Chapter 1

Foundation and Context

Purpose: Foundation materials for the other chapters: a quick review of Earth’s geography, definitions of some common terms, and intuitive properties of the atmosphere and ocean circulations.

1.1 Orientation

1.1.1 Overview

Understanding the general circulation spans the broad spectrum of subjects in meteorology. To study the general circulation requires much more than dynamics, one must include radiation, synoptics, moist processes, even chemistry. This grand circulation is global in scope and so this book begins with a brief review of Earth’s basic geography. Many elements of the circulation can be understood from fundamental principles, such as conservation of mass, energy, and momentum. Those core principles, along with some observations allow one to deduce some broad properties of the global circulations and by doing so, in this chapter, the intent is to provide a meaningful context to being discussion of those features and why they occur. Since the atmosphere interacts strongly with the ocean, some very simple properties of the ocean are reviewed.

What is the general circulation? Monin (1986) defines it as “a statistical ensemble of large-scale components of state of the atmosphere.” Wallace and Hobbs (2006) have a similar definition; the general circulation is “the statistical properties of the large-scale atmospheric motions, as viewed in a global context.” In the earliest works relating to the subject, general circulation meant the time and zonal average circulation. (The term “zonal” refers to the longitudinal direction.) However, mechanisms are needed to drive the motions, and those mechanisms draw in all the other commonly studied variables: pressure, density, temperature,
and moisture. In addition, the mechanisms maintaining the zonal average circulation include zonally varying phenomena (eddies) and time varying (transient) phenomena. So, the lack of specificity in current definitions is a necessity.

Scale analyses of the equations provide insight into the different balances in the equations that predominate for different scales in time and space of different specified phenomena. It has been popular practice to identify distinct scales of motion in the atmosphere based upon such scale analyses (e.g., Fujita, 1981, and 14 references in his article). Unfortunately, plenty of phenomena bridge these tidy categories. For example, in the tropics the large-scale atmospheric motions are composed of, and driven by, organized small-scale convection. So, the definition of the general circulation is ambiguous in length and time scales. Our subject is not completely undefined, the emphasis is upon understanding the largest scales, greater than 1000km and lasting longer than several weeks; it is just that phenomena with shorter length and time scales are often needed to reach an understanding.

What causes the general circulation is easy to answer: the atmosphere develops motions in an attempt to reach thermal equilibrium between net sources and sinks of radiant energy. One can see these source regions (latitudes <38 degrees) and sink regions (latitudes >38 degrees) in Figure 1.1a. From that figure one deduces that there must be a circulation transporting heat because the solar radiation absorbed has greater latitudinal variation than the terrestrial radiation emitted. In a similar vein, the zonal and annual average precipitation (Figure 1.1b) has a markedly different variation than does evaporation which requires transport of moisture to be sustained. While the oceans do accomplish some heat and moisture transports, the ocean alone cannot explain the precipitation minus evaporation difference. Beyond that, matters become very complicated.

The properties of the circulation are a harder story to tell and they are the focus of this book. To answer the question properly one must look at a variety of things that are intertwined with the winds, such as radiation, moisture, and landform. The challenge is keeping the discussion manageable. A jigsaw puzzle of a picture may be a good metaphor. One does not need all the pieces to be in place but one does need enough that sample the main parts of the image in order to see the general picture made by the pieces. Similarly, this book gathers most of the largest pieces and assembles them in a way that makes the giant multivariate jigsaw puzzle of the general circulation understandable. Of necessity though, some phenomena must be neglected or discussed only superficially.

1.1.2 Basic Physical Geography

The zonal average weather patterns are rather symmetric between the Northern and Southern Hemispheres. Most differences that do occur between the two hemispheres can be
ascribed to the differing fraction of land area as well as major topographic features of each hemisphere. Figure 1.2a also shows the major land masses along with the major topographic features. The major mountain ranges are clearly evident, as are broad plateaus of: Tibet and western China, southern Africa, and western North America. The major ice sheets (Greenland and Antarctica) have high elevation as well. The higher topography can block or divert the flow that is apparent on a time mean as a standing wave pattern. Examples include the long wave upper air pattern and the south Pacific subtropical high. High topography has other important effects, such as introducing surface heating at a high level, as might be important for the Asian Monsoon.

The percentage of land and ocean areas in each five-degree latitude belt is diagrammed in Figure 1.2b. Land interacts differently with the atmosphere than does the ocean. Thermodynamic properties are different: heat capacity, availability of moisture, and surface reflectivity. The ocean moderates the seasonal change because the water can mix the heat convectively (as mentioned in the quotation by Benjamin Franklin). The solar heating is input at the top which stabilizes the temperature profile, but the winds cause a turbulent ‘convection’ that mixes the heating through a considerable depth. This wind mixed layer is on the order of 50m. In contrast, heat must be conducted into and out of subsurface soil, a much slower process than wind mixing. In addition, water has higher specific heat than dry soil. The heat required to raise the temperature of a kg of material by one degree Celsius is the heat capacity. The specific heat capacity varies with soil composition and wetness, but some general numbers are these: for dry soil C_{sd} ≈ 800 J/(kgK) and for wet soil the heat capacity is C_{sw} ≈ 1500 J/(kgK). For water the heat capacity is C_{w} ≈ 4200 J/(kgK). These two factors coupled together, heat mixed through greater volume (and thus more mass) of the ocean multiplied by the higher specific heat of water, give the ocean areas much greater capacity to buffer temperature changes due to heating and cooling than the land. The Northern Hemisphere has much more land area, and consequently the seasonal variability is much greater in that hemisphere. Of course, land is at the South Pole and ocean is at the North Pole. But the high reflectivity of year-round snow cover diminishes seasonal change over Antarctica. Arctic sea ice is also highly reflective, but Arctic sea ice is much thinner (about 3 m thick, on average). Heat conduction, through Arctic sea ice and at leads (breaks in the ice pack), reduces the seasonal variability of the Arctic and maintains moderate polar temperatures in contrast to the Antarctic.

The different radiative responses of land and sea are notable in Figure 1.3. The figure shows annual average absorbed shortwave (solar) radiation and corresponding emitted terrestrial (longwave) radiation. The absorbed and emitted radiation is nearly zonally symmetric across each ocean basin except near the continental edges where surface currents parallel the coastlines and cause the sea surface temperature and cloud cover to differ from the oceanic zonal mean.
Extremely simple arguments are given later in this chapter to deduce these boundary currents in subtropical and middle latitudes. Mountainous regions also disrupt the zonal symmetry, as might be expected since the emission temperature and surface vegetation respond to elevation.

1.1.3 Some Common Jargon

The previous paragraph referred to subtropical and middle latitudes, which demand defining. Overlapping latitudinal ranges will be labeled as follows. ‘Tropics’ refer to latitudes between 30S and 30N. The ‘subtropical’ latitudes lie between 20 and 40 degrees. The ‘middle’ latitudes mean the same as ‘mid-latitudes’ and refer to 30-60 degrees. ‘Polar’ regions are poleward of 60 degrees. Overlapping ranges are used in part because phenomena characteristic of one latitude band or another do not always fit within neatly separate latitude ranges, particularly if one considers different longitudes.

Zonal was defined above as referring to the east-west direction. ‘Meridional’ refers to the north-south direction, generally measured with the origin at the equator. ‘Eddies’ refer to deviations from a zonal average; such deviations are often similar in appearance to transverse waves. Long waves refers to waves with wavelengths greater than ~8,000 km. The ‘stationary’ part of the general circulation often has a long wave pattern, where ‘stationary’ refers to the time mean and generally contrasts with ‘transient’ eddies that are time varying, often due to their moving horizontally.

Another useful bit of jargon is to imagine two types of zonal mean ‘meridional circulations’ or ‘meridional cells’. In the tropics one imagines a pair of overturning circulations labeled ‘Hadley cells’, one mainly in the Northern, and the other mainly in the Southern Hemisphere. Each Hadley cell has rising near the equator, poleward motion aloft, sinking in the subtropics, and low level return flow towards the equator. They are named in honor of George Hadley, who in 1735 first described this overturning circulation, including an application of conservation of angular velocity (not momentum). Others, such as Dalton in 1793, rediscovered Hadley’s ideas. The other useful concept is of a meridional cell that circulates in the opposite direction in the middle latitudes of each hemisphere. This cell is labeled the Ferrel after William Ferrel, who published his idea first in 1856. A similar circulation was postulated by Maury in 1855 and independently by Thomson in 1857, but Ferrel’s name is associated with the circulation perhaps because he was more active in publicizing and refining his ideas (Ferrel, 1859, 1893). Further discussions of the historical development of general circulation ideas can be found elsewhere (e.g. Lorenz, 1967; Grotjahn, 1993).

1.2 Basic balances of the General Circulation
Despite the great complexity of the atmospheric and oceanic processes, one can deduce the general properties of the large scale atmospheric circulation by applying some balances augmented by a few empirical facts. By exploiting the concept of balances, one can deduce much about the general circulation. Certain balances are believed to hold and these give rise to formulas and understanding as to why the observed circulation looks like it does. The discussion in this section reproduces and updates a similar presentation in Grotjahn (2007) portions are reprinted with permission from Elsevier.

