

*I'm singing in the rain, Just singing in the rain,
What a glorious feelin', I'm happy again.
I'm laughing at clouds, So dark up above,
The sun's in my heart, And I'm ready for love.*

Lyrics by Arthur Freed, music by Nacio Herb Brown, *Singin' in the Rain*, 1929.

CHAPTER 18

Water Vapour and the Tropical Atmosphere

WATER IS AN ORDINARY SUBSTANCE WITH EXTRAORDINARY EFFECTS. The most obvious is that oceans themselves are made of water, and if our planet were dry this book would perforce be much shorter (if only). Leaving aside the dynamical effects of the oceans, water covers over two-thirds of Earth's surface and because it is warm in some places and cold in others, and because the atmosphere is in motion, water evaporates into the atmosphere in one place and condenses from it elsewhere. The condensation leads to rain, one of the most talked-about aspects of weather and climate. Water also freezes to form ice, so that at any given time water exists on Earth in all three phases. Radiatively, water vapour is a greenhouse gas, meaning that it absorbs infrared radiation that might otherwise be lost to space and so maintains the surface of the planet at a temperature over 20 K higher than an equivalent dry planet. Dynamically, the condensation of water vapour in the atmosphere releases energy, warming the air and tending to make it more unstable than otherwise and leading to convection. Further, the net transport of water vapour from low to high latitudes is effectively a meridional transport of energy.

In this chapter we focus on a small number of these issues, mainly on the kinematics and dynamics of water vapour itself and on some aspects of the dynamics of the tropical atmosphere, where the effects of water vapour are most manifest. The tropics would certainly differ from the mid-latitudes even if the atmosphere were dry — its Coriolis parameter is small among other things — so our attention there is by no means confined to the effects of water vapour. Nevertheless, tropical convection and the attendant ‘radiative-convective equilibrium’ are greatly influenced by the presence of water. We begin with a discussion of the thermodynamic properties of water vapour itself. We then move on to an essentially kinematic description of the factors determining the large scale distribution of relative humidity, before finally looking at convection and at tropical dynamics more generally.¹

18.1 A MOIST IDEAL GAS

Water is the compound of hydrogen and oxygen with the chemical formula H_2O , although in informal conversation water is often understood to mean only the liquid form of the compound. Water vapour is a gas made up of molecules of H_2O , and ice is the solid form of water. Steam, in common

parlance, is a mixture of air, water vapour and suspended droplets of water, usually at a very high temperature. Steam is formed when water vapour at temperatures above boiling point cools under contact with air and some of the water vapour condenses, forming a fine mist. Cloud and fog are also mixtures of dry air, water vapour and water droplets, but need not be at high temperature. Our focus will be on water vapour which, as we will see, can exist over a wide range of temperatures; let us first say how we quantify it and how it affects the equation of state. A number of thermodynamic derivations are also given in Appendix A (page 720). Those derivations are more systematic but less pedagogical than those below, and may appeal to some.

18.1.1 Ideal Gas Equation of State

The thermal equation of state for an ideal gas is conventionally written in the form

$$pV = Nk_B T = nR^*T, \quad (18.1)$$

where N is the total number of molecules in the volume V , n is the number of moles in that volume, and $N = nN_A$ where N_A is Avogadro's number. A mole is the amount of a substance that contains the same number of elementary entities, usually atoms or molecules, as there are atoms in 12 grams of carbon-12, that number being Avogadro's number ($N_A \approx 6.02 \times 10^{23}$). Two moles of a substance contains two times Avogadro's number of elementary units. The constants in the above equation are Boltzmann's constant, k_B , and the universal gas constant, R^* , where $R^* \equiv N_A k_B = 8.314 \text{ J mol}^{-1} \text{ K}^{-1}$. As noted in Chapter 1, for any particular gas it is convenient to define the specific gas constant by $R = R^*/\mu$ where μ is the molar mass (mean molecular weight in kg/mol). For a single component gas we then divide (18.1) by the total mass $M = n\mu$ to obtain

$$p = \rho RT. \quad (18.2)$$

Throughout this chapter we will be concerned only with 'simple ideal gases', or 'perfect gases', for which the gas constants at constant composition are, in fact, constant.

For a multi-component ideal gas the partial pressure of each component is independent of the presence of the other components (because the volume of the molecules is negligible) and so is equal to the hypothetical pressure of that gas if it alone occupied the volume of the mixture. The total pressure is therefore the sum of the partial pressures of each gas, a dictum known as Dalton's law of partial pressures. The partial pressure of each constituent in a mixture is proportional to the number of molecules of that constituent, and therefore also proportional to the number of moles. Because of Dalton's dictum we can obtain a simple expression for the equation of state of a mixture, as follows. Denoting the constituents by subscript i , the total pressure is given by

$$p = \sum_i p_i = \sum_i \frac{1}{V} n_i R^* T = \sum_i \left(\frac{M}{V} \right) \frac{n_i \mu_i}{M \mu_i} R^* T. \quad (18.3)$$

Let us define the effective molar mass, μ_e , by

$$\frac{1}{\mu_e} = \sum_i \frac{n_i \mu_i}{M \mu_i} = \sum_i \frac{\varphi_i}{\mu_i}, \quad (18.4)$$

where $\varphi_i = (n_i \mu_i)/M$ is the mass fraction of the i -th constituent. We then have

$$p = \rho RT, \quad \text{where} \quad R = \frac{R^*}{\mu_e} = \sum_i \varphi_i R_i, \quad (18.5a,b)$$

and $R_i = R^*/\mu_i$. The effective gas constant of the mixture is thus the mass-weighted mean of the specific gas constants of its constituents. Any given gas has a specific gas constant that is inversely

proportional to its molecular weight. Thus, for a given fluid density and temperature a gas with a higher molecular weight will exert a lower pressure than one that has a smaller molecular weight, because it will have fewer molecules per unit mass. Similar expressions apply to the heat capacities c_p and c_v , so that a heavier gas (higher molecular weight) has a smaller specific heat capacity.

18.1.2 Application to Moist Air

Dry air has virtually constant composition and its mean molar mass is $\mu^d = 29.0 \times 10^{-3} \text{ kg mol}^{-1}$, giving $R^d = R^*/\mu^d = 287 \text{ J kg}^{-1} \text{ K}^{-1}$. Water vapour has a molar mass of $\mu^v = 18.014 \times 10^{-3} \text{ kg mol}^{-1}$ giving $R^v = 461.5 \text{ J kg}^{-1} \text{ K}^{-1}$. The two gas constants are related by

$$\frac{R^v}{R^d} = \frac{\mu^d}{\mu^v} \equiv \frac{1}{\epsilon} \approx 1.608. \quad (18.6)$$

Now consider a mixture of dry air and water vapour.

Measures of moisture

When mixtures are present we use superscripts d , v and l to denote thermodynamic quantities associated with dry air, water vapour and liquid water. The *absolute humidity* is the amount of water vapour per unit volume, with units of kg m^{-3} , or informally g m^{-3} . The *mixing ratio*, w , is the ratio of the mass of water vapour, m^v , to that of dry air, m^d , in some volume of air and is thus

$$w \equiv \frac{m^v}{m^d} = \frac{\rho^v}{\rho^d}. \quad (18.7)$$

It is a nondimensional measure but it is often expressed in terms of grams per kilogram. In the atmosphere values range from close to zero to about 20 g kg^{-1} (2×10^{-2}) in the tropics on a humid day.

The *specific humidity*, q , is the ratio of the mass of water vapour to the total mass of air — dry air plus water vapour — and so is

$$q \equiv \frac{m^v}{m^d + m^v} = \frac{w}{1 + w} \quad \text{and} \quad w = \frac{q}{1 - q}. \quad (18.8a,b)$$

The specific humidity is just the mass concentration of water vapour in air. In most circumstances in Earth's atmosphere $m^v \ll m^d$ so that $q \approx w$, usually to an accuracy of about one percent. In most of this chapter we will ignore the differences between w and q , but this is not appropriate for all planetary atmospheres.

The partial pressure of water vapour in air, e , is the pressure exerted by water molecules and is proportional to the number of moles of water vapour in the volume. It is given by

$$e = \frac{n^v}{n^d + n^v} p = \frac{m^v/\mu^v}{m^d/\mu^d + m^v/\mu^v} p, \quad (18.9)$$

where n^v and n^d are the number of moles of water vapour and dry air in the mixture and p is the total pressure. Using (18.7) we can write (18.9) as

$$e = \frac{wp}{w + \epsilon} \quad \text{or} \quad w = \frac{\epsilon e}{p - e}. \quad (18.10)$$

In terms of q instead of w these expressions are

$$e = \frac{qp}{q + \epsilon(1 - q)} \quad \text{and} \quad q = \frac{\epsilon e}{p - e(1 - \epsilon)}, \quad (18.11)$$

In Earth's atmosphere $w \ll 1$ so that

$$e \approx w \frac{P}{\epsilon} = 1.61wp \quad \text{and} \quad q \approx w \approx \epsilon \frac{e}{P}. \quad (18.12)$$

If the mixing ratio of water vapour is 10 g kg^{-1} (a typical tropical value) and $p = 1000 \text{ hPa}$ then $e \approx 16 \text{ hPa}$.

The *relative humidity*, \mathcal{H} , is the ratio of the actual vapour pressure to the saturation vapour pressure, e_s , which is the maximum vapour pressure that can occur at a given temperature before condensation occurs, as will be discussed in Section 18.1.4. Thus, $\mathcal{H} = e/e_s \approx q/q_s$ where q_s is the specific humidity at saturation.

18.1.3 Equation of State and Virtual Temperature

Using (18.5b) the effective gas constant of moist air varies with humidity according to

$$R = \frac{m^d R^d + m^v R^v}{m^d + m^v} = (1 - q)R^d + qR^v = R^d \left[1 + q \left(\frac{1}{\epsilon} - 1 \right) \right], \quad (18.13)$$

with similar expressions for c_p and c_v . In humid air, with $R^d = 287 \text{ J kg}^{-1} \text{ K}^{-1}$ and $q = 0.02$ say, we have $R = R^d(1 + 0.02 \times 0.61) = 290.5 \text{ J kg}^{-1} \text{ K}^{-1}$,

The heat capacity of water vapour can be estimated from its molecular properties. Water vapour is a triatomic molecule with three translational and three rotational degrees of freedom. If these were the only degrees of freedom then the internal energy would be given by $I = 6R^v T/2$, whence $c_v^v \approx 3R^v = 1384 \text{ W m}^{-2}$ and $c_p^v = R^v + c_v^v = 1846 \text{ J kg}^{-1} \text{ K}^{-1}$, where c_v^v and c_p^v are the specific heat capacities for water vapour at constant volume and pressure, respectively. In fact vibrational degrees of freedom can sometimes be excited and the measured values are a little higher, namely $c_v^v = 1397 \text{ J kg}^{-1} \text{ K}^{-1}$ and $c_p^v = 1859 \text{ J kg}^{-1} \text{ K}^{-1}$ (at 273 K, increasing very slightly with temperature). The heat capacity of moist air is thus slightly higher than that of dry air, but since values of q are small the difference is only about 1%.

The variation of gas constant with humidity can be inconvenient in numerical models. A workaround is to define a so-called virtual temperature, T_v , which is the temperature that dry air would need to be in order to have the same density and pressure as moist air. That is, by definition,

$$p = \rho RT = \rho R^d T_v, \quad (18.14)$$

where R is given by (18.13). Using (18.13) we obtain

$$p = \rho R^d T_v, \quad \text{where} \quad T_v = T \left[1 + q \left(\frac{1}{\epsilon} - 1 \right) \right] \approx T(1 + 0.61q). \quad (18.15)$$

The virtual temperature, T_v , increases with specific humidity and if $q = 20 \text{ g kg}^{-1}$ then T_v is about 12%, or 3 K, larger than the actual temperature. Such a temperature is often used in numerical models of the atmosphere because it enables various thermodynamic equations to keep their original form, with gas constants that actually are constant.

Because the concentration of water vapour in Earth's atmosphere is so small, the variations of the heat capacities are small and constant values are often used to calculate quantities such as the potential temperature and the adiabatic lapse rate. This is not always appropriate, and Appendix A of this chapter indicates how, in principle, more accurate calculations could be made.

18.1.4 Saturation Vapour Pressure

Vapour pressure is the partial pressure of water vapour in the atmosphere. At any given temperature, there is a maximum value of that vapour pressure beyond which condensation normally

occurs and this is known as the *saturation vapour pressure*. Why should this be so, and why don't other atmospheric gases, such as oxygen or carbon dioxide, also condense? To understand this, we have to understand the thermodynamic equilibrium between a liquid and a gas.

Equilibration of the Gibbs function

Consider a system that consists of an enclosed, insulated container partially filled with liquid, and with vapour above it. The two subsystems can exchange mass and energy, with liquid potentially evaporating into vapour and vapour condensing into the liquid. Energy is required to evaporate the liquid into a vapour to overcome the molecular forces in the liquid and this is given by $M(h^v - h^l)$, where M is the mass that has evaporated, h^v is the specific enthalpy of the vapour and h^l is the specific enthalpy of the liquid. The enthalpy of vaporization, more commonly called the *latent heat of evaporation*, is defined by the difference between the two enthalpies at a temperature T , namely,

$$L(T) \equiv h^v - h^l, \quad (18.16)$$

with L having units of J kg^{-1} . It is a function of temperature, because the enthalpies of liquid water and water vapour are both functions of temperature, and a very weak function of pressure — see Appendix A for details. For water, L diminishes almost linearly by about 10% going from 0°C to 100°C, from 2.5×10^6 to $2.26 \times 10^6 \text{ J kg}^{-1}$.

Now suppose we leave the container alone for a long time so that the liquid and vapour come into equilibrium at a temperature T . If a mass M is to evaporate from liquid into vapour then the energy required, E , can be related to the entropy difference between the liquid and vapour phases,

$$E = ML = M(h^v - h^l) = MT(\eta^v - \eta^l), \quad (18.17)$$

where η^v and η^l are the specific entropies of the vapour and liquid, and the temperature is fixed because all the energy put into the liquid is used for evaporation. Re-arranging we find

$$h^v - T\eta^v = h^l - T\eta^l \quad \text{or} \quad g^v = g^l, \quad (18.18a,b)$$

where $g^l \equiv h^l - T\eta^l$ and $g^v \equiv h^v - T\eta^v$ are the specific Gibbs functions for the liquid and vapour (Section 1.5.2). That is, *the specific Gibbs functions for the liquid and water phases of a substance are the same at equilibrium*. The result follows directly from (18.17): the energy required to evaporate a mass of liquid, which is equal to the mass times the specific enthalpy difference between the vapour and the liquid, is also equal to the mass times the specific entropy difference between the vapour and liquid. The equality is true only at equilibrium, when temperature remains fixed, so that (18.18b) is an equation, not an identity.

Another way to derive the above result is to begin with the fact that the total Gibbs function, for the entire system, must remain constant. That is, if M_l and M^v are the masses of liquid and water, then

$$G = M_l g^l + M^v g^v \quad (18.19)$$

and

$$\delta G = M_l \delta g^l + M^v \delta g^v + (g^v - g^l) \delta M = 0, \quad (18.20)$$

where δM is the mass exchanged between liquid and vapour arising from a small fluctuation. Now, from (1.78) changes in Gibbs functions arise because of changes in temperature and pressure; that is, in general,

$$\delta g = -\eta \delta T + \alpha \delta p, \quad (18.21)$$

and given this, (18.20) becomes

$$\delta G = M_l(-\eta^l \delta T + \alpha^l \delta p) + M^v(-\eta^v \delta T + \alpha^v \delta p) + (g^v - g^l) \delta M = 0, \quad (18.22)$$

where α_l and α^v are the specific volumes (the inverse density) of the liquid and vapour, respectively. But the temperature and pressure are fixed, and thus, in order that $\delta G = 0$ we must have that $g^v = g^l$. The reason for equality of the two Gibbs functions — as opposed to the equality of some other thermodynamic potential — stems from the fact that the Gibbs function is the only potential for which the natural variables are intensive, namely temperature and pressure. The derivation we have given exploits this directly, for we kept p and T fixed.

18.1.5 Clausius–Clapeyron Equation

Now suppose that the temperature of the liquid-vapour system changes by an amount δT , leading to a change in the vapour pressure and in the specific Gibbs functions for the liquid and water vapour. Using (18.21) the change in the two Gibbs functions is given by

$$\delta g^l = -\eta^l \delta T + \alpha^l \delta p, \quad \delta g^v = -\eta^v \delta T + \alpha^v \delta p, \quad (18.23)$$

and, since the two changes must be the same, $\delta g^l = \delta g^v$. Re-arranging (18.23) and taking the limit of small changes then gives

$$\frac{dp}{dT} = \frac{\eta^l - \eta^v}{\alpha^v - \alpha^l}. \quad (18.24)$$

Using (18.16) and (18.17) this equation can be written

$$\frac{dp}{dT} = \frac{h^l - h^v}{T(\alpha^v - \alpha^l)} = \frac{L}{T(\alpha^v - \alpha^l)}, \quad (18.25)$$

where to obtain the rightmost expression we use the definition of L . The quantity p is the vapour pressure of the vapour above the liquid, which we are denoting e . Furthermore, it is the *saturation* vapour pressure, e_s , because the vapour is in equilibrium with the liquid: if more vapour were added it would immediately condense. Using this notation, the *saturation vapour pressure* of a condensable gas above a liquid is given by

$$\frac{de_s}{dT} = \frac{L}{T(\alpha^v - \alpha^l)}. \quad (18.26)$$

This is the *Clausius–Clapeyron* equation, and it tells us how the pressure of a vapour that is in thermodynamic equilibrium with an adjacent liquid varies with temperature. If for some reason the vapour pressure is higher than this value, and if there is an adjacent surface of liquid water, then the vapour will condense into a liquid — a common manifestation of which is the formation of clouds and rain. Evidently, since $L > 0$ and $\alpha^v > \alpha^l$, the saturation vapour pressure increases with temperature so that a reduction in temperature can lead to saturation. At the temperature at which the saturation vapour pressure of a substance equals that of the ambient pressure then any liquid present will *boil*. For water at a pressure of 1000 hPa this occurs at about 100°C, with a lower temperature needed at a lower pressure, which is why it takes longer to properly boil an egg at high altitude.

The presence of a liquid surface is crucial to the derivation, and it means that the equation only applies to a condensable. For gases such as carbon dioxide or oxygen at temperatures encountered on Earth, the saturation vapour pressure is very much higher than the actual pressure at the Earth's surface and the gas never condenses; the partial pressure of the gas is then determined by the ideal gas relation and not by (18.26). On Mars, temperatures are sufficiently low that carbon dioxide (the main constituent of the Martian atmosphere) is a condensate and as much as 25% of the Martian atmosphere will condense in winter. On Titan temperatures are even lower and methane is a condensate, and methane lakes are scattered over the dystopian surface.

On Earth, the partial pressure of water vapour, however, is constrained by (18.26), meaning that the partial pressure will often reach the saturated value and the vapour will then normally condense. Condensation is not, however, guaranteed, and to see this imagine a container that contains unsaturated water vapour and no liquid water, and suppose its temperature is then lowered; the pressure of the vapour will then fall following the ideal gas law. The saturation vapour pressure falls more quickly than this and so at some temperature the vapour pressure will exceed the saturation vapour pressure, but the vapour will not automatically condense since there is no liquid present and (18.26) does not apply. The vapour is then said to be *supersaturated*. A supersaturated state is unstable and condensation will eventually occur and liquid water will form, and subsequent changes in temperature induce pressure changes that satisfy the Clausius–Clapeyron equation. Supersaturated water vapour is fairly rare in Earth's atmosphere because there is usually no shortage of condensation nuclei.

At the other end of the temperature scale, liquid water can exist at temperatures well below freezing when there is insufficient water for the molecules to become organized into a crystalline structure, and super-cooled water droplets rather than ice then result — a common situation in clouds. If the liquid that is present is in the form of small spherical droplets then the saturation vapour pressure will differ slightly from that when the vapour is over a flat surface, because surface tension will affect the energy required for a molecule to escape from the droplet and so the latent heat of vaporization will differ. If the vapour is in contact with ice instead of water then its saturation vapour pressure will differ again, because the specific enthalpy of ice is different from that of liquid water.

Application to an ideal gas

In a mixture of ideal gases the partial pressure of one gas is unaffected by the presence of the other gases (because the volume of the gas molecules is assumed to be negligible) and in particular the saturation vapour pressure for a particular component is independent of the presence of other components. We can then, at least approximately, integrate (18.26) to see how the saturation vapour pressure of a particular component varies with temperature. Let us assume that density of the vapour is much less than that of the liquid, so that $\alpha^v \gg \alpha^l$. Given that the partial pressure of the vapour satisfies the ideal gas law, namely $e^v \alpha^v = R^v T$, where R^v is the specific gas constant for the vapour, the Clausius–Clapeyron equation becomes

$$\frac{de_s}{dT} = \frac{Le_s}{R^v T^2}. \quad (18.27)$$

This is the form of Clausius–Clapeyron equation that is normally used in atmospheric applications. If we further assume that L is a constant then (18.27) can be integrated to give

$$e_s = e_0 \exp \left[\frac{L}{R^v} \left(\frac{1}{T_0} - \frac{1}{T} \right) \right], \quad (18.28)$$

where e_0 and T_0 are constants, for example $T_0 = 273\text{ K}$ and $e_0 = 6.12\text{ hPa}$. Equation (18.28) is a good approximation if the temperature range is not too wide, and as seen in Fig. 18.1 the saturation vapour pressure of water increases approximately exponentially over commonly encountered terrestrial temperatures. Note that the expression for saturation vapour pressure does not depend on the presence or otherwise of dry air.

The fact that water vapour content cannot normally exceed the saturation value distinguishes the distribution of water from other tracers in the atmosphere, even without taking the heating effects of condensation into account. Note finally that if the atmosphere were motionless it would everywhere be in thermodynamic equilibrium with the moist surface and the surface layers would be saturated, and diffusion of water vapour upwards would then saturate the rest of the atmosphere. Thus, *the relative humidity of the atmosphere is determined by its circulation*, as we now discuss.

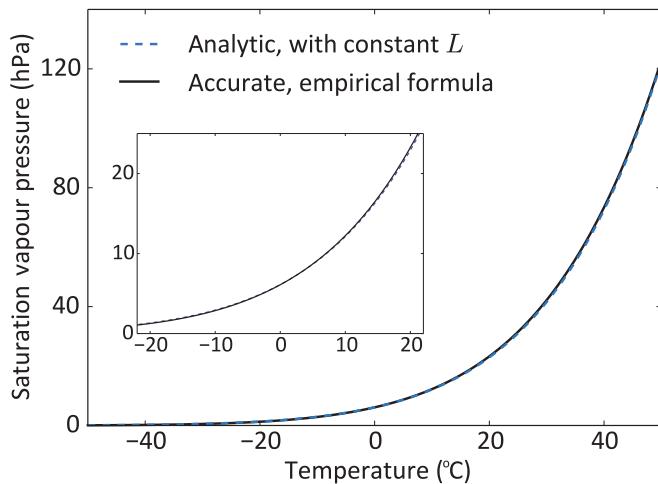


Fig. 18.1 The saturation vapour pressure of water vapour, calculated using the analytic formula (18.28), which assumes that L is constant (dashed line), and using a more accurate, semi-empirical formula (solid line).² The inset is the same plot over a smaller range.

The parameters used in the analytic formula are $T_0 = 273\text{ K}$, $e_0 = 6.12\text{ hPa}$, $L = 2.44 \times 10^6\text{ J kg}^{-1}$ and $R^v = 462\text{ J kg}^{-1}\text{ K}^{-1}$.

18.2 THE DISTRIBUTION OF RELATIVE HUMIDITY

To a first approximation, the distribution of water in the atmosphere is determined by the distribution of temperature. This is because of the near-exponential dependence of absolute humidity on temperature through the Clausius–Clapeyron equation, so that variations of relative humidity of even an order of magnitude, from 10% to 100% say, are barely noticeable in the specific humidity distribution, as seen in Fig. 18.2. It is, however, the relative humidity that determines such basic quantities as rainfall, and its distribution (Fig. 18.3 and Fig. 18.4) shows a quite different picture, with the following features evident:

- High, near-saturated values close to the ground.
- Low relative humidity in the subtropics at latitudes between 15° to 40° (depending on season) in both hemispheres.
- High values near the equator extending up to the tropopause.
- Vertically near-uniform values in mid- and high latitudes, increasing with latitude close to the pole in some cases.
- Very low values over much of the stratosphere.

