Chapter 19

Atmospheric General Circulation

ABSTRACT
This chapter briefly reviews the principal factors controlling the climate on our planet. We first summarize the global heat budget and then describe the large-scale convective cells and review the major wind systems. The chapter ends with weather forecasting and the particular challenge of simulating cloud dynamics. Ingredients of modern operational weather-forecast models are described.

19.1 CLIMATE VERSUS WEATHER

Climate is to be distinguished from *weather*. Whereas *weather* includes the detailed behavior of the atmosphere on a timescale of a day to a week, *climate* represents the prevailing or average weather conditions over a period of years. In other words, the climate of the earth can be regarded as the basic state of the atmosphere, subject to variations over years, centuries, millennia, and beyond, whereas the weather corresponds to its incessant and short-lived instabilities. The engine of climate is a global convection carrying heat from the warmer tropical belt to the much colder polar regions, and its primary manifestation is the distribution of prevailing winds over the globe.

Numerous books have been written on climate and weather dynamics. For texts presenting materials at a slightly deeper level than presented here, the reader is referred to the classic book by Gill (1982) and the highly readable textbook by Marshall and Plumb (2008), each written from the perspective of geophysical fluid dynamics.

19.2 PLANETARY HEAT BUDGET

Because the long-term gradual cooling of the earth's core contributes insignificantly to the heat input near the surface, the incoming solar radiation can be considered as the sole source of heat. From its hot surface ($T \simeq 5750$ K), the sun emits most of its energy in short wavelengths (200–4000 nm; 1 nm $= 10^{-9}$ m), of which about 40% is in the visible range (400–670 nm). According to the

Introduction to Geophysical Fluid Dynamics, Volume 101, ISSN 0074-6142, DOI: 10.1016/B978-0-12-088759-0.00019-5

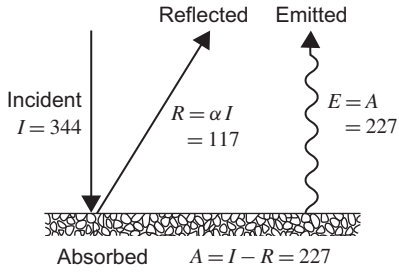

FIGURE 19.1 Simplest possible model of the earth's budget. Straight lines indicate short-wave radiation, whereas the wavy line represents long-wave radiation. (Fluxes are in watts per square meter.) Under this scenario, which does not account for the atmosphere, the earth's average temperature would be a freezing −21°C.

Stefan–Boltzmann law, a so-called *black body* (a perfect emitter and absorber of radiation) emits a radiative flux F depending on its temperature

$$F = \sigma T^4, \tag{19.1}$$

where σ is a constant equal to $5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^{-4}}$, and T is the absolute temperature. Idealizing the sun to a black body, we obtain $F_{\mathrm{sun}} = 6.2 \times 10^7\ \mathrm{W/m^2}$ as the outgoing energy flux from the sun's surface. Given the size of the sun, the sun–earth distance, and the earth's area exposed to the sun, the earth receives only a minute fraction of the sun's output: 1376 $\mathrm{W/m^2}$. Averaged over the entire earth's surface (equal to four times the projected area facing the sun), this incident flux amounts to $I = 344\ \mathrm{W/m^2}$.

Let us, at this point, first discard the thickness of the atmosphere and idealize the earth's land and sea surface plus atmosphere as a thin sheet insulated from below. Of the incident radiation, a fraction is reflected out to space by snow, ice, some types of clouds, and everything else that is bright. With α as the reflection coefficient, called the *albedo* ($\alpha \simeq 0.34$), the amount of radiation reflected is $R = \alpha I = 117\ \mathrm{W/m^2}$. The difference is the amount absorbed by the earth's surface: $A = I - R = (1-\alpha)I = 227\ \mathrm{W/m^2}$ (Fig. 19.1). Because the earth is in overall thermal equilibrium[1] (its temperature is not constantly rising), its outgoing radiation matches absorption, and the earth emits a radiative flux E equal to A. This outgoing radiation is in the form of longer wavelengths than the incoming solar radiation and is termed long-wave radiation. Assuming as for the sun that the earth behaves as a black body and using the preceding values, we state

$$\sigma T^4 = E = 227\ \mathrm{W/m^2} \tag{19.2}$$

and deduce a mean temperature for the earth to be $T = 251\ \mathrm{K} = -21°\mathrm{C}$. This value is obviously much less than the average temperature of the earth as we know it (about 15°C). The failure of this simple model resides in the neglect

[1] Some of the heat received by the sun is transformed into mechanical (wind) and chemical (photosynthesis) energies, but these eventually dissipate and turn back into heat.

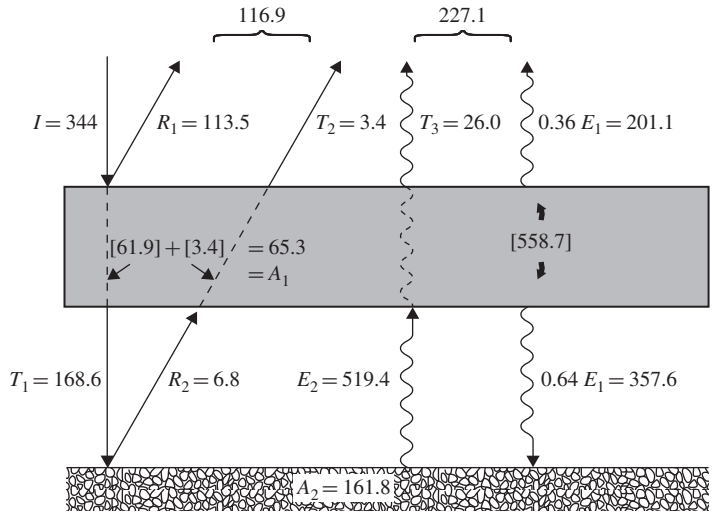

FIGURE 19.2 A second model of the earth's budget, which distinguishes the atmospheric layer from the earth's surface. All flux values are in watts per square meter. Under this scenario, the earth's average temperature would be a very warm +36°C. Here the greenhouse effect (flux loop between the earth's surface and the atmosphere) is present and exaggerated. Note how this effect causes the long-wave radiative fluxes from the earth and atmosphere to exceed the incident short-wave radiative flux from the sun.

of the atmospheric layer. The preceding value is more representative of the temperature at the top of the atmosphere than at ground level.

As a next step, we distinguish the atmosphere from the earth's surface (Fig. 19.2). The incident short-wave radiation from the sun is unchanged ($I = 344$ W/m^2); of it, the fraction α_1 (= 0.33) is reflected back to space, primarily by clouds and secondarily by particulate matter ($R_1 = \alpha_1 I = 113.5$ W/m^2), the fraction β_1 (= 0.49) is transmitted to the earth's surface ($T_1 = \beta_1 I = 168.6$ W/m^2), and the rest is absorbed by the atmosphere. The earth's surface (snow, ice, and so on) reflects a fraction α_2 (= 0.04) of what it receives ($R_2 = \alpha_2 T_1 = 6.8$ W/m^2) and absorbs the rest ($A_2 = T_1 - R_2 = 161.8$ W/m^2). Of the portion R_2 reflected from the earth's surface, the fraction β_1 is transmitted through the atmosphere and out to space ($T_2 = \beta_1 R_2 = 3.4$ W/m^2), whereas the rest is absorbed by the atmosphere. Thus the atmosphere absorbs short-wave radiation directly from the sun ($I - R_1 - T_1$) and indirectly from the earth below ($R_2 - T_2$), and the net is

$$\begin{aligned} A_1 &= (I - R_1 - T_1) + (R_2 - T_2) \\ &= [1 - \alpha_1 - \beta_1 + \beta_1\alpha_2(1 - \beta_1)]I \\ &= 65.3 \text{ W/m}^2. \end{aligned} \tag{19.3}$$

Then both the atmosphere and the earth's surface emit long-wave radiation, in amounts equal to their total intakes of both short- and long-wave radiation. If the atmosphere emits a flux E_1, some of it goes upward into space and the rest goes downward to the earth. Because the top of the atmosphere, where the outgoing radiation originates, is colder than its lower layers, where the earthbound radiation originates, the two amounts are not equal; a representative split is 36% to space and 64% to the earth. Thus, the earth receives $0.64E_1$ of long-wave radiation from the atmosphere in addition to the amount A_2 received in short waves, and its emission E_2 must equal their sum:

$$E_2 = A_2 + 0.64E_1. \tag{19.4}$$

At this point, we still do not know either E_1 and E_2, but we can already conclude that the presence of atmospheric radiation toward the earth's surface establishes a loop, whereby the earth's surface emits some radiation, a portion of which returns to the earth. As a consequence, the earth's surface must emit more radiation in the presence of an atmosphere than in its absence and (according to the Stefan–Boltzmann law) must be correspondingly warmer. This is the *greenhouse effect*, so called because of its believed similarity with the trapping of long-wave infrared radiation by the glass panes of a greenhouse.[2]

Of the amount E_2 radiated by the earth's surface and entering the atmosphere, a fraction β_2 (= 0.05) is transmitted and lost to space ($T_3 = \beta_2 E_2$), with the remainder being absorbed by the atmosphere ($E_2 - T_3$). If the atmosphere absorbs the amounts A_1 and $E_2 - T_3$ in short- and long-wave radiations, respectively, its total emission must be equal to their sum; that is,

$$\begin{aligned} E_1 &= A_1 + E_2 - T_3 \\ &= A_1 + (1 - \beta_2)E_2. \end{aligned} \tag{19.5}$$

From Eqs. (19.4) and (19.5), we can obtain the emission fluxes E_1 and E_2 to find $E_1 = 558.7$ W/m^2 and $E_2 = 519.4$ W/m^2. Note that both are higher than the incident flux $I = 344$ W/m^2. Then, using the Stefan–Boltzmann law (19.1), we estimate the mean temperature of the earth's surface to be $T = (519.4/\sigma)^{1/4} = 309$ K $= 36^\circ$C. This temperature is higher than the first estimate, thanks to the capping effect of the atmosphere (greenhouse effect) but is unrealistically high.