One can invoke physical concepts; two examples being: radiative and angular momentum balance. Radiative balance states that when averaged over the surface of the globe, there will be no net heating up (or cooling down) of the Earth if the same amount of energy is emitted as is absorbed. While there appears to be evidence for global warming, that rate of temperature increase is still far less than the seasonal cycle at most places on Earth. Angular momentum balance has several implications for the general circulation. One implication is surface winds need a mixture of easterlies and westerlies so as not to apply a net torque upon the solid Earth. As a trivial example, if the surface winds were westerly everywhere, then the atmosphere would be applying a net torque upon the oceans and solid earth, as one consequence the rotation of the Earth would speed up.

One can use balances, physical reasoning, and the observed precipitation to deduce some basic properties. Various laws, such as ideal gas (C.1) and conservation of potential temperature (C.2) and (C.3) for adiabatic motions may be applied. Mathematical simplifications, such as hydrostatic and geostrophic balances are useful. An overarching assumption shall be that the atmosphere can be considered annually-stationary: any trends that remain when fields are averaged over a year are small and can be ignored for this discussion.

1.2.1. Radiation and temperature

Excluding global warming, the Earth does not heat up in the net over a year. So, one may expect a balance between absorption and emission of radiant energy. Figure 1.3a shows the net annual average absorption of solar radiation by the Earth (including its atmosphere and ocean). The maximum is near the equator with a strong decrease towards the poles. Absorption is complicated by variations in the reflectance, absorbance, and transmittance of various terrestrial constituents. Despite these complications, the orbital geometry is a primary cause for more absorption averaged over the tropics than averaged over polar regions. That geometry is illustrated in Figure 1.4a. This fact creates a meridional temperature gradient. The loss of radiation back to space is complicated by several factors; for example, the infrared emission from the earth is not from a single surface, but includes all levels in the atmosphere as well as the ocean and earth’s surface.
If the rate at which the atmosphere cools down by infrared emission were large compared to the temperature of the earth divided by the length of a day, then the primary temperature gradient of the atmosphere would be from the subsolar point to the shadow side of the planet. Radiative cooling rates in the troposphere are around 2 degrees per day, which suggests that temperatures would vary longitudinally by only a few degrees, an amount much less than the roughly 210-310 Kelvin range of temperatures over the troposphere. (This thermal inertia is largely maintained by radiative emission from the Earth’s surface.) As a consequence, the diurnal cycle of temperature is small compared to the daily mean temperature. Also, the terrestrial longwave emission is similar to the long term average (Figure 1.3b). Similarly, the 20-70 C range in temperature observed from equator to pole in the troposphere exceeds the temperature change from midday to midnight. (The equator to pole temperature range varies with elevation in the atmosphere.) So, the temperature gradient is primarily meridional rather than zonal. And, longer term (seasonal) average differences in radiation absorbed and emitted therefore dominate the temperature distribution.

One might ask, can the meridional temperature gradient be maintained even while air is moving meridionally? One might assume that air moving poleward is cooling by 2K/day. Across the 30 degrees of the middle latitudes, the zonal and time mean temperature changes by about 50K during winter months. These assumptions imply 25 days are needed to make the temperature change. From that time and the latitudinal distance, one estimates a meridional motion of ~1.5 m/s. Though small, this meridional speed is comparable to observed values.

Defining net radiation as the local difference of absorption minus emission, net radiation is positive over much of the tropics, and negative in the polar regions (averaged over a year). This pattern of net radiation implies a meridional transport of heat, perhaps by some meridional motions. Such motions could conceivably remove the implied meridional temperature gradient, but if they did, the motions would also remove their driving mechanism.

In the stratosphere the temperature pattern is a bit different. Stratospheric temperatures are dominated by radiative and transport processes, whereas tropospheric temperatures are also influenced by latent heat release, and exchanges with the Earth’s surface. The orbital tilt and spherical geometry of the earth combine to give a rather flat distribution of incoming solar radiation during summer in high latitudes. Hence, temperatures have little meridional gradient in much of the upper stratosphere in mid-summer. At upper stratospheric levels, temperatures tend to be coldest near the winter pole. In the lower stratosphere, the coldest temperatures are technically in the tropics, owing to the much higher tropopause there than in polar regions. The pattern is illustrated schematically in Figure 1.4c. However, in mid and upper stratosphere elevations, temperatures tend to decrease from the summer polar region to the winter polar region during those seasons. In other seasons, the mid and upper level temperatures tend to
decrease from the tropics towards both poles. The radiation and temperature fields are covered in more depth in Chapter 3.

1.2.2. Pressure and geopotential

Pressure decreases with increasing elevation more strongly than it varies in the horizontal. At sea level, a vigorous midlatitude frontal cyclone has a minimum pressure of 960 hPa while a very strong high has a maximum pressure of 1060 hPa. Such values are 1000’s of km apart in the horizontal. In the vertical, the pressure drops 100 hPa from its surface values in about 1 km. A similar drop in pressure over a horizontal distance would generate tornadic winds that would rapidly reduce the horizontal gradient. While horizontal gradients are small at sea level, the horizontal gradient can become larger far above the surface due to density differences caused by horizontal temperature gradients.

Both water vapor content as well as temperature in the troposphere decrease from tropics to the polar regions. Hence the virtual temperature decreases in similar fashion. From the hypsometric equation (C.13) in Appendix C, the meridional virtual temperature gradient thereby creates a meridional geopotential gradient. The meridional temperature gradient applies for each layer (above the boundary layer, say) in the troposphere so the thickness variation in each layer is additive. Thus, the meridional gradient of geopotential increases with elevation, as illustrated in Figure 1.4c. Near the ground, the pressure pattern has comparable longitudinal and latitudinal variations. However, the pressure at elevations above roughly 2 km (and geopotential height of pressure surfaces for P < 850 hPa) also decrease from tropics to polar regions.

By similar reasoning, in the stratosphere, the meridional gradient of geopotential decreases with increasing elevation in the summer hemisphere and tends to increase with elevation in the winter hemisphere.

Where there is a horizontal gradient of pressure there will be a pressure gradient force (PGF). In the troposphere the PGF is directed towards the poles and provides the impetus to initiate the circulation. More information is found in Chapter 4.

1.2.3. Winds

The Earth rotates such that observed velocities outside the tropics are subject to a Coriolis force comparable to the poleward PGF. Hence, the meridional gradient of pressure (and geopotential) implies westerly winds from geostrophic balance (C.36). Since the gradient increases with height then so must the westerly winds in the troposphere. It was established above that heat must be transported poleward to explain the latitudinal distribution of net radiation. If the heat is transferred in a way that employs meridional motion, then zonal winds are created to conserve angular momentum during such meridional motions. Are the zonal winds
implied by the virtual temperature distribution consistent with zonal winds created by large scale angular momentum conserving meridional motions? Presumably a balance can be struck whereby stronger meridional motions more effectively remove the equator to pole temperature gradient thereby lessening the zonal thermal winds, and vice versa.

Conservation of angular momentum by meridional motions can create very large zonal winds. (Chapter 6 treats the subject more fully.) For example, if an air parcel having zero zonal motion at the equator were brought to 30 N latitude, an angular momentum conserving zonal wind, \( U_{amc} \) would equal 134 m/s. This is about twice the time mean speed of the strongest portion of the subtropical jet stream found at a similar latitude. Obviously, one can reduce this acceleration by internal viscosity and vertical mixing of momentum by turbulence or convection. The greater the length of time over which friction and mixing can act, the greater the reduction possible. For realistic estimates of internal friction, \( \text{meridional} \) wind speeds of a few cm/s keep \( \text{zonal} \) winds within observed bounds for motions spanning the tropics and subtropics (See Chapter 6, or Grotjahn, 1993; § 6.2.3.2) In the tropical upper troposphere observed meridional velocities average 10-100 times larger, suggesting that either the friction is too weak and/or other mechanisms are important. Another mechanism might be convection that mixes momentum vertically. This may be happening within ‘tropical plumes’ created in advance of upper level troughs that penetrate deeply into the tropics; air is accelerated poleward, but actively growing clouds moderate the angular momentum by vertical mixing. Much more discussion of the wind fields is found in Chapter 4.

1.2.4. Heat transport and mass balance

Surface pressure values in polar regions do not have a significant trend. So, a meridional circulation that transports heat poleward would be expected to have compensating equatorward motion, otherwise mass would accumulate in the polar regions.

The previous three subsections suggest that zonal average meridional velocities are small based on observations and radiative and angular momentum considerations. Can such small meridional motion still transport sufficient heat? Such a heat transport is proportional to the meridional wind, \( v \) times the dry static energy, DSE. DSE is defined by (C.10). (Latent heat processes from water evaporation, transport, and condensation are being ignored for the moment.) DSE is the sum of internal and gravitational energy. As elevation increases the internal energy (temperature-dependent) decreases while the gravitational energy increases (from higher elevation). For a statically stable atmosphere, such as Earth’s, DSE is positive and increases with elevation.

