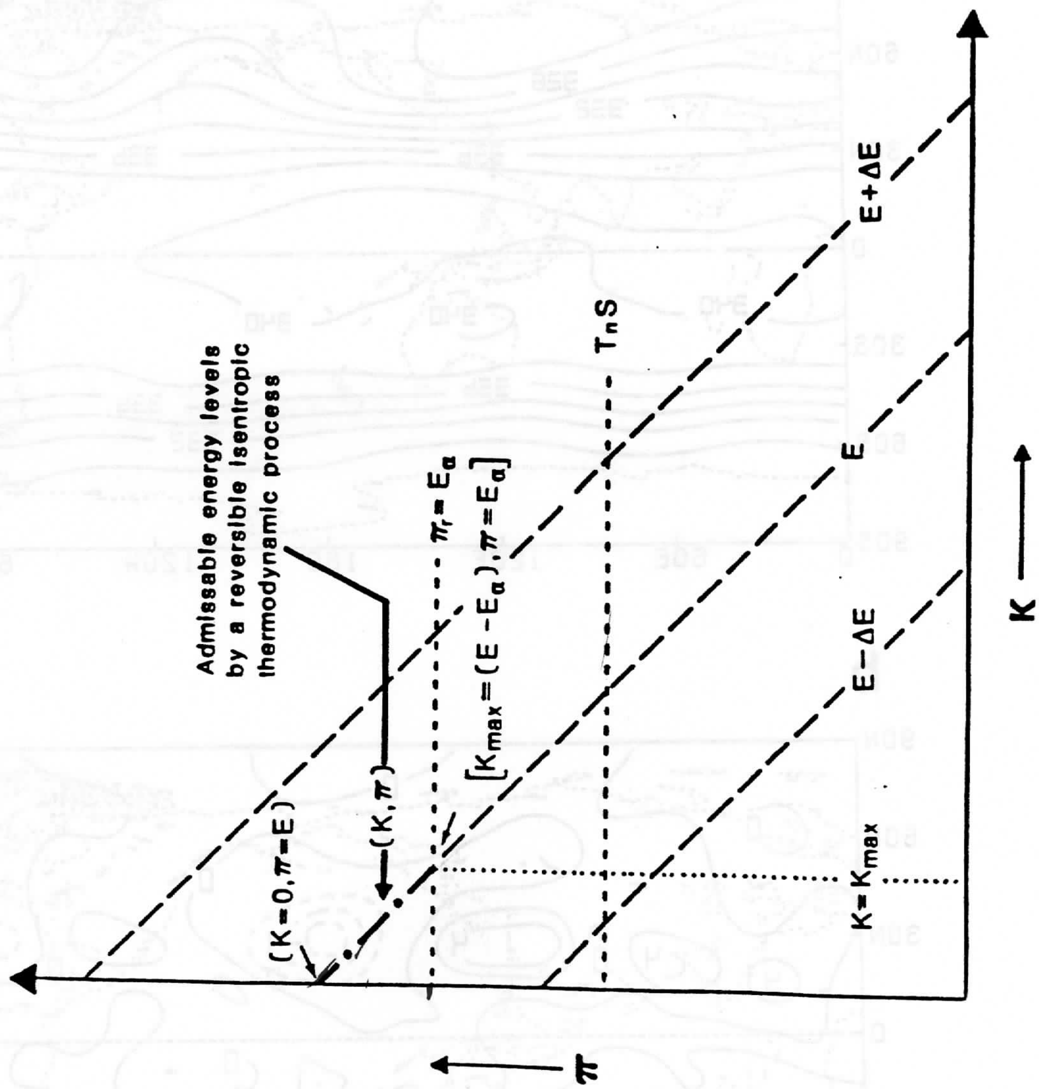


Fig. 16 An energy phase space diagram illustrating the relations between total, total potential, kinetic and available potential energies. The abscissa is kinetic energy, the ordinate is total potential energy, while the family of sloping lines designates isopleths of total energy. The region of admissible energy levels under reversible isentropic energy transformation between kinetic and available potential energies lies above the isopleth of π_r.

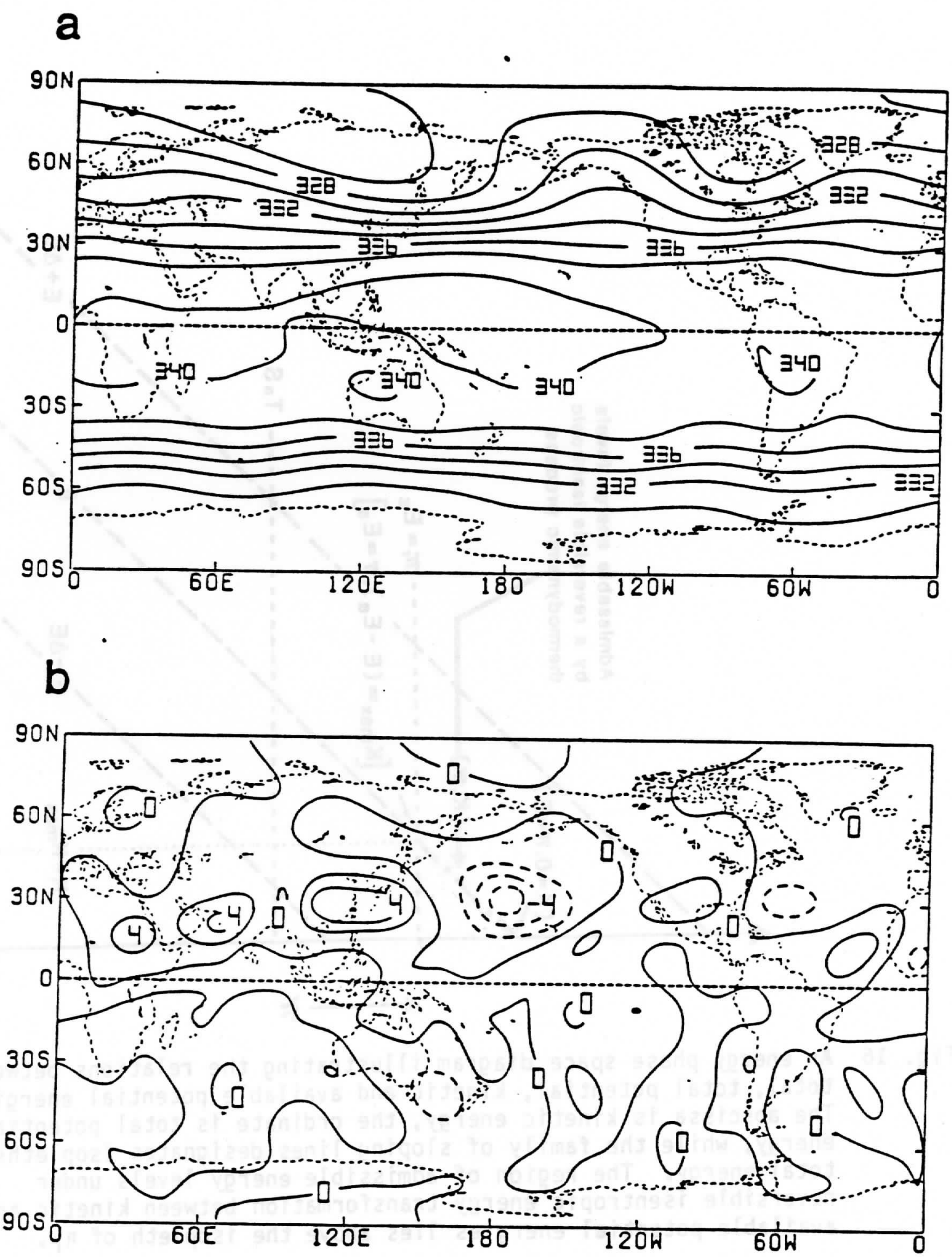
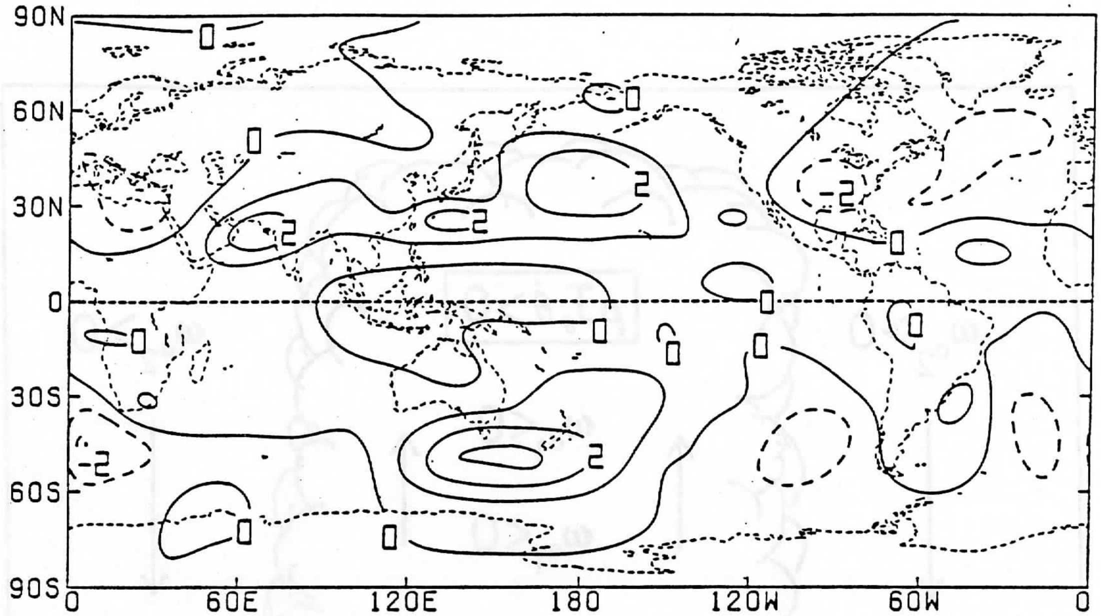