The gross distribution of zonally-averaged temperature can be understood, at least in a rough way, using fairly simple arguments. The incoming solar radiation at the top of the atmosphere would lead, in the absence of atmospheric motion, to a strong meridional radiative equilibrium temperature gradient, as in Fig. 14.1. The meridional transport of heat by the Hadley Cell and by mid-latitude baroclinic eddies flattens that temperature gradient, and one might crudely model this transport as a diffusion. In the vertical heat is transported upwards by convection and baroclinic eddies, which might be modelled as a relaxation back to some specified neutrally stable profile or specified isentropic slope. However, arguments of this type cannot capture some basic features of the relative humidity distribution, and diffusive arguments in particular can be quite misleading. The effects of advection, either explicitly or as represented by a stochastic process, are crucial, and in this section we consider advection-diffusion-condensation models of the general form

$$\frac{\partial q}{\partial t} + \mathbf{v} \cdot \nabla q = \nabla \cdot \kappa \nabla q - S, \quad (18.29)$$

where q is the specific humidity, \mathbf{v} is a specified velocity field, κ is a diffusion coefficient and S is the condensational sink.⁴ Condensation in the atmosphere involves complicated microphysical processes but the basic effect is to remove liquid water once the volume becomes saturated and to

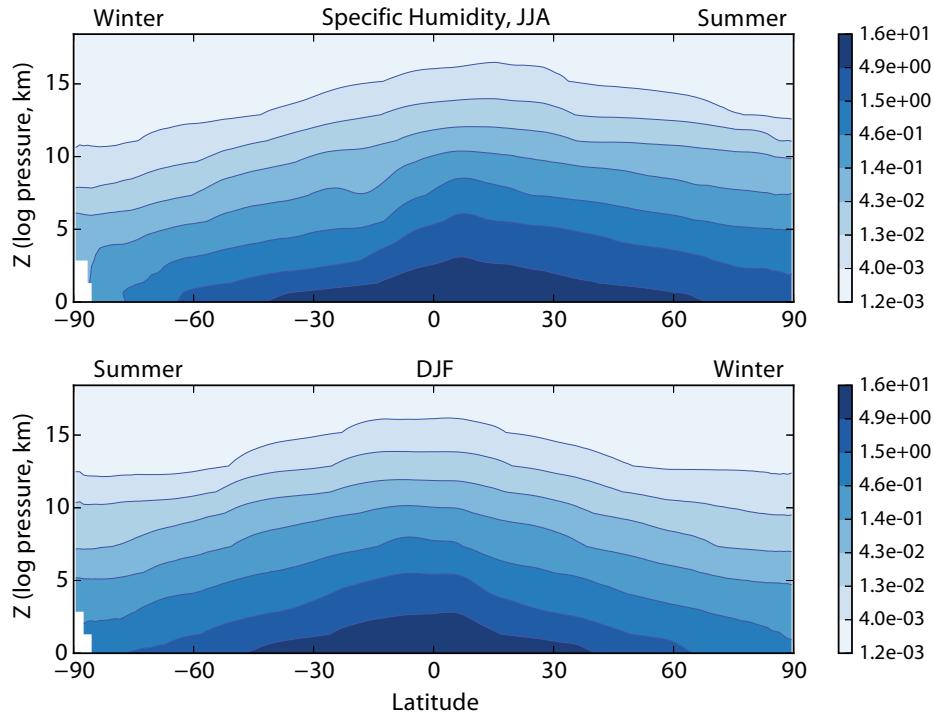


Fig. 18.2 Zonally-averaged specific humidity distribution (g/kg) in the atmosphere, as retrieved from a microwave satellite, for boreal summer in 2008 (top) and boreal winter 2008–2009 (bottom).³ Note the logarithmic scale.

largely prevent relative humidity from exceeding 100% (although supersaturation can occur locally if there are no particulates in the air onto which the water vapour may condense). A simple analytic way to represent such a condensation process is to let

$$S = \begin{cases} 0, & q \leq q_s, \\ (q - q_s)/\tau, & q > q_s, \end{cases} \quad (18.30)$$

where the time τ is much smaller than any large scale diffusion time, L^2/κ , where L is a characteristic length. Such a sink effectively prevents q from exceeding q_s except by a tiny amount. In (18.29) we may consider the advection and diffusion terms as representing larger scale processes and the sink as representing small-scale microphysical processes. We begin our exploration by omitting advection, and a summary can be found on page 683.

18.2.1 A Diffusion-Condensation Model

If we omit advection in (18.29) we have a simple diffusion-condensation model

$$\frac{\partial q}{\partial t} = \nabla \cdot \kappa \nabla q - S. \quad (18.31)$$

Although superficially plausible, such a model is too-often unable to reproduce locally unsaturated regions. For simplicity consider the one-dimensional case satisfying

$$\frac{\partial q}{\partial t} = \kappa \frac{\partial^2 q}{\partial x^2} - S, \quad (18.32)$$

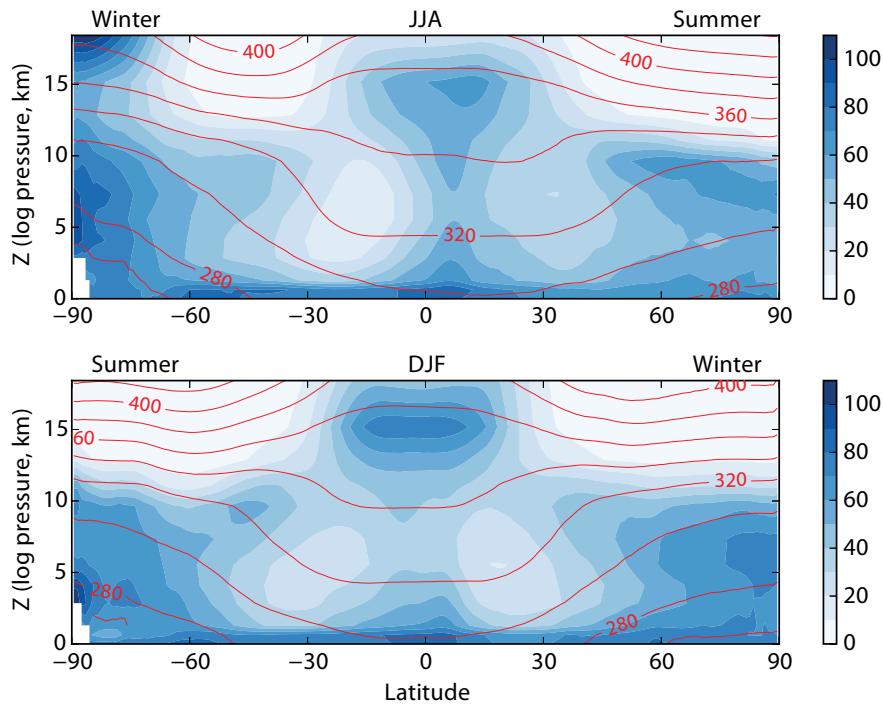


Fig. 18.3 Zonally-averaged relative humidity distribution in the atmosphere (in percent, shading), and isolines of equivalent potential temperature (contours), inferred from satellite as in Fig. 18.2. (Equivalent potential temperature is a modification of potential temperature to account for water vapour, and is approximately an adiabat; see Section 18.3.2.)

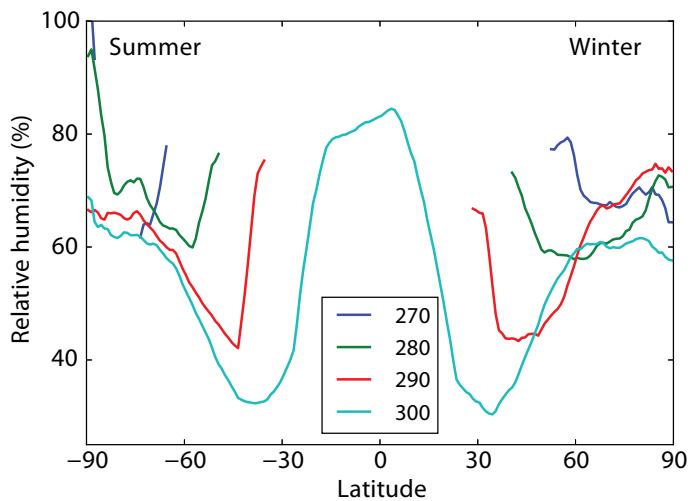


Fig. 18.4 Values of zonally-averaged relative humidity in boreal winter plotted along isolines of equivalent potential temperature (red contours in Fig. 18.3), with values of θ_{eq} as indicated in the legend.

Water Vapour Transport and Relative Humidity

- Water vapour in the atmosphere is primarily transported by advection, much of this on large, near planetary scales but also by convection and smaller scale turbulence. Specific humidity is materially conserved in the absence of condensation and diffusion.
- If the vapour pressure exceeds the saturated value (as given by the Clausius–Clapeyron relation) condensation will occur provided condensation nuclei or liquid water are present. The condensation normally occurs much more quickly than large-scale advective processes, and this situation is known as the *fast condensation limit*.
- When dealing with large-scale flows in Earth's atmosphere the limit is a good approximation. Levels of relative humidity are then determined mainly by advective processes rather than the microphysical details of the condensation process. Specifically, the relative humidity of a parcel is determined by the temperature at the location of last saturation, as in (18.35).
- If the advection is not fully resolved — for example if there is some small-scale turbulence in the flow — then introducing some diffusion seems natural, as is commonly done for tracers, but a large diffusivity can give unrealistic results because of the irreversible nature of condensation. Diffusion is then not a good representation of small-scale quasi-random flow and is overly prone to produce saturation.
- In Earth's atmosphere some of the large-scale features of the relative humidity distribution may be explained as follows:
 - High levels of relative humidity close to the surface. These are due to transport from a saturated surface, especially over the ocean and moist ground.
 - High levels of relative humidity in the ascending branch of the Hadley Cell. These are due to upward advection from a nearly saturated surface. The branch is, however, not saturated on the zonal-average, because of the presence of smaller scale motion such as downdrafts that unsaturate the air.
 - A subtropical minimum of relative humidity. This largely arises because of the mean descending motion, advecting water vapour into a warmer region and decreasing its relative humidity.
 - Variable relative humidity in mid- and high latitudes, with locally strong gradients. Chaotic advection by baroclinic eddies takes moisture upwards and polewards into cooler regions where it becomes saturated, but also downwards and equatorwards so reducing relative humidity.
 - Very low levels of relative humidity in the stratosphere. The tropopause is a cold trap and so relative humidity is very low beyond it. The cold-trap effect occurs in both advective and diffusive models. In Earth's atmosphere, the little water vapour that is in the stratosphere mainly enters advectively through the tropical tropopause.

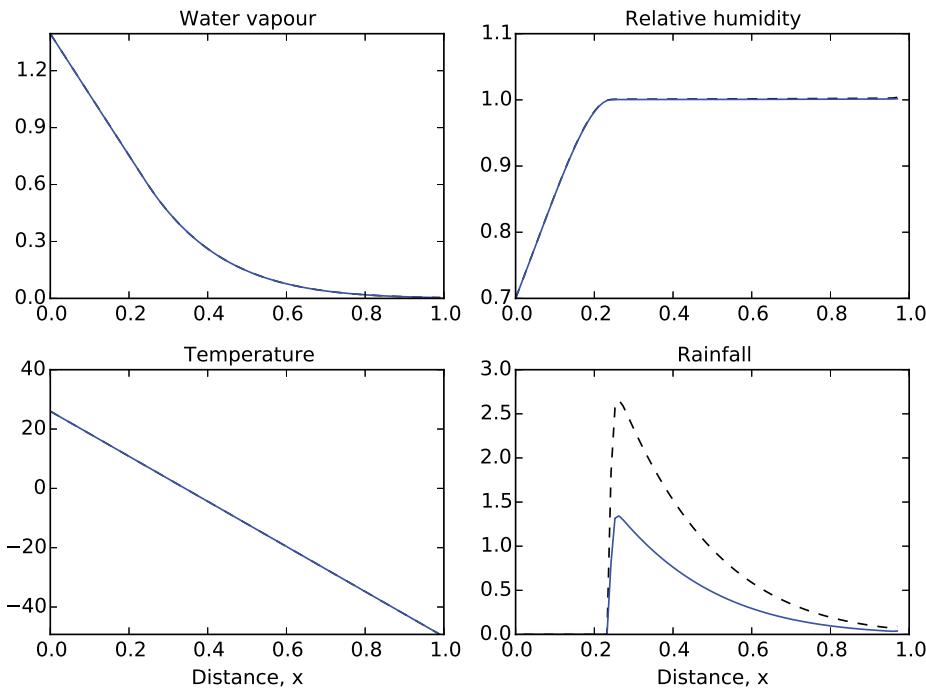


Fig. 18.5 Steady solution of the diffusion-condensation model (18.32) and (18.33), with $\mathcal{H}_b = 0.7$ at $x = 0$ and temperature ($^{\circ}\text{C}$) falling linearly from $x = 0$ as shown. Water vapour falls linearly away from the boundary at $x = 0$ until it becomes saturated at $x = 0.23$, after which the region is saturated. Two solutions are plotted, the one with dashed lines showing the solution with twice the diffusivity as that with solid lines. The two solutions are nearly identical except for the sink term, the rainfall.

with constant κ . In this model, any interior minimum of q_s will lead to saturation in that neighbourhood. To see this, note that in any region where moisture is present and that has $\partial^2 q / \partial x^2 > 0$ and $q < q_s$, there will be a net flux of water vapour into that region. Saturation must eventually occur, at which point q remains very close to the value of q_s . If $\partial^2 q_s / \partial x^2 > 0$ then the diffusive flux will maintain the saturated state. Given the monotonic dependence of q_s on temperature this result means that, in a moist atmosphere in which water vapour is transported diffusively, the neighbourhood of an interior minimum of temperature will become saturated.

A corollary of this result is that, unless there is a source of moisture at a boundary, a region with $\partial^2 q_s / \partial x^2 > 0$ everywhere will under many conditions eventually lose *nearly all* its moisture. Suppose that $\partial q / \partial x = 0$ at $x = 0$ and that q_s has a maximum at $x = 0$, and that the region extends to infinity and is initially saturated. (Envision a semi-infinite domain with temperature decreasing linearly away from a no-flux boundary, and therefore with no source of moisture, at $x = 0$.) Moisture is transported to higher values of x where condensation occurs and moisture is removed. As time progresses the region of saturation moves to higher and higher values of x , but nevertheless water is continuously removed. In a finite domain, with no flux boundary conditions at either end, a small amount of moisture will remain in the system for all time because a finite amount of water vapour is needed for condensation to occur.

Now consider the more atmospherically relevant situation with a moisture source at $x = 0$, with q_s decreasing monotonically away and with the other boundary either extending to infinity or being a no-flux boundary at finite x . Such a situation might represent an atmosphere sitting

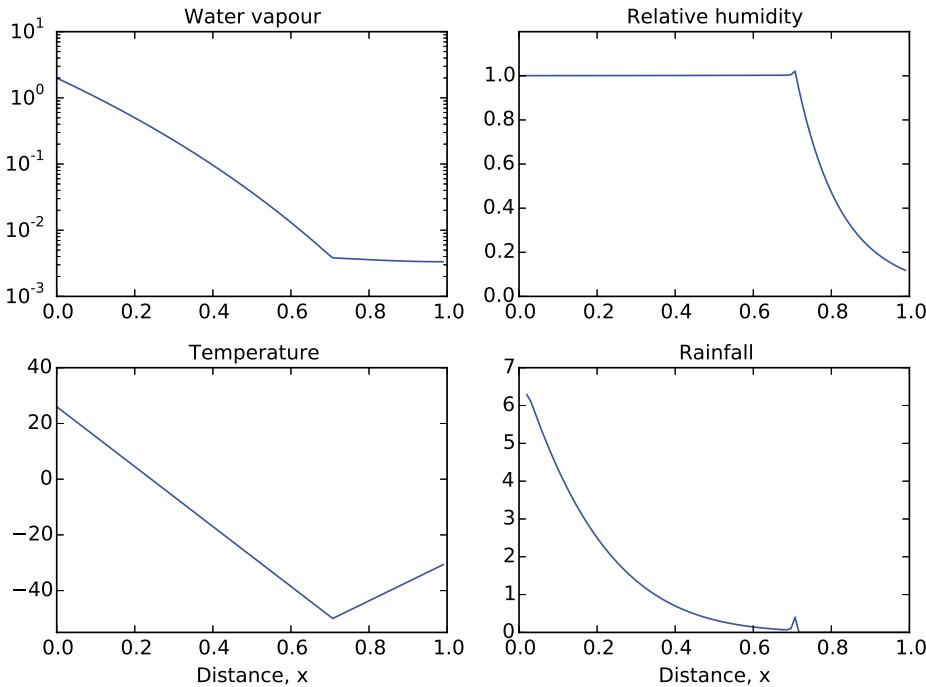


Fig. 18.6 As for Fig. 18.5, but now with $\mathcal{H}_b = 1$ and an interior temperature minimum, a ‘cold trap’, at $x = 0.7$ (see temperature panel), and water vapour has a log scale. Relative humidity falls rapidly beyond the cold trap and the rainfall there is zero.

atop a moist surface with temperature decreasing with height. Consider the case

$$\begin{aligned} q &= \mathcal{H}_b q_s(T_0), & x = 0, \\ \frac{\partial q}{\partial x} &= 0, & x = 1, \end{aligned} \tag{18.33}$$

where \mathcal{H}_b is a parameter such that if the boundary is effectively saturated then $\mathcal{H}_b = 1$, and $\mathcal{H}_b < 1$ otherwise, and we suppose that T decreases between T_0 at $x = 0$ and T_1 at $x = x_1$. Consider first the case with $H_b = 1$. If T falls linearly then the value of q_s falls approximately exponentially between $x = 0$ and $x = 1$, with $\partial^2 q_s / \partial x^2 > 0$, and the steady solution of this problem is that the domain is saturated everywhere. To see this, suppose that $q < q_s$ is some region so that $S = 0$. Water vapour will then diffuse into that region until condensation begins, maintaining $q \approx q_s$ everywhere, with, if (18.30) applies, q in fact exceeding q_s by a very small amount so that the diffusion into the region is balanced by condensation. If $H_b < 1$ then the region next to the surface will not be saturated and in steady state the water vapour content will decrease linearly, to satisfy $\partial^2 q / \partial x^2 = 0$, until at some value of x , x_s say, the atmosphere becomes saturated and remains so for $x > x_s$. The actual solution is

$$q(x) = \begin{cases} \mathcal{H}_b q_s(0) + x \Delta q / x_s, & x < x_s, \\ q_s(x), & x \geq x_s, \end{cases} \tag{18.34}$$

where $\Delta q = (q_s(x_s) - \mathcal{H}_b q_s(0))$ and x_s is such that the flux is continuous there. A moment’s thought reveals that $x_s = \Delta q / (\partial q_s / \partial x)_{x=x_s}$ and $x_s = 0$ if $\mathcal{H}_b = 1$. It is interesting that the values of water vapour in the solution (plotted in Fig. 18.5) do not depend upon κ and only very weakly on τ . In the fast condensation limit (in which τ is small compared to the diffusion time) variations of τ determine only the tiny amount by which q exceeds q_s . The amount of condensation is actually

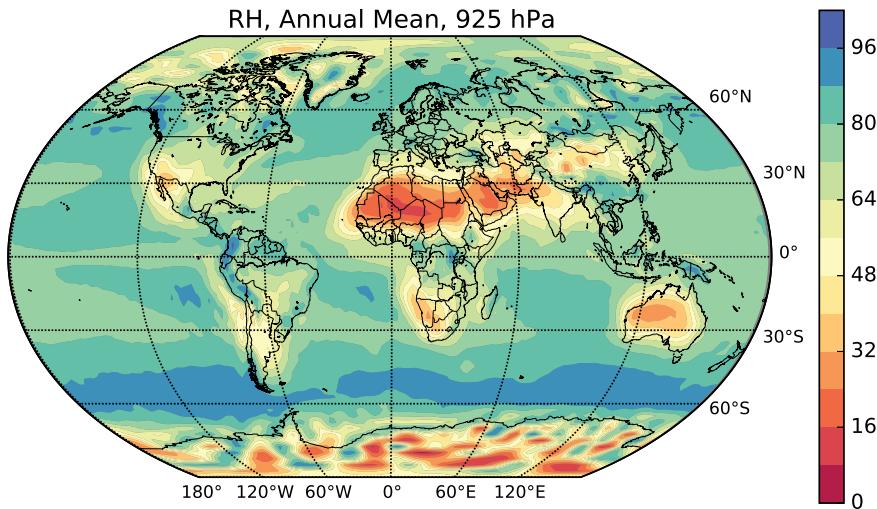


Fig. 18.7 Annually-averaged relative humidity (in percent) at 925 hPa, about 750 m above sea level. The contrast over land and ocean is apparent especially in the subtropics. (In regions of high topography values are interpolated.)

largely determined by the value of κ : large values of κ lead to a stronger diffusion of water vapour into dryer regions where it is almost immediately removed by condensation. This result illustrates the tenet that on large scales precipitation is at leading order determined by the motion of the fluid (here represented by diffusion), with variations in τ (crudely representing complex microphysical processes) being of less import. Microphysical processes are nevertheless important in many ways — a weather forecast model with poor microphysics would likely have little skill in forecasting the onset of precipitation, even if the climatology of the model were good.

Although the above model is over-simple in some respects, the dependence of the solution on H_b does capture the dependence of relative humidity in the lower atmosphere on the nature of the surface beneath, as seen in Fig. 18.7 showing relative humidity at 925 hPa. Over the desert regions the relative humidity is unsurprisingly low. Perhaps what is surprising is that the general dryness of the subtropics cannot be seen over the oceans — the surface moisture source simply overwhelms the drying effects of descending air (discussed more below).

A variation on the above theme introduces a temperature minimum, or ‘cold trap’, in the interior of the domain, as at $x = 0.7$ in Fig. 18.6. This configuration is a crude model of the tropopause, with temperatures increasing in the stratosphere beyond. Water vapour has to pass through the cold trap and so, since the specific humidity cannot be higher than the saturated value at the cold trap, the atmosphere will be unsaturated beyond it with relative humidity decreasing rapidly, as is seen in the real atmosphere in Fig. 18.3.

Although informative, diffusive-condensation models are fundamentally limited in what they can achieve, because of the deficiencies of diffusion in parameterizing the motion of a tracer in the presence of condensation. In particular, in the absence of a cold trap, diffusive models are prone to produce saturation everywhere. If the atmosphere obeyed (18.31) with a saturated surface, then the atmosphere would become saturated everywhere up to the tropopause, which from Fig. 18.3 is manifestly not the case. To remedy this we turn our attention to the effects of advection.

18.2.2 Advection-Diffusion-Condensation models

Consider now the effects of advection. Neglecting diffusion, the specific humidity of a parcel is conserved unless condensation occurs. If a moist parcel travels into a region of decreasing temperature

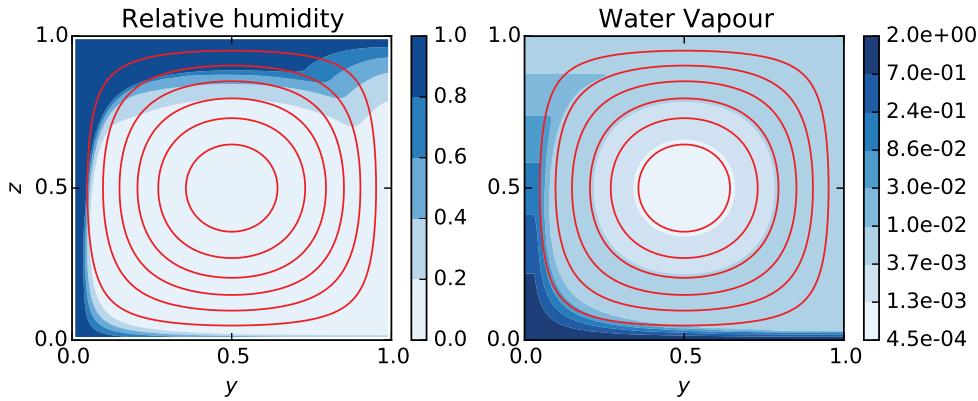


Fig. 18.8 Steady-state distributions of relative humidity (filled contours, left) and water vapour (right) in a single cell as defined by the streamfunction (red contours, clockwise flow, so air rising at the 'equator' at $y = 0$) in a closed domain. The domain boundary is saturated at the bottom, and temperature decreases linearly with height. The diffusivity $\kappa = 0.001$ and so the Peclet number, $Pe \equiv UL/\kappa \sim 1000$.