If the air is moving both poleward and equatorward, then net sensible heat transport depends on the difference in dry static energy (DSE), between the north and southward moving
air. Consequently, to have the same net heat flux, a smaller meridional velocity must be balanced by a larger difference in DSE between the north and southward moving air. If the larger DSE difference were due to temperature changes, then that means larger static stability. Alternatively, if the vertical separation were made larger between opposing meridional motions then the DSE difference would be larger and the meridional motions need not be as strong. So in theory the small meridional velocities identified above can transport heat if the DSE change is large enough. Clearly the DSE variation depends on the temperature variation, both horizontally and vertically.

If zonally and vertically-integrated mass transport in the meridional is to be zero, then a poleward DSE heat flux occurs when the poleward moving air has higher DSE than the equatorward moving air. Hence, if the upper level flow has larger DSE than the low level flow, the upper level flow would need to be poleward. Figure 1.4d illustrates the point. (Further development is in Chapter 7.)

Water mass conservation is a different issue because water exists in 3 different states on Earth. Water could transport heat poleward if it were evaporated in lower latitudes (removing heat there), transported poleward by the winds, and condensed at higher latitudes (releasing the heat there). The mass of atmospheric water in vapor form would not be conserved. Water could conceivably be condensed, fall as precipitation, and run off, returning to the lower latitude in liquid form. While this may seem like a reasonable way to accomplish the heat transport, a major drawback is that this mechanism works against the dry heat transport by a zonally-averaged circulation. To see this, one defines a moist static energy (MSE) by adding latent heat content to the DSE. MSE is defined by (C.11). The scale height of water vapor (~2km) is much smaller than that of dry air (~8.5km). So for high enough moisture content, the moist static energy could conceivably decrease with height and if so, poleward heat transport would be accomplished if the poleward moving air were at low levels – opposite to the pattern deduced for DSE. In much of the tropics, MSE is observed to decrease with height in the lower troposphere (p>700). It is conceivable that a circulation with surface poleward motion and mid to lower tropospheric equatorward return flow might transport the necessary heat. Such a circulation might also have a lot of boundary layer friction to control the development of excessive zonal winds from angular momentum conservation. There are potential problems with this idea. First, a mechanism (vertical mixing or solar absorption perhaps) may be needed to overcome a meridional gradient of MSE that may oppose the motion. Second, it may conflict with the observed tropical precipitation pattern.

The observed distribution of precipitation (Pw, Figure 1.1b) is crucial to deducing the Earth’s general circulation. The primary maximum in precipitation is near the equator and relative minima occur in the subtropics. The distribution of evaporation, Ew is opposite to
precipitation in having its maximum in the subtropics and relative minimum near the equator. The precipitation and evaporation patterns exclude the zonal mean meridional ‘boundary layer’ cell postulated in the previous paragraph. Furthermore, secondary maxima of precipitation in northern and southern middle latitudes must to be explained and that selects for a different midlatitude circulation type than in the tropics.

1.2.5. Surface torques

When wind blows across the surface of the earth friction slows down the wind while imparting a torque upon the earth. If the surface winds were to blow the same direction as the rotation everywhere on a planet with no topographic features, then the winds would be accelerating the planetary rotation. Such reasoning suggests that surface winds on the earth are unlikely to be westerly (or easterly) everywhere. The simplest conclusion might be that there is a balance between the torque applied in regions of surface easterlies versus the torque applied by regions of surface westerlies.

Surface torques are applied not only by winds blowing across the surface of the earth but also by surface pressure differences across mountain ranges. Restricting our attention to zonal flows, then westerly winds impart an eastward torque upon the earth. A surface pressure pattern of low on the windward side and high on the lee side of a topographic feature imparts a “mountain torque” that opposes the torque by the winds.

How much of a surface pressure difference across the mountain is needed to balance a given zonal wind? Is such a pressure difference “realistic”? The earth has several major mountain ranges that may provide such a balance, of particular interest for westerly winds are the north-south oriented ranges such as the Andes and Rockies. To gain a sense of the magnitudes involved, it is instructive to make a crude calculation that balances surface stress $\tau$ from the winds against mountain torque. Specifically, for the Andes, the balance (see Peixoto and Oort, 1992; § 11.1.4) is roughly:

$$
\int_0^{2\pi} \int_0^{-z/3} \tau r^3 \cos^2(\phi) d\phi d\lambda = \int_0^{z_{\text{top}}} \{P_\text{E} - P_\text{W}\} r^2 \cos(\phi) d\phi dz
$$

(1.1)

Where: $r$ is the earth’s radius (~6370 km), $\phi$ is latitude, $\lambda$ is longitude, $z_{\text{top}}$ is an elevation above the highest point of the Andes, $P_\text{E}$ is the surface pressure on the east side of the Andes, while $P_\text{W}$ is the corresponding pressure on the west side. (Where $z$ exceeds the local height of the Andes, the pressure difference is zero.) Other Southern Hemisphere topography is ignored. To illustrate, one imagines surface winds of the Southern Hemisphere that are a uniform 10 m/s over all longitudes and over latitudes from equator to 60S. One further assumes the pressure difference is uniform with elevation and latitude and important only for the lowest 3 km. One approximates the surface stress using a drag coefficient of $10^{-3}$ and average surface density of 1.25 kg m$^{-3}$ such
that $\tau \sim 1.25 \times 10^{-1} \text{ kg m}^{-1} \text{ s}^{-2}$. The resultant pressure difference $dp (= P_E - P_W)$ is $\sim 24 \text{ hPa}$. Such a pressure change is within the range of sea level pressure observed across South America and therefore is not unrealistically large. Unfortunately, the observed sea level pressure change across most of the Andes has the opposite sign, implying that surface pressure (and above) across the Andes might not be able to balance surface westerlies. (Note: such westerlies would exert stress upon the ocean, causing a zonal change in the height of sea level, but the change would largely pass the wind torque along to the solid earth.) Another problem is that westerlies over the Andes would be up, not down the sea level pressure gradient.

To summarize, the discussion of balances and key empirical facts implies the following seven properties.

(i) Incoming radiation makes the tropics warmer than the poles, which in turn creates an equatorward pressure gradient (and hence poleward pressure force) above some level in the atmosphere.
(ii) The pressure gradient and Earth's rotation lead the T gradient to have associated westerly winds that increase with elevation.
(iii) Net radiation implies that there must be poleward transport of heat.
(iv) Mass balance requires the poleward heat transport to have “equal” mass flux components directed north and south.
(v) Angular momentum balance anticipates “roughly equal” areas of easterly and westerly surface winds; and/or high pressure on the lee side and low on the windward side of major north-south oriented mountain ranges.
(vi) Angular momentum conservation leads to westerly acceleration of poleward moving air (deceleration of equatorward-moving air).
(vii) Water mass balance and transport are needed to explain precipitation maxima in the midlatitudes and tropics.

1.3. Deducing the atmospheric general circulation from thought experiments

To help understand the context of what follows in this book, it is useful to have an idea of what circulations one might expect. One can make quite a bit of progress by performing a series of “thought experiments.” The discussion that follows is qualitative, but it expands and makes visible the quantitative principles listed in the previous section. Some candidate, “imaginary” circulations are proposed until one qualitatively satisfies those seven properties discussed and listed above.

1.3.1. Idea #1: pure northward motion
While this motion transports sensible heat poleward, by bringing warm air from the tropics towards the poles, it is immediately rejected due to a lack of mass conservation. Furthermore, angular momentum is not balanced, unless the motion were heavily slowed by friction. Also, there is no mechanism to create the observed precipitation pattern.

1.3.2. Idea #2: Alternating meridional flows

In an effort to solve the mass balance problem of idea #1, one might propose meridional circulations that alternate between northward and southward motion (Figure 1.5a) through the depth of the troposphere (or perhaps even more of the atmosphere). The arrows shown in Figure 1.5a are trajectories. Heat is now mixed both directions, cooling tropical regions by importing cooler air from higher latitudes while warming polar regions with poleward flow of warmer air. Mass balance is satisfied by matching the masses of poleward and equatorward moving air.

Unfortunately, angular momentum balance might require slow meridional motion to occur mainly within a boundary layer. The motion above the boundary layer would be strongly ‘tilted’ in the zonal direction as shown in Figure 1.5a. In order to avoid a net torque upon the earth, the strong westerlies (from angular momentum conservation) in high latitudes would be balanced by strong easterlies in low latitudes. If there is vertical mixing of momentum (by convection, if not a large scale overturning) then these wind speeds are somewhat moderated. This candidate circulation would develop large upper level winds similar to the jet streams, however, it has no obvious mechanism by which to limit the jet to subtropical or midlatitude locations.

This candidate would have trouble explaining other important observed features. The midlatitude precipitation maximum (Figure 1.1b) might be explainable as follows. The poleward moving air in the boundary layer might be quite moist, as that air migrates poleward it cools (by net radiation, and by surface exchange) to a point where drizzle might occur. Explaining the midlatitude precipitation maxima as drizzle is a bit of a stretch! This scheme also has no obvious explanation for the tropical precipitation maximum. Finally, the predicted strong upper level equatorial easterlies are really not seen; the closest observed phenomenon is the Indian Ocean monsoon jet.

While this circulation is rejected, it does suggest trying alternating motions in the vertical and in the horizontal.

