Chapter 5

The concept of balance

The concept of balance is discussed in the framework of the primitive equations in ln-p coordinates. Particular insight is provided by the Eliassen problem which teaches us how balanced flows responds to a weak external perturbation.

5.1 General idea

A wind field $\mathbf{u}_b$ is a *balanced wind field* if it is diagnostically related to the mass field, i.e.

$$\mathbf{u}_b = \mathcal{F}(\text{mass field}) \; , \tag{5.1}$$

where the symbol $\mathcal{F}$ may involve spatial derivatives, but no time derivaties. In other words, knowledge of the mass distribution at any particular instance allows one to compute the balanced wind $\mathbf{u}_b$ at that instance.

Newton's second law relates the acceleration of a fluid element to the sum of all forces working on it. Let us assume that the distribution of mass allows us to compute the force acting on each particle (which is possible, e.g., in a hydrostatic atmosphere). Knowledge of the location of every fluid parcel, i.e. knowledge of the mass distribution, is generally not sufficient to compute unique solution, because the solution of a second order differential equation requires two initial conditions (typically initial location *and* initial momentum) for any fluid parcel. It is, therefore, generally not possible to uniquely diagnose a wind field from the mass distribution alone; rather, given the initial location of all fluid parcels, there is still a large variety of different wind fields corresponding to different initial conditions for the wind. The definition of a specific balance through $\mathcal{F}$ picks one particular solution out of many possible ones. One generally aims to specify such a balance condition $\mathcal{F}$ that the resulting balanced flow $\mathbf{u}_b$ carries (to a good approximation) all the information on the scales of interest while filtering out other types of motion which are of less interest.

5.2 Geostrophic balance

In this section we introduce the geostrophic wind as the simplest example for the general concept of a balanced wind field. Starting point are the primitive equation in log-pressure coordinates (4.76).

The geostrophic wind

Let us consider frictionless ($\mathbf{X} = \mathbf{0}$) flow. Regarding the remaining terms in the momentum equation (4.76a) on synoptic scales, the Coriolis force and the pressure force are much larger than $D\mathbf{v}/Dt$. More formally, on synoptic scales the Rossby number is very small ($Ro \ll 1$) and the momentum balance approximately reads

$$f\,\mathbf{k} \times \mathbf{v} \approx -\nabla_h \Phi \; . \tag{5.2}$$

The *geostrophic wind* $\mathbf{v}_g$ is defined to be that wind for which the above relation is exactly satisfied with f replaced by f_0, i.e.

$$\boxed{f_0\,\mathbf{k} \times \mathbf{v}_g = -\nabla_h \Phi} \; , \tag{5.3}$$

and this balance is called *geostrophic balance*. Forming $\mathbf{k} \times$ (5.3), we get

$$\boxed{\mathbf{v}_g = \frac{1}{f_0}\,\mathbf{k} \times \nabla_h \Phi} \; . \tag{5.4}$$

By definition the geostrophic wind is strictly horizontal and parallel to the isolines of geopotential Φ. On the Northern Hemisphere ($f_0 > 0$) the geostrophic wind blows counterclockwise around a pressure minimum, and on the Southern Hemisphere ($f_0 < 0$) the opposite is true. Introducing the *geostrophic stream function*

$$\psi := \frac{1}{f_0}\Phi \; , \tag{5.5}$$

the geostrophic wind can be written as

$$\mathbf{v}_g = \mathbf{k} \times \nabla\psi \; , \quad \text{or} \quad \begin{pmatrix} u_g \\ v_g \end{pmatrix} = \begin{pmatrix} -\psi_y \\ \psi_x \end{pmatrix} \; . \tag{5.6}$$

Right at the equator, where $f_0 = 0$, the geostrophic wind is undefined. Nevertheless, geostrophic balance turns out to hold approximately even close to the equator for the larger scales of the flow. This does not necessarily imply very large values of $\mathbf{v}_g$ (as (5.4) might suggest at first sight), rather the horizontal gradients of temperature and, hence, of geopotential become very small as one approaches the equator such that (5.4) yields reasonable values for the vertical wind despite the smallness of the Coriolis parameter.