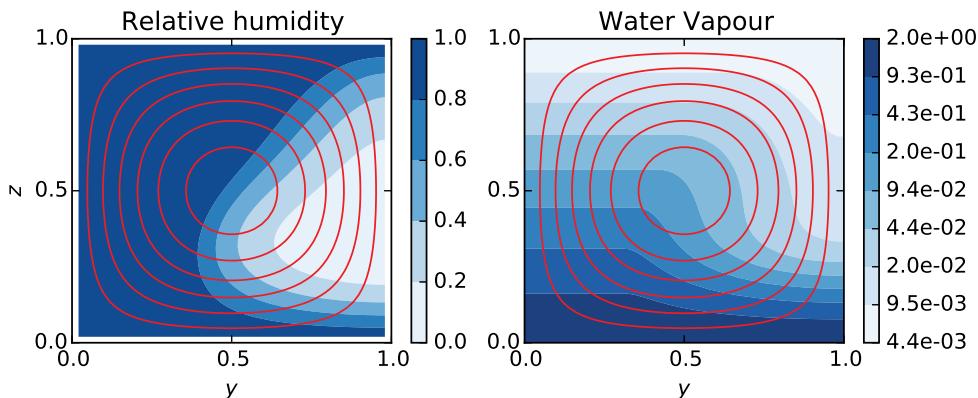


Fig. 18.9 As for Fig. 18.8, but with bigger diffusivity, $\kappa = 0.1$ and $Pe \sim 10$. Most of the domain is now much wetter.

it will eventually become saturated; however, if it passes into a region of increasing temperature then its relative humidity drops. Consider, for example, a cold trap with a temperature minimum at $x = x_{ct}$, as in the previous section. A parcel advected through the trap becomes saturated at x_{ct} , beyond which its specific humidity is constant, so that relative humidity is given by

$$\mathcal{H}(x) = \frac{q_s(T(x_{ct}))}{q_s(T(x))}. \quad (18.35)$$

That is to say, the relative humidity of a parcel is given by the value of the saturated vapour pressure at the point of last saturation, divided by the saturated vapour pressure at its current location.

Let us see the extent to which a simple advection-diffusion-condensation model can explain some of the features of the zonally-averaged relative humidity seen in Fig. 18.3. Consider a two-dimensional model in the y - z plane obeying (18.29), with a divergence-free advecting velocity, a saturated surface at $z = 0$ and no flux boundary conditions elsewhere. A simple model representing some of the features of the Hadley Cell is that of a single cell in a unit-sized square domain with a

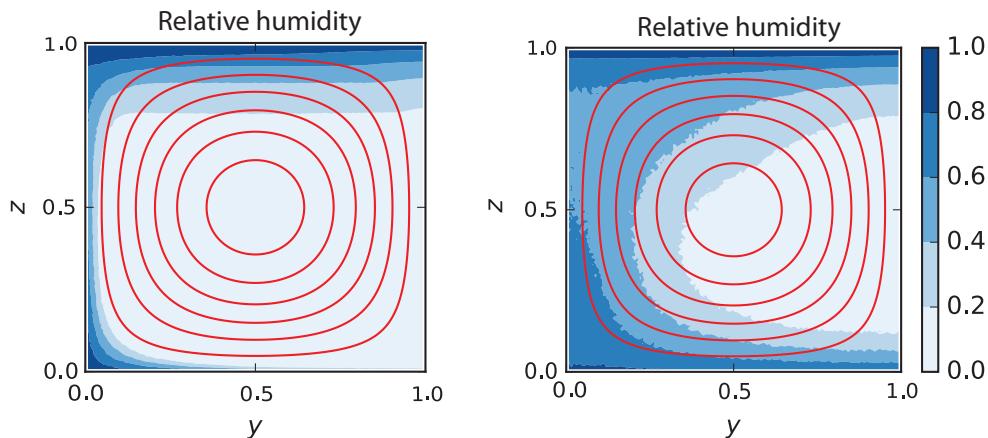


Fig. 18.10 The relative humidity produced using a stochastic advection-condensation model with the same imposed temperature, mean streamfunction (red contours) and boundary conditions as Fig. 18.8 and Fig. 18.9, with a larger stochasticity in the right panel. Specifically, the stochastic components of the left and right panels correspond to diffusivities of Figs. 18.8 and 18.9 respectively.

streamfunction of the form $\psi = \sin \pi y \sin \pi z$. We suppose the lower boundary is saturated, that there are no-flux boundary conditions on the other boundaries, and that temperature is uniform in latitude and falls linearly with height from 26°C to -50°C .

Solutions are illustrated in Fig. 18.8 and Fig. 18.9, the latter having diffusivity κ that is a hundred times larger. The upward branch in both cases is completely saturated. In the low diffusivity case the interior of the cell is quite dry, with a relative humidity that drops almost to zero in the centre. Increasing the diffusivity simply tends to moisten the interior, but does little to reduce the relative humidity in the upward branch of the cell. However, the upward branch of the real Hadley Cell is *not* saturated on the zonal mean (Fig. 18.3). Let us try to remedy this model failing.

Stochastic effects

The upwards branch of the Hadley Cell is not saturated because it is not steady — there is both ascending and descending motion, with the relative humidity of parcels falling as they descend so reducing the average value. We can mimic this effect by introducing a stochastic component into a Lagrangian model and performing ‘Monte Carlo’ simulations.⁵ The model advects infinitesimal parcels of water vapour — one million in the simulations shown in Fig. 18.10 — by a prescribed mean field plus a random component. Whenever the parcels touch the ground they become saturated and whenever their relative humidity exceeds one condensation occurs. We show results with two levels of stochastic motion, one comparable to the diffusivity used in Fig. 18.8 and the other comparable to that of Fig. 18.9, all with the same temperature, boundary condition and mean velocity fields, and the results shown in Fig. 18.10 are coarse-grained time averages. Although the results with weak stochasticity do resemble those with small diffusion (compare the left-hand panel of Fig. 18.10 with Fig. 18.8), those with large stochasticity have less resemblance to those with large diffusivity. Most notably, the explicit randomness reduces the relative humidity in the ascending branch without saturating the interior, a property that diffusion is unable to capture.

Relative humidity in mid-latitudes

The diffusive steady overturning model fails qualitatively when applied to an entire hemisphere, even with more realistic overturning circulation and temperature fields. To see this we construct a model in which the overturning streamfunction (representing a residual circulation) resembles that of the bottom panel of Fig. 15.22, as illustrated (red contours) in Fig. 18.11. To this we add

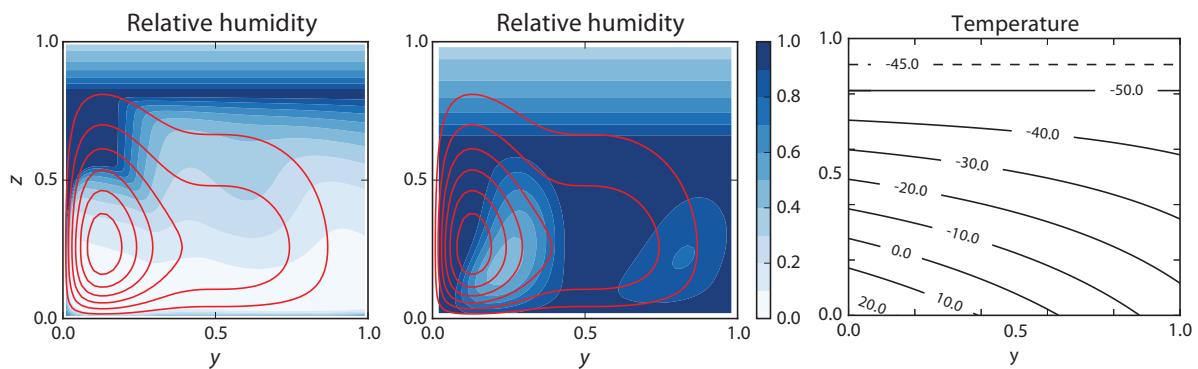


Fig. 18.11 Relative humidity (shading) for small diffusivity (left, $\kappa = 0.001$, $Pe = 1000$) and larger diffusivity (centre, $\kappa = 0.05$, $Pe = 20$), and a clockwise overturning circulation as contoured. The imposed temperature (right) falls linearly with height to a tropopause at $z = 0.8$ with $T = -50^\circ\text{C}$, then rises again to -40°C , and diminishes meridionally at the surface from 26°C at the equator ($y = 0$) to -15°C at the pole ($y = 1$).

a temperature distribution that decreases polewards and upwards until it reaches a tropopause at $z = 0.8$, then increases upwards again in the stratosphere, and we assume that the surface is saturated. With a small diffusivity (which is more realistic) the lower tropopause dries out too readily because of the sinking motion, and with larger diffusivity the atmosphere becomes more saturated but with little of the observed structure. Two features that the model *can* reproduce are the dry stratosphere, because of the tropospheric cold trap, and the subtropical minimum, a robust feature of the downwelling in the Hadley Cell.

The main failings of this model arise from the fact that in mid-latitudes water vapour is not primarily advected by the overturning circulation, and diffusion is a poor model of the meridional transport. Rather, moisture is transported from the boundary layer into the free atmosphere by convection and polewards by larger-scale baroclinic eddies in a quasi-horizontal fashion, roughly along moist isentropic surfaces. Mid-latitude relative humidity is thus highly variable, as can be seen in Fig. 18.12, with sharp gradients between high and low values and with variations being strongly correlated with the parcel trajectory — note for example the swath of high relative humidity just west of the UK associated with moisture moving poleward and to a lower temperature. As previously noted, on the zonal average relative humidity is high in the boundary layer, diminishes as one moves upwards, then stays roughly constant as one moves polewards along a moist isentrope before increasing again at very high latitudes (Fig. 18.3 and Fig. 18.4).

We can understand some of these features using a simple advective-diffusive-condensation model, with the unsteadiness explicitly incorporated into the advection, and with no other stochastic component. Consider (18.29) in two horizontal dimensions in a channel of size (L_x, L_y) , periodic in x , with a saturated boundary at $y = 0$ and no flux at $y = L + y$, and with imposed meridional temperature gradient ($^\circ\text{C}$) and advecting streamfunction of the forms

$$T(y) = 20 - 30y/L_y, \quad \psi(x, y, t) = -Uy + \Psi_0 \sin(\pi y/L_y) e^{i(kx - \omega t)}. \quad (18.36)$$

We will choose $|\Psi_0| = 1$, $U = 10$, $k = 4\pi/L_x$, $\kappa = 0.005$ (so $Pe \sim 200$ on the domain scale), and a solution is shown in Fig. 18.13. The details of the solution depend on the parameters chosen, but there are a couple of quite robust results:

- (i) Relative humidity at a point depends on where a parcel has come from, and hence is correlated with the streamlines, and is generally higher for parcels moving to colder regions.
- (ii) Relative humidity is close to one (i.e. 100%) next to the saturated boundary, then decreases

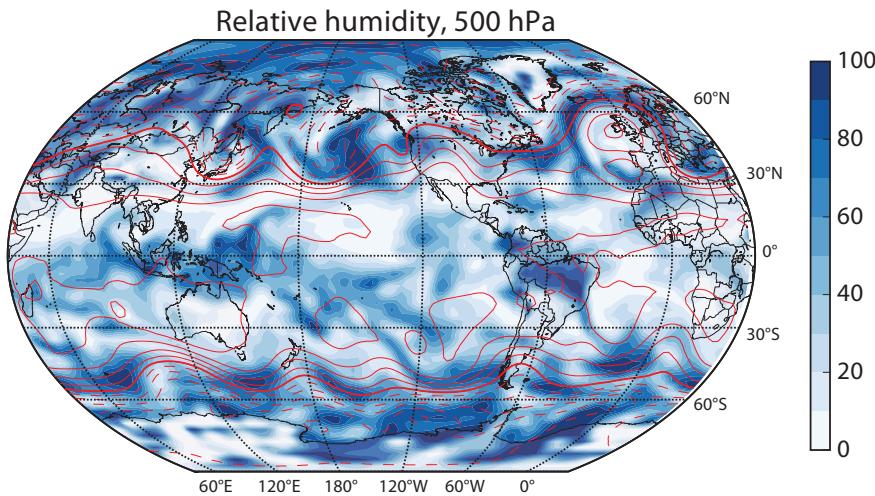


Fig. 18.12 Snapshot of relative humidity at 500 hPa on 9 February, 2015 (shading, in percent), with red contours of geopotential height, from reanalysis.

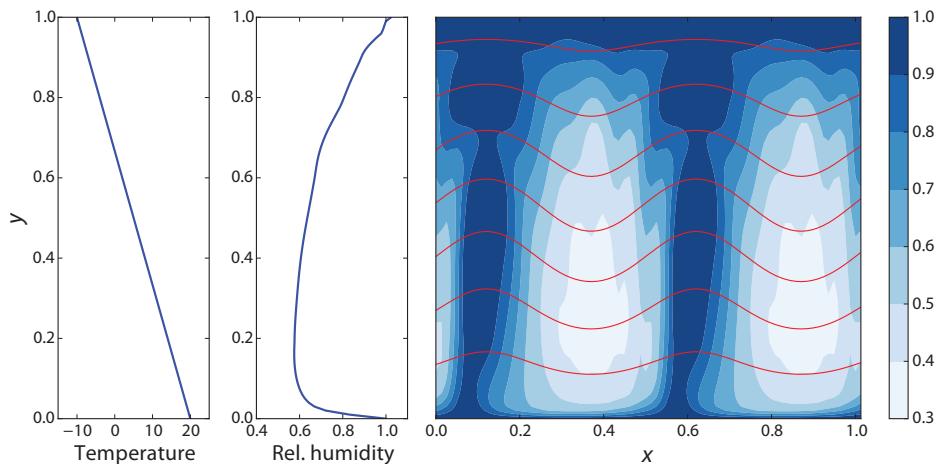


Fig. 18.13 A snapshot of relative humidity (right panel, shading) obtained by a numerical solution of the advective-diffusive-condensation model (18.29) and (18.36) in the horizontal (x - y) plane. The contours are the streamfunction. The middle panel is the zonally-averaged relative humidity and the left panel is the imposed temperature gradient ($^{\circ}$ C).

to a minimum in the interior, rising again at high latitudes (cf., the observations shown in Fig. 18.4). The width of the saturated region near the saturated boundary diminishes as the Peclet number increases, and the mid-domain minimum arises because of the drying effects of advection, which diminish near either boundary.

There are many limitations to such a model, one being that in reality temperature itself is advected by the flow and another being that advection is three-dimensional, but exploring these effects is beyond the scope of our story.⁶ A summary of relative humidity transport is provided in the box on page 683, but we now move on to the dynamical effects of water vapour, and in particular convection.

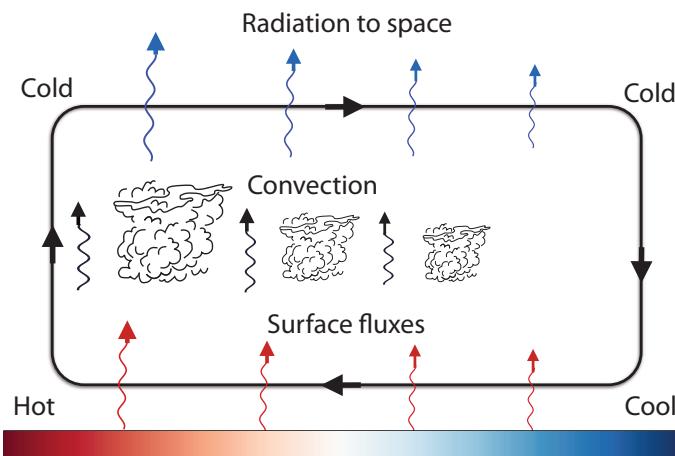


Fig. 18.14 Idealization of small-scale convection embedded within a large scale circulation. The fluid is in contact with a warm surface below, and cools by radiation to space from above. The horizontal temperature gradient drives a large-scale meridional overturning circulation, embedded in which convection forms, especially at the warmer end.

18.3 ATMOSPHERIC CONVECTION

As introduced in Section 2.10, convection occurs when a parcel that is displaced upwards (downwards) finds itself lighter (heavier) than its surroundings, and hence becomes subject to a buoyancy force that amplifies the initial displacement. Convection is particularly prevalent in the tropics for reasons that are, at lowest order, independent of the presence of water vapour; rather, it is simply that the surface is warmer there, as sketched in Fig. 18.14. Consider, rather heuristically, a fluid that is heated by a warm surface from below and cools by radiation from above. If the surface is warmer at one end, as illustrated, a large-scale circulation will arise, transporting energy from warmer regions to cooler ones. In the absence of convection, there will be a larger vertical temperature gradient where the surface temperature is warm, and so these regions are more prone to convection. Having said this, moisture *does* affect convection in rather profound ways, one being that moisture is a *destabilizing* influence, as we now discover.

18.3.1 Generalized Enthalpy and Adiabatic Lapse Rates

A column of air that is convectively stable when dry can be made convectively unstable by the presence of water vapour, because when water vapour condenses heat is released, warming the parcel and potentially making it more buoyant than its surroundings. To explore this we begin with a derivation of the conditions under which a moist column of air may be unstable — essentially an extension of Section 2.10 to include the effects of condensation.

An adiabatic (and convectively neutral) profile is one in which the entropy is constant. Recalling the arguments in Sections 1.6 and 1.10.3, the entropy may be related to the enthalpy by writing the fundamental relation as

$$dh = Td\eta + \alpha dp, \quad (18.37)$$

and where, if the column is hydrostatically balanced, $\alpha dp = -g dz$. Thus, the condition that $d\eta = 0$ is equivalent to

$$d(h + gz) = 0. \quad (18.38)$$

That is, the dry static stability or generalized enthalpy, $h^* = h + gz$, is constant in an adiabatic, hydrostatic, profile. The enthalpy of a dry parcel of ideal gas is given by $h^d = c_p T$ and thus,

$$h^* \equiv h^d + gz = c_p T + gz. \quad (18.39)$$

Since an adiabatic profile has $dh^* = 0$ we recover the dry adiabatic lapse rate (Section 2.10.2), to wit

$$\Gamma^d = -\frac{dT}{dz} \Big|_{ad} = \frac{g}{c_p}. \quad (18.40)$$

The adiabatic lapse rate for a moist parcel is given using similar reasoning, but we need an appropriate expression for the enthalpy. A parcel of total mass M^t composed of dry air, water vapour and liquid water such that $M^t = M^v + M^l$ has an enthalpy given by

$$H = h^d M^d + h^v M^v + h^l M^l, \quad (18.41)$$

where h denotes specific enthalpy and the superscripts denote dry air, water vapour and liquid water respectively. Dividing by the total mass, the specific enthalpy of the parcel is thus

$$h = (1 - q^w)h^d + q^v h^v + q^l h^l, \quad (18.42)$$

where

$$q^w = q^v + q^l, \quad q^v = \frac{M^v}{M^t}, \quad q^l = \frac{M^l}{M^t}. \quad (18.43)$$

As in (18.16) the latent heat of evaporation is defined to be $L = h^v - h^l$ so that (18.42) becomes

$$h = (1 - q^w)h^d + q^w h^l + L q^v. \quad (18.44)$$

The generalized enthalpy is equal to this quantity plus a potential ϕ , and if that potential is equal to gz the quantity is the *moist static energy*,

$$h_m^* = (1 - q^w)h^d + q^w h^l + L q^v + gz. \quad (18.45)$$

Since $h^d = c_p^d T$ and $h^l = c_l^l T$ (to a very good approximation and where c_l^l is the heat capacity of liquid water) we have

$$h_m^* = c_p^{dl} T + L q^v + gz, \quad (18.46)$$

where $c_p^{dl} = (1 - q^w)c_p^d + q^w c_l^l$. This quantity varies with liquid water content but in Earth's atmosphere the variation is small and in the derivation below we will take c_p^{dl} to be a constant and denote it c_p (its value is very similar to that of c_p^d). For adiabatic motion, the moist static energy is a constant, whether or not water is evaporating or condensing: the latent heat of evaporation or condensation is merely exchanged with the dry generalized enthalpy.

The saturated adiabatic lapse rate

In a moist atmosphere the moist static energy is conserved as a parcel ascends and $dh_m^*/dz = 0$. Using (18.46), an ascending parcel then has a lapse rate given by

$$c_p \frac{dT}{dz} = -L \frac{dq^v}{dz} - g. \quad (18.47)$$

If the air is moist but not saturated then an ascending parcel will follow the dry adiabatic lapse rate (because $dq^v/dz = 0$) but if it is saturated (and $q^v = q_s$) then as a parcel ascends it will cool and some water vapour will condense. Since $q_s \approx \epsilon e_s / p$, and so is a function of temperature and pressure, we have

$$\frac{dq_s}{dz} = \left(\frac{\partial q_s}{\partial T} \right)_p \frac{dT}{dz} + \left(\frac{\partial q_s}{\partial p} \right)_T \frac{dp}{dz} = \left(\frac{\partial q_s}{\partial T} \right)_p \frac{dT}{dz} + \left(\frac{q_s}{p} \right) \rho g, \quad (18.48)$$

using hydrostatics. Using the Clausius–Clapeyron equation and the ideal gas equation of state gives

$$\frac{dq_s}{dz} = \frac{L q_s}{R^v T^2} \frac{dT}{dz} + \frac{g q_s}{R^d T}. \quad (18.49)$$

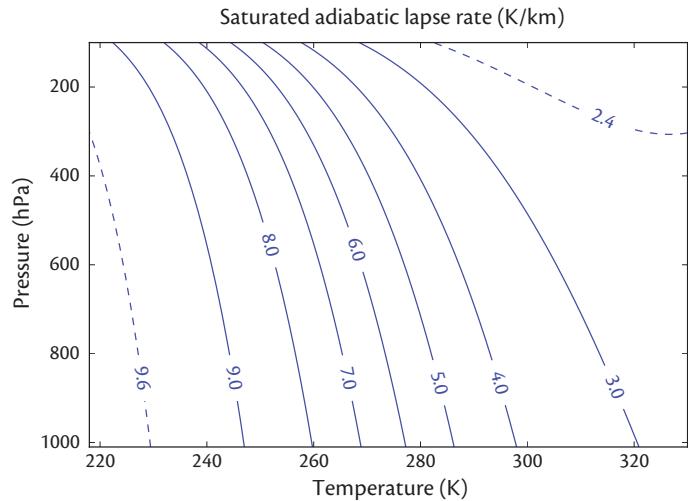


Fig. 18.15 Contours of saturated adiabatic lapse rate (K/km) as a function of pressure and temperature, calculated using (18.50) with $q_s = ee_s/p$.

Using (18.47) and (18.49) we obtain an expression for the lapse rate of an adiabatically ascending saturated parcel,

$$\Gamma_s = -\left.\frac{dT}{dz}\right|_{ad} = \frac{g}{c_p} \frac{1 + Lq_s/(R^d T)}{1 + L^2 q_s/(c_p R^v T^2)}. \quad (18.50)$$

The quantity Γ_s is the *saturated adiabatic lapse rate*, plotted in Fig. 18.15. (This quantity is often called the moist adiabatic lapse rate, but that name is better given to the lapse rate of moist air that does not condense, and which differs slightly from the dry adiabatic lapse rate. In fact Γ_s is properly a *pseudo-adiabatic* lapse rate, because liquid water is not accounted for — see Appendix A.) The lapse rate is a function of temperature and pressure because $q_s = ee_s/p$ and e_s is given by the solution of the Clausius–Clapeyron equation, (18.28). Values of Γ_s are typically around 6 K km^{-1} in the lower atmosphere although since dq_s/dT is an increasing function of T , Γ_s decreases with increasing temperature and can be as low as 3 K km^{-1} . The second term in the numerator of (18.50) is usually quite small (around 0.1, although it can become large at very high temperatures) but the second term in the denominator is positive and order unity. Since g/c_p is the dry adiabatic lapse rate, the saturated adiabatic lapse rate is smaller than the dry adiabatic lapse rate, as can be seen directly from (18.47) since $dq^v/dz < 0$.

