Fronts and Frontogenesis

Jonathan E. Martin
Department of Atmospheric and Oceanic Sciences, University of Wisconsin-Madison, Madison, WI, USA

Synopsis

Extratropical cyclones are characterized by relatively narrow zones of sharp contrast in temperature and moisture content that extend hundreds, even thousands, of kilometers outward from the center of lowest sea-level pressure (SLP). These zones are manifest as elongated strips of cyclonic vorticity indicative of their association with coherent winds shifts along their lengths. Such features have been known as "fronts" for over a century. A large fraction of the cloud and precipitation associated with extratropical cyclones is generated in the vicinity of these fronts by virtue of their connection to the dynamics of frontogenesis. In this chapter the structural characteristics of these fronts are connected to the dynamics that govern their production and sustenance to reveal fundamental physical explanations for this important and ubiquitous relationship.

Keywords

Front, frontal zone, frontogenesis, vertical circulation, thermally direct, thermally indirect, cloud and precipitation distribution

Key Points

• Frontal zones are characterized by large temperature contrasts, large vorticity, and large static stability.

• The process of frontogenesis (i.e. increasing the horizontal temperature contrast associated with a front) is necessarily associated with a vertical circulation

• Divergent motions associated with this vertical circulation also contribute to intensifying the temperature contrast – a reality that is better represented in the semi-geostrophic than in the quasi-geostrophic equations

Introduction

A defining structural feature of the extratropical cyclone is its asymmetric thermal structure manifest most clearly in the concentrated zones of thermal contrast, known as fronts, that characterize the cyclone. The existence of such features had been suggested by a number of thinkers as far back as the late 18th century (see the historical review by Kutzbach (2016)), but only systematically confirmed in the last ~100 years. Just after the end of WWI, meteorologists at the University of Bergen in Norway, under the leadership of Vilhelm Bjerknes, developed the Polar Front Theory of the structure and life cycle of mid-latitude cyclones, now known colloquially as the Norwegian Cyclone Model (NCM, Bjerknes and Solberg, 1922). The essential genius of this conceptual model, which represented a grand synthesis of prior insights concerning the cyclone, was that it described the instantaneous structure of the cyclone while placing that structure into an identifiable life cycle. The central organizing feature of the NCM was the existence of a tropospheric deep, globe-girdling, knife-like boundary known as the polar front which separated cold polar air from warm tropical air. It was believed that perturbation vortices occasionally developed along this polar front and that their attendant circulations would then serve to locally deform it, ushering tropical air poleward and polar air equatorward. The physical mechanism by which such perturbation vortices might grow in intensity was not well explained in the NCM, but the continued growth of the perturbation was thought to lead to further deformation of the polar front as well as a lower sea-level pressure (SLP) at the center of the perturbation. By the so-called mature stage of the life cycle, the deformation of the polar front had become so extreme as to lend the cyclone its now familiar characteristic frontal structure; a cold front extending equatorward and a warm front extending eastward from the sea-level pressure minimum. The region of nearly homogeneous temperature between the two fronts

was known as the warm sector. Continued intensification of the cyclone and its circulation compelled the cold front to encroach upon, and subsequently overtake, the warm front. Two important results of this process were 1) that the sea-level pressure minimum was removed from the peak of the warm sector and 2) a 3D thermal structure developed that left a surface boundary known as the occluded front to connect the SLP minimum to the peak of the warm sector. It was thought that this process could result in the development of two varieties of occluded structures in cyclones. One of these was the so-called warm occlusion in which the cold front would ascend the warm front upon overtaking it. Conversely, a so-called cold occlusion would result if the encroaching cold front was able to undercut the warm front. The warm (cold) occlusion was thought to occur when the air poleward of the warm front was more (less) dense than the air west of the cold front. Note that in either case, the development of the occluded structure was associated with the denser air lifting the less dense air aloft. In so doing, the horizontal density contrast originally characterizing the cyclone (manifest in the horizontal temperature gradient associated with the Polar Front) was reduced and a stable vertical stratification near the cyclone center was gradually put in place. Such a transformation reduces the center of gravity of a fluid system gradually driving the system to its lowest potential energy state[1]. Based upon this type of energetics argument, the NCM proposed that the development of the occluded front heralded the post-mature phase for a mid-latitude cyclone, a cessation of intensification and the commencement of cyclone decay.

The NCM accounted for the typical cloud and precipitation distributions associated with a mid-latitude cyclone with reference to the vertical structure of the fronts themselves. The cold

[1] Stoelinga et al. 2002 make clear that it is static stability differences, not density differences, that determine the type of occluded structure that will develop.

front was described as a steeply sloped boundary between polar and tropical air masses whose steady advance into the tropical air produced upgliding motions along the boundary itself. As a consequence of its steep slope, the updrafts were vigorous and horizontally restricted leading to a narrow, sometimes squally precipitation distribution. The warm front, on the other hand, was a less steeply sloped boundary between advancing tropical air and gradually retreating polar air. The upgliding motions along the warm frontal surface, less intense as a consequence of its shallower slope, were thought to underlie the more horizontally widespread cloudiness and more benign precipitation associated with the warm front.

Thus, since the introduction of the NCM it has been well known that, in addition to their ubiquity, these fronts are also vested with considerable sensible weather relevance as large variations of meteorological conditions exist across them and the precipitation distribution associated with a typical extratropical cyclone is often concentrated in their vicinity. Figure 1a shows analyses of the sea-level pressure and surface potential temperature for a typical mid-latitude cyclone. The characteristic comma shaped cloud pattern from the same storm (Fig. 1b) is anchored by the frontal structure identified in Fig. 1a. Daily inspection of surface, upper air, and satellite observations easily confirms that the structural relationship demonstrated in Fig. 1 is quite common in the mid-latitudes.

Importantly, the across-front dimension of the cold frontal zone in Fig. 1a (on the order of 100 km) is much smaller than its along-front dimension (on the order of 1000 km). Considering characteristic velocities given such length scales it is safe to assume that geostrophic balance exists in the along-front direction (where the Rossby Number (R_o) is given by R_o=10 m s^{-1}/(10^{-4} s^{-1})(10^6 m) = 0.1. However, the environment in the across-front direction strays from geostrophy as R_o = (10 m s^{-1})/(10^{-4} s^{-1})(10^5 m) = 1.0. Thus, mid-latitude fronts are hybrid

phenomena characterized by along-front geostrophy but a fair degree of across-front ageostrophy. This mixture of scales that characterizes fronts makes them the focus of important scale interactions in the mid-latitude cyclone. For this reason, a purely quasi-geostrophic diagnostic perspective is insufficient as a means to investigate fronts and must be extended in order to incorporate additional, physically relevant processes that are fundamental to the frontal environment.

In the face of the coincidence of fronts and the cloud and precipitation distribution in mid-latitude cyclones suggested by Fig. 1, a fundamental question is *what underlies this relationship*? Fashioning a satisfying answer requires gaining an understanding of the essential elements of frontal structure and dynamics. Subsequently, examination of *frontogenesis*, the process of strengthening a horizontal temperature contrast, leads to an understanding of the nature and origin of the vertical motions that characterize fronts. Adoption of a *semi-geostrophic* perspective in the Sawyer-Eliassen frontal circulation equation formally incorporates the previously mentioned interplay between the geostrophic and ageostrophic flows that characterize the frontal environment. Finally, an investigation of fronts that form at the tropopause, known as upper-level jet/front systems or upper-level fronts, will round out this chapter. First, analysis of simple conceptual models can assist in identifying the essential characteristics of fronts.

The Structural and Dynamical Characteristics of Mid-latitude Fronts

A front is a boundary whose primary structural characteristic is the larger-than-background temperature (or density) contrast associated with it. In order to determine some basic characteristics of fronts, from which a working definition of a front can be formulated, it is

useful to consider the somewhat unphysical case of the *zero order front*; that is, one that is characterized by *discontinuities* in the temperature and density across the frontal boundary (as shown in Fig. 2). Real fronts actually more closely resemble a First Order Front, in which *gradients* of temperature and density, not the variables themselves, are discontinuous across the front. For heuristic purposes, we will demand that pressure be continuous across the zero order front in Fig. 2 (so that the geostrophic winds are not infinite along the front). If we take the *x*-axis as the along-front direction and further assume 1) that there is no along-front variation in any variable, and 2) that the pressure is steady state (i.e. $\frac{\partial p}{\partial t} = 0$), then the differential of pressure is given by

$$dp = (\frac{\partial p}{\partial y})dy + \left(\frac{\partial p}{\partial z}\right) dz \qquad (1)$$

which can be written, separately, for both the warm and the cold sides of the front as

$$dp_w = (\frac{\partial p}{\partial y})_w dy + \left(\frac{\partial p}{\partial z}\right)_w dz \quad \text{and} \quad dp_c = (\frac{\partial p}{\partial y})_c dy + \left(\frac{\partial p}{\partial z}\right)_c dz \,,$$

respectively. The hydrostatic equation can be substituted for $\frac{\partial p}{\partial z}$ in both expressions and the expressions for dp can be set equal to one another. Upon rearrangement the result is,

$$0 = \left[\left(\frac{\partial p}{\partial y}\right)_c - \left(\frac{\partial p}{\partial y}\right)_w\right] dy - (\rho_c - \rho_w) g dz. \qquad (2)$$

This can be solved for $\frac{dz}{dy}$, the slope of the zero order front:

$$\frac{dz}{dy} = \frac{(\frac{\partial p}{\partial y})_c - (\frac{\partial p}{\partial y})_w}{g(\rho_c - \rho_w)}. \qquad (3)$$

Since more dense fluid must lie beneath less dense fluid, as portrayed in Fig. 2, in order that the frontal structure be statically stable and therefore sustainable, we note that $\frac{dz}{dy} > 0$. Considering the along-front geostrophic winds, which are related to the across-front pressure gradients, (3) can be recast into

$$\frac{dz}{dy} = \frac{f(\rho_w u_{gw} - \rho_c u_{gc})}{g(\rho_c - \rho_w)} \qquad (4)$$

from which it is clear that, in order for $\frac{dz}{dy} > 0$, $u_{gw} > u_{gc}$; in other words the front must be characterized by positive geostrophic relative vorticity (i.e. $\frac{\partial u_g}{\partial y} < 0$). Thus, a fundamental *dynamical* characteristic of mid-latitude fronts is that they are characterized by positive geostrophic relative vorticity. In fact, (4) suggests that the stronger the density (temperature) contrast across the front, the more intense is the frontal vorticity.