In reality, the warming influence of the greenhouse effect is partially short-circuited by the hydrological cycle. As water evaporates over the ocean and land, latent heat is extracted from the earth's surface. (Latent heat is the heat required to change the phase of a substance, here to transform liquid water into water vapor. The latent heat of water is 2.5×10^6 J/kg.) This water vapor rises through the atmosphere, where it condenses in clouds before returning

[2] In fact this belief is incorrect because heat retention in a physical greenhouse is chiefly because of the elimination of convection by the glass barrier. Greenhouses covered with polyethylene plastic are about as effective as glass-covered greenhouses.

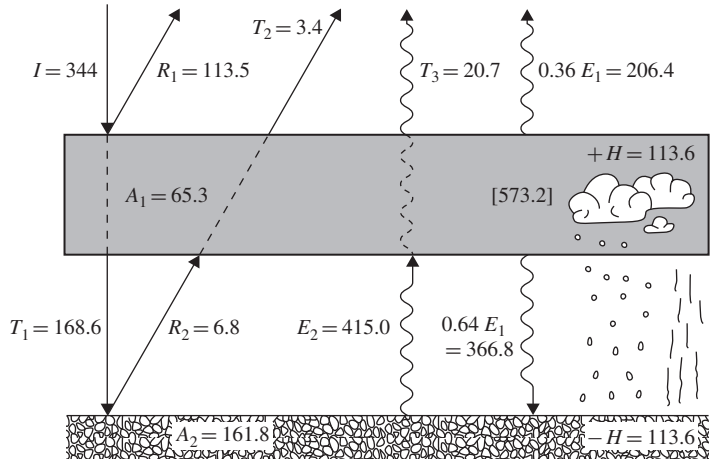

FIGURE 19.3 A third model of the earth's budget, with atmosphere and hydrological cycle. All flux values are in watts per square meter. This scenario includes the greenhouse effect tempered by the hydrological cycle, resulting in a realistic average temperature at the earth's surface of $+19^\circ$C.

to the earth's surface as rain (liquid phase). Thus, the latent heat extracted from the earth's surface is released in the atmosphere, causing a net heat flux from the earth to the atmosphere that is not in the form of radiation. To this latent-heat flux is added a convective heat transfer. With an estimated total non-radiative heat flux $H = 113.6$ W/m^2, the earth and atmospheric balances, (19.4) and (19.5), must be amended as (Fig. 19.3):

$$E_2 = A_2 + 0.64\,E_1 - H, \tag{19.6a}$$

$$E_1 = A_1 + E_2 - T_3 + H, \tag{19.6b}$$

yielding $E_1 = 573.2$ W/m^2 and $E_2 = 415.0$ W/m^2. From the radiation law, we deduce a corrected estimate of the mean temperature at the earth's surface: $T = (415.0/\sigma)^{1/4} = 292$ K $= 19^\circ$C. This third estimate is in good agreement with the seasonally and globally averaged temperature on the earth's surface. All in all, we conclude that the greenhouse effect resulting from the presence of the atmosphere (especially with regard to its near opacity to long-wave radiation) raises the temperature of the earth's surface and that the impact of this effect is partially canceled by the hydrological cycle.

19.3 DIRECT AND INDIRECT CONVECTIVE CELLS

The preceding considerations exposed the globally averaged heat budget, glossing over all spatial variations. However, that the tropical regions of the globe receive a disproportionate amount of solar radiation, because of their better exposure, is not to be overlooked. The earth receives considerably more heat

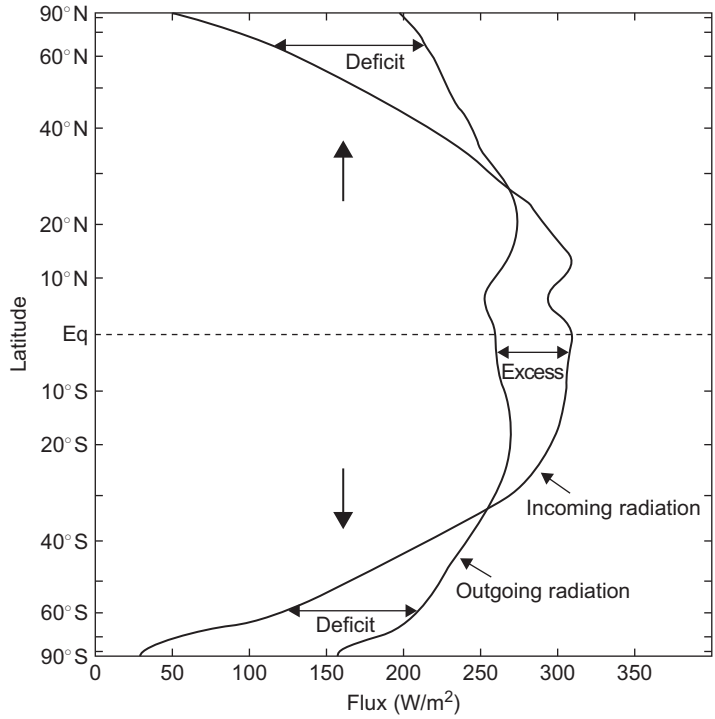

FIGURE 19.4 Averaged radiation flux by latitude, calculated from satellite data over the period 1974–1978. The latitude scale simulates the amount of surface area between latitude bands. Incoming radiation is the short-wave solar radiation absorbed by the earth and atmosphere. Outgoing radiation is the long-wave radiation leaving the atmosphere. (*From Winston et al., 1979*)

at low latitudes than near the poles, but its outgoing radiation is more uniformly spread, decreasing only slightly with latitude (Fig. 19.4). The resulting heat excess at low latitudes and deficit at high latitudes call for a poleward heat transfer. George Hadley[3] hypothesized that this transfer is accomplished by a giant thermally driven circulation: Warm tropical air rises and flows toward each pole, where it cools and sinks, returning to the tropics along the surface (Hadley, 1735). As it turns out, Hadley was partly correct, insofar as such convective circulations exist on both sides of the equator, and partly incorrect, insofar as these meridional circulations extend only to 30° of latitude.

Indeed, a single Hadley cell spanning equator to pole is unlikely to exist because of conservation of angular momentum. In the absence of friction, a torus of equatorial air mass m at rest with respect to the earth conserves its

[3]British physicist and meteorologist (1685–1768) who first explained the trade winds.

absolute angular momentum. Hence, when progressing poleward to latitude φ, it ought to conserve $ma^2\Omega = mr(\Omega r + u) = ma\cos\varphi\ (\Omega a\cos\varphi + u)$, where a is the earth's radius and $r = a\cos\varphi$ is the distance of the torus to the earth's axis of rotation. This would lead to unrealistically high wind velocities u at high latitudes.

North of 30°N and south of 30°S, different circulations are observed, up to 60°, beyond which circulations in the sense predicted by Hadley are again found. Because the convective circulations theorized by Hadley follow our intuition, they are generally called *direct cells*. Those direct cells bordering the equator are also called *Hadley cells*. In contrast, the reverse circulations found at midlatitudes bear the name of *indirect cells*. Our purpose here is to explain, in some qualitative manner, why such oppositely directed meridional circulations exist. The story is not simple, invoking the aggregate effect of the transient weather systems (storms) of the midlatitude regions.

To begin, we note that, although a single direct convective cell could theoretically span an entire hemisphere, such would be unstable. The strong zonal flow in thermal-wind balance with the large meridional temperature gradient would be baroclinically unstable. In fact, the more moderate zonal winds accompanying the alternating circulation structure that exists on the earth are themselves unstable, as the vagaries of the midlatitude weather show so well. According to our discussion of baroclinic instability (Section 17.4), such instabilities develop into coherent vortex systems, called cyclones and anticyclones, that are capable of transferring heat meridionally (Section 17.5). At midlatitudes, therefore, the transfer does not take place in a vertical loop, as in a Hadley cell, but through the horizontal circulation of each vortex moving warm air poleward on one side and cold air equatorward on the other. We will now show how the cumulative action of these weather systems at midlatitudes can perform the required poleward transfer of heat so effectively as to reverse the meridional circulation in the vertical plane.