Fig. 17 Time-averaged Montgomery Stream Function (a), kinetic energy generation (b) and standing irrotational (c) and rotational (d) modes of kinetic energy generation for the 340-350 K isentropic layer of January 1979 (Units, a (10^3 m^2 s^{-2}); b, c; and d (10^{-1} kg s^{-3} K^{-1}).

c



d

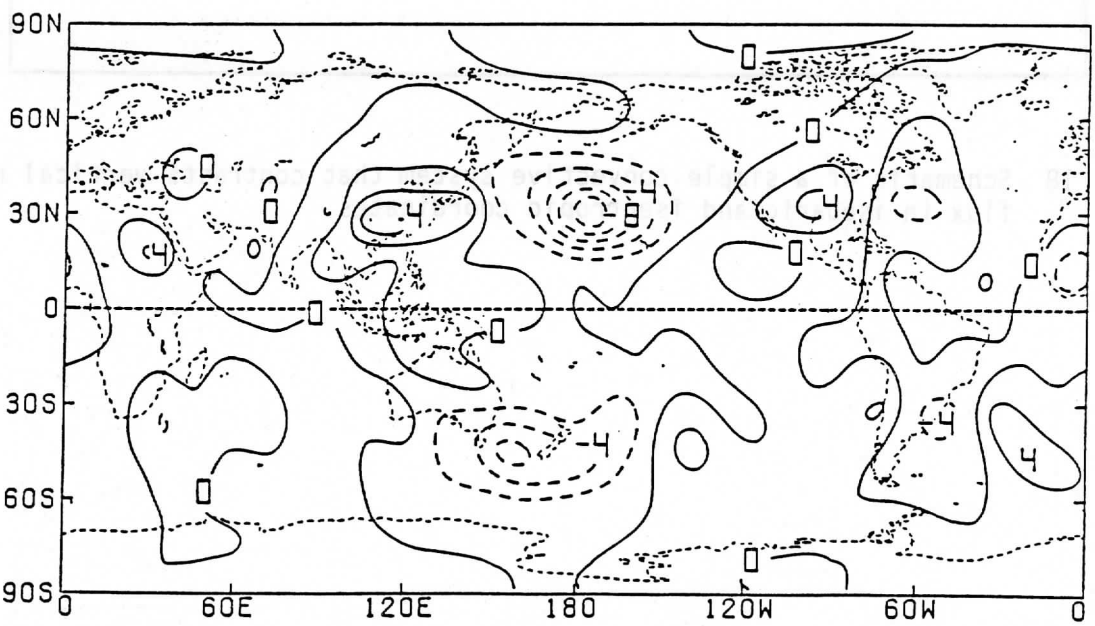


Fig. 17 (Continued)

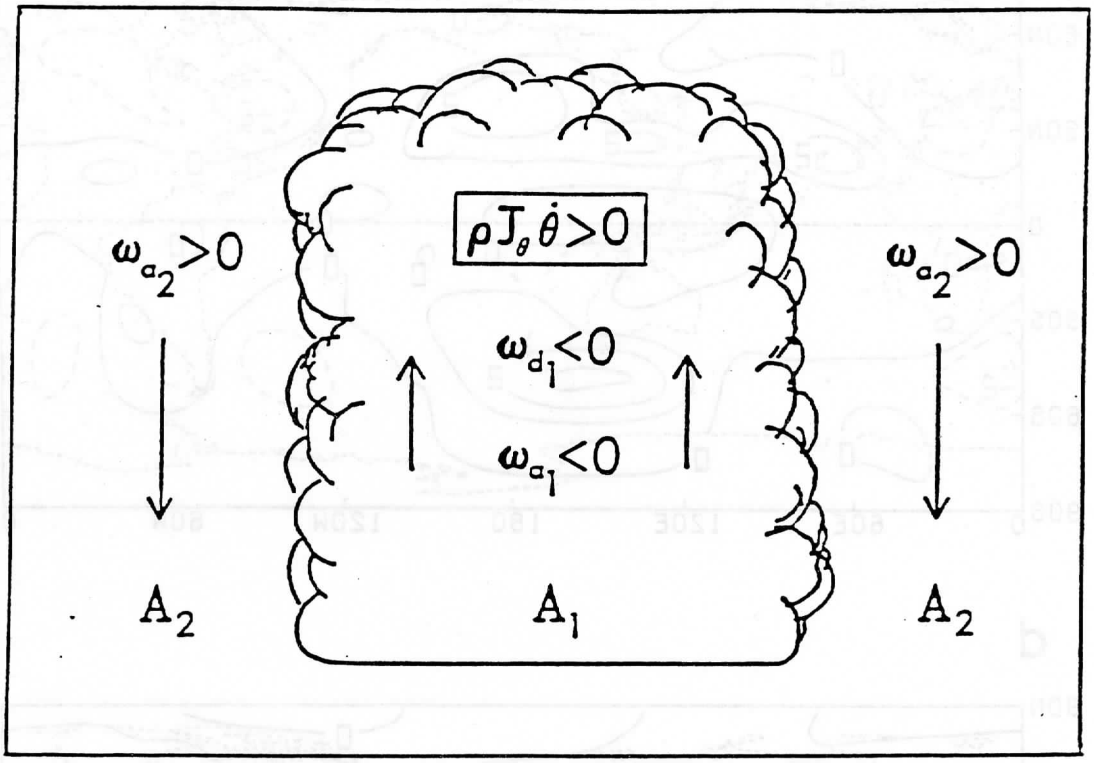


Fig. 18 Schematic of a simple convective system that contrasts vertical mass flux in isobaric and isentropic coordinates.

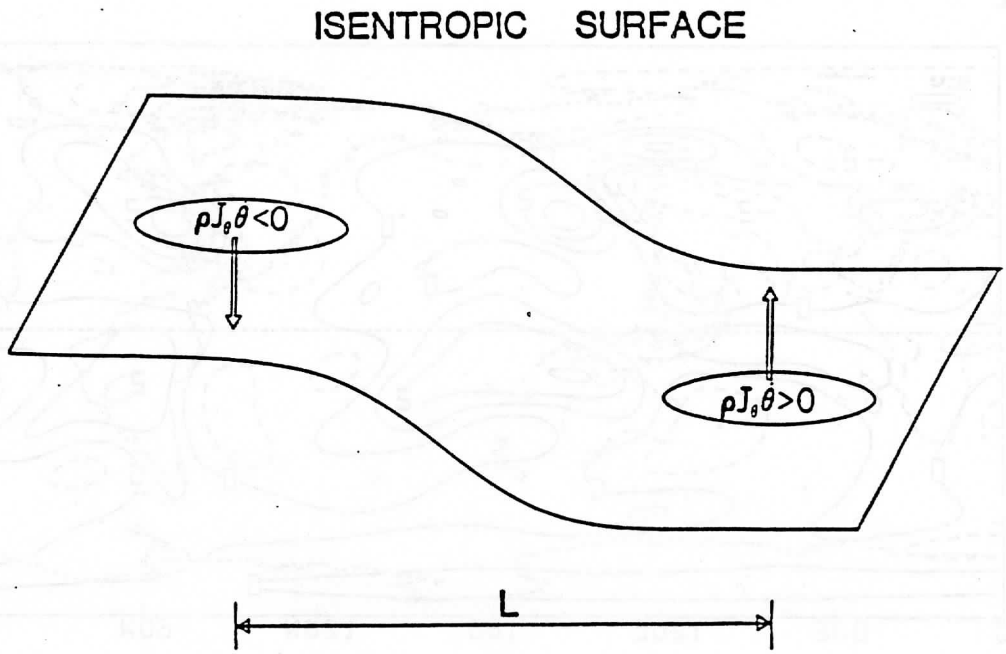


Fig. 19 Schematic to illustrate the determination of length scale of mass transport in isentropic coordinates.