In reality, the temperature cannot be discontinuous at a front, but the temperature *gradient* can be. In this more realistic case of the First Order Front the isentropes must appear as in Fig. 3. Careful examination of the isentropes in the requisite *frontal zone* reveals that the frontal zone is also characterized by larger static stability ($-\frac{\partial \theta}{\partial p}$) than either the cold or warm side of the boundary. Thus, frontal zones are characterized by 1) larger-than-background horizontal temperature (density) contrasts, 2) larger-than-background relative vorticity, and 3) larger-than-background static stability.

> Based upon these essential characteristics a reasonable working definition of a front is:
>
> *The leading edge of a transitional zone, whose length is significantly greater than its width, that separates advancing cold (warm) air from warm (cold) air. The zone is characterized by high static stability, larger-than-background vorticity and a larger-than-background gradient in temperature.*

Fronts, thus defined, come in varying degrees of intensity in nature but every front shares these fundamental physical and dynamical characteristics. Thus, the lack of a threshold numerical designation is not an oversight but instead an attempt to distinguish features in the mid-latitude atmosphere that ought to be called fronts from those that should not. Of course, the intensity of a

front *is* a meaningful distinction to make both for scientific and operational interests. One way of measuring the strength of one front against another is by considering the *magnitudes* of their respective horizontal temperature gradients and whether or not the environmental flow is serving to alter the gradient magnitude via differential temperature advection.

Frontogenesis and Vertical Motions

The thermal wind relationship dictates that fronts, by virtue of their large thermal contrasts, be associated with strong vertical shear of the geostrophic wind. Shown in Fig. 4 is an idealized vertical cross-section of a frontal zone. Notice that the temperature gradient is largest near the surface and that the frontal zone is characterized by the strongest vertical shear. As a consequence of the latter point, the leading edge of the zone (i.e. the front) is a maximum in geostrophic relative vorticity as the prior analysis of the zero order front suggested it should be. In the absence of vertical advection and tilting, the frictionless vorticity equation suggests that vorticity can change only as a result of divergence ($\frac{d\eta}{dt} = -f(\nabla \cdot \vec{V})$). The presence of divergence is accompanied by vertical motions ($\nabla \cdot \vec{V} = -\frac{\partial \omega}{\partial p}$) according to the continuity equation. These two relationships form the basis of an important logical argument. If, by some horizontal advective process, for instance, the magnitude of ∇T increases, then the wind shear and jet core wind speed necessarily increase as well. A more intense jet results in increased vorticity. Increased vorticity implies that some divergence is operating in the fluid. If divergence is operating, there must be some vertical motion as well. Therefore, generally, *any change in the magnitude of* ∇T *requires the production of a vertical circulation in an atmosphere in approximate thermal wind balance*. Various physical/mathematical formulations that seek to

quantify this important physical relationship will be considered next after first examining how an increase in the magnitude of ∇T can be accomplished.

Any process that acts to increase the magnitude of ∇T is referred to as "*frontogenetic*". Such a process in action is known as *frontogenesis*. More precisely, any horizontal advective process that acts to increase the magnitude of ∇T will be referred to as *horizontal frontogenesis*. Some schematic illustrations of horizontal frontogenetical environments are given in Fig. 5. This verbal definition of frontogenesis has a corresponding mathematical one (known as the *frontogenesis function*),

$$\mathcal{F} = \frac{d|\nabla_p \theta|}{dt}, \tag{5}$$

that represents the Lagrangian rate of change of the magnitude of $\nabla_p \theta$ (the potential temperature gradient measured on an isobaric surface). Though it looks deceptively compact, (5) is a rather bulky expression. However, consideration of a simpler, one-dimensional version of (5) can, without loss of physical insight, provide a fundamental understanding of the nature of frontogenesis. Such an expression is given by

$$\mathcal{F}_x = \frac{d}{dt}\left(\frac{\partial \theta}{\partial x}\right) = \frac{\partial}{\partial x}\left(\frac{d\theta}{dt}\right) - \frac{\partial u}{\partial x}\frac{\partial \theta}{\partial x} - \frac{\partial v}{\partial x}\frac{\partial \theta}{\partial y} - \frac{\partial \omega}{\partial x}\frac{\partial \theta}{\partial p} \tag{6}$$

Thus, four physical processes, represented by the four terms on the RHS of (6), can contribute to an increase in $\frac{\partial \theta}{\partial x}$. The first of these processes is the effect of across-front gradients in diabatic heating, represented by $\frac{\partial}{\partial x}\left(\frac{d\theta}{dt}\right)$. Consider the meridionally oriented isentropes illustrated in Fig. 6. If there is latent heat release in ascending air on the warm side of this potential temperature gradient, then $\frac{\partial}{\partial x}\left(\frac{d\theta}{dt}\right) > 0$. Consequently, such a distribution of latent heat release is frontogenetical (i.e. renders $\mathcal{F}_x > 0$). The effect of differential cloud cover on frontal strength can be assessed utilizing the same expression. With the warm side of Fig. 6 cloudy and the cold

side clear, then differential insolation during the day renders $\frac{\partial}{\partial x}\left(\frac{d\theta}{dt}\right) < 0$ and, accordingly, daytime heating is frontolytic under such circumstances. Under the same distribution of clouds overnight, however, the cold side would cool more rapidly than the warm side so that $\frac{\partial}{\partial x}\left(\frac{d\theta}{dt}\right) > 0$ revealing that such nocturnal cloud cover promotes frontogenesis.

The effect of confluent flow superimposed upon a temperature gradient (Fig. 7a) is represented by the second term on the RHS of (6), $-\frac{\partial u}{\partial x}\frac{\partial \theta}{\partial x}$. Note, in Fig. 7a, that $\frac{\partial \theta}{\partial x} > 0$ while the winds render $\frac{\partial u}{\partial x} < 0$. The overall effect of the scenario depicted in Fig. 7a is to promote frontogenesis. This is mathematically consistent with the physical implication of the winds acting to push the isentropes closer together in the horizontal, thereby increasing $\left|\frac{\partial \theta}{\partial x}\right|$.

The effect of horizontal shear on $\frac{\partial \theta}{\partial x}$ is represented by the third term on the RHS of (6), $-\frac{\partial v}{\partial x}\frac{\partial \theta}{\partial y}$, and is illustrated in Fig. 7b, where isentropes are aligned at a slight angle to both the *x*- and *y*-axes such that $\frac{\partial \theta}{\partial y} < 0$. The schematic winds render $\frac{\partial v}{\partial x} > 0$ meaning that the entire shearing term is positive. Note that the increase in $\frac{\partial \theta}{\partial x}$ produced by the shearing is accomplished by rotating the isotherms into a more meridional orientation. *The resulting increase in* $\frac{\partial \theta}{\partial x}$ *does not, however, represent a decrease in the absolute distance between successive isentropes* (as was the case for both of the prior physical mechanisms).

Finally, the process of vertical tilting is represented by the fourth term on the RHS of (6), $-\frac{\partial \omega}{\partial x}\frac{\partial \theta}{\partial p}$. Figure 7c is a schematic vertical cross-section in which a frontal bundle of isentropes is coincident with a thermally direct vertical circulation. In a statically stable atmosphere, $\frac{\partial \theta}{\partial p}$ must be negative. Since $\omega = \frac{dp}{dt}$, upward vertical motion is consistent with negative omega and vice

versa. Therefore, $\frac{\partial \omega}{\partial x} < 0$ for the scenario depicted in Fig. 7c. Thus, the entire vertical tilting term is negative consistent with the implication that, by rotating the isentropes into a more nearly horizontal orientation, a thermally direct vertical circulation acts to decrease $\frac{\partial \theta}{\partial x}$ (e.g. acts *frontolytically*). Additional support for this interpretation arises through consideration of temperature rather than potential temperature. From that perspective, the rising warm air cools by expansion while the sinking cold air warms by compression; making the originally warm side colder while making the originally cold side warmer - clearly frontolytic.

Precisely similar physical interpretations can be made of the various terms in the more complicated, 3 dimensional frontogenesis function given by

$$\mathcal{F}_{3D} = \frac{d}{dt}|\nabla\theta| = \frac{d}{dt}\left[\left(\frac{\partial\theta}{\partial x}\right)^2 + \left(\frac{\partial\theta}{\partial y}\right)^2\right]^{0.5}$$

$$= \frac{1}{|\nabla\theta|}\left[\left(-\frac{\partial\theta}{\partial x}\right)\left(\frac{\partial u}{\partial x}\frac{\partial\theta}{\partial x} + \frac{\partial v}{\partial x}\frac{\partial\theta}{\partial y}\right) - \left(\frac{\partial\theta}{\partial y}\right)\left(\frac{\partial u}{\partial y}\frac{\partial\theta}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial\theta}{\partial y}\right)\right.$$

$$\left. -\left(\frac{\partial\theta}{\partial p}\right)\left(\frac{\partial\omega}{\partial x}\frac{\partial\theta}{\partial x} + \frac{\partial\omega}{\partial y}\frac{\partial\theta}{\partial y}\right)\right] \qquad (7)$$

In this expression, all terms with $\frac{\partial u}{\partial x}$ or $\frac{\partial v}{\partial y}$ are confluence terms, all terms with $\frac{\partial v}{\partial x}$ or $\frac{\partial u}{\partial y}$ are shearing terms, and all terms with derivatives of ω are tilting terms. For many instances of frontal development it is sufficient to consider the 2-D version of (7) in which the tilting terms are neglected;

$$\mathcal{F}_{2D} = \frac{1}{|\nabla\theta|}\left[\left(-\frac{\partial\theta}{\partial x}\right)\left(\frac{\partial u}{\partial x}\frac{\partial\theta}{\partial x} + \frac{\partial v}{\partial x}\frac{\partial\theta}{\partial y}\right) - \left(\frac{\partial\theta}{\partial y}\right)\left(\frac{\partial u}{\partial y}\frac{\partial\theta}{\partial x} + \frac{\partial v}{\partial y}\frac{\partial\theta}{\partial y}\right)\right]. \qquad (8)$$