The saturated adiabatic lapse rate determines the stability of a saturated profile. Using parcel theory, just as in Section 2.10, a profile will be stable or unstable depending on whether the lapse rate is less than or greater than (18.50); that is

$$\text{Stability : } -\frac{\partial \tilde{T}}{\partial z} < \Gamma_s, \quad \text{Instability : } -\frac{\partial \tilde{T}}{\partial z} > \Gamma_s, \quad (18.51a,b)$$

where \tilde{T} is the environmental temperature. The lapse of the atmosphere rarely exceeds the saturated adiabatic lapse rate, as was seen in Fig. 15.25. In the tropics and subtropics the lapse is, on average, very close to the saturated adiabat up to about 300 hPa (about 9 km), whereas in mid-latitudes it is considerably less, and so more stable, because of the upward transport of heat by baroclinic eddies. Convection only directly determines the lapse rate over a small fraction of the tropics where convection actually occurs, but nevertheless the average lapse rate is close to moist neutral. This is because the gravity waves emanating from convective regions adjust the tropics to have weak horizontal temperature gradient, in a process akin to geostrophic adjustment, hence maintaining approximately the same vertical profile even away from regions of active convection, as discussed more in Sections 18.8 and 18.9.

18.3.2 Equivalent Potential Temperature

In a dry atmosphere adiabatic motion is characterized by the material conservation of potential temperature, θ , a surrogate for entropy. Thus in adiabatic flow $D\theta/Dt = 0$, and the dry adiabatic lapse rate can be characterized equivalently either as $dT/dz = -g/c_p$ or $d\theta/dz = 0$. We can construct a similar quantity for moist air. Suppose that a parcel is lifted, and so cooled, until all its moisture condenses, and that all the latent heat released goes into heating the parcel. The *equivalent potential temperature*, θ_{eq} , is the potential temperature that the parcel then achieves.⁷ If the parcel is then moved along a dry adiabat to a reference pressure p_R (commonly 1000 hPa) the actual temperature it will then have is θ_{eq} . As a consequence of the near adiabatic nature of the process, θ_{eq} is an approximate measure of the entropy of the parcel, as we will show (see also Appendix A).

We may obtain an approximate analytic expression for θ_{eq} by noting that the first law of thermodynamics, $dQ = T d\eta$, implies, by definition of potential temperature,

$$-L dq = c_p T d \ln \theta, \quad (18.52)$$

during the condensation process, where dq is the change in water vapour content. Integrating gives, by definition of equivalent potential temperature,

$$-\int_q^0 \frac{L}{c_p T} dq = \int_\theta^{\theta_{eq}} d \ln \theta. \quad (18.53)$$

Here, q is the initial amount of water vapour contained in the parcel and T is the temperature at which condensation occurs. We might imagine lifting a parcel from its initial position to a temperature T at which saturation first occurs. To remove *all* the water vapour the parcel must be lifted to great height, because all of the water vapour will only condense if the final temperature is very low, with condensation occurring continuously and so with varying temperature along the way. But if we assume that temperature is constant during condensation, and that L and c_p are also constants, then (18.53) gives

$$\theta_{eq} = \theta \exp \left(\frac{Lq}{c_p T} \right) = T \left(\frac{p_R}{p} \right)^{R/c_p} \exp \left(\frac{Lq}{c_p T} \right). \quad (18.54)$$

The equivalent potential temperature so defined is approximately conserved during condensation, the approximation arising going from (18.53) to (18.54). It is a useful expression for diagnostic purposes and in constructing theories of convection, but it is not accurate enough to use as a primary prognostic variable in a numerical model that aims to be realistic. In saturated adiabatic flow a parcel will follow a ‘moist isentrope’ (an isoline of equivalent potential temperature) more closely than a dry isentrope.

To see that the equivalent potential temperature is an approximate measure of the entropy of a saturated parcel we begin with the entropy of dry air, which referring to (1.107) is

$$\eta^d = c_p \ln \theta = c_p \ln(T/T_0) - R \ln(p/p_0), \quad (18.55)$$

where T_0 and p_0 are constants. If we add to this a contribution from liquid water (see Appendix A on page 720 for a more complete treatment) then the saturation entropy, η_s , is approximately given by

$$\eta_s \approx \eta^d + \frac{Lq_s}{T} = c_p \ln(T/T_0) - R \ln(p/p_0) + \frac{Lq_s}{T}, \quad (18.56)$$

where q_s is the saturation specific humidity and final liquid water content. If we now define θ_{eq} by

$$c_p \ln(\theta_{eq}/T_0) \equiv \eta_s, \quad (18.57)$$

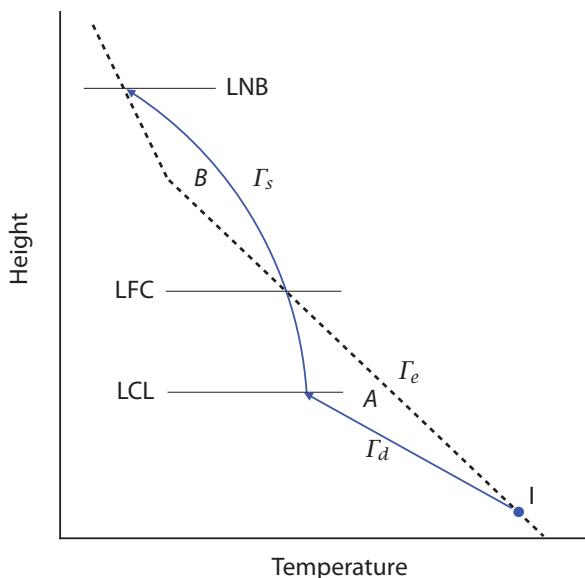


Fig. 18.16 Schematic of a conditional instability in an atmosphere with environmental lapse rate Γ_e (dashed line). A parcel at I is forced to rise, and it does so along the dry adiabat, Γ_d , until it is saturated at the lifting condensation level, LCL. It will then rise along the saturated adiabat, Γ_s , and after reaching the level of free convection (LFC) it is convectively unstable. The parcel continues to rise along the saturated adiabat, without any external forcing, until it reaches the level of neutral buoyancy (LNB). Not to scale.

then using (18.56) and (18.57) we recover (18.54) with $q = q_s$ — meaning that the logarithm of the equivalent potential temperature is, approximately, the entropy. To relate this result to the lapse rate, consider a parcel ascending into a colder region and condensing. The saturation moist entropy will stay the same: the last term on the right-hand side of (18.56) will fall, but the ensuing heating causes the temperature to increase. If we differentiate (18.56) with respect to z and set $\partial\eta_s/\partial z = 0$ then, using hydrostasy and a little algebra, we recover the saturated adiabatic lapse rate given by (18.50), which, therefore, has constant θ_{eq} and constant η_s .

18.4 CONVECTION IN A MOIST ATMOSPHERE

Over a wide range of temperatures and pressures the saturated adiabatic lapse rate is considerably less than the dry lapse rate (Fig. 18.15), and thus a moist atmosphere can be stable to dry convection but unstable to moist convection. If an environmental profile is unstable then convection ensues, transporting energy upward, until the atmosphere nearly stabilizes. But if the atmosphere is not saturated it is the dry adiabatic lapse rate that is the relevant one (or, strictly, the adiabatic lapse rate of moist, non-condensing air, which is very slightly different). The observed environmental profile in convecting situations may thus be a combination of the dry adiabatic and saturated adiabatic profiles. An unsaturated parcel that is unstable by the dry criterion will rise and cool following a dry adiabat, Γ_d , until it becomes saturated, above which it will rise following a saturated adiabat, Γ_s . The temperature then does not fall as rapidly as the dry case because the latent heat release warms the parcel as it rises.

These facts give moist convection a particular flavour and lead to the notion of *conditional stability*, whereby a parcel is stable to an infinitesimal perturbation but unstable to a finite perturbation. Consider an environment in which the lapse rate lies between the moist and dry rates, as in Fig. 18.16, and consider a parcel near the surface at position I. Suppose the parcel is adiabatically lifted (perhaps mechanically) then its temperature profile follows the dry adiabat until it is saturated at, by definition, the *lifting condensation level* (LCL). However, the parcel is actually negatively buoyant and so would sink unless the parcel is forced to continue rising, but if it does continue to rise it will be along a saturated adiabat and eventually, at the *level of free convection* (LFC) it will become buoyant and convectively unstable. It will continue to rise until its buoyancy no longer exceeds that of the environment, at the *level of neutral buoyancy*, which may well be at or close to

the tropopause where the temperature starts to increase again. (The height of the tropopause is not independent of the convection itself, a matter we discuss in Section 18.6.) Evidently, to trigger such an instability a *finite* perturbation is needed (for example a large scale flow over a hill forcing a parcel upwards) and this is known as *conditional instability*. A parcel that is initially sitting at the LCL will only be unstable if it is saturated, which is by no means usually the case: the water has to get there somehow.⁸

Convection in the atmosphere need not be ‘conditional’, in the sense described above, nor is the atmosphere necessarily always in a conditionally unstable state. Indeed, on average the tropical atmosphere is closer to a moist profile than a dry profile. It is also important to remember that convection ultimately arises because the radiative forcing produces a vertical profile that is convectively unstable, with large-scale horizontal temperature gradients confining the convection to warmer regions, as in Fig. 18.14.

18.4.1 Energetics of Convection

How much energy is available for a parcel in convection? One way is to calculate the work released by, or required for, a parcel as it moves through an environment that has a different density. The upward force per unit mass on a parcel of density ρ_p in an environment with density ρ_e is just the familiar buoyancy force,

$$F_b = g \frac{\rho_e - \rho_p}{\rho_p} = g \frac{\alpha_p - \alpha_e}{\alpha_e}. \quad (18.58)$$

Thus, the energy released as a parcel ascends from z_1 to z_2 is

$$\text{Energy} = \int_{z_1}^{z_2} g \frac{\alpha_p - \alpha_e}{\alpha_e} dz = - \int_{p_1}^{p_2} (\alpha_p - \alpha_e) dp, \quad (18.59)$$

if the environment is in hydrostatic balance. If we take the limits of integration to be the level of free convection and the level on neutral buoyancy, the energy is known as the *convective available potential energy*, or CAPE. Using the ideal gas relation, and assuming that the parcel has the same pressure as the environment, (18.59) becomes

$$\text{CAPE} = -R \int_{\text{LFC}}^{\text{LNB}} (T_p - T_e) \frac{dp}{p}. \quad (18.60)$$

If the height axis in Fig. 18.16 were log pressure then the CAPE would be proportional to the area B . If the limits were taken from the initial height to the LFC then the integral would be proportional to the area A and would be negative, and is known as the ‘convective inhibition’, or CIN. That is

$$\text{CIN} = -R \int_{p_{bot}}^{\text{LFC}} (T_p - T_e) \frac{dp}{p}, \quad (18.61)$$

where p_{bot} is the pressure at the bottom or in a boundary layer, and CIN is the amount of energy that must be supplied to initiate convection. In some accounts the definition of CAPE includes the CIN, in which case CAPE may be negative (but this is unusual). It is often the case that observed profiles of temperature in the tropics exhibit both CIN and CAPE, and a parcel in the lower atmosphere is stable and the instability is only conditional. That a finite perturbation may be needed to initiate convection is a distinguishing feature of moist convection.

Given the initiation of convection, one may envision an unstable parcel of air experiencing a buoyant force and accelerating through the unstable region, gaining kinetic energy, and this energy

can be quite significant. Suppose that $T_p - T_e = 1^\circ\text{C}$ and that the limits of integration are 250 hPa and 750 hPa. Then we estimate

$$\text{CAPE} \sim R(T_p - T_e) \frac{\Delta p}{p} = R \times 1 \times \frac{750 - 250}{500} = 286 \text{ J kg}^{-1}. \quad (18.62)$$

This is actually an underestimate, by a factor of a few, of the value of CAPE often found in tropical atmospheres, but nevertheless if we translate it to a vertical velocity using $w = 2\sqrt{\text{CAPE}}$ then we find $w \approx 24 \text{ m s}^{-1}$! In fact, CAPE is a significant overestimate of the kinetic energy that is imparted to a buoyant parcel: a parcel will not keep its identity because of mixing or entrainment with its surroundings and the environmental profile itself becomes altered, and even if the parcel did keep its identity it would give up some of its kinetic energy to the environment in pushing it out of the way. Nevertheless, the estimate gives the sense that atmospheric convection is a vigorous process, occurring on a fast timescale compared to large-scale dynamics. Let us suppose that the above estimate is too big by a factor of 10. The time taken for a parcel to travel half the height of the troposphere in the tropics is then approximately $7500 \text{ m} / 2.4 \text{ m s}^{-1} = 3000 \text{ s}$, or less than one hour. *Thus, the timescales of atmospheric convection are measured in hours, not days.*

18.4.2 Effects of Convection

We have seen that, with even small differences in between an environmental profile and a stable profile, convection will be vigorous and act on rather fast timescales. What will the result of that convection be? Parcels will ascend until they are no longer unstable, at which point they can mix irreversibly with their environment, warming the environment aloft. (Indeed the parcels will mix as they ascend, but they have more time to mix as the ascent slows.) The convection and the mixing will proceed until the environmental profile is no longer unstable, a process that, if it occurs on a very short timescale, is called *convective adjustment* or, if on a longer timescale, *convective relaxation*. One might suppose that the profile to which the environment will adjust will be the saturated adiabatic or dry adiabatic lapse rates, depending on whether the atmosphere is saturated or not, although this is something of an oversimplification.⁹ The convection will also transfer moisture vertically, and if this transfer causes the profile to become saturated then condensation and precipitation may occur.

Let us denote the adjusted reference profiles of temperature and specific humidity as $T_r(z)$ and $q_r(z)$ (or equivalents in pressure). If the vertical structure of these profiles is given, the magnitude of the profile will be determined by the fact that convection will conserve enthalpy. If, for example, atmosphere is far from saturated then the reference profiles will satisfy

$$\int_{p_B}^{p_T} c_p (T_r - T_i) dp = 0 \quad \text{and} \quad \int_{p_B}^{p_T} (q_r - q_i) dp = 0, \quad (18.63)$$

where the subscript i denotes the initial environmental profile and p_B and p_T denote the levels at the base and top of the convection. If precipitation does occur then only the total enthalpy is conserved and we have

$$\int_{p_B}^{p_T} (h_r - h_i) dp = 0, \quad (18.64)$$

where $h = c_p T + Lq$ is the (moist) specific enthalpy. One then needs a second constraint to determine the profiles of T_r and q_r separately.

The specification of these profiles, and the levels of the base and the top of the convection, is not an exact science since the notion of adjustment is an approximation. As a first estimate in a moist atmosphere one might suppose that T_r is the dry adiabat below saturation and the saturated adiabatic lapse rate above, and that the reference profile of q_r is such that the atmosphere remains saturated, or nearly so, after adjustment, and that p_B and p_T are the levels of free convection and of

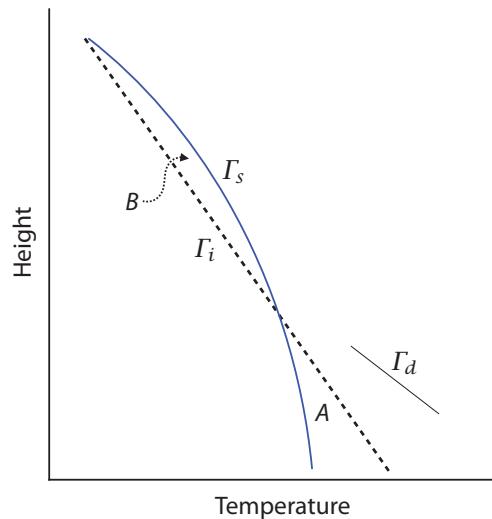


Fig. 18.17 Schematic of convective adjustment in a saturated atmosphere. The initial environment, with profile Γ_i , is unstable to moist convection and adjusts to the saturated adiabatic lapse rate Γ_s , with the dry adiabat Γ_d shown for reference. The average temperature of the final profile is such that the enthalpy is conserved, which provides a relationship between the areas A and B.

neutral buoyancy, respectively. Application of (18.64) then uniquely determines the final profiles of temperature and humidity, as schematically illustrated for a saturated atmosphere in Fig. 18.17. Details apart, two robust, key points should be emphasized:

- (i) To a first approximation, convection determines the *profiles* of temperature and humidity rather than the actual heating rate or the precipitation. The heating and precipitation released by convection are a function of the environmental profile that in turn is largely determined by the larger-scale processes, although convection feeds back onto this profile.
- (ii) After convection has occurred the local profile is, by construction, stable. However, this profile is not necessarily in equilibrium with the large-scale motion, or with the radiative processes that initially gave rise to the unstable profile. This disequilibrium may cause the profile to evolve further, possibly giving rise to more convection.

The above state of affairs — a putative initial profile set up by radiation and/or large-scale dynamics, modified quickly by convection, leading to a further slow evolution and convection — is called *quasi-equilibrium*.

18.4.3 † Convective Quasi-Equilibrium

Convective quasi-equilibrium is a posited state in which the forcing of a convectively unstable profile by large-scale dynamics and/or radiation is statistically balanced by convection.¹⁰ The large-scale forcing may change, and the system may evolve, but in quasi-equilibrium the convection is assumed to occur on a faster timescale than the other processes and the fluid quickly enters into a statistical equilibrium state. Even as the large-scale evolves, the vertical profile at each step is largely determined by convection. That such a state exists is an assumption, but in Earth's tropical atmosphere it seems a fairly good one when dealing with timescales of longer than a few days, because convective timescales are relatively short and so convection is able to significantly alter the stratification. The assumption of quasi-equilibrium is marginally satisfied over the diurnal cycle.

The most obvious consequence of quasi-equilibrium is that the temperature profile is constrained to be close to being neutrally stable, which generally means a saturated-adiabatic profile except possibly in some regions in the lower atmosphere where the appropriate profile may be dry adiabatic, and in the boundary layer where properties are well mixed. A second important consequence concerns CAPE, convective available potential energy. Quasi-equilibrium implies that the

Convection, Quasi-Equilibrium and Radiative-Convective Equilibrium

- Atmospheric convection ultimately arises because the radiative forcing tends to produce temperature profiles, and hence buoyancy profiles, that are statically unstable. Large-scale advection, moving cold air over warm air, also leads to convection.
- Convection is most prevalent in the tropics because it is here that the surface is warmest, and the radiative-equilibrium lapse rate would be most unstable (Section 18.5). This would be true even in the absence of moisture, but moisture is most important in the tropics because it is warmest, and because water is imported via the Hadley Cell.
- The presence of moisture may cause the critical lapse rate for convection to be lower than that for dry air, because of the release of heat when water vapour condenses. The effect only arises if a parcel is saturated. Thus, a column of air may be *conditionally unstable* if it requires a finite perturbation to raise a parcel to a level where it saturates (Fig. 18.16).
- Atmospheric convection tends to occur on much faster timescales than those associated with the large-scale circulation. Once convection occurs the lapse rate is closely constrained to the critical lapse rate in the regions of convection.
- In a process analogous to geostrophic adjustment, internal waves propagating away from convection maintain weak horizontal temperature gradients and ensure that the lapse rate does not deviate too far from saturated adiabatic anywhere in the tropics (Fig. 15.25). In mid-latitudes, in contrast, energy is also transported upward by large-scale baroclinic eddies and the lapse rate is lower (more stable).
- When the large-scale dynamical or radiative forcing changes, the convective flux changes to constrain the lapse rate, and the production of CAPE by the large scale is *approximately* balanced by the relaxation of CAPE by convection. This state of affairs is called quasi-equilibrium, and its occurrence is really a hypothesis that appears to be fairly well-satisfied, if the larger timescales are sufficiently long, in Earth's tropical atmosphere.
- Quasi-equilibrium, if and where closely satisfied, leads to considerable simplifications regarding the dynamics of large-scale flow for it de-emphasizes the need to explicitly calculate vertical energy fluxes because the lapse rate is constrained. Equations resembling the shallow-water equations, with a small number of vertical modes, then become a decent approximation for large-scale baroclinic flow, with diabatic effects forcing the first baroclinic mode.
- A radiative-convective equilibrium state arises when the radiative-equilibrium state is convectively unstable. Convection then balances the radiative forcing, producing at first approximation a convectively active state up to some finite height, with a radiative equilibrium state beyond (Fig. 18.18).
- The height to which the constrained lapse rate extends is determined by the need to maintain an overall radiative balance — the incoming solar radiation must equal the outgoing infrared radiation — and this is the leading-order determinant of the tropopause height (Section 18.6.2). A similar calculation could be applied in mid-latitudes, with a lapse rate determined as much or more by baroclinic instability as by convection.
- Quasi-equilibrium is a useful idealization when convection occurs quickly and efficiently, but it does not mean that the large scale always controls the convection. There are many situations when there is no causal separation between convection and the large scale — indeed the convection may create its own large scale flow, the MJO being one example.

CAPE is released by convection at about the same rate as it is generated by the forcing. Thus, at the same time as radiative forcing may be creating an unstable profile with a finite amount of CAPE, convection is destroying that potential energy, largely by converting it to kinetic energy where it is dissipated. The amount of CAPE in a profile may (and does) vary on slow timescales, the variation being produced by the weak imbalance between production and destruction; observations indicate it can vary by a factor of a few over the course of many days, and the environmental profile may also depart from the saturated adiabat away from convectively active regions. However, it turns out that the *average* stratification over the tropics will not deviate too much from moist neutral because internal waves extend the moist-neutral profile in the convective regions to the cloud-free regions, and horizontal pressure and temperature gradients are constrained to be relatively small. This ‘weak temperature gradient’ effect is described in Section 18.8.

Finally, albeit a little less tangibly, quasi-equilibrium gives rise to a *point of view* of tropical dynamics, that convection does not act as a primary determinant of the heat source for the large-scale circulation. Rather, convection controls the temperature and water vapour profiles, with the heating associated with the convection being determined by the requirement that such a profile be maintained, and thus ultimately arising from the larger scale forcing. (This process is analogous to the production of condensation in the fast-condensation models of Section 18.2, in which the condensation criteria determine the profile of water vapour but not the amount of condensation.) The resulting environmental profile then lies on the edge of stability, with a non-zero CAPE because of the finite time-scale and finite efficiency of convection and with the intensity of convection being just what is required to maintain that profile. The value of CAPE then evolves slowly, compared to convective timescales, if and as external factors change.

Having said this, quasi-equilibrium should not be thought of as being exact or even especially profound — it is a way to gain an understanding of the structure of the tropics but is not a universal recipe. There are many situations — the diurnal cycle, aspects of boundary layer development, weak convection — in which the quasi-equilibrium assumptions do not hold. Furthermore, the recipe may mislead if misapplied: even when quasi-equilibrium holds in so far as the buoyancy profile is constrained to be moist neutral, it should not be thought that convection and the associated condensation is always a quasi-passive process. That is, the scale separation between convective events and the large-scale does not mean that the large-scale always ‘controls’ the convection and the convection merely ‘feeds back’ on the large scale. For example, a localised source of convection near the equator may itself create a large-scale flow pattern similar to that of the Matsuno–Gill problem (Section 8.5), and this effect may be at the heart of the Madden–Julian oscillation (Section 18.10). Here one might say the convection controls the large-scale and the large-scale feeds back on the convection! Also, in the moist model of the Hadley Cell (Section 14.2.7) the release of latent heat in the upward branch serves to change the horizontal distribution of heating and greatly intensify the overturning circulation. The convection should not be thought of as the primary driver of the Hadley Cell, but it has a lowest-order effect.

Caveats aside, quasi-equilibrium is a very useful concept and in following sections we look at two particularly important consequences. First (albeit after an introduction to radiative equilibrium) we discuss radiative-convective equilibrium, and we then use quasi-equilibrium to simplify the equations of motion for the larger-scale circulation.

18.5 RADIATIVE EQUILIBRIUM

In order to understand the effects of convection we must first determine what the profile of temperature would be in its absence — the *radiative-equilibrium* state. The electromagnetic radiation in the Earth’s atmosphere may usefully be divided into two types, solar (or shortwave) radiation and infrared (or longwave) radiation. We will assume the atmosphere is semi-grey (i.e., grey in the infrared) and, in some instances, transparent to solar radiation. Neither assumption is quantitatively good, but they capture the essence. In Appendix B to this chapter we show that the upward,

U , and downward, D , streams of longwave radiation then satisfy the radiative-transfer equations,

$$\frac{dD_L}{d\tau} = B - D_L, \quad \frac{dU_L}{d\tau} = U_L - B. \quad (18.65a,b)$$

Here, $B = \sigma T^4$ is the radiative flux emitted by a black body, τ is the optical depth and σ is Stefan's constant. The optical depth is related to the geometric height z by $d\tau = -e_L dz$ where e_L is the longwave emissivity. The net flux of longwave radiation is $N_L = U_L - D_L$ and the longwave heating is proportional to the net flux divergence, $-\partial N_L / \partial z$.