1.3.3. Idea #3: “one big Hadley cell”

A circulation having alternating motion in the vertical is the “Hadley cell”. The Hadley cell circulation could be similar at all longitudes and form a loop: with rising in the tropics, poleward motion in the upper troposphere, sinking at higher latitudes, and equatorward motion in
the lower troposphere. This Hadley cell, is depicted in Figure 1.5b. The schematic trajectories depict how poleward motion has westerly wind acceleration while equatorward motion in a boundary layer (subject to greater frictional dissipation) has less easterly acceleration. If mountain torques are absent, the wind might have no net torque upon the earth’s rotation by a careful adjustment of where the location lies of zero surface zonal wind. Mass is conserved by matching northward and southward mass fluxes. Heat transport is accomplished if the poleward flow has higher MSE than the equatorward flow. This is possible if the poleward flow is at tropospheric levels high enough so that the gravitational potential in MSE overcomes any high moisture content at low levels. Another reason for having a high troposphere poleward flow is to develop strong westerlies at jet stream level. The rising motion in the equatorial region provides an obvious explanation for the equatorial precipitation maximum (though details are missing like overcoming the vertical gradient of MSE within thunderstorm ‘hot towers’).

Some observed features are problematic. If this ‘big Hadley cell’ extends to polar regions, how would the jet streams occur where they do and not much further poleward? The zonal wind speeds reached would be unrealistically large in mid- and higher latitudes. Finally, there is no obvious explanation for the observed midlatitude precipitation maximum.

So while this circulation has promise in the tropics, it can’t explain all the general circulation.

1.3.4. Idea #4: barotropic eddies

The alternating meridional motions of Figure 1.5a lead to very large horizontal wind shears. At some point, large shears may become hydrodynamically unstable. A simple approximate determinate for (barotropic) instability is that the meridional gradient of absolute vorticity $\Omega$, change sign in the domain. In spherical geometry (Baines, 1976) $\Omega_{ay}$ becomes:

$$\Omega_{ay} = \frac{\partial \Omega_u}{r \partial \phi} = \beta - \frac{1}{r} \frac{\partial}{\partial \phi} \left( \frac{1}{r \cos(\phi)} \frac{\partial}{\partial \phi} (u \cos(\phi)) \right).$$

Where, $r$ is the Earth’s radius, $\phi$ is latitude, $\beta$ is the meridional gradient of the Coriolis parameter. ($\Omega_{ay}$ is the absolute vorticity gradient for a purely zonal flow, the contribution by meridional motion to $\Omega_{ay}$ has been ignored.) $\beta$ decreases while the horizontal shear of $u$ increases for increasing magnitude of latitude. However, since shear increases with latitude, the relative vorticity of such a zonal flow is negative. For the Earth, a zonal flow constructed from angular momentum conservation would have $\Omega_{ay} = 0$ at all latitudes! This result is a consequence of that flow also conserveing absolute vorticity. If some internal viscosity is added, then $\Omega_{ay} < 0$ at all latitudes and the flow is stable to barotropic eddies according to classical reasoning, even if large horizontal shear is still present. But these assumptions are too simplistic.
If topography is added, then waves of a scale similar to the topography may be forced and these would interact and grow, releasing some horizontal shear instability. Also, viscosity may lead to form drag instability (e.g. Frederiksen and Frederiksen, 1990). If moisture is included, perhaps through moist convective heating, then the condition that $\zeta_{\text{v}}$ change sign may be “neither a necessary nor a crucial condition for instability for a barotropic flow” (Mak, 1983; p. 2355). In short, it may be that barotropic eddies could form, by several different mechanisms.

This idea (fig. 1.5c) might appear to transport heat poleward by motion in both directions (warm air poleward, cold air equatorward). However, the horizontal winds around the (equivalent) barotropic eddies blow parallel to temperature contours. Also, the colder air would have smaller thickness (between isobaric surfaces) which in turn creates zonal geopotential gradients which in turn creates geostrophic winds. The consequence of those winds is unclear, perhaps the whole pattern migrates westward, perhaps it breaks down into multiple eddies, perhaps it leads to vertical axis tilts (baroclinic eddies).

The scheme can satisfy mass balance. A mean horizontal zonal velocity can be added to the flow in such a way as to create compensating amounts of easterly and westerly surface torque. Friction in the boundary layer of barotropic lows implies low level convergence with consequent rising motion that may lead to precipitation, though it is unclear how this gets focused to midlatitudes except by restricting eddies to higher latitudes where some threshold of horizontal shear is crossed.

This idea shows some promise since it recognizes that hydrodynamic instability can play a role. However, there are some critical flaws. First, how the necessary meridional heat transport occurs is unclear. In of themselves, barotropic eddies do not transport heat. However, one might argue that radiative gain in the tropical and loss in the polar regions leads to a net heat flux if the resident time in each region is long enough for sufficient radiative temperature changes. Second, the energetics of such lows would be a bit strange; while radiation-induced temperature changes drive meridional motion that creates zonal wind shear that is unstable to barotropic eddies, kinetic energy is lost by friction and conversion to eddy potential energy. (Potential energy increases because cold air at a low is rising while warm air at a high is sinking. Such vertical motion might arise from planetary boundary layer friction, approximated by an Ekman layer, for example.) Third, a mechanism for the tropical precipitation maximum is missing.

3.5. Idea #5: vertically tilted eddies:

The barotropic eddies thought experiment invoked instability of horizontally-sheared flow. The meridional temperature gradient requires vertical shear if thermal wind balance holds. As for horizontal shear, there is a classical necessary condition for instability of vertically sheared flows. Reasoning by analogy to Cartesian geometry, an appropriate governing equation
for stability of a zonal mean flow may be deduced from a potential vorticity conserving governing equation (e.g. Grotjahn and Castello, 2002). Using a streamfunction ($\Psi$) to represent the zonal mean flow, then

$$Q_{QGy} = \frac{\partial}{\partial \mu} V^2 \Psi + 4\mu^2 \frac{\partial}{\partial \mu} \left\{ \frac{\partial}{\partial \phi} \left( \frac{P}{\sigma \cdot \partial \phi} \right) \right\} + 2 + \text{Boundary Contributions.} \quad (1.3)$$

Where $\mu$ equals $\sin(\phi)$, $p$ is vertical coordinate pressure and static stability, $\sigma$ can vary in the vertical. Boundary Contributions refers to contributions to $Q_{QGy}$ from the bottom and top boundaries. In Cartesian geometry a change in sign of the mean flow potential vorticity gradient is necessary to have unstable waves (Charney and Stern, 1962); one might reasonably extend this classical analysis to the corresponding potential vorticity gradient in spherical geometry.

Simplifying the schematic circulation to be described by flow in two tropospheric layers, Baines and Frederiksen (1978) extend the Phillips (1954, Cartesian geometry) condition for stability to spherical geometry as:

$$[u_U] - [u_L] \geq \frac{62.25 \cos(\phi)}{r \Omega \sin^2(\phi)} \left( \theta_{UL} - \theta_{UL} \right) \quad (1.4)$$

Where the subscripts U and L refer to upper and lower layers, respectively. Also, $[u]$ is the zonal mean zonal flow (positive for eastward motion) while $\theta_s$ is the horizontal mean potential temperature. It is clear that this criterion decreases nonlinearly with latitude. Static stability helps control the latitude at which the instability criterion is exceeded in two ways. The greater the static stability then i) more vertical shear is needed for instability but ii) the vertical difference of DSE is larger. The heat transport is more efficient (smaller meridional motion is needed to transport the same amount of heat). But smaller meridional motion reduces the strength of the $U_{amc}$ (by more time for friction to act).

If zonal mean flow has vertical shear from meridional motion that conserves angular momentum (zonal wind equals $U_{amc}$) at upper levels, but is strongly damped by friction at low levels, then one might approximate the shear as some fraction of $U_{amc}$. For illustrative purposes, if the shear is 70% of $U_{amc}$ and the potential temperature difference from 750 to 250 hPa is 30K (roughly the observed value in the subtropics) then the instability criterion is met for latitudes higher than 32 degrees.

The amplifying waves feed upon the vertical shear and thereby provide a mechanism to control the unrealistically large zonal velocities created by a purely angular momentum conserving flow. The resultant waves will be tilted against the vertical shear. That can occur if the temperature and mass fields are offset, and if so, there will be a net heat flux across the mean wind, i.e. poleward. The needed poleward heat transport is accomplished by mixing cold air equatorward and warm air poleward. The motion around the eddies can satisfy mass balance and it is also possible to balance surface westerly and easterly torques. The poleward heat fluxes are
accompanied by vertical motions across isentropes; such that warm air rises to a higher potential
temperature surface. Since such warm air is expected to have moisture, the amplifying waves
will also have areas where condensation and precipitation are likely to occur, thus providing a
reasonable mechanism to explain the midlatitude precipitation maximum. Surface friction further
leads to low level convergence in the troughs of such waves, further amplifying the midlatitude
precipitation maximum.

This idea seems to work well in middle and high latitudes. The crudely estimated stability
criterion suggests that such baroclinic waves are not likely to be the main factor in the tropical
latitudes. That result, coupled with the ability of the Hadley cell to explain tropical features
suggests that the tilted eddy and Hadley cell ideas be applied in different latitude bands with
some sort of matching in between.