As discussed previously, there is a physical relationship between changes in $|\nabla\theta|$ and the production of vertical circulations in the middle latitudes. Consideration of an alternative version of (8) in which all winds are geostrophic, provides some mathematical rigor to that argument:

$$\mathcal{F}_{2D_g} = \frac{1}{|\nabla\theta|}\left(\left(\frac{\partial\theta}{\partial x}\right)\left[-\frac{\partial u_g}{\partial x}\frac{\partial\theta}{\partial x} - \frac{\partial v_g}{\partial x}\frac{\partial\theta}{\partial y}\right] + \left(\frac{\partial\theta}{\partial y}\right)\left[-\frac{\partial u_g}{\partial y}\frac{\partial\theta}{\partial x} - \frac{\partial v_g}{\partial y}\frac{\partial\theta}{\partial y}\right]\right). \qquad (9)$$

The terms inside square brackets on the RHS of (9) are equal to $\frac{1}{f\gamma}Q_1$ and $\frac{1}{f\gamma}Q_2$, the components of the Hoskins et al. (1978) $\vec{Q}$-vector, respectively. Thus, (9) can be expressed as

$$\mathcal{F}_{2D_g} = (\frac{1}{f\gamma})\frac{\vec{Q}\cdot\nabla\theta}{|\nabla\theta|} \tag{10}$$

which is also a scalar multiple of the magnitude of the across-isentrope component of $\vec{Q}$. Shown in Fig. 8 is a set of $\vec{Q}$-vectors and isentropes at 700 hPa. From (10), any place where $\vec{Q}$-vectors point across the isentropes from cold to warm air will be associated with geostrophic horizontal frontogenesis (i.e. $\mathcal{F}_{2D_g} > 0$). In such locations, the geostrophic winds are advecting θ in such a way as to increase $|\nabla\theta|$ and a thermally direct vertical circulation is the response. Figure 8 illustrates that where the geostrophic deformation is acting frontogenetically, the $\vec{Q}$-vectors will be convergent somewhere, and to some degree, on the warm side of the baroclinic zone. This implies warm air rising and cold air, in which the $\vec{Q}$-vectors are divergent, sinking thereby manifesting the expected thermally direct vertical circulation.

Importantly, the geostrophic frontogenesis function only references the influence of *geostrophic advection* on forcing the secondary circulation. Recalling the large Rossby number that characterizes the across-front flow, it is reasonable to wonder whether this is enough to accurately describe nature. It is entirely possible that across-front ageostrophic advections of temperature and geostrophic momentum might accomplish a considerable amount of frontal intensification. In fact, the geometric form of the frontogenesis equation[2] demonstrates that some fraction of the total horizontal frontogenetical forcing is provided by the divergence (*D*) of the ageostrophic wind.

[2] The geometric form of the frontogenesis equation is given by

$$\mathcal{F}_{2D} = \frac{|\nabla\theta|}{2}(Fcos2\beta - D)$$

Considering the influence of *just* geostrophic confluence on the evolution of a given temperature contrast would lead to the simple formulation,

$$\frac{d}{dt}\left(\frac{\partial\theta}{\partial x}\right) = -\frac{\partial u_g}{\partial x}\frac{\partial\theta}{\partial x} = k\frac{\partial\theta}{\partial x} \qquad (11)$$

where k is a constant, commonly observed value of geostrophic confluence ($k = -\frac{\partial u_g}{\partial x} = 10^{-5}s^{-1}$). This expression can be solved explicitly to yield,

$$(\frac{\partial\theta}{\partial x})_t = (\frac{\partial\theta}{\partial x})_o e^{kt} \qquad (12)$$

suggesting that, for typical conditions at mid-latitudes, it takes 10^5 seconds (~1 day) for pure geostrophic confluence to increase the intensity of a frontal temperature contrast by a factor of e. Such an intensification rate is ***much*** slower than is actually observed in nature in many instances. Why should nature be able to accomplish frontogenesis so much faster than a geostrophic confluence model? The underlying reason is that the confluence model adopts a view of frontal intensification in which the secondary circulation forced by the geostrophic frontogenesis does not feedback upon the across-front advection of temperature (or momentum). Consequently, this is not a truly *dynamical* approach to frontogenesis as the advection of temperature that is accomplished by the across-front ageostrophic wind that develops in response to the geostrophic frontogenesis has been neglected. This across-front ageostrophic advection is able to further intensify the temperature contrast, further increase the vertical shear, and further enhance the thermally direct circulation associated with positive horizontal frontogenesis. Thus, in order to more accurately describe nature, these across-front, ageostrophic advections of temperature and momentum must be included in any diagnostic equation for frontogenesis. The so-called *semi-geostrophic equations*, developed independently by Sawyer

---

where F is the total deformation of the flow and D is the divergence. β is the angle between the axis of dilatation of the total deformation and the isentropes. A derivation of this form can be found in Martin (2006), p. 198-200.

(1956) and Eliassen (1962) include these important missing processes and consequently provide an analytic tool by which to develop a more comprehensive and physically accurate picture of frontogenesis.

The Semi-Geostrophic Equations

Sawyer (1956) noted that, though bands of considerable baroclinicity can be produced by differential horizontal advection in non-divergent (i.e., geostrophic) deformation fields, such non-divergent flows cannot account for the production of the characteristic frontal vorticity which can only be produced by divergent (i.e. ageostrophic) motions. Consider a 2-D front with the *x*-axis along the front (i.e. along the isentropes) and the *y*-axis points directly into the cold air as depicted in Fig. 9. Given the isobaric expressions of the geostrophic winds ($u_g = -\frac{1}{f}\frac{\partial \phi}{\partial y}$ and $v_g = \frac{1}{f}\frac{\partial \phi}{\partial x}$) and the fact that the hydrostatic equation can be written as $\frac{1}{f}\frac{\partial \phi}{\partial p} = -\gamma\theta$ where $\gamma = \frac{R}{f p_o}(\frac{p_o}{p})^{c_v/c_p}$, with p_o=1000 hPa, the thermal wind components can be written as

$$\frac{\partial U_g}{\partial p} = \gamma \frac{\partial \theta}{\partial y} \text{ and } \frac{\partial V_g}{\partial p} = -\gamma \frac{\partial \theta}{\partial x}. \qquad (13)$$

Since the hypothetical front is assumed to be 2-D there is no along-front geopotential height gradient (i.e. $\frac{\partial \phi}{\partial x} = 0$)so the equation of motion in the along-front (x) direction is given by

$$\frac{dU_g}{dt} + \frac{du}{dt} = fv, \qquad (14)$$

where u and v (U_g and V_g) are the x- and y-direction agesotrophic (geostrophic) winds, respectively. If the along-front flow is assumed to be nearly geostrophic (i.e. u is small with respect to U_g) then the *geostrophic momentum approximation,* in which it is assumed that there is no systematic increase in

the magnitude of the along-front ageostrophic wind (i.e. $\left|\frac{dU_g}{dt}\right| \gg \left|\frac{du}{dt}\right|$) can be made. In this case, (14) is reduced to

$$\frac{dU_g}{dt} = fv, \tag{15}$$

or

$$\frac{dU_g}{dt} = \frac{\partial U_g}{\partial t} + U_g\frac{\partial U_g}{\partial x} + u\frac{\partial U_g}{\partial x} + V_g\frac{\partial U_g}{\partial y} + v\frac{\partial U_g}{\partial y} + \omega\frac{\partial U_g}{\partial p} = fv. \tag{16}$$

Similarly, the thermodynamic energy equation becomes,

$$\frac{d\theta}{dt} = \frac{\partial \theta}{\partial t} + U_g\frac{\partial \theta}{\partial x} + u\frac{\partial \theta}{\partial x} + V_g\frac{\partial \theta}{\partial y} + v\frac{\partial \theta}{\partial y} + \omega\frac{\partial \theta}{\partial p}. \tag{17}$$

In accord with the scaling discussed earlier, the along-front flow is presumed to be largely in geostrophic balance. Consequently, the along-front ageostrophic advection terms (i.e. $u\frac{\partial}{\partial x}$ terms) are also neglected. Additionally, a new conservative variable, the absolute geostrophic momentum (*M*) defined as

$$M = U_g - fy \tag{18}$$

is introduced noting that M is conserved, according to (15). Using (18), (16) can be rewritten as

$$\frac{\partial U_g}{\partial t} + U_g\frac{\partial U_g}{\partial x} + V_g\frac{\partial U_g}{\partial y} + v\frac{\partial M}{\partial y} + \omega\frac{\partial M}{\partial p} = 0. \tag{19}$$

Now, taking $\frac{\partial}{\partial p}$(19) and adding it to -$\gamma\frac{\partial}{\partial y}$(17), using the thermal wind relationships and the non-divergence of the geostrophic wind yields

$$-\frac{\partial}{\partial y}\left(\gamma v\frac{\partial \theta}{\partial y} + \gamma\omega\frac{\partial \theta}{\partial p}\right) + \frac{\partial}{\partial p}\left(v\frac{\partial M}{\partial y} + \omega\frac{\partial M}{\partial p}\right) =$$

$$-2\left(\frac{\partial U_g}{\partial p}\frac{\partial U_g}{\partial x} + \frac{\partial V_g}{\partial p}\frac{\partial U_g}{\partial y}\right) - \gamma\frac{\partial}{\partial y}\left(\frac{d\theta}{dt}\right). \tag{20}$$