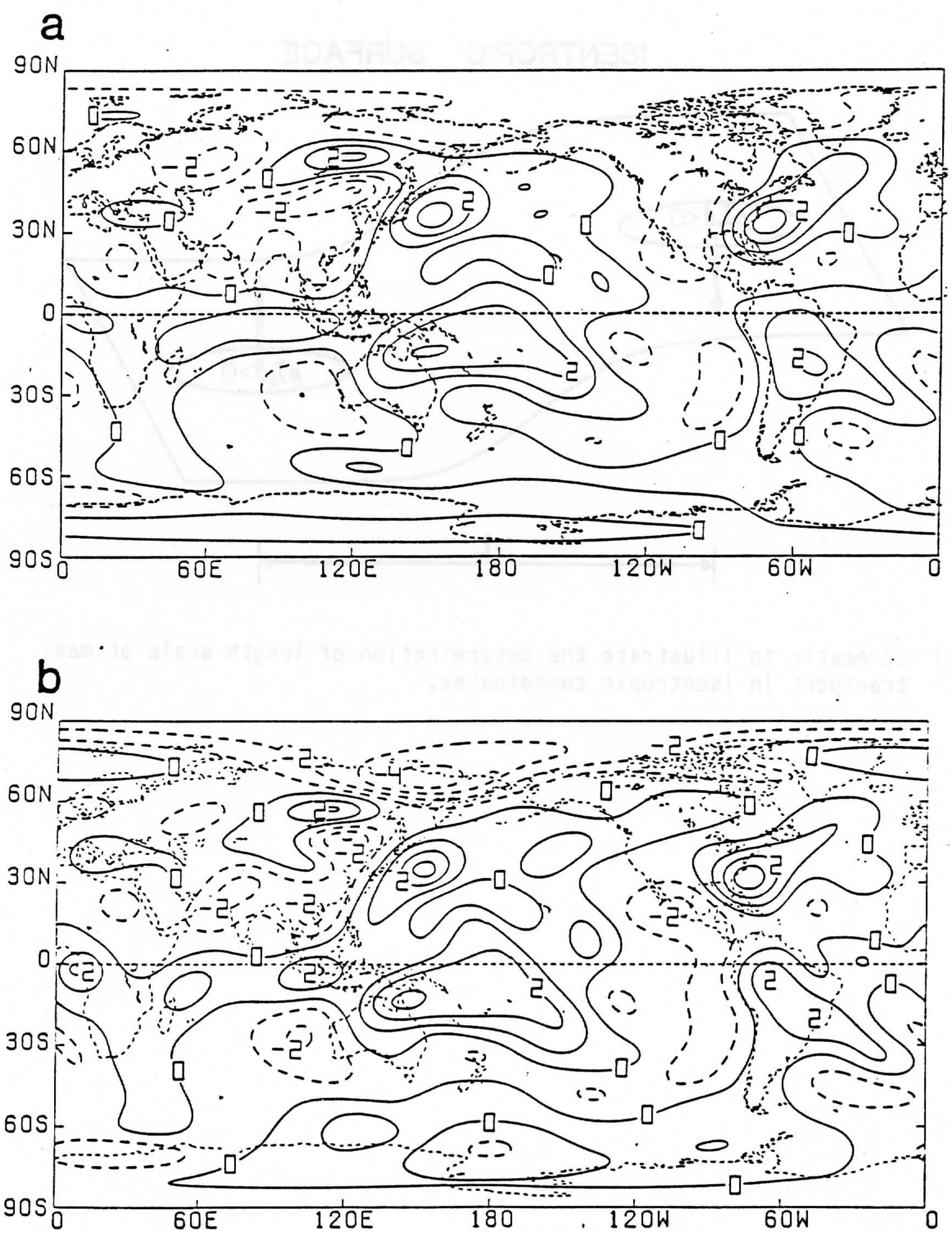


Fig. 20 Global distributions of vertically integrated divergence of energy transport for January 1979 from ECMWF Level III analyses. (a) Integration in isentropic coordinates, (b), (c) and (d) show integration in isobaric coordinates for total divergence, standing component and transient component (10^2 watts m^{-2}), respectively.

c

d

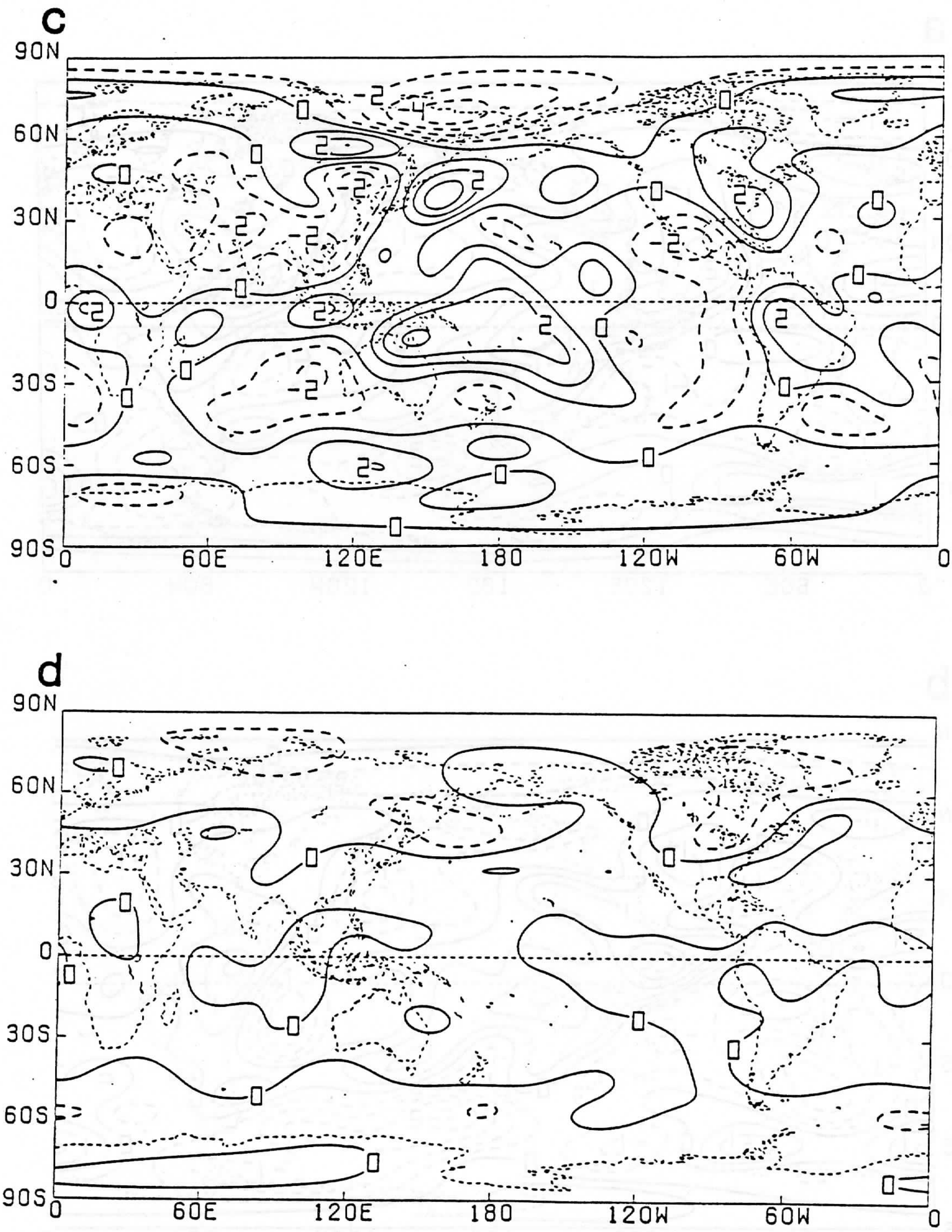


Fig. 20 (Continued)