18.5.1 Solutions

Formal solution to radiative transfer equations

If the temperature profile is known then B is known and (18.65) is a pair of first order differential equations in the two unknowns U and D . The solution requires two boundary conditions and in atmospheric problems these might be provided at the top of the atmosphere, for example by requiring that the downward infrared radiation be zero (i.e., $D_L(\tau = 0)$), and by specifying the upward radiation (i.e., let $U_L(\tau = 0) = U_0$ where U_0 is given). Alternatively, the upward radiation at the bottom of the atmosphere could be specified if the ground temperature were known. The solution in physical space is found by specifying the form of $\tau(z)$; that is, by specifying the emissivity. To obtain the solution we multiply (18.65) by the integrating factors $\exp(\tau)$ and $\exp(-\tau)$ to give

$$\frac{d}{d\tau}(D_L e^\tau) = B e^\tau, \quad \frac{d}{d\tau}(U_L e^{-\tau}) = -B e^{-\tau}. \quad (18.66a,b)$$

Integrating between 0 and τ' straightforwardly gives

$$D_L(\tau') = e^{-\tau'} \left[D_L(0) - \int_0^{\tau'} B(\tau) e^\tau d\tau \right], \quad U_L(0) = U_L(\tau') e^{-\tau'} + \int_0^{\tau'} B(\tau) e^{-\tau} d\tau. \quad (18.67a,b)$$

There are other ways to write the solutions that may be appropriate depending on the boundary conditions, but in any case the solutions are in general *non-local*, for they depend on the temperature along the path. The terms in (18.67) represent the attenuation of radiation as it travels along its path, as well as the cumulative emission. However, in some important special cases we can get a local solution, as we now see.

Radiative equilibrium solution

A *radiative equilibrium* state has, by definition, no radiative heating. If the atmosphere is transparent to solar radiation then the condition implies that the vertical divergence of the longwave radiation is zero:

$$\frac{\partial(U_L - D_L)}{\partial z} = 0 \quad \text{implying} \quad \frac{\partial(U_L - D_L)}{\partial \tau} = 0. \quad (18.68a,b)$$

This condition is normally *not* satisfied in the atmosphere because the air is in motion. If it were satisfied then (18.65) and (18.68b) form three equations in three unknowns, U_L , D_L , and B , and a solution can be found as follows.

Consider an atmosphere with net incoming solar radiation S_{net} and suppose the planet is in radiative equilibrium with the incoming solar radiation balanced by outgoing infrared radiation. That is, $U_{Lt} \equiv U_L(\tau = 0) = S_{net}$ where U_{Lt} is the net outgoing longwave radiation (OLR) at the top of the atmosphere. The downward infrared radiation at the top of the atmosphere is zero, so that the boundary conditions on the radiative transfer equations at the top of the atmosphere are

$$D_L = 0, \quad U_L = U_{Lt} \quad \text{at} \quad \tau = 0. \quad (18.69)$$

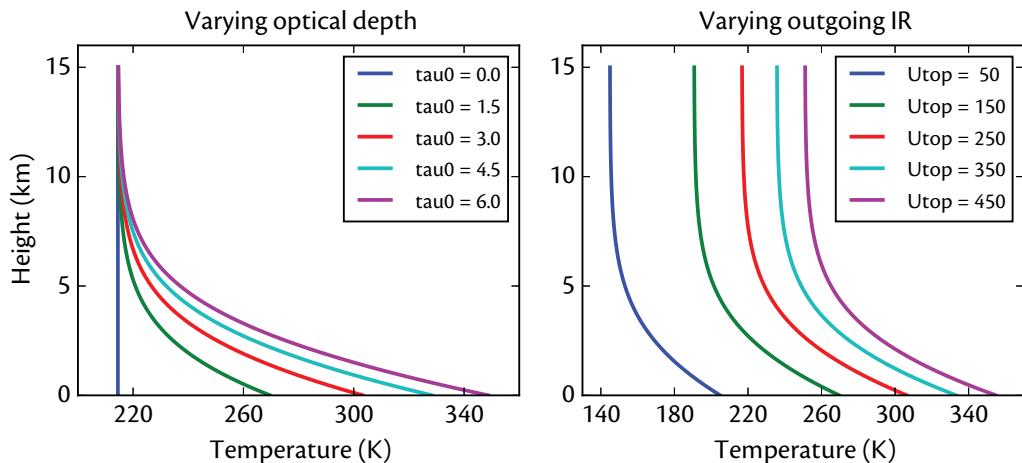


Fig. 18.18 Radiative equilibrium temperature calculated using (18.73), with $H_a = 2 \text{ km}$. Left: outgoing IR (and incoming solar) radiation is 240 W m^{-2} and surface optical depth varies from 0 to 6. Right: the optical depth is 3.0 and outgoing IR radiation varies from 50 to 450 W m^{-2} .

To obtain a solution we rewrite (18.65) as

$$\frac{\partial}{\partial \tau}(U_L - D_L) = U_L + D_L - 2B, \quad \frac{\partial}{\partial \tau}(U_L + D_L) = U_L - D_L. \quad (18.70a,b)$$

and a little algebra reveals that a solution of these equations that satisfies (18.68b) is

$$D_L = \frac{\tau}{2} U_{Lt}, \quad U_L = \left(1 + \frac{\tau}{2}\right) U_{Lt}, \quad B = \left(\frac{1 + \tau}{2}\right) U_{Lt}, \quad (18.71a,b,c)$$

as can be easily verified by substitution back into the equations.

It remains to explicitly relate τ to z , and one approximate recipe is to suppose that τ has an exponential profile,

$$\tau(z) = \tau_0 \exp(-z/H_a), \quad (18.72)$$

where τ_0 is the optical depth at $z = 0$ and H_a is the scale height of the absorber. In the Earth's atmosphere the optical depth is determined by the concentrations of water vapour (primarily) and carbon dioxide (secondarily) and τ_0 (the scaled optical depth) typically varies between 2 and 4, depending on the water vapour content of the atmosphere, and $H_a \approx 2 \text{ km}$, this being a typical scale height for water vapour. From (18.71c) the temperature then varies as

$$T^4 = U_{Lt} \left(\frac{1 + \tau_0 e^{-z/H_a}}{2\sigma} \right), \quad (18.73)$$

which is illustrated in Fig. 18.18. The following aspects of the solution deserve mention:

- (i) Temperature increases rapidly with height near the ground.
- (ii) The upper atmosphere, where τ is small, is nearly isothermal.
- (iii) The temperature at the top of the atmosphere, T_t is given by

$$\sigma T_t^4 = \frac{U_{Lt}}{2}. \quad (18.74)$$

Thus, if we define the emitting temperature, T_e , to be such that $\sigma T_e^4 = U_{Lt}$, then $T_t = T_e/2^{1/4} < T_e$; that is, the temperature at the top of the atmosphere is *lower* than the emitting temperature.

- (iv) Related to the previous point, $B_t/U_{Lt} = 1/2$. That is, the upwards long wave flux at the top of the atmosphere is twice that which would arise if there were a black surface at a temperature T_t . The reason is that there is radiation coming from all heights in the atmosphere.

Interestingly, in obtaining a solution we have not imposed a boundary condition at the ground — in fact there is no ground at all in this problem! What happens if we add one? That is, suppose that we declare that there is a black surface at some height, say $z = 0$, and we require that the atmosphere remain in radiative equilibrium with the same temperature profile. What temperature does the ground take? From (18.71) the upward irradiance and temperature at any height z are related by

$$U_L(z) = \left(\frac{2 + \tau(z)}{1 + \tau(z)} \right) \sigma T^4(z). \quad (18.75)$$

At $z = 0$ the ground will have to supply upwards radiation equal to that given by (18.75), and therefore its temperature, T_g is given by

$$\sigma T_g^4 = \left(\frac{2 + \tau_0}{1 + \tau_0} \right) \sigma T_s^4, \quad (18.76)$$

where T_s is the temperature of the fluid adjacent to the ground (the ‘surface temperature’). That is, $T_g > T_s$ and there is a temperature discontinuity at the ground, especially if optical depth is small. In fact, in very still and clear conditions a very rapid change of temperature near the ground can sometimes be observed, but usually the presence of conduction and convection, as we discuss below, ensures that T_g and T_s are equal.

We note that in the limit in which $\tau = 0$ in the upper atmosphere we have

$$D_L = 0, \quad U_L = U_{Lt}, \quad B = \frac{U_{Lt}}{2}. \quad (18.77)$$

That is, the upper atmosphere is isothermal, there is no downwelling irradiance and the upward flux is constant. The upper-atmosphere temperature, T_{ua} say, and the emitting temperature are related by $T_{ua} = T_e/2^{1/4}$.

The above results imply that at a given optical depth, the temperature difference between the surface and the ground increases with temperature. From (18.73) we have that

$$T_t^4 = \frac{U_{Lt}}{2\sigma}, \quad T_s^4 = \frac{U_{Lt}(1 + \tau_0)}{2\sigma}, \quad (18.78)$$

whence

$$T_s^4 - T_t^4 = T_t^4 \tau_0 \quad \text{or} \quad \Delta T \approx \frac{\tau_0}{4} T_t, \quad (18.79)$$

where $\Delta T = T_s - T_t$ and we assume $\Delta T \ll T_s, T_t$. Thus, in the absence of other effects, higher temperatures lead to larger average lapse rates, and so are potentially more conducive to convection.

18.6 RADIATIVE-CONVECTIVE EQUILIBRIUM

In the radiative equilibrium solution the temperature gradient near the ground varies so rapidly that $-\partial T/\partial z$ may exceed even the dry adiabatic lapse rate. If so, the radiative equilibrium solution is convectively unstable and convection will be triggered and energy (and moisture) will be redistributed throughout the column. What is the resulting temperature profile? The answer is given in Fig. 18.19, which we now explain.

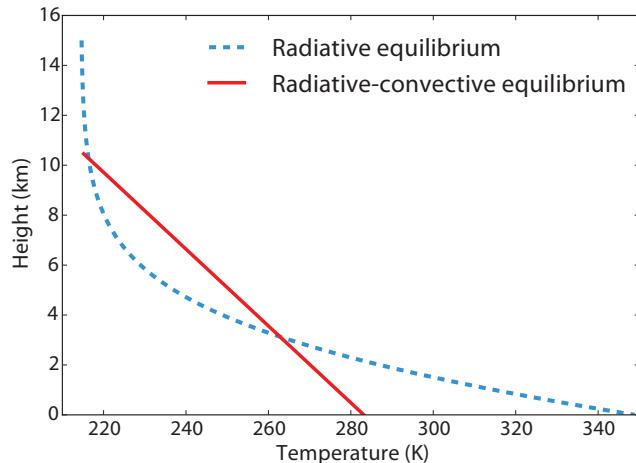


Fig. 18.19 Radiative and radiative-convective equilibrium profiles. The initial radiative equilibrium temperature adjusts to a specified profile (in this instance a constant lapse rate) that extends to a finite height, beyond which radiative equilibrium holds. This height (the tropopause) is determined by the requirement of radiative balance, and the overall adjustment process is not adiabatic. The curves are results of numerical calculations with $\tau_0 = 6$, $\Gamma = 6.5 \text{ K km}^{-1}$ and outgoing radiation of 240 W m^{-2} (see Section 18.6.2).

18.6.1 ♦ The General Case

We first approach the problem without making any simplifying quasi-equilibrium assumptions, supposing only that we have a balance between longwave and solar radiation and the effects of convection. Let us assume that upward and downward shortwave fluxes, U_S and D_S , obey the two-stream radiative equations with no thermal emission and no scattering, known as the Schwarzschild equations, namely

$$\frac{dD_S}{d\tau_S} = -D_S, \quad \frac{dU_S}{d\tau_S} = U_S. \quad (18.80\text{a,b})$$

Here, τ_S is the shortwave optical depth, which is related to the physical height by $d\tau_S = -e_S dz$, where e_S is the shortwave emissivity, and we can then write the radiative equations as

$$\frac{dD_S}{d\tau_L} = -\frac{e_S}{e_L} D_S, \quad \frac{dU_S}{d\tau} = \frac{e_S}{e_L} U_S. \quad (18.81\text{a,b})$$

The solar heating is proportional to the convergence of the net solar fluxes, $-\partial N_S / \partial z$ where $N_S = U_S - D_S$. The convection also provides a local heating, and this is proportional to the convergence of the upwards enthalpy flux. Thus, in a steady state,

$$\frac{\partial}{\partial z} (N_S + N_L + \mathcal{H}) = 0 \quad \text{implying} \quad \frac{\partial}{\partial \tau} (N_S + N_L + \mathcal{H}) = 0, \quad (18.82\text{a,b})$$

where \mathcal{H} is the enthalpy flux. Using (18.70a) we write (18.82b) as

$$U_L + D_L - 2B + \frac{1}{e_L} Q_S + \frac{\partial \mathcal{H}}{\partial \tau} = 0, \quad (18.83)$$

where Q_S is the solar heating given by $Q_S = -\partial N_S / \partial z = e_L^{-1} \partial N_S / \partial \tau$. If we had a theory for the enthalpy fluxes, \mathcal{H} , then (18.83), in conjunction with the Schwarzschild equations and knowledge of the solar radiation and atmospheric emissivity, would determine the temperature profile with height. We have no such theory, but we can still make progress.

Effects of solar fluxes

Suppose first (and taking a diversion from convection) that the enthalpy fluxes are zero, much as in the stratosphere. We then have a radiative equilibrium profile that satisfies

$$\frac{\partial}{\partial z} (N_L + N_S) = 0, \quad (18.84)$$

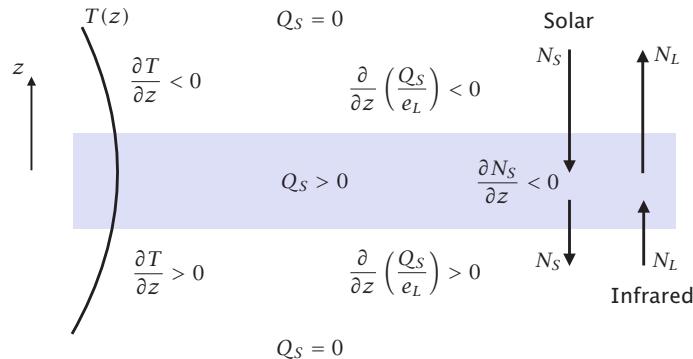


Fig. 18.20 Effects of a band of solar radiation absorption. Shading marks the absorbing region where $\partial N_s / \partial z < 0$, with consequences for the heating, Q_S , and temperature profile, $T(z)$, determined by (18.86b), as indicated.

implying

$$\frac{\partial N_L}{\partial z} = Q_S \quad \text{or} \quad \frac{\partial N_L}{\partial \tau} = -\frac{Q_S}{e_L}. \quad (18.85a,b)$$

If we differentiate (18.85b) with respect to τ and use (18.70a,b) we obtain

$$2 \frac{\partial B}{\partial \tau} = N_L + \frac{\partial}{\partial \tau} \left(\frac{Q_S}{e_L} \right) \quad \text{or} \quad 8\sigma T^3 \frac{\partial T}{\partial z} = -e_L N_L + \frac{\partial}{\partial z} \left(\frac{Q_S}{e_L} \right). \quad (18.86a,b)$$

The middle of the stratosphere in Earth's atmosphere contains a layer of ozone that absorbs solar radiation. Thus, $\partial N_s / \partial z < 0$ and Q_S is positive. The net longwave radiation N_L is positive in this region, as there is much more upwelling radiation than downwelling, and this tends to produce a negative vertical temperature gradient, but the effect is weak because e_L is small. However, the solar radiation can have a strong effect because e_L appears in the *denominator* of the term involving Q_S . Just below the heating region $\partial Q_S / \partial z$ is positive and this produces a positive value of $\partial T / \partial z$, with the converse above the ozone layer. Thus, the region of solar heating corresponds to a maximum of temperature, or at least a region of small vertical temperature gradient, with negative values of $\partial^2 T / \partial z^2$, as schematically illustrated in Fig. 18.20. The result is intuitively reasonable, but (18.86) provides the explicit solution. A non-intuitive result that follows from (18.86b) concerns the effects of an increasing infrared emissivity, such as happens with global warming. Suppose that $\partial T / \partial z > 0$, as in the stratosphere because of the presence of a layer of ozone, and suppose further that e_L then increases, because of an increased concentration of greenhouse gases. Changes in the longwave term are small, because e_L is small, but changes in the solar term are large, and these cause $\partial T / \partial z$ to diminish. That is, an increase in the concentration of greenhouse gases will cause stratospheric temperatures to fall.¹¹ The result arises because of the balance, in this case, between solar heating and infrared cooling, as in (18.85a) and Fig. 18.20. If the emissivity increases there will still be the same infrared loss to space, to achieve radiative balance, and if the emissivity is higher the loss can be achieved at a lower temperature.

Effects of enthalpy fluxes

Now consider the effects of an enthalpy flux, supposing that the atmosphere is transparent to solar radiation. The equilibrium condition, (18.83), becomes

$$U_L + D_L - 2B + \frac{\partial \mathcal{H}}{\partial \tau} = 0. \quad (18.87)$$

Differentiating with respect to τ and using (18.70b) gives

$$2 \frac{\partial B}{\partial \tau} = N_L + \frac{\partial^2 \mathcal{H}}{\partial \tau^2}. \quad (18.88)$$

Now, from (18.82a) we must have $N_S + N_L + \mathcal{H} = 0$ for all z , if the atmosphere is in an overall radiative balance and if $\mathcal{H} = 0$ at the atmosphere's top. If the solar flux is constant with height then, also for all z ,

$$N_L = S_{net} - \mathcal{H}, \quad (18.89)$$

where S_{net} is the net incoming solar radiation, a positive quantity. Using this expression (18.88) becomes

$$2\frac{\partial B}{\partial \tau} = S_{net} - \mathcal{H} + \frac{\partial^2 \mathcal{H}}{\partial \tau^2} \quad \text{or} \quad \frac{8\sigma T^3}{e_L} \frac{\partial T}{\partial z} = -S_{net} + \mathcal{H} - \frac{1}{e_L} \frac{\partial}{\partial z} \left(\frac{1}{e_L} \frac{\partial \mathcal{H}}{\partial z} \right). \quad (18.90a,b)$$

These expressions explicitly tell us how the lapse rate is affected by enthalpy fluxes. If we knew \mathcal{H} we could integrate (18.90) to give us the temperature at every level, and if $\mathcal{H} = 0$ the equations are equivalent to the derivative of (18.71c). Enthalpy fluxes are usually positive (i.e., upwards) in the atmosphere and so they tend to increase $\partial T/\partial z$ and reduce the lapse rate; that is, they tend to warm the upper atmosphere and cool the lower atmosphere and the surface. Since we do not have a quantitative theory for these fluxes let us invoke some ideas of quasi-equilibrium and specify the resulting lapse rate rather than the fluxes themselves.

18.6.2 † Convective Adjustment and the Height of the Tropopause

Let us assume that convection occurs with sufficient efficiency so that it will establish a convectively neutral lapse rate, Γ say, which for simplicity we here assume is constant. We may also assume that the convection equalizes the ground temperature and the surface temperature (the temperature of the layer of air immediately above the ground). Such a model simplifies the problem enormously, for we avoid completely the problem of solving the radiative transfer equations. However, radiation is still present, and we must still satisfy an overall radiative balance, and this determines the height to which the convection extends — which is, effectively, the height of the tropopause. In the first instance we can suppose that the convection occurs quickly and adiabatically, so that the re-arrangement of the temperature profile conserves energy. However, such a profile will not necessarily be a solution of the radiative transfer equations (18.67) and the profile must adjust until a solution is found. If the lapse rate is fixed, the only degree of freedom is the height to which convection reaches — that is, the height of the tropopause — above which radiative equilibrium holds.

The computation of this height follows a straightforward algorithm. If the top of the convecting layer occurs at a height, H_T , at which the optical depth is small, then outgoing long-wave radiation there will be approximately equal to the outgoing radiation at the top of the atmosphere, $U_L(\tau = 0)$, which is equal to the incoming solar radiation and therefore known. This gives the tropopause temperature (from (18.76), namely $T_T^4 = U_L(\tau = 0)/2\sigma$) and therefore, if we know H_T and the lapse rate, the temperature at all heights. We can then calculate the upwards radiative flux using a variant of (18.67b), which we may write as

$$U_L(z = 0) = U_L(z = H_T) e^{\tau_s} + \int_0^{H_T} B(\tau) e^{\tau_s - \tau} \frac{d\tau}{dz} dz, \quad (18.91)$$

where τ_s is the optical depth at the surface. However, for an arbitrary tropopause height, the up-welling radiation at the bottom, $U_L(z = 0)$, as given by (18.91), will *not* equal σT_s^4 , which it must do if radiative balance is to be satisfied. The height of the tropopause must therefore adjust, and an iterative algorithm for finding the equilibrium solution goes as follows:

- (i) Obtain the radiative equilibrium temperature profile.
- (ii) Make a guess for the height of the tropopause, and using the given lapse rate obtain the temperature all the way down to the ground.

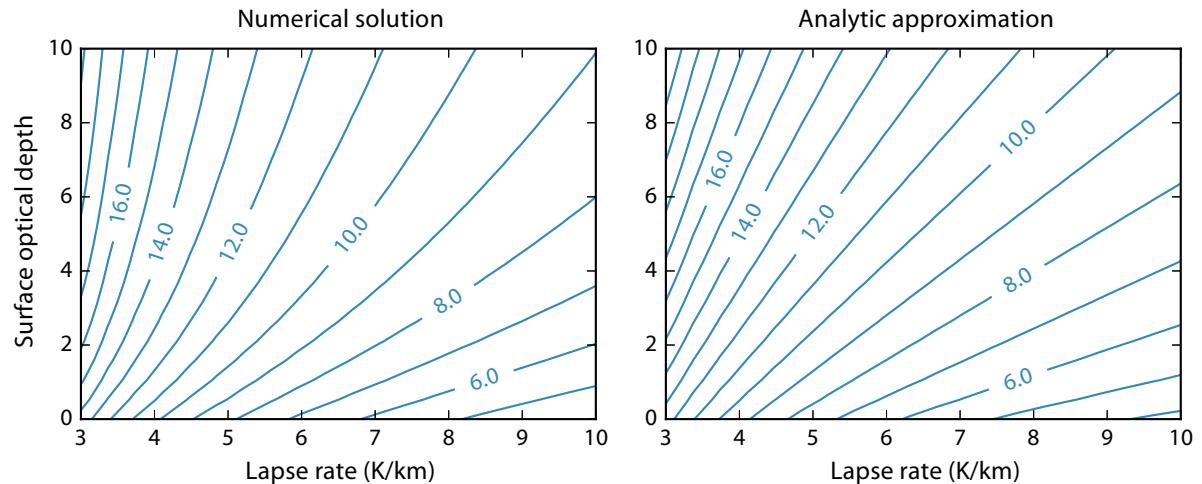


Fig. 18.21 Contours of tropopause height (km) as a function of lapse rate and surface optical depth in a grey atmosphere, calculated numerically by iterating the Schwarzschild equations (left) and using an analytic approximation (right), with $H_a = 2\text{ km}$ and outgoing longwave radiation of 242 W m^{-2} , or tropopause temperature of 215 K .

- (iii) Integrate the radiative transfer equations, (18.91), down from the top. The outgoing radiative balance is achieved this way but there is no balance at the surface if temperature is continuous. That is, $\sigma T_s^4 \neq U_L(z=0)$.
- (iv) Change the height of the tropopause, find another solution, and iterate until the surface radiative balance is properly achieved.

A profile of temperature so calculated, along with the radiative-equilibrium temperature, is shown in Fig. 18.19, and solutions as a function of lapse rate and temperature are given in Fig. 18.21.

