2 Energy and Radiation in Planetary Atmospheres

2.1 Energy Sources and Fluxes on Planets

2.1.1 Planetary Energy Sources

All planetary atmospheres are affected by the supply of free energy (i.e., energy that can do work and be dissipated), which includes the following possibilities: (1) radioactive decay, (2) accretional energy, (3) energy released by internal differentiation of the planet, such as core formation, (4) tidal energy, and (5) light from a parent star. A combination of energy sources (1)–(4) tends to make planetary interiors hotter than planetary surfaces, so that heat is transferred outwards to a planet's exterior. But this heat flux at the planets' exterior is usually much smaller than absorbed stellar insolation, except for gas giants such as Jupiter (where internal heat is nearly comparable) or extrasolar bodies with great tidal heating. Despite the general difference in the magnitude of fluxes, stellar radiation does not induce the high temperatures found in planetary interiors because internal heat only escapes slowly through solid material whereas stellar energy is balanced by infrared radiation that is usually readily emitted back to space. Exceptions are steam atmospheres, which are highly opaque to infrared, and strongly irradiated exoplanets.

Radioactive energy from planetary interiors is produced by the decay of radiogenic elements, of which the main three on Earth are uranium (U), thorium (Th) and potassium (K). By considering conservation of energy, the change in internal heat must equal the heat input from radioactive decay minus the heat loss Q_{hf}, from the surface of the planet. On Earth, the global average for Q_{hf} is 0.087 W m^{-2}, whereas on the Moon, Q_{hf} is about 0.015 W m^{-2} (Schubert *et al.*, 2001).

Geothermal activity and outgassing should scale with heat flow Q_{hf}. However, the scaling with Q_{hf} may be different for each planet, depending on its tectonic regime and the thickness of its lithosphere, which is the outer rigid layer within which heat is transferred by conduction rather than convection. For the Earth, volcanic outgassing is thought to scale roughly with Q_{hf} to a power between 1 and 2. The upper value comes from the idea that outgassing depends on the rate of seafloor area creation, which theory suggests should scale as Q_{hf}^2 (Schubert *et al.*, 2001, p. 597). Since Q_{hf} was certainly greater in the past, then the rate of volcanism and outgassing would have been correspondingly greater, with consequences for atmospheric composition.

Accretional energy comes from the conversion of kinetic energy during giant impacts when planets were formed by the coalescence of smaller bodies. Accretional heat was enough to melt terrestrial planets (Newsom and Sims, 1991). When a terrestrial planet melts, its molten exterior becomes a magma ocean, and this overlies an iron-rich core that is at least partly molten. The surface of this magma ocean should chill and solidify because of radiation of heat to space. However, the skin would be continuously ruptured by the vigorous convection of hot material underneath.

Differentiation, such as core formation. During terrestrial planetary formation, iron sinks to a planet's core because iron is dense, and this causes heating as gravitational energy is converted to thermal kinetic energy. We consider this source of internal energy in more detail in Ch. 6 when we discuss the origin of the Earth and planets.

Tidal heating. Just as the Earth's oceans experience tides due to the pull of the Moon and the Sun, so the solid Earth itself is distorted by these forces and produces tides with amplitude up to 10 cm. This squeeze and release generates a small amount of heat that is often neglected for the Earth. However, tidal energy dissipated in the oceans due to lunar and solar tides is ~10% of the surface heat flow. In Earth's long-term evolution, tidal dissipation from ocean tides is important. Because of oceanic tidal friction, the Earth's rotation rate is slowing down (see Sec. 6.8.2).

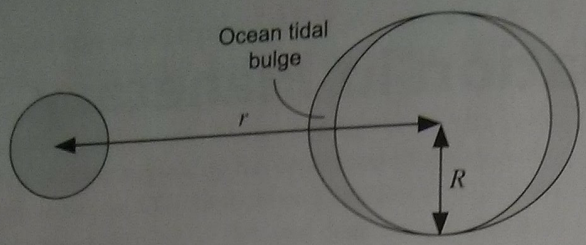


Figure 2.1 The Moon and Earth separated by distance *r*, illustrating tidal forces. Both the Earth and Moon orbit a common center of mass. In a co-rotating frame the Moon pulls the ocean on the Earth's near side more than the centrifugal force of orbital motion about the center of mass pulls the ocean away. On the far side of the Earth, the centrifugal force of orbital motion about the center of mass pulls the ocean outward relative to the lunar gravitational force. The net effect produces tidal bulges on both near and far sides.

For satellites of the outer planets, tidal heating can dominate the heat flow from the interior. Tidal heat makes Jupiter's moon, Io, the most volcanically active body in the Solar System, while Europa's tidal heating provides an energy source of potential importance for possible life in the Europa's probable subsurface ocean.

For bodies separated by radial distance, r, from center to center of each body (Fig. 2.1), the tidal amplitude, which arises from a differential force, is proportional to $1/r^3$. The gravitational force (F) produced by a body of mass M on a 1 kg test mass is given by Newton's law of gravity, $F = GM/r^2$, where G is the universal gravitational constant. The dependence of tidal force on distance r is given by differentiation:

$$\delta F = \frac{dF}{dr}\delta r = \frac{d}{dr}\left(\frac{GM}{r^2}\right)\delta r \propto \frac{M}{r^3} \tag{2.1}$$

Because the size of the tidal force varies as $1/r^3$, a small change in separation distance r has a large effect on the size of the tidal bulge.

The amount of energy dissipated by tides can be big. For exoplanets with very eccentric orbits close to their parent stars, tidal heating can be large enough to significantly affect climate and drive atmospheric escape (Barnes *et al.*, 2013). Tidal energy dissipation $\dot{E}_{tide}$ (in Watts) for a synchronously orbiting satellite is a much stronger function of separation distance than the tidal force and depends on $1/r^6$ (Ferraz-Mello *et al.*, 2008; Goldreich and Soter, 1966; Murray and Dermott, 2001, p. 173)

$$\dot{E}_{tide} = \frac{\omega E_{\max}}{Q_{factor}} \propto \frac{GM^2 r_{sat}^5 e^2}{Q_{factor} r^6} \tag{2.2}$$

Here, ω is the angular frequency of flexing (often denoted in the literature as the mean motion "n," which is 2π divided by the orbital period), $E_{\max}$ is the peak potential energy stored in the tidal distortion, e is the eccentricity of an orbit, r_{sat} is the radius of a satellite about the large body of mass M, and Q_{factor} is the tidal dissipation parameter (or "quality factor") of the satellite. The Q_{factor} parameter describes the inverse ratio of the energy lost per cycle to the maximum potential energy stored in the tidal distortion:

$$Q_{factor} = \frac{2\pi E_{\max}}{\Delta E} \tag{2.3}$$

The Q_{factor} parameter increases with lower loss and is the number of cycles for the energy to decrease by $1/e$. For example, the Q_{factor} for Io is 36 and for Jupiter it is ~3×10^4; for dry bodies the Q_{factor} tends to be around 50–100 (Segatz *et al.*, 1988).

Stellar radiation is the fifth energy source, which we discuss at length below because light from the host star strongly influences a planet's climate.

2.1.2 Radiation From the Sun and Other Stars

2.1.2.1 *Spectral Types*

Our Sun is a star. Stars differ in their luminosity and color, and the latter indicates the temperature of their outer atmospheres. Stellar temperatures and spectra are important when we consider planetary habitability in other solar systems in Ch. 15. From hot, blue-white to relatively cold infrared, the letters OBAFGKM[1] designate stellar spectral types. Cooler R- and N-type stars (today referred to as carbon, or C-type, stars) and S-type stars are additional stellar classes. But the C- and S-type stars are invariably giant or supergiant stars and differ substantially from the others in having peculiar heavy-metal abundances.

Three spectral types, L stars (very low mass red dwarfs and brown dwarfs) and T and Y stars (brown dwarfs) have been added since the 1990s. Sky surveys found that L and T stars were abundant and that even cooler Y stars exist (Burgasser *et al.*, 2002; Cushing *et al.*, 2011; Kirkpatrick *et al.*, 1999). These stars are comparatively cool (<2500 K) and emit radiation mostly at infrared wavelengths. L stars have deep H_2O absorption bands, but no CH_4 absorption except for late L stars; T dwarfs have H_2O and CH_4 features, while Y stars (~300 K to 500 K) have H_2O, CH_4 and NH_3 bands (Helling and Casewell, 2014). Overall, the complete sequence of spectral types is now OBAFGKMLTY.

Amongst the spectral types are some famous stars visible to the naked eye. Sirius, the brightest star in our sky is A-type, Polaris (the pole star) is an F star, Aldebaran is a K star, and Betelgeuse is an M supergiant.

[1] Generations of students have remembered stellar spectral types with the mnemonic: "Oh Be A Fine Girl/Guy Kiss Me". The addition of LTY dwarfs changes the ending to "My Lips Tonight, Yahoo!".

But we cannot see the most common type of star with the naked eye, namely M dwarfs, which are red dwarfs.

Types O–M are subdivided into ten subcategories. For example, the hottest B star is a B0, followed by B1, B2, and so on; B9 is followed by A0. Our Sun is a G2-type star. For historical reasons, the hotter stars are called early types and the cooler stars late types, so that "B" is earlier than "F," for example. Table 2.1 shows how an *effective temperature* characterizes each spectral type (see Box 2.1). This temperature applies to a star's

Table 2.1 Effective temperature (T_{eff}) as a function of the spectral type for main sequence stars.

Spectral type	O5	B0	A0	F0	G0	K0	M0	L0
T_{eff} (K)	40000	25000	11000	7600	6000	5100	3600	2200

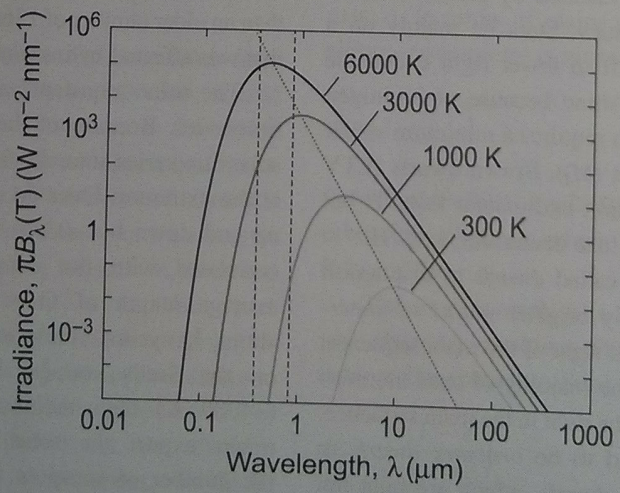
Box 2.1 Blackbody Spectra and "Effective Temperature"

Every object gives out a spectrum of radiation by virtue of its finite temperature. We can see this in a coal fire: the hottest parts are blue and emit a wavelength that is shorter than that coming from the parts that glow red or even cooler parts that radiate in the infrared.

A blackbody spectrum is an idealized spectrum that depends only on the temperature of the source. The temperature uniquely determines the intensity of emitted radiation as a function of wavelength, including the wavelength at which the intensity is a maximum. Moreover, when the radiation output is summed over all wavelengths, the total flux in Watts per square meter depends only on temperature.

Historically, the concept of blackbody radiation derived from considering the radiation field within a closed cavity that perfectly absorbs and emits radiation at all wavelengths. At equilibrium, production and loss of radiation balance and the intensity of the radiation field is uniform. Max Planck (1858–1947) found the spectrum to be uniquely related to the cavity wall temperature, T, by a mathematical expression, the *Planck function*, denoted by the symbol, $B_\lambda(T)$ where the λ subscript indicates that B_λ is specified at a particular wavelength, λ.

The graph below shows blackbody spectra at different temperatures from planets (300 K) to the Sun (~6000 K). A higher temperature shortens the wavelength at which the flux is maximal (where curves intersect the dotted line) and increases the total flux, which is the area under each curve. Dashed lines enclose visible wavelengths.



As a first approximation, celestial bodies such as the Earth or Sun are often considered as blackbodies, i.e., objects that give out blackbody spectra of radiation characterized by a single radiating temperature. This temperature is 5780 K for the Sun and 255 K for a mean radiating temperature of the Earth's troposphere. In reality, the spectral output of the Sun and Earth are complicated by species that absorb and emit radiation in their atmospheres. But the total flux output can still be characterized by an equivalent blackbody temperature, which we call the *effective temperature*.

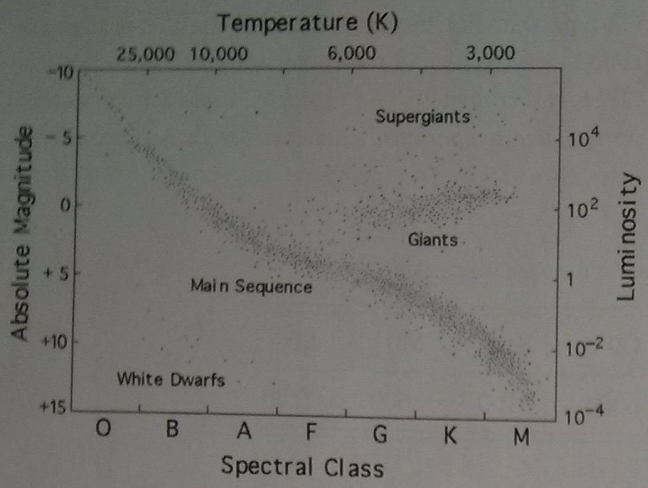


Figure 2.2 Hertzsprung–Russell diagram for nearby stars, with the stellar luminosity plotted relative to a solar luminosity on the right-hand vertical axis of 1. The absolute magnitude is a scale proportional to $\log_{10}$(luminosity), where one step corresponds to a change in brightness of factor $100^{1/5} \approx 2.512$. (Source: NASA.)

photosphere, which is the outer region of the star's atmosphere where most of the emitted electromagnetic energy originates. Although a photosphere is gaseous, the photosphere is also called a star's "surface."

A plot of stellar luminosity against effective temperature is the *Hertzsprung–Russell* (H-R) *diagram*, where temperature is always plotted "backwards" from high to low on the horizontal axis so that it runs in the same sense as the series of spectral types. Figure 2.2 shows an H-R diagram. Nearby stars are plotted because distances to nearby stars can be readily measured by parallax, which gives their absolute luminosities. Stars lie mainly on a diagonal band from upper left to lower right called the *main sequence*. These stars shine because of hydrogen fusion in their interiors, which requires a minimum stellar mass of 75–80 Jupiter masses (M_J). Brown dwarfs (LTY types) are below this mass limit and cannot fuse ^{1}H but those heavier than $13M_J$ can fuse deuterium, i.e., ^{2}H.

Main sequence stars are called *dwarfs* to distinguish them from larger, intrinsically brighter *giants* or *supergiants* that lie above and to the right of the main sequence. The Sun is a G dwarf, and the majority of stars are even smaller dwarfs and farther to the right in the main sequence. Main sequence stars are said to be *ordinary dwarfs* to distinguish them from *white dwarfs*, which are faint hot objects located below and to the left of the main sequence. The white dwarfs are stellar remnants of mass less than 1.4 solar masses. Thus, there are G giants, G ordinary dwarfs, and even G white dwarfs all with similar effective temperature. However, the giants are much more luminous than the dwarfs because of greater stellar radii.

Main sequence stars with approximately solar composition have an empirical *mass–luminosity relation* between stellar mass M_* and luminosity L_*, which Salaris and Cassisi (2005, p. 139) give as

$$\begin{aligned} L_* &\propto M_*^{4.5} \quad (\sim 0.5M_\odot < M_* < 2M_\odot) \\ L_* &\propto M_*^{3.6} \quad (2M_\odot < M_* < 20M_\odot) \end{aligned} \tag{2.4}$$

where $M_\odot$ is the solar mass. The Sun is made up of about 73% hydrogen by mass with most of the remainder helium, and only ~2% of heavier elements.

2.1.2.2 *The Solar Constant: Solar Radiation at the Top of Planetary Atmospheres*

The Sun's properties and location are important in the way they affect planetary atmospheres of the Solar System. Some key solar parameters are follows:

$$\left.\begin{aligned} &\text{mass} = M_\odot = 1.99 \times 10^{30}\ \text{kg} \\ &\text{luminosity} = L_\odot = 3.83 \times 10^{26}\ \text{W} \\ &\text{radius} = R_\odot = 6.96 \times 10^{8}\ \text{m} \end{aligned}\right\} \tag{2.5}$$

At one Astronomical Unit (AU), the mean distance of the Earth from the Sun, the Sun's energy spreads over the surface area of a sphere, $4\pi r^2$, with radius $r = 1$ AU $= 1.495\ 98 \times 10^{11}$ m. Consequently, the *solar constant*, which is the solar flux at Earth's mean orbital distance, is $L_\odot/(4\pi r^2) = (3.83 \times 10^{26}\ \text{W})/(4\pi \times (1.495\ 98 \times 10^{11}\ \text{m})^2) = 1.36 \times 10^3\ \text{W m}^{-2}$. More precisely, satellites orbiting the Earth measure the solar constant as $1360.8 \pm 0.5\ \text{W m}^{-2}$ (Kopp and Lean, 2011), consistent with ground-based inferences (Chapman *et al.*, 2012). This value is somewhat less than an older number of $1366\ \text{W m}^{-2}$ (Lean and Rind, 1998) that was affected by instrumental issues of stray light.

The solar "constant" also changes over a variety of timescales. Because of the eccentricity of Earth's orbit, solar flux varies annually from $1316\ \text{W m}^{-2}$ to $1407\ \text{W m}^{-2}$ at the extremes. Over decades, the solar constant cycles up and down by ~0.1%, a small amount. The cycle is correlated with the number of sunspots, where the average length of time between sunspot maxima is about 11 years. *Sunspots* are relatively dark regions on the Sun's surface. Because sunspots are cooler (~4000 K) than their surroundings (~6000 K), one might expect the solar luminosity to decrease with the number of sunspots. However, the opposite occurs (Fig. 2.3) because bright regions known as *plages* (pronounced "plah'jes") surround sunspots and bright, hot patches called *faculae* accompany sunspots.

On timescales of millennia, solar variability (the sunspot number) is deduced from isotope proxies. The flux of cosmic rays is modulated by solar activity and carbon-14

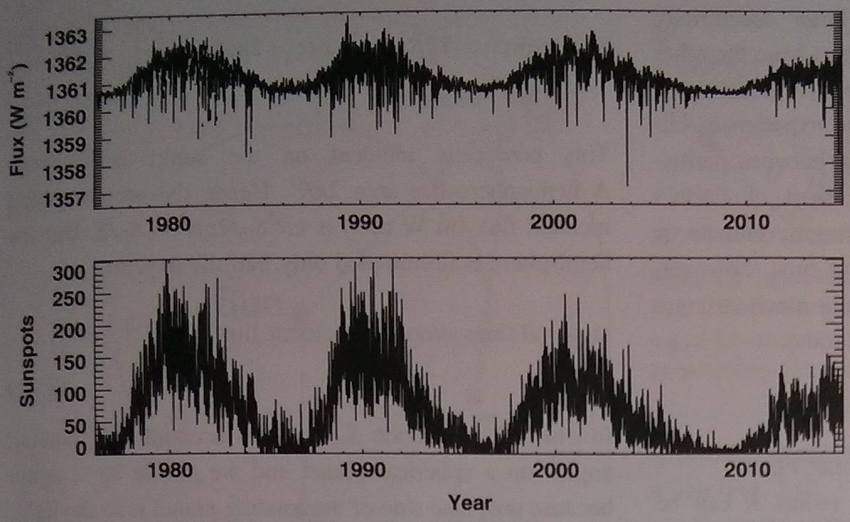


Figure 2.3 The correlation of daily solar irradiance at 1 AU and sunspot number for the years 1976–2014. Flux data are from PMOD (Physikalisch-Metorologisches Observatorium Davos), Switzerland, using measurements from the ESA/NASA SoHO (Solar and Heliospheric Observatory) mission. Sunspot data are from SILSO (Sunspot Index and Long-term Solar Observations), Royal Observatory of Belgium.

and beryllium-10 form when these rays hit the atmosphere (Beer, 2000). Solar activity is inferred from ^{14}C in tree rings and ^{10}Be in ice cores over the past 11 000 years (Solanki *et al.*, 2004), but the effect on the solar constant and Earth's climate is modest (Solanki and Krivova, 2003). However, over tens of thousands of years, the solar flux at the top of a planet's atmosphere is affected by changes in orbital parameters of a planet, according to Milankovitch theory. This is important for the climate of Earth (Ch. 11) and Mars (Ch. 12). On yet longer timescales of millions and billions of years, the solar luminosity (and luminosity of other main sequence stars) gradually increases, as described in Ch. 11.

2.1.2.3 *The Solar Spectrum*

The spectrum of sunlight consists of a continuum from gamma rays to radio waves with a peak flux around 0.5 μm and a superimposed line structure. In the visible and infrared, there are thousands of absorption lines, known as *Fraunhofer lines*, characteristic of chemical elements in the photosphere. For ultraviolet wavelengths shorter than 185 nm, lines tend to be seen in emission. With a solar constant of 1361 W m^{-2}, a blackbody spectrum at an effective temperature of 5780 K corrected for the distance of a planet provides a good approximation to the visible and infrared portions of the solar spectrum. For ultraviolet wavelengths of 210–300 nm, the effective temperature of the Sun tends to be lower, ~5000 K. However, a spike in flux occurs at 121 nm, well above the Sun's nominal blackbody flux, which is the *Lyman-α* wavelength, corresponding to the transition between ground and first excited state of hydrogen atoms. Lyman-α radiation is important for providing energy to the upper atmospheres of planets and for driving chemistry, e.g., breaking up methane molecules in the stratospheres of giant planets, Titan, or the early Earth.