1.3.6. Idea #6: hybrid circulation

The preceding discussion suggests combining ideas #3 and #5. The challenge is to sort
out how the two quite different ideas can be matched. The resultant ‘hybrid’ circulation is more
difficult to draw than the other candidates since the middle and high latitude baroclinic eddies
are traveling and they change their size and other properties as they move. Hence, Figure 1.5d is
not so much showing trajectories, or even a snapshot of the motion, but is instead a ‘collage’ of
some aspects of the circulation.

Beginning with the Hadley circulation, sinking is predicted at the poleward end of the
Hadley cell. The analysis in idea #5 suggests that the midlatitude eddies become unstable and
dominant in the middle latitudes (say, 35 degrees latitude). This will define our conceptual
boundary between Hadley cell and baroclinic eddy dominated circulations; for compactness the
HCBE boundary. It is no coincidence that this HCBE boundary is not far from the observed
location in different seasons of the surface subtropical highs (in sea level pressure). The
subtropical highs fit nicely with the Hadley circulation in several ways: winds on the
equatorward side of those highs are consistent with the low level Hadley circulation, and from a
boundary layer friction argument, those highs would have surface divergence that can support
the sinking by the Hadley circulation. (Sinking happens if some mechanism, such as radiational
cooling allows air parcels to sink to lower potential temperatures).

The same boundary layer friction argument fosters rising motion for surface lows so
some mechanism is needed to separate the baroclinic lows from the edge of the Hadley cell.
Furthermore, the baroclinic eddies of higher latitudes extract energy from the zonal flow as well
as mix it vertically; both processes weaken the upper level flow. This argument suggests that the
HCBE boundary is also likely to be a region where the zonal flow is a maximum and that could
be a prototype for the subtropical jet stream.
The Hadley cell may extend a bit further poleward in one region due to such things as topographic deflection of the flow, interaction with higher latitude eddies, and land-sea contrasts. Angular momentum conservation implies a stronger jet (and presumably greater vertical shear) in that extended portion of the Hadley cell. Enhanced tropical convection (Sardeshmukh and Hoskins, 1988) in one region (such as a landmass) may also enhance the subtropical jet in a distant region. It is a bit simplistic, but one might argue that such a region would favor the baroclinic eddy growth compared with a region where the Hadley cell did not extend as far. If one allows this simplistic argument, then a remarkable chain of events unfolds. The eddies take time to grow, and they tend to move with that zonal flow when their amplitude is small. This stage of growth is termed “linear” in the sense that the eddy wind speeds are “small” compared to the mean flow. The linear stage is also the time when the stability criteria mentioned above have any relevance. While the eddy may grow by means that alter its shape (primarily the amount of upstream tilt) during this stage, the orientation and motion of the eddy along the wind flow maximum is sustained. When the eddy grows to such amplitude that its winds are comparable to the mean flow, then nonlinear effects become dominant. Nonlinearity leads to such things as: enhanced upper level amplitude, net zonal average eddy momentum fluxes, and less efficient poleward heat transport (though still large). Of most interest to the problem of matching circulations is that the baroclinic highs and lows tend to migrate to the sides of the jet-like flow having the matching sign of vorticity. The lows migrate to the poleward side of the jet, while the highs migrate to the equatorward side. The baroclinic highs can thereby reinforce the subtropical highs while the lows carry precipitation to the middle latitudes where precipitation has its secondary maxima.

The primary function of the circulation is to transport heat poleward and that also needs to be matched. Developing baroclinic lows and highs at the poleward edge of the Hadley circulation help to carry further the poleward heat transport by the Hadley cell. Baroclinic highs that migrate equatorward to reinforce the subtropical highs also assist the heat transport. Lagging behind (upstream) the surface baroclinic lows are upper level troughs which also assist in picking up the heat transport by the Hadley cell and carrying it to higher latitudes.

In summary, the principal balances and fundamental relationships were described for the large scale atmospheric circulation. Some implications of these properties were discussed. For example, a qualitative discussion illustrated how a mountain torque could balance a uniform westerly wind, though the resultant pressure distribution is opposite to that observed in subtropical latitudes. From this foundation a variety of candidate circulations were proposed and tested against the six principal balances and observed facts.

A hybrid circulation combining a lower latitude “Hadley” cell paired with middle and higher latitude baroclinic eddies can satisfy the criteria specified. This circulation looks a lot like
the observed circulation, which is no surprise since our proposed circulations were strongly influenced by foreknowledge of the actual circulation, in the empirical knowledge of the primary maxima and minima of precipitation.

There are several points to draw from the discussion. First, one is reminded that satisfying known balances places several constraints upon what circulation is workable. Second, that one can avoid any mention of a Ferrel cell circulation, instead placing appropriate emphasis upon baroclinic eddies dominating the midlatitude circulation. Third, that understanding the general circulation requires looking at fields other than the wind field. Finally, there are many details that have been glossed over in this discussion. Such “details” can be subtle and challenging and constitute important subjects of the general circulation.

1.4. Anticipating the ocean circulation

The success of the thought problems above emboldens one to anticipate some key properties of the ocean. Three highly idealized models can provide some insight to some properties of the ocean that are relevant to the atmospheric general circulation.

1.4.1 Western intensification by a wind-driven circulation

The first model considers how the winds blowing across the surface of the ocean might drive a circulation of that ocean. In driving a circulation, water of one thermodynamic property can be carried by ocean currents to another latitude and thereby set up variation along a latitude circle. Such variation across the subtropical ocean basins was hinted at when discussing the radiation fields shown in Figure 1.3. The wind-driven circulation model sufficient for our purposes here is a variation on the model described by Stommel (1948).

If the wind blows across the surface, what is the response by the ocean? If there were no continents, one might expect the wind to accelerate the flow at the top surface until some balance with friction is reached. Much of the ocean is 4 km deep and one might also assume that internal viscosity could cause the current to diminish with increasing depth, perhaps becoming nearly zero at the bottom. The loss of kinetic energy might create a slight viscous heating. It would take some time to establish the ocean circulation and in this thought problem one consequence would be ocean surface motion that mirrors the long term time mean atmospheric surface winds. A big problem with this analysis is that Earth’s continents get in the way. The component of the ocean currents normal to the boundary must vanish at those continental shores.

To proceed further, it is useful to consider the simplest model that can describe a realistic circulation in the ocean. This model describes a stationary ocean circulation resulting from wind forcing and contains just three terms:

i. a term indicative of the wind driving the water motion, proportional to the wind stress
ii. a term proportional to the Coriolis force since the scale of motion is very large

iii. a term to remove energy, namely friction

Friction appears because something is needed to extract energy at a rate that balances the input from the wind stress in order to maintain a stationary (steady in time). These three terms appear in the vorticity equation as follows:

\[
\vec{\beta} \cdot \nabla \times \vec{r} = \frac{\partial \tau_x}{r \cos \phi \lambda} - \frac{\partial \tau_y}{r \rho \phi} = \beta v + k_R \zeta = \beta \frac{\partial \psi}{\partial x} + k_R \nabla^2 \psi
\]  

(1.5)

\(\beta\) is the meridional derivative of the Coriolis parameter, \(v\) is the meridional wind component, \(\zeta\) is the vertical component of relative vorticity (in the ocean), and \(\psi\) is a corresponding streamfunction for the ocean current horizontal components.

\[u = \frac{\partial \psi}{\partial y} \quad \text{and} \quad v = \frac{\partial \psi}{\partial x}\]  

(1.6)

A simple Rayleigh friction formulation is used for friction, coefficient \(k_R\) is assumed to be a positive constant. The wind stress curl provides the wind forcing. The subscripts indicate the horizontal components; \(\tau = (\tau_x, \tau_y)\) is surface wind stress of the atmosphere upon the ocean.

Figure 1.5d suggests that large highs in sea level pressure (SLP) would be prominent in the subtropical latitudes. Such subtropical highs are observed, especially during summer months (and after the zonal average SLP values are removed). Figure 1.5d also suggests that the equatorial surface winds would have an easterly component while the midlatitude winds would have a westerly component. These properties are illustrated in Figure 1.6a.

The specified wind stress might be proportional to \(-\cos(a \phi)\) where \(a\) is chosen to make the function zero in the subtropics; the wind stress curl is thus proportional to \(-a \sin(a \phi)\) which has a maximum in the subtropics. If friction was not operating, then the Coriolis term needed to balance this wind stress must be negative, and since \(\beta>0\), the current must be equatorward. The flow could not be equatorward through the whole depth of the ocean from mass continuity. The Rayleigh friction assumed in (1.5) requires a positive relative vorticity (\(\zeta>0\)) in order to balance a poleward (\(v>0\)) current. The sign of the vorticity in turn indicates that the frictional balance can only occur on the left (west) side of the ocean basin. When the motion is poleward, \(\beta v\) and \(k \zeta\) have the same sign, so the friction term must cancel both. Hence, the vorticity must have large magnitude. Vorticity is a local measure of the rotation present in the flow and has contributions from the curvature and the horizontal shear in the flow. Along that straight western boundary, shear is the primary contributor and it must be large. Since the current is parallel to the streamfunction and inversely proportional to the spacing between isolines of streamfunction, a large viscous term \((k_R \zeta\) large\) requires a strong current and close spacing of the isolines, as seen in Figure 1.6b. In contrast, on the east side the vorticity is relatively small since the Coriolis and
wind stress terms have opposite signs. Hence the streamfunction isolines are widely spaced on the east side of Figure 1.6b.