18.6.3 Approximate Analytic Solution

In Appendix C (page 725) we show that an approximate analytic solution of the above algorithm for the height of the tropopause, H_T , is

$$8GH_T^2 - CH_TT_T - \tau_s H_a T_T = 0, \quad (18.92)$$

or

$$H_T = \frac{1}{16\Gamma} \left(CT_T + \sqrt{C^2 T_T^2 + 32\Gamma \tau_s H_a T_T} \right), \quad (18.93)$$

where $C = \log 4 \approx 1.4$, Γ is the lapse rate, T_T is the temperature at the tropopause, τ_s is the surface optical depth and H_a is the scale height of the main infrared absorber. For Earth's atmosphere, $H_a \approx 2\text{ km}$, $\tau_s \approx 5$ and $\Gamma \approx 6.5 \text{ K km}^{-1}$. All three terms in (18.92) are then approximately the same size and (18.93) gives $H_T \approx 11\text{ km}$. We can now be a little more precise about what it means for an atmosphere to be optically thin or thick. Using (18.93) and approximating $C^2 = 2$ we see that the optically thick limit arises when

$$\tau_s H_a \gg \frac{T_T}{16\Gamma} \quad \text{whence} \quad H_T \approx \sqrt{\frac{T_T \tau_s H_a}{8\Gamma}}. \quad (18.94)$$

The optically thin case has

$$\tau_s H_a \ll \frac{T_T}{16\Gamma} \quad \text{whence} \quad H_T \approx \frac{CT_T}{8\Gamma}. \quad (18.95)$$

With parameters appropriate for Earth's atmosphere both of the above limits give estimates in the range 6–15 km, and they are additive effects. Plots of the tropopause height as a function of lapse rate and optical depth, calculated using (18.93), as well as numerical solutions using the algorithm of the previous subsection, are given in Fig. 18.21, and the agreement is fairly good in Earth's parameter regime. A robust result is that as the lapse rate diminishes ($-\partial T/\partial z$ becomes smaller) the tropopause height increases, essentially to maintain the same tropopause temperature.

The main quantitative deficiency of this argument is that the atmosphere is not grey — the absorption of radiation is a function of wavelength. A less severe approximation is to suppose that the infrared radiation occurs in two bands — a 'window' region and the remainder. In the window region the atmosphere is fairly transparent, meaning that a fraction ($\sim 1/4$ or $1/3$) of the radiation emitted by the surface goes straight to space. This has the effect of slightly decoupling the tropopause temperature from the effective emitting temperature. Nonetheless, the general arguments leading to (18.93) remain valid and the way in which the tropopause height varies with optical depth and stratification will carry through in the more realistic case.¹²

Effects of lateral energy transport and application to mid-latitudes.

Although we have set the above calculation in the context of convection, it applies in any situation where the lapse rate can be specified; the argument would equally well apply if the lapse rate were set by baroclinic instability. We can also apply a modified argument even when there is a horizontal transport of energy, for such a transport effectively just acts to change the radiative emitting temperature. If the lateral convergence of energy and the incoming solar radiation are known then the amount of radiation that the column must emit to space in order to maintain energy balance is easily calculated. For example, if there is a convergence of energy at high latitudes because of the transport of energy by baroclinic eddies, then the column needs to emit less infrared radiation to space than it would otherwise in order to maintain radiative equilibrium. In actuality, in mid-latitudes the lapse rate cannot be regarded as being specified independently of the lateral heat transport, and the determination of the lapse rate and the tropopause height become intertwined. The reader is referred back to Section 15.5 to follow this argument.

18.7 VERTICALLY-CONSTRAINED EQUATIONS OF MOTION FOR LARGE SCALES

Let us now segue toward the large scale dynamics of the tropics. We first show how having a nearly constant lapse rate constrains the vertical degrees of freedom, so reducing the equations of motion to something akin to the shallow water equations, and readers may wish to skim Section 3.4 before continuing.

18.7.1 Reduction of Vertical Degrees of Freedom

The assumption that convection maintains a moist-adiabatic profile everywhere constrains the vertical structure of the horizontal temperature gradient and, at least in so far as the momentum dynamics are linear, reduces the equations of motion governing the large-scale to a set similar to the shallow-water equations.¹³ To see this we begin with hydrostasy in pressure coordinates applied to the fluctuating fields,

$$\frac{\partial \phi'}{\partial p} = -\alpha', \quad (18.96)$$

where ϕ' and α' are variations in the geopotential and the specific volume. The latter can be taken to be a function of the state variables pressure and saturation entropy, η (dropping the subscript s). Variations in α at constant pressure thus obey

$$\alpha' = \left(\frac{\partial \alpha}{\partial \eta} \right)_p \eta' = \left(\frac{\partial T}{\partial p} \right)_\eta \eta', \quad (18.97)$$

using one of Maxwell's equations (page 19). If we integrate (18.96) in the vertical and use (18.97) we obtain

$$\phi'(x, y, p, t) = - \int \alpha dp = \int \left(\frac{\partial T}{\partial p} \right)_\eta \eta' dp = \eta' \int dT, \quad (18.98)$$

where we can take η' out of the integral because it is constant if the atmosphere has a saturated adiabatic profile. Integrating gives

$$\phi(x, y, p, t) - \bar{\phi}(x, y, t) = (\bar{T}(x, y, t) - T(x, y, p, t)) \eta(x, y, t), \quad (18.99)$$

where we have written the constant of integration such that $\bar{\phi}$ and \bar{T} are the vertical means, in pressure coordinates, of ϕ and T . Following conventional usage, we call the vertically integrated components 'barotropic' and the deviations 'baroclinic'.

The temperature difference $(\bar{T}(x, y, t) - T(x, y, p, t))$ in (18.99) is determined by the saturated adiabatic lapse rate and in the horizontal this varies quite weakly with the temperature variations found in the tropics. Thus, the horizontal variations of the terms on the right-hand side are dominated by variations in entropy, and to a good approximation we can write

$$\nabla \phi(x, y, p, t) = \nabla \bar{\phi}(x, y, t) + M(p) \nabla \eta(x, y, t), \quad (18.100)$$

where $M(p) = \bar{T}(x, y, t) - T(x, y, p, t)$ and which we take to be a function of pressure alone. The profile of $M(p)$ has a single node in the vertical, and this constrains the horizontal velocity to have a similar variation in the vertical, and the vertical integral of $M(p)$ vanishes.

Velocity decomposition

First consider the vertical velocity. In pressure coordinates the mass continuity equation is

$$\frac{\partial \omega}{\partial p} = -\nabla_p \cdot \mathbf{u}, \quad (18.101)$$

where $\bar{\mathbf{u}} = (u, v)$. Let $\mathbf{u} = \bar{\mathbf{u}} + \mathbf{u}^*$, where $\bar{\mathbf{u}}$ and \mathbf{u}^* are the barotropic and baroclinic components of \mathbf{u} , and ω is the vertical (pressure) velocity. We suppose that the fluid is confined between the ground at $p = p_g$ and a rigid tropopause at p_t , with $\omega = 0$ at both. The baroclinic components of the velocity will, as a consequence of their definition, vanish when integrated over the troposphere. An integration of (18.101) over the troposphere implies the barotropic flow is divergence-free:

$$\nabla_p \cdot \bar{\mathbf{u}} = 0. \quad (18.102)$$

The vertical velocity is thus related to the baroclinic flow by,

$$\frac{\partial \omega}{\partial p} = -\nabla_p \cdot \mathbf{u}^*. \quad (18.103)$$

The horizontal flow itself is given using the momentum equation which, using (18.100), we write as

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{v} \cdot \nabla \mathbf{u} + \mathbf{f} \times \mathbf{u} = -\nabla \bar{\phi} + M(p) \nabla_p \eta, \quad (18.104)$$

where \mathbf{v} is the three-dimensional velocity and we omit forcing and viscous terms, and ∇ without a subscript denotes horizontal derivative. The above equation suggests that it will be a good approximation to suppose that the horizontal velocity has the same vertical structure as $M(p)$ and we let

$$\mathbf{u}(x, y, p, t) = \mathbf{u}_0(x, y, t) + m(p) \mathbf{u}_1(x, y, t), \quad (18.105)$$

where \mathbf{u}_0 ($= \bar{\mathbf{u}}$) and \mathbf{u}_1 are the ‘barotropic’ and ‘first baroclinic’ components of the flow, and $m(p) = M(p)/N$ where N is a normalizing coefficient, say $[\int M^2 dp/(p_t - p_g)]^{1/2}$, so that $m(p)$ is a dimensionless basis function. The reader will appreciate the similarity of this approach to that of Section 3.4, now with the quasi-equilibrium constraints determining the form of the vertical eigenfunctions for us.

Momentum equations

Neglecting the advective term the horizontal momentum equation decomposes exactly into separate barotropic and baroclinic equations which, omitting frictional terms, are

$$\frac{\partial \mathbf{u}_0}{\partial t} + \mathbf{f} \times \mathbf{u}_0 = -\nabla \phi_0, \quad \frac{\partial \mathbf{u}_1}{\partial t} + \mathbf{f} \times \mathbf{u}_1 = -\nabla \phi_1, \quad (18.106a,b)$$

where $\phi_0 = \bar{\phi}$ and $\phi_1 = N\eta$. These have the same form as the linearized shallow water equations. We may obtain a diagnostic equation for the barotropic pressure by taking the divergence of (18.106a), whence the time derivatives disappear because $\nabla \cdot \mathbf{u}_0 = 0$. Alternatively, take the curl of (18.106a) to give the linear barotropic vorticity equation,

$$\frac{\partial \zeta_0}{\partial t} + \beta v_0 = 0. \quad (18.107)$$

This equation is closed because, since the flow is divergence-free, there exists a streamfunction ψ such that $(u_0, v_0) = (-\partial\psi/\partial y, \partial\psi/\partial x)$ and $\zeta_0 = \nabla^2\psi$.

If we add nonlinearity back in then the baroclinic and barotropic modes interact. Thus, the terms $\mathbf{u}_1 \cdot \nabla \mathbf{u}_1$ and $\mathbf{u}_0 \cdot \nabla \mathbf{u}_0$ both project onto the barotropic flow, and the terms $\mathbf{u}_1 \cdot \nabla \mathbf{u}_0$ and $\mathbf{u}_0 \cdot \nabla \mathbf{u}_1$ project onto the baroclinic flow, in a way that is analogous to that occurring in two-layer quasi-geostrophic flow (Section 12.2.2). A baroclinic–barotropic interaction also arises because of vertical advection and we obtain

$$\frac{\partial \mathbf{u}_0}{\partial t} + \mathbf{u}_0 \cdot \nabla \mathbf{u}_0 + D_0(\mathbf{u}_1, \mathbf{u}_1) + \mathbf{f} \times \mathbf{u}_0 = -\nabla \phi_0, \quad (18.108a)$$

$$\frac{\partial \mathbf{u}_1}{\partial t} + \mathbf{u}_1 \cdot \nabla \mathbf{u}_0 + \mathbf{u}_0 \cdot \nabla \mathbf{u}_1 + D_1(\mathbf{u}_1, \mathbf{u}_1) + \mathbf{f} \times \mathbf{u}_1 = -\nabla \phi_1, \quad (18.108b)$$

where D_0 and D_1 are nonlinear (largely advective) operators whose exact form is not of concern here. The barotropic flow is divergence free and this determines the pressure ϕ_0 , but ϕ_1 is as yet undetermined.

Thermodynamics

To close the equation for baroclinic flow we use an equation for the entropy or for the equivalent potential temperature in conjunction with (18.100). Since entropy is assumed invariant with height there is no vertical dependence and the equation might be written in the general form,

$$\frac{\partial \eta}{\partial t} + \mathbf{u} \cdot \nabla \eta = S, \quad (18.109)$$

where S are various source and sink terms of both heat and moisture. This term hides a multitude of sins, for it potentially includes radiative, condensational, evaporative and turbulent flux terms with complex flow-dependent specifications, possibly needing separate equations for dry air and moisture. However, since entropy is (it is assumed) constant with height, we need a thermodynamic equation only at one level, which we might take to be near the surface, or we may use a vertical average.

In summary, the constraints on the vertical temperature structure imposed by convection lead to a relatively simple vertical structure of all the dynamical fields, and as a consequence a heat

source will primarily force the first baroclinic mode. In the linear approximation the baroclinic flow is uncoupled from the barotropic flow. The procedure provides some justification for using shallow water-like equations in tropical dynamics, as in Chapter 8, with the velocity fields to be interpreted as being those of the first baroclinic mode, an approach that is especially useful if the dynamics are predominantly linear. More generally, it is perhaps useful to think of this type of quasi-equilibrium model in a similar light as quasi-geostrophy: neither is quantitatively accurate for flow evolution (we would use neither model for an accurate weather forecast) but they may be able to provide insight where the full equations are too complex.

The equations derived above are still unbalanced, and are simpler than the primitive equations only in their constrained vertical structure. Let us try to go a little further and see if we can incorporate any balances in the horizontal that might simplify matters.

18.8 SCALING AND BALANCED DYNAMICS FOR LARGE-SCALE FLOW IN THE TROPICS

In earlier chapters we looked at flow with small Rossby number and, by performing a scale analysis, determined what the dominant balance of terms was in the equations of motion. This allowed us to derive the quasi-geostrophic equations, which are the basis for much of the theory of mid-latitude motion. Is such a program possible for the tropics? The answer is, ‘well, in part’. It *is* possible to derive some reduced sets of equations for the tropics, and that will be the main topic of this section. However, these reduced sets have not proven nearly as useful for the tropics as quasi-geostrophy has been for the mid-latitudes, in part because it is harder to make relevant equations that are simple, or simple equations that are relevant. In any case, let us begin with the stratified primitive equations, without explicitly invoking a constrained vertical structure.

18.8.1 Balanced, Adiabatic Flow

We will present a scaling for tropical flow side-by-side with the corresponding scaling for mid-latitude flow.¹⁴ We begin with the hydrostatic primitive equations for adiabatic, frictionless flow, which, reprising (5.15b), may be written as,

$$\frac{D\mathbf{u}}{Dt} + \mathbf{f} \times \mathbf{u} = -\nabla_z \phi, \quad \frac{\partial \phi}{\partial z} = b, \quad (18.110a,b)$$

$$\frac{Db}{Dt} + N^2 w = 0, \quad \nabla \cdot (\tilde{\rho}\mathbf{v}) = 0. \quad (18.110c,d)$$

These are nominally the anelastic equations in height coordinates in our standard notation, but an entirely equivalent derivation could use pressure coordinates as these have similar form. The reader will recall that $b = g\delta\theta/\theta_0$ is the buoyancy and $N^2 = d\tilde{b}/dz$ where $\tilde{b}(z)$ is a reference stratification. We will suppose that the basic variables scale according to

$$(x, y) \sim L, \quad z \sim H, \quad (u, v) \sim U, \quad w \sim W, \quad t \sim \frac{L}{U}, \quad \phi \sim \Phi, \quad b \sim B, \quad f \sim f_0. \quad (18.111)$$

The quantity B is representative of horizontal variations in buoyancy. Vertical variations scale differently, hence their separate representation in (18.110c). By choosing the time t to scale advectively we are implicitly eliminating gravity waves. The nondimensional numbers that will arise are the Rossby, Burger and Richardson numbers,

$$Ro = \frac{U}{f_0 L}, \quad Bu = \left(\frac{L_d}{L} \right)^2 = \left(\frac{NH}{f_0 L} \right)^2, \quad Ri = \left(\frac{NH}{U} \right)^2, \quad (18.112)$$

and evidently

$$Bu = Ri \times Ro^2. \quad (18.113)$$

The Rossby number Ro is generally small in mid-latitudes for large-scale flow, but in the tropics it is $\mathcal{O}(1)$ or larger. The Richardson number (which is the inverse square of a Froude number) is usually large in both mid-latitudes and tropics, *except* in regions of active convection where N is very small. If we take $N = 10^{-2} \text{ s}^{-1}$, $H = 10^4 \text{ m}$ and $U = 10 \text{ m s}^{-1}$ then $Ri = 100$. In fact, for large-scale flow the Richardson number is usually sufficiently large that $1/(Ro Ri)$ is small in both mid-latitudes and tropics.

The difference between the tropics and mid-latitudes is apparent from the dominant balance in the momentum equation. At small Rossby number we have the familiar geostrophic balance and with hydrostatic balance we obtain the scaling:

$$\mathbf{f} \times \mathbf{u} \approx -\nabla_z \phi, \quad \frac{\partial \phi}{\partial z} = b, \quad \Rightarrow \quad \Phi = f_0 U L, \quad B = \frac{f_0 U L}{H}. \quad (18.114)$$

In the tropics the advective term, or the advectively-scaled time derivative, balances the pressure gradient meaning that $D\mathbf{u}/Dt \sim \nabla_z \phi$ and, since we still have $\partial \phi / \partial z = b$ we find

$$\Phi = U^2, \quad B = \frac{U^2}{H}. \quad (18.115)$$

If U is of similar magnitude in the tropics and mid-latitudes (and in the absence of a dynamical analysis this is an assumption), then *variations of pressure and temperature are smaller in the tropics than in mid-latitudes*. This is an important and not-quite obvious result and it is the essence of the *weak temperature gradient approximation*, discussed further in the next section. If we were to carry through the derivation in pressure co-ordinates (see the shaded box on page 81) with geopotential and temperature as the variables we would find

$$\text{Mid-latitudes: } \Phi = f_0 U L, \quad T_s = \frac{f_0 U L}{R}, \quad (18.116a)$$

$$\text{Tropics: } \Phi = U^2, \quad T_s = \frac{U^2}{R}, \quad (18.116b)$$

where Φ is now the scaling for geopotential, T_s is the scaling for temperature and R is the ideal gas constant. For $U = 10 \text{ m s}^{-1}$, $L = 10^6 \text{ m}$ and, for mid-latitudes only, $f_0 = 10^{-4} \text{ s}^{-1}$, we obtain

$$\text{Mid-latitudes: } \Phi \sim 1000 \text{ m}^2 \text{ s}^{-2}, \quad T_s \sim 3 \text{ K}, \quad (18.117a)$$

$$\text{Tropics: } \Phi \sim 100 \text{ m}^2 \text{ s}^{-2}, \quad T_s \sim 0.3 \text{ K}. \quad (18.117b)$$

The implications of these results are illustrated in Fig. 18.22, which shows a snapshot of observed contours of geopotential height, temperature and zonal wind; the large-scale variability of geopotential and temperature is evidently much smaller in the tropics than mid-latitudes.

Vertical velocity

The obvious scaling for the vertical velocity is that suggested by the mass continuity equation, namely $W = UH/L$. However, as we know from Chapter 5, the vertical velocity may be much less than this estimate in a stratified, rotating fluid, and here we start with the thermodynamic equation for adiabatic flow,

$$\mathbf{u} \cdot \nabla b + w N^2 = 0, \quad \Rightarrow \quad W = \frac{UB}{LN^2}. \quad (18.118)$$

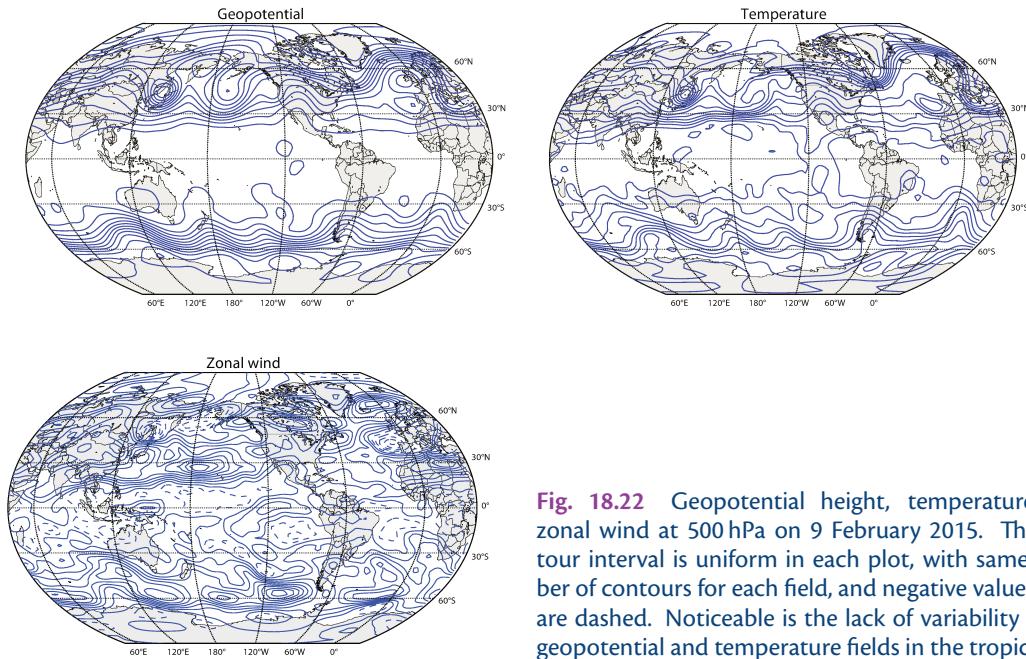


Fig. 18.22 Geopotential height, temperature and zonal wind at 500 hPa on 9 February 2015. The contour interval is uniform in each plot, with same number of contours for each field, and negative values for u are dashed. Noticeable is the lack of variability of the geopotential and temperature fields in the tropics.

Using (18.114) and (18.115) gives

$$\text{Mid-latitudes: } W = \frac{f_0 U^2}{HN^2} = \frac{f_0 L}{U} \frac{U^2}{N^2 H^2} \frac{UH}{L} = (Ro \, Ri)^{-1} \left(\frac{UH}{L} \right), \quad (18.119a)$$

$$\text{Tropics: } W = \frac{U^3}{LHN^2} = \frac{U^2}{N^2 H^2} \frac{UH}{L} = Ri^{-1} \left(\frac{UH}{L} \right), \quad (18.119b)$$

where $Ri \equiv N^2 H^2 / U^2$ is the Richardson number, with typical values of $\mathcal{O}(10\text{--}100)$ for large-scale flow. Thus, again perhaps non-intuitively, the vertical velocity is, for adiabatic flow, *smaller* in the tropics than in mid-latitudes, by order of a mid-latitude Rossby number. In mid-latitudes the scaling for W is more commonly written as

$$W = \frac{f_0 U^2}{HN^2} = Ro \frac{L^2}{L_d^2} \frac{UH}{L} = \frac{Ro}{Bu} \frac{UH}{L}, \quad (18.120)$$

so that for scales comparable to the mid-latitude deformation radius (i.e., with $Bu \sim 1$) the vertical velocity is order Rossby-number smaller than mass-continuity scaling suggests.

Vorticity

Cross-differentiating the horizontal momentum equation gives, as in (4.66) but without a baroclinic term, the vertical component of the vorticity equation with associated scalings,

$$\frac{D}{Dt}(\zeta + f) = -(\zeta + f) \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) + \left(\frac{\partial u}{\partial z} \frac{\partial w}{\partial y} - \frac{\partial v}{\partial z} \frac{\partial w}{\partial x} \right), \quad (18.121a)$$

$$\text{Tropics: } \left(\frac{U}{L} \right)^2 \sim \left(\frac{U}{L} + \frac{1}{Ro} \frac{U}{L} \right) \left(\frac{1}{Ri} \frac{U}{L} \right) \left(\frac{U}{H} \right) \left(\frac{1}{Ri} \frac{UH}{L^2} \right), \quad (18.121b)$$

$$\text{Mid-latitudes: } \left(\frac{U}{L} \right)^2 \sim \left(\frac{U}{L} + \frac{1}{Ro} \frac{U}{L} \right) \left(\frac{Ro}{Bu} \frac{U}{L} \right) \left(\frac{U}{H} \right) \left(\frac{Ro}{Bu} \frac{UH}{L^2} \right). \quad (18.121c)$$

In the tropical case with $Ro = \mathcal{O}(1)$ or larger *all* the terms on the right-hand side are much smaller than the terms on the left-hand side, whereas in the mid-latitude case vortex stretching by the Coriolis term, $f(\partial u/\partial x + \partial v/\partial y)$, is the same order (because f is big and the divergence is small). Thus, in the tropical case the vorticity equation simplifies severely and at lowest order becomes the two-dimensional vorticity equation,

$$\frac{D}{Dt}(\zeta + f) = 0. \quad (18.122)$$

The large-scale velocity is, at this order, purely rotational and is given by a streamfunction ψ such that

$$\nabla^2\psi = \zeta, \quad (u, v) = \left(-\frac{\partial\psi}{\partial y}, \frac{\partial\psi}{\partial x} \right). \quad (18.123)$$

This equation nominally holds independently at each vertical level, although the assumptions that give rise to it assume that the depth scale is large. The pressure is not needed to step forward (18.123) but it may be obtained diagnostically. To do so we go back to the horizontal velocity equation written in the vector-invariant form,

$$\frac{\partial \mathbf{u}}{\partial t} - \mathbf{u} \times (\zeta + f) = -\nabla \left(\phi + \frac{1}{2} \mathbf{u}^2 \right), \quad (18.124)$$

neglecting the vertical and divergent velocities. Taking the divergence gives the so-called nonlinear balance or gradient-wind equation,

$$\nabla^2\phi = \nabla \cdot \left[(f + \zeta) \nabla\psi - \frac{1}{2} \nabla(\nabla\psi)^2 \right], \quad (18.125)$$

which is similar to (2.238) and expresses a balance between the pressure gradient, Coriolis and centrifugal forces.