Rewriting the hydrostatic balance (4.76b) as

$$\frac{\partial \Phi}{\partial z} = \frac{1}{\rho}\frac{p_0}{R}e^{-z/H} \; , \tag{5.7}$$

we see that knowldege of the mass distribution $\rho(\mathbf{x})$ is sufficient to compute $\Phi(\mathbf{x})$ at any instance of time through a simple integration, and from this we obtain $\psi(\mathbf{x})$ via (5.5) and $\mathbf{v}_g$ through differentiation. It follows that the geostrophic wind $\mathbf{v}_g$ is a balanced wind as defined in the previous subsection, i.e. related to the mass field in a purely diagnostic manner.

In summary, assuming both hydrostatic and geostrophic balance, the knowledge of the mass distribution $\rho(\mathbf{x})$ alone allows one to diagnostically compute $\Phi(\mathbf{x})$ and, from this, the temperature $T(\mathbf{x})$ via (4.76b) through a vertical derivative and the balanced geostrophic wind $\mathbf{v}_g$ via (5.4) through horizontal derivatives.

The thermal wind

Using (5.4), the vertical shear of the geostrophic wind reads

$$\frac{\partial \mathbf{v}_g}{\partial z} = \frac{1}{f_0} \mathbf{k} \times \nabla_h \frac{\partial \Phi}{\partial z} \, , \tag{5.8}$$

which together with the hydrostatic relation (4.76b), $\partial_z \Phi = gT/T_s$, becomes

$$\boxed{\frac{\partial \mathbf{v}_g}{\partial z} = \frac{g}{T_s f_0} \mathbf{k} \times \nabla_h T} \, . \tag{5.9}$$

This is the *thermal wind relation*, which tells us that a horizontal temperature gradient $\nabla_h T$ is associated with a vertical shear $\partial_z \mathbf{v}_g$ of the geostrophic wind. The thermal wind relation is a diagnostic relation between the wind and the temperature field; it is a result of assuming both hydrostatic and geostrophic balance. In components, it reads

$$\begin{pmatrix} \partial_z u_g \\ \partial_z v_g \end{pmatrix} = \frac{g}{T_s f_0} \begin{pmatrix} -\partial_y T \\ \partial_x T \end{pmatrix} \, . \tag{5.10}$$

The thermal wind relation can be used to explain one of the key features of the Earth's general circulation. Generally, the troposphere is warm in the tropics and subtropics, while it is cold in polar regions, and this is obviously related to solar forcing and astronomical properties like the inclination of the Earth's axis of rotation. It follows that there is a significant equatorward temperature gradient in the extratropics of both hemispheres. Since the effect of the lower boundary is such as to reduce the wind to values close to zero, the upper tropospheric wind is given by the tropospheric wind shear $\partial \mathbf{v}_g / \partial z$ or, more formally,

$$\mathbf{v}_g(z) \approx \int_0^z \frac{\partial \mathbf{v}_g}{\partial z} \, dz = \int_0^z \frac{g}{T_s f_0} \mathbf{k} \times \nabla_h T \, dz = \frac{g}{T_s f_0} \mathbf{k} \times \int_0^z \nabla_h T \, dz \, . \tag{5.11}$$

Thus, in both hemispheres an equatorward temperature gradient in the troposphere is related to westerly winds in the upper troposphere, and these westerlies are called *jet streams* or simply *jets*. Note that in the lower extratropical stratosphere the temperature gradients are generally polewards rather than equatorward leading to negative wind shear for the zonal wind. It follows that the jets have their maximum strength close to the tropopause.

A batrotropic atmosphere was defined to be characterized by $\rho = \rho(p)$. For an ideal gas with

$$\rho = \frac{p}{RT} \tag{5.12}$$

this is equivalent to requiring $T = T(p)$. It follows that in a barotropic atmosphere the horizontal gradients of temperature in pressure coordinates vanish

$$\nabla_h T|_p = 0 \quad \Leftrightarrow \quad \text{barotropic atmosphere .} \tag{5.13}$$

Using the thermal wind relation (5.9), this becomes

$$\frac{\partial \mathbf{v}_g}{\partial z} = 0 \quad \Leftrightarrow \quad \text{barotropic atmosphere .} \tag{5.14}$$