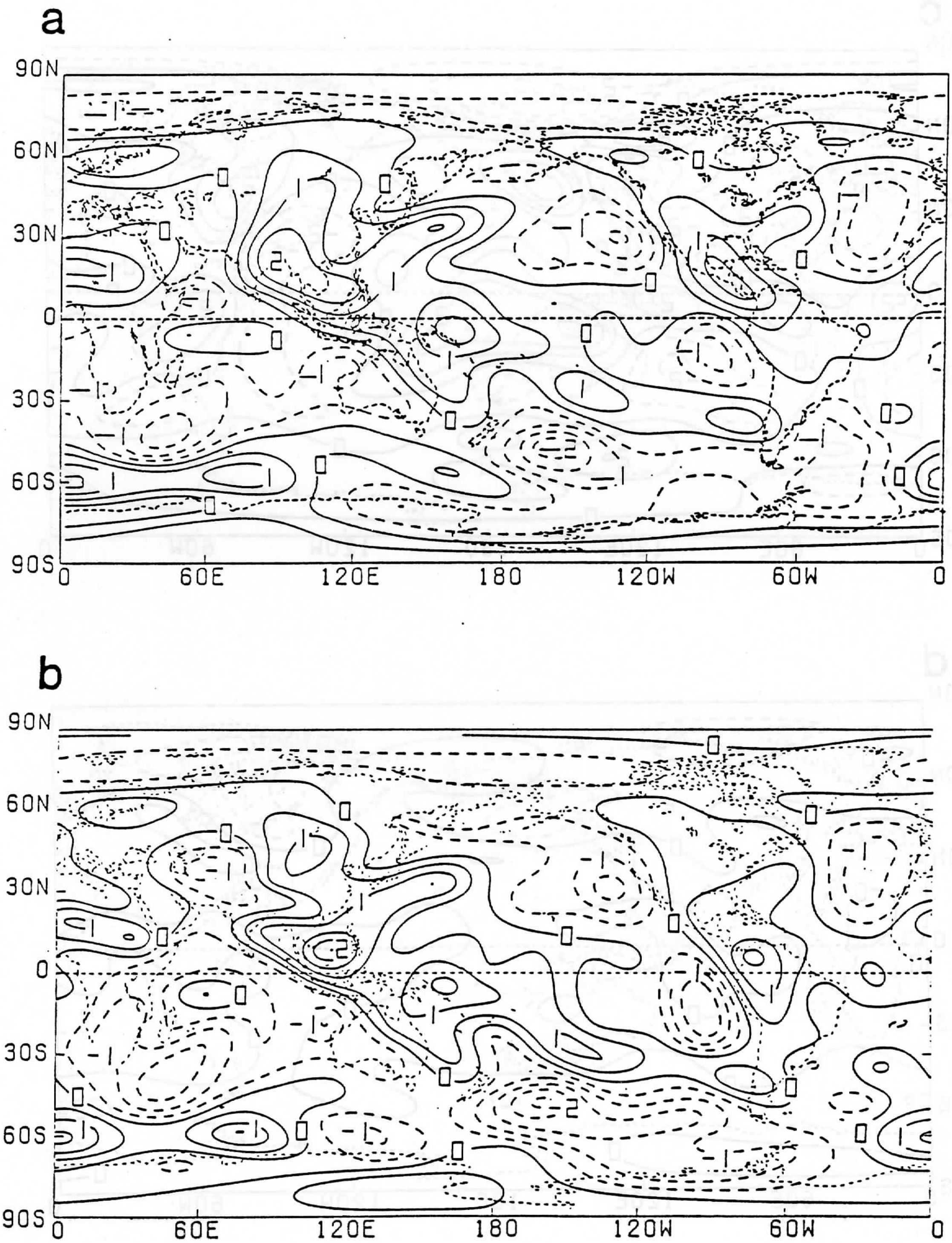
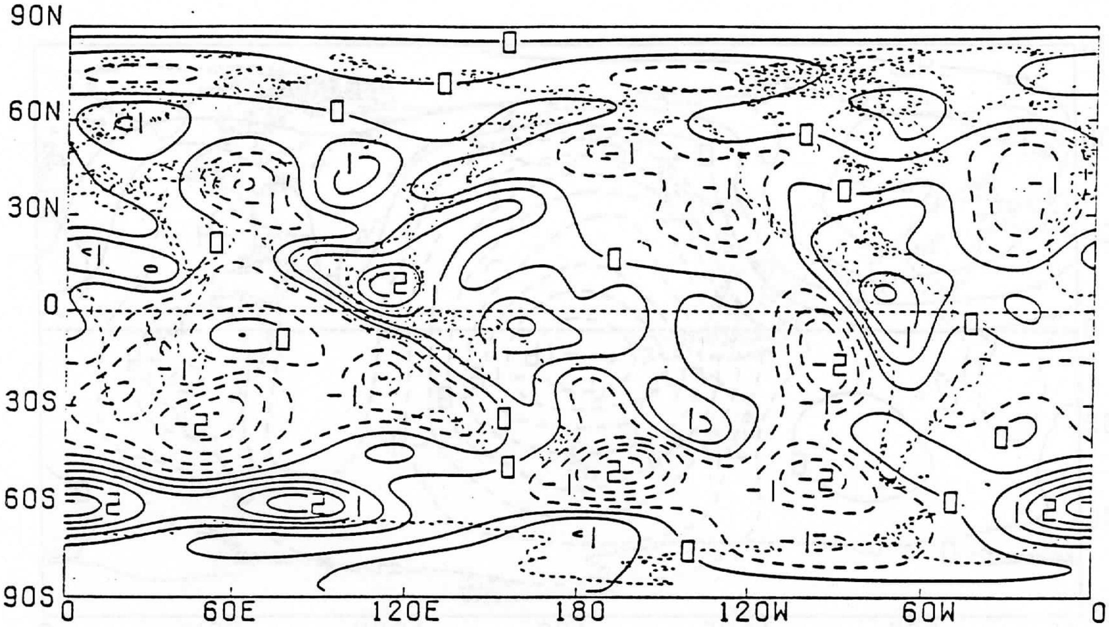


Fig. 21 Same as Fig. 20 except for July 1979.

c



d

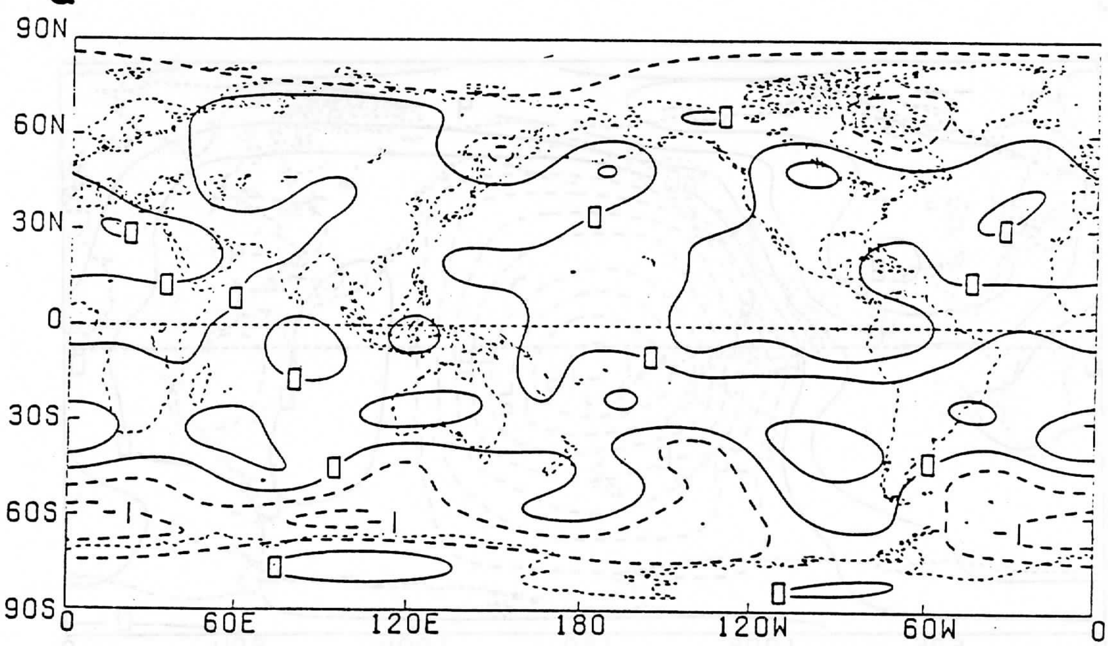
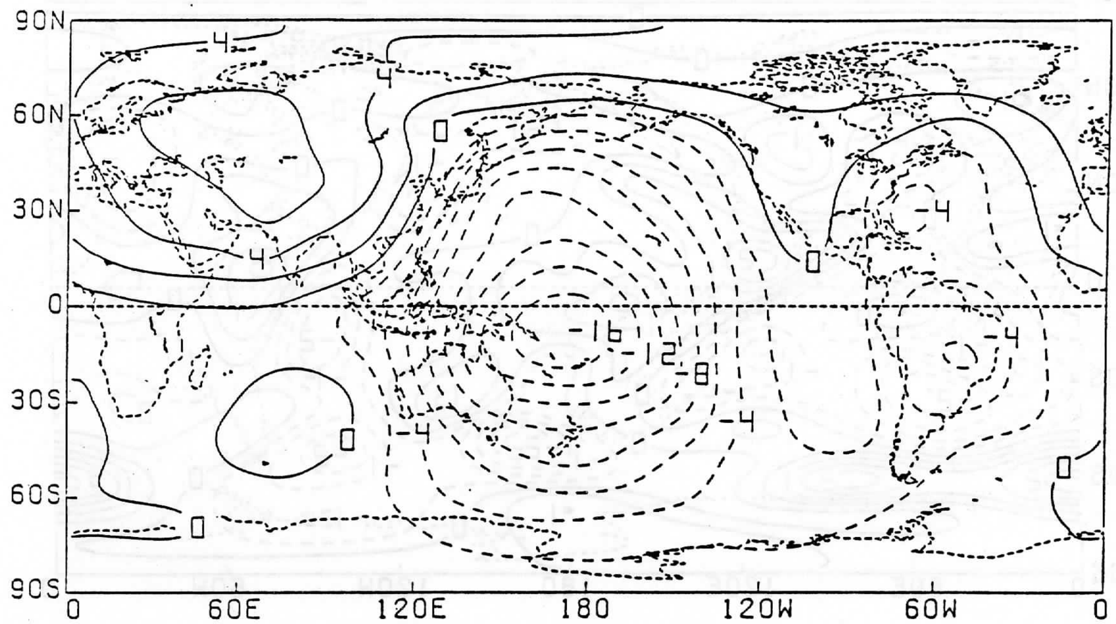