18.8.2 A few Remarks

The above derivations and results require some comment:

- The relative weakness of large-scale horizontal gradients of pressure (or geopotential) and temperature in the tropics, for a given velocity field, is a robust result of the scaling analysis and borne out in observations.
- The smallness of the vertical velocity requires that the Richardson number, $N^2 H^2 / U^2$ be large; that is, stratification is strong. This is only true in regions that are not actively convecting; in convective regions N may be small.
- Relatedly, the scaling does not take into account diabatic sources, which may be expected to be particularly important in tropical regions.
- Equally tellingly, (18.123) tells us nothing about the vertical structure and so, unlike quasi-geostrophy in mid-latitudes, is not sufficiently complete to be a useful prognostic, or even diagnostic, equation for tropical motion.

Let us now look at how diabatic effects might affect large-scale motion.

18.9† SCALING AND BALANCE FOR LARGE-SCALE FLOW WITH DIABATIC SOURCES

In the tropics we might expect that heat sources, for example condensational heating, would be important to the extent that they should be explicitly included in any development of reduced equations. Let us see if and how this is possible, using the shallow water equations for illustration.

Using the standard nonlinear shallow water equations for the tropical atmosphere is rather ad hoc (and not justified by quasi-equilibrium except in the linear case) but our intention here is primarily illustrative. The underlying physical assumption is that in an air column adiabatic cooling associated with vertical motion balances diabatic heating, which in the shallow water equations becomes a balance between the heating and the divergence.

18.9.1 Diabatic Balanced Shallow Water Equations

On an f -plane and in conventional notation the equations may be written in vorticity-divergence form as

$$\frac{\partial h}{\partial t} + \nabla \cdot (\mathbf{u}h) = Q, \quad (18.126a)$$

$$\frac{\partial \zeta}{\partial t} + \nabla \cdot [\mathbf{u}(\zeta + f_0)] = -r\zeta, \quad (18.126b)$$

$$\frac{\partial \delta}{\partial t} + \nabla^2 \left(\frac{1}{2} \mathbf{u}^2 + gh \right) - \mathbf{k} \cdot \nabla \times [\mathbf{u}(\zeta + f_0)] = -r\delta, \quad (18.126c)$$

where $\delta = \partial u / \partial x + \partial v / \partial y$ is the divergence, Q is the mass or heating source, r is a frictional coefficient and other notation is standard. The height field is a proxy for both pressure and temperature and we will assume that horizontal gradients are weak, by which we mean that the dominant balance in (18.126a) is characterized by the scaling

$$H\Delta = Q_0, \quad (18.127)$$

where H is the mean thickness of the layer, Δ is a scaling for the divergence and Q_0 is the magnitude of the heating. We then choose the velocity scale, U , and the vorticity scale, Z , to be

$$U = \frac{Q_0 L}{H}, \quad Z = \Delta = \frac{Q_0}{H}. \quad (18.128)$$

Finally, the magnitude of horizontal deviations in the height field, \mathcal{H} , are determined from the divergence equation. Depending on whether rotation is or is not important we deduce

$$\mathcal{H} = \frac{f_0 U L}{g} = \frac{Q_0 f_0 L^2}{g H}, \quad \text{or} \quad \mathcal{H} = \frac{U^2}{g} = \frac{Q_0^2 L^2}{g H^2}. \quad (18.129a,b)$$

The height field may then be separated into a mean and deviation, $h = H + \eta$, where $\eta = \mathcal{H}\hat{\eta}$. These scalings involve the heating in an essential way and so are fundamentally different from the adiabatic scaling of the previous section.

Using the above scalings, with rotation, (18.126) may be written in nondimensional form as,

$$Bu^{-1} \left[\frac{1}{f_0 T} \frac{\partial \hat{\eta}}{\partial \hat{t}} + Ro \nabla \cdot (\hat{\mathbf{u}}\hat{\eta}) \right] + \hat{\delta} = \hat{Q}, \quad (18.130a)$$

$$\frac{1}{f_0 T} \frac{\partial \hat{\zeta}}{\partial \hat{t}} + \nabla \cdot [\hat{\mathbf{u}}(\hat{\zeta} + \hat{f}_0)] = -\frac{r}{\hat{f}_0} \hat{\zeta}, \quad (18.130b)$$

$$\frac{1}{f_0 T} \frac{\partial \hat{\delta}}{\partial \hat{t}} + \nabla^2 \left(\frac{1}{2} \hat{\mathbf{u}}^2 + \hat{\eta} \right) - \mathbf{k} \cdot \nabla \times [\hat{\mathbf{u}}(\hat{\zeta} + \hat{f}_0)] = -\frac{r}{\hat{f}_0} \hat{\delta}, \quad (18.130c)$$

where T is the scaling for time, $Bu = (L_d/L)^2$ with $L_d = \sqrt{gH}/f_0$, and $\hat{f}_0 = 1$. The Rossby number is given by $Ro = Q_0/f_0 H$ and is not necessarily small. In the non-rotating case a similar set of equations can be derived but with different coefficients.

Reduced equations

Let us suppose that the mass source determines the divergence in (18.130a). The condition for this is that

$$\max\left(\frac{1}{f_0 T}, Ro\right) Bu^{-1} \ll 1, \quad (18.131)$$

which means that the scale of motion cannot be too large and the time scale cannot be too short. If (18.131) is satisfied then the dimensional height equation, (18.126a), becomes

$$\nabla \cdot \mathbf{u} = \frac{Q}{H}. \quad (18.132a)$$

This value of the divergence is used in (18.126b) and (18.126c) which, retaining all terms since none are obviously small, become

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla(\zeta + f_0) + (\zeta + f) \frac{Q}{H} = -r\zeta, \quad (18.132b)$$

$$g\nabla^2 h = \mathbf{k} \cdot \nabla \times [\mathbf{u}(\zeta + f_0)] - \frac{1}{H} \frac{\partial Q}{\partial t} - r\delta - \nabla^2 \frac{\mathbf{u}^2}{2}. \quad (18.132c)$$

The equation set (18.132) has but one prognostic equation, namely (18.132b), and so is truly balanced and may be thought of as a generalization of (18.122) and (18.125) to the case with non-zero heating. The divergence equation is a nonlinear balance equation, similar to (18.125), except now with a diabatic term on the right-hand side. The divergent flow itself is computed using the height equation, by an assumed balance between adiabatic cooling and diabatic heating. The relationship between velocity and geopotential (or pressure) is the same as in the adiabatic case, because this arises through the momentum equation. Thus, even in the presence of a heating, gradients of geopotential and temperature remain relatively weak, a result that ultimately arises from the smallness of the Coriolis parameter. The arguments that lead to (18.132), along with the scaling of Section 18.8, provide us with the *weak temperature-gradient approximation*.¹⁵ The importance of the result lies in what it implies about the response of the atmosphere to a localized heating: without making any linear approximation, the equations provide a scaling for the response of the velocity, and suggest that the response may become spread out over a sufficient area to keep the temperature gradients small.

Gravity wave adjustment and weak temperature gradients

Although convection certainly constrains the lapse rate where it is occurring, it does not directly do so elsewhere. Nevertheless, the observed tropical atmosphere has an *average* lapse rate close to the saturated adiabat. Why should this be so? The reason is that a process akin to geostrophic adjustment (Section 3.9) brings the atmosphere close to a state with weak horizontal temperature gradients. Suppose that a particular column undergoes moist convection and, perhaps in a matter of hours, adjusts to a moist neutral profile. Away from the cloud the buoyancy profile will in general be different from that, and so there will be unbalanced horizontal pressure (and hence temperature) gradients. Gravity waves, initiated by the motion surrounding the convection ('compensating subsidence'), will spread from the cloud and will adjust the environmental buoyancy to the same profile as that of the convecting region. The timescale for the adjustment is determined by the time that gravity waves take to propagate horizontally between convecting regions. Internal gravity wave speeds are typically about 10 m s^{-1} or somewhat faster and so will spread a distance 50 km from a cloud in a couple of hours, and the adjustment time on this scale will be of this order. This buoyancy adjustment time is much less than the time it would take a passive tracer to homogenize between clouds by advective or mixing processes. The actual process of gravity wave initiation and subsequent adjustment is complex and beyond our scope, and the reader is referred to the literature.¹⁶

18.9.2 ♦ Weak Temperature Gradient and Stratified Flow

The weak temperature gradient approximation is formally independent of any quasi-equilibrium argument, so in this section we briefly discuss its application to the stratified, three-dimensional equation of motion. The thermodynamic equation in pressure coordinates (see Equation (P.5) on page 81) may be written,

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T + \omega \frac{\partial s}{\partial p} = Q, \quad (18.133)$$

where $s = T + gz/c_p$ is the dry static energy divided by c_p and Q represents heating terms. In the weak temperature gradient approximation this equation becomes

$$\omega \frac{\partial s}{\partial p} = Q. \quad (18.134)$$

If the stratification s is known (e.g., moist neutral) the above equation becomes a diagnostic for the vertical velocity, namely $\omega(p) = Q/\partial_p s$. Even with this approximation the remaining stratified equations (e.g., momentum equation) are somewhat complex, with both nonlinearity and continuous vertical structure.

If we are willing in addition to make the quasi-equilibrium assumption then further simplifications are possible because all the fields are assumed to have a simple vertical structure and the flow may be described, as before, by a system similar to the shallow water equations. The vertical and horizontal velocities are related by the mass continuity equation, (18.101). The barotropic flow has zero horizontal divergence so that the vertical (pressure) velocity and the horizontal are related by

$$\omega(x, y, p, t) = - \int_{p_s}^p \nabla \cdot \mathbf{u} \, dp = -n(p) \nabla \cdot \mathbf{u}_1, \quad (18.135)$$

where $n(p) = \int m(p) \, dp$ and we take $\omega = 0$ at the bottom boundary. Thus, we may write $\omega = n(p)\omega_1(x, y, t)$ and $\omega_1 = -\nabla \cdot \mathbf{u}_1$. If we divide (18.134) and vertically integrate we obtain

$$\bar{s} \nabla \cdot \mathbf{u}_1 = \bar{Q}, \quad (18.136)$$

where \bar{Q} is a vertically integrated heating term, weighted by $1/n(p)$. This equation is the direct analogue of (18.132a) and it provides a predictive equation for the divergence of the baroclinic flow if the heating and stratification are known. Moisture may be added to the mix, in which case a moist static energy appears, and the effects of evaporation and condensation must be included in Q . The precise form of (18.136) and other particulars of implementation will depend on the basis functions chosen for the vertical structure, but the form is generically $\nabla \cdot \mathbf{u}_1 = [Q]/[s]$, where $[Q]$ is a measure of heating and $[s]$ a measure of stratification.¹⁷ Equation (18.136) may be used to close the momentum equation (18.108b). If we take the curl of that equation we obtain a vorticity equation analogous to (18.132b), and if we take its divergence we obtain a diagnostic equation for $\nabla^2 \phi_1$ analogous to (18.132c). The combination of a weak temperature gradient approximation and shallow-water-like equations arising from a constrained vertical stratification may be as close as the tropical atmosphere allows us to get to tractable and understandable equations of motion.

18.10 † CONVECTIVELY COUPLED GRAVITY WAVES AND THE MJO

We now look at some of the manifestations of the theoretical development of this chapter and Chapter 8, and three phenomena suggest themselves, namely the Walker circulation, monsoons and the Madden–Julian Oscillations (MJO). We leave monsoons for another day, in part because the subject is large and its theoretical development is a moving target. And we defer discussion of the Walker circulation — an atmospheric overturning circulation largely in the zonal (i.e., $y-z$)

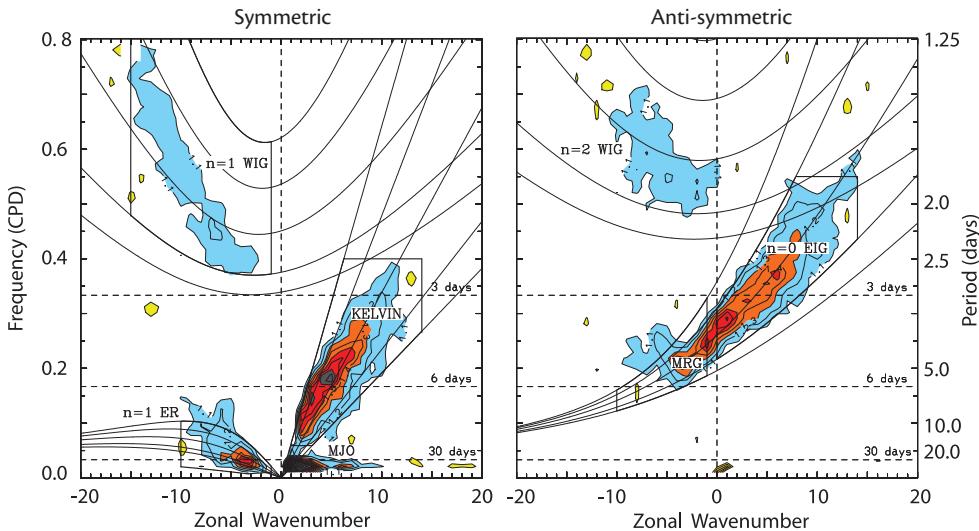


Fig. 18.23 A power spectrum of cloud brightness (colouring) from 15° S to 15° N, measured from satellite and Fourier-transformed into frequency-wavenumber space. Left panel is the symmetric spectrum (Northern plus Southern Hemispheres) and right panel the anti-symmetric spectrum. The solid lines are the corresponding dispersion relations with multiple equivalent depths ranging from 8 m to 90 m, with the best fit at about 25 m. Compare with Fig. 8.6 and Fig. 8.7.²⁰

plane — to Chapter 22 because the phenomenon is closely tied to the equatorial ocean. So let us briefly discuss the MJO, although we cannot give it a wholly crisp explanation because at the time of writing none exists.¹⁸

What is the MJO? It is a pattern of precipitation and winds that takes shape across the western tropical Indian Ocean and drifts eastward at about $4\text{--}8 \text{ m s}^{-1}$ into the western Pacific before dying out over the cooler waters of the eastern tropical Pacific. It often recurs roughly every 30–60 days or so, although it does not oscillate like a conventional wave. Rather, it is more like a somewhat coherent, drifting pattern several thousand kilometres across, consisting of a wet, rainy region of ascending air flanked by dryer regions on either side.

18.10.1 The Observations

The MJO has a characteristic spectral signature that can be obtained from satellite measurements of cloud brightness. Figure 18.23 shows the symmetric (Northern Hemisphere plus Southern Hemisphere) and antisymmetric power spectrum (that is, the intensity of the field, Fourier analysed in zonal wavenumber and frequency) of tropical satellite brightness temperature, after filtering away some background noisiness.¹⁹ Power spectra of other fields, including velocity fields from re-analysis, show similar features. The main shaded regions in the figure fall nicely on the theoretical dispersion relations for Rossby, Kelvin and mixed Rossby-gravity waves as derived in Section 8.2. This is perhaps somewhat startling to see, but evidently equatorial waves do exist!

The theoretical curves match the observations best if the unadorned gravity wave speed, c , is in the region of 10 to 20 m s^{-1} . This speed is given by $c = \sqrt{gH_e}$ where H_e is the equivalent depth (Section 3.4.2), so theoretical values match the observations with $H_e \approx 10\text{--}50 \text{ m}$, perhaps with the best match around 25 m, somewhat smaller than the first equivalent depth computed for the atmosphere in Section 3.4.²¹ Part of the difference may come from the fact that the tropopause is not a rigid lid, and part from the fact that the presence of moisture may reduce the effective static stability of the atmosphere below that implied by the value of N^2 computed using dry potential temperature ($g/\theta_e \partial \theta_e / \partial z$). In a saturated atmosphere the (smaller) value of $g/\theta_e \partial \theta_e / \partial z$ may be

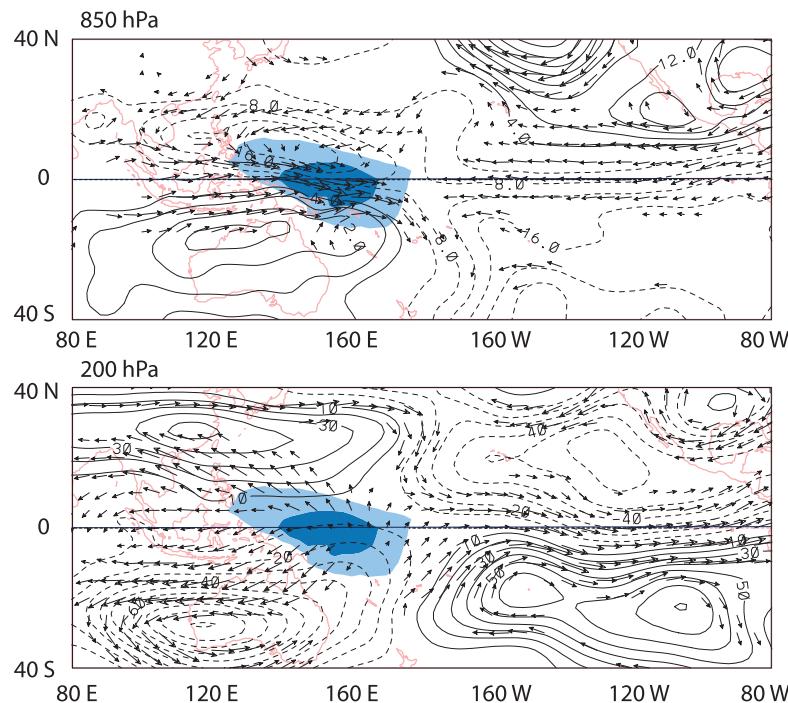


Fig. 18.24 Composite of an observed pattern during an MJO period. Shading indicates anomalously low outgoing IR radiation (less than 16 and 32 W m^{-2} below normal) and indicates the presence of high clouds and precipitation. The solid lines are streamfunctions and the arrows indicate velocities.²²

more relevant for the generation and propagation of gravity waves. The Rossby wave speed is given by $c_R \approx -\beta k / [(2m + 1)\beta/c + k^2]$, so it too depends, albeit more weakly, on c .

In addition to the relatively well understood gravity and Rossby waves, there appears to be a prominent spectral signature in Fig. 18.23 at a small positive zonal wavenumber (between 1 and 5 in the left-hand panel) and a timescale of about 40 days, and this is a signature of the MJO. If we look in physical space, we get the sense that the MJO resembles a drifting Matsuno–Gill pattern more than an oscillating wave — compare Fig. 18.24 with Fig. 8.11 on page 325. Consistently, the observed vertical structure largely has a first baroclinic mode structure, as expected if the lapse rate is constrained to be nearly constant by convection — note the oppositely directed velocity fields in Fig. 18.24, just as in Fig. 8.12. The shaded region corresponds to a region of heating in the Matsuno–Gill model and without too much imagination one may see a Kelvin wave response to its east, along the equator, and off-equatorial Rossby lobes to the west over Australia and South East Asia. A defining characteristic of the MJO is that this pattern drifts eastward at a speed of a few metres per second, much lower than the Kelvin wave speed at that equivalent depth.

18.10.2 † The Mechanism

Precisely why this pattern should move eastward remains unclear, and there is little certainty as to the underlying mechanism except in so far as it likely has to do with the interaction of moist convection with the large-scale circulation — perhaps moist convection *producing* a large-scale circulation that then modulates the convection. Thus, a locally warm region near the equator (initiated by a sea-surface temperature anomaly, for example) will produce a pattern that evolves towards that of the Matsuno–Gill solution (Fig. 8.11), with a low-level convergence of moisture and condensation amplifying the initial pattern. Furthermore, the convergence will produce a dryer region on either side of the heat source (as observed). However, the timescales of condensation are short and the process is unsteady, and it is unlikely that the condensation can produce a self-sustained stationary pattern — in the initial value problem it takes several days for a heating anomaly to settle into a steady Matsuno–Gill pattern in the tropical atmosphere.²³

Thus, instead of a truly steady pattern, the convection will initiate an eastward moving Kelvin wave, along with slower westward moving Rossby waves trapped nearer the source. Now, the initial moisture convergence arises because the Kelvin wave draws air in from the west, and the Rossby wave similarly draws air from the east. If the Kelvin waves moves east the region of low-level moisture convergence, and thus the region of condensational heating, will also move eastward, and the process begins again. However, the convergence region will move eastward much more slowly than the Kelvin wave, since it is the advective convergence that provides the heat source that is the source of both the Kelvin and Rossby waves. A Kelvin wave that breaks free of the convergence will eventually decay since it will not produce the moisture convergence to feed itself. Rather, the MJO is a self-sustained interaction between moisture convergence and the forced-dissipative Kelvin and Rossby waves it produces, with the requirement for moisture convergence greatly slowing the Kelvin wave. Radiative effects tied to the variable cloud field also almost certainly play a role in the way the pattern is damped.

Although wavelike, in this picture the MJO is not a conventional linear wave with a dispersion relation; rather it is related to the translation of a forced quasi-steady pattern with a longer timescale than either of those waves themselves, with the forcing *maintained by the pattern it itself produces*. Evidently, this will be a delicate balance for the interaction of moist convection with Kelvin and Rossby waves is a complex process, so it is not surprising that even the most sophisticated numerical models have trouble reproducing the phenomenon properly, even if the underlying mechanism can seemingly be described in relatively simple terms. Note also that it is not helpful to think of the convection as being in quasi-equilibrium with a more slowly and independently evolving external environment — the convection helps produce its own external environment and the two evolve in synchrony.

APPENDIX A: MOIST THERMODYNAMICS FROM THE GIBBS FUNCTION

Many of the thermodynamic quantities of interest for moist air can be obtained from knowledge of the Gibbs function, in a manner analogous to that followed for dry air in the appendix on page 47. The benefits are that the approach is systematic, explicit formulae for thermodynamic variables can be given, and approximations can be made consistently as needed. Here we outline the methodology and provide some examples.²⁴ We use superscripts d , v and l to denote dry air, water vapour, and liquid water, respectively, and a subscript 0 always denotes a constant, for any quantity. Quantities with multiple superscripts are mixtures of the quantities, but the superscripts are occasionally omitted. A subscript s denotes the saturated value.

Given (1.217), the specific Gibbs functions for dry air and water vapour are, respectively,

$$g^d(p, T) = R^d T \ln(p^d/p_0) + c_p^d T [1 - \ln(T/T_0)], \quad (18.137a)$$

$$g^v(p, T) = R^v T \ln(p^v/p_0) + c_p^v T [1 - \ln(T/T_0)] - L_0^v \frac{T}{T_0} + L_0^v, \quad (18.137b)$$

where p^d and p^v are the partial pressures of dry air and water vapour. A constant, L_0^v , appears in the expression for the Gibbs function of water vapour because we will be concerned with water in its liquid phase. We take the analogous constant to be zero for dry air.

The two fluids are at the same temperature, T , and the total pressure, p , is given by the sum of the partial pressures. Let q be the mass fraction of moist air in the mixture (i.e., the specific humidity), and let ϵ be the ratio of the molecular weights of water vapour and dry air, μ^v/μ^d . The two partial pressures are given by

$$p^v = \frac{qp}{q + \epsilon(1-q)}, \quad p^d = \frac{\epsilon(1-q)p}{q + \epsilon(1-q)}, \quad (18.138)$$

satisfying $p = p^d + p^v$, and p^v is often denoted e in the literature. The Gibbs function for the mixture is

$$g^{dv}(p, T, q) = (1 - q)g^d + qg^v, \quad (18.139)$$

whence

$$\begin{aligned} g^{dv}(p, T, q) &= (qc_p^v + (1 - q)c_p^d)T[1 - \ln(T/T_0)] \\ &\quad + T \left[qR^v \ln \left(\frac{qp/p_0}{q + \epsilon(1 - q)} \right) + (1 - q)R^d \ln \left(\frac{\epsilon(1 - q)p/p_0}{q + \epsilon(1 - q)} \right) \right] \\ &\quad + qL_0^v \left(1 - \frac{T}{T_0} \right). \end{aligned} \quad (18.140)$$

This expression is symmetric in dry air and water vapour, and valid for any mass fraction of moisture in the air (given that $q \ln q \rightarrow 0$ as $q \rightarrow 0$). We will also need the Gibbs function for liquid water, which for our purposes may be written

$$g^l(p, T) = c^l T [1 - \ln(T/T_0)] - \eta_0^l T + \alpha^l p + g_0^l. \quad (18.141)$$

This expression is a simplified version of (1.146) with slightly different notation, with c^l being the heat capacity and α^l the inverse density.