In a barotropic atmosphere the geostrophic wind does not have vertical shear (in pressure coordinates). In a baroclinic atmosphere, on the other hand, the geostrophic wind has a nontrivial vertical shear

$$\frac{\partial \mathbf{v}_g}{\partial z} \neq 0 \quad \Leftrightarrow \quad \text{baroclinic atmosphere .} \tag{5.15}$$

5.3 Nonlinear balance

to be written

5.4 Weakly disturbed balance: the Eliassen problem

Consider zonally symmetric flow ($\partial/\partial x = 0$) on the f-plane in a channel which has solid walls at the northern and southern boundary and which is periodic in the zonal direction. In this case the primitive equations reduce to

$$\boxed{\begin{aligned}
&\partial_t u + v\,\partial_y u + w\,\partial_z u - fv = X\ , && (a)\\
&\partial_t v + v\,\partial_y v + w\,\partial_z v + \underline{fu} = \underline{-\partial_y \Phi} + Y\ , && (b)\\
&\underline{\partial_z \Phi} = \underline{\frac{g}{T_s} T}\ , && (c)\\
&\rho_0\,\partial_y v + \partial_z(\rho_0 w) = 0\ , && (d)\\
&\partial_t T + v\,\partial_y T + Sw = \tilde{Q}\ , && (e)
\end{aligned}} \tag{5.16}$$

with $\tilde{Q} = J/c_p$ and $S = \partial_z T + \kappa T/H$.

Assume, first, purely conservative conditions with $X = Y = 0$ and $\tilde{Q} = 0$. In this case there is a very simple balanced stationary solution with $v = w = 0$, for which all terms in the above equations vanish except those which are underlined. We are left with $u(y,z)$, $\Phi(y,z)$ and $T(y,z)$ satisfying

$$fu = -\partial_y \Phi\ , \tag{5.17}$$

$$\partial_z \Phi = \frac{g}{T_s} T\ . \tag{5.18}$$

Forming $\partial_z(5.17)$, $\partial_y(5.18)$, and eliminating $\partial^2_{yz}\Phi$ yields

$$\boxed{f\,\partial_z u = -\frac{g}{T_s}\,\partial_y T}\ , \tag{5.19}$$

which is the thermal wind equation for this special situation.

The problem we want to address was originally studied by A. Eliassen in 1952, which is why we call it the *Eliassen problem.* Let us start with a balanced *primary flow* flow as given by $u(y,z)$ and $T(y,z)$ satifsying (5.19). Then, at time t_0 we switch on some weak external forcing X, Y, and $\tilde{Q}$. As a result the solution ceases to be stationary after t_0, and in addition there will be a weak circulation (v,w) in the meridional plane. The latter is called *Eliassen secondary circulation*, in order to distinguish it from the primary flow. The problem which we want to solve is to find an expression for the secondary circulation (v,w) in terms of the weak forcing terms X, Y, and $\tilde{Q}$.

As long as the external forcing is weak, we can assume that those terms in (5.16) which are not underlined are small in comparison with those which are underlined.[1] In particular, we can assume that (5.16b) and (5.16c) can be approximated by (5.17) and (5.18) such that the thermal wind relation (5.19) remains vaild despite the external forcing. Forming $\partial_t(5.19)$ gives

$$f\,\partial^2_{zt}u = -\frac{g}{T_s}\,\partial^2_{yt}T\ , \tag{5.20}$$

which will be used below.

First we note that continuity (5.16d) allows us to introduce a mass stream function ψ as follows

$$\boxed{\begin{pmatrix}\rho_0 v\\ \rho_0 w\end{pmatrix} = \begin{pmatrix}-\partial_z\psi\\ \partial_y\psi\end{pmatrix}}\ , \tag{5.21}$$

which is defined such that (5.16d) is satisfied automatically. The Eliassen problem reduces to finding a relation between the external forcing and ψ.