The continuity equation in isobaric coordinates ($\nabla \cdot \vec{V} = 0$) can be approximated as $\frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p} \approx 0$ by assuming that the along-front derivative of the along-front ageostrophic flow ($\frac{\partial u}{\partial x}$) is negligible. Then,

setting $v = -\frac{\partial \psi}{\partial p}$ and $\omega = \frac{\partial \psi}{\partial y}$, (20) can be rewritten in terms of a streamfunction, ψ, for the ageostrophic flow in the *y*-*p* plane as

$$\left(-\gamma \frac{\partial \theta}{\partial p}\right)\frac{\partial^2 \psi}{\partial y^2} + \left(2\frac{\partial M}{\partial p}\right)\frac{\partial^2 \psi}{\partial p \partial y} + \left(-\frac{\partial M}{\partial y}\right)\frac{\partial^2 \psi}{\partial p^2} = Q_g - \gamma \frac{\partial}{\partial y}\left(\frac{d\theta}{dt}\right) \qquad (21)$$

where

$$Q_g = -2\left(\frac{\partial U_g}{\partial y}\frac{\partial V_g}{\partial p} - \frac{\partial V_g}{\partial y}\frac{\partial U_g}{\partial p}\right) \qquad (22)$$

is the geostrophic forcing function. Equation (21), known as the *Sawyer-Eliassen Circulation equation*, is a linear, second order partial differential equation for the 2-D, transverse (i.e. in the across-front plane) ageostrophic streamfunction, ψ. Solutions for ψ will arise entirely as a consequence of the frontogenetic forcing provided that the discriminate, $B^2 - 4AC < 0$ (i.e. $4\left(\frac{\partial M}{\partial p}\right)^2 - 4\gamma \frac{\partial \theta}{\partial p}\frac{\partial M}{\partial y}$), which translates to $\gamma\left(\frac{\partial \theta}{\partial p}\frac{\partial M}{\partial y} - \frac{\partial \theta}{\partial y}\frac{\partial M}{\partial p}\right) > 0$. Physically, this represents the condition that the quasi-geostrophic potential vorticity is greater than zero in the solution domain. If this condition is not met, then there is either 1) inertial instability, or 2) static instability somewhere in the domain. The presence of either instability in the solution domain will allow growth of non-unique solutions, arising from the release of the instability, thus prohibiting clear attribution of the resulting ageostrophic motions to the process of frontogenesis.

The Sawyer-Eliassen equation is a complicated looking expression and yet it avails itself of rather simple conceptual interpretations. Since it is a second order partial differential equation, the RHS forcing in a given portion of the domain yields an opposite signed ψ in the same portion of the domain. Consequently, based on the definitions of *v* and *w*, this means that $Q_g > 0$ ($Q_g < 0$) is associated with a thermally direct (indirect) vertical circulation.

Finally, considerable physical insight into the nature of fronts and the process of frontogenesis can be garnered by considering each term on the LHS of (21) in some detail. The first such term is

$\left(-\gamma\frac{\partial\theta}{\partial p}\right)\frac{\partial^2\psi}{\partial y^2}$. This term represents the product of the static stability ($-\gamma\frac{\partial\theta}{\partial p}$) and across-front gradients in ω ($\frac{\partial^2\psi}{\partial y^2}$, since $\omega = -\frac{\partial\psi}{\partial y}$). The only way that across-front gradients in vertical motion can have any effect on $|\nabla\theta|_h$ is if they act upon the static stability via the tilting term. The second term on the LHS of (21) represents the product of the across-front baroclinicity ($2\frac{\partial M}{\partial p}$) and the across-front, ageostrophic divergence ($\frac{\partial^2\psi}{\partial p\partial y}$, or $-\frac{\partial v}{\partial y}$ since $v = -\frac{\partial\psi}{\partial p}$). Clearly, if there is ageostrophic convergence in the presence of baroclinicity, the frontal intensity is increased. Finally, the third term on the LHS of (21) represents the product of the vorticity ($-\frac{\partial M}{\partial y}$) and the across-front vertical shear of the ageostrophic wind ($\frac{\partial^2\psi}{\partial p^2}$, or $-\frac{\partial v}{\partial p}$ since $v = -\frac{\partial\psi}{\partial p}$). The tilting of vortex tubes by the across-front vertical shear of the ageostrophic wind will modulate the tilt of the frontal zone, an observable characteristic of fronts in nature. Each of the coefficients of these three terms represents one of the three essential dynamical characteristics of a front and each of the three terms themselves represents an aspect of the secondary ageostrophic circulation that responds to the frontogenetical forcing. This suggests that in solving (21) for ψ, a successive overrelaxation procedure (SOR) would execute the following solution steps; 1) assess the RHS forcing in (21), 2) make a first guess for ψ in the solution domain, 3) use ψ to compute the first guess ageostrophic circulation, 4) allow the ageostrophic circulation to advect temperature and momentum (as in the terms just discussed on the LHS of (21), 5) iterate to a state of balance between the RHS and LHS. In this way, the ageostrophic secondary circulation feeds back into the final frontogenesis process as originally intended with the introduction of the increased complexity of the Sawyer-Eliassen equation. Thus, solution of the Sawyer-Eliassen equation mimics nature and suggests that frontogenesis is a two-step process. First, the non-divergent, geostrophic deformation tightens the temperature gradient resulting in the production of a secondary, ageostrophic transverse

circulation. Second, the ageostrophic circulation itself advects temperature and momentum in the frontal zone, produces the characteristic vorticity, and further intensifies the temperature contrast leading to the sometimes rapid development of the sharp frontal boundaries observed in the mid-latitude atmosphere.

Though the sharpest such boundaries tend to develop at the Earth's surface, fronts are not constrained to form only at physical boundaries. An extremely important class of fronts forms routinely at the tropopause, a thermodynamic boundary across which there is very little mixing.

Upper-Level Frontogenesis

A vertical cross-section through a modest local wind speed maxima (labeled with a *J*) in the upper troposphere and lower stratosphere is shown in Fig. 10a. The tropopause is marked by the sudden increase in the vertical gradient in potential temperature corresponding to increased static stability. As a consequence of the thermal wind relationship, the wind speed maxima sits atop a column of air in which a modest horizontal temperature contrast exists, particularly in the upper troposphere. The associated vertical wind shear produces *x*-direction vortex tubes with the indicated spin beneath the maxima. Given these environmental ingredients, consider the consequences of imposing a thermally indirect circulation straddling the wind speed maxima (Fig. 10b).

First, since the lower stratosphere is characterized by high static stability, a thermally indirect circulation would tilt a bundle of closely packed isentropes from their originally horizontal orientation into a more vertical orientation. Such tilting, recall, does not change the absolute distance between isentropes in the bundle. As the isentropes acquire a more vertical orientation, the magnitude of the horizontal θ gradient ($|\nabla\theta|_h$) increases in the upper troposphere. The thermally indirect circulation

would also act upon the *x*-direction vortex tube, gradually tilting a component of its vorticity into the vertical direction. Thus, a local increase in cyclonic vorticity occurs precisely where a local increase in $|\nabla\theta|_h$ is accomplished. These ingredients, along with the high static stability present in the developing baroclinic zone owing to its lower stratospheric origin, constitute the three essential dynamical characteristics of a front. Thus, imposition of a thermally indirect circulation onto the environment illustrated in Fig. 10a would lead to the development of a frontal zone along the tropopause boundary. Such frontal zones are known as *upper-level fronts*. Owing to their characteristic association with intensifying wind speed maxima, they are also often referred to as *upper-level jet/front systems*.

In contrast to surface based frontal zones which separate air masses with different origins in the horizontal, upper-level fronts separate tropospheric air (beneath them) from stratospheric air (above them). In fact, the tropopause boundary often becomes "folded" over itself in the vicinity of the upper-level front as a result of the differential vertical advection of potential vorticity (PV) at the tropopause boundary implied in Fig. 10c. A number of observable quantities in the atmosphere serve as "tracers" for distinguishing between tropospheric and stratospheric air. In the late 1940's and 1950's it was considered safe to test new weapons systems by exploding nuclear devices at high altitude in the stratosphere. Underlying this folly was the notion that the dangerous radioactive by-products of the devices would rain out in the highly stratified stratosphere and so not quickly or easily mix into the troposphere. Some of the pioneering work done in identifying and diagnosing upper-level fronts (i.e. Reed and Danielsen, 1959; Danielsen, 1964) used analysis of radioactivity, an indisputable criteria for establishing the presence of stratospheric air in upper-level fronts. Upper-level fronts are also characterized by high ozone mixing ratios which, again, testify to the stratospheric origin of the air in upper-level frontal zones. Finally, given the dramatically larger static stability in the lower

stratosphere, it is also a region of high PV as compared to the upper-troposphere. Thus, the fact that upper-level frontal zones are characterized by high PV is yet another clear indication that upper-level fronts separate stratospheric from tropospheric air. Since PV is a well conserved variable, when high PV from the lower stratosphere is subducted into the less stably stratified middle and upper troposphere during upper frontogenesis, the large PV is manifest in large *vorticity*. In fact, a sizeable fraction of the middle- and upper-tropospheric vorticity maxima that serve as dynamical precursors to lower tropospheric cyclogenesis are produced as by-products of the upper frontogenesis process. Some of these analysis elements are illustrated in Fig. 11, a vertical cross-section through an upper-level front. It is clear from Fig. 10 that a thermally indirect vertical circulation in the vicinity of an upper tropospheric wind speed maxima is of vital importance to the development of such upper-level frontal zones. Naturally, then, determining the synoptic-scale conditions under which such a thermally indirect vertical circulation might be produced is of great interest.

Numerous studies have shown that the subsidence portion of such a circulation is particularly effective in producing an upper front if it occurs beneath the jet core (e.g. Shapiro, 1981; Shapiro, 1983; Keyser and Pecnick, 1985; Shapiro and Keyser, 1986; Reeder and Keyser, 1988; Keyser et al., 1992, Keyser, 1999). In fact, much of this work focused on the forcing function, Q_g, of the Sawyer-Eliassen equation and suggested that geostrophic cold air advection in cyclonic shear tended to shift the thermally direct circulation at the jet entrance region to the warm side of the attendant baroclinic zone thereby placing the descent near the jet core.