Density and the thermal equation of state

The inverse density of the gas mixture is given by

$$\alpha^{dv} = \left(\frac{\partial g}{\partial p} \right)_{q,T} = \frac{T}{p} [qR^v + (1 - q)R^d], \quad (18.142)$$

or equivalently $p = \rho RT$ where $R = qR^v + (1 - q)R^d$ is the specific gas ‘constant’ for the mixture (which of course varies with q).

Entropy

The entropy is given by $\eta = -(\partial g / \partial T)_{p,q}$ giving

$$\begin{aligned} \eta^{dv} &= (qc_p^v + (1 - q)c_p^d) \ln(T/T_0) \\ &\quad - \left[qR^v \ln \left(\frac{qp/p_0}{q + \epsilon(1 - q)} \right) + (1 - q)R^d \ln \left(\frac{\epsilon(1 - q)p/p_0}{q + \epsilon(1 - q)} \right) \right] + \frac{qL_0^v}{T_0}. \end{aligned} \quad (18.143)$$

The entropy of the mixture may also be written as

$$\eta^{dv} = (1 - q)\eta^d + q\eta^v, \quad (18.144)$$

where the specific entropies of dry air and water vapour are

$$\begin{aligned} \eta^d &= c_p^d \ln(T/T_0) - R^d \ln(p/p_0) - R^d \ln(1 - p^v/p), \\ \eta^v &= c_p^v \ln(T/T_0) - R^v \ln(p/p_0) - R^v \ln(1 - p^d/p) + L_0^v/T_0. \end{aligned} \quad (18.145)$$

The third terms on the right-hand sides are called the entropies of mixing.

The entropy of liquid water is given by

$$\eta^l = \eta_0^l + c^l \ln(T/T_0). \quad (18.146)$$

The entropy of a mixture of dry air, liquid water and water vapour is given by

$$\eta^{dvl} = (1 - q^v)\eta^d + q^v\eta^v + q^l\eta^l, \quad (18.147)$$

where q^v and q^l are the mass concentrations of vapour and liquid in the total mixture and $q^v + q^l = q^v + q^l$.

Heat capacities

The heat capacity at constant pressure of moist air (i.e., dry air and water vapour, with no liquid content) is given by

$$c_p^{dv} \equiv T \left(\frac{\partial \eta}{\partial T} \right)_{p,q} = qc_p^v + (1-q)c_p^d. \quad (18.148)$$

That is, it is the mass weighted heat capacity of water vapour and dry air, as expected. The heat capacity at constant volume is given by

$$c_v^{dv} \equiv T \left(\frac{\partial \eta}{\partial T} \right)_{\alpha,q} = \frac{T(g_{PT}^2 - g_{PP}g_{TT})}{g_{PP}} = qc_v^v + (1-q)c_v^d, \quad (18.149)$$

where $c_v^v = c_p^v - R^v$ and $c_v^d = c_p^d - R^d$, and details are left to the reader. Alternatively, eliminate p in favour of α in (18.143) using (18.142) and then directly evaluate $T(\partial \eta / \partial T)_{\alpha,q}$.

Potential temperature

The potential temperature is the temperature that a parcel has if moved adiabatically and at constant composition to a reference pressure, and so satisfies

$$\eta(p_R, \theta, q) = \eta(p, T, q). \quad (18.150)$$

For moist air we use (18.143) for the entropy and (18.150) becomes

$$\begin{aligned} c_p \ln \theta - & \left[qR^v \ln \left(\frac{q p_R / p_0}{q + \epsilon(1-q)} \right) + (1-q)R^d \ln \left(\frac{\epsilon(1-q) p_R / p_0}{q + \epsilon(1-q)} \right) \right] \\ & = c_p \ln T - \left[qR^v \ln \left(\frac{q p / p_0}{q + \epsilon(1-q)} \right) + (1-q)R^d \ln \left(\frac{\epsilon(1-q) p / p_0}{q + \epsilon(1-q)} \right) \right]. \end{aligned} \quad (18.151)$$

With a couple of lines of algebra this expression simplifies to

$$c_p \ln \theta - [qR^v + (1-q)R^d] \ln(p_R/p) = c_p \ln T \quad \text{or} \quad \theta = T \left(\frac{p_R}{p} \right)^{R/c_p}, \quad (18.152)$$

where $R = qR^v + (1-q)R^d$ and c_p is given by (18.148).

Latent heat of evaporation and condensation

The latent heat of evaporation (or condensation) is the amount of energy that must be supplied to evaporate a unit mass of liquid to a vapour, or equivalently the amount of energy released when water vapour condenses. It is therefore equal to the difference between the specific enthalpy of water vapour and liquid water at a given temperature and also known as the enthalpy of vaporization. Using (18.141) the entropy and enthalpy of liquid water are given by

$$\eta^l = - \left(\frac{\partial g^l}{\partial T} \right)_p = \eta_0^l + c^l \ln(T/T_0), \quad h^l = g^l + T\eta = g_0^l + \alpha^l p + c^l T. \quad (18.153)$$

Using (18.137b) the enthalpy of water vapour is given by

$$h^v = g - T \left(\frac{\partial g^v}{\partial T} \right)_p = c_p^v T + L_0^v. \quad (18.154)$$

The enthalpy of vaporization, L , is therefore given by

$$L = h^v - h^l = L_0^v - (g_0^l + \alpha^l p) + (c_p^v - c^l)T. \quad (18.155)$$

The pressure term is negligibly small: $\alpha^l p \sim 100 \text{ J kg}^{-1}$ whereas, by experiment, $L = 2.501 \times 10^6 \text{ J kg}^{-1}$ at 0°C . Using this value for L_0^v (and setting $g_0^l = 0$), and with $c_p^v = 1859 \text{ J kg}^{-1} \text{ K}^{-1}$, $c^l = 4218 \text{ J kg}^{-1} \text{ K}^{-1}$, we obtain $L \approx (2.501 \times 10^6 - 2359 T) \text{ J kg}^{-1}$, with T in Celsius. The temperature dependence of L arises because we have to expend the same amount of energy evaporating water at (say) 10°C and then raising the vapour temperature to 20°C as we do in raising the water temperature from 10° to 20° and then evaporating it.

Using the latent heat provides a convenient expression for the entropy of a saturated mixture of dry air, water vapour and liquid water. If the mixture is in equilibrium then $q^v = q_s^v$ where q_s^v is given by the Clausius–Clapeyron relation and is a function of p and T . Furthermore, the differences in entropy between vapour and liquid water are then related to the latent heat of vaporization by

$$T(\eta^v - \eta^l) = h^v - h^l = L, \quad (18.156)$$

and (18.147) becomes

$$\eta_s^{dvl} = (1 - q^{vl})\eta^d + q^{vl}\eta^l + \frac{Lq_s^v}{T}. \quad (18.157)$$

Equivalent potential temperature

There are various definitions of equivalent potential temperature, θ_{eq} , with different names and small quantitative differences.⁷ For qualitative uses the differences don't really matter as they are small, and for an exact calculation one rarely needs to use equivalent potential temperature (entropy is often better). Nevertheless, it is a commonly used thermodynamic variable. The definition used in the main text was that equivalent potential temperature, θ_{eq} , is the potential temperature a parcel reaches after it is lifted adiabatically and at constant composition to a level at which it becomes saturated, and then all the water vapour is condensed and all of the latent heat released is used to heat the parcel. To obtain an analytic expression for θ_{eq} so defined we assume the condensation occurs at constant temperature and so obtain (18.54), namely

$$\theta_{eq}^p = \theta \exp\left(\frac{Lq}{c_p T}\right) = T \left(\frac{p_R}{p}\right)^{R/c_p} \exp\left(\frac{Lq}{c_p T}\right). \quad (18.158)$$

Here we give it a superscript p and call it the pseudo-adiabatic equivalent potential temperature. It is related to entropy, but it is not a true measure of it. One problem is that the above process is not physically realizable. In order to condense all the water vapour the temperature must be taken to absolute zero (or in practice we must lift the parcel to a great height where the temperature is very low), but we cannot simply condense all the water at a constant temperature. Second, we have neglected the contribution of liquid water. To see this, begin with an expression for the entropy of a mixture of dry air, water vapour and liquid water in equilibrium, with an entropy given by (18.157). An equivalent potential temperature, θ_{eq} , may then be defined by

$$(1 - q^{vl})c_p^d \ln(\theta_{eq}/T_0) = (1 - q^{vl})\eta^d + q^{vl}\eta^l + \frac{Lq_s^v}{T}. \quad (18.159)$$

Using the expressions for η^d and η^l given in (18.145) the above expression becomes

$$(1 - q^{vl})c_p^d \ln\left(\frac{\theta_{eq}}{T_0}\right) = (1 - q^{vl})c_p^d \left[\ln\left(\frac{\theta^d}{T_0}\right) - \frac{R^d}{c_p^d} \ln\left(1 - \frac{p^v}{p}\right) \right] + q^{vl}c^l \ln\left(\frac{T}{T_0}\right) + \frac{Lq_s^v}{T}. \quad (18.160)$$

Solving this for θ_{eq} gives

$$\theta_{eq} = \theta^d \exp\left(\frac{Lw_s}{c_p^d T}\right) \left(\frac{T}{T_0}\right)^{w^{vl}c^l/c_p^d} \left(1 - \frac{p^v}{p}\right)^{-R^d/c_p^d}, \quad (18.161)$$

where $w = q/(1-q^l)$ and $\theta^d = T(p_R/p)^{R^d/c_p^d}$. Equation (18.161) is a true measure of the entropy of a moist parcel, albeit a convoluted one. It differs slightly from (18.54) because θ_{eq}^p is not obtained by an adiabatic process, since the entropy of the liquid water is lost during condensation. If the water content is small ($q \ll 1$, $p^v/p \ll 1$) then (18.161) reduces to the more commonly used expression (18.158), with $q = q_s^v$.

Various other temperature-like quantities may be defined, notably the dew-point temperature and the wet-bulb temperature. The dew-point temperature is the temperature at which moist air, when cooled at constant pressure, becomes saturated. The wet-bulb temperature is the temperature that a parcel of air would have if cooled to saturation by the evaporation of water into it, with the energy required being supplied by the air parcel itself; wet-bulb temperature is directly measurable, for it is the temperature that a thermometer shows when wrapped in a wet cloth in a breeze. These quantities are both useful and have instinctive and human appeal. Potential temperature also has an intuitive attraction, but there are rarely objective reasons to use it in a quantitative calculation — entropy itself is usually a more straightforward alternative. This comment also applies to seawater.

APPENDIX B: EQUATIONS OF RADIATIVE TRANSFER

Consider a beam of radiation propagating through a thin slab of gas. Some of the incoming radiation may be absorbed, some may be scattered, and the slab may emit radiation of its own. Scattering is the change in direction of the radiation, so that it may reduce or — if radiation from other directions is scattered into the beam — amplify the beam's intensity. The difference between the incoming and outgoing radiation is

$$dI_\nu \equiv I_\nu^{\text{in}} - I_\nu^{\text{out}} = -d\tau_\nu I_\nu + dJ_\nu. \quad (18.162)$$

In this expression I_ν is the spectral radiance (power per unit area, per unit solid angle, per unit frequency interval) of the radiation, the term $d\tau_\nu I_\nu$ is the extinction (the absorption plus the radiation scattered away) and dJ_ν is the emission plus the scattering into the beam. The quantity $d\tau_\nu$ is the nondimensional *optical depth*; it may be written as $d\tau_\nu = k_\nu \rho ds$ where ds is the slab thickness, ρ is its density and k_ν is the extinction coefficient, a property of the gas in question. The minus sign on $d\tau_\nu$ is appropriate when τ increases in the direction of the beam, and all of the above quantities depend on the frequency, ν , of the radiation.

Suppose there is no scattering, which is a good approximation for infrared radiation. The emission of radiation is, in thermal equilibrium, then given by the Planck function, B_ν , multiplied by the optical depth. Equation (18.162) becomes

$$dI_\nu = -d\tau_\nu(I_\nu - B_\nu) \quad \text{or} \quad \frac{dI_\nu}{d\tau_\nu} = -(I_\nu - B_\nu). \quad (18.163)$$

This equation is the foundation of much of radiative transfer. If radiation is propagating in all directions we must integrate over solid angle to obtain the upward and downward spectral irradiances (power per unit area per unit frequency interval). This is a complicated procedure in general but, to a good approximation for infrared radiation in Earth's atmosphere, the simple upshot is the multiplication of the optical depth by a geometric, order-one (for example, 5/3) factor of γ — because most of the radiation is passing slantwise through the medium — and the multiplication of B by π because of the integration over a hemisphere, giving the two-stream approximation.²⁵ Equation (18.163) becomes

$$\frac{dF_\nu}{d\tau_\nu^*} = -F_\nu - \pi B_\nu, \quad (18.164)$$

where F_ν is the spectral irradiance along a vertical path along which τ increases and τ_ν^* is the 'scaled' optical depth given by $\tau^* = \gamma \tau_\nu$. We will drop the asterisk on τ_ν^* and we will absorb the factor π into the definition of B_ν .

In Earth's atmosphere it is common to choose τ increasing downwards, from 0 at the top of the atmosphere, although no physical result depends on this choice. The downwards (D_ν) and upwards (U_ν) irradiances are then

$$\frac{dD_\nu}{d\tau_\nu} = B_\nu - D_\nu, \quad \frac{dU_\nu}{d\tau_\nu} = U_\nu - B_\nu. \quad (18.165a,b)$$

These are the *two-stream equations*, without scattering, commonly known as the *Schwarzchild equations*. The downward and upward fluxes are uncoupled, because of the absence of scattering.

If there is no dependence of the optical depth on frequency then the medium is said to be *grey* and we may integrate (18.165) over frequency to give

$$\frac{dD}{d\tau} = B - D, \quad \frac{dU}{d\tau} = U - B, \quad (18.166a,b)$$

where U and D are the upward and downward (total) irradiances, $B = \sigma T^4$ (all three with units of W m^{-2}), and $\sigma = 5.6704 \times 10^{-8} \text{ W m}^{-2} \text{ K}^{-4}$ is Stefan's constant. The grey assumption is not accurate for Earth's atmosphere. Nevertheless, for conceptual or approximate calculations it is often useful to suppose that the atmosphere is grey in the infrared, in which case (18.166) applies to infrared radiation with separate equations (that might include scattering or reflection, but that might also be grey) for solar radiation; this is the *semi-grey* approximation.

APPENDIX C: ANALYTIC APPROXIMATION OF TROPOAUSE HEIGHT

Here we provide an approximate, analytic, expression for the height of the tropopause, given the optical depth and lapse rate. The idea is to solve for a self-consistent radiative-convective state, with a specified lapse rate extending upward to a tropopause and then transitioning to a radiative-equilibrium state, with the tropopause height being determined by the requirement of overall radiative balance.²⁶

Instead of trying to use the formal solutions to the Schwarzchild equations, it is easier to approximately solve the radiative-transfer equations for the upward long wave irradiance, U_L , *ab initio*. We make one other approximation, that the value of B/U_L varies linearly from the tropopause (where its value is 0.5) to its value at the surface (where $B/U_L = 1$). Thus,

$$\frac{B}{U_L} = 1 - \frac{z}{2H_T}. \quad (18.167)$$

Numerical calculations suggest this is a decent approximation. Proceeding, we write (18.65b) as

$$\frac{d \log U_L}{d\tau} = 1 - \frac{B}{U_L} = \frac{z}{2H_T}. \quad (18.168)$$

Using $\tau(z) = \tau_s \exp(-z/H_a)$ we obtain

$$\frac{d \log U_L}{dz} = -\frac{z}{2H_T H_a} \tau_s \exp(-z/H_a). \quad (18.169)$$

We can integrate this expression by parts to obtain a value of the upwelling radiation at the tropopause $U_L(H_T)$, namely

$$\log \left(\frac{U_L(H_T)}{U_L(0)} \right) = -\frac{\tau_s}{2H_T} \int_0^{H_T} \exp(-z/H_a) dz \approx -\frac{\tau_s H_a}{2H_T}, \quad (18.170)$$

for $H_T \gg H_a$. This is an expression for the upwelling longwave radiation, $U_L(H_T)$, and the unknowns in the equation are $U_L(0)$, the upwelling radiation at the surface, and the tropopause height,

H_T . The value of $U_L(0)$ is given by the surface temperature, which is a function of the tropopause temperature, T_T , the specified lapse rate, Γ , and the tropopause temperature, T_T . Inverting the argument, if we are given $U_L(H_T)$ and T_T then we can calculate H_T .

Let us assume that the stratosphere is optically thin, in which case $U_L(H_T)$ is the outgoing longwave radiation, equal to $2\sigma T_T^4$, and this is known because it equals the incoming solar radiation. Thus, we can in principle now solve (18.170) for H_T . To do this note that the upwelling radiation at the surface is given by $U_L(0) = \sigma T_g^4 = \sigma T_s^4$, where $T_s = T_T + \Gamma H_T$. The left-hand side of (18.170) then becomes

$$\log\left(\frac{2\sigma T_T^4}{\sigma T_s^4}\right) = \log 2 + 4 \log \frac{T_T}{T_s} = \log 2 + 4 \log\left(\frac{T_T}{T_T + \Gamma H_T}\right) \approx \log 2 - \frac{4\Gamma H_T}{T_T}. \quad (18.171)$$

The rightmost terms in (18.171) and (18.170) are approximately equal so that

$$\log 2 - \frac{4\Gamma H_T}{T_T} = -\frac{\tau_s H_a}{2H_T} \quad \text{or} \quad 8\Gamma H_T^2 - C H_T T_T - \tau_s H_a T_T = 0, \quad (18.172)$$

where $C = 2 \log 2 \approx 1.39$. The solution of this equation is

$$H_T = \frac{1}{16\Gamma} \left(CT_T + \sqrt{C^2 T_T^2 + 32\Gamma \tau_s H_a T_T} \right). \quad (18.173)$$

The tropopause height given by this equation is fairly close to the actual solution of the radiative-convective equations, obtained by numerically integrating the Schwarzschild equations and iterating to obtain the correct tropopause height following the algorithm of Section 18.6.2, as seen in Fig. 18.21. The analytic approximation may be further improved with a bit of effort, but even in the form above it captures the essential aspects of the true solution.

Notes

- 1 Many thanks to Adam Sobel for a number of conversations and notes that informed the sections on convection and quasi-equilibrium, and to Will Beeson and his colleagues in Chicago for many useful comments. I am also grateful to Brian Mapes for a detailed critique of this chapter; I was unable to address all of his concerns but his point of view was salutary.
- 2 I use a formula given by Bolton (1980), which is a variant on the original Magnus formula, *aka* the August–Roche–Magnus formula or the Tetens formula — see Lawrence (2005) for discussion.
- 3 Measurements come from the hybrid advanced microwave sounding unit, Atmospheric Infrared Sounder (AIRS), as in Sherwood *et al.* (2010).
- 4 To read more about models of this form see Pierrehumbert *et al.* (2007), with extensions and applications by O’Gorman *et al.* (2011), Sukhatme & Young (2011), Tsang & Vanneste (2016) and others.
- 5 The simulations here used a model developed by Dr. Yue-Kin Tsang, and I am grateful to him for discussions and help.
- 6 The interested reader might start with Sherwood *et al.* (2010) or Schneider *et al.* (2010) and go forward and back from there.
- 7 This definition of θ_{eq} is sometimes called the ‘pseudo-equivalent potential temperature’, because the condensation product, liquid water, is assumed to fall out of the air parcel and the process is ‘pseudo-adiabatic’, not adiabatic. Other names with slightly different definitions exist (Betts 1973, Emanuel 1994, Ambaum 2010).
- 8 The notion of conditional instability has been with us for many years — it appears in Haurwitz (1941) for example — and an influential form was introduced Ooyama (1963) (see Ooyama 1982) and Charney & Eliassen (1964). They proposed models of a cooperative mechanism between the

convection and larger scale flow, with the convection producing a convergence at low levels, leading to ascent, more latent heat release and convection and so on. This mechanism and variations about it became known as 'Conditional Instability of the Second Kind', or cISK, to distinguish it from more conventional conditional instability ('of the first kind') which does not involve such a feedback with the large scale. The cISK mechanism tends to produce vertical updraughts and may be important for hurricane growth, but this remains a topic of some debate (e.g., Raymond 1995, Emanuel 1994, Smith 1997). Other theories of hurricanes tend to de-emphasize the cISK mechanism in favour of model involving a feedback between wind speed and evaporation, called wind-induced surface heat exchange or WISHE, with stronger winds giving more evaporation, leading to saturation and thence convection. The convection then mainly serves to establish a moist adiabatic lapse rate (as in the quasi-equilibrium ideas of Section 18.4.3) which ties the boundary layer to a warm core extending upward (e.g., Craig & Gray 1996).

- 9 Convective adjustment is a great simplification over what actually occurs, but nevertheless it is a useful concept and the basis of many early parameterization schemes for numerical models of the atmosphere. Convective adjustment was introduced into modelling by Manabe & Strickler (1964) and a theoretical and observational discussion of the general problem was given by Ludlam (1966). A popular variation (a relaxation rather than an adjustment) was proposed by Betts (1986) and Betts & Miller (1986). These days (c. 2017) GCMs rarely use simple convective adjustment or relaxation schemes, although the underlying ideas endure.
- 10 The nature of a 'quasi-equilibrium' between convection and large-scale forcing was made explicit by Betts (1973) and Arakawa & Schubert (1974), with precedents to be found in Scorer & Ludlam (1951) and Ludlam (1966). Many papers have since followed, with, for example, an extended discussion in Emanuel *et al.* (1994), a counterpoint, examples and further discussion in Mapes (1997, 1998, 2000), and application to the boundary layer in Raymond (1997). More references can be found in the reviews by Arakawa (2004) and Emanuel (2007), and debates continue about applicability and efficacy.
- 11 This behaviour is seen in many comprehensive General Circulation Models of the atmosphere, going back to Manabe & Wetherald (1980) and beyond. The explanation given here follows Vallis *et al.* (2015) but has much earlier roots.
- 12 Numerical calculations of the radiative constraint with more realistic treatments of radiation were carried out by Thuburn & Craig (2000).
- 13 We roughly follow Emanuel (1987). Neelin & Zeng (2000) and Zeng *et al.* (2000) give details of how a functional reduced model, including diabatic and frictional terms, may be constructed. Lindzen & Nigam (1987) have another, albeit related, take on the problem.
- 14 The tropical scaling was presented by Charney (1963).
- 15 The derivation of the weak temperature-gradient approximation given here is more-or-less that of Sobel *et al.* (2001), and may be regarded as an extension of Charney's (1963) ideas to include diabatic effects. A number of authors previously used the approximation, implicitly or explicitly, in one form or another (e.g., Neelin 1988, Browning *et al.* 2000), and various extensions and rigour have been added by Majda & Klein (2003) and others.
- 16 See Bretherton & Smolarkiewicz (1989) and Mapes (1997) and go from there.
- 17 Details are described in Bretherton & Sobel (2002) and Neelin & Zeng (2000).
- 18 A review of convectively coupled equatorial waves is provided by Kiladis *et al.* (2009). The MJO was first described by Madden & Julian (1971, 1972), and a review may be found in Zhang (2005). Schubert & Masarik (2006) discuss aspects of a moving Matsuno–Gill pattern and its relation to the MJO, and Raymond & Fuchs (2009), Majda & Stechmann (2009) and Sobel & Maloney (2013), among others, offer theoretical models of the MJO.
- 19 Diagrams such as these are known as Wheeler–Kiladis diagrams, after Wheeler & Kiladis (1999). Filtering the noise is required to obtain clean plots and requires some attention.
- 20 Adapted from Kiladis *et al.* (2009).

- 21 Dias & Kiladis (2014) discuss reasons why the value is smaller than might be expected.
- 22 Adapted from Kiladis *et al.* (2005).
- 23 Heckley & Gill (1984).
- 24 Feistel *et al.* (2010) and Thuburn (2017) give more details. I am grateful to John Thuburn for a number of very useful conversations on this matter.
- 25 There are various versions of the two-stream approximation; see Goody & Yung (1995) or Pierrehumbert (2010).
- 26 The contents of this appendix are the results of joint work with Pablo Zurita-Gotor. A seemingly casual question to me by Rich Kerswell led to the theoretical development.