Fig. 21 (Continued)

a



b

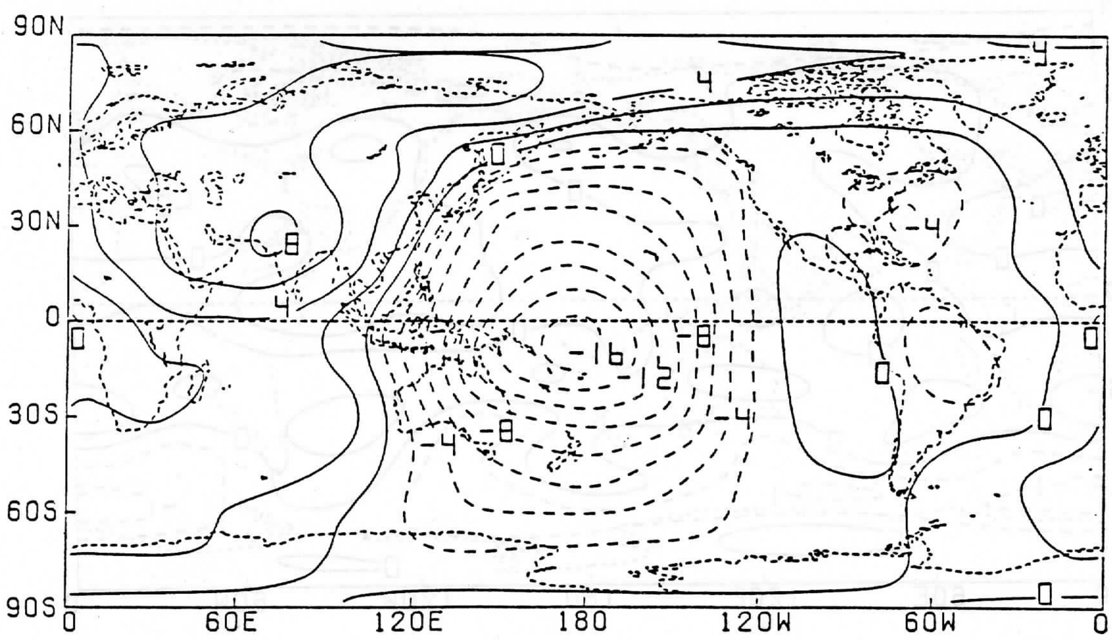
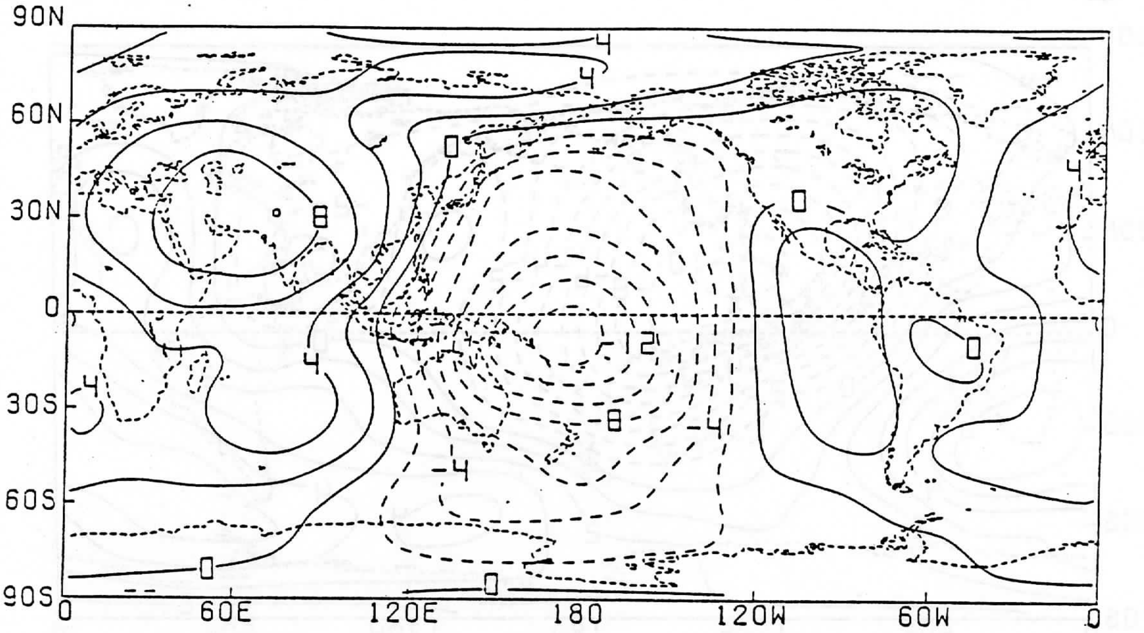


Fig. 22 Potential functions determined for January 1979 from the corresponding global distribution of isentropic and isobaric energy divergence fields shown in Fig. 20 (units, a, b and c (10^{14} J s^{-1}), d (10^{13} J s^{-1})).

c



d

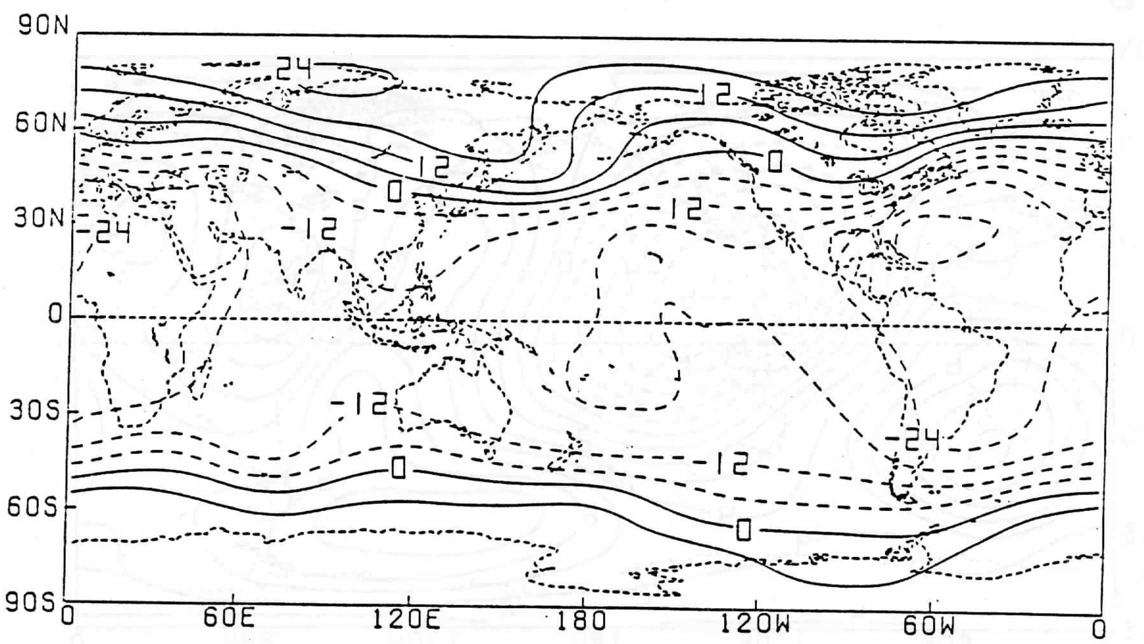


Fig. 22 (Continued)