The analysis starts with a few modifications of the governing equations. First, the density departure from the reference ρ_0 is expressed in terms of a temperature anomaly T measured from the temperature corresponding to the reference density: $\rho = -\rho_0\alpha T$, where $\alpha = 1/T_0$ is the thermal-expansion coefficient. Then, viscosity and heat diffusivity are neglected, but a heat source or sink term is added in the temperature equation to represent the heat gain in the tropics and the heat loss at high latitudes. From Eqs. (4.21), we have

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} - fv = -\frac{1}{\rho_0}\frac{\partial p}{\partial x} \tag{19.7a}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} + fu = -\frac{1}{\rho_0}\frac{\partial p}{\partial y} \tag{19.7b}$$

$$\frac{\partial p}{\partial z} = \rho_0\alpha gT \tag{19.7c}$$

$$\frac{\partial u}{\partial x}+\frac{\partial v}{\partial y}+\frac{\partial w}{\partial z}=0 \tag{19.7d}$$

$$\frac{\partial T}{\partial t}+u\frac{\partial T}{\partial x}+v\frac{\partial T}{\partial y}+w\frac{\partial T}{\partial z}=\frac{Q}{\rho_0 C_p}, \tag{19.7e}$$

where Q is the aforementioned thermal forcing (in watts per cubic meter). Focusing exclusively on the northern hemisphere, we take Q positive in the tropics (at lower values of y, the northward coordinate) and negative at high latitudes (higher values of y). Thus, the gradient $\partial Q/\partial y$ is negative. The choice of beta-plane equations based on a Cartesian coordinate system over more accurate equations in spherical coordinates is justified in the spirit of a highly simplified analysis aimed at highlighting physical processes in a qualitative way.

We next define the zonal average as the mean over the values of x at any given y and z levels and time t. The zonal averages of the linear equations (19.7c) and (19.7d) are immediate:

$$\frac{\partial \bar{p}}{\partial z}=\rho_0 \alpha g \bar{T} \tag{19.8}$$

$$\frac{\partial \bar{v}}{\partial y}+\frac{\partial \bar{w}}{\partial z}=0, \tag{19.9}$$

where the overbar denotes this zonal average. With a prime denoting the departure from the average (e.g., $u=\bar{u}+u'$ etc.) and with some use of Eq. (19.7d), the zonal average of Eq. (19.7b) can be expressed as

$$\frac{\partial \bar{v}}{\partial t}+\bar{v}\frac{\partial \bar{v}}{\partial y}+\bar{w}\frac{\partial \bar{v}}{\partial z}+f\bar{u}=-\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial y}-\frac{\partial}{\partial y}\overline{v'^2}-\frac{\partial}{\partial z}\overline{v'w'}. \tag{19.10}$$

The large meridional pressure gradient ($\partial \bar{p}/\partial y$) associated with the important northward decrease in temperature ($\partial \bar{T}/\partial y<0$) is balanced by a significant zonal flow ($\bar{u}$). In contrast, the meridional cell ($\bar{v}$, $\bar{w}$) is much weaker, as are the corresponding eddy fluxes ($\overline{v'^2}$, $\overline{v'w'}$). Thus, the preceding may be reduced to

$$f\bar{u}=-\frac{1}{\rho_0}\frac{\partial \bar{p}}{\partial y}. \tag{19.11}$$

Together, the hydrostatic balance, Eq. (19.8), and the geostrophic relation, Eq. (19.11), provide the thermal-wind relation

$$f\frac{\partial \bar{u}}{\partial z}=-\alpha g\frac{\partial \bar{T}}{\partial y}, \tag{19.12}$$

which relates the vertical shear of the average zonal wind to the average meridional temperature gradient. With the temperature decreasing northward in the northern hemisphere ($\partial \bar{T}/\partial y<0, f>0$), the wind shear is positive ($\partial \bar{u}/\partial z>0$), indicating that the winds must become more westerly (eastward) with altitude.

Finally, we apply the zonal average to the remaining two equations, (19.7a) and (19.7e), to obtain:

$$\frac{\partial \bar{u}}{\partial t} + \bar{v}\frac{\partial \bar{u}}{\partial y} + \bar{w}\frac{\partial \bar{u}}{\partial z} - f\bar{v} = -\frac{\partial}{\partial y}\overline{u'v'} - \frac{\partial}{\partial z}\overline{u'w'} \tag{19.13a}$$

$$\frac{\partial \bar{T}}{\partial t} + \bar{v}\frac{\partial \bar{T}}{\partial y} + \bar{w}\frac{\partial \bar{T}}{\partial z} = \frac{\bar{Q}}{\rho_0 C_p} - \frac{\partial}{\partial y}\overline{v'T'} - \frac{\partial}{\partial z}\overline{w'T'}. \tag{19.13b}$$

According to our previous remarks, the eddy fluxes of momentum and heat associated with the horizontal circulations of the weather systems ($\overline{u'v'}$ and $\overline{v'T'}$) are anticipated to be important, and the corresponding terms are retained. However, the vertical eddy fluxes ($\overline{u'w'}$ and $\overline{w'T'}$) are neglected. Except for the mean vertical advection of temperature ($\bar{w}\partial\bar{T}/\partial z$) because there is a substantial vertical stratification, mean meridional and vertical advection is unimportant, compared with the meridional eddy transports. In the light of these considerations, the leading terms of the preceding two equations are

$$\frac{\partial \bar{u}}{\partial t} - f\bar{v} = -\frac{\partial}{\partial y}\overline{u'v'} \tag{19.14}$$

$$\frac{\partial \bar{T}}{\partial t} + \frac{N^2}{\alpha g}\bar{w} = \frac{\bar{Q}}{\rho_0 C_p} - \frac{\partial}{\partial y}\overline{v'T'}. \tag{19.15}$$

Here, we have introduced the stratification frequency N through $N^2 = \alpha g \partial\bar{T}/\partial z$. We shall assume that it does not vary significantly with y.

Forming f times the z-derivative of the first equation plus αg times the y-derivative of the second, to eliminate the time derivatives through Eq. (19.12), we obtain

$$\underbrace{\frac{\partial \bar{w}}{\partial y} - \frac{f^2}{N^2}\frac{\partial \bar{v}}{\partial z}}_{=\omega} = \frac{\alpha g}{\rho_0 C_p N^2}\frac{\partial \bar{Q}}{\partial y} - \frac{\alpha g}{N^2}\frac{\partial^2}{\partial y^2}\overline{v'T'} - \frac{f}{N^2}\frac{\partial^2}{\partial y \partial z}\overline{u'v'}. \tag{19.16}$$

In this last equation, the sign of the left-hand side ω is directly related to the direction of the average circulation in the vertical plane. For simplicity, let us restrict our attention again to the northern hemisphere. In a direct cell (Fig. 19.5a), $\bar{w}$ decreases northward and $\bar{v}$ increases upward, together yielding a negative ω. On the other hand (Fig. 19.5b), an indirect cell corresponds to a positive left-hand side.

According to the right-hand side of Eq. (19.16), there are three competing mechanisms influencing the sense of the circulation. In the tropical regions, away from the midlatitude eddy activity, the dominant factor is heating ($\bar{Q}$ term). Because the rate of heating decreases northward ($\partial\bar{Q}/\partial y < 0$), this term is negative, and the circulation in the vertical plane is in the direct sense (as in

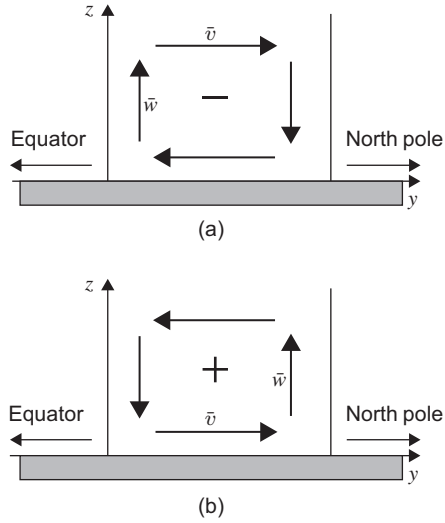

FIGURE 19.5 Atmospheric circulation in the meridional-vertical plane: (a) direct cell, also called Hadley cell, with $\partial\bar{w}/\partial y < 0$, $\partial\bar{v}/\partial z > 0$, and $\omega < 0$, and (b) indirect cell, also called Ferrel cell, with opposite circulation and positive ω.

Fig. 19.5a). This occurs up to about 30°N, and the circulation driven by thermal convection is the Hadley cell. The northerly (equatorward) winds along the surface ($\bar{v} < 0$) veer to the right under the action of the Coriolis force, resulting in easterly (westward) zonal winds ($\bar{u} < 0$). These form the *trade winds*.