2.2 Planetary Energy Balance and the Greenhouse Effect

2.2.1 Orbits and Planetary Motion

Variations in *solar insolation*, which is the flux of energy per unit area (W m^{-2}), affect the radiation balance that sets planetary temperatures. Insolation changes with orbital position. Figure 2.4 shows a general planetary orbit, with a semi-major axis a, and semi-minor axis b. From Fig. 2.4, it can be seen that the *orbital eccentricity* e is defined by the ratio of the distance $a \times e$ from the center of the ellipse to either focus divided by the semi-major axis, a. In terms of major and minor axes of the orbital ellipse, eccentricity is given by

$$e^2 = 1 - \left(\frac{b}{a}\right)^2 \tag{2.6}$$

For circular orbits, $e = 0$. By Kepler's first law, planetary orbits are elliptical, which means that $b/a < 1$ so that $0 < e < 1$.

The *true anomaly* f_t is the angle between perihelion and planetary position (Fig. 2.4). This angle sets the heliocentric (planet-to-Sun) distance $r_\odot$, as follows

$$\left.\begin{aligned} r_\odot &= \frac{a(1-e^2)}{1+e\cos f_t} \\ r_{\odot,perihelion} &= a(1-e) \\ r_{\odot,aphelion} &= a(1+e) \end{aligned}\right\} \tag{2.7}$$

The flux at a planet varies inversely with the square of the heliocentric distance, $r_\odot$.

The orbital eccentricity value of different Solar System planets varies widely, and larger values have

consequences for climates. Pluto's large eccentricity $e = 0.25$ sets this Kuiper Belt Object apart from the other planets, and Pluto's orbital position drastically affects its atmosphere. Mars, with $e = 0.093$, also experiences climatically significant solar flux variation between perihelion and aphelion. In contrast, the effect of Earth's eccentricity of $e = 0.017$ on our current climate is modest. Earth's eccentricity varies over time, however, from 0 to 0.06, so eccentricity has had a much stronger effect on climate in the past than it has today.

2.2.2 Time-Averaged Incident Solar Flux

The flux of sunlight on a planet of radius R can be calculated by considering Fig. 2.5(a). In this diagram, an elemental ring of planetary surface is defined by angle ϕ, which goes from 0 to $\pi/2$ radians. The element of surface has area $2\pi(R\sin\phi)Rd\phi$, and the component of sunlight normal to the surface is $S_\odot \cos\phi$, where $S_\odot$ is the solar flux at the top of the atmosphere. Consequently, the power integrated over a hemisphere is

$$\text{total power [Watts]} = 2\pi R^2 S_\odot \int_0^{\pi/2} \cos\phi \sin\phi \, d\phi \qquad (2.8)$$

We can substitute $x = \sin\phi$ and $dx = \cos\phi \, d\phi$, so that eq. (2.8) can be rewritten as:

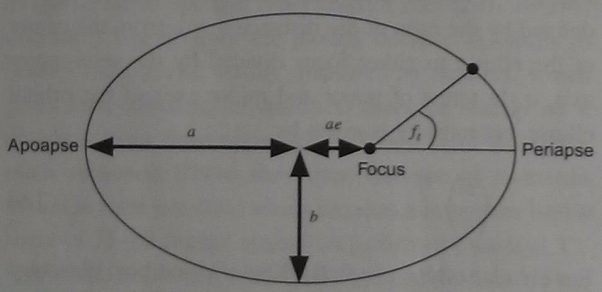


Figure 2.4 Basic elements of an elliptical orbit, showing semi-major axis a, semi-minor axis b, and true anomaly f_t. The eccentricity of the orbit is e.

$$\text{total power} = 2\pi R^2 S_\odot \int_0^1 x\,dx = 2\pi R^2 S_\odot \left[\frac{x^2}{2}\right]_0^1 = \pi R^2 S_\odot \qquad (2.9)$$

This power is incident on the sunlit hemisphere. A hemisphere has area $2\pi R^2$. Hence the area-averaged incident flux (in W m^{-2}) is $\pi R^2 S_\odot / 2\pi R^2 = S_\odot/2$. But the hemisphere is illuminated only half the day, so

$$\text{area-and time-averaged incident flux } \left[\text{W m}^{-2}\right] = \frac{S_\odot}{4} \qquad (2.10)$$

In essence, we divide $S_\odot$ by 2 to account for glancing angles on a spherical planet and we divide by 2 again because only one side of the rotating planet is in daylight.

Another way to understand eq. (2.10) recognizes that the factor of 1/4 represents the ratio of disk to sphere areas ($\pi R^2/4\pi R^2$). The instantaneous solar flux intercepted by a planet given by eq. (2.9) is equivalent to that on a projected disk of area πR^2 (proved by eq. (2.9)) compared with a total planetary area of $4\pi R^2$ (Fig 2.5(b)). For the Earth, the globally averaged insolation at the top of the atmosphere given by eq. (2.10) is (1360.8±0.5 W m^{-2})/4 = 340.2±0.1 W m^{-2}, where $S_\odot$ is from Kopp and Lean (2011).

2.2.3 Albedo

Solar System planets reflect sunlight by different amounts. Consequently, the energy flux given by eq. (2.10) is not entirely absorbed by a planet's surface or atmosphere. The fraction of incident power that is reflected is the *albedo*. There are different ways of expressing the albedo. The *monochromatic albedo* is the fraction of incident power that gets reflected or scattered back to space at a given frequency of light:

$$A_\nu = \frac{(\text{reflected or scattered power at frequency } \nu)}{(\text{incident radiation power at frequency } \nu)} \qquad (2.11)$$

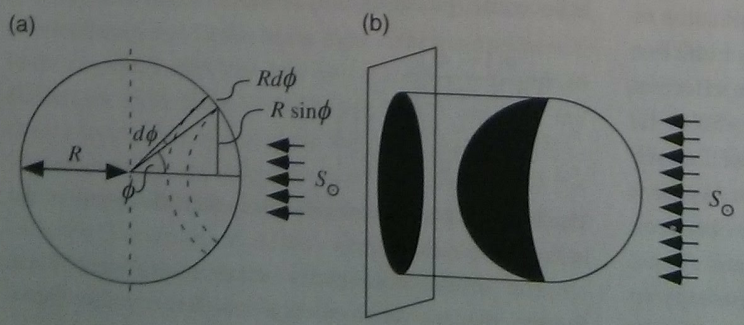


Figure 2.5 (a) An elemental ring of surface defined by angle ϕ on a planet of radius R. The incident solar flux is $S_\odot$. (b) The equivalent area of intercept for the solar flux is a projected disk of area πR^2 compared with a total sphere area of $4\pi R^2$.

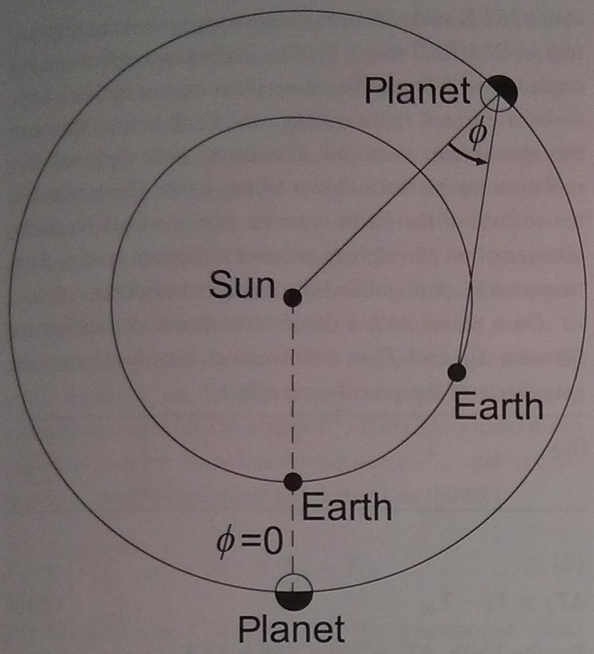


Figure 2.6 The phase angle ϕ for determining geometric albedo is the angle between lines connecting the observer, planet and Sun. *Quadrature* is when the angle subtended at the Earth by the directions to the Sun and a superior planet is 90°.

If we integrate this over all frequencies, we get the *Bond albedo* named after astronomer George Bond (1825–1865) and synonymous with *planetary albedo*,

$$A_b = \frac{\text{(total reflected or scattered radiation power)}}{\text{(incident radiation power)}} \tag{2.12}$$

For climate calculations, we use the Bond albedo.

In astronomy, a *geometric albedo* is often used. For Solar System planets viewed from the Earth, the phase angle ϕ, as shown in Fig. 2.6, is the angle between incident sunlight and radiation received on Earth. When ϕ = 0, sunlight is observed in pure backscatter. The geometric albedo is defined by

$$A_g = \frac{F(\phi = 0)}{F_{\text{Lambert-disk}}} = \frac{\text{(reflected flux at zero phase angle)}}{\text{(flux reflected from a Lambertian disk of the same cross-section)}} \tag{2.13}$$

Here, $F_{\text{Lambert-disk}}$ is the flux reflected by a disk with a Lambertian surface of the same cross-section as the planet at the same distance from the Sun. A Lambertian surface is one that reflects all incident radiation isotropically, as illustrated in Fig. 2.7, such that its brightness is the same in all directions of view. For example, a wall painted in matt white is very roughly Lambertian. Thus, the geometric albedo is the fraction of incident light reflected in the direction of the observer of an outer planet measured at *opposition*, meaning when the Sun, Earth and planet form a line (Fig. 2.6).

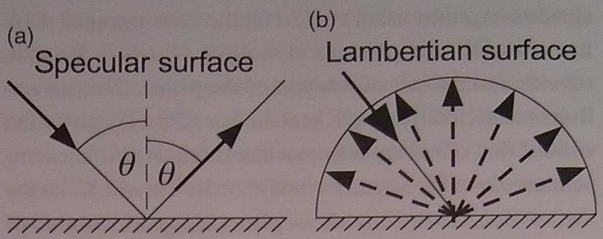


Figure 2.7 Different types of reflection. (a) Pure *specular reflection*, where the angle of incidence equals the angle of reflection, e.g., a mirror. (b) A *Lambertian* surface reflects radiation evenly in all directions. This surface is to be *diffusely* reflecting.

Geometric albedo is related to the Bond albedo by the expression:

$$A_b = qA_g \tag{2.14}$$

where q is a *phase integral* which reflects the variation of the intensity of radiation over the phase angle ϕ. Karttunen (2007), pp. 149–151, and Seager (2010), pp. 33–39, provide a step-by-step derivation of the phase integral. We note that if the sphere is perfectly reflecting (Bond albedo A_b is unity) and Lambertian (isotropic scatterer) the phase integral q is 1.5 and the geometric albedo is 2/3. Thus 1/3 of the radiation is scattered out of the line-of-sight. In general, even if the Bond albedo is not unity, the geometric albedo is 2/3 of the Bond albedo if the planet is a spherical Lambertian scatterer. Bond and geometric albedos are equal when q =1, which means that reflection from the planet behaves like a Lambertian disk of the same diameter.

Finally, in photometry, reflectance observations of planetary surfaces and atmospheres are often given as a *radiance factor*, *I*/*F*, said as "I over F." The radiance factor is the reflected intensity (radiance) I at a given wavelength and viewing angle for an incident solar flux density (irradiance) defined as "πF" in this case (see Sec. 2.4.1.1 for the definition of "flux density"). A perfectly reflecting Lambertian surface viewed at normal incidence has *I*/*F* = 1. Hapke (2012) reviews *I*/*F* and similar quantities.

2.2.4 Planetary Equilibrium Temperature

The Bond albedo enables us to calculate the effective temperature (i.e., blackbody equivalent temperature) that a planet attains in equilibrium with sunlight. If the Bond

albedo is A_b, then using eq. (2.10), the time-averaged solar flux absorbed by a planet is $(1 - A_b)S_p/4$, where S_p is the solar flux at the orbital position of the planet. The Stefan–Boltzmann Law (derived later in Sec. 2.4.1.5) gives the emitted flux of radiation from a blackbody at absolute temperature T as σT^4, where $\sigma = 5.6697 \times 10^{-8}$ W m^{-2} K^{-4} is the Stefan–Boltzmann constant. A planet is a good absorber of infrared radiation and therefore a good emitter, so a planet can be approximated by a blackbody at infrared wavelengths. If we assume that a planet's internal heat can be neglected, an equilibrium temperature, T_{eq}, is achieved from a balance of incoming and outgoing radiation, as follows:

$$\text{absorbed solar radiation flux} = \text{outgoing infrared radiation flux}$$

$$(1 - A_b)\frac{S_p}{4} = \sigma {T_{eq}}^4 \tag{2.15}$$

We can apply eq. (2.15) to the inner planets. Earth's Bond albedo is 0.3, so the left-hand side of eq. (2.15) is ~238 W m^{-2} = (1 – 0.3)×(1361 W m^{-2} /4). At equilibrium, this energy flux must be emitted back to space by radiation from the Earth. Solving eq. (2.15) gives an equilibrium temperature, T_{eq} = 255 K. This temperature is an effective blackbody temperature, defined in Box 2.1, and is clearly not equal to Earth's actual surface temperature. Table 2.2 shows actual mean surface emission temperatures of Earth, Venus, and Mars compared to equilibrium temperatures calculated from eq. (2.15). The difference between the surface temperature and equilibrium temperature in Table 2.2 arises because of the *greenhouse effect* caused by the presence of an atmosphere around each planet.

2.2.5 The Greenhouse Effect

On planets with atmospheres, the atmosphere warms up and provides a source of thermal infrared radiation to the surface in addition to the shortwave solar flux. Thus, the global average surface temperature of the Earth, T_s, is about 288 K rather than the blackbody emission temperature of 255 K (Table 2.2). The *greenhouse effect* can be explained as follows. The atmosphere warms up because it absorbs infrared radiation from the Earth below. Because the atmosphere is warm it radiates, and some of this radiation travels back down to the Earth. Consequently, the surface of the Earth is hotter than it would be in the absence of an atmosphere because it receives energy from two sources: sunlight and the heated atmosphere.

On a planet with a dense atmosphere, the difference between T_{eq} and T_s is a convenient way to express the magnitude of the greenhouse effect

$$\Delta T_g \equiv T_s - T_{eq} \tag{2.16}$$

For the Earth, ΔT_g = 288 – 255 = 33 K.

We now qualify the above considerations. On an airless body, such as the Moon, there is no greenhouse effect, so ΔT_g = 0 K. The Moon's surface receives solar radiation only, so we might expect its mean surface temperature to be the equilibrium temperature, according to eq. (2.15). However, the arithmetic mean global temperature is lower. A slowly rotating airless body does not have an atmosphere or ocean to transport heat efficiently. Because surface conduction is inefficient, the lateral transfer of heat is minimal and each local surface rapidly reaches thermal equilibrium with incoming solar radiation in relative isolation. Although global energy balance must still apply, a global average temperature is not a good measure of radiation balance because local emission has a T^4 scaling by the Stefan–Boltzmann Law, so the actual temperature representative of mean surface emission is biased towards warmer areas. For airless bodies, one must

Table 2.2 A comparison of equilibrium temperatures calculated using eq. (2.15) and the mean global surface temperatures of inner planets. (1 AU = 1.496×10^{11} m.)

Planet	Bond albedo (dimensionless)	Semi-major axis (AU)	Equilibrium temperature (K)	Mean global surface temperature of emission (K)	Greenhouse effect (K)
Venus	0.76[a]	0.7233	229	737[d]	508
Earth	0.3[b]	1.00	255	288	33
Mars	0.25[c]	1.5236	210	215[e]	5

[a] Moroz *et al.* (1985). [b] Palle *et al.* (2003). [c] Pleskot and Miner (1981). [d] At the mean radius of Venus (Fegley, 2014). [e] For Mars, this value is computed from the global and annual mean value of T_s^4, where T_s is the surface temperature; the temperature 215 K is fourth root of this value via eq. (2.17) (Haberle, 2013). See Sec. 2.2.5 for why a simple arithmetic global mean is less than this value for Mars.

use a time-dependent, spatially resolved model (Pierrehumbert, 2010 p. 152; Vasavada *et al.*, 2012; Vasavada *et al.*, 1999) and the average emission temperature must be calculated with a T^4 weighting.

Although Mars is not an airless body, its atmosphere is so thin that heat transfer is quite inefficient too. The arithmetic global mean surface temperature of Mars is estimated to be 202 K (Haberle, 2013), whereas the equilibrium temperature is 210 K (Table 2.2). This apparent paradox is resolved by realizing that one needs to compute the fourth power of the surface temperature T_s^4 at all locations on the planet and average this quantity globally and annually to obtain $\overline{T_s^4}$. Then, the global mean effective *surface emission temperature* T_{se} and corresponding greenhouse effect are defined as follows:

$$T_{se} = \left[\overline{T_s^4}\right]^{1/4}, \qquad \Delta T_g \equiv T_{se} - T_{\rm eq} \tag{2.17}$$

For Mars, T_{se} is 215 K so that the greenhouse effect is ~5 K by eq. (2.17).

For the Earth two factors allow the global mean surface temperature to be an accurate reflection of the global mean emission temperature. First, there is efficient lateral heat transfer by the atmosphere and oceans. Second, a feedback effect of water vapor in the atmosphere causes the emission of radiation from the surface to have a net linear scaling with surface temperature, which we explain in Sec. 2.3.1.

For Earth, there is debate about whether a greenhouse effect should be expressed in terms of radiative fluxes rather than a global mean temperature enhancement, ΔT_g in eq. (2.16) (Schmidt *et al.*, 2010; Zeng, 2010). In our view, using ΔT_g is no problem if we remember that it is directly related to fluxes of IR radiation. When we say 33 K warming compared to an Earth without an atmosphere, we mean, "all other things being equal in a hypothetical scenario," i.e., if you imagined the same albedo and a mean isothermal surface. Of course, an imaginary Earth with no atmosphere and ocean would not have the same albedo and, even if it did, hypothetically, the mean surface temperature would differ from 255 K because of the nonlinear T^4 scaling of local thermal emission on an airless world. We use the convention of ΔT_g because it is rooted in two observables: it indirectly quantifies the roughly 150 W m^{-2} difference in flux between ~390 W m^{-2} from a measurable 288 K mean global surface (emission) temperature and ~240 W m^{-2} at the top of the atmosphere from a globally equivalent blackbody temperature of 255 K that can be measured by satellites.

In Earth's modern atmosphere, the two most important greenhouse gases are H_2O and CO_2. In clear skies, H_2O is responsible for ~2/3 of the 33 K warming, although it acts in a different manner than CO_2 because it is near its condensation temperature in the troposphere. Effectively, H_2O is a slave to CO_2 and other greenhouse gases: if CO_2 levels increase and warm the Earth, then warmer surface waters produce a larger vapor pressure of H_2O, which amplifies the greenhouse effect (Lacis *et al.*, 2013; Lacis *et al.*, 2010). In Earth's current climate, doubling CO_2 itself leads to ~1.2 K mean global temperature rise. However, various feedbacks in the climate system can cause the actual temperature to eventually increase 2.0–4.5 K. Based on observations, the positive feedback due to water vapor is ~2 W m^{-2} K^{-1} and 0.3 W m^{-2} K^{-1} in the troposphere (Dessler *et al.*, 2008) and stratosphere, respectively (Dessler *et al.*, 2013). More will be said about water vapor feedback in Sec 2.3.

After water vapor, CO_2 accounts for most of the remaining 1/3 of the greenhouse effect in a clear sky atmosphere (Kiehl and Trenberth, 1997; Stephens and Tjemkes, 1993). Lesser contributions, around 2–3 K total, come from CH_4, O_3, N_2O, and various anthropogenic chlorofluorocarbons (CFCs). These trace constituents are powerful greenhouse gases on a molecule per molecule basis in Earth's present atmosphere. For example, in the current atmosphere, each CH_4 molecule increases radiative forcing 34 times more than a CO_2 molecule over a century (Stocker *et al.*, 2013). This is partly because CH_4, N_2O, O_3, and CFCs absorb within the ~8–12 μm wavelength region, an otherwise mostly transparent *window* through which most of the Earth's thermal-IR radiation energy fluxes to space. But it is also partly because these last four trace gases are all much less abundant in Earth's atmosphere than is CO_2. When one compares equal amounts of CH_4 and CO_2, CO_2 is by far the better greenhouse gas because the shorter wavelength wing of its strong absorption band at 15 μm is closer to the Wien peak of the outgoing thermal-IR radiation (see Sect. 2.4.1.5) than is the 7.6-μm CH_4 band (Kiehl and Dickinson, 1987).

Equations (2.15) and (2.16) show that the mean surface temperature, T_s, of a rocky planet with an atmosphere depends on three factors:

(a) solar flux, which is set by astronomical geometry and solar physics,

(b) Bond albedo A_b, which is set by clouds, aerosols, and the nature of the surface, and

(c) the greenhouse effect, which is set by the composition of the atmosphere.

For the evolution of climate, any changes in T_s (e.g., low-latitude glaciation) can only be understood by appealing to changes in one or more of these factors.