Forming $f\,\partial_z(5.16a)$ we get

$$f\,\partial^2_{tz}u + f\partial_z\left[v(\partial_y u - f)\right] + f\,\partial_z(w\,\partial_z u) = f\,\partial_z X\ , \tag{5.22}$$

and forming $gT_s^{-1}\partial_y(5.16e)$ yields

$$\frac{g}{T_s}\,\partial^2_{yt}T + \frac{g}{T_s}\,\partial_y(v\,\partial_y T) + \frac{g}{T_s}\,\partial_y\,(Sw) = \frac{g}{T_s}\,\partial_y\tilde{Q}\ . \tag{5.23}$$

Adding these two equations and using (5.20) to cancel the time derivatives leaves us with

$$f\partial_z\left[v(\partial_y u - f)\right] + f\,\partial_z(w\,\partial_z u) + \frac{g}{T_s}\,\partial_y(v\,\partial_y T) + \frac{g}{T_s}\,\partial_y\,(Sw) = f\,\partial_z X + \frac{g}{T_s}\,\partial_y\tilde{Q}\ . \tag{5.24}$$

[1] In fact, this latter condition can be taken as a definition for the external forcing to be weak.

Using (5.21) to write v and w in terms of ψ and multiplying by ρ_0, this becomes

$$\begin{aligned} \rho_0 \frac{\partial}{\partial z}\left[(f\,\partial_y u - f^2)\left(-\frac{\partial_z \psi}{\rho_0}\right) + f\,\partial_z u\left(\frac{\partial_y \psi}{\rho_0}\right)\right] + \\ + \frac{\partial}{\partial y}\left[-\frac{g}{T_s}\,\partial_y T\,\partial_z \psi + \frac{g}{T_s} S\,\partial_y \psi\right] = \\ = \rho_0\left(f\,\partial_z X + \frac{g}{T_s}\,\partial_y \tilde{Q}\right)\,. \end{aligned} \tag{5.25}$$

As an abbreviation we introduce

$$\boxed{N^2 := \frac{g}{T_s} S \equiv \frac{g}{T_s}\left(\frac{\partial T}{\partial z} + \frac{\kappa}{H} T\right) \equiv \frac{g}{T_s}\frac{\partial \theta}{\partial z}\, e^{-\frac{\kappa z}{H}} \equiv \left(\frac{T}{T_s}\right)\frac{g}{\theta}\frac{\partial \theta}{\partial z}}\,, \tag{5.26}$$

which is a natural measure of static stability in ln-p coordinates (see (4.87) in section 4.3). As long as $N^2 > 0$ or $S > 0$, the atmosphere is statically stable. Using this new piece of teminology in the above equation, we get

$$\boxed{\mathcal{L}\,\psi = \mathcal{F}}\,, \tag{5.27}$$

where the linear partial differential operator $\mathcal{L}$ is defined through

$$\mathcal{L}\,\psi = \frac{\partial}{\partial y}\left(N^2 \frac{\partial \psi}{\partial y} + f\,\partial_z u\,\frac{\partial \psi}{\partial z}\right) + \rho_0 \frac{\partial}{\partial z}\left[\frac{1}{\rho_0} f\,\partial_z u\,\frac{\partial \psi}{\partial y} + \frac{1}{\rho_0} f(f - \,\partial_y u)\frac{\partial \psi}{\partial z}\right]\,, \tag{5.28}$$

and the forcing $\mathcal{F}$ is given by

$$\mathcal{F} = \rho_0 f\,\frac{\partial X}{\partial z} + \rho_0 \frac{g}{T_s}\frac{\partial \tilde{Q}}{\partial y}\,. \tag{5.29}$$

For constant zonal flow $u = const$ and constant $N^2 = const$, equation (5.27) reduces to

$$N^2\,\frac{\partial^2 \psi}{\partial y^2} + f^2\,\rho_0 \frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial \psi}{\partial z}\right) = \mathcal{F}\,. \tag{5.30}$$

Generally, however, the terms with the mixed partial derivatives do not vanish. Rewriting the operator $\mathcal{L}$ in its canonical form

$$\mathcal{L}\,\psi = A\,\frac{\partial^2 \psi}{\partial y^2} + 2B\,\frac{\partial^2 \psi}{\partial y\,\partial z} + C\,\frac{\partial^2 \psi}{\partial z^2} + \ldots \tag{5.31}$$

where the dots indicate a number of further terms, we get

$$\mathcal{L}\,\psi = N^2\,\frac{\partial^2 \psi}{\partial y^2} + 2f\,\partial_z u\,\frac{\partial^2 \psi}{\partial y\,\partial z} + f(f - \partial_y u)\,\frac{\partial^2 \psi}{\partial z^2} + \ldots\,. \tag{5.32}$$