North of approximately 30°N, where the eddy activity is most intense, the corresponding terms ($\overline{v'T'}$ and $\overline{u'v'}$) dominate the right-hand side of Eq. (19.16). Both induce an indirect circulation. This is easy to see with the $\overline{v'T'}$ term and a little harder with the $\overline{u'v'}$ term. The average product $\overline{v'T'}$ is proportional to the meridional heat flux of the eddies. Because this net heat flux must be northward, warm anomalies ($T' > 0$) are preferentially moved northward ($v' > 0$), while cold anomalies, are advected southward ($T' < 0$, $v' < 0$), both yielding a net positive $\overline{v'T'}$ correlation. Because the storm activity is most intense at midlatitudes, the term $\overline{v'T'}$ reaches a maximum there. Thus, the second derivative $\partial^2\overline{v'T'}/\partial y^2$ must be negative. Preceded by a minus sign in Eq. (19.16), the corresponding term is positive.

The convergence of warm and cold air masses aloft creates a locally intensified gradient of temperature. In thermal-wind balance with this gradient is the polar-front jet stream (Fig. 18.1) that flows eastward. The maintenance of this jet in spite of eddy activity requires a continuous influx of eastward momentum (i.e., positive u' anomalies must be transported to that latitude). This is effected by the eddies, which import positive momentum anomalies from the south ($u' > 0$, $v' > 0$) and from the north ($u' > 0$, $v' < 0$). Thus, the average $\overline{u'v'}$ is

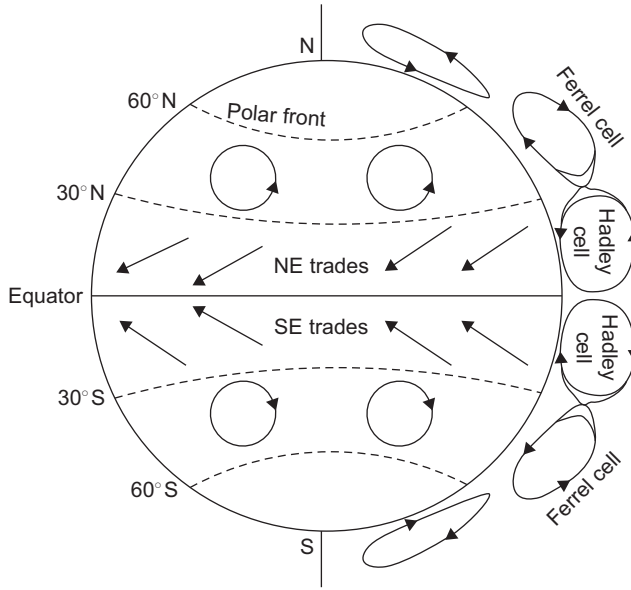

FIGURE 19.6 Sketch of the general atmospheric circulation, composed of direct (Hadley) and indirect (Ferrel) cells in the meridional direction and alternating winds in the zonal direction.

positive south of the jet and negative north of it, and the derivative $\partial\overline{u'v'}/\partial y$ must be negative. At the surface, where the jet stream is not found, the correlation $\overline{u'v'}$ is much less important, and we conclude that $\partial\overline{u'v'}/\partial y$ is increasingly negative with altitude, namely, $\partial^2\overline{u'v'}/\partial y\partial z$ is negative. Preceded by a minus sign in Eq. (19.16), this term adds to the positive contribution of the other eddy-flux term, and together these terms overcome the $\bar{Q}$ term. The result is an indirect cell, called the *Ferrel cell*. A corresponding indirect cell is found in the southern hemisphere. These Ferrel cells extend to approximately 60°; beyond that latitude, eddy activity yields to a thermal circulation in the vertical and direct cells exist (Fig. 19.6).

The alternation of direct and indirect cells across latitudes leads to a similar alternation in surface zonal winds: from the easterly trades to the prevailing westerlies, to the polar easterlies (Fig. 19.6).

19.4 ATMOSPHERIC CIRCULATION MODELS

Atmospheric circulation models are generally at the forefront of developments in both parameterization of subgrid-scale processes and numerical aspects (see Section 1.9) and continue to change in terms of included physics and numerical discretizations. From the first operational models using a single-layer quasi-geostrophic approach, models have progressed to solving the primitive

equations at ever shorter scales. The equations solved numerically are the governing equations of Section 4.4, to which are added turbulence closure, cloud parameterization, radiation budgets, and tracer evolution. An account of the improvements in atmospheric models during the last couple of decades lies beyond the scope of this book (see Randall, 2000, for a coverage of the progress made during the 1990s), and we focus here on the distinguishing features of atmospheric models compared to oceanic models or general geophysical-flow models.

The most widely used models for weather prediction at global scale include those of the European Centre for Medium-Range Weather Forecasts (ECMWF, EU) and the National Centers for Environmental Prediction (NCEP, USA). Both models are adapted to the atmosphere by including a series of physical models or parameterizations, specific to the air that we have on Earth. Radiation budgets are more complicated than those of Section 19.2, and in reality, the heat equation should contain a local source term due to radiation, the behavior of which should depend among other things on the orientation of the sun, the wavelength of the radiation, humidity and presence of aerosols, the latter being transported with the winds and hence governed by an advection-diffusion equation. Because radiation is behaving differently for each wavelength, a separate equation for radiative transfer should be used for each wavelength. In practice, wavelengths are lumped into spectral bands, and, at a minimum, two groups are distinguished: short- and long-wave radiation, as we did earlier in our simple globally averaged models. The models include absorption of radiation (and hence local heating) by water vapor, ozone, carbon dioxide, and clouds within the atmosphere itself and on the lower boundary, the earth's surface. Not only absorption must be calculated at each grid point, but also the scattering of radiation by aerosols and clouds, as well as the reflection by the earth's surface and clouds. One must further take care of the re-emission of long-wave radiation by ozone. These processes involve a series of parameterizations which make up the particular radiative transfer equations of the models.

The ECMWF model uses for example a radiation scheme based on Orcrette (1991): for clear-sky conditions, short-wave radiation is mainly constrained by aerosol scattering and the effects of the absorption by water vapor, ozone, oxygen, carbon monoxide, methane, and nitrous oxide. Clear-sky, long-wave radiation is modeled using absorptive properties of water vapor, carbon dioxide, and ozone, which are temperature and pressure dependent. Cloudy skies are handled separately, and their parameterization includes absorption and scattering properties of cloud droplets, with clouds being characterized by optical thickness and their scattering properties. Cloud physics are not only important for studies of radiative transfer but also in precipitation forecasts. Because of their short dimensions compared with typical model grid sizes, clouds need to be parameterized in some ways. This requires specific treatment (Section 19.6).

When weather is calculated over the entire planet, Atmospheric General Circulation Models (AGCMs) have to take into account the spherical nature of

the domain and hence governing equations are most naturally written in spherical coordinates (Appendix A). This is both a source of complications and of simplifications. Complications arise because of the more complex nature of the equations (coefficients now depend on latitude φ) and of the mathematical singularities at each pole because of the presence of $1/\cos\varphi$ in certain coefficients. The latter causes numerical stability problems.

To understand why numerical stability problems may arise, let us imagine that discretization is performed by using a rectangular grid in the longitude–latitude (λ, φ) coordinates with $\Delta\lambda = 2\pi/M$ where M is the number of grid points used in the west–east direction. Then, the Euclidean distance between two neighboring grid points is $\Delta x = a\cos\varphi\,\Delta\lambda$, where a is the earth's radius. It is clear that this distance Δx vanishes at the poles. So, if there is any numerical stability condition of the type $U\Delta t \leq \Delta x$, where U is a physical propagation velocity of similar magnitude across latitudes, the stability condition will be much more stringent near the pole than near the equator, even if the underlying physical process acts similarly in both locations. The overall numerical efficiency is then drastically reduced if a pole imposes its stability condition on the rest of the domain. This is called the problem of *convergence of meridians*, and it must be addressed. If finite-difference grids in longitude–latitude are used, an implicit scheme or filtering must be used in the vicinity of the poles.

A distinct simplification for any AGCM covering the planet is the absence of any lateral boundary, avoiding the need for open-boundary conditions. Regional models (so-called *Limited Area Models*, nicknamed LAMs) may also avoid the open-boundary problem by appropriate nesting within an AGCM running alongside. The ALADIN model (*Aire Limitée Adaptation dynamique Développement InterNational*[4]) is one such LAM used for downscaling processes from the global scale to the desired regional scale. ALADIN simulates smaller scale features such as the sea breeze and thunderstorms, using high-resolution orography and less parameterization.

Even if they do not have lateral boundaries, AGCMs need conditions at the vertical edges of the domain. The upper boundary of atmospheric models is generally taken at a given pressure level (e.g., 0.25 hPa) or at a given height (e.g., 70 km), well above the troposphere in which most weather phenomena are confined. It is also placed well above the tropopause to avoid unphysical reflections of waves at the boundary. This, however, is still an artificial boundary because air gradually rarefies with height, and there exists no distinct level where the atmosphere ends and space begins. Nevertheless, rigid-lid conditions are commonly assumed. At the lower boundary, the atmosphere interacts with oceans, land masses, and ice covers, which all demand specific definitions of fluxes of heat and momentum.

[4]Translation: Limited area, dynamical adaptation, international development.