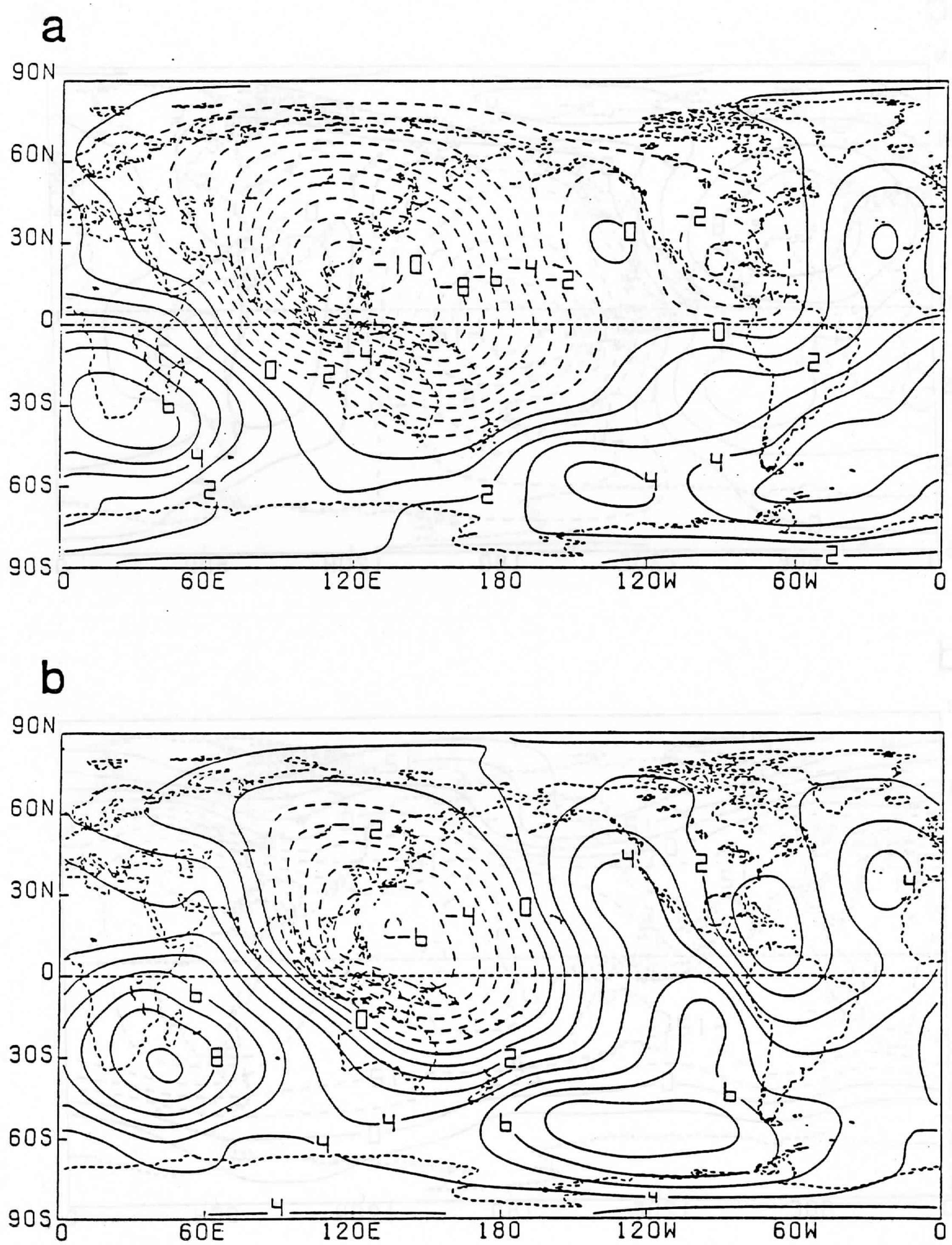
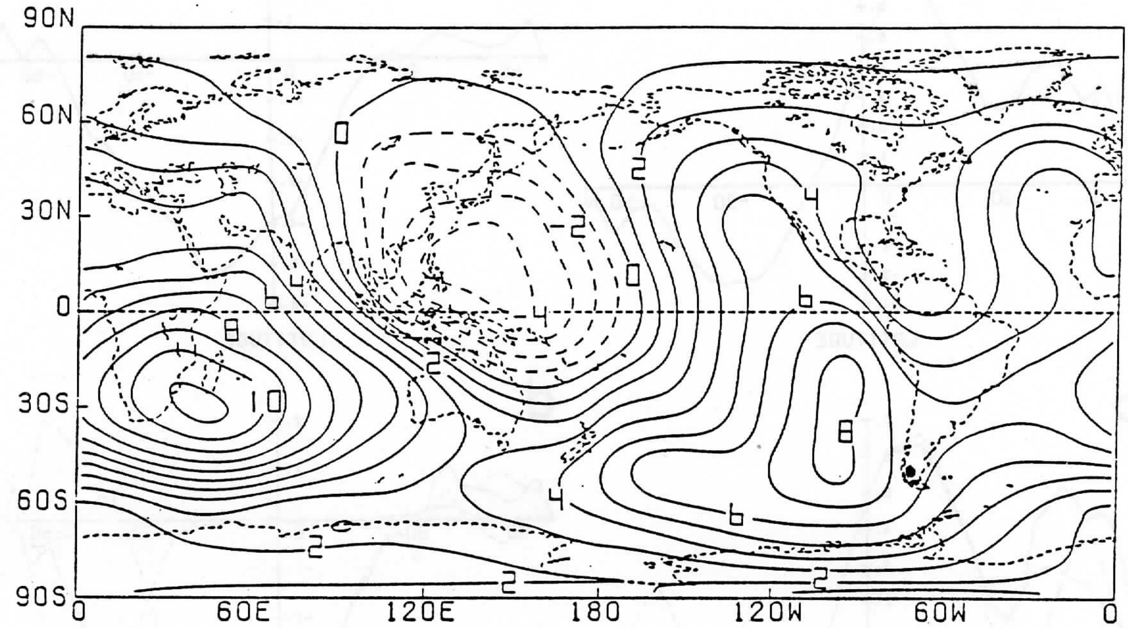


Fig. 23 Potential functions for July 1979 determined from the corresponding global distribution of isentropic and isobaric energy divergence fields shown in Fig. 21 (units a, b and c (10^{14} J s^{-1}), d (10^{13} J s^{-1})).

c



d

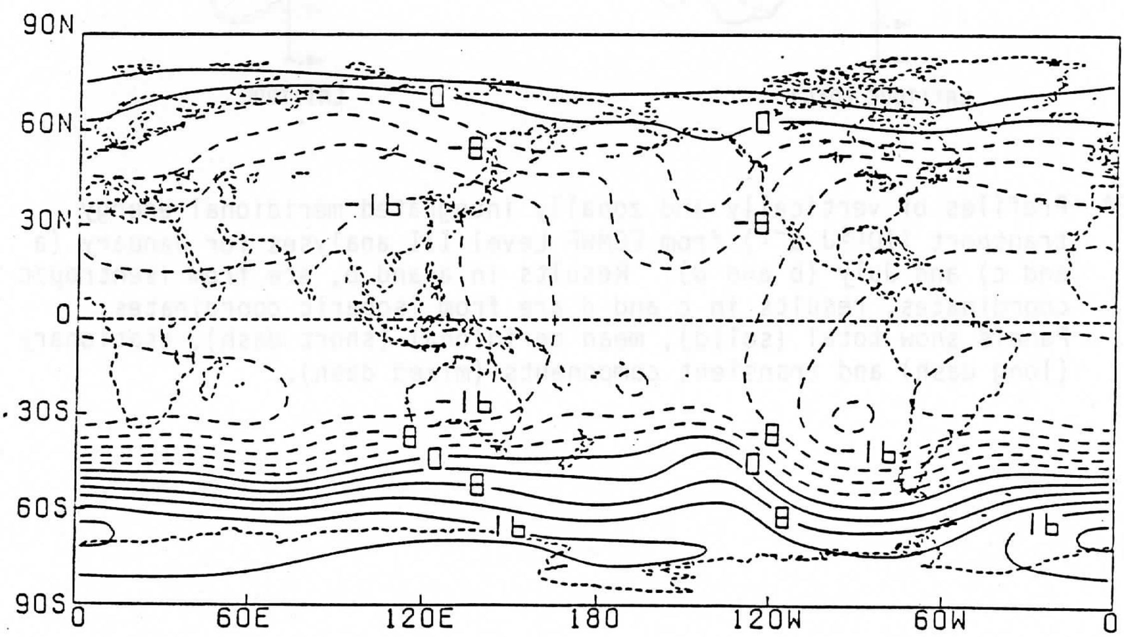


Fig. 23 (Continued)