The key feature to note is the western intensification of the current. The streamfunction contours are greatly compressed in the west leading to a western boundary current or WBC. More elaborate formulations of the friction also have a WBC in the solution. The WBC requires the variation of the Coriolis parameter (i.e. \( \beta \) nonzero), without a variable Coriolis force, the solutions do not have western intensification. More discussion of the WBCs in extensions of this type of wind-driven model is found in other books such as Pedlosky (1987) and Vallis (2006). Such currents are observed on the west side of every ocean basin. For the north Pacific, the WBC is labeled the Kuro Shio; for the north Atlantic the WBC is called the Gulf Stream. These WBCs are very important in several aspects of the general circulation. First, the WBCs transport heat poleward, removing some of that burden from the atmosphere. Second, large contrasts in thermodynamic properties (especially during winter) are set up between the warm waters of the WBC and adjacent cold continent. Those thermal contrasts have implications for the development of mid-latitude eddies.

Even though the wind forcing does not vary with longitude, the oceanic response is asymmetric. Generally higher heights are on the western side of the domain. From the similarity between a streamfunction and geostrophic current elevations, one can convert the streamfunction values into sea surface elevations. Observations find peak elevations near the west edge of the Pacific to be about a half meter higher than near the east edge (Levitus and Oort, 1977; assuming the 1000m depth is ‘flat’).

Finally, the WBC exists in a ‘boundary layer’ where friction is prominent. Vallis (2006) estimates the width of that boundary layer. If a wide spacing and thus weak flow on the east side of the basin is sufficient to balance the wind stress curl. A much stronger meridional flow on the west side of the basin means \( \beta \psi \) is much greater than \( k \cdot \vec{v} \times \vec{e} \) in that boundary layer. The primary balance is thus between \( \beta \frac{\partial \psi}{r \cos \phi \lambda} \) and \( k \nabla^2 \psi \). For a friction coefficient value of \( k=1/20 \) days\(^{-1} \) a length scale of \( k/\beta \) is \(~60\)km, approximately the width of the Gulf Stream. The other implication of this primary balance is a relative insensitivity of the WBC flow to the local atmospheric wind stress.

### 1.4.2 Ekman layers

The wind-driven model assumed a uniform distribution of atmospheric winds across the ocean basin. In reality, the subtropical highs often have finite extent and the winds on the east side of the ocean basin can have a meridional component. For a subtropical high limited to the
ocean basin, the winds on the east side would be generally equatorward. Winds parallel to the coast set up currents with a component normal to the coast and the second model illustrates that principle. In addition, this model will have wider application to atmospheric circulations at several scales.

A balance between friction, pressure gradient force, and Coriolis force in a boundary layer can be defined. Such a layer is commonly termed an ‘Ekman’ layer based on an original analysis by Ekman (1905). From the equations of motion (see Appendix C) in height coordinates, (C.26) and (C.27) with (C.22), the two component equations are:

\[ -fv + \frac{1}{\rho} \frac{\partial p}{\partial x} = F_1 = k_z \frac{\partial^2 u}{\partial z^2} - \frac{\partial \tau_x}{\partial z} \]  \hspace{1cm} (1.7)  
\[ fu + \frac{1}{\rho} \frac{\partial p}{\partial y} = F_2 = k_z \frac{\partial^2 v}{\partial z^2} - \frac{\partial \tau_y}{\partial z} \]  \hspace{1cm} (1.8)

where the velocities are those within the Ekman layer. A second order definition of friction is assumed. In the interior of the fluid and outside the Ekman layer geostrophic balance is assumed to hold (p is continuous across the interface between interior and start of the Ekman layer). Since geostrophic balance holds outside the Ekman layer, the left hand sides of (1.7) and (1.8) are a Coriolis forcing by the ageostrophic current.

In the case of an equatorward stress applied by the atmosphere on the east side of an ocean basin, the stress is applied by the winds at the surface \((z_T = 0)\) and is transmitted downwards and assumed to reach zero at the bottom of the ocean \((z = z_B)\). Multiplying (1.8) by an average density \(\rho_o\) and integrating from the top of the ocean to the bottom obtains, after rearranging, an estimate of the mass flux within the Ekman layer. In general:

\[ \int_{z_o}^{z_B} \rho_o \tilde{v}_d z = \frac{1}{f} \int_{z_o}^{z_B} k \frac{\partial \tilde{v}}{\partial z} d z = \frac{1}{f} \int \tilde{v} \times \{ \tilde{v} (Z_T) \} - \tilde{v} (Z_B) \]  \hspace{1cm} (1.9)

One may define a total wind within the Ekman layer as having two parts: \(\tilde{v} = \tilde{v}_E + \tilde{v}_g\) where \(\tilde{v}_g\) is the geostrophic flow vector based on the pressure field. Generally, the geostrophic flow is that outside the Ekman layer. More specifically, an estimate for the ocean would assume the geostrophic flow is zero at the bottom. The Ekman layer would be at the ‘top’ of the ocean, since the wind stress is applied there. Hence, for an equatorward wind stress, the mass transport is:

\[ \int_{z_o}^{z_B} \rho_o u dz = \frac{\tau_y (Z_T)}{f} \]  \hspace{1cm} (1.10)

For the northern hemisphere, \(f>0\) and equatorward motion means \(\tau_y<0\) so the mass flux \(M_E<0\) in the Ekman layer. Hence, the northerly winds drive the water westward, i.e. offshore as illustrated in Figure 1.7a. That water is replenished by water upwelled from deeper in the ocean. Similarly, for the Southern Hemisphere, \(f<0\) but equatorward motion has \(\tau_y>0\) so again \(M_E<0\) and water is
driven offshore. The upwelled water is much colder than the water it replaced and thus a thermodynamic contrast is created on the east side of each ocean basin relative to the adjacent continent especially during summer.

A highly useful, but slight digression is to consider Ekman balance within the atmosphere. For the atmosphere, the stress is that applied by the Earth’s surface upon the atmosphere, hence the stress has maximum magnitude at the surface and decreases upward. Hence, the integral of the stress in (1.9) will have a contribution with opposite sign. For example, the surface stress on the east side of a low has only a meridional component, assuming a geostrophic wind around that low. Because there is a geostrophic flow at the edge of the Ekman layer, unlike the oceanographic problem just done, the mass transport on the east side of a low has poleward (geostrophic) and east-west (ageostrophic) contributions. The latter interests us, so only the east-west component of the mass transport is considered. One can deduce the component normal to the geostrophic wind by using the form of (1.8) with diffusion of $v$. The east-west mass transport on the east side of a low becomes:

$$M_{E_{1}} = \int_{z_{a}}^{z_{b}} \rho_{o} u_{E_{1}} dz = \int_{z_{a}}^{z_{b}} \rho_{o} k_{Z} \frac{\partial v}{\partial z} dz = \rho_{o} k_{Z} \left( \frac{\partial v(Z_{T})}{\partial z} - \frac{\partial v(Z_{B})}{\partial z} \right) = -\rho_{o} k_{Z} \frac{\partial v(Z_{B})}{\partial z}$$

In the Northern Hemisphere the $f>0$, the southerly flow is positive ($v>0$) so the stress acting to slow down the atmosphere to have zero velocity at $z = Z_{B} = 0$. The meridional component increases most rapidly with height near the bottom as that component asymptotes to the geostrophic value at the top of the Ekman layer as illustrated by the inset in Figure 1.7b. The vertical shear of $v$ is positive as are the average density $\rho_{o}$ and the diffusion coefficient $k_{Z}$. Hence, $M_{E}<0$. On the west side of the low the geostrophic flow is in the opposite direction ($v<0$) so the stress and the diffusion have opposite sign to before and so $M_{E}>0$. A similar analysis is found for meridional mass flux components induced by the zonal stresses on the north and south sides of the low. The net effect is a convergence of the ageostrophic winds within the atmospheric Ekman layer. From mass continuity (C.29) there must be upward motion out of the Ekman layer for a low. In the Southern hemisphere the sign of $f$ reverses but so does the direction of the geostrophic motion and the result is again a convergence within the Ekman layer for a low and consequent rising motion at the boundary layer top. For a high pressure the flow is reversed but the sign of $f$ is not and so the mass transport is outward from the surface high. Consequently, the Ekman layer for a high has sinking at the boundary layer top. The vertical motion out of the Ekman layer, ‘pumping’ above a surface low and the ‘suction’ above a high, can be expressed as proportional to the geostrophic relative vorticity. These results are an important factor in understanding properties of middle latitude eddies near the surface. The Ekman ‘pumping’ out of the boundary layer for a low pressure system can encourage cloud cover and precipitation.
1.4.3 A buoyancy-driven circulation thought experiment

In the tropical regions there is net heating of the ocean by radiative processes. One might expect heating at the top to be stabilizing the ocean since the warming (slightly) decreases the water density at the surface. However, the situation is a bit more complicated than that conclusion might imply. In polar regions, the ocean is cooled in the net by radiation and surface heat fluxes, a process that is destabilizing by (slightly) increasing the density of the water at the surface which may begin to sink. Since the heat input and output occur along a horizontal surface, this circulation is sometimes called ‘horizontal convection’ (Stern, 1975). Hughes and Griffiths (2008) review the topic. Sea water contains salt and salt content contributes to the density of sea water as well. Since temperature (‘thermal’) and saltiness (‘haline’) both play important roles in this circulation it is sometimes called the ‘thermohaline’ motion. Clearly, this density, or buoyancy driven circulation is not directly captured by either the wind-driven or the Ekman models discussed above since neither temperature nor density was driving the circulations created by those models. So a third type of model, driven by buoyancy, gives some insight into the vertical structure of the ocean.