The theory of linear partial differential equations tells us that (5.27) is an *elliptic partial differential equation* if and only if $AC - B^2 > 0$, i.e.

$$\boxed{\mathcal{L}\ \text{ is an elliptic operator} \quad \Leftrightarrow \quad N^2\,f(f - \partial_y u) > f^2(\partial_z u)^2}\,. \tag{5.33}$$

As long as the atmosphere is statically stable ($N^2 > 0$) and both the meridional and vertical shear of the zonal wind are not too strong, (5.27) is an elliptic partical differential equation. Given the external forcings X and $\tilde{Q}$ and assuming that appropriate boundary conditions for ψ are specified, it can be solved for ψ, from which we get the *Eliassen secondary circulation* (v, w) using (5.21). Linearity of (5.27) implies that the strength of the secondary circulation is proportional to the strength of the external forcing.

What happens if (5.33) is not satisfied, i.e. if $\mathcal{L}$ ceases to be elliptic? Later in chapter 9 we shall see that (5.33) is precisely the condition for "symmetric stability". In other words, if the ellipticity condition is violated, a completely new process takes over, namely symmetric instability and there is a sudden transition to unbalanced motions. In such an unstable situation we cannot expect any longer that the flow adjusts to a balanced state.

Let us consider a specific example. Assume $u = const$, $X = 0$, and weak heating $\tilde{Q} > 0$ somewhere in the interior of the domain. Owing to the meridional derivative in $\mathcal{F}$, this term has a maximum at low latitudes and a minimum a higher latitudes in the middle troposphere. Qualitatively the pattern of ψ resembles a smeared-out version of the pattern of $\mathcal{F}$, so we get a minimum at low latitudes and a maximum at high latitudes in the middle troposphere. The Eliassen secondary circulation is upward in the center of the heating ($w \propto \partial_y \psi$), downward at high and low latitudes, and above and below the forcing there is meridional flow for mass continuity.

Effect of the Eliassen secondary circulation

Intially, i.e. for $t \leq t_0$, the primary vortex is characterized by $u(y, z)$ and $T(y, z)$, which are in balance meaning that (5.19) is satisfied. At $t = t_0$ we switch on the forcing and immediately do we get the secondary circulation which is diagnostically related to the external forcing X and $\tilde{Q}$. The Eliassen secondary circulation is thus an immediate (albeit weak) response to (weak) forcing. We can now insert v and w into (5.16a) and (5.16e) to compute the ensueing (slow) evolution of the primary vortex as follows

$$\partial_t u = X + (f - \partial_y u)v - \partial_z u \, w \; , \tag{5.34}$$

$$\partial_t T = \tilde{Q} - f \, \partial_y T - Sw \; . \tag{5.35}$$

The slowness of this evolution is related to the weakness of the forcing: to the degree that X and $\tilde{Q}$ are weak, so is $\mathcal{F}$ and, hence, ψ and (v, w). Although X and $\tilde{Q}$ are unrelated and may be specified arbitrarily, the rate of change $\partial_t u$ and $\partial_t T$ is such that thermal wind balance (5.19) is satisfied at any time during the (slow) evolution of the primary flow. This is not a miracle but rather happens by design, since we deliberately determined the secondary circulation such as to ensure perpetual validity of (5.19).

Let us consider, again, the above example: $u = const$ at initial time, $X = 0$, and $\tilde{Q} > 0$). In this case we get at $t = t_0$

$$\partial_t u = fv \; , \tag{5.36}$$

$$\partial_t T = \tilde{Q} - Sw \; , \tag{5.37}$$

which means that the heating $\tilde{Q} > 0$ is partly compensated by the adiabatic cooling $-Sw$ owing to the Eliassen secondary circulation. The latter is associated with rising motion in the heating center, which can be considered as an indirect result of the forcing. On the other hand, although $X = 0$, the zonal wind u is affected indirectly through the Coriolis force $-fv$ acting on the induced meridional wind v. Thus, the Eliassen secondary circulation reduces the direct effect of $\tilde{Q}$ on the temperature and "redistributes" part of the effect to the momentum equation. This must be so, for if only temperature T were affected with the zonal flow u remaining unmodified, this would immediately destroy the balance (5.19). The secondary circulation helps to keep (5.19) satisfied at any time by, first, reducing the direct effect of $\tilde{Q}$ on T and, second, transfering part of the effect to the momentum field u.