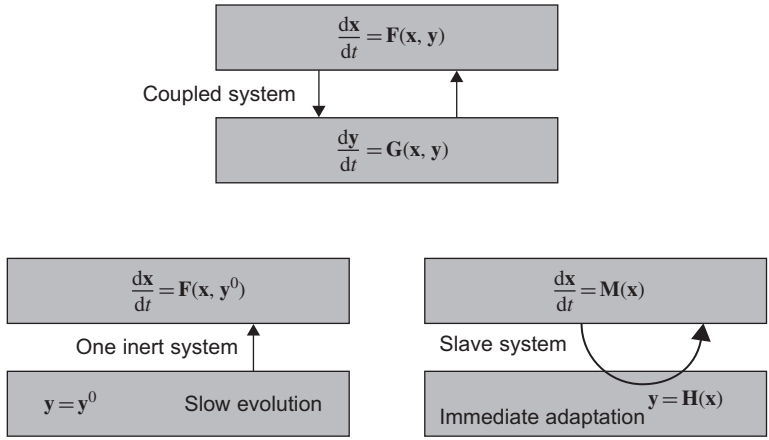

FIGURE 19.7 Depending on its intrinsic timescale compared with that of the atmosphere, the effect of an interacting system (such as vegetation or clouds) may be reduced to persistence or instantaneous adaptation. Only when the timescales of both systems are comparable is the fully coupled version to be retained.

Depending on the timescale of the processes at hand, those interactions between systems may be simplified (Fig. 19.7). If the system coupled to the atmosphere (e.g., glaciers or vegetation) is reacting comparatively slowly, there is no need to take into account temporal variations in the feedback mechanism, and persistence of the slower system can be assumed while the atmospheric model is moved forward for some time. For example, the ice cover of Antarctica does not change significantly within the few days covered by a weather forecast model, and the observation of the ice cover at the beginning of the forecast is sufficient to constrain the atmospheric evolution during the forecast period.

If on the contrary the coupled component reacts relatively fast compared with atmospheric conditions, it is often possible to establish quasi-equilibrium laws predicting those fast adaptations directly in terms of atmospheric parameters. For example, the albedo of the land surface changes in time following alterations to surface characteristics and should normally require a comprehensive land-use model, but, if the atmospheric model predicts snow fall, the albedo may be immediately updated for use in the atmospheric model so as to accommodate the change in reflection.

When the non-atmospheric system possesses a similar timescale to, and interacts with, the atmosphere, both models must be run in parallel, and outputs from each must inform the other. A good example of such a situation is the forecast of El Niño events (see Section 21.4). Such coupling, however, is not easy, and the first attempts at coupling AGCMs to ocean models for climate calculations demanded unphysical discontinuities in fluxes between the atmosphere and ocean. Ocean models are sensitive to errors in wind distribution over the sea

surface, requiring that the wind field provided by the atmospheric model include the same timescales (spectral window) as those active in the ocean model. When fluxes were not subjected to one correction or another, the atmospheric and ocean models drifted away from each other, leading to unrealistic situations. To enable simulation of past climatic variations, it was then deemed necessary to inject information on the climatic average within the flux formulations, usually by way of relaxation toward a known state. This meant that a piece of the solution had to be incorporated into the model formulation. Such an unsatisfactory approach of feeding models with *a priori* knowledge was cause for objecting that climate forecasts based on such models could not be trusted. At this time, improvements have been made, and *flux correction* is no longer necessary. Coupled ocean-atmospheric models are now the core of so-called *global climate models* incorporating a vast number of physical, biological, and chemical components such as wind, ocean currents, ice cover, hydrological cycle, vegetation, land use, carbon and nutrients cycles, and so on.

The coupling of models allows feedback to be taken into account such as the melting of ice due to heating followed by a change in the albedo itself modifying heat budgets. Other feedback loops are possible through chemical reactions. The ozone layer, for example, changes under modified climate conditions, itself changing radiative budgets. Integrated models that include feedback loops are called *Earth simulators*. They incorporate as many processes as possible instead of considering them as forcing functions. Note that because of high computational demands, most of the submodels are optimized to exploit particular computer hardware (parallel, distributed, or shared), and making them work together is not trivial. Because of the physical coupling, this requires information exchange between models, called message passing on parallel computers. The implementation of exchanges is quite a challenging task both in terms of physical-interaction modeling as well as technical programming. Needless to say, such integrated models are also more and more demanding in terms of understanding the simulation results, particularly when grid resolution increases and more and more physical processes are resolved. After all, the model should behave almost as the real world does, which, as we know, is extremely difficult to comprehend. Hence, associated with most models, there is now a suite of statistical and graphical analysis tools to help the modeler grapple with the huge amount of information produced.

When speaking about atmospheric models, it should always be specified if the model is used for weather forecast, seasonal forecast, or climate-change scenarios. Indeed, a frequent argument invoked to disqualify climate-change studies is the inability to predict weather beyond a few days while germane models are used for investigating climate change. This argument simply disregards the difference between weather and climate (see Section 19.1). We may well be unable to predict next week's weather in New York but still be able to predict an increase of temperature over the United States over the next few years. The situation is similar to the unpredictability of individual eddies in a

turbulent flow not preventing us from making meaningful statements about their aggregate effect on pollutant transport while tapping from the same family of governing principles and models. The problem is simply a question of scales of interest, and the reader may benefit from meditating on Fig. 1.7. As long as the model possesses prediction capability for the processes at temporal and spatial scales of interest, it does not matter whether its prediction capabilities are reduced for processes at other scales.

19.5 BRIEF REMARKS ON WEATHER FORECASTING

For weather forecasts, namely for forecasts not beyond of few days, most of the feedbacks with systems other than the atmosphere itself can be simplified and rendered relatively passive. For example, air–sea heat fluxes depend on sea surface temperature (SST), and the situation should ideally call for a fully coupled ocean-atmosphere model, but for weather forecast, the atmospheric model can rely on estimates of sea surface temperature, such as climatological sea surface temperature, observed SST over previous days, a simple mixed-layer model, or any combination of these.

The general approach of using observations to prescribe forcings explains why weather forecasting needs to rely on dense observational networks. To improve over time, however, prediction capabilities demand not only denser observational networks and better data-assimilation techniques but also more sophisticated physical parameterizations, better performing numerical methods, and ever more powerful computers to permit increased spatial resolution.

Nowadays, weather predictions are accompanied by calculated error bars around the predicted values. Thus, the weather reporter may speak about probability of occurrence of events, such rain probability. The error bars are not restricted to precipitation, however; they also apply to temperature, dew point, wind speed, pressure, cloud cover, snow fall, visibility, radiation, and so on. Although forecast capability has improved significantly over the last decades, weather remains somewhat elusive and difficult to predict, especially in the case of extreme events.

19.6 CLOUD PARAMETERIZATIONS

Probably the most difficult parameterizations encountered in atmospheric models are those related to clouds because several difficulties arise simultaneously: The physical processes at play are multiple and complex, and cover a broad range of scales. Clouds involve microphysics at crystal and water-drop level (millimeter scale), convection related processes at larger scales (hundreds of meters), and turbulence at all scales. Yet, none of these can possibly be represented in an explicit way in a global model. In one grid cell of a global model, even the largest cumulus clouds barely cover a fraction of the space.

At the shortest scale, physical processes leading to precipitation are complicated but can be explained by thermodynamics. The description involves water condensation on tiny solid particles (aerosols), collisions of incipient droplets, and interaction between water droplets and ice crystals (*Bergeron process*). The problem for models, of course, is that they are not able to represent these processes for each individual droplet, but models must nonetheless incorporate the aggregate effect at the level of the variables retained in the model. Mixing ratios[5] of cloud water content and ice content are typical variables to be computed. Because drop formation is controlled by the presence of condensation nuclei, aerosol concentration is also a pertinent variable. Then, in the calculation of water budget, ice and vapor content within a grid cell, exchanges between the three phases of water (ice, liquid water, and water vapor) must appear in the governing equations. The problem is to extrapolate these microphysical processes to the scale of the grid cell. For example, condensation and evaporation can coexist depending on the vapor pressure distribution within the cell (saturated vapor). Yet, the model can only calculate one value for the entire grid cell (Fig. 19.8).

Outlining a series of cloud parameterization techniques lies beyond the scope of the present introductory text, and the reader is referred to Randall, Khairoutdinov, Arakawa and Grabowski (2003) for a helpful review and further references. For examples of actual parameterizations, one can consult Sundqvist, Berge and Kristjansson (1989). The variables most commonly involved include humidity and temperature in cloud-free and cloud-covered regions, as well as fractional stratiform cloud coverage, mixing ratios of cloud water and ice, mixing ratios of rain and snow, and so on.

For some cloud types, parameterization is easier than for others. *Stratus*-type clouds are associated with large-scale global upward motion and stratiform condensation so that they can be captured by model grids. *Cumulus*-type clouds, on the contrary, are formed by shorter-scale convective motions that are nonhydrostatic. Towers of ascending buoyant air, involving *thermals*, cannot be captured by the grid and need a heavy dose of parameterization, which generally assumes that surrounding air is uniform with properties specified by the model solution at the resolved scale so that exchange laws between convective and ambient air can be formulated in terms of temperature, moisture, and so on of the ambient air and cloud.