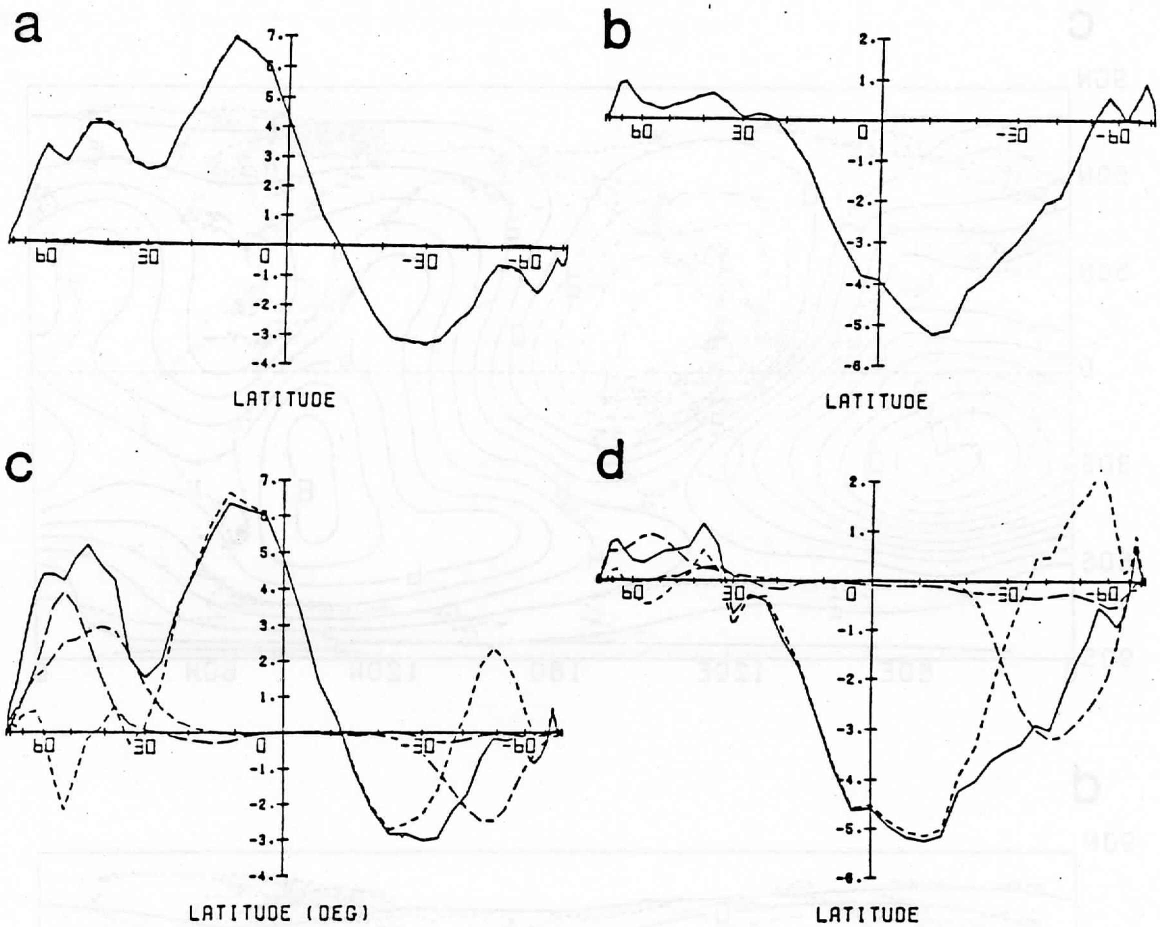


Fig. 24 Profiles of vertically and zonally integrated meridional energy transport (10^{15}J s^{-1}) from ECMWF Level III analyses for January (a and c) and July (b and d). Results in a and b, are from isentropic coordinates, results in c and d are from isobaric coordinates. Panels show total (solid), mean meridional (short dash), stationary (long dash) and transient components (mixed dash).

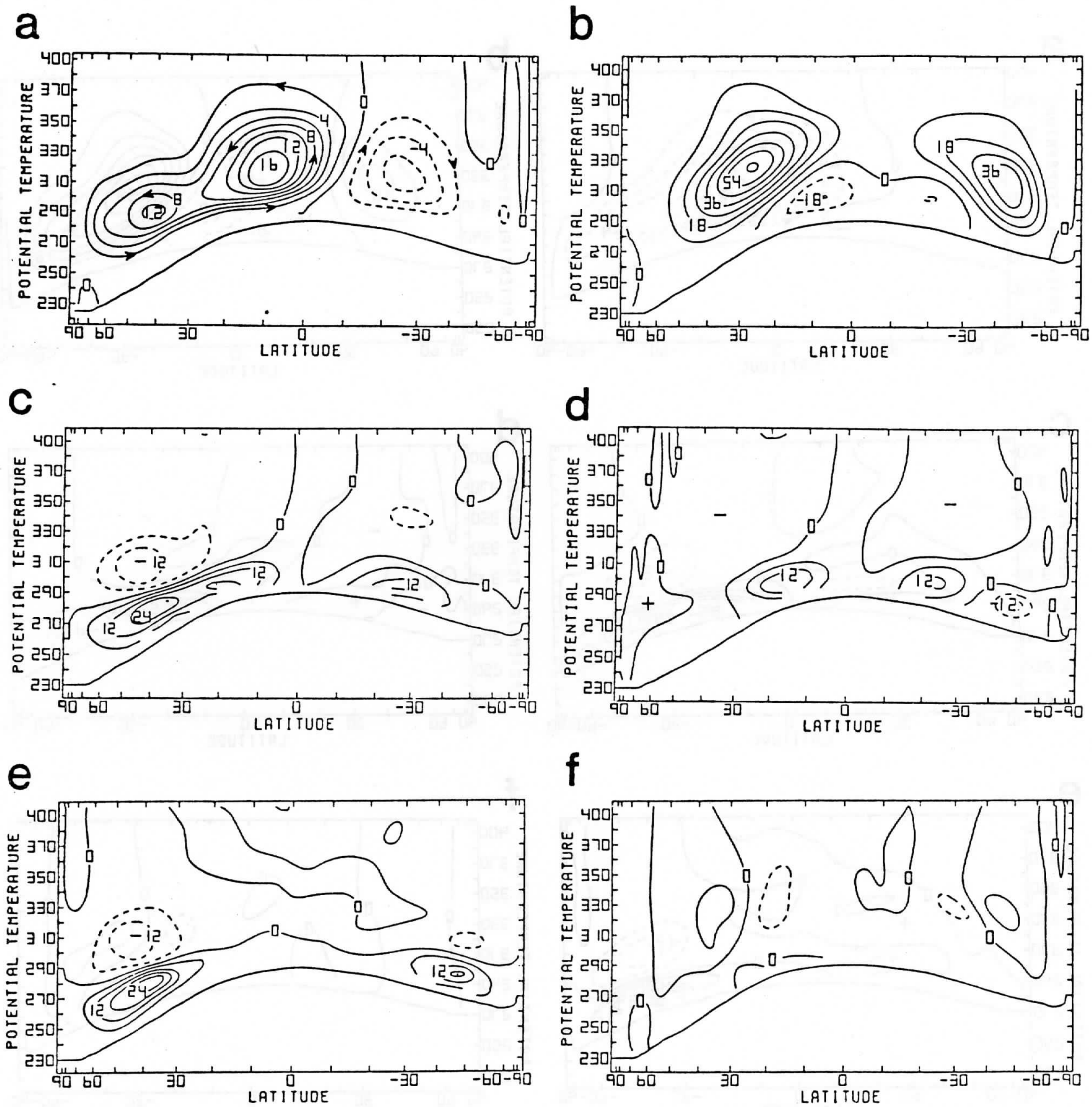


Fig. 25 Meridional cross sections of: (a) mean meridional circulation (10^{10} kg s^{-1}), (b) zonally integrated relative angular momentum (10^{16} kg m s^{-1} K^{-1}) and zonally integrated (c) total torques, (d) frictional torque, (e) pressure torque, and (f) forcing by horizontal and vertical divergence of eddy angular momentum transport (10^{11} kg m s^{-2} K^{-1}) for January of the FGGE year from ECMWF Level III analyses.

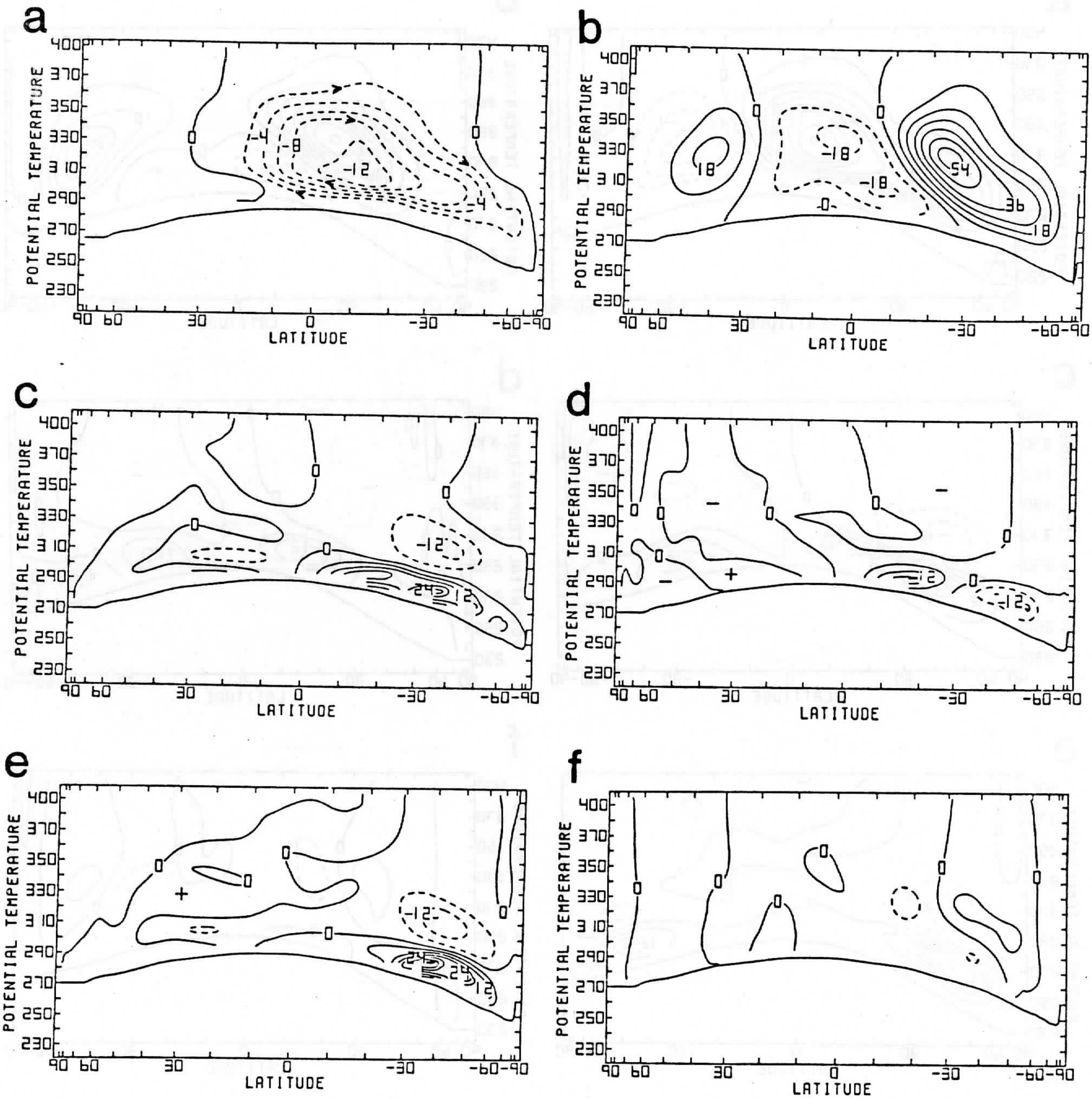


Fig. 26 Same as Fig. 25 except for July 1979.