It is equally instructive to study the opposite extreme: $u = const$ at initial time, $\tilde{Q} = 0$, and $X < 0$, i.e. we force the primary flow to locally reduce its momentum (like X being an effective friction force). In this case the pattern of $\psi(y, z)$ is a dipole with the positive center below the forcing, the negative center above the forcing, and with poleward flow in the center of the forcing. For the rate of change of the primary flow we get at $t = t_0$

$$\partial_t u = X + fv \, , \tag{5.38}$$

$$\partial_t T = -Sw \, , \tag{5.39}$$

Again, $X < 0$ is partly compensated by $fv > 0$ such that the total effect $\partial_t u$ is reduced in comparison with the hypothetical effect of X alone. At the same time the induced vertical wind w modifies the temperature such that both T and u remain balanced at any time.

In a sense both $\partial_t u$ and v can be viewed as caused by X, and

$$\partial_t u - fv = X \tag{5.40}$$

can be interpreted as the effect of X being partly a reduction $\partial_t u$ of the primary flow and and partly a secondary flow v. The secondary flow v is at right angles to the force X, quite like the effect of a torque on a spin, and indeed the Eliassen secondary circulation expresses this *gyroscopic effect* in the continuum formulation appropriate to the stably stratified atmosphere. This analogy is not coincidental. Rather, balanced flow on a rotating sphere is by definition dominated by the effect of rotation — quite like the motion of a spin.

To understand why the response is (partly) at right angles to the direction of the forcing X, is is helpful to consider the full equation (5.16b). Neglecting those terms which are nonlinear in the meridional circulation (and assuming $Y = 0$), we get

$$\partial_t v = -\partial_y \Phi - fu \, . \tag{5.41}$$

Initially ($t \leq t_0$) the two terms on the right hand side cancel each other and we have $\partial_t v = 0$. At time $t > t_0$ the frictional force reduces u but lets $\partial_y \Phi$ unaffected, thus producing an imbalance on the right hand side of the above equation with the result that $\partial_t v > 0$. Thus, the small external force X disturbs the balance between two large forces perpendicular to the primary flow. It is the small residuum which accelerates the flow a direction at right angles to the original force X.

"Downward control"

The zonal mean circulation in the middle atmosphere is affected by both Rossby and gravity waves which propagate from below and which dissipate in the stratosphere and mesosphsere. Let us for the moment view this as an external drag force $X = const < 0$ (although this is conceptionally rather problematic, see Egger 1996). Let us furthermore assume that the heating is given by Newtonian relaxation towards an equilibrium temperature $T_{eq}(y,z)$ with a relaxation constant α, i.e.

$$\tilde{Q} = -\alpha(T - T_{eq}) \ . \tag{5.42}$$

Let at initial time be $T = T_{eq}$ and u in balance with this temperature. Switching on X immediately induces a Eliassen secondary circulation (v,w), where at initial time the forcing $\mathcal{F}$ is given by

$$\mathcal{F} = \rho_0 f \frac{\partial X}{\partial z} \ . \tag{5.43}$$

The ensueing evolution of both u and T drives the temperature away from T_{eq} and we obtain external heating in addition to momentum forcing such that at later times the forcing includes both X and $\tilde{Q}$, i.e.

$$\mathcal{F} = \rho_0 f \frac{\partial X}{\partial z} - \alpha\,\rho_0 \frac{g}{T_s}\frac{\partial(T - T_{eq})}{\partial y} \ . \tag{5.44}$$