The situation is even more complicated with heat budgets associated with phase changes. Warming by condensation and cooling through evaporation may take place side by side, over a distance shorter than the width of a grid cell. This creates thermal gradients and movement at the level of unresolved dynamics that are important because they modify cloud behavior, which in turn affects radiation, heat content, and so on. In other words, the system includes feedback

[5]Mixing ratios are expressed in gram or kilogram of the variable per kilogram of air.

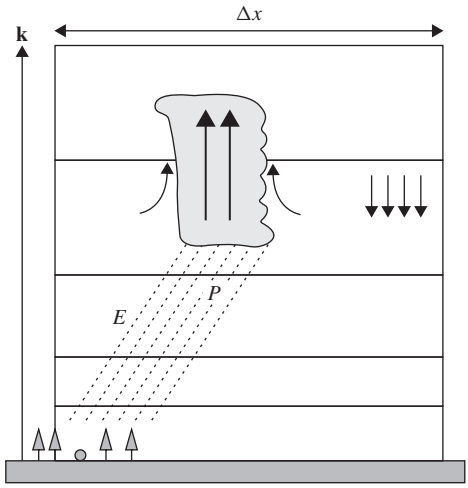

FIGURE 19.8 Numerical grids of global atmospheric models are too wide to represent individual clouds. The localized upward motion inside a cumulus cloud and the subsidence that surrounds it cannot be resolved by the grid, yet the cloud formed by condensation in upward motion of moist air and associated precipitation affects the water budget. The situation is further complicated by possible entrainment of surrounding unsaturated air, leading to evaporation, cooling, and weakening updraft. Precipitation originating in saturated air is normally modeled as a loss of water in the model. Yet, during its journey down to the earth surface, it can pass through unsaturated air and evaporate. This cools the air and, in the presence of downward entrainment creates a downdraft. These and other processes all take place on a spatial scale shorter than the size of a single grid cell.

mechanisms at unresolved scales. This is not only important for day-to-day prediction of rainfall, but clouds also play an essential role in climate dynamics. Because clouds provide shadow in the day but act as a thermal blankets during night, they are crucial in the global heat budget. Climate variations in turn modify the hydrological cycle and cloud coverage and yet another feedback exists. Because of the extreme complexity of cloud dynamics, the Intergovernmental Panel on Climate Change (IPCC, 2001) identified possible changes in cloud cover as one of the major uncertainties in predicting future states of the climate.

19.7 SPECTRAL METHODS

Numerical methods used in atmospheric models have evolved from quasi-geostrophic models using Arakawa grids for the discretization of the Jacobian operator and inversion of Poisson equations toward more sophisticated spectral models based on the primitive equations (i.e., no longer any QG approximation) with semi-Lagrangian tracer advection. Most modern global models are based on this approach, which we now outline.

Spectral models are based on the same technique as the spectral models presented in the quasi-geostrophic framework (see Section 16.7). They use a truncated series of orthogonal basis functions spanning the domain of interest. For global models, spherical coordinates do not lend themselves to a classical Fourier expansion of the solution in terms of trigonometric (sine and cosine)

functions, and more complicated spherical functions must be used. Assuming that the vertical dependence is handled by standard finite-volume or finite-difference techniques, the dependence of a field u on longitude λ and latitude φ is expressed as a series of functions $\mathrm{Y}_{m,n}$ called spherical harmonics:

$$u(\lambda, \varphi, t) = \sum_m \sum_n a_{m,n} \mathrm{Y}_{m,n}(\lambda, \sin\varphi) \tag{19.17}$$

in which

$$\mathrm{Y}_{m,n}(\lambda, \sin\varphi) = \mathrm{P}_{m,n}(\sin\varphi) \mathrm{e}^{\mathrm{i}m\lambda}. \tag{19.18}$$

The expansion series is of Fourier type in longitude but involves Legendre functions $\mathrm{P}_{m,n}$ of $\sin\varphi$:

$$\mathrm{P}_{m,n}(x) = \sqrt{\frac{(2n+1)(n-m)!}{2(n+m)!}(1-x^2)^m} \ \frac{\mathrm{d}^m}{\mathrm{d}x^m}\mathrm{P}_n(x). \tag{19.19}$$

These Legendre functions are in turn defined in terms of Legendre polynomials of degree n:

$$\mathrm{P}_n(x) = \frac{1}{2^n n!}\frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[(x^2-1)^n\right]. \tag{19.20}$$

Because P_n is a polynomial of degree n, Legendre functions differ from zero only when $m \leq n$. Extension toward negative values of m is desirable to correspond to the full set of Fourier modes $\exp(\mathrm{i}m\lambda)$, and so we extend the definition of Legendre functions with $\mathrm{P}_{-m,n}(x) = (-1)^m \mathrm{P}_{m,n}(x)$. With the well-chosen coefficient in front of Eq. (19.19), Legendre functions are orthonormal:

$$\int_{-1}^{1} \mathrm{P}_{m,n}(x)\mathrm{P}_{m,k}(x)\mathrm{d}x = \delta_{n,k}, \tag{19.21}$$

which is $=1$ if $n=k$ and $=0$ otherwise.

On the surface $\mathcal{S}$ of the sphere of radius r, the elementary surface element $r^2\cos\varphi\ \mathrm{d}\varphi\mathrm{d}\lambda$ may be written as $r^2\ \mathrm{d}\xi\mathrm{d}\lambda$ with $\xi = \sin\varphi$ so that

$$\frac{1}{r^2}\int_{\mathcal{S}} \mathrm{Y}_{m,n}\mathrm{Y}^*_{p,k}\mathrm{d}\mathcal{S} = \int_{-1}^{1}\int_{0}^{2\pi} \mathrm{Y}_{m,n}\mathrm{Y}^*_{p,k}\mathrm{d}\lambda\mathrm{d}\xi = 2\pi\,\delta_{m,p}\,\delta_{n,k}, \tag{19.22}$$

where the symbol * stands for the complex conjugate. The horizontal Laplacian of the basis functions is in spherical coordinates:

$$\nabla^2 \mathrm{Y}_{m,n} = -\frac{n(n+1)}{r^2}\,\mathrm{Y}_{m,n} \tag{19.23}$$

so that inversion of the Poisson equation can be performed algebraically in the transformed space. Note that, surprisingly, the pseudo[6] wavenumber $\sqrt{n(n+1)}/r$ is independent of m.

Orthogonality of spherical harmonics can be used to obtain separate evolution equations for the amplitudes $a_{m,n}(t)$ by multiplying the governing equations by $\mathrm{Y}_{m,n}{}^*$ and integrating over the global surface. To do so, we exploit the inverse transform (19.17) and the associated forward transformation that reads

$$a_{m,n} = \int_{-1}^{1} \left[\int_{0}^{2\pi} u(\lambda, \xi, t) e^{-\mathrm{i}m\lambda} \mathrm{d}\lambda \right] \mathrm{P}_{m,n}(\xi) \mathrm{d}\xi \tag{19.24}$$

as it follows from the orthogonality property.

Truncation of the sums in numerical schemes can be achieved in several ways (Fig. 19.10), the only constraint being that in all cases $|m| \leq n$ which can be achieved by

$$\tilde{u} = \sum_{m=-M}^{M} \sum_{n=|m|}^{N(m)} a_{m,n} \mathrm{Y}_{m,n} \tag{19.25}$$

The structure of spatial resolution depends on the value taken for $N(m)$. When $N(m) = M$, called a *triangular truncation*, uniform resolution is achieved on the sphere. Other truncations provide increased resolution in particular regions (Figs. 19.9 and 19.10).

The various possibilities are denoted according to the chosen truncation. For example, T256L60 signifies a triangular truncation with $M = 256$ spectral components.[7] The qualifier L60 stands for the vertical grid using 60 discrete levels.

In the ECMWF model, vertical levels are distributed according to a hybrid coordinate system where the vertical coordinate s depends on pressure p and surface pressure p_{surf} by $s = s(p, p_{\mathrm{surf}})$ scaled so that $s(0, p_{\mathrm{surf}}) = 0$ (top of the atmosphere) and $s(p_{\mathrm{surf}}, p_{\mathrm{surf}}) = 1$ (land or ocean surface). This is a generalization of a pressure coordinate, called σ coordinate, introduced by Phillips (1957) with $s = p/p_{\mathrm{surf}}$ and which is used in NCEP. Because more general hybrid vertical coordinates are now used in ocean models, we will postpone the corresponding discussion until Section 20.6.1.

---

[6] In Cartesian coordinates with the Fourier decomposition $u = U \exp(\mathrm{i}k_x x + \mathrm{i}k_y y)$, we would have $\nabla^2 u = -(k_x^2 + k_y^2)\, u$, hence the analogy.

[7] Note that because m takes negative values, the actual wavenumbers resolved are indeed M, contrary to the standard FFT presentation, which uses only positive values and hence half of the wavenumbers (see Appendix C).