The most problematic factor is the way clouds affect both planetary albedo A_b and the greenhouse effect. Clouds cause most of the albedo on some planets, such as Earth,[2] and all of the albedo on planets such as Venus. (However, if you could magically take the clouds away from Venus, the planet would still have a high albedo of ~0.6 because dense CO_2 is good at scattering shortwave photons). On Earth, clouds contribute about 0.15 of the 0.3 Bond albedo, which cools the Earth. However, icy cirrus clouds also contribute to greenhouse warming because they are largely transparent to visible radiation but opaque to thermal IR. When clouds are taken into account, it is found that the global contributions to the greenhouse effect are ~50% from water vapor, ~25% from clouds, ~20% from CO_2, and the rest from other gases (Schmidt *et al.*, 2010). Here the percentages refer to fractions of the total longwave (LW) greenhouse forcing of 155 W m^{-2}: the global annual average upwelling LW flux ($\sigma T^4 = 390$ W m^{-2}, with $T = 288$ K) minus the top-of-atmosphere mean longwave flux, currently estimated as 235 W m^{-2} (Kiehl and Trenberth, 1997). However, when we look at the climate as a whole with satellite observations, while clouds enhance the greenhouse effect by ~30 W m^{-2}, they reduce Earth's absorbed shortwave radiation by –47 W m^{-2} (Loeb *et al.*, 2009). The net effect is that clouds *cool* the Earth by –17 W m^{-2}, which is large (Loeb *et al.*, 2009; Ramanathan and Inamdar, 2006).

Clouds can be observed and parameterized in Earth's present atmosphere, but their exact properties are difficult to predict for atmospheres other than our own (Marley *et al.*, 2013). For this reason, climate calculations for early Earth or for other Earth-like planets are subject to considerable uncertainty.

Clouds can also be influenced by biology. Clouds consist of water vapor that condenses on cloud condensation nuclei (CCN), which are submicroscopic particles. In Earth's atmosphere, a large contribution to CCN comes from sulfate salts, which can be affected by biology in the following ways. Some marine algae creates a gas, dimethlyl sulfide (CH_3SCH_3), or "DMS," that is subsequently oxidized and hydrated to form CCN (Charlson *et al.*, 1987). In the troposphere, where there is biological ammonia, most of the sulfate aerosols are mixtures of ammonium sulfate ($(NH_4)_2SO_4$) and bisulfate ($(NH_4)HSO_4$). For example, if sulfuric acid is derived from the oxidation of a sulfur-bearing gas, ammonia can neutralize the sulfuric acid derived as follows,

$$2NH_3 + H_2SO_4 \rightarrow (NH_4)_2SO_4 \qquad (2.18)$$

There is a correlation between marine stratus clouds and phytoplankton chlorophyll (Falkowski *et al.*, 1992), but whether this system has positive, negative or zero feedback on climatic temperature is currently unknown. However, other feedbacks, which we discuss below in Sec. 2.3, are more constrained.

In the stratosphere, sulfate aerosols are sourced from volcanic sulfur (SO_2 or H_2S that is subsequently oxidized) as well as biological gases, mainly carbonyl sulfide (OCS). Oxidation and hydration causes sulfuric acid (H_2SO_4) aerosols to form in the stratosphere, where their scattering of solar radiation contributes to planetary albedo.

2.2.6 Giant Planets, Internal Heat, and Equilibrium Temperature

As we discussed in Sec. 2.2.4, planets absorb radiation from the Sun and lose energy by radiating to space in equal amounts at equilibrium, but this idea assumes that we can ignore internal heat sources. For terrestrial planets, the internal heat flux is small. For example, the Earth's average heat flow is only ~90 mW m^{-2} (Schubert *et al.*, 2001) compared to the globally averaged net solar flux of 238 W m^{-2}. But in the Solar System Jupiter, Saturn and Neptune radiate significantly more energy than they absorb. For example, Jupiter radiates ~13.6 W m^{-2} while it absorbs only ~8.2 W m^{-2} from the Sun. An internal energy flux of 5.44 ± 0.43 W m^{-2} makes up the difference (Hanel *et al.*, 1981; Pearl and Conrath, 1991).

Internal heat on giant planets comes from accretion and differentiation. Gravitational energy is released in these processes. Saturn has higher internal heat than expected from its size and is believed to have been in state of internal differentiation for billions of years, in which immiscible helium "rains out" of a deep interior layer of metallic hydrogen towards Saturn's core, with release of gravitational energy (Stevenson and Salpeter, 1977b). Support comes from a relatively low He/H ratio of 0.6 in Saturn's atmosphere relative to solar compared to 0.81 on Jupiter. Because Saturn is smaller and cooler in its interior than Jupiter, the phase separation of helium has been going for far longer than inside Jupiter, where it began more recently. The fact that Uranus has little internal heat flow, 42 ± 47 mW m^{-2} (Pearl *et al.*, 1990), unlike the other giant planets is puzzling and may be due to different internal structure.

[2] Terrestrial clouds cover $68 \pm 0.3\%$ of the surface for >0.1 in optical depth clouds or $\sim 56 \pm 0.3\%$ for >2 in optical depth. Also, there is 10%–15% more cloud cover over oceans than land (Stubenrauch *et al.*, 2013).

For giant planets with significant internal heat fluxes, the *effective temperature* T_{eff} (of an equivalent blackbody) is greater than the *equilibrium temperature*, T_{eq} where outgoing IR balances incoming sunlight. Consequently, there is an internal heat flux, F_i, given by

$$F_i = \sigma\left({T_{eff}}^4 - {T_{eq}}^4\right) \tag{2.19}$$

The total emitted flux is modified to include this internal heat,

outgoing radiation flux = absorbed solar flux + internal heat flux

$$\sigma T_{eff}^4 = (1 - A_b)\frac{S_p}{4} + F_i \tag{2.20}$$

2.3 Climate Feedbacks in the "Earth System"

Feedbacks in the Earth system can amplify the mean temperature of the planet in positive feedback or stabilize it in negative feedback. There are many potential feedbacks for climate, of which four important ones are:

(1) positive feedback from atmospheric water vapor;
(2) positive feedback from ice–albedo;
(3) negative feedback from outgoing radiation on short timescales; and
(4) negative feedback from the carbonate–silicate cycle on geological timescales.

Water vapor acts as a positive feedback on the climate system because it is near its condensation temperature. As the climate cools, the saturation vapor pressure drops. If the relative humidity (the ratio of vapor pressure to saturation vapor pressure) remains constant, then the water vapor concentration in the atmosphere will decrease proportionately. Less water vapor results in a smaller greenhouse effect, which in turn results in further cooling. Just the opposite happens if the climate warms: atmospheric H_2O increases, thereby increasing the greenhouse effect and amplifying the initial warming. This particular feedback loop is very important for Earth's climate. It essentially doubles the effect of climatic perturbations, such as changes in solar flux or in atmospheric CO_2 (Stevens and Bony, 2013).

The second feedback, the *ice–albedo feedback*, describes the interaction between surface temperature and the fraction of Earth's surface covered by snow and ice. An increase in surface temperature causes a decrease in snow and ice cover, which decreases the albedo and so causes further temperature increase in *positive* feedback. Conversely, a decrease in surface temperature increases the albedo due to snow and ice cover, which amplifies the decline in surface temperature. The snow-and-ice albedo feedback loop played a major role in the advances and retreats of the *Pleistocene* (~2.6 Ma–10 ka) ice sheets (Lorius *et al.*, 1990) and is an important positive feedback in anthropogenic climate change.

The climate system must also contain negative feedbacks, or else it would be unstable. The most basic negative feedback is the interaction between surface temperature T_s, and the outgoing infrared flux, F_{IR}. As T_s increases, F_{IR} increases. Earth cools itself by emitting infrared radiation; thus, as F_{IR} increases, T_s decreases. This creates a negative feedback loop. This feedback loop is the reason that Earth's climate is stable on short timescales. On long timescales ~1 m.y., we will see in Ch. 11 that an important climate feedback involves the interaction between atmospheric CO_2 and surface temperature, called the carbonate–silicate cycle (Berner, 2004; Walker *et al.*, 1981). In this feedback, if the climate warms, the weathering rate on land increases and removes carbon dioxide, which acts in negative feedback on surface temperature.

2.3.1 Climate Sensitivity

An illustration of the way the Earth's climate works in response to feedback comes from considering a simple energy balance, following Budyko (1969). Imagine the Earth as a blackbody with a surface temperature in Kelvin of $T = T_0 + \Delta T$. If T_0 is a reference temperature, such as water's freezing point (273.15 K), then we can write $\Delta T = T_s$, the surface temperature in °C. Thus, the emitted thermal-IR flux by the Stefan–Boltzmann Law (Sec. 2.4.1.5) is

$$F_{IR} = \sigma T^4 = (T_0 + T_s)^4$$

We can expand the right-hand side of this equation as follows:

$$\sigma(T_0 + T_s)^4 = \sigma(T_0 + T_s)^2(T_0 + T_s)^2 = \sigma\left({T_0}^2 + 2T_sT_0 + {T_s}^2\right)\left({T_0}^2 + 2T_sT_0 + {T_s}^2\right)$$
$$= \left(\sigma {T_0}^4\right) + \left(4\sigma {T_0}^3 T_s\right) + \cdots$$

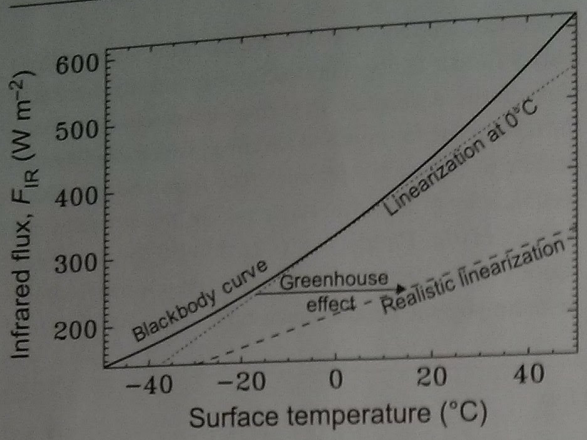


Figure 2.8 Outgoing infrared radiation at the top of the terrestrial atmosphere F_{IR}, as a function of global mean surface temperature. The solid line is the blackbody case. A linearization to this curve about 0 °C gives a surface temperature of –16.6 °C for an outgoing flux of 239 W m^{-2}. A more realistic linear model has a mean surface temperature of 15 °C and a significantly different slope or climate sensitivity. The arrow at F_{IR} = 239 W m^{-2} shows the magnitude of the greenhouse effect, which causes the temperature offset between the two linearizations.

By taking only the first two terms, we can linearize the emitted IR flux as

$$F_{IR} = \underbrace{a}_{\sigma T_0^4} + \underbrace{bT_s}_{4\sigma T_0^3 T_s}; \qquad T_s \text{ in } ^\circ\text{C} \tag{2.21}$$

If the Earth behaved truly as an airless blackbody, then this linearization would give $a = \sigma(273.15)^4 = 315.659$ W m^{-2} and $b = 4\sigma(273.15)^3 = 4.623$ W m^{-2} °C^{-1}. Given an incoming or outgoing radiation of ~239 W m^{-2}, the value of T_s would be –16.6 °C, close to 255 K for the blackbody curve.

For Earth's actual climate, the values of a and b are empirically closer to $a = 206$ W m^{-2} and $b = 2.2$ W m^{-2} °C^{-1}. Inserting these numbers in eq. (2.21), we derive:

$$239 \text{ W m}^{-2} = 206 \text{ W m}^{-2} + (2.2)(T_s) \tag{2.22}$$

From eq. (2.22), we calculate T_s = 15 °C. Equation (2.21) is plotted in Fig. 2.8. Because the value of b is smaller in the real climate than for a blackbody, the surface temperature, T_s, is more sensitive to changes in F_{IR} (or the solar constant S_0, since F_{IR} is proportional to S_0). Thus, the surface temperature increases more for a given change in F_{IR} than it would in the blackbody case. The main reason for the higher sensitivity is water vapor feedback. If the climate warms, the water vapor increases and prevents IR radiation from escaping.

Climate sensitivity λ, is defined as the ratio of the change in global mean surface temperature ΔT_s to the change in climate forcing ΔQ (where Q has units of W m^{-2}).

$$\Delta T_s = \lambda \Delta Q \tag{2.23}$$

The change in forcing could be a change in the solar constant or greenhouse effect. In the case of the latter, an accurate value of λ is important for predicting the anthropogenic influence on climate change.

We can evaluate the climate sensitivity to changes in the solar constant for a blackbody approximation and compare our simple but more realistic linearized model of Earth's climate system represented by eq. (2.21). For the blackbody case, let us ignore any temperature dependence of albedo and assume that the effective temperature and surface temperature are linearly related. Radiation balance gives us

$$F_{IR} = \sigma T^4 = \frac{(1 - A_b)}{4} S_\odot$$

Then

$$\frac{d(F_{IR})}{dS_\odot} = \frac{(1 - A_b)}{4} = \left(\frac{d(F_{IR})}{dT}\right)\left(\frac{dT}{dS_\odot}\right) \tag{2.24}$$

Hence,

$$\lambda = \frac{dT}{dS_\odot} = \frac{(1 - A_b)}{4}\left(\frac{1}{d(F_{IR})/dT}\right) = \frac{(1 - A_b)}{4(4\sigma T^3)} = \frac{T}{4}\left(\frac{(1 - A_b)}{4\sigma T^4}\right) = \frac{T}{4S_\odot} \tag{2.25}$$

Consequently, λ = (255 K)/(4 × 1361 W m^{-2}) = 0.047 K (W m^{-2})$^{-1}$. This means that a 1% change in the solar constant $S_\odot$ of 0.01 × 1361 = 13.61 W m^{-2} would produce a 13.61 × 0.047 = 0.64 K change in surface temperature. However, this Stefan–Boltzmann feedback is too small and unrealistic. If we consider our linearized model, we have

$$F_{IR} = a + bT = \frac{(1 - A_b)}{4} S_\odot \tag{2.26}$$

If we now apply eq. (2.24), we obtain

$$\lambda = \frac{dT}{dS_\odot} = \left(\frac{d(F_{IR})/dS_\odot}{d(F_{IR})/dT}\right) = \frac{[(1 - A_b)/4]}{b} \tag{2.27}$$

Consequently, in this case the climate sensitivity is almost double that before, λ = (1– 0.3)/(4 × 2.2 W m^{-2} K^{-1}) = 0.080 K (W m^{-2})$^{-1}$. In this case, a 1% change in the solar

constant would produce a change in surface temperature of about 1.1 K. The almost factor of 2 difference between this and the Stefan–Boltzmann case comes about because we have included an empirical water vapor positive feedback, which doubles climate sensitivity.

We can also estimate how much temperature is raised by an increase in the greenhouse effect. Radiative model calculations suggest that a doubling of CO_2 from 300 ppmv to 600 ppmv would cause an extra 4 W m^{-2} in the greenhouse effect and warm the surface. From eq. (2.26), $dT/d(F_{IR}) = 1/b = 0.45$ K (W m^{-2})$^{-1}$, so that the surface temperature would be expected to change by 4 W m^{-2} × 0.45 K (W m^{-2})$^{-1}$ ~2 K, assuming no other positive or negative feedbacks. This simple calculation produces an answer that is near the lower end of the range predicted by sophisticated, three-dimensional simulations of global warming. The IPCC (Intergovernmental Panel on Climate Change) suggests a temperature increase of (1.5–4.5) K for doubled CO_2, based on calculations by a suite of different climate models (Stocker *et al.*, 2013). Additional positive feedbacks from decreased sea ice and changes in cloud cover generally suggest ~3 K warming for doubled CO_2 (Hansen *et al.*, 2008).

2.3.2 The Emission Level and Radiative Time Constants

For planetary thermal equilibrium, it is useful to think of a level at some height, $z_{emission}$, in the atmosphere where the temperature T equals the effective temperature T_{eff}. We call this level the *emission level*, i.e.,

$$T_{surface} = T_{eff} + \Gamma z_{emission} \tag{2.28}$$

where Γ is the lapse rate. Given that the Earth's troposphere has a typical mean lapse rate $\Gamma = 6$ K km^{-1}, $T_{surface} = 288$ K, and $T_{eff} = 255$ K, the emission level is at a height of about 5 km on Earth.

We can think of energy absorbed from the Sun as being balanced by emission at the emission level at temperature T_{eff}. The column mass at this level, M_{ce}, is given by eq. (1.19) as $M_{ce} = p_{emission}/g$, where $p_{emission}$ is the pressure at the emission level. If a temperature change $\Delta T_{emission}$ is forced by the absorption of solar radiation, the heat change due to heat capacity of the column of atmosphere is given by

$$\text{heat change per m}^2 = M_{ce}c_p(\Delta T) = \frac{p_{emission}c_p(\Delta T_{emission})}{g}$$

where c_p is the specific heat capacity and g is gravitational acceleration. For sunlight,

$$\text{power absorbed per m}^2 = \frac{(1 - A_b)S_p}{4}$$

Consequently, we can define a time constant as follows:

$$\tau_e = \left(\frac{p_{emission}c_p(\Delta T_{emission})}{g}\right) \Big/ \left(\frac{(1 - A_b)S_p}{4}\right) = \frac{4p_{emission}c_p(\Delta T_{emission})}{(1 - A_b)S_p g} \tag{2.29}$$

If τ_{day} = 1 day, we obtain a typical diurnal temperature variation at the emission level associated with average solar input of

$$\Delta T_{emission} = \frac{(1 - A_b)S_p g \tau_{day}}{4p_{emission}c_p} \tag{2.30}$$

Typical values of $\Delta T_{emission}$ are 2 K (Venus), 2 K (Earth), 80 K (Mars), and 0.001 K (Jupiter). Diurnal radiative effects are clearly important on Mars. In contrast, terrestrial radiative processes act slowly and can be largely ignored in short-term atmospheric developments. In the long-term, however, they dominate sources and sinks of energy. Also from eq. (2.30), we can find the pressure level where $\Delta T_e = T_{eff}$ as follows: 700 mbar (Venus), 8 mbar (Earth), 2 mbar (Mars), and 0.05 mbar (Jupiter).

The radiative time constant is sometimes defined differently from above. For example, if the pressure in eq. (2.29) is taken as surface pressure, and we consider heating a full column of air to temperature T_{eff}, then using eq. (2.15) to substitute for $\sigma T_{eff}^4 = S(1 - A_b)/4$, we get

$$\tau_{rad1} = \frac{M_c c_p}{\sigma T_{eff}^3} = \frac{p_s c_p}{g\sigma T_{eff}^3} \tag{2.31}$$

where M_c is the total columnar mass. Equation (2.31) applies to planets with surfaces.

2.4 Principles of Radiation in Planetary Atmospheres

The way that solar radiation interacts with atmospheres and surfaces of planets is essential for understanding planetary habitability, given that liquid water is needed for Earth-like life. The solar energy absorbed is balanced by the transmission of thermal infrared radiation through a planet's atmosphere and back to space. The extent to which thermal infrared radiation is absorbed depends upon the greenhouse gases in the atmosphere, and these

gases link the evolution of atmospheric composition to climate. In order to appreciate how climate depends upon atmospheric composition, we need to understand *radiative transfer*, the physics of how electromagnetic radiation passes through media. For our purposes, the medium is an atmosphere of matter that scatters, absorbs, and emits radiation. However, before getting into details, we start with some basic definitions.

2.4.1 Basic Definitions and Functions in Radiative Transfer

2.4.1.1 Solid Angles and Spherical Polar Coordinates

The study of radiation often involves considering the amount of energy in a solid angle. A *solid angle*, expressed in units of steradians (sr), is defined as:

$$\text{solid angle}, \Omega = \frac{\text{area } \sigma_A \text{ of spherical surface}}{r^2} \quad [\text{sr}] \tag{2.32}$$

Thus, for a complete sphere, $\Omega = 4\pi r^2/r^2 = 4\pi$ steradians. *Spherical polar coordinates* are also often used in radiative transfer where a *zenith angle* θ, and an *azimuthal angle* ϕ, are defined according to Fig. 2.9(a). This figure also shows an elemental area, $d\sigma_A = (rd\theta)(r\sin\theta d\phi)$, from which we define a differential solid angle, $d\Omega$ as

$$d\Omega = \sin\theta d\theta d\phi \tag{2.33}$$

2.4.1.2 Radiometric Quantities

In radiative transfer, it is important to understand the distinction between radiance, irradiance, and net flux. Briefly, these are defined as follows.

(a) *Radiance = intensity*, I, is the power per unit frequency interval crossing unit area in unit solid angle in a given direction at some point in space. It has units of W m^{-2} Hz^{-1} sr^{-1} (or it can be per unit wavelength, e.g., W m^{-2} nm^{-1} sr^{-1}). Blackbody radiation is usually expressed as intensity denoted by B.

(b) *Flux density = irradiance*, F, is radiance integrated over a solid angle, usually a hemisphere, i.e., total power per unit frequency crossing unit area (a scalar quantity). In a spectral region, it is specified in units of W m^{-2} Hz^{-1} (or per unit wavelength, e.g. W m^{-2} nm^{-1}). Strictly speaking, "flux" only refers to the power in units of Watts, but that definition is almost never used. In most literature, "flux" is identical to "flux density." If we integrate over the whole spectrum, the result is a total flux density, expressed in W m^{-2}. A mnemonic for the difference between *radiance* and *irradiance* is that the "i" in the latter helps you to remember that irradiation involves integration of radiance over a solid angle.

(c) *Net flux* ($\underline{F}$) takes into account energy transfer in all directions, and $\underline{F}$ refers to a unit area in a specific direction, i.e., $\underline{F}$ is a vector. $F^+ - F^-$ is the component of $\underline{F}$ in that direction. Net flux is specified in W m^{-2} for a given frequency interval.