The equations for the rate of change of the zonal mean flow read

$$\partial_t u = X + (f - \partial_y u)v - \partial_z u\, w \approx X + fv \ , \tag{5.45}$$

$$\partial_t T = -\alpha(T - T_{eq}) - f\,\partial_y T - Sw \approx -\alpha(T - T_{eq}) - Sw \ , \tag{5.46}$$

where in the second expression on the right hand side we neglected the terms related to the shear of the zonal wind u.[2] Owing to the relaxational character of $\tilde{Q}$, we may expect the solution to approach a steady state for which the effect of the induced secondary circulation exactly balances that of the external forcings,

$$0 = X + fv \ , \tag{5.47}$$

$$0 = \tilde{Q} - Sw \ . \tag{5.48}$$

The stationary scenario represents the extreme case of total compensation between the nonconservative terms (X and $\tilde{Q}$, respectively) and the effect of the Eliassen secondary circulation (fv and Sw, respectively). The way this experiment is set up, we can view X as the forcing and v, w, and $\tilde{Q}$ as the response of the system.

In steady state the response to X in terms of v is a simple local relation

$$\boxed{v(y,z) = \frac{1}{f}\,X(y,z)} \ . \tag{5.49}$$

It is not much harder to find an expression for w in terms of X. Forming $\rho_0 \partial_y \ldots$ of the former equation and integrating from z through infinity, we get

$$-\int_z^\infty \rho_0\,\partial_y X\,dz' = \int_z^\infty f\,\rho_0\,\partial_y v\,dz' = -\int_z^\infty f\,\partial_z(\rho_0 w)\,dz' = +f\rho_0\,w(z) \ , \tag{5.50}$$

[2]The derivation can be done without neglecting the terms associated with the shear of the zonal wind u, see Haynes et al. (1991).

where we used the continuity equation (5.16d) and assumed $w = 0$ at $z = \infty$. Hence, the vertical wind at z can be expressed as

$$\boxed{w(y,z) = -\frac{1}{f\rho_0}\frac{\partial}{\partial y}\left(\int_z^\infty \rho_0 X(y,z')\,dz'\right)}\,. \tag{5.51}$$

In steady state, the vertical wind at level z is diagnostically related to the zonal forcing X above that level. Haynes et al. (1991) introduced the term *downward control* in connection with this relation, although — as was pointed out by Egger (1996) — it is dangerous to invoke the concept of cause and effect ("control") in a diagnostic relationship such as (5.51). To be sure, it is meaninful to talk about cause and effect during the transient evolution outlined above, which eventually leads to the steady state. A major criticism then refers to the assumption that X can be viewed as externally given, which in the real atmosphere is not a priori a useful view since wave propagation and dissipation does depend on the state of the zonal mean atmosphere. To the degree that we are willing to view w as contolled by the X as quantified by (5.51), so can we view the heating $\tilde{Q}$ as contolled by X via

$$\tilde{Q}(y,z) = Sw = -\frac{S}{f\rho_0}\frac{\partial}{\partial y}\left(\int_z^\infty \rho_0 X(y,z')\,dz'\right)\,. \tag{5.52}$$

The above can be used to explain why the stratospheric polar vortex is much colder in the Southern Hemisphere than in the Northern Hemisphere. Owing to the stronger orography and larger thermal contrasts in the longitudinal direction, the wave forcing X is about twice as strong in the Northern Hemisphere than in the Southern Hemisphere. As a consequence, the *dynamical heating* as computed from (5.52) is about twice as strong in the Northern Hemisphere rendering the Northern polar vortex significantly warmer than the Southern polar vortex. It is not surprising to observe a much stronger ozone hole in the Southern Hemisphere, because the heterogeneous chemistry which eventually destroys ozone requires very low temperatures.

As a caveat we point out that for some relevant applications the middle atmosphere cannot be assumed to be in steady state regarding the momentum balance; in these cases the full transient evolution needs to be considered and the concept of downward control may be misleading.

References

Egger, J., 1996: Comments on "On the "Downward Control" of Extratropical Diabatic Circulation by Eddy-Induced Zonal Forces". *J. Atmos. Sci.*, **53**, 2103–2104.

Eliassen, A., 1952: Slow thermally or frictionally controlled meridional circulation in a circular vortex. *Astrophys. Norv.*, **5**, No 2, 19–60.

Haynes, P. H., C. J. Marks, M. E. McIntyre, T. G. Shepherd, and K. P. Shine, 1991: On the "Downward Control" of Extratropical Diabatic Circulations by Eddy-Induced Mean Zonal Forces. *J. Atmos. Sci.*, **48**, 651–678.