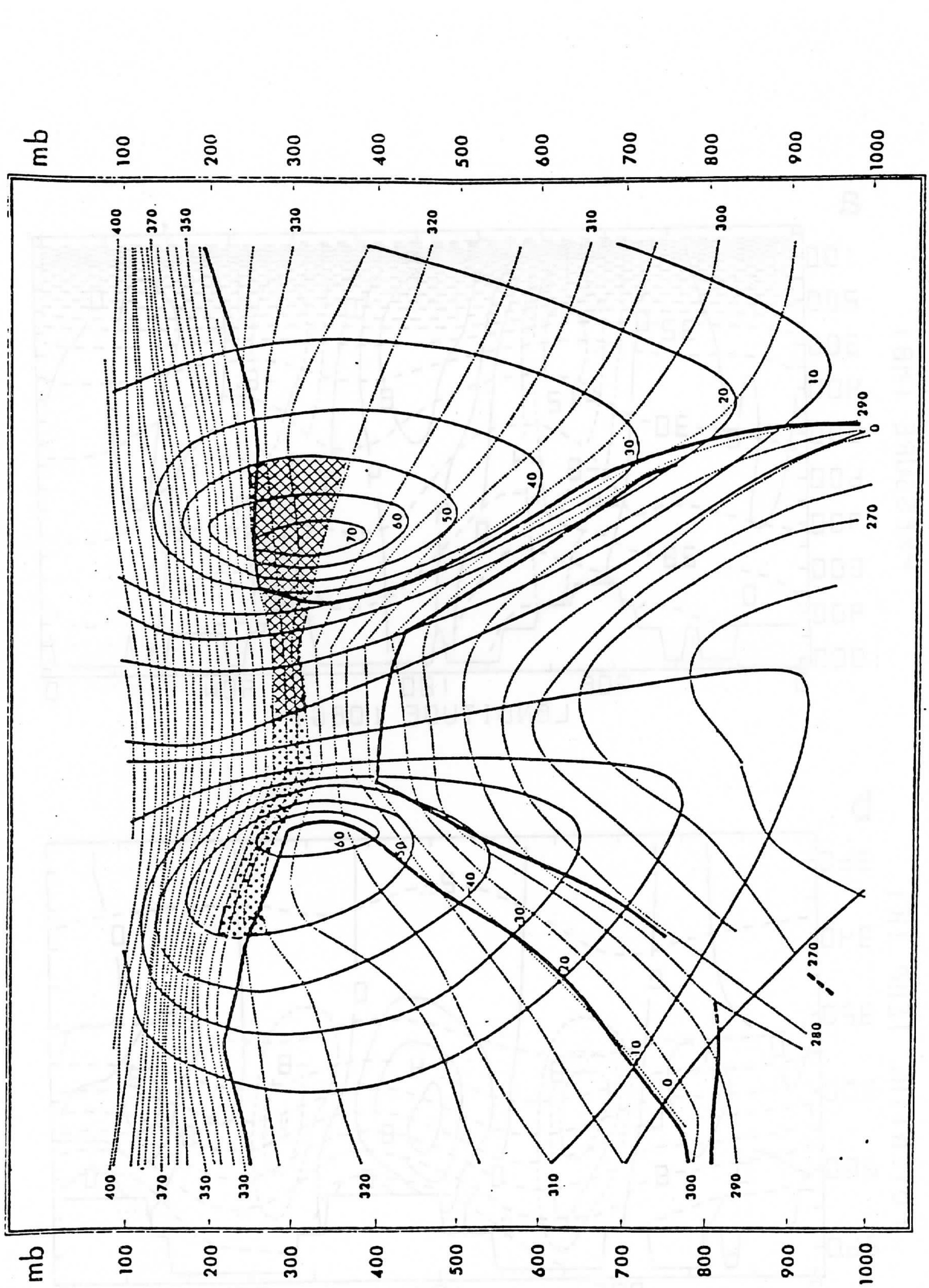


Fig. 27 Vertical cross section from Tucson to Bermuda through a large amplitude wave studied by Newton and Palmén (1963). Dotted lines are isentropes, solid lines are isotachs of the meridional wind while heavy lines are frontal boundaries and tropopause. Within the 320-335 K isentropic layer the cross hatched area denotes poleward mass transport while the dotted area denotes equatorward mass transport. The larger area of cross-hatching indicates the positive covariance of mass and meridional motion embedded in wave structure.

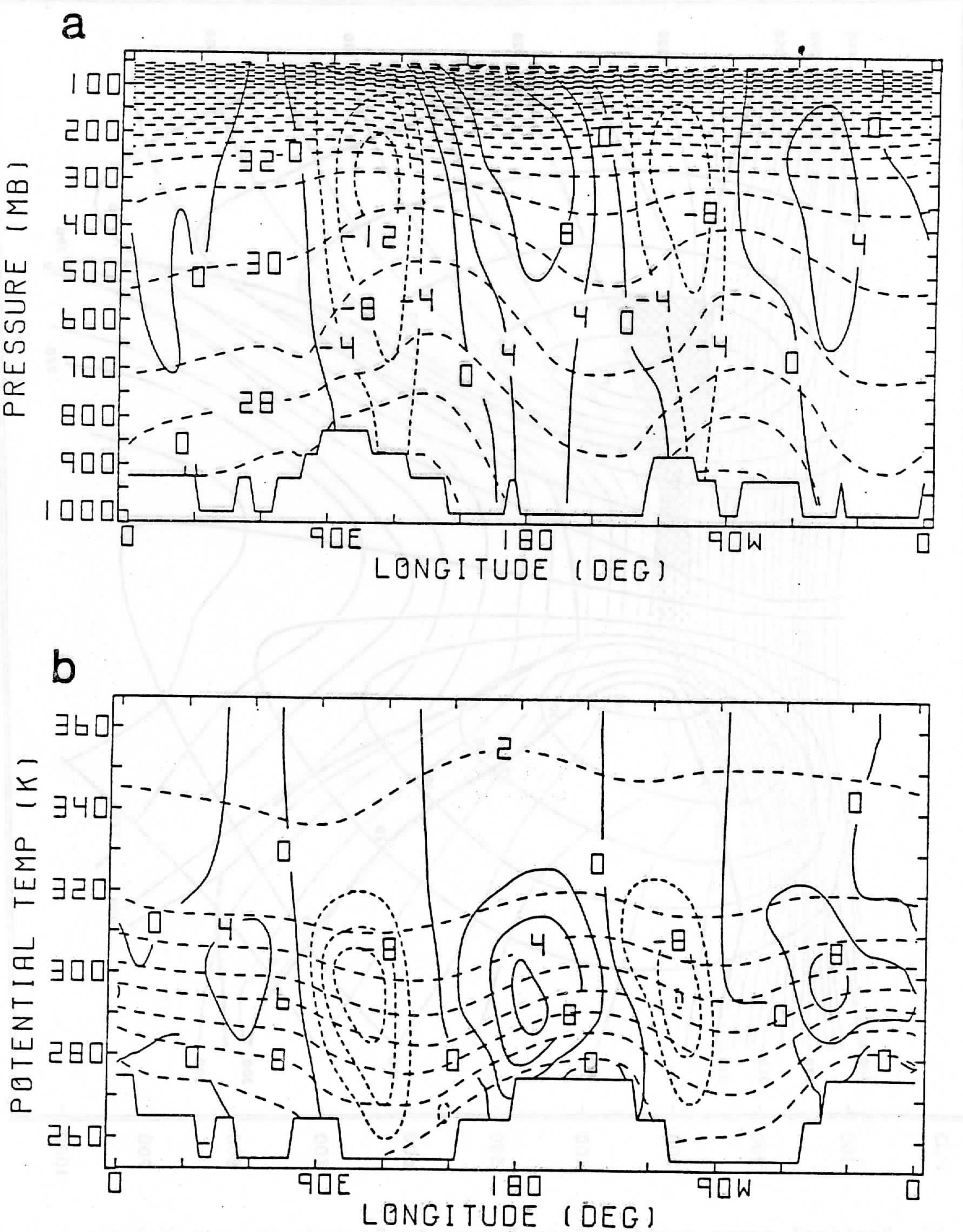


Fig. 28 Zonal cross sections along 50°N of (a) Time-averaged potential temperature, $\overline{\theta}^t$ (units, 10 K), and meridional velocity $\overline{v^t}$ (units, m s^{-1}) in isobaric coordinates; and (b) Time-averaged pressure, $\overline{p}^t$ (units, 10^2 mb), and meridional mass transport $\overline{\rho J_\theta v^t}$ (units, 10^2 kg m^{-1} K^{-1} s^{-1}), in isentropic coordinates for December, January and February of the GWE year.