The emphasis on vertical motions forced by horizontal frontogenesis may have been overly myopic as an earlier analysis by Mudrick (1974) had emphasized the importance of negative vorticity advection by the thermal wind in producing a similar distribution of subsidence. A recent analysis by Martin (2014) considered the combined effects of negative vorticity advection (NVA) by the thermal

wind along with horizontal frontogenesis in forcing the necessary subsidence. Specifically, he showed that the vorticity advection and frontogenetic forcings, which are dynamically distinct processes by virtue of the fact that their respective forcings arise from orthogonal components of the *Q*-vector, always yield coincident regions of subsidence beneath the jet core in environments characterized by geostrophic cold air advection along the jet axis (Fig 12). In fact, it appears that the predominant sinking driving upper frontogenesis in such synoptic environments is related to negative shear vorticity advection by the thermal wind, the phenomenological equivalent of geostrophic cold air advection in cyclonic shear. Thus, the most comprehensive way to conceptualize the forcings that promote rapid upper level jet-front development in such regions is to consider *both* of these processes simultaneously.

Concluding Statement

This chapter has considered fronts both from the perspective of their essential thermodynamic and dynamic characteristics and in terms of the limited set of physical processes that conspire to produce them and modify their structure and intensity in the middle latitudes of Earth. Other thermal boundaries, often occurring on smaller scales, can also be influential in shaping the weather. Such features as gust fronts and sea-breeze fronts are among this variety. Other chapters in this section address these features in some detail.

References

Bjerknes, J., and H. Solberg, 1922: Life cycle of cyclones and the polar front theory of atmospheric circulation. *Geofys. Publ*., **3**(1), 1-18.

Danielsen, E.F., 1964: Project Springfield report. *Defense Atomic Support Agency, DASA*, *1517*, p.97.

Eliassen, A., 1962: On the vertical circulation in frontal zones. *Geofys. Publ*., **24**. 147-160.

Kutzbach, G., 2016. *The thermal theory of cyclones: A history of meteorological thought in the nineteenth century*. Springer.

Hoskins, B. J., I. Draghici, and H. C. Davies, 1978: A new look at the ω-equation. *Quart. J. Roy. Meteor. Soc*., **104**, 31-38.

Keyser, D., 1999: On the representation and diagnosis of frontal circulations in two and three dimensions. *The Life Cycles of Extratropical Cyclones*, 239-264.

_____, and M. J. Pecnick, 1985: A two-dimensional primitive equation model of frontogenesis forced by confluence and horizontal shear. *J. Atmos. Sci., **42***(12), 1259-1282.

_____, and M. A. Shapiro, 1986: A review of the structure and dynamics of upper-level frontal zones. *Mon. Wea. Rev.*, **114**, 452-496.

_____, B. D. Schmidt, and D. G. Duffy, 1992: Quasigeostrophic vertical motions diagnosed from along- and cross-isentrope components of the *Q* vector. *Mon. Wea. Rev*., **120**, 731-741.

Martin, J. E., 2006: *Mid-Latitude Atmospheric Dynamics: A First Course*. Wiley, 324 pp

_____, 2014: Quasi-geostrophic diagnosis of the influence of vorticity advection on the development of upper level jet-front systems. *Quart. J. Roy. Meteor. Soc.*, **140**, 2658-2671.

Mudrick, S. E., 1974: A numerical study of frontogenesis., *J. Atmos. Sci.,* **39**, 869 – 892.

Reed, R. J. and E. F. Danielsen, 1959: Fronts in the vicinity of the tropopause. *Arch. Meteor. Geophys. Bioklim*., **A11**, 1-17.

Reeder, M. J. and D. Keyser, 1988: Balanced and unbalanced upper-level frontogenesis. *J. Atmos. Sci.,* **45**, 3366–3386.

Sawyer, J. S., 1956: The vertical circulation at meteorological fronts and its relation to frontogenesis. *Proc. R. Soc*., **A234**, 246-262.

Shapiro, M. A., 1981: Frontogenesis and geostrophically forced secondary circulations in the vicinity of jet stream – frontal zone systems. *J. Atmos. Sci.,* **38**: 954–973.

Shapiro, M. A., 1983: Mesoscale weather systems of the central United States. In *The National STORM Program: Scientific and Technological Bases and Major Objectives*, R. A. Anthes (ed.): 3.1–3.77. University Corporation for Atmospheric Research: Boulder, CO.

Stoelinga, M. T., J. D. Locatelli, and P. V. Hobbs, 2002: Warm occlusions, cold occlusions, and forward-tilting cold fronts. *Bull. Amer. Meteor. Soc.,* **83,** 709-721.

Further Reading

Margules, M., 1906: Uber temperaturschichtung in stationar bewegter und ruhender luft Hann-Band. *Meteor. Z*., 243-254.

Miller, J. E., 1948: On the concept of frontogenesis. *J. Meteor*., **5**, 169-171.

Petterssen, S., 1936: Contribution to the theory of frontogenesis. *Geofys. Publ*., **11**, 27 pp.

Reed, R. J., 1955: A study of a characteristic type of upper-level frontogenesis. *J. Meteor*., **12**, 226-237.

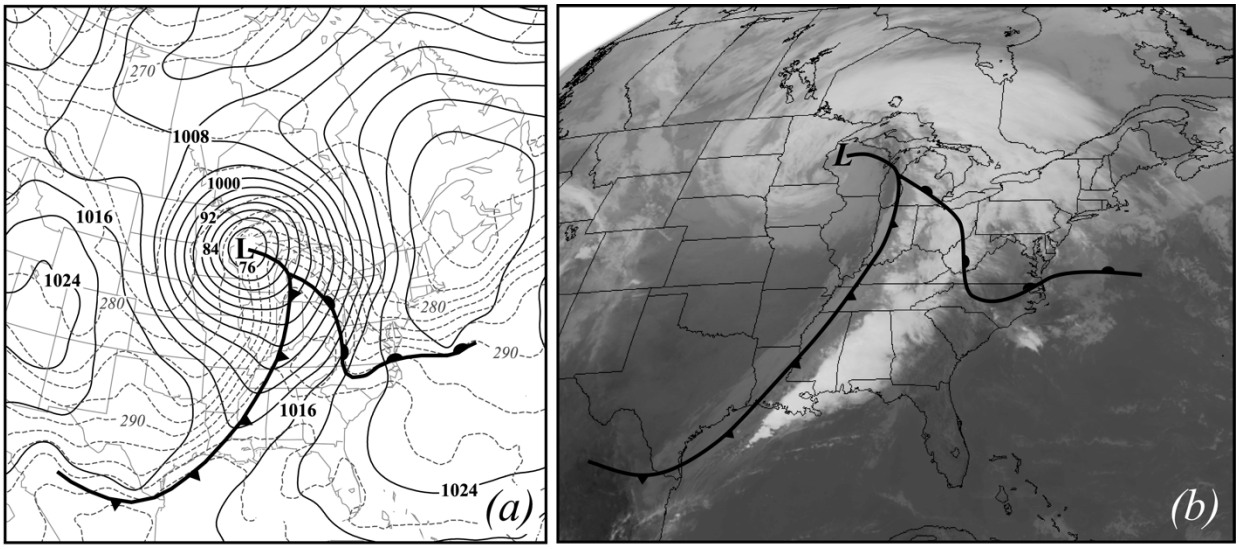

Fig. 1 (a) Sea-level pressure and 950 hPa potential temperature analysis at 1800 UTC 12 November 1998. Solid lines are sea-level isobars, labeled in hPa and contoured every 4 hPa. Dashed lines are 950 hPa isentropes labeled in K and contoured every 2.5 K. Standard frontal symbols identify the cold and warm fronts while the occluded front is indicated as a thick black line. (b) Infra-red satellite image of the same storm at 1815 UTC 12 November 1998. Frontal symbols as in Fig. 1a. (Adapted from Martin 2006).

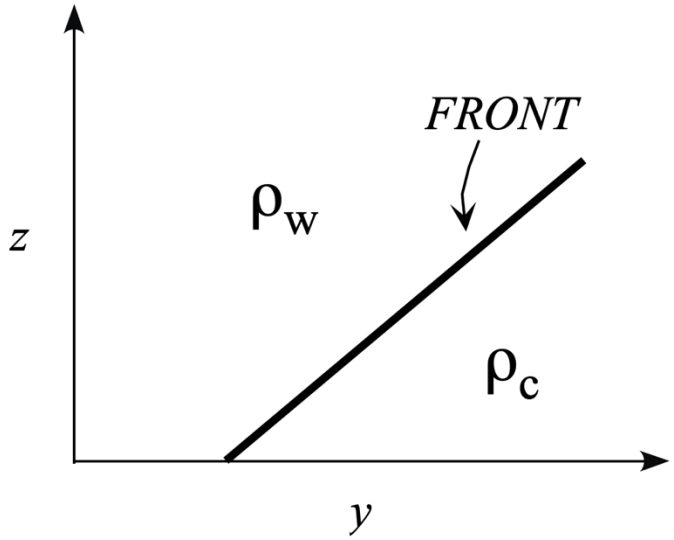

Fig. 2 Vertical cross-section through the zero order front.
(Adapted from Martin 2006)

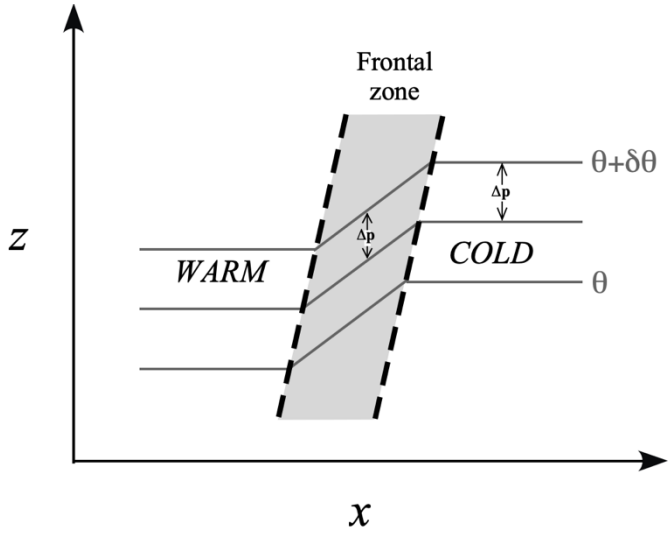

Fig. 3 Isentropes associated with a first order front.
Note that the static stability is largest in the frontal zone.
(Adapted from Martin 2006)