A variety of ‘simple’ models exist to describe the buoyancy-driven circulation. Early efforts include the Wellander (1959) and Robinson and Stommel (1959) models. Another approach is to divide the ocean into ‘boxes’ such as a tropical box, a high latitude box, and an abyssal box for the waters beneath those boxes (Stommel, 1961). Rooth (1982) also includes another high latitude box in the opposite hemisphere to explore equatorially asymmetric circulations. These simple models are nonetheless mathematically complex enough to discourage complete development here. Instead a thought problem based on the thermodynamic equation is sufficient for our discussion.

The purpose of this discussion is to estimate what portion of the ocean interacts with the atmosphere on a seasonal time scale. For this purpose it is sufficient to construct a zonally-averaged analysis based on the primary balances of a temperature equation. While a buoyancy parameter (proportional to the acceleration of gravity and the density departures from an average density) is often used in oceanographic models, a simple temperature equation is adequate for this discussion. Variation of saltiness will be ignored for this presentation. Dropping the time tendency for a conservation equation for temperature with two diabatic processes: second order diffusive mixing and diabatic heating $Q$ yields:
where $k_z$ is a vertical diffusion coefficient. The terms are labeled to facilitate comparison with Figure 1.8, which shows the basic geometry of this thought problem.

The conceptual model has net radiational heating in the tropics and net cooling at high latitudes, both properties are consistent with Figure 1.3. Observations as well as simple models consistently show a circulation like that in the figure. Intuitively, the circulation can be understood as resulting from cold water sinking at high latitudes that draws warm surface water from low latitudes which in turn is replaced by upwelled water. The upwelling is consistent, if not explained by surface Ekman mass transport away from the equator for equatorial easterly winds like those in Figure 1.5d.

This is not the only explanation for the circulation, but it is sufficient for the purpose here. In the tropical region there is net heating that warms the ocean surface layer. In higher latitudes there is net cooling. The result is a temperature gradient, $\frac{\partial T}{\partial \phi} < 0$. One might expect the ocean to develop a poleward surface motion ($v > 0$) to reduce this temperature gradient, for example by postulating a pressure gradient caused by the slight expansion of the warm water relative to the cold water that initiates the motion. At some subtropical or middle latitude one might expect the net heating $Q$ to be zero on a time average. Similarly, one might expect the vertical motion of the thermohaline circulation to be zero as well. Therefore, the balance remaining from (1.12) is between the meridional temperature advection and the vertical diffusion.

Based on observations, it is reasonable to assume a temperature distribution with depth that asymptotes to a constant value in the abyss. For example, one might assume temperature can be expressed as a separable function with form that is exponential in the vertical and linear in latitude: $T = T_b + G \phi \exp(az)$. $G$ is a constant with units K rad$^{-1}$ and it is understood that latitude $\phi$ is in radians. When substituted into (1.13) one finds:

$$a = \left( \frac{v}{k_z r \phi} \right)^{\frac{1}{2}}$$

One might assume $v = 2 \text{ cm/s}$ (a basin-wide magnitude, not a value typical of a WBC) and an eddy diffusivity coefficient value of $k_z = 10^{-5} \text{ m}^2/\text{s}$, both from Vallis (2006). Assuming that $Q \approx 0$ at 40 degrees latitude, one obtains $a = 0.0212 \text{ m}^{-1}$. Hence the e-folding depth is about 200 m.
Compared with the 4-5 km depth of most of the ocean, the layer satisfying this balance is quite shallow. An important result is thus that only a comparatively shallow layer of the ocean interacts with the atmosphere on seasonal time scales.

In the tropics and high latitudes where the net heating and cooling are strong, one might expect vertical motion to occur. Beneath the surface waters where the heating and cooling takes place, there must be a vertical motion that communicates with the deeper waters. Diabatic heating (Q) will be small at those levels and one assumes the meridional advection (HA) is also small there. The terms remaining assume a balance between vertical advection and diffusion. From terms VA and D, Vallis (2006) estimates an e-folding depth of around 100 m, so again, the temperature rapidly asymptotes to the abyssal value.

The observed temperature variation in the ocean is more complex. Winds blowing across the surface cause mechanical mixing of the water. A typical depth for this wind-mixed layer might be 50 m. The thermohaline circulation balances also lead to a rapid change of temperature with depth beneath the wind mixed layer. Where there are western boundary currents, the depth of water moving poleward is deeper. Simple models of the thermohaline circulation also find a current flowing the opposite direction beneath the WBC. Figure 1.9 illustrates a complex pattern of currents that are thought to comprise a world ocean thermohaline circulation.

In summary, this chapter established a context for viewing and understanding the general properties of the atmosphere’s general circulation. Basic physical constraints and some key empirical information were presented. Tropical circulations differ from middle and higher latitude motions with subtropical circulations responding to both. Some properties of the ocean were discussed. Western boundary currents, upwelling, and the mixed layer depth of the ocean were also deduced and they impact understanding the atmospheric circulation. The imaginary circulations deduced in this chapter foreshadow much of the material covered in the remainder of the book both from the perspective of what the circulation looks like (including the interconnectedness of many variables) and from what mechanisms drive what is seen.
Problems

   a. Equatorial Indian Ocean. From Figure 1.3a assume 310 W/m^2 of solar radiation is absorbed (clouds and scattering ignored) while from Figure 1.3b 260 W/m^2 is emitted. Assume that winds mix this net heating into a column of sea water 50m tall. Find the rate of temperature change in K/s. How many degrees would the temperature of the water change if this net heating continued for 100 days? Assume the specific heat for water is \( C_w = 4180 \, \text{J K}^{-1} \, \text{kg}^{-1} \) and the density of water \( \rho_w = 1025 \, \text{kg/m}^3 \). Hint: 1 W = 1 J/s = 1 kg m^2 s^{-3}

   b. 60N Atlantic Ocean. From Figure 1.3a assume 160 W/m^2 of solar radiation is absorbed (clouds and scattering ignored) while from Figure 1.3b, 230 W/m^2 is emitted. Assume that winds mix this net cooling into a column of sea water 70m tall. How many degrees would the temperature of the water change if this net cooling continued for 30 days? Assume the specific heat for this cooler water is \( C_w = 4200 \, \text{J K}^{-1} \, \text{kg}^{-1} \) and the density of water \( \rho_w = 1025 \, \text{kg/m}^3 \).

   c. 60N Western Canada. From Figure 1.3a assume 140 W/m^2 of solar radiation is absorbed (clouds and scattering ignored) while from Figure 1.3b, 210 W/m^2 is emitted. Assume that conduction mixes this net cooling into a column of soil 6m tall. How many degrees would the temperature of the soil change if this net cooling continued for 30 days? Assume the heat capacity of the soil is: \( C_s = 1000 \, \text{J K}^{-1} \, \text{kg}^{-1} \). Assume the soil has a density of 1500 kg/m^3.

   d. The net radiative cooling is the same in parts b and c. Comment on the difference in the rate of temperature change (in K/s) at the two locations.

2. Heat transport to sustain specified heating and cooling. Assume that the cooling rate over the North Atlantic is sustained by a poleward heat transport across 35N, a width assumed to be 60 degrees of longitude. Assume the transport takes place within a 100m deep layer at the top of the ocean. Assume a net solar absorption of 50 W/m^2 over an area from the equator to 20N that is 42 degrees of longitude wide.

   a. Find the surface area of ocean between the equator and 20N, over 42 degrees of longitude. (Hint: appendix C has the general formula for the area.)
b. Find the total number of watts of energy gained by the tropical region of part a.

c. Find the area of an imaginary ‘gateway’ at 35 N that is 100m deep and 60 degrees of longitude wide.

d. Find the meridional velocity, \( v \) through the ‘gateway’ in part c such that the energy gained in part b passes through that gateway at an equivalent rate. Define \( dT = 18 \text{ K} \) as the temperature difference between the surface layer and abyssal temperatures. (To conserve mass, return flow is assumed to occur in an abyssal layer 18 K colder than in this surface layer.) Hint: a heat flux per unit area is \( \nu C_w \rho_w dT \) where the density of seawater is 1024 kg/m\(^3\) and heat capacity of water is given in Appendix B.

e. How does the meridional velocity obtained this way compare with the 2 cm/s meridional velocity assumed in section 1.4.3 of this chapter?