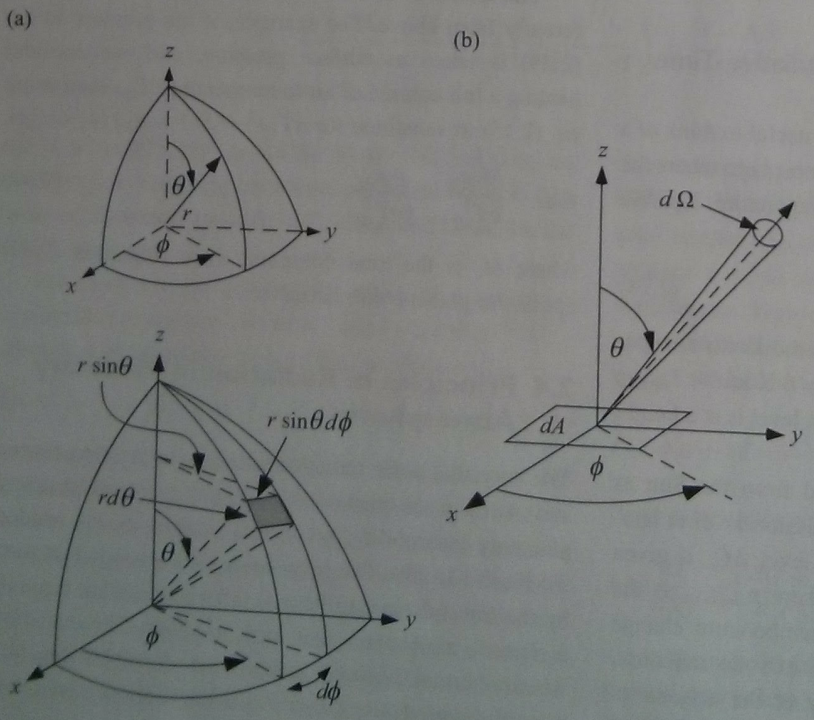


Figure 2.9 (a) Spherical coordinates: zenith angle θ, azimuthal angle ϕ, and radius r. (b) Illustration of the angles that define the radiance fluxing through an element of area dA through elemental solid angle $d\Omega$ in a direction defined by zenith angle θ and azimuthal angle ϕ.

2.4.1.3 *Heating Rate*

The heating rate in a planetary atmosphere is the net power dissipated (or removed from) the medium per unit volume. It is often expressed in units of K day^{-1}. The heating rate depends on the divergence of $\underline{F}$, which is the gradient of $\underline{F}$ as reflects inflow and outflow. Thus the heating rate due to the difference in upward and downward fluxes in the vertical (z) direction is given by

$$\frac{dT}{dt}\rho_{\text{air}}c_p = -\text{div}\underline{F}, \qquad \text{which, in the } z \text{ direction, } = -\frac{d}{dz}(F^+ - F^-) \tag{2.34}$$

where F^+ is in the same direction as z. (For understanding divergence, recall that for a vector $\underline{v} = v_1\mathbf{i} + v_2\underline{\mathbf{j}} + v_3\mathbf{k}$, we define $\text{div}\,\underline{v} = \partial v_1/\partial x + \partial v_2/\partial y + \partial v_3/\partial z = \nabla \cdot \underline{v}$, and that divergence measures outflow minus inflow). Alternatively, sometimes arrows are used to indicate the direction of fluxes, in which case we can rewrite eq. (2.34) as

$$\frac{dT}{dt}\rho_{\text{air}}c_p = -\left(\frac{dF\uparrow}{dz} - \frac{dF\downarrow}{dz}\right) = \frac{d}{dz}(F\downarrow - F\uparrow) \tag{2.35}$$

where $F\uparrow$ is the upward flux.

2.4.1.4 *Hemisphere Radiation and the Relationship Between Flux and Intensity*

The geometry for a pencil beam of radiation passing through an elemental area, dA, is shown in Fig. 2.9(b). The radiant energy dE_ν in time interval dt and frequency range ν to $\nu + d\nu$ crossing elemental area dA is given by

$$dE_\nu = I_\nu \cos\theta\, dA\, d\Omega\, d\nu\, dt \tag{2.36}$$

where $dA\cos\theta$ is the effective area of intercept. From earlier, an element of solid angle is $d\Omega = \sin\theta\, d\theta\, d\phi$, which we can substitute. The *monochromatic flux density*, i.e., the flux density at a particular frequency, is defined by the normal of the radiance I_ν integrated over the entire upper hemisphere,

$$F_\nu^+ = \frac{E_\nu}{At} = \int_0^{\phi=2\pi}\int_0^{\theta=\pi/2} I_\nu \cos\theta \sin\theta\, d\theta\, d\phi = 2\pi\int_0^{x=1} I_\nu x dx \tag{2.37}$$

In the last step, we have substituted $x = \sin\theta$ and $dx = \cos\theta\, d\theta$ and assumed I_ν has no azimuthal dependence. Note that we can also write flux density for the hemisphere on the opposite side as:

$$F_\nu^- = \int_0^{\phi=2\pi} d\phi \int_\pi^{\theta=\pi/2} I_\nu \cos\theta \sin\theta\, d\theta \tag{2.38}$$

Considering the upper hemisphere, if the irradiance is isotropic (i.e., $I_\nu(\theta) = I_\nu$), we can integrate the expression in eq. (2.37) as follows:

$$F_\nu^+ = 2\pi I_\nu \left[\frac{x^2}{2}\right]_0^1 = \pi I_\nu \quad \left[\text{unit : W m}^{-2}\text{Hz}^{-1}\right] \tag{2.39}$$

Note that integration over a hemisphere is why we frequently see a "pi" written in front of radiances. Note also that the net flux transported through an area will be $F_\nu = F_\nu^+ - F_\nu^-$.

Incidentally, the factor of pi is a perennial source of confusion for generations of students of radiative transfer because of a peculiar convention in a foundational textbook by Chandrasekhar. He chose to absorb the factor of pi into his definition of net flux as "$\pi F_{A\nu}$" instead of using F_ν (Chandrasekhar, 1960, p. 3). Here, we use an "A'" subscript in $F_{A\nu}$ to denote Chandrasekhar's "astrophysical radiative flux". In Chandrasekhar's notation, eq. (2.37) would be written $F_{A\nu}^+ = 2\int_0^{x=1} I_\nu x dx$, for comparison.

The *total flux density* or *irradiance* (which has units of W m^{-2}) is obtained by integrating over all frequencies ν, i.e.,

$$F = \int_0^\infty F_\nu d\upsilon \tag{2.40}$$

2.4.1.5 *Blackbodies: the Planck Function, the Stefan–Boltzmann Law, and Wien's Law*

The Planck function is an intensity that derives from Planck's Law for the energy of a photon,

$$E = h\nu \tag{2.41}$$

where $h = 6.626 \times 10^{-34}$ J s. It describes the blackbody radiation spectrum (Box 1.2), as a function of temperature. There are three forms of the Planck function that are commonly encountered. First, when the spectrum is in frequency units, we have

$$B_\upsilon d\upsilon = \frac{2h\upsilon^3}{c^2}\frac{d\upsilon}{(e^{h\upsilon/kT} - 1)}, \quad \text{where } B_\upsilon \text{ is in W m}^{-2}\,\text{Hz}^{-1}\,\text{sr}^{-1} \tag{2.42}$$

Second, using $c = \upsilon\lambda$, $\Rightarrow$ $d\upsilon = -(c/\lambda^2)d\lambda$, we can rewrite eq. (2.42) in terms of a wavelength spectrum

$$B_\lambda d\lambda = \frac{2hc^2}{\lambda^5}\frac{d\lambda}{(e^{hc/\lambda kT}-1)}, \quad \text{where } B_\lambda \text{ is in W m}^{-2}\text{ m}^{-1}\text{ sr}^{-1} \tag{2.43}$$

Third, we can rewrite the function in terms of wavenumber units (usually cm^{-1}), which are used in spectroscopy. Since $\tilde{\upsilon} = 1/\lambda$, $d\tilde{\upsilon} = -(1/\lambda^2)d\lambda$, so

$$B_{\tilde{\upsilon}} d\tilde{\upsilon} = 2hc^2\frac{\tilde{\upsilon}^3 d\tilde{\upsilon}}{(e^{hc\tilde{\upsilon}}/kT-1)}, \quad \text{where } B_{\tilde{\upsilon}} \text{ is in W m}^{-2}\left(\text{cm}^{-1}\right)^{-1}\text{ sr}^{-1} \tag{2.44}$$

There are two key inferences from the Planck function: the Stefan–Boltzmann Law and Wien's Law.

The Stefan–Boltzmann Law results from eqs. (2.39) and (2.40) applied to the Planck function (2.42). We can use the substitution $x = h\upsilon/kT$ and the standard integral

$$\int_0^\infty \frac{x^3 dx}{e^x - 1} = \frac{\pi^4}{15}$$

Thus,

$$F = \int_0^\infty \pi B_\upsilon(T) d\upsilon = \frac{2\pi h}{c^2}\int_0^\infty \frac{\upsilon^3 d\upsilon}{(e^{h\upsilon/kT}-1)} = \frac{2\pi k^4 T^4}{c^2 h^3}\int_0^\infty \frac{x^3 dx}{(e^x-1)} = \left(\frac{2\pi^5 k^4}{15c^2h^3}\right)T^4 \tag{2.45}$$

The group of constants in front of T^4 on the right-hand side comprise the Stefan–Boltzmann constant, σ (= 5.67 $\times 10^{-8}$ W m^{-2} K^{-4}), as follows,

$$F = \sigma T^4 = \left(\frac{2\pi^5 k^4}{15c^2h^3}\right)T^4 \tag{2.46}$$

Wien's Law gives the spectral location of the maximum emission of the Planck function,

$$\frac{\partial B_\lambda}{\partial \lambda} = 0 \quad \Rightarrow \lambda_{\max} T = 2.897 \times 10^{-3}\text{ m K} \tag{2.47}$$

For the Sun, characterized by an effective temperature of 5780 K, Wien's Law indicates that the peak emission occurs at a wavelength of $\lambda_{\max}$ = (2.897 $\times 10^{-3}$ m K)/(5780 K) = 0.5 μm, in the visible. One nuance of Wien's Law is that when we work in wavenumbers, the maximum occurs at a wavenumber corresponding to a wavelength $\lambda_{\max}$ given by $\lambda_{\max} T = 5.100 \times 10^{-3}$ m K, which is not the same as in the case of the wavelength curve. Thus, strictly speaking, the position of the Planck maximum has no meaning unless we say which spectral units we are using.

2.4.1.6 Kirchoff's Law

Media often differ from perfect blackbodies. One defines an absorptivity, α_υ, which is the fraction of energy at frequency υ per unit frequency falling on a body that is absorbed. In the case of a blackbody, it perfectly absorbs and $\alpha_\upsilon = 1$ for all frequencies. Generally, $\alpha_\upsilon < 1$ for all frequencies and some of the radiation is reflected or transmitted rather than absorbed.[3] When we assume that α_υ is some fixed value less than 1 for all frequencies, we define what is known as a *gray body*. For example, a *gray atmosphere* is an approximation using a broadband value of α_υ less than 1 for radiative calculations.

In thermodynamic equilibrium, with a steady temperature of the medium, a body emits as much as it absorbs. We can introduce the emissivity, which is the ratio of the spectral radiance at a particular frequency or wavelength to that of a blackbody, i.e., ε_υ = (spectral radiance at frequency υ)/ $[B_\upsilon(T)]$. Kirchoff recognized a curious fact: in thermodynamic equilibrium, the absorptivity at a particular frequency depends on the temperature of the substance. Kirchoff's Law says that

$$(\text{emissivity})_\upsilon = (\text{absorptivity})_\upsilon, \quad \text{or } \varepsilon_\upsilon = \alpha_\upsilon \tag{2.48}$$

Thus, a fractional absorber at a particular frequency and given temperature is also a fractional emitter at the same frequency in thermodynamic equilibrium. Strictly, eq. (2.48) is valid only for a specific viewing direction, unless ε_υ (and so α_υ) is independent of direction.

2.4.2 Radiative Transfer in the Visible and Ultraviolet

We now have enough basic definitions to consider the passage of radiation through an atmosphere and we start with shortwave (ultraviolet (UV) and visible) because it has simpler physics than that in the thermal infrared (IR). In the IR, the complication is that a planetary atmosphere itself, by virtue of its finite temperature, generates radiation at IR wavelengths in addition to the incident

[3] Absorptivity (and hence emissivity) can exceed unity in the case of small particles that have a larger interaction cross-section with radiation than their geometric cross-section. The rule that absorptivity < 1 applies strictly only to extensive objects.

sunlight. In the visible and UV parts of the spectrum, we can neglect radiative emission at temperatures typical of planetary surfaces and atmospheres, which greatly simplifies calculations.

We can easily justify our neglect of shortwave emission by using eqs. (2.43) and (2.39). At the short wavelength end of the planetary infrared spectrum in the near-infrared at 1 μm, the monochromatic flux radiance falling on a horizontal plane just above the Earth's surface from radiation from the Earth is πB_λ (288 K, 1 μm) = 7.1×10^{-8} W m^{-2} m^{-1}. In contrast, the monochromatic flux radiance at 1 μm from the Sun is

$$(1 - A_b)\left(\frac{R_\odot}{r_{AU}}\right)^2 \pi B_\lambda(5780\ \text{K}, 1\ \mu\text{m}) \tag{2.49}$$

which computes to 2.76×10^8 W m^{-2} m^{-1}. Here, A_b is the Bond albedo of the Earth (≈ 0.3), $R_\odot$ is the radius of the Sun (6.96×10^8 m), and r_{AU} is the Earth–Sun distance of 1 AU or 1.49598×10^{11} m. Because the visible flux is ~10^{15} greater, we can neglect the planetary thermal radiation at this wavelength; solar radiation clearly dominates, as seen in Fig. 2.10.

Given that solar and thermal-IR radiation fields have little overlap, we often treat them independently. In this approximation, the atmosphere is assumed to be transparent to solar radiation but largely opaque to thermal radiation from the Earth. This was the approach we used in estimating a planetary equilibrium temperature earlier in Sec. 2.2.4. However, planetary atmospheres are not completely transparent to solar radiation: they absorb in the UV and near-IR parts of the solar spectrum. Since relatively little energy is contained in the UV and near-IR parts of the visible spectrum, the energy balance calculations are not affected much.

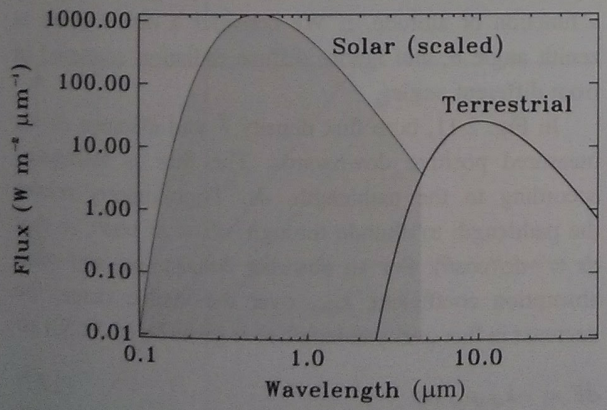


Figure 2.10 Blackbody curves of the Sun and Earth, where the solar flux density is that at 1 AU distance scaled by a factor of $1-A_b$, where A_b is Earth's Bond albedo. The curves cross near ~4 μm. Solar and thermal spectra peak at ~0.5 and 10 μm, respectively.

The structure and color of atmospheres are strongly affected by shortwave absorption and scattering by gases. In the terrestrial atmosphere, H_2O vapor absorbs near-IR solar radiation (1–4 μm region) and contributes 1 K per day heating to the mid-troposphere at ~5 km altitude. Stratospheric ozone absorbs UV (0.3–0.2 μm), and all gases absorb EUV (i.e., <100 nm) in the thermosphere. Ozone also weakly absorbs in the visible spectrum in the Chappuis bands continuum (~450–850 nm), with peak absorption in the yellow-orange at ~600 nm (which is why pure ozone gas is a pale blue color).

The blue color of Earth's sky comes from Rayleigh scattering of the solar spectrum by N_2 and O_2 molecules along with the response of the human eye to color (Bohren and Fraser, 1985; Hoeppe, 2007; Smith, 2005). At very low Sun angles during twilight and dusk, transmitted light from the Earth's horizon should be yellow, which is the color remaining after Rayleigh scattering out of the light; but absorption of yellow by the ozone layer ensures a blue hue. Edward Hulburt first realized that the long pathlength through ozone dominates over Raleigh scattering under these circumstances (Hulburt, 1953). Yellow-orange absorption in the transmission of starlight through exoplanet atmospheres is also a possible way to detect ozone (e.g., Yan *et al.*, 2015). Such a measurement is of interest because the ozone on Earth is derived from biogenic O_2.

The color of the Earth when viewed as a single pixel from Saturn's orbit or beyond has been described as a *Pale Blue Dot*, which arises from the scattering and absorption properties of the Earth (Sagan, 1994). Rayleigh scattering of sunlight back to space from the atmosphere primarily causes the blue color, not the ocean (Crow *et al.*, 2011; Krissansen-Totton *et al.*, 2016b). Clouds, which reflect fairly uniformly in the visible, cause the paleness. In detail, ozone absorption in the Chappuis bands sits at the flat bottom of a U-shape in the 400–1000 nm reflectance spectrum, reducing the reflectivity by 7%–8% (Robinson *et al.*, 2014a). An upward slope in reflectance occurs from 600 nm to 900 nm because of vegetated land surfaces (Arnold *et al.*, 2002; Fujii *et al.*, 2010; Seager *et al.*, 2005; Tinetti *et al.*, 2006).

In the giant planet atmospheres, parts of the solar visible and near-IR spectrum are absorbed. In particular, Jupiter, Saturn, Uranus, Neptune, and Titan all have strong CH_4 bands in the visible and near-IR (543 nm, 619 nm, 725 nm, 844 nm, 865 nm, 888 nm) (Karkoschka, 1994, 1998). The light we see coming from Neptune is sunlight scattered by its uppermost atmosphere. CH_4 absorbs in the orange-red part of the spectrum so that scattered sunlight is depleted in red light, and Neptune

appears bluish-green. The color of Uranus is attributable to the same process.

2.4.2.1 *The Beer–Lambert–Bouguer "Extinction Law" and the Meaning of Opacity*

The *Beer–Lambert–Bouguer Law* describes how a beam of light is exponentially attenuated with distance in an optically homogeneous medium. Pierre Bouguer (pronounced "boo-gair") first described the law in 1729, followed by Johann Lambert in 1760 and August Beer in 1852. Despite Bouguer being first, the law is commonly called *Beer's Law* or *Lambert's Law*. For simplicity, we call it *the Extinction Law*.

When a beam of light passes through a substance, it is found experimentally that the loss of intensity depends linearly on both the incident intensity I_ν, and the amount of optically active matter (i.e., absorbing or scattering material) along the beam direction, proportional to an elemental path, ds:

$$dI_\nu \propto I_\nu ds \qquad \Rightarrow dI_\nu = k_\nu' I_\nu ds \tag{2.50}$$

The constant of proportionality is the *extinction coefficient* k_ν'. In general, extinction of radiant energy is caused by absorption and scattering. Bouguer's insight was that the exact process doesn't matter for the formulation of the law. The extinction coefficient is the sum of an absorption coefficient and a scattering coefficient, and generally depends on temperature and pressure because these affect the absorption of gases.

The extinction coefficient can be defined in three ways:

(i) the **extinction coefficient** or **volume extinction coefficient**, in terms of pathlength:

$$k_\nu' = \frac{dI_\nu}{I_\nu ds} \qquad [\mathrm{m}^{-1}] \tag{2.51}$$

(ii) the **mass extinction coefficient**, in terms of mass path, $dM = \rho_a ds$, where ρ_a is the mass density of the absorber or scatterer [kg m^{-3}]:

$$k_\nu = \frac{dI_\nu}{I_\nu \rho_a ds} \qquad [\mathrm{m}^2\ \mathrm{kg}^{-1}] \tag{2.52}$$

(iii) the **extinction cross-section**, in terms of column number density, $dN_c = n_a ds$, where n_a is the number density [m^{-3}] of the optically active gas:

$$k_\nu'' = \frac{dI_\nu}{I_\nu n ds} \qquad [\mathrm{m}^2] \tag{2.53}$$

In fact, for particles, a dimensionless extinction efficiency can be defined in terms of k_ν'', as follows

$$Q_e = \frac{k_\nu''}{A_{\mathrm{xs}}} \tag{2.54}$$

where A_{xs} is the geometric cross-section, such as πr^2 for a sphere of radius r. For terrestrial cloud droplets, $Q_e \sim 2$ in the visible part of the spectrum, for example.

If we integrate eq. (2.50), we get the usual form of the Extinction Law:

$$I_\nu = I_{\nu 0} e^{-\tau_\nu} \Rightarrow \text{transmissivity } T_\nu = e^{-\tau_\nu} \tag{2.55}$$

The law indicates that intensity decays exponentially along the path. Here $I_{\nu 0}$ is the intensity at distance $s = 0$ and τ_ν is the (extinction) *optical path* or *optical thickness* or *opacity* along the beam direction, which is a dimensionless quantity defined as:

$$\tau_\nu = \int_0^s k_\nu' ds' = \int_0^s k_\nu \rho_a ds' = \int_0^s k_\nu'' n_a ds' \tag{2.56}$$

Sometimes, particularly in astronomy, the extinction coefficient k_ν is called "opacity," with potential for confusion. Optical thickness τ_ν is a measure of the strength and number of *optically active* particles (meaning those that scatter and/or absorb) along a beam. Equation (2.56) also shows that factors of ρ_a or n_a relate all three extinction coefficients.