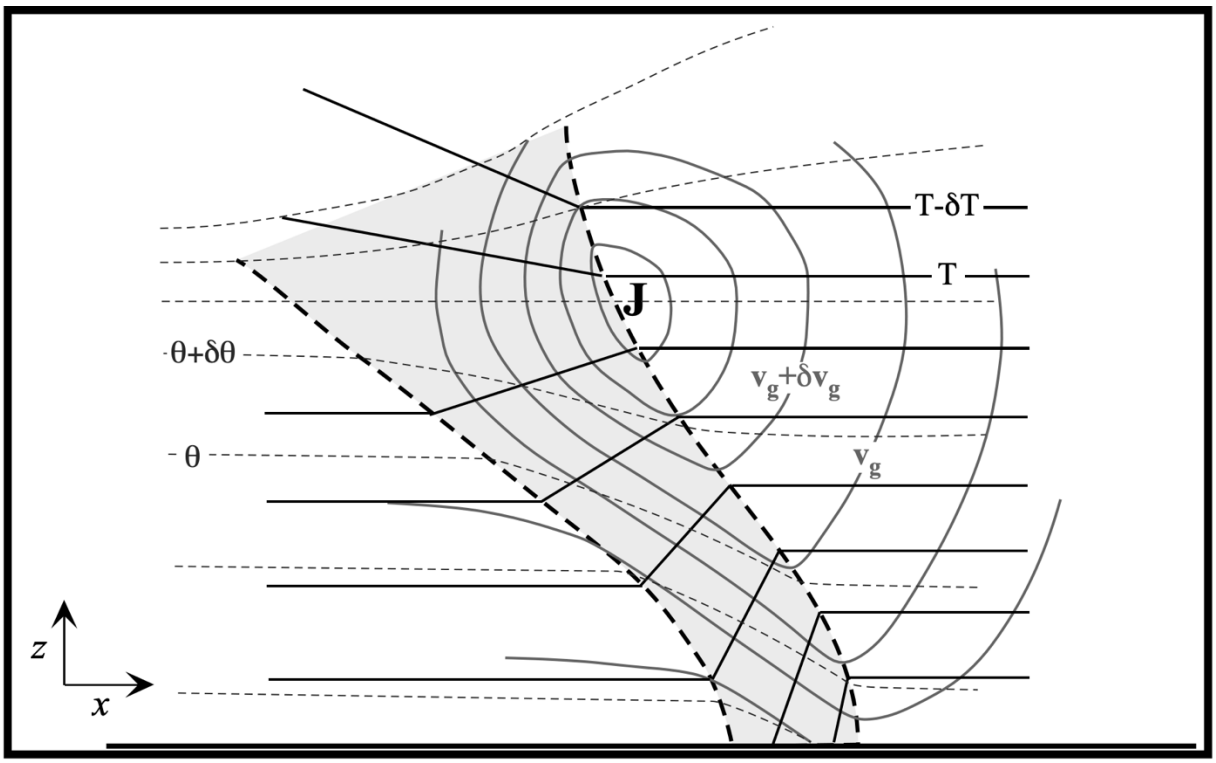

Fig. 4 Idealized vertical cross-section through a frontal zone. Gray solid lines are isotachs of the geostrophic wind into the page with "J" indicating the position of the wind maxima. Black solid lines are isotherms and thin dashed lines are isentropes. Gray shaded region with thick dashed border represents the idealized frontal zone. (Adapted from Martin 2006)

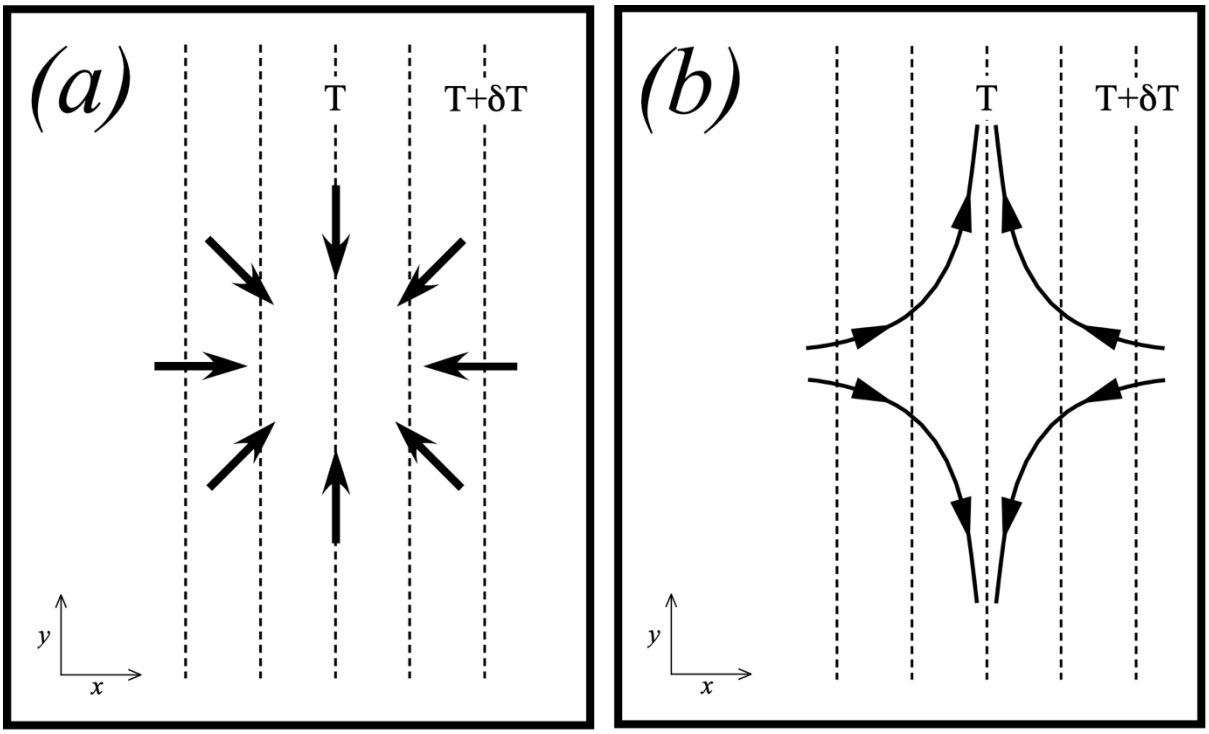

Fig. 5 (a) Pure convergence superimposed upon a field of isotherms. (b) Horizontal deformation superimposed upon a field of isotherms. In both cases the horizontal wind will tend to intensify $|\nabla_T|$. (Adapted from Martin 2006)

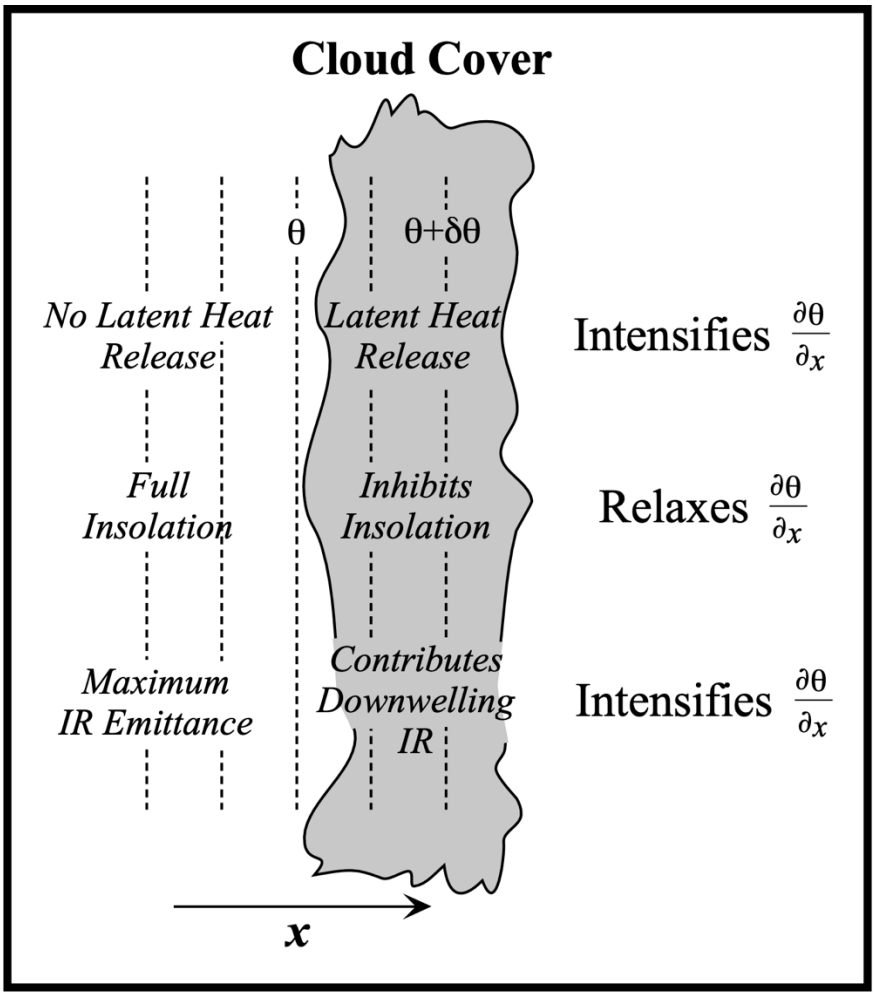

Fig. 6 The diabatic effects of cloud cover on $\frac{\partial\theta}{\partial x}$. The effect of differential latent heat release can occur at any time of day. Differential insolation and IR emittance are specific to day and night, respectively. (Adapted from Martin 2006)