3. \textit{Ekman balance winds and implied vertical motion}. Using (1.7) and (1.8) for guidance, Ekman balance from the two components of horizontal momentum equations can be written using the geostrophic wind components in stead of pressure. To simplify the appearance without loosing the generality, assume that a nondimensionalization has been made such that the equations can be written:

\[
\frac{1}{2} \frac{\partial^2 u}{\partial z^2} + v = v_g \\
\frac{1}{2} \frac{\partial^2 v}{\partial z^2} - u = -u_g
\]

(1.15) (1.16)

a. Find the general formulas for \( u \) and \( v \) for this Ekman balance assuming the values asymptote to the geostrophic values as height approaches infinity. Hints: combine the equations to obtain a fourth order equation in \( u \), for example. The solution will have the general form:

\[
\{ u = u_g + A_u \exp(-z) \{ \cos z + A_g \sin z \} + B_u \exp(z) \{ \cos z + B_g \sin z \} \}
\]

(1.17)

From the asymptotic boundary condition, \( B_u = 0 \). Find the value of \( A_u \) from the bottom boundary condition that the total wind must be zero. The solution for \( v \) is similar.

b. Calculate the divergence from the winds \( u \) and \( v \) obtained in part a. From a scaled nondimensional continuity equation, assume that:

\[
\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = -\frac{\partial w}{\partial z}
\]

(1.18)
Noting that geostrophic winds satisfy:

\[ \vec{v} = (u_g, v_g) = \left( -\frac{\partial P}{\partial y}, \frac{\partial P}{\partial x} \right) \]  

(1.19)

where \( P \) is a nondimensional pressure, the divergence simplifies considerably. Noting that the vertical component of vorticity equals the curl of the velocity \( \zeta_g = \frac{\partial v_g}{\partial x} - \frac{\partial u_g}{\partial y} \) in this nondimensional derivation, express the divergence in terms of vorticity.

c. Integrate the divergence found in part b in the vertical subject to \( w=0 \) at the bottom boundary. The result should express vertical velocity at some level \( z \neq 0 \) in terms of the relative vorticity, \( \zeta_g \). Discuss the sign of the vorticity and \( w \) for a circular low pressure system.

d. Assume a ‘square wave’ low pressure center can be defined from a pressure field \( p = G \sin(x) \sin(y) \) in the Cartesian domain \( 0 \leq x \leq \pi \) and \( 0 \leq y \leq \pi \) where \( G<0 \). Derive the formula for the geostrophic horizontal winds based on \( P \). Express the velocities found in part a using these formulas for the geostrophic winds. Create plots of: \( P, u_g, v_g, \zeta_g \), and \( w \) for level \( z=1 \) for the indicated domain and \( G= -1 \).

4. Equatorial Ekman layers. In these problems use the general analysis of the mass transport in an Ekman layer as in § 1.4.2.

a. Using Figure 1.5d as a guide, assume that the atmospheric surface wind stress is purely zonal in both the Northern and Southern Hemispheres.

(i) Derive the formula and deduce the direction of the integrated mass flux in the ocean at 5N and at 5S. Imposed a \( \tau_x \) surface stress in the starting equation (1.7).

(ii) Given your answer for each hemisphere, deduce the likely mass flux at the equator.

(iii) Describe whether the Ekman mass transports create or do not create vertical motion along the equator.

(iv) Discuss whether this motion opposes or reinforces the vertical motion deduced for the tropics in a thermohaline circulation (§1.4.3)?

(v) Explain how this stress creates a zonal gradient in the height of the ocean surface and where this enters into the starting equation (1.7).

b. Assume that a purely easterly flow similar to that shown in Figure 1.5d occurs at the top of a boundary layer at 10N (and similarly at 10S) and is therefore a ‘geostrophic’ flow.
(i) Deduce the atmospheric mass transport direction in each hemisphere with reasoning analogous to (1.11) except for a zonal frictional diffusion.
(ii) From the indicated winds, explain whether the sea level pressure is lower, the same, or higher at the equator than at 10 N.
(iii) Explain whether the implied vertical motion reinforcing or opposing the vertical motion expected near the equator for a Hadley cell?

5. *Barotropic instability*. In these problems the instability of the horizontal shear is examined.

a. Determine whether solid body rotation, given by _______. Discuss the solution.

b.
Figure. 1.1 (a) Annual average total precipitation, $P_w$ and evaporation $E_w$ during 2001-2009 in mm/day. These ERA-interim data are a model product used in the reanalysis based on the 12 hour forecast rate. (b) Annual average net downward shortwave (solar) radiation and upward longwave radiation (terrestrial) estimated for the top of the atmosphere during the 1979-2009 period. These data are from the NCEP/DOE AMIP-II reanalysis datasets.
Figure 1.2  (a) Topographic elevations where the shading generally becomes lighter for higher elevations. Map based upon 30 arc second data provided by the EROS Data Center of the US Geological Survey. (b) Percentage of land (solid line) and ocean (dashed line) areas within each five-degree-wide latitudinal belt. Land fraction data courtesy of Dr. Steve Warren.
Figure 1.3 Global radiative balance for the Earth. (a) Time average incoming, downward, solar, absorbed radiation estimated at the top of the atmosphere. (b) Time average upwelling, longwave radiation from the Earth. ECMWF Reanalysis 1 (ERA-15) monthly mean data from January 1979 through December 1993. Contour interval 20 W/m².
Figure 1.4  (a) Geometry of unit areas on the Earth relative to solar radiation. More radiant solar energy reaches an area at the equator than a corresponding area at 40N. (b) Rough illustration of how temperature varies with latitude in the troposphere and stratosphere. The tropopause, indicated by a dashed line, separates the troposphere and stratosphere. Warmth and coldness indicated is relative to other temperatures at that elevation. (c) The spacing between pressure surfaces is larger for warmer temperatures and the effect is additive in the vertical so that the slope of a constant pressure surface $p_2$ is greater than for a higher pressure surface $p_1$. (d) Schematic diagram showing a possible variation of MSE and DSE with height. Dashed arrows indicate direction and relative magnitudes of the meridional flux of DSE needed for mass conservation and net poleward heat transport.
Figure 1.5 Schematic diagrams of selected candidate circulations generally showing proposed air parcel trajectories. Solid lines show upper level flow, dashed lines show low level flow. Vertical dimension shown on each horizon is greatly exaggerated. Circles enclosing a dot suggest upward motion; circled X for downward. (a) ‘idea #2’ alternating poleward and equatorward flows. (b) ‘idea #3’, ‘one big Hadley cell’ showing acceleration caused by meridional motion (slowed by friction for low level flow. (c) ‘idea #4’ barotropic eddies that form spontaneously when the strong shears of the ‘one big Hadley cell’ become too large. (d) ‘idea #6’ a ‘hybrid’ circulation that combines baroclinic eddies in high latitudes with a ‘Hadley’ circulation in low latitudes. The J indicates a subtropical jet created between the ‘Hadley’ cell and a midlatitude eddy. Also included are equatorward, sinking, cold air (left horizon) and poleward, rising, warm air (right horizon) both in middle latitudes.
Figure 1.6 Schematic diagrams of wind forcing and oceanic response for a simple balance between the curl of the wind stress ($\tau$), Coriolis term, and Rayleigh friction from the vorticity equation. (a) Surface winds (dashed arrows) associated with a large atmospheric SLP high pressure (H), zonal component of wind stress (solid arrows), curl of the wind stress (dotted arrows) for a hypothetical ocean basin from the equator to 60N. (b) Schematic contours of stream function, $\Psi$ with the indicated balances between the three terms. Arrows on the contours show the clockwise ocean circulation. To have strong enough friction to balance the Coriolis and wind forcing terms, the current must be concentrated in a western boundary current.
Figure 1.7  Schematic diagrams of mass transport $M_E$ in an Ekman layer for a specified stress $\tau$. (a) northerly flow over eastern side of a Northern Hemisphere ocean basin results in a net offshore mass flux in the ocean. (b) geostrophic flow around a low pressure center results in a mass flux in the atmospheric Ekman boundary layer directed towards the low center. The southerly flow, $v>0$ shown is at the top edge of the Ekman layer while $\tau_y$ is the stress upon the atmosphere by the ground. Inset is a schematic profile of the meridional wind component which asymptotes to the geostrophic value as elevation (z) increases. Only the zonal component of the mass flux is shown.
Figure 1.8  Schematic diagram of a highly idealized buoyancy-driven circulation deduced from four temperature equation terms: horizontal advection (HA) and vertical advection (VA) of temperature, net diabatic heating or cooling (Q), and second order diffusion (D). Thin contours are of a streamfunction in the meridional plane. The temperature variation with elevation is assumed to be an exponential proportional to exp(az). Large cooling (Q<0) at high latitudes and large heating at low latitudes drive a circulation that is shallow (large a) in the upper, poleward moving portion.
Figure 1.9 A schematic depiction of the thermohaline circulation for the world ocean. Darker grey lines indicate upper ocean currents while light grey lines indicate deep currents. Arrows show the direction of the flow. Sinking is concentrated in high latitudes and especially in key regions (ovals) where particularly dense water is formed in the north Atlantic and near Antarctica. Upwelling is shown in the Indian Ocean, other upwelling zones are not depicted. Shaded areas indicate regions of comparatively higher and lower salinity in the upper ocean. Higher salinity occurs over most of the subtropics (exceptions being off the Central America west coast and Southeast Asia). Those exceptions and the higher latitudes have lower salinity. Not all major currents are depicted. For example, the Kuro Shio WBC is missing. Reproduced from Rahmstorf (2002) and printed in black and white by permission of the Nature Publishing Group.