2.4.2.2 *Direct Beam Solar Flux and the Meaning of "Optical Depth"*

The Extinction Law can be formulated for the case of solar irradiance, i.e., flux, noting that the above is for radiance. Consider a *plane-parallel* atmosphere, which is an approximation that ignores the spherical nature of a planet and assumes that atmospheric properties are only a function of altitude, z. We consider a direct beam at zenith angle θ, and ignore diffuse radiation scattered in from different angles.

In Fig. 2.11, both flux density F and distance ds are measured positive downwards. The flux is attenuated according to the pathlength, ds. Trigonometry relates the pathlength to altitude through $-dz = ds \cos\theta$, so that $ds = -dz/(\cos\theta)$. For an absorber density, ρ_a, and mass absorption coefficient, k_{abs}, over the visible range, the decrease in flux over pathlength ds is given by eq. (2.52) as:

$$dF = -k_{\mathrm{abs}} \rho_a F ds \tag{2.57}$$

By substituting for ds in terms of dz, we obtain

$$\cos\theta \frac{dF}{F} = k_{\mathrm{abs}} \rho_a dz \tag{2.58}$$

We define *optical depth, τ, as the optical path (or optical thickness) measured vertically downwards from the top of the atmosphere* to some altitude,

$$\tau = \int_z^\infty k_{abs}\rho_a dz' = \int_{\tau=\tau}^{\tau=0} -d\tau' \tag{2.59}$$

Note that we need to be careful with the sign in eq. (2.59). Thus, (2.58) becomes:

$$\cos\theta \frac{dF}{F} = -d\tau \tag{2.60}$$

which integrates to

$$F = F_\infty e^{-\tau/\cos\theta} \tag{2.61}$$

Here, F_∞ is the downward flux density at the top of the atmosphere where $z = \infty$. In (2.61), the flux density decreases exponentially along the slant path (inclined at angle θ to the vertical) with a total optical thickness given by $\tau/\cos\theta$.

2.4.2.3 Direct and Diffuse Solar Fluxes

The optical depth of atmospheres can often be relatively high. For example, on Mars the broadband visible optical depth in global dust storms rises to ~5 (Colburn *et al.*, 1989; Kahn *et al.*, 1992), which means that direct beam sunlight is attenuated by a factor of $e^{-5} \approx 1/148$. However, this apparently large factor bears little relationship to the sunlight reaching the ground because much of the visible light is scattered by the dust and reaches the ground as diffuse light (Haberle *et al.*, 1993a).

In reality, the total flux at the ground is the sum of two components (Monteith, 2012):

$$\text{total solar flux at ground} = (\text{vertical component of direct flux}) + (\text{diffuse flux}) \tag{2.62}$$

The *direct flux* is the flux normal to the surface of the beam from the solar disk (plus a small component scattered forward), i.e., eq. (2.61) gives F and then $F\cos\theta$ gives the vertical component of F. The *diffuse flux* is all other scattered radiation from the sky and from clouds, either scattered or transmitted. (The sky component includes a bright white glow around the Sun, known as the *aureole*, which is caused by forward scattering of sunlight by aerosol particles.) Another term for the total flux at the ground is the *global flux*, which is generally considered as the solar radiation (0.3–3 μm) received from a solid angle of 2π steradians on a horizontal surface.

On Earth during daylight, the direct beam rarely exceeds 75% of the solar constant, i.e. ~1030 W m^{-2}, with the remaining 25% attributable to scattering and absorption in roughly equal proportions. In a cloudless sky, the diffuse flux is maximally ~200 W m^{-2} at noon and no more than ~10%–15% of the total flux. If the sky is obscured by dense cloud, the diffuse flux is the total flux at the ground.

The calculation of the diffuse flux is more complicated than the direct flux. Typically, *aerosols* (defined as suspended solid or liquid particles) are involved in the scattering and absorption, so we must know their properties.

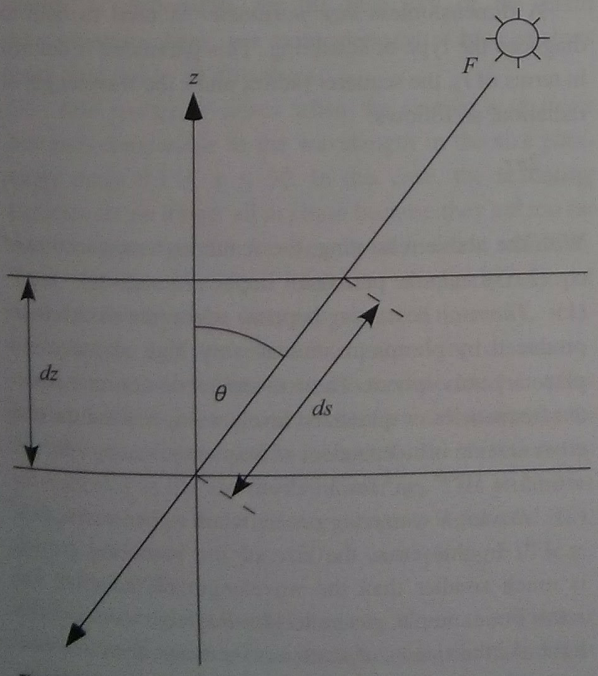


Figure 2.11 The slanted extinction path *ds*, of a direct solar beam through a slab of air in a plane-parallel atmosphere. The Sun is at a zenith angle of θ.

2.4.2.4 Extinction, Absorption, and Scattering

The extinction coefficient mentioned previously is the sum of an absorption coefficient and a scattering coefficient, i.e.,

$$k_\nu = a_\nu + s_\nu \tag{2.63}$$

Both a_ν and s_ν typically depend on temperature, so that k_ν is determined from lab measurements. Absorption or scattering can be due to gases or aerosols. The ratio of s_ν to extinction is known as the *single scattering albedo*:

$$\tilde{\omega} = \frac{s_\nu}{a_\nu + s_\nu} \tag{2.64}$$

The value of $\tilde{\omega}$ is <0.5 for strongly absorbing substances and 1 for purely scattering.

Absorption is when a molecule or atom absorbs a photon, producing an excited state. The lifetime of excitation varies from ~10^{-8} s in the shortwave to ~0.1–10 s in the infrared, and at sufficiently high pressures, molecular collisions generally occur before the excited atom or molecule is able to undergo *radiative decay* and emit a photon in a random direction. Thus, the energy gets transferred to translation, rotation, or vibration that readily interchange energy. Excess vibrational or rotational energy gets converted into kinetic energy by collisions amongst the molecules and is released as heat. The process of changing the photon energy into heat is *thermalization* or *quenching*. For short wavelength photons, such as EUV (i.e., <120 nm) in the upper atmosphere, absorption can also cause *photo-ionization* or *photo-dissociation*. Aerosols can also absorb photons directly and warm the atmosphere during daytime, e.g., dust on Mars (Pollack *et al.*, 1979) or black carbon on Earth (Ramanathan and Carmichael, 2008).

Scattering is when the direct beam of radiation is attenuated because light is redirected out of the original direction of propagation as a result of an interaction with aerosols or when radiative decay of an excited molecule or atom emits a photon in a random direction. Gas molecules scatter solar radiation (*Rayleigh scattering* – see below). Aerosol particles (dust, smoke or cloud droplets) with dimensions comparable to the wavelength of radiation scatter radiation according to *Mie scattering*. Mie theory solves Maxwell's equations for the electromagnetic field around a dielectric particle. Although scattering of photons can result in loss of radiation from the beam, photons can also be scattered into the beam. This is the problem of *multiple scattering*, which occurs in atmospheres loaded with aerosols.

An insight into various forms of scattering can be gleaned from the classical physics of a molecule acting as a simple harmonic oscillator. We consider an electron bound to a nucleus with a "spring constant," which undergoes oscillations at a single natural frequency when irradiated. The electric field of the incoming radiation drives the electron up and down and the electron radiates because it is accelerated. In this case, we use the mathematics of a forced oscillator. The following expression for the resulting total scattering cross-section is derived in *The Feynman Lectures on Physics*, pp. 32–36 (Feynman *et al.*, 1963), where it is discussed in the context of light scattering. For simplicity, the expression omits a damping term in the denominator:

$$\sigma_s = \frac{8\pi {r_e}^2}{3}\left[\frac{\upsilon^4}{\left(\upsilon^2 - {\upsilon_0}^2\right)^2}\right] \tag{2.65}$$

Here υ_0 is the resonant frequency and r_e is the *classical electron radius*, which is derived by equating the electrostatic energy of an electron with its relativistic rest energy, giving $r_e = ({q_e}^2/4\pi\varepsilon_0)/(m_e c^2) = 2.82 \times 10^{-15}$ m, where q_e is the electron charge, ε_0 is the permittivity of free space, m_e is the mass of an electron, and c is the speed of light.

How does eq. (2.65) apply? If we know the electric field E, the average of the square of this field times $\varepsilon_0 c$ gives the flux (in W m^{-2}) through a surface normal to the direction in which the radiation is going, i.e.

$$F = \varepsilon_0 c\langle E^2\rangle \tag{2.66}$$

Molecules or atoms in a gas are randomly located so that the power radiated by the total number of molecules in any direction is just the sum of the power scattered by each molecule. So we can multiply the radiated power by the number of free gas molecules. But if N molecules are joined together in a lump such as a tiny aerosol, which is much smaller than the wavelength of radiation, then the electric field of the aggregated molecules will move in phase in response to incident radiation. The amplitude of the scattered radiation will be the sum of the N radiated electric fields in phase, so that the power will be N^2 times the scattered power of a single molecule by eq. (2.66).

A dimensionless *size parameter* is used to help to diagnose the type of scattering. This parameter is defined in terms of r_s, the scatterer radius, and λ the wavelength of radiation, as follows:

$$x = \frac{2\pi r_s}{\lambda} \tag{2.67}$$

With the above reasoning, the scattering cross-section of eq. (2.65) and size parameter implies four cases.

(1) *Thomson scattering* happens when free electrons are produced by photoionization at very high altitudes in a planetary atmosphere. These electrons do not have resonant frequencies or quantized levels, so $\upsilon_0 = 0$ and the total cross-section is independent of frequency, i.e., $\sigma_s = 8\pi {r_e}^2/3 = 6.65 \times 10^{-25}$ cm^2 for an electron.

(2) *Rayleigh scattering* occurs when $\upsilon_0 \gg \upsilon$, and $\sigma_s \propto \upsilon^4 \propto \lambda^{-4}$. In this case, the size of the scattering particle is much smaller than the wavelength of radiation, i.e., $x \ll 1$. For example, gas molecules Rayleigh scatter visible light. As the size r_s, of scatterers increases, the number of molecules they contain scales as ${r_s}^3$, so the scattering power of the particles and effective cross-section grows in proportion to ${r_s}^6$ (see Fig. 2.12). Rayleigh scattering, of

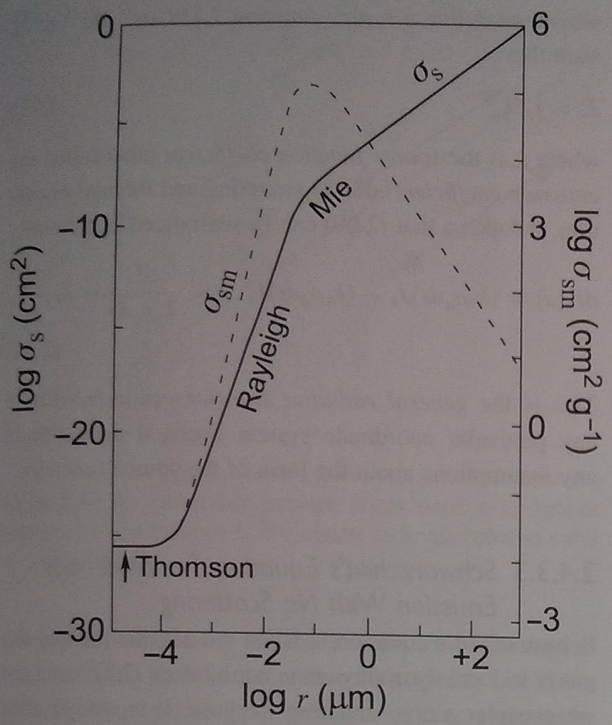


Figure 2.12 The scattering cross-section σ_s and cross-section per unit mass σ_{sm}, as a function of the effective radius r, of a scattering particle, for 0.6 μm radiation (orange visible light). Four regimes of scattering occur: Thomson, Rayleigh, Mie, and geometric. Radius r ranges from the size of an electron to 1 mm. (Redrawn, based on a similar diagram by Lewis (2004), p.185.)

course, is responsible for the blue sky of the Earth because atmospheric gas molecules scatter blue photons more than red in all directions.

(3) *Mie scattering* arises when the size of a scatterer becomes comparable to the wavelength in the size parameter range $0.1 \leq x \leq 50$. In this case, the scattering particles are no longer all in phase because they are too far apart compared to the wavelength of incident radiation. Then, the situation is no longer so simple because of complex relationships between scattered intensity, scattering angle, and the resultant superposition of electromagnetic radiation. Fine dust, smoke, or cloud droplets produce *Mie scattering*, often with forward scattering. In models, a parameter used to account for the directionality of scattering is the asymmetry factor, g_s. This parameter indicates whether the scattering direction is forward ($0 < g_s \leq 1$), backward ($-1 \leq g_s < 0$) or isotropic ($g_s = 0$) (e.g., see Petty (2006)).

(4) *Geometric scattering* happens when the scatterer size is much greater than the wavelength of incident radiation, $x > 50$. Then, the scattering cross-section grows in proportion to the geometric area, πr_s^2, as shown in the solid line of Fig. 2.12.

2.4.3 Radiative Transfer in the Thermal Infrared

In calculating the transmission of purely visible and UV radiation passing through an atmosphere, we saw that we did not need to consider the thermal emission in the atmosphere because thermal emission at planetary temperatures has negligible radiative output in the shortwave. For example, with a visible–UV spectrometer on the surface of Earth or Mars looking upwards, we could calculate the column abundance of a species from measured atmospheric absorption across a particular band without having to know the thermal structure of the atmosphere.

In the thermal-IR, we need to consider the attenuation of IR radiation through a path *and* the temperature-dependent emission in the IR at points along the path. Consequently, radiative transfer in the IR is inherently more complicated than in the visible and UV and forms the subject of a large literature.

2.4.3.1 *Spectral Visualization of the Greenhouse Effect*

Satellite measurements of the emission spectrum of the Earth show that the Earth–atmosphere system is not a blackbody in reality (Fig. 2.13). In comparing the level of a blackbody spectrum corresponding to the N. African surface temperature in Fig. 2.13 of ~320 K, one sees "holes" or dips in the spectrum due to gas absorption. The base of the dips corresponds to the effective temperature of the emission. Photons from the bottom of the dips emanate not from the surface but from higher up in the atmosphere where it is colder. In particular, the 20 μm water band has emission at the temperature of the lower troposphere. The bottom of the absorption dip for carbon dioxide emits at the temperature near the tropopause, shown schematically in Fig. 2.13(b).

In Fig. 2.13(a), the total area of the emitted spectrum would be the total emitted power per unit surface area. If we had a globally averaged version of Fig. 2.13(a), the total emitted power would correspond to that of a blackbody at 255 K, which we calculated in Sec. 2.2.4 from energy balance. Gaseous absorption (the area of the "holes" in the real spectrum) obviously decreases the total area and the total power emitted. So in order to maintain equilibrium between power received from the Sun and power emitted, the surface temperature (which defines the blackbody curve where there is an atmospheric window, e.g., 8–12 μm) must increase well above 255 K. Thus, Fig. 2.13(a) or any similar spectrum obtained from orbit can be interpreted as a visualization of Earth's greenhouse effect.

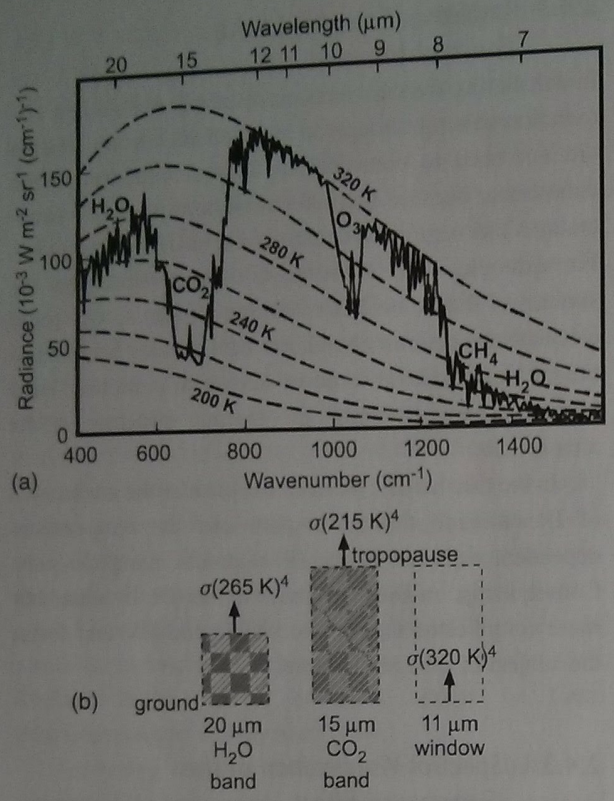


Figure 2.13 (a) The spectral emission function for noon over a vegetated region of the Niger Valley in N. Africa. Dashed lines show blackbody curves for particular temperatures. (Adapted from Hanel *et al.* (1972).) (b) A schematic showing how to interpret the meaning of parts of the curve in (a). The arrows indicate from where blackbody fluxes originate, according to the Stefan–Boltzmann Law. (Part (b) follows a concept from Jacob (1999), p. 132.)

2.4.3.2 *General Equation of Radiative Transfer*

We now derive the general equation of radiative transfer, which describes how radiation passes though a medium in any coordinate system. If a beam of monochromatic radiance I_ν [W m^{-2} Hz^{-1} sr^{-1}] passes through an elemental path ds, the change of intensity of the beam will be as follows:

$$\text{intensity change} = \text{emission} - \text{extinction}$$
$$dI_\nu = dI_\nu(\text{emitted}) - dI_\nu(\text{extinguished}) \tag{2.68}$$

Using our previous expression (2.52) for the mass extinction coefficient (k_ν [m^2 kg^{-1}]), we can express the extinction component using the Extinction Law, so that

$$dI_\nu(s) = dI_\nu(\text{emitted}) - k_\nu \rho_a I_\nu(s) ds \tag{2.69}$$

where ρ_a is the density of the absorbing and/or scattering gas. The increase in intensity due to emission and multiple scattering is defined as:

$$dI_\nu(\text{emitted}) = j_\nu \rho_a ds = k_\nu J_\nu(s) \rho_a ds \tag{2.70}$$

where we define a *source function* J_ν[W m^{-2} Hz^{-1} sr^{-1}] such that

$$J_\nu = j_\nu / k_\nu \tag{2.71}$$

where j_ν is the *source function coefficient* (also called the *emission coefficient*) due to scattering and thermal excitation. It follows that (2.69) can be rearranged as follows:

$$dI_\nu(s) = (k_\nu \rho_a ds) J_\nu - (k_\nu \rho_a ds) I_\nu \quad \Rightarrow \quad \frac{dI_\nu}{k_\nu \rho_a ds} = J_\nu - I_\nu \tag{2.72}$$

This is the *general radiative transfer equation* without any particular coordinate system imposed and without any assumptions about the form of the source function.

2.4.3.3 *Schwarzchild's Equation: For Blackbody Emission With No Scattering*

Schwarzchild's equation is when we assume (a) that the gas is in local thermodynamic equilibrium (LTE) and (b) we consider a non-scattering medium. It is named after astrophysicist Karl Schwarzchild who first considered such a solution to the radiative transfer equation for the Sun's atmosphere in 1914. LTE means that the source function defined by (2.71) is given by the Planck function, i.e.,

$$J_\nu = B_\nu(T) \tag{2.73}$$

Hence, the equation of radiative transfer can be rewritten as

$$\frac{dI_\nu}{k_\nu \rho_a ds} = B_\nu(T) - I_\nu \quad \text{Schwarzchild's equation} \tag{2.74}$$

Because we are neglecting scattering, k_ν is now the mass *absorption* coefficient rather than the mass extinction coefficient.

We shall consider two solutions to eq. (2.74) (Schwarzchild's equation): (1) a general form for the solution; (2) the case of a plane parallel atmosphere.