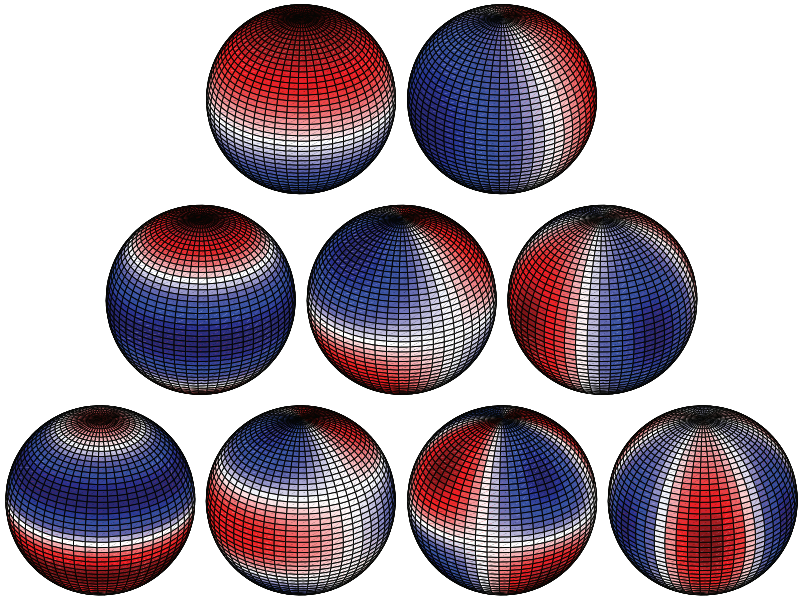

FIGURE 19.9 Real part of spherical harmonics $Y_{m,n}$ for (m, n) taking values (0,1),(1,1) on the first row, (0,2),(1,2),(2,2) on the second row, and (0,3),(1,3),(2,3),(3,3) on the bottom row. The white color marks the separation between positive and negative values. Mode (0,0) is not shown because its value is constant on the sphere.

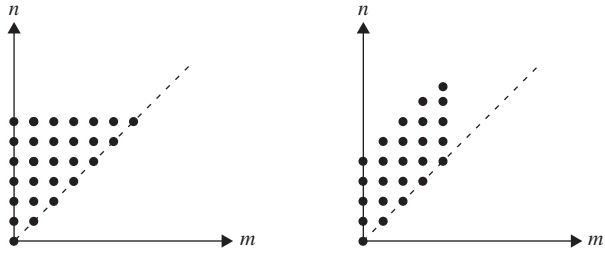

FIGURE 19.10 Triangular and rhomboidal truncations stipulating which values of m are retained for every value of n. Only positive m values are shown, the negative ones being symmetric. Triangular truncation is now more popular than the originally preferred rhomboidal version.

With discrete versions of integrals, orthogonality of basis functions is no longer ensured, and a direct transformation followed by an inverse transformation does not guarantee to a perfect return to the original function. For the discrete Fourier transform, orthogonality is maintained (see Appendix C and Section 18.4), so that we can evaluate the inner integral of the forward transform through an FFT and the corresponding inverse transform by an inverse

FFT. There only remains to ensure that the numerical treatment of the outer integral of Eq. (19.24) conserves orthogonality.

Unfortunately, there exists no numerical tool similar to the FFT that allows to perform transforms on the Legendre expansion as quickly, and we must resort to numerical quadrature of the integrals. First we can perform an FFT at a series of given latitudes $\varphi_j, j=1,\ldots,J$ with $\xi_j=\sin\varphi_j$ to obtain the Fourier coefficients

$$b_m(\xi_j,t)=\int_0^{2\pi} u(\lambda,\xi_j,t)\mathrm{e}^{-\mathrm{i}m\lambda}\mathrm{d}\lambda \tag{19.26}$$

defined at locations ξ_j. Then the time-dependent coefficients $a_{m,n}$ can be estimated by a numerical quadrature using the value of the integrand at those locations φ_j:

$$a_{m,n}=\sum_{j=1}^{J} \mathrm{w}_j b_m(\xi_j,t)\mathrm{P}_{m,n}(\xi_j). \tag{19.27}$$

The weights w_j and locations ξ_j can be chosen so as to reduce integration errors. Gaussian quadrature can be shown to produce exact results when integrating polynomials of degree $(2J-1)$ if the J points on which the integrand is evaluated coincide with the zeros of the Legendre polynomial of degree J, that is, $\mathrm{P}_J(\xi_j)=0$ and if the weights are taken as

$$\mathrm{w}_j=\frac{2}{(1-\xi_j^2)\left[\frac{\mathrm{dP}_J}{\mathrm{d}\xi}(\xi_j)\right]^2}. \tag{19.28}$$

It would appear that the integrands are not polynomials because Legendre functions involve square roots. What matters, however, is that transforms of the nonlinear terms, such as $u\partial u/\partial\lambda$, are treated correctly, and these, fortunately, involve products of Legendre functions which turn into polynomials that can be integrated exactly. The number of points J must then be taken so as to integrate correctly polynomials of the highest degree as a consequence of nonlinear terms. The transform of such a term would require the evaluation of triplet of Legendre functions (one for each appearance of u and then the application of the transform itself involving the third Legendre function). For a highest degree, $m=M$, of Legendre functions, a polynomial of degree $3M$ appears, and we must therefore use $J>(3M+1)/2$ points to integrate it exactly.

For the Fourier transform in longitude, we also face an aliasing problem, and the same analysis as in Section 18.4 applies. Because M modes use $(2M+1)$

grid points,[8] avoidance of aliasing requires the use of $(3M+1)$ evaluation points in longitude. Hence a model with 42 modes will use typically an underlying longitude–latitude grid of 128×64 points for the evaluation of the nonlinear terms (note the rounding toward powers of 2 to take advantage of efficient FFTs). This grid is called the *Gaussian grid* or *transform grid*. The calculation of grid spacing based on the number of Gaussian grid points overestimates actual resolution because it is designed to avoid aliasing on nonlinear interactions, and the actual, lower resolution is that which corresponds to the wavenumbers associated with the spectral decomposition.

The transform methods thus allows us to calculate some terms in spectral space (linear terms) and others (quadratic advection terms and nonlinear terms stemming from various parameterizations) in the transformed space so as to use the most appropriate technique for each term. In practice, it means that the model utilizes both spectral and grid representations of each variable.

The high convergence rate of spectral methods is inherited with the spherical harmonics, as long as the physical solution is sufficiently smooth. When fronts or jumps are present in the solution, however, spatial oscillations emerge near the place of rapid variation. This is known as Gibb's phenomena. The associated over- or undershooting on the physical grid can lead to spurious physical results. An overshooting of specific humidity, for example, may lead to the poetically named *spectral rain*.

Because of the calculation of some terms on the physical grid, the geometrical convergence of meridians toward the poles may also be a problem. For the advection part, this can be overcome by the semi-Lagrangian approach, which we describe next.

19.8 SEMI-LAGRANGIAN METHODS

To deal with advection, we again turn our attention to the passive-tracer concentration c, which is conserved along a trajectory of a parcel of fluid as long as diffusion remains negligible. The Lagrangian approach ensures exact conservation of the tracer value at the price of calculating its trajectory in time (see Section 12.8). As we have seen, however, the pure Lagrangian method leads sooner or later to an impractical distribution of particles, and it becomes impossible to determine concentration values in regions nearly void of particles. This is what happens when we follow the same set of particles over time: some of which flow out of the system or are caught in stagnation points. *Semi-Lagrangian* methods avoid this problem by using a different set of particles at each time step. The set is chosen at t^n so that at t^{n+1} the chosen particles arrive at the nodes of the numerical grid. This amounts to integrating trajectories

[8] Remember that the sum runs from $-M$ to M.

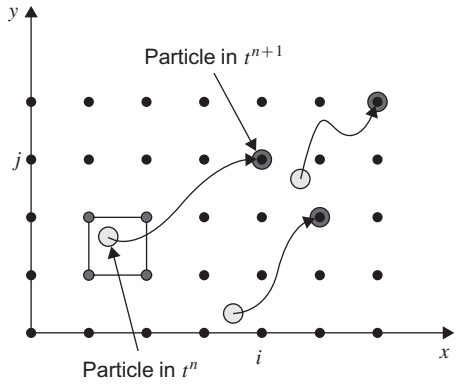

FIGURE 19.11 In a semi-Lagrangian method, trajectories are integrated backward in order to find an earlier location of the fluid particle that reaches grid node (i, j) by time t^{n+1}. Once this location is known, the value of the variable of interest, such as a temperature or concentration, at that location can be obtained by interpolation of nearby values. The interpolated value is then translated by advection to the new location (i, j) at time t^{n+1}.

backward for one time step in order to find where they originate. Once the past locations are determined, at t^n, the concentration in those locations is then determined by interpolation among known values on the grid (Fig. 19.11).