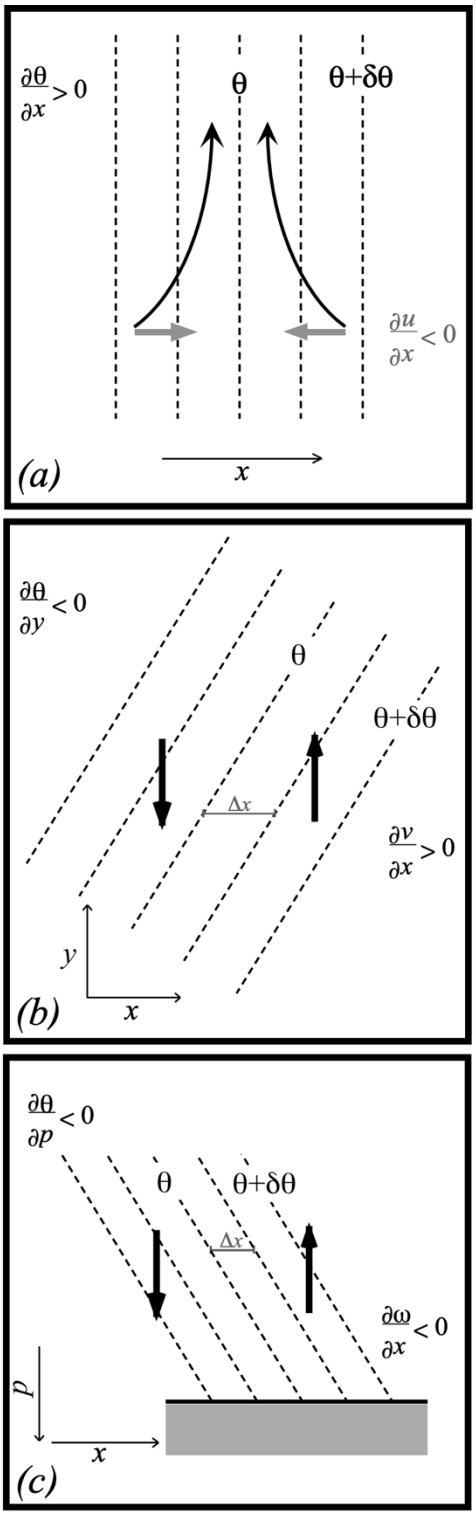

Fig. 7 (a) Confluent horizontal flow acting on meridionally oriented isentropes. Thegray arrows represent the *x*-direction wind. (b) Effect of horizontal shear on $\frac{\partial \theta}{\partial x}$. The black arrows represent the *y*-direction wind. (c) Effect of tilting on $\frac{\partial \theta}{\partial x}$. The black arrows represent upward and downward vertical motions.

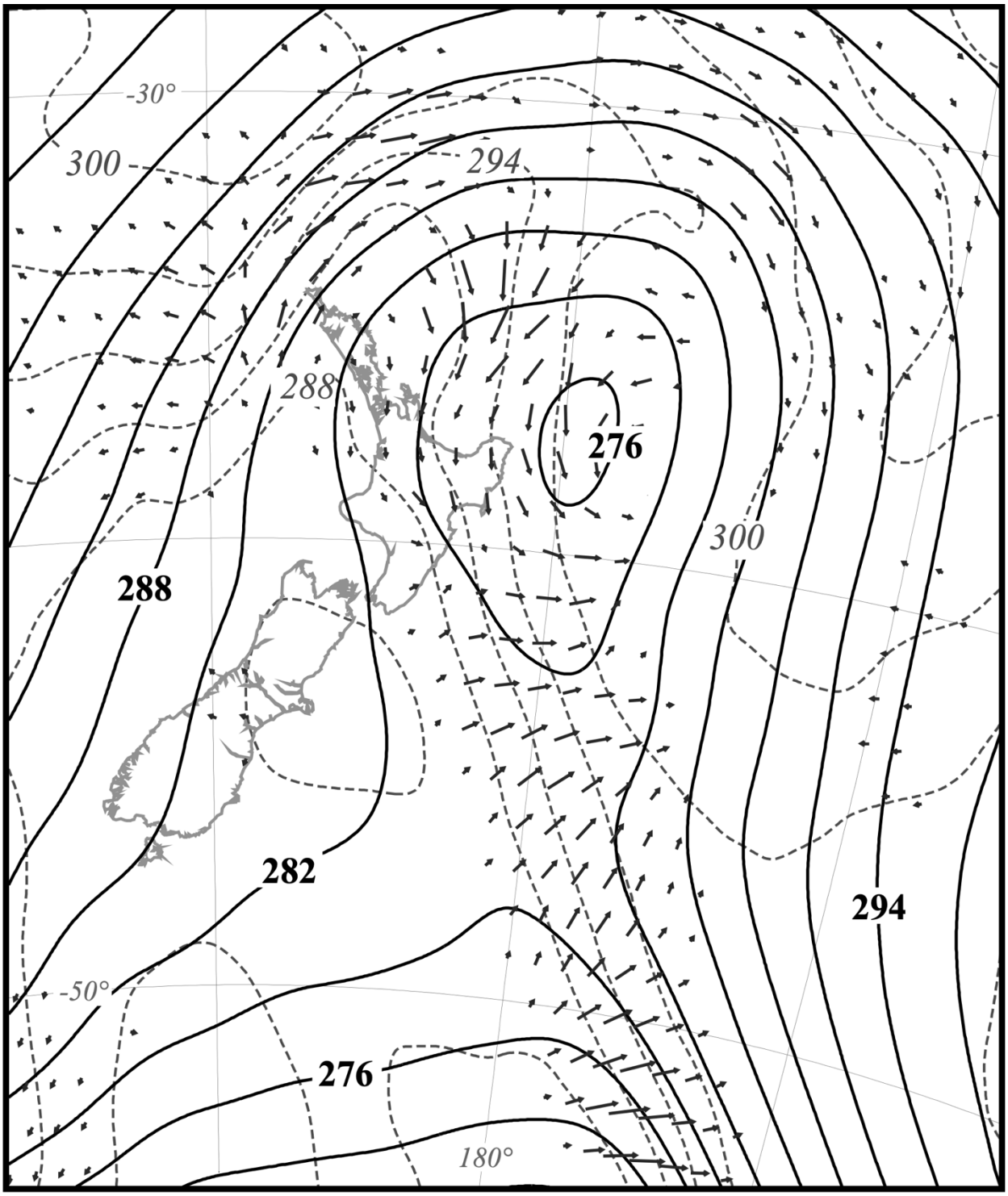

Fig. 8 700 hPa geopotential heights, isentropes and ***Q***-vectors near New Zealand at 0600 UTC 16 August 2004. Geopotential heights are labeled in dam and contoured every 3 dam. Isentropes are labeled in K and contoured every 3K. For clarity, only ***Q*** vectors larger than 2 x 10^{-10} m^2 kg^{-1} s^{-1} are plotted. (Adapted from Martin 2006)

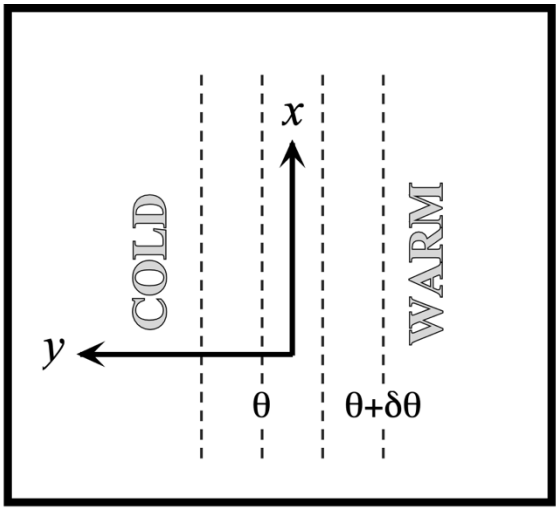

Fig. 9 Coordinate system orientation to isentropes adopted for development of the Sawyer-Eliassen equation. (Adapted from Martin 2006)

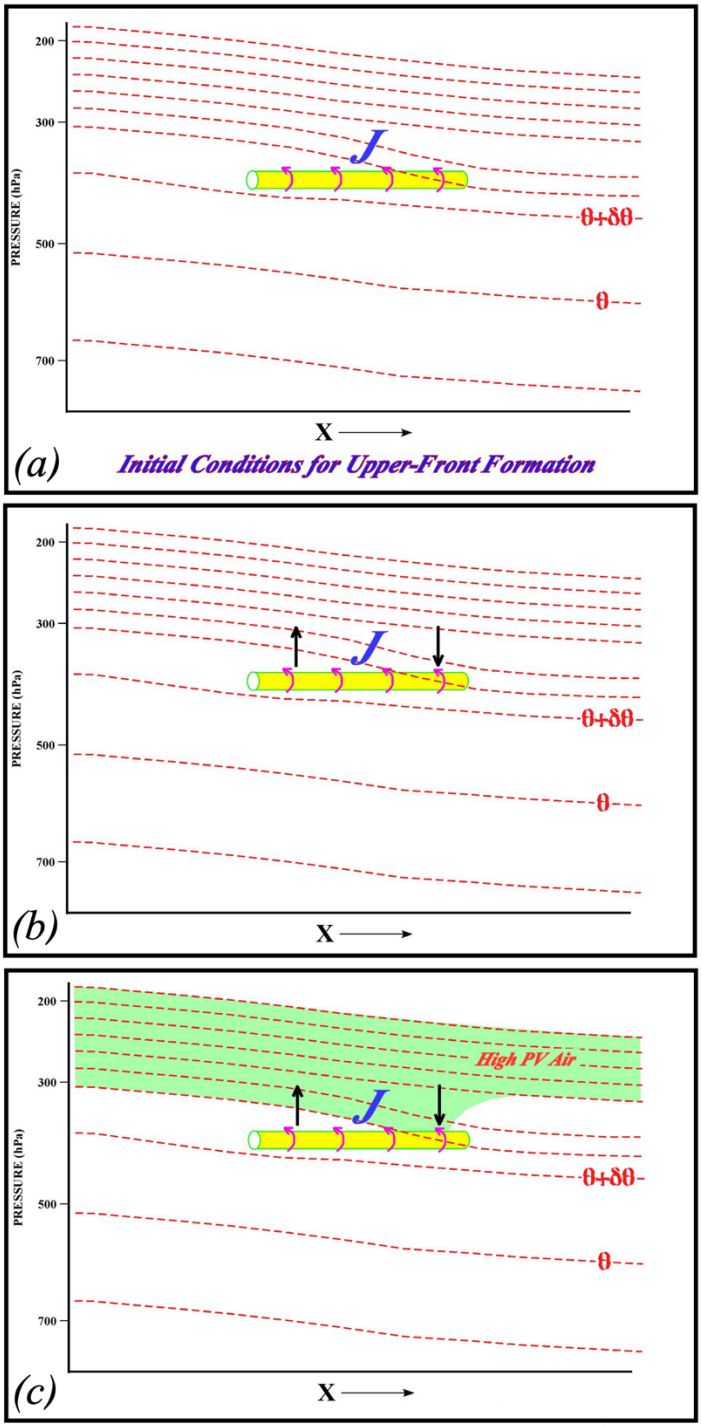

Fig. 10 (a) Schematic vertical cross-section through the upper troposphere/lower stratosphere in the vicinity of a local wind speed maximum near the tropopause. Red dashed lines are isentropes, "J" represents the tropopause-level wind speed maximum, yellow shaded tube represents a segment of an x-direction tube with a rotation imposede upon it by the vertical shear associated with the modest temperature gradient just below the tropopause. (b) Black arrows represent hypothetical thermally indirect circulation that, under certain conditions, might be imposed upon this environment. (c) Green shading represents high potential vorticity (PV) of the lower stratosphere that is differentially advectedin the vertical near the wind speed maximum, forcing the folding of the tropopause in its vicinity. See text for further explanation.