2.4.3.4 *A General Solution to Schwarzchild's Equation*

We obtain a general solution to Schwarzchild's equation by considering a path for radiation without a specific coordinate system and integrating the equation. We define a monochromatic optical path between points s and s_1 (Fig. 2.14), as

$$\tau_\nu = \int_s^{s_1} k_\nu(s') \rho_a(s') ds' \tag{2.75}$$

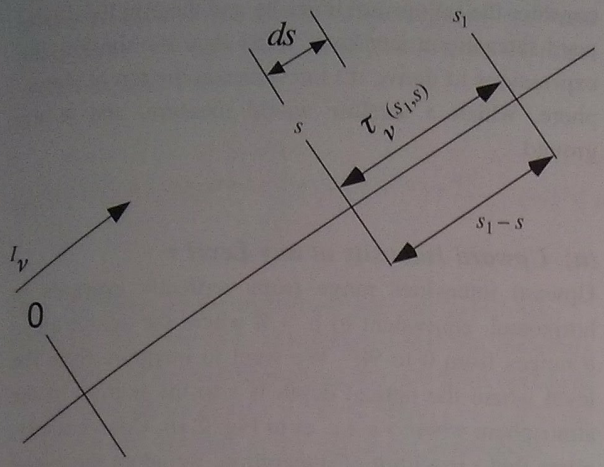


Figure 2.14 An optical path between points s and s_1 for light of frequency ν and radiance I_ν. The optical thickness between s and s_1 is $\tau_\upsilon(s_1, s)$.

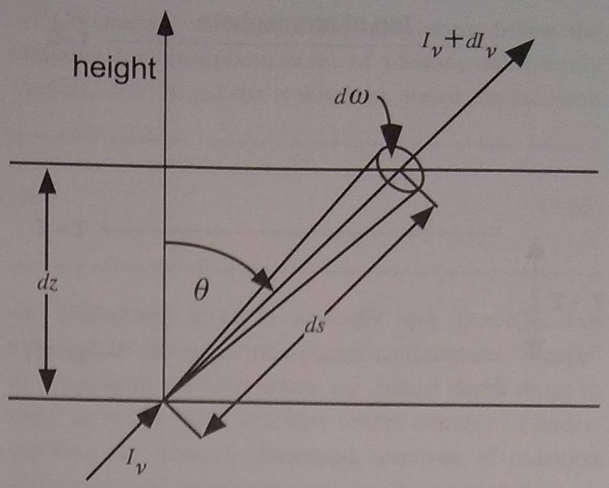


Figure 2.15 A beam of monochromatic radiance I_ν passes through a medium at angle θ from the vertical. An incremental distance ds is traversed and the beam has a solid angle $d\omega$. The vertical incremental distance is $dz = ds\cos\theta$.

using a primed dummy variable s' because s is the start of the path and a limit of integration. We specify our direction such that optical depth decreases as s increases (to be analogous to altitude), so that the elemental optical depth between s and s_1 is

$$d\tau_\upsilon = -k_\upsilon(s')\rho_a(s')ds' \tag{2.76}$$

Then eqn. (2.74) becomes

$$\frac{dI_\nu(s)}{d\tau_\nu(s_1,s)} = I_\nu(s) - B_\nu[T(s)] \tag{2.77}$$

We can multiply this equation by an integrating factor, $e^{-\tau_\nu(s_1,s)}$ and integrate between 0 and s_1, as follows:

$$dI_\nu(s)e^{-\tau_\nu(s_1,s)} + I_\nu(s)\left(-d\tau_\nu(s_1,s)e^{-\tau_\nu(s_1,s)}\right) = -B_\nu[T(s)]e^{-\tau_\nu(s_1,s)}d\tau_\nu(s_1,s)$$

$$\int_0^{s_1} d\left[I_\nu(s)e^{-\tau_\nu(s_1,s)}\right] = -\int_0^{s_1} B_\nu[T(s)]e^{-\tau_\nu(s_1,s)}d\tau_\nu(s_1,s) \tag{2.78}$$

which has solution (substituting in the integral for $d\tau_\upsilon$ from eq. (2.76)),

$$I_\nu(s_1) = I_\nu(0)e^{-\tau_\nu(s_1,0)} + \int_0^{s_1} B_\nu[T(s)]e^{-\tau_\nu(s_1,s)}k_\upsilon\rho_a ds$$

$$I_\nu(s_1) = I_\nu(0)e^{-k_\upsilon\rho_a s_1} + \int_0^{s_1} B_\nu[T(s)]e^{-k_\upsilon\rho_a(s_1-s)}k_\upsilon\rho_a ds \quad \text{for constant } k_\upsilon \text{ and } \rho_a \tag{2.79}$$

Here the first term is the absorption attenuation due to the absorber from 0 to s_1, in the form of the Extinction Law. The second term is the emission contribution along the path 0 to s_1. Note that to integrate (2.79), we need to know T, ρ_a, and k_ν along the path.

2.4.3.5 *Plane-Parallel Solutions to the Radiative Transfer Equation: Upward and Downward Intensities*

We now consider radiative transfer in a plane-parallel atmosphere. We define our coordinate system as a beam of radiance I_ν passing upward through a layer of depth dz at an angle to the vertical of θ. Then we get $dz = ds\cos\theta$ (Fig. 2.15). Returning to generality by using the source function, J_ν, we have from eq. (2.74),

$$\cos\theta\frac{dI_\nu}{k_\nu\rho_a dz} = J_\nu - I_\nu \tag{2.80}$$

Let us re-introduce the optical depth measured downwards from the top of the atmosphere (eq. (2.59)),

$$\tau = \int_z^\infty k_\nu\rho_a dz' = \int_{\tau=\tau}^{\tau=0} -d\tau' \tag{2.81}$$

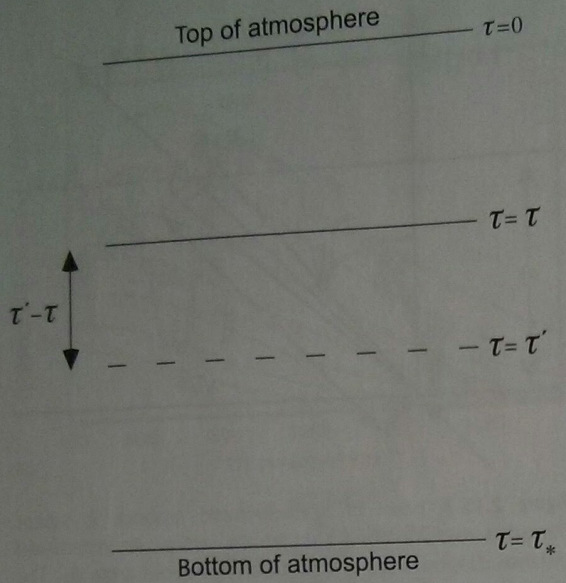


Figure 2.16 Optical depth coordinate system for a solution to the radiative transfer equation.

consider the (a) upward intensity and then (b) the downward intensity at any level. After that, we simplify the expressions to derive (c) intensities at the top of atmosphere, which a satellite would measure, and at the ground.

(a) Upward Intensity at any Level τ

Upward intensities range from vertically upwards to horizontal, equivalent to $\mu > 0$ where the zenith angle θ ranges from 0 to 90°. We want to integrate from the level where the optical depth is τ to the bottom of the atmosphere where $\tau = \tau_*$, as in Fig. 2.16. Consequently, given that τ is a limit of integration, we adopt the use of the dummy variable $\tau'-\tau$, according to scheme of Fig. 2.16. To obtain the upward intensity at any level where the optical depth is τ, we multiply (2.82) by an integrating factor $e^{-(\tau'-\tau)/\mu}$. Here we have dropped the ν subscript on τ, and assume it is implicit. So we have the following:

$$e^{-(\tau'-\tau)/\mu} dI_\nu(\tau') + I_\nu(\tau')\left(-\frac{d\tau'}{\mu} e^{-(\tau'-\tau)/\mu}\right) = -J_\nu(\tau') e^{-(\tau'-\tau)/\mu} \frac{d\tau'}{\mu}$$

$$\Rightarrow d\left[I_\nu(\tau') e^{-(\tau'-\tau)/\mu}\right] = -J_\nu(\tau') e^{-(\tau'-\tau)/\mu} \frac{d\tau'}{\mu} \quad \Rightarrow d\left[I_\nu(\tau') e^{-(\tau'-\tau)/\mu}\right]\Big|_\tau^{\tau_*} = -\int_\tau^{\tau_*} J_\nu(\tau') e^{-(\tau'-\tau)/\mu} \frac{d\tau'}{\mu}$$

$$I_\nu(\tau) \uparrow = I_\nu(\tau_*) e^{-(\tau_*-\tau)/\mu} + \int_\tau^{\tau_*} J_\nu(\tau') e^{-(\tau'-\tau)/\mu} \frac{d\tau'}{\mu} \quad (1 \geq \mu > 0) \tag{2.83}$$

We also define $\mu \equiv \cos\theta$, the cosine of the zenith angle. Then, we can rewrite (2.80) as

$$\mu \frac{dI_\nu}{d\tau_\nu} = I_\nu - J_\nu \tag{2.82}$$

This is the basic equation for the problem of radiative transfer in plane-parallel atmospheres. Here J_ν[W m^{-2} Hz^{-1} sr^{-1}] and I_ν[W m^{-2} Hz^{-1} sr^{-1}] are functions of μ and τ, and in the most general form, they are also functions of azimuth angle.

Here, the terms on the right-hand side are as follows.

First term: the absorption attenuation due to the absorbing medium of the radiation along the path from the ground where $\tau = \tau_*$ to the $\tau = \tau$ level. Here $I_\nu(\tau_*)$ represents the upward intensity at the ground.

Second term: upwards emission contribution along the path from the ground to the $\tau = \tau$ level. Here $J_\nu(\tau')$ corresponds to emission in the upward direction.

We could work in terms of altitude, but if we transform eq. (2.83) in terms of z, it looks more complicated, as follows:

$$I_\nu(z) \uparrow = I_\nu(z=0) e^{-\frac{1}{\mu}\int_0^z k_\nu(z')\rho_a(z')dz'} + \int_0^z k_\nu(z')\rho_a(z') J_\nu(z') e^{-\frac{1}{\mu}\int_{z'}^z k_\nu \rho_a dz} \frac{dz'}{\mu} \tag{2.84}$$

We can solve (2.82) using the same procedure that we used to solve (2.77). We take the atmosphere as being bounded at the top and bottom (Fig. 2.16). We

(b) Downward Intensity at any Level τ

To obtain the solution for the downward intensity ($\mu < 0$) at any level where the optical depth is τ, we replace μ

by $-\mu$, multiply (2.82) by an integrating factor $e^{\tau/\mu}$and integrate from the top of the atmosphere where $\tau = 0$ to the $\tau = \tau$ level. This yields

$$I_\nu(\tau)\downarrow = I_\nu(\tau=0)e^{-\tau/\mu} + \int_0^\tau J_\nu(\tau')e^{-(\tau-\tau')/\mu}\frac{d\tau'}{\mu} \quad (-1 \le \mu < 0) \tag{2.85}$$

Here, the terms on the right-hand side are as follows.

First term: the absorption attenuation due to the absorbing medium of the radiation along the path from the top of the atmosphere to the $\tau = \tau$ level. Here, $I_\nu(\tau = 0)$is the downward intensity at the top of the atmosphere.

Second term: the emission contribution along the path from the top of the atmosphere to the $\tau = \tau$ level. Here, $J_\nu(\tau')$ corresponds to emission in the downward direction.

(c) *Intensities Seen from Space or from the Surface*

Frequently, one measures the emergent intensity at the top of the atmosphere with a satellite or the emergent intensity at the bottom of an atmosphere with a surface-based instrument. Emerging intensity at the top of the atmosphere is derived from (2.83) by setting $\tau = 0$ in the integral limit, i.e.,

$$I_\nu(\tau=0)\uparrow = I_\nu(\tau_*)e^{-\tau_*/\mu} + \int_0^{\tau_*} J_\nu(\tau')e^{-\tau'/\mu}\frac{d\tau'}{\mu} \tag{2.86}$$

Emerging intensity at the bottom of the atmosphere is derived from (2.85) if we set $\tau = \tau_*$, i.e.,

$$I_\nu(\tau_*)\downarrow = I_\nu(\tau=0)e^{-\tau_*/\mu} + \int_0^{\tau_*} J_\nu(\tau')e^{-(\tau_*-\tau')/\mu}\frac{d\tau'}{\mu} \tag{2.87}$$

Recall from previous discussion (Sec 2.3.1) that we can derive a hemispherical flux from the intensity at a particular frequency by integrating over a hemisphere. Then we derive a total flux density [W m^{-2}] by integrating over all frequencies.

2.4.4 Level of Emission and the Meaning of "Optically Thick" and "Optically Thin"

Below we show that absorbed radiation and emitted radiation come from a region in the atmosphere near the region of unit optical depth for a particular wavelength or frequency.

Commonly, we talk of the atmosphere below the emission (or absorption) level of photons as *optically thick* ($\tau_\nu >> 1$) and the atmosphere above the emission (or absorption) level as *optically thin* ($\tau_\nu << 1$). For example, IR photons from a planet's atmosphere "escape" to space from the level where the optical depth drops to unity or below. This is a very useful concept in understanding the remotely measured spectrum of radiation emitted from planetary atmospheres. We applied this idea in discussing Fig. 2.13.

In the shortwave, incoming solar UV photons are absorbed where the level reaches unity optical depth for a particular UV wavelength. Hence, in the UV, this level can be thought of as where action occurs, such as heating or ionization. We consider shortwave absorption and infrared emission levels, in turn.

2.4.4.1 *The Level of Shortwave Absorption*

At what level do some atmospheric gases, such as ozone in the Earth's atmosphere or methane in giant planet atmospheres, absorb shortwave sunlight and heat the atmosphere? The vertical distribution of any monochromatic component of the solar flux absorbed in a planetary atmosphere can be calculated as follows. The monochromatic flux crossing a surface normal to the solar beam is

$$F_{\nu\odot} = I_{\nu\odot}\delta\Omega \tag{2.88}$$

where $I_{\nu\odot}$is the monochromatic solar radiance in the small solid angle subtended by the Sun at the planet, $\delta\Omega$. Within the planetary atmosphere at height z, the direct flux can be deduced from the Extinction Law, which accounts for absorption by overlying atmospheric layers,

$$F_\nu = \mu F_{\nu\odot}e^{-\tau_\nu/\mu} \tag{2.89}$$

where $\mu \equiv \cos\theta$ is the cosine of the zenith angle. The contribution to volume heating rate Q is given by $Q = -(dF_{net}/dz)$ where F_{net} is the net flux, that is, the upward minus the downward flux [see eqns. (2.34) and (2.35)]. In our case the solar flux is downward, so

$$Q = \frac{dF_\nu}{dz} = \frac{dF_\nu}{d\tau}\frac{d\tau_\nu}{dz} = \left(-\frac{1}{\mu}\mu F_{\nu\odot}e^{-\tau_\nu/\mu}\right)(-k_\nu\rho_\nu) = (k_\nu\rho_\nu F_{\nu\odot})e^{-\tau_\nu/\mu} \tag{2.90}$$

where we have used our definition of optical depth $d\tau_\nu/dz = -k_\nu\rho_\nu$. Note how function (2.90) will change with height.

- At the top of the atmosphere, $\tau_\nu(z)$ is small and the exponential factor remains near 1, i.e., $e^{-\tau_\nu/\mu} \to e^0$. Consequently, the overall function increases downward as the absorption term $k_\nu\rho_\nu$ in the curved brackets increases.
- Eventually, the increase in optical depth becomes significant, so the term $\exp(-\tau_\nu/\mu)$begins to dominate over the downward increase of $k_\nu\rho_\nu$, and the volume heating rate begins to decrease exponentially downwards.

Consequently, the volume heat rate, Q, has a single maximum in the planetary atmosphere.

To find the maximum of Q, we differentiate eq. (2.90) w.r.t. z and set $dQ/dz = 0$. We use the product rule of calculus, where the first term $(k_\nu\rho_\nu)$ multiplies the second term $F_{\nu\odot}e^{-\tau_\nu/\mu}$. Hence,

$$\frac{dQ}{dz} = 0 = (k_\nu\rho_\nu)\left(\cancel{F_{\nu\odot}e^{-\tau_\nu/\mu}}\frac{(k_\nu\rho_\nu)}{\mu}\right) + \cancel{F_{\nu\odot}e^{-\tau_\nu/\mu}}\frac{d(k_\nu\rho_\nu)}{dz}$$

so

$$\frac{(k_\nu\rho_\nu)^2}{\mu} = -\frac{d(k_\nu\rho_\nu)}{dz} \tag{2.91}$$

In this differentiation, we have used $d\tau_\nu/dz = -k_\nu\rho_\nu$ from our definition of optical depth. Using this definition yet again, we note that $d(k_\nu\rho_\nu)/dz = -d^2\tau_\nu/dz^2$ and $(k_\nu\rho_\nu)^2 = (d\tau_\nu/dz)^2$. So eq. (2.91) can be rewritten (swapping left and right sides) as

$$\frac{d^2\tau_\nu}{dz^2} = \sec\theta\left(\frac{d\tau_\nu}{dz}\right)^2 \tag{2.92}$$

The density of absorbing gases usually drops approximately exponentially with height, and the absorption coefficient is usually either constant or a function of pressure that can be expressed as an approximate power law relationship. (The pressure dependence is because of *pressure* or *collision-induced broadening* of absorption lines that we meet later in Sections 2.5.6 and 2.5.7.2.) Optical depth also usually varies approximately exponentially with height, so that

$$\tau_\nu = \tau_{\nu 0}\exp\left(-\frac{z}{H_\tau}\right) \tag{2.93}$$

where H_τ is a scale height for optical depth. Consequently,

Substituting these approximations into eq. (2.92), we find the following:

$$\frac{1}{\cancel{H_\tau^2}}\cancel{\tau_\nu} = \sec\theta\left(\frac{1}{\cancel{H_\tau^2}}\cancel{\tau_\nu}\right)$$

$$\Rightarrow \tau_\upsilon \sec\theta = 1, \qquad \tau_\upsilon = 1 \text{ when } \theta = 0 \tag{2.94}$$

This important result tells us the following.

> *At any wavelength, the maximum volume heating rate, or the maximum of any related physical effect such as ionization or dissociation, occurs at the level where the optical depth is unity for normal incidence radiation, or at a level that is higher by a factor of sec θ for solar radiation incident at any other zenith angle θ.*

The level of absorption forms a layer with a characteristic width. From eqs. (2.90) and (2.93), we can deduce that the volume absorption rate decreases by a factor of e within a height range H_τ above the maximum, and decreases even more rapidly below the maximum. So the monochromatic absorption peaks near or above $\tau_\nu = 1$ and has a characteristic width ~ H_τ, the scale height for optical depth, which is comparable to or smaller than a density scale height. Such a vertical structure is called a *Chapman layer*. These atmospheric layers where action is occurring (heating, ionization, photodissociation) correspond to specific wavelengths of solar radiation.

2.4.4.2 *The Level of Emission of Infrared*

A matter of practical utility is the measurement by spacecraft of IR radiation emitted from a planetary atmosphere or surface. Consequently, we consider monochromatic radiation emitted at zenith angle θ according with the Planck function and detected by a spacecraft outside the atmosphere at $z = \infty$. To simplify matters, we neglect the radiance emitted from the lower boundary of the atmosphere, which is a good approximation for atmospheres with large optical depth to the surface. Then, in our expression for upward intensity (eq. (2.84)) we neglect the first term, which is due to surface radiation, and keep the second. We also adjust the integral limits to bound the whole atmosphere and assume blackbody emission, i.e., $J_\nu = B_\nu$. Hence,

$$I_\nu(\infty) \uparrow\approx \sec\theta\int_0^\infty k_\nu\rho_a B_\nu \exp\left[-\sec\theta\int_z^\infty k_\nu\rho_a dz'\right]dz$$

$$= \sec\theta\int_0^\infty B_\nu W_\nu dz \tag{2.95}$$

where

$$W_\nu(z) \equiv k_\nu(z)\rho_a(z)\exp\left[-\sec\theta\int_z^\infty k_\nu(z')\rho_a(z')dz'\right] \tag{2.96}$$

is a "weighting function." Thus according to eq. (2.95), the outgoing monochromatic radiance is a weighted vertical integral of the Planck function, with a weight factor given by eq. (2.96).

In fact, the form of the weighting function means that radiation of a certain frequency is emitted from a distinct layer. Function W_υ has the same form as the factor multiplying the monochromatic solar flux in eq. (2.90) for the volume absorption rate, Q. Consequently, except for a possible additional lower boundary term, the radiance measured by the satellite is equal to the secant of the zenith angle times the weighted integral of the Planck function, where the weighting factor has all the properties of the vertical distribution of incoming absorbed monochromatic solar radiation. In particular, the emitted monochromatic radiance comes from a layer with characteristic thickness of order of a scale height in optical depth or smaller centered at the level where

$$\tau_\upsilon \sec\theta = 1, \qquad \tau_\upsilon = 1 \text{ when } \theta = 0$$

This is the same condition as eq. (2.94), which shows some unity between infrared and shortwave radiation.

2.4.5 Radiative and Radiative–Convective Equilibrium

The average vertical temperature profile and climate of a planet depends, principally, on radiation and convection. The temperature profile calculated according to radiative equilibrium in an atmosphere like the Earth's is found to be superadiabatic near the surface, which is unstable to vertical motion. Hence convection ensues and establishes a stable adiabatic, subadiabatic, or moist adiabatic vertical temperature profile. This lapse rate will persist up to some level, the *radiative–convective boundary*, above which radiative equilibrium tends to dominate the average vertical profile of temperature. To show that atmospheres behave like this, we consider some simple equilibrium models.

2.4.5.1. *Atmospheric "Skin Temperature"*

To discuss atmospheric thermal structure, we start with *atmospheric skin temperature*, which is the asymptotic temperature at high altitudes of an upper atmosphere that is optically thin in the thermal-IR and transparent to shortwave radiation.