For simplification, let us consider first the one-dimensional case with positive velocity u and uniformly spaced grid (Fig. 19.12). The particle that lands at grid node x_i by time t^{n+1} was at the earlier time $t^n = t^{n+1} - \Delta t$ at position

$$x = x_i - u\Delta t. \tag{19.29}$$

On a uniform grid with spacing Δx, this position x most likely lies within a grid interval rather than, per chance, at an other grid point. This grid interval is given by

$$x_{i-1-p} \le x = x_i - u\Delta t \le x_{i-p}, \quad \text{with} \quad p = \text{integer part of } u\frac{\Delta t}{\Delta x}. \tag{19.30}$$

By virtue of advection without diffusion, the value of $\tilde{c}_i^{n+1}$ is none other than $\tilde{c}^n$ at x, a value which we can obtain by interpolation. Performing a linear interpolation, we obtain

$$\tilde{c}_i^{n+1} = \frac{(x_{i-p} - x)}{\Delta x}\tilde{c}_{i-1-p}^n + \frac{(x - x_{i-1-p})}{\Delta x}\tilde{c}_{i-p}^n = \tilde{C}\tilde{c}_{i-1-p}^n + (1-\tilde{C})\tilde{c}_{i-p}^n, \tag{19.31}$$

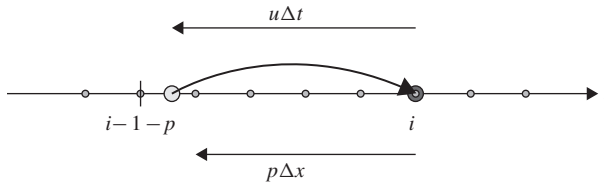

FIGURE 19.12 The semi-Lagrangian method in one dimension. The particle in light gray moves during time interval $[t^n, t^{n+1}]$ over a distance $u\Delta t$ to reach grid node labeled i at time t^{n+1}.

for which we define

$$\tilde{C} = \left(\frac{u\Delta t}{\Delta x} - p \right). \tag{19.32}$$

The scheme is monotonic and thus of first order. We can easily see that, as long as $u\Delta t \leq \Delta x$ (and thus $p = 0$), the scheme is equivalent to the upwind scheme. However, contrary to the upwind scheme, no stability condition is necessary here because the method uses the correct grid interval from which to interpolate. Numerical diffusion persists, however, although it is reduced in the sense that large time steps can be used and the total number of time steps can be decreased. For a given simulation time, fewer time steps mean less numerical diffusion.

To decrease the amount of diffusion introduced by the interpolation, a better than linear interpolation can be used. A second-order, parabolic interpolation yields a scheme equivalent to the Lax–Wendroff method.[9]

In two dimensions, the approach is readily generalized with backward trajectories for a single time step followed by 2D spatial interpolation (either bilinear or biparabolic). The trajectory calculation needs to take into account that the flow field (u, v), and this may become quite complicated if $U\Delta t \geq \Delta x$. If the velocity field varies over the trajectory on a scale comparable to the grid scale Δx, intermediate time steps are necessary for the calculation of the backward trajectory in order to maintain accuracy, and the calculation cost increases rapidly.

However, if the velocity is relatively smooth on the scale of the numerical grid (which ought to be the case and will necessarily be the case near the pole), that is, $\Delta x \ll L$, simple trajectory integrations will suffice. In fact, if $\Delta x \ll U\Delta t \ll L$ the semi-Lagrangian approach is much more efficient than the Eulerian method because during each time step, a large number of grid points can be "jumped over" by advection. Hence, interpolation (and associated diffusion) is less frequent, and the spatial scale of the trajectories is correctly captured. This is the way to reap the maximum benefit for the Semi-Lagrangian approach.

If $\Delta x \sim L$, time steps are similar to the Eulerian approach. The major advantage in this case is the stability of the method in the face of the occasionally excessive time step. For higher accuracy, however, one should not use longer time steps than allowed by $U\Delta t \sim L$. If there are many different tracers to be advected simultaneously, as in air pollution studies or ecosystem modeling, the method presents considerable advantages because a single, common trajectory needs to be calculated for all tracers.

For the nonadvective terms, such as source/sink and diffusion terms, a fractional-step approach is possible, for example, by first using a semi-

[9]Note the difference: In Eulerian methods, we spoke about interpolating for flux calculations to be discretized subsequently; here we speak about interpolation of the solution itself.

Lagrangian advection scheme followed by a Eulerian diffusion scheme either on the physical grid or in spectral space. Alternatively, the evolution of source terms may be taken into account along the trajectory (e.g., McDonald, 1986). Contrary to the finite-volume approach of Eulerian methods, global conservation properties are more difficult to handle but can be respected (e.g., Yabe, Xiao & Utsumi, 2001).

ANALYTICAL PROBLEMS

19.1. Consider the regular gardening greenhouse and idealize the system as follows: The air plays no role, the ground absorbs all radiation and reradiates it as a black body, and the glass is perfectly transparent to short-wave (visible) radiation but totally opaque to long-wave (heat) radiation. Further, the glass emits its radiation inward and outward in equal parts. Compare the ground temperature inside the greenhouse with that outside. Then, redo the exercise for a greenhouse with two layers of glass separated by a layer of air.

19.2. Consider the long-wave radiation fluxes of Figs. 19.2 and 19.3. In each case, the upward flux from the ground (E_2) is greater than the downward flux from the atmosphere ($0.64\, E_1$). Can you explain why?

19.3. Consider the crudest heat budget for the earth (without atmosphere and hydrological cycle) and assume the following dependency of the albedo on temperature: At low temperatures, much ice and clouds cover the earth, yielding a high albedo, whereas at high temperatures, the absence of ice and clouds reduce the albedo to zero. Taking the functional dependence as

$$\begin{aligned} \alpha &= 0.5 \quad \text{for} \quad T \leq 250\ \text{K} \\ \alpha &= \frac{270 - T}{40} \quad \text{for} \quad 250\ \text{K} \leq T \leq 270\ \text{K} \\ \alpha &= 0 \quad \text{for} \quad 270\ \text{K} \leq T, \end{aligned} \tag{19.33}$$

solve for the earth's average temperature T. Discuss the several solutions.

19.4. Using the global heat budget of the earth model, complete with an atmospheric layer and hydrological cycle, explore a worst-case scenario, whereby elevated concentrations of greenhouse gases completely block the transmission of long-wave radiation from the earth's surface, the intensity of the hydrological cycle is unchanged, and the anticipated global warming has caused the complete melting of all ice sheets, effectively eliminating all reflection by the earth's surface of short-wave solar radiation. What would then be the globally averaged temperature of the earth's surface? (Except for those transmission and reflection coefficients that need to be revised, use the parameter values quoted in the text.)

NUMERICAL EXERCISES

19.1. What is the spatial resolution (in kilometers) along the equator for a T256 spectral model? How many grid points must the underlying Gaussian grid have in order to avoid aliasing in the advection terms?

19.2. Use `spherical.m` to consider other basis functions $Y_{m,n}$ than those of Fig. 19.9.

19.3. Estimate the numerical cost of the forward and inverse transform associated with spectral harmonics.

19.4. In addition to the problem of decreasing grid spacing near the poles, which other problem can you identify at the poles for models that do not work with a spectral decomposition? (*Hint*: Think about boundary conditions for an AGCM, for longitude first and then for latitude.)

19.5. Exploiting properties of Legendre polynomials, given for example in Abramowitz and Stegun (1972), find the spectral coefficients of spatial derivatives, knowing the spectral coefficients of the function to be differentiated.

Edward Norton Lorenz
1917–2008

Edward Lorenz began his career as a weather forecaster for the U.S. army during World War II, before obtaining his doctorate in meteorology from MIT and later becoming professor of meteorology at the same institution. In the early 1960s, while using early numerical models of weather systems, Lorenz once introduced a small error in the value of a parameter that led to a completely different weather scenario. Puzzled by this, he proceeded to reduce the complex problem to an apparently simple set of equations (see Section 22.1) and sensitivity to very small perturbations persisted. The result was the discovery of chaos in atmospheric dynamics, published in a highly acclaimed paper (Lorenz, 1963). This contribution, among the most often cited scientific papers in the world, initiated a brand new field of research, that of deterministic chaos and strange attractors. Since then, the poetically named "butterfly effect" has become a standard metaphor for describing the sensitive dependence of a system's evolution on its initial conditions. (*Photo by Jane Loban*)

Joseph Smagorinsky
1924–2005

A native of New York City, Joseph Smagorinsky studied meteorology and began a career with the U.S. Weather Bureau. In 1955, he founded the *Geophysical Fluid Dynamics Laboratory*, which was first established in Washington and later relocated to Princeton University. The 1950s were exciting years when the prospect of computers gave hope that weather could be predicted by machines. Recognizing this opportunity, Smagorinsky developed numerical methods for predicting weather and climate, and by so doing profoundly influenced the practice of weather forecasting. In particular, he made the first attempt in 1955 to predict precipitation, and this led him to include compensating effects like radiation and to argue for the inclusion of a comprehensive "physics package," which became standard in all operational models.

Besides numerical methods and models, Smagorinsky also contributed to weather prediction by assuming a leading role in the establishment of a global observational network. While setting high goals for himself, Smagorinsky had an excellent sense of humor and a common touch. (*Photo by Michael Oort*)