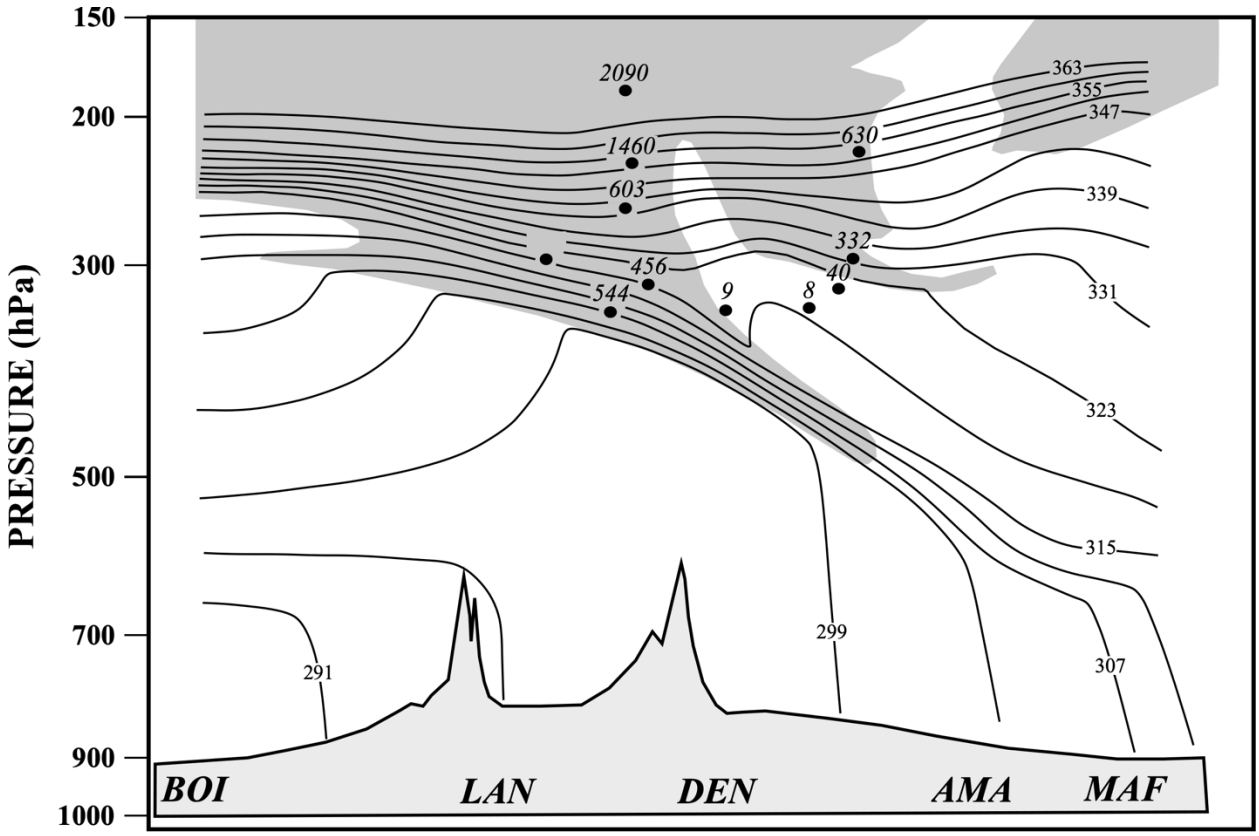

Fig. 11 Vertical cross-section taken at 0000 UTC 22 April 1963 from Boise, Idaho (BOI), to Lander, WY (LAN), to Denver, CO (DEN), to Amarillo, TX (AMA), to Midland, TX (MAF). Solid lines are isentropes labeled in K and contoured every 4 K. Shading highlights regions in which potential vorticity (PV) is greater than or equal to 2.5 PVU (1 PVU= 10^{-6} K m^2 kg^{-1} s^{-1}) safely indicating stratospheric air. Dots are measurements of radioactive decay of strontium (Sr_{90}) in units of disintegrations per minute per 1000 cubic feet of air. Note the high radioactivity that exists within the high PV stratospheric air as well as within the upper frontal zone. Adapted from Danielsen (1964). (Adapted from Martin 2006)

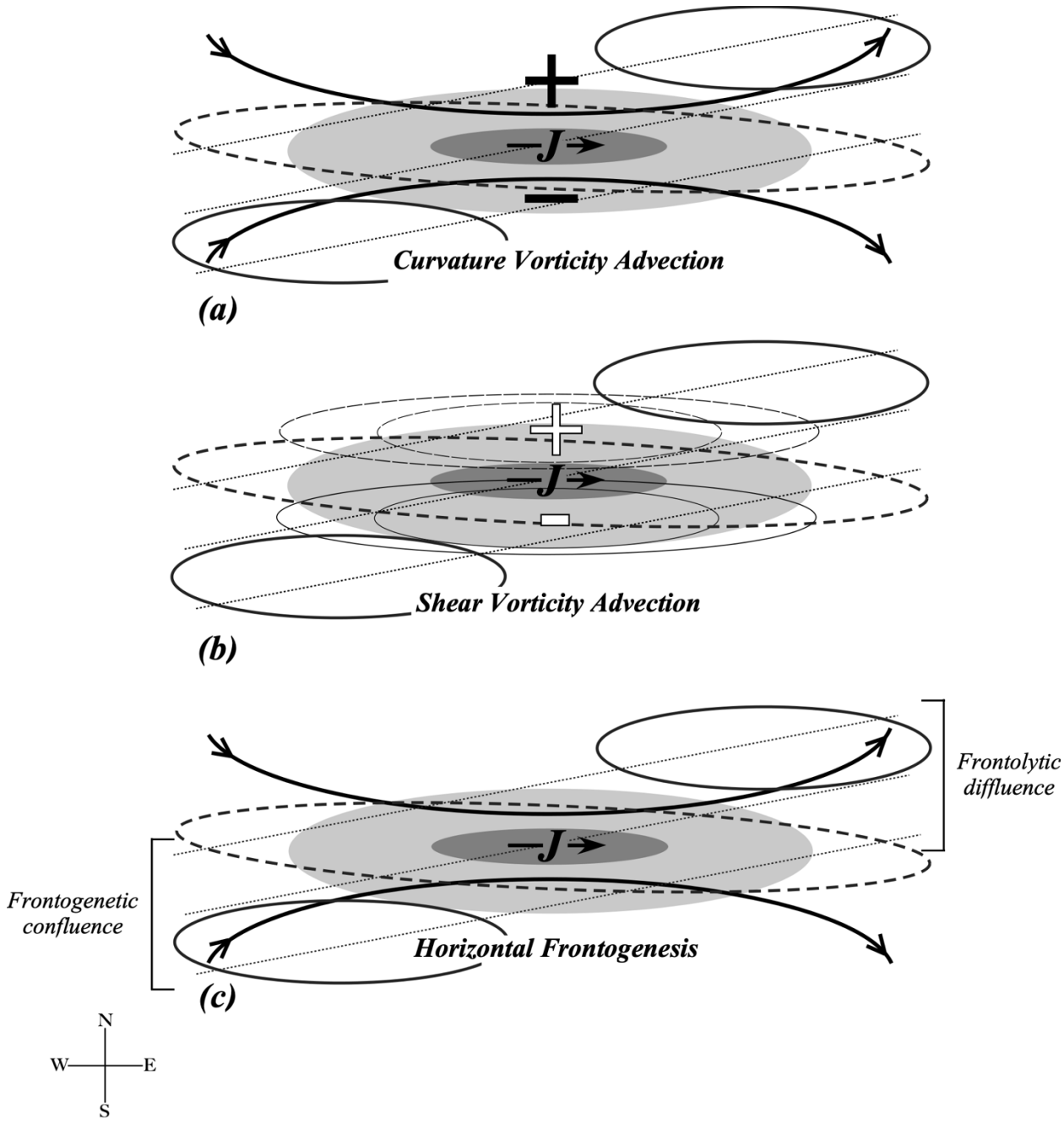

Fig. 12 Schematic straight jet streak with along-flow geostrophic air advection illustrating the vertical circulations associated with (a) geostrophic curvature advection by the thermal wind, (b) geostrophic shear vorticity advection by the thermal wind, and (c) horizontal frontogenesis. Shading represents isotachs with "**J**" indicating the jet core. Dashed lines are isentropes with colder air to the north. Medium solid (dashed) lines are regions of quasi-geostrophic (QG) ascent (descent). In (a) and (c) the bold solid lines are geopotential height lines while in (b) the thin dashed (solid) lines represent schematic contours of positive (negative) geostrophic shear vorticity.

(Adapted from Martin 2014)