To derive an equation for skin temperature, we consider a two-layer model of an atmosphere on a planet with an effective temperature T_{eff}. A flux of IR radiation given by σT_{eff}^4 (eq. (2.46)) goes upwards from an emission level in the lower atmosphere (where IR photons "escape") into an overlying slab of optically thin gas at temperature T_{skin} (Fig. 2.17). We assume that the slab is a gray absorber with an absorptivity equal to the emissivity, ε, by Kirchoff's Law (eq. (2.48)). In equilibrium, the outgoing IR flux absorbed by the slab must equal the IR flux absorbed, as follows:

$$2\varepsilon\sigma T_{skin}^4 = \varepsilon\sigma T_{eff}^4 \quad \Rightarrow T_{skin} = \frac{T_{eff}}{2^{1/4}} \tag{2.97}$$

Thus, the skin temperature in an optically thin, gray upper atmosphere is colder than the effective temperature by a factor of $2^{-1/4}$ or ~0.84. For the Earth with $T_{eff} = 255$ K, we calculate $T_{skin} = 214$ K.

In planetary atmospheres that absorb little shortwave radiation, the skin temperature is the asymptotic

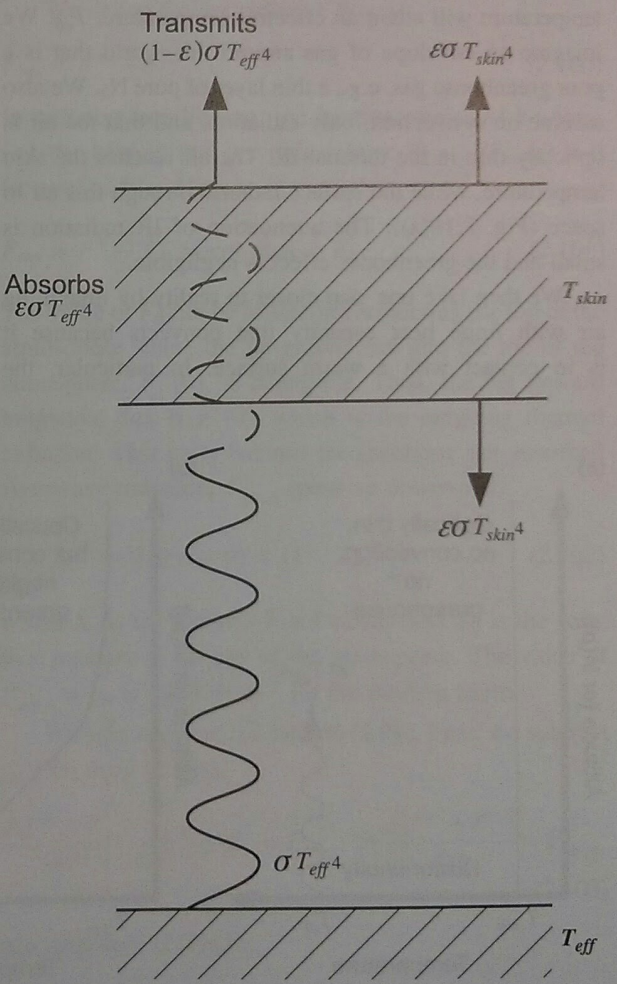


Figure 2.17 A two-layer atmospheric model to understand the concept of atmospheric *skin temperature*. The lower atmosphere is represented by layer that is optically thick in the infrared and emits at the effective temperature, T_{eff}. The upper atmosphere is represented by a layer that is optically thin in the infrared and has a temperature T_{skin} and emissivity ε, which is equal to absorptivity by Kirchoff's Law. By considering conservation of energy in the upper layer, we determine that $T_{skin} = 2^{-1/4} T_{eff}$ (see text).

temperature of a nearly isothermal stratosphere. For real atmospheres with stratospheric inversions, the skin temperature provides an estimate of the tropopause minimum temperature. For example, the Earth's observed global mean tropopause temperature is 208 K (Han *et al.*, 2011), which differs from T_{skin} by only 3%.

2.4.5.2 *Illustrative Radiative–Convective Thought Experiments*

To understand radiative–convective equilibrium, it is useful to consider some thought experiments. In Fig. 2.18(a), we imagine a planet that is rotating rapidly so that we can apply time- and spatial-average energy balance, in which the planet attains an isothermal mean temperature in equilibrium with sunlight. The surface temperature will attain an effective temperature, T_{eff}. We imagine an envelope of gas around this world that is a poor greenhouse gas, e.g., a thin layer of pure N_2. We also assume no convection, only radiation, and that the air is optically thin in the thermal-IR. The air reaches the skin temperature, while the surface radiates through this air to space (Fig. 2.18(a)). The attenuation of IR radiation is small and the greenhouse effect is negligible.

We then take one step closer to reality by imagining air with finite heat capacity that convects because it is in contact with a warm surface. In particular, the discontinuous T gradient in Fig. 2.18(a) is unstable to convection. The result is the temperature profile in Fig. 2.18(b).

Finally, we imagine air that contains greenhouse gases and is optically thick to thermal-IR near the surface. Then the emission level will be at the altitude where the mass of air between the emission level and space becomes optically thin to thermal-IR (Fig. 2.18(c)). The emission level is no longer at the surface pressure but at a lower pressure, p_{emission}. This *emission level* is located where a column of air from space down to p_{emission} has an optical depth of about unity. If there is a dimensionless mass mixing ratio of the absorber, μ_a, which is constant with height, and if the absorber has a broadband (gray), pressure-independent mass absorption coefficient k_a (m^2/kg), the optical depth τ_{emission} at the emission level is

$$\tau_{\text{emission}} = k_a \mu_a \left(\frac{p_{\text{emission}}}{g} \right) \approx 1 \qquad (2.98)$$

Here, we have combined eq. (2.59) and eq. (1.19), where g is gravitational acceleration. We can think of the region below pressure level p_{emission} as sliced up into blackbody slabs each of unity optical depth (Goody and Walker, 1972, p. 56). Each slab absorbs thermal-IR radiation, and IR photons only escape to space from the top slab.

If there is an even greater mixing ratio of greenhouse gas, μ_a, the IR emission level moves up in altitude to a

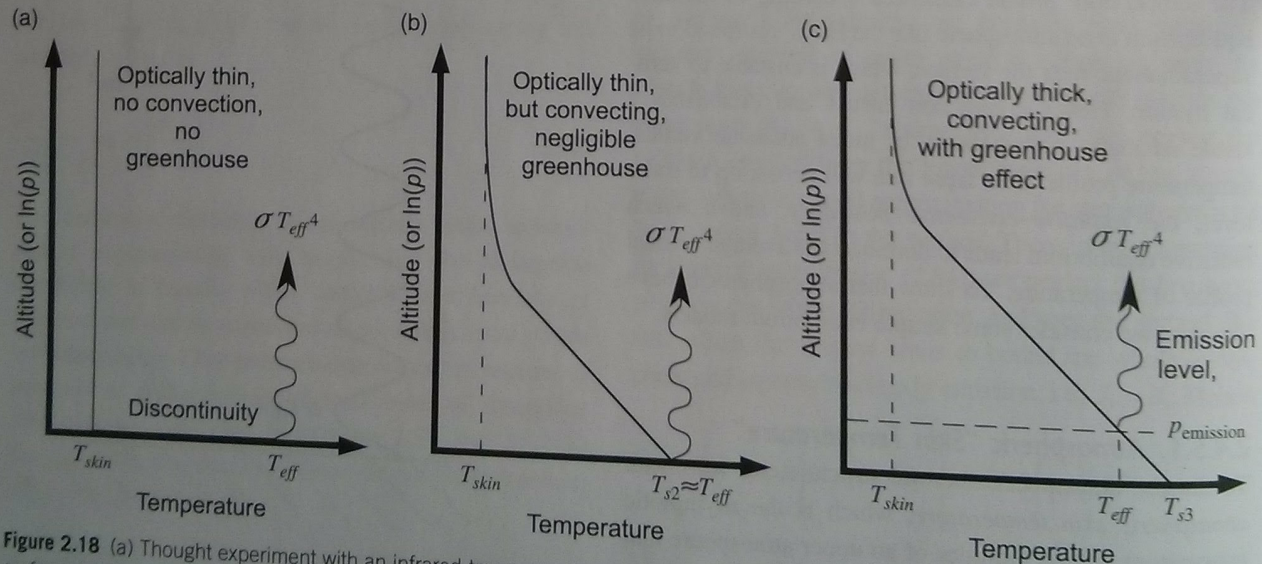


Figure 2.18 (a) Thought experiment with an infrared-transparent atmosphere and no convection. The surface reaches a temperature of the effective temperature, T_{eff}, while the optically thin atmosphere attains the skin temperature, T_{skin}. There is a temperature discontinuity at the surface. (b) Thought experiment with negligible greenhouse gases but convection. A convective profile sets in. The surface temperature T_{s2} is T_{eff}. (c) Thought experiment with greenhouse gases and convection. The near-surface atmosphere is optically thick, so that emission at the effective temperature takes place at a higher level where the pressure is p_{emission}. The surface temperature T_{s3} exceeds T_{eff} because of a greenhouse effect.

lower pressure and the surface temperature increases. The skin temperature remains the same.

Modeling the climate of any planet involves more realism for increasing accuracy. The complexity includes actual gas mixtures, p- and T- and wavelength-dependent absorption and scattering, clouds and their feedbacks, three dimensions, atmospheric dynamics and heat transport, ocean dynamics and heat transport, ice sheet dynamics, and feedbacks involving biogeochemical cycles. To generally understand planetary climates, we need to consider these issues for all possible atmospheres. This is part of what makes discoveries of exoplanets with atmospheres a very grand challenge.

2.4.5.3 *Radiative Equilibrium in a Gray Atmosphere*

To be more quantitative, we can use a simple radiative equilibrium model. Several assumptions make the problem amenable to analytical solution.

(1) We consider a gray atmosphere, which means that the extinction coefficient is independent of frequency or, equivalently, wavelength.

(2) We assume an atmosphere transparent to solar radiation, i.e. visible and other shortwave radiation.

(3) We adopt a so-called *two-stream* approximation, in which we treat the upward (F^+) and downward (F^-) fluxes independently.[4]

(4) We also use the *diffuse approximation* in which we replace radiances I and B in Schwarzchild's radiative transfer equation (i.e., eq. (2.82) of the form $dI/d\tau = I - B$) with irradiance, i.e. flux, provided that we scale the optical depth τ by an appropriate factor as $\tau^* = D\tau$, where D is the *diffusivity factor*. Basically, D takes account of the slant paths and allows us to work with fluxes instead of radiances. An approximation justified by comparison to numerical results is $\tau^* = 1.66\tau$ (Armstrong, 1968; Rodgers and Walshaw, 1966), where τ is the actual optical depth and τ^* is *the scaled optical depth* with $D = 1.66$. In other treatments, D is taken as 3/2 (Goody and Yung, 1989; Weaver and Ramanathan, 1995). We take the scaled optical depth τ^* as our vertical coordinate, which increases downward from the top of the atmosphere ($\tau^* = 0$) to a value at the ground of $\tau^* = \tau_0$.

With the above assumptions, Schwarzchild's equation (2.74) becomes

$$\frac{dF^+}{d\tau^*} = F^+ - \pi B \qquad -\frac{dF^-}{d\tau^*} = F^- - \pi B, \quad \text{where } \pi B(\tau^*) = \sigma T^4(\tau^*) \tag{2.99a,b}$$

Here, upward and downward fluxes and the spectrally integrated Planck function B depend on τ^*. Signs in eq. (2.99) are such that (by convention) the IR optical depth decreases along the upward path. The flux density is πB by eq. (2.39). Thus, if $T(\tau^*)$ is known, upward (F^+) and downward (F^-) radiative fluxes can be calculated.

We define the net radiative flux as:

$$F_{\text{net}} = F^+ - F^- \tag{2.100}$$

In radiative equilibrium, the heating rate must be zero, i.e., the heating rate per unit volume $Q = -dF_{\text{net}}/dz = 0$ (see eq. (2.34)). Hence,

$$F_{\text{net}} = F^+ - F^- = \text{constant} = F_0 \tag{2.101}$$

The value of F_0 can be found by considering the top of the atmosphere where the longwave flux into the top of the atmosphere, $F^-(0)$, is negligible. Thus, the net upward longwave flux is $F^+(0)$, which is the outgoing thermal radiation. This must balance the incoming net absorbed shortwave radiation, $F^{\odot}_{\text{net,a}}$ (positive downward):

$$F^{\odot}_{\text{net,a}}(\tau^* = 0 \text{ or } \tau = 0) = (1 - A_b)\frac{S_{\odot}}{4} \tag{2.102}$$

where A_b is the planet's Bond albedo and $S_{\odot}$ is the solar flux incident at the top of the atmosphere. The value of $F^{\odot}_{\text{net,a}} = F_0$ is ~239 W m^{-2} for the modern Earth.

We next apply eq. (2.101) to (2.99). First, we subtract (2.99b) from (2.99a),

$$\frac{d(F^+ + F^-)}{d\tau^*} = F^+ - F^- = F_{\text{net}} \quad \Rightarrow \frac{d(F^+ + F^-)}{d\tau^*} = F^{\odot}_{\text{net,a}} \tag{2.103}$$

and then sum (2.99a,b):

$$\frac{d(F_{\text{net}})}{d\tau^*} = \frac{dF^{\odot}_{\text{net,a}}}{d\tau^*} = F^+ + F^- - 2\pi B = 0 \tag{2.104}$$

Hence $F^+ + F^- = 2\pi B$, which can be substituted into (2.103) to give

$$2\frac{d(\pi B)}{d\tau^*} = F^+ - F^- = F^{\odot}_{\text{net,a}} \tag{2.105}$$

[4] There are many other forms of two-stream approximation. A common one is Eddington's approximation in which the intensity field is represented by a Legendre polynomial series, truncated at the second term. This approximation treats the upward and downward fluxes as independent.

We can integrate this equation. For a limit on the integral, we note that at the top of the atmosphere where $\tau^* = 0$, the downward longwave flux $F^-(0) \approx$ zero, so eq. (2.104) gives $\pi B(\tau^* = 0) = F^+(0)/2$, which equals $F^{\odot}_{\text{net,a}}/2$ by eqs. (2.101) and (2.102). Hence

$$\int_{F_0/2}^{\pi B} d(\pi B) = \frac{F^{\odot}_{\text{net,a}}}{2}\int_0^{\tau^*} d\tau^* \quad \Rightarrow \pi B = \frac{1}{2}F^{\odot}_{\text{net,a}}\tau^* + \frac{1}{2}F^{\odot}_{\text{net,a}} \Rightarrow \pi B = \frac{1}{2}F^{\odot}_{\text{net,a}}(\tau^* + 1) \tag{2.106}$$

$$\Rightarrow \boxed{\sigma T^4(\tau) = \frac{F^{\odot}_{\text{net,a}}}{2}(1 + D\tau)}, \qquad T(\tau) = \left[\frac{F^{\odot}_{\text{net,a}}}{2\sigma}(1 + D\tau)\right]^{1/4}$$

The boxed expression in eq. (2.106) is a common result of simple, purely radiative equilibrium models given in textbooks (e.g., Goody and Yung (1989), p. 392; Andrews (2010), p. 85) and reviews (Weaver and Ramanathan, 1995).

We can also derive upward and downward fluxes, including the upward flux at the surface. Since $F^+ + F^- = 2\pi B$ (eq. (2.104)) then eq. (2.106) can also be written as

$$F^+ + F^- = F^{\odot}_{\text{net,a}}(1 + \tau^*) \tag{2.107}$$

Adding (2.107) and (2.101), gives a solution for the upward radiative flux,

$$F^+ = \frac{1}{2}F^{\odot}_{\text{net,a}}(2 + \tau^*) \quad \Rightarrow F^+_{\text{ground}} = \frac{1}{2}F^{\odot}_{\text{net,a}}(2 + \tau_0{}^*) = F^{\odot}_{\text{net,a}}\left(1 + \frac{\tau_0{}^*}{2}\right) \tag{2.108}$$

where F^+_{ground} is the flux from the ground in the lowest layer at scaled optical depth $\tau_0{}^*$. We see a greenhouse effect because F^+_{ground} is greater than the flux to space ($F^{\odot}_{\text{net,a}}$). Substitution of eq. (2.108) into eq. (2.107) gives the downward radiative flux:

$$F^- = \frac{1}{2}F^{\odot}_{\text{net,a}}\tau^* \tag{2.109}$$

The flux F^+_{ground} is $\sigma T^4_{\text{ground}}$, where T_{ground} is surface temperature:

$$\pi B(T_{\text{ground}}) = \sigma T^4_{\text{ground}} = F^{\odot}_{\text{net,a}}\left(1 + \frac{\tau_0{}^*}{2}\right) \tag{2.110}$$

If we apply eq. (2.110) to the Earth with $F^{\odot}_{\text{net,a}}$ ~239 W m^{-2} and T_{ground} ~288 K, the optical depth at the ground is $\tau_0 = \tau_0{}^*/1.66 = 0.8$. This value is consistent with other purely radiative models of Earth's total gray IR optical depth (Stephens and Tjemkes, 1993). However, a gray radiative–convective calculation gives a broadband IR optical depth ~2 for modern Earth (Robinson and Catling, 2012, 2014). The value of 0.8 is smaller in the purely radiative case of eq. (2.110) because we are using an observed value of ground temperature that, in fact, is cooler than a purely radiative model would predict because of heat transfer by tropospheric convection in the real atmosphere. If we apply eq. (2.110) to Venus, with $F^{\odot}_{\text{net,a}}$ ~163 W m^{-2} (lower than for Earth because of Venus' high albedo of 0.76) and $T_{\text{ground}} = 735$ K, the value of $\tau_0 = \tau_0{}^*/1.66 = 121$. This big optical depth indicates an extreme greenhouse effect at the bottom of Venus' thick atmosphere, but again a more realistic broadband optical depth would be even larger.

A characteristic of purely radiative equilibrium models is a temperature discontinuity between the ground and the atmosphere immediately above. In the lowest layer of the atmosphere, we have scaled optical depth $\tau_0{}^*$, so that eq. (2.106) gives

$$\pi B(T_{\text{a-ground}}) = \sigma T^4_{\text{a-ground}} = \frac{1}{2}F^{\odot}_{\text{net,a}}(1 + \tau_0{}^*) \tag{2.111}$$

Thus,

$$\pi B(T_{\text{ground}}) - \pi B(T_{\text{a-ground}}) = \frac{1}{2}F^{\odot}_{\text{net,a}}$$

or, in terms of temperature,

$$2\sigma\left(T^4_{\text{ground}} - T^4_{\text{a-ground}}\right) = F^{\odot}_{\text{net,a}} \tag{2.112}$$

As we described in Sec. 2.4.5.2, a temperature discontinuity is unphysical. Convection will ensure and dominate the temperature profile in the lower atmosphere.

Fluxes from the simple radiative model are shown in Fig. 2.19(a). The net thermal IR flux at the top of the atmosphere is $F^{\odot}_{\text{net,a}}$ upwards, showing equilibrium with absorbed incoming sunlight. The fact that there is a linear dependence of πB on scaled optical depth $\tau^* = 1.66\tau$, means that there is an exponential increase of πB as we go down through the atmosphere because of increasing absorber mass. Recall eq. (2.57), $d\tau = -k_{\text{abs}}\rho_a dz$, where k_{abs} and ρ_a are the absorption coefficient and absorber mass density. This relationship implies an isothermal stratosphere because the straight line of πB vs. τ converts into a curved line on a height vs. temperature graph.

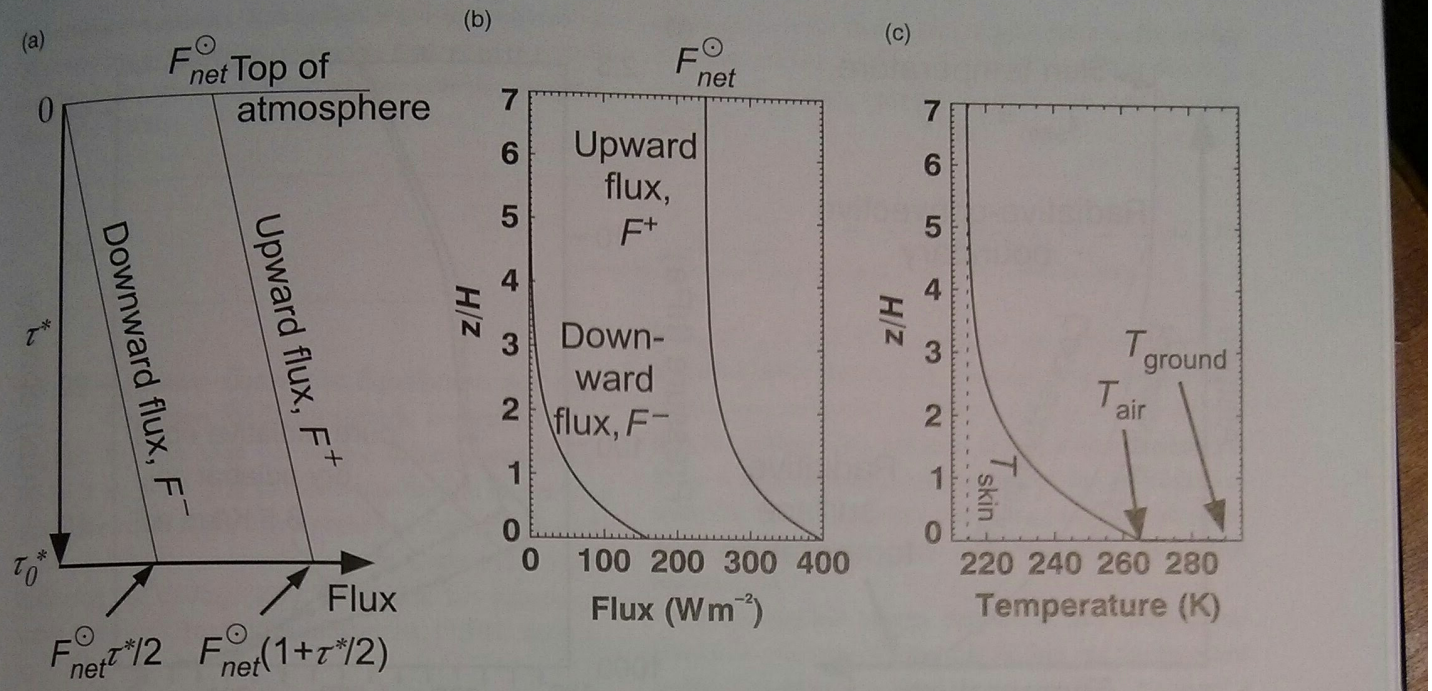


Figure 2.19 Fluxes and temperatures for the purely radiative model described in the text. (a) Downward irradiance F^- and upward irradiance F^+ as a function of scaled optical depth τ^*. (b) The irradiance plotted as a function of altitude z in units of scale height H. A total optical depth τ_0 of 0.8 is assumed, equivalent to a total scaled optical depth $\tau_0^* = D\tau_0 = 1.328$ for diffusivity parameter $D = 1.66$. The net absorbed solar flux is taken as $F^{\odot}_{net} = 239$ W m^{-2}. (c) The corresponding temperature profile. For an effective temperature $T_{eff} = 255$ K, the skin temperature of the upper atmosphere is $T_{skin} = 214.4$ K, the near-surface air temperature is $T_{air} = 264.9$ K, and the ground temperature is $T_{ground} = 289.6$ K.

We can plot fluxes and temperature in terms of altitude if we make the hydrostatic assumption and assume a value for the surface optical depth. We assume that k_{abs} is constant and that ρ_a varies with altitude z according to the hydrostatic equation as $\rho_a(z) = \rho_a(0)\exp(-z/H_a)$. Then the scaled optical depth will vary as $\tau^*(z) = \tau_0^* \exp(-z/H_a)$. Substituting for $\tau^*(z)$, the equations for the upward (2.108) and downward (2.109) radiative fluxes as a function of altitude are:

$$F^+(z) = \frac{1}{2}F^{\odot}_{net,a}\left(2 + \tau_0^* e^{-(z/H_a)}\right), \qquad F^-(z) = \frac{1}{2}F^{\odot}_{net,a}\tau_0^* e^{-(z/H_a)} \tag{2.113}$$

Similarly, the temperature profile and ground temperature from eqs. (2.106) and (2.110) are

$$T(z) = \left[\frac{F^{\odot}_{net,a}}{2\sigma}\left(1 + \tau_0^* e^{-(z/H_a)}\right)\right]^{1/4} = T_{skin}\left(1 + \tau_0^* e^{-(z/H_a)}\right)^{1/4},$$
$$T_{ground} = \left[\frac{F^{\odot}_{net,a}}{2\sigma}\left(2 + \tau_0^*\right)\right]^{1/4} = T_{skin}\left(2 + \tau_0^*\right)^{1/4} \tag{2.114a,b}$$

In eq. (2.114), we used the fact that $\left(F^{\odot}_{net,a}/2\sigma\right)^{0.25}$ is equivalent to $\left(\sigma T_{eff}^4/2\sigma\right)^{0.25} = T_{eff}/2^{0.25}$, the skin temperature, T_{skin}, given earlier in eq. (2.97). In eq. (2.114a), at the top of the atmosphere $\tau^* = 0$, so the temperature asymptotes to T_{skin}. Also, the effective temperature occurs at an emission level where $T(z) = \left(F^{\odot}_{net,a}/\sigma\right)^{0.25} = T_{eff}$, which is when $\tau^* = 1$. Radiative fluxes and temperature as a function of altitude are shown in Fig. 2.19, assuming a gray optical depth $\tau_0 = 0.8$, so that the scaled optical depth is $\tau_0^* = D\tau_0 \sim 1.3$. Given a net absorbed solar flux of 239 W m^{-2}, these values give a ~290 K ground temperature.

Further realism can be introduced by non-gray analytic radiative models (Parmentier and Guillot, 2014; Parmentier *et al.*, 2015), analytic radiative–convective models (Robinson and Catling, 2012), and finally numerical radiative-convective models, which were pioneered by Manabe and Strickler (1964). The basic difference between radiative and radiative-convective profiles can be seen in the conceptual diagram of Fig. 2.20(a). Figure 2.20(b) shows results from a numerical radiative–convective model for the Earth, which unlike Fig. 2.20(a) incorporates a stratospheric inversion due to ozone

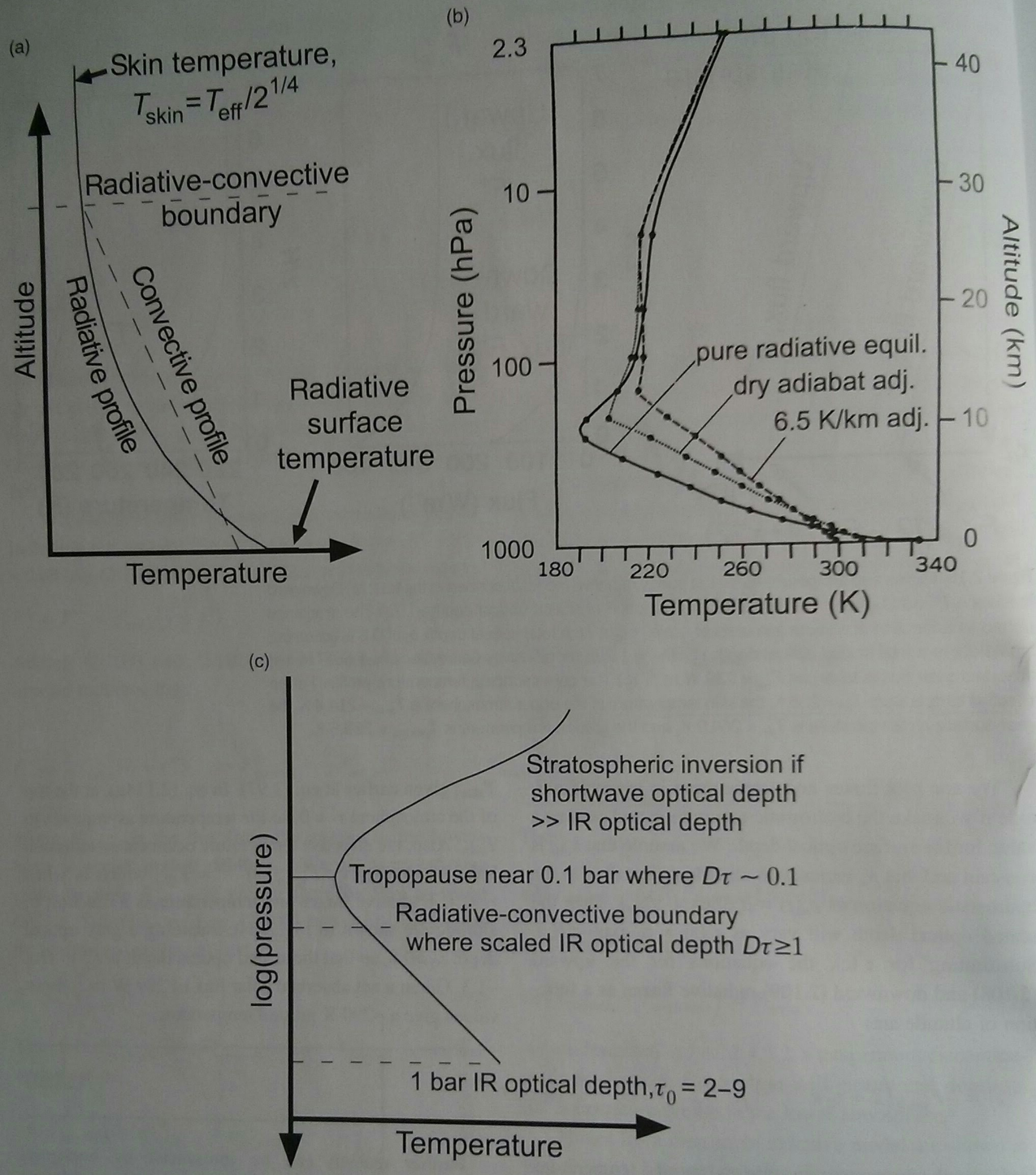


Figure 2.20 (a) A schematic radiative profile (solid) is unstable to convection, so turbulence ensues and an adiabatic, convective profile (dashed) is established. This intersects the radiative profile at a radiative–convective boundary and is close to the real atmospheric profile. In the upper stratosphere, the isothermal skin temperature, T_{skin} is a simple function of the effective temperature, T_{eff} of the planet. (b) Results from a radiative model for the Earth that includes the effects of stratospheric ozone causing an inversion above a tropopause minimum. (From Manabe and Strickler (1964). Reproduced with permission. Copyright 1964, American Meteorological Society.) (c) Schematic diagram following Robinson and Catling (2014), showing the typical thermal structure of dense atmospheres in the Solar System (Earth, Titan, and the giant planets) that have stratospheric inversions. Such atmospheres are optically thick in the IR at 1 bar, but higher up is a radiative–convective boundary where the scaled IR optical depth $D\tau$ is about unity (or greater for tropospheres that absorb significantly in the shortwave). Further up in height, a tropopause temperature minimum occurs at an optically thin IR optical depth near 0.1 bar pressure.

absorption. Appendix A describes a one-dimensional (1-D) numerical radiative–convective code that we have used in many studies of planetary atmospheres, including the early Earth.

2.4.5.4 *Radiative–Convective Equilibrium and the Condition for Stratospheric Inversions*

We can do better than our simple radiative equilibrium model if we join a gray radiative equilibrium model for a stratosphere and upper troposphere to a convective profile below, following treatments given by Catling (2015) and Robinson and Catling (2012, 2014). The key equation to solve is derived from differentiating eq. (2.104) w.r.t. τ,

$$\frac{d^2F_{\rm net}}{d\tau^2} = D\left(\underbrace{\frac{d(F^+ + F^-)}{d\tau}}_{DF_{\rm net}} - 2\pi\frac{dB}{d\tau}\right)$$
$$\Rightarrow \frac{d^2F_{\rm net}}{d\tau^2} - D^2F_{\rm net} = -2\pi D\frac{dB}{d\tau} \qquad (2.115)$$

Here, we have used eq. (2.100) for $F_{\rm net}$. The Stefan–Boltzmann Law gives the profiles of blackbody radiance $B(\tau)$ and temperature $T(\tau)$ from eq. (2.115) if $F_{net}(\tau)$ is known and if we use a top-of-atmosphere boundary condition on the net flux from eq. (2.102).

To consider the profile of net flux, $F_{\rm net}(\tau)$, we allow the atmosphere to absorb some of the incoming shortwave flux. Shortwave absorption can produce stratospheric temperature inversions, while absorption in tropospheres can stabilize upper tropospheres against convection, which is notable on Titan. Exponential attenuation is a reasonable approximation for stellar radiation, so the profile of net absorbed stellar flux $F^{\odot}_{\rm net}$ can be described as follows (Robinson and Catling, 2012):

$$F^{\odot}_{\rm net}(\tau) = \underbrace{F^{\odot}_{tropo}e^{-k_{tropo}\tau}}_{\text{tropospheric (and surface) solar absorption}} + \underbrace{F^{\odot}_{strato}e^{-k_{strato}\tau}}_{\text{stratospheric solar absorption}} \qquad (2.116)$$

Here, $F^{\odot}{}_{strato}$ and $F^{\odot}{}_{tropo}$ are the top-of-atmosphere net absorbed stellar fluxes in the stratosphere and troposphere, respectively, where $F^{\odot}{}_{tropo}$ also includes surface absorption for planets with surfaces. Dimensionless parameters $k_{strato} = \tau_{\rm sws}/\tau$ and $k_{tropo} = \tau_{\rm swt}/\tau$ control the attenuation of solar flux in two channels, where $\tau_{\rm sws}$ and $\tau_{\rm swt}$ are shortwave optical depths and τ is the gray thermal-IR optical depth.

A generalized energy balance at any atmospheric level is between the net thermal IR flux and the absorbed stellar flux plus any internal energy flux, F_i from a planet's interior (which is important on giant planets or certain tidally heated rocky exoplanets):

$$F_{\rm net}(\tau) = F^{\odot}_{\rm net}(\tau) + F_i \qquad (2.117)$$

By combining eq. (2.116) with (2.117) and inserting into eq. (2.115), we obtain

$$\frac{d\sigma T^4}{d\tau} = \frac{1}{2D}\left[\left(D^2 - k^2_{tropo}\right)F^{\odot}_{tropo}e^{-k_{tropo}\tau} + \left(D^2 - k^2_{strato}\right)F^{\odot}_{strato}e^{-k_{strato}\tau} + D^2F_i\right] \qquad (2.118)$$

which integrates to:

$$\sigma T^4(\tau) = \sigma T^4(0) + \frac{1}{2D}\left[\frac{D^2 - k^2_{tropo}}{k_{tropo}}\left(1 - e^{-k_{tropo}\tau}\right)F^{\odot}_{tropo} + \frac{D^2 - k^2_{strato}}{k_{strato}}\left(1 - e^{-k_{strato}\tau}\right)F^{\odot}_{strato} + D^2F_i\tau\right] \qquad (2.119)$$

The flux at the top of the atmosphere in this equation ($\sigma T^4(0)$) can be eliminated to derive a general expression in three steps. First, we use the definition of $F_{\rm net}$ (eq. (2.100) to rewrite eq. (2.104) as

$$\frac{dF_{\rm net}}{d\tau} = D\left(F_{\rm net} + 2F^- - 2\sigma T^4\right) \qquad (2.120)$$

Second, we solve for σT^4 using eqs. (2.116) and (2.117):

$$\sigma T^4(\tau) = F^-(\tau) + \frac{1}{2}\left[\frac{D + k_{tropo}}{D}F^{\odot}{}_{tropo}e^{-k_{tropo}\tau} + \frac{D + k_{strato}}{D}F^{\odot}{}_{strato}e^{-k_{strato}\tau} + F_i\right] \qquad (2.121)$$

Third, we obtain $\sigma T^4(0)$ by inserting $\tau = 0$ in eq. (2.121) and noting that $F^-(0)$ is zero. Inserting $\sigma T^4(0)$ into eq. (2.119) gives a final expression for the temperature

profile in the radiative part of a gray atmosphere that lies above a radiative–convective boundary:

$$\sigma T^4(\tau) = \frac{F^{\odot}_{strato}}{2}\left[1 + \frac{D}{k_{strato}} + \left(\frac{k_{strato}}{D} - \frac{D}{k_{strato}}\right)e^{-k_{strato}\tau}\right] + \frac{F^{\odot}_{tropo}}{2}\left[1 + \frac{D}{k_{tropo}} + \left(\frac{k_{tropo}}{D} - \frac{D}{k_{tropo}}\right)e^{-k_{tropo}\tau}\right] + \frac{F_i}{2}(1 + D\tau) \tag{2.122}$$

If there is no internal heat flux ($F_i = 0$) and no solar attenuation ($k_{strato} = k_{tropo} \rightarrow 0$), eq. (2.122) reduces to simple radiative equilibrium, eq. (2.106), using L'Hôpital's rule.

Equation (2.122) allows us to derive the condition under which an atmosphere has a pronounced tropopause temperature minimum and thus a stratospheric temperature inversion. We remove terms in tropospheric shortwave attenuation by taking $k_{tropo} << 1$, which is true for most real atmospheres. Then, setting the derivative w.r.t. τ of the resulting expression to zero, the optical depth at the tropopause minimum τ_{tp}, is (Robinson and Catling, 2014):

$$\tau_{tp} = \underbrace{\frac{1}{k_{strato}}}_{\sim 1/100} \underbrace{\ln\left[\underbrace{\frac{F^{\odot}_{strato}}{F^{\odot}_{tropo} + F_i}}_{\sim 1/10}\underbrace{\left(\frac{{k_{strato}}^2}{D^2} - 1\right)}_{\sim 10^3}\right]}_{\sim 5} \tag{2.123}$$

The numbers under parts of this equation are typical values in Solar System atmospheres, which yield a small gray IR optical depth at the tropopause temperature minimum of $\tau_{tp} \sim 0.05$. The tropopause minimum is at a much lower pressure than the radiative–convective boundary, which is where the optical depth of ~1 or greater, in the case of atmospheres with considerable tropospheric shortwave absorption such as Titan's. Furthermore, eq. (2.123) is only valid if

$$k^2_{strato} > D^2\left[1 + \frac{(F^{\odot}{}_{tropo} + F_i)}{F^{\odot}{}_{strato}}\right] \tag{2.124}$$

Equation (2.124) is the general condition for a temperature minimum and stratospheric inversion to exist. Remembering that $k_{strato} = \tau_{sws}/\tau$, the inequality indicates that the shortwave optical depth τ_{sws} in a stratosphere must be large compared to the thermal-IR optical depth, τ. In Solar System atmospheres with well-developed tropopause temperature minima, typical fluxes in eq. (2.124) give $k_{strato} \sim 10^2$.

The inequality of eq. (2.124) indicates the radiative properties that cause stratospheric temperature inversions in some atmospheres but not others. Earth's stratospheric ozone ensures that $k_{strato} \sim 90$, so shortwave heating is large relative to radiative cooling from CO_2, water vapor, and ozone. Consequently, Earth has a stratospheric inversion. On the giant planets and Titan, k_{strato} is $\sim 10^2$ because stratospheric aerosols absorb shortwave radiation strongly and methane absorbs near-IR solar radiation. Thermal-IR emission cools outer planet stratospheres inefficiently through a collision-induced absorption continuum of H_2–H_2 and H_2–He on the giant planets, and the bands of acetylene (C_2H_2, 13.7 μm), ethane (C_2H_6, 12.1 μm) and methane (7.6 μm) (Yelle *et al.*, 2001; Zhang *et al.*, 2013) on the giant planets and Titan. In contrast, CO_2 in the stratospheres of Mars and Venus has weak shortwave absorption compared to strong IR cooling from CO_2 emission. Consequently, the global mean temperature profiles of Mars and Venus lack stratospheric inversions.

To join a convective temperature profile to the radiative equilibrium profile described by eq. (2.122) as $T(\tau)$, we must convert optical depth to pressure or vice versa. Temperature as a function of pressure in the convective part of the troposphere is given by eq. (1.35). Combining the definition of differential optical IR depth $d\tau = -k_a \rho_a dz$ (eq. (2.59), where k_a is a gray mass absorption coefficient, ρ_a is the absorber mass density, and dz is the differential altitude) with hydrostatic equilibrium $dp/dz = -g\rho$ (where g is gravitational acceleration and ρ is atmospheric density), we get $d\tau \propto k_a dp$ if the absorber is well-mixed. Pressure-broadening and collision-induced absorption (see Sec. 2.5.7) cause a dependence where $k_a \propto p$ below middle stratospheres. Thus, integration of $d\tau \propto k_a dp \propto pdp$ gives $\tau \propto p^{n_p}$ where $n_p = 2$. More generally, a gray IR optical depth is often approximated by a power law relationship with pressure,

$$\tau = \tau_0\left(\frac{p}{p_{ref}}\right)^{n_p} \tag{2.125}$$

where τ_0 is the optical depth at reference pressure p_{ref}, such as 1 bar. As noted, $n_p = 2$ in a troposphere and lower stratosphere applies for a well-mixed IR absorber. The terrestrial atmosphere is more complex in detail: the 8–12 μm IR window, $n_p = 2$ scaling for the 15 μm CO_2 band, and $n_p = 4$–5 scaling for water bands at 6.3 μm and

beyond ~20 μm. However, calculations show that $n_p = 2$ is a good overall gray approximation for the tropospheres and lower stratospheres of Earth, Titan, and the giant planets (Robinson and Catling, 2012, 2014).

A radiative–convective temperature profile is obtained using the scaling relationship between pressure and optical depth (eq. (2.125)). The τ–p scaling is inserted into the equation for the T–p profile in the convective part of the troposphere (eq. (1.35)) to give a temperature profile in terms of optical depth rather than pressure. In turn, an analytic expression for the convective upwelling flux can be derived, which gives an analytic solution to radiative–convective equilibrium (Robinson and Catling, 2012). The temperature in an atmosphere follows a convective profile up to a radiative–convective boundary. Above the boundary, the radiative temperature profile of eq. (2.122) applies. The treatment can be extended to include scattering (Heng *et al.*, 2014).

The physics just described allows us to understand why dense atmospheres in the Solar System with stratospheric inversions tend to have global mean tropopause temperature minima within a factor of ~2 of 0.1 bar pressure (Fig. 1.1, Fig. 2.20(c)). If an atmosphere satisfies the temperature minimum condition (eq. (2.124)) then the IR optical depth at the minimum is close to $\tau_{tp} \sim 0.05$. The dense atmospheres of Earth, Venus, the giant planets, and Titan are all optically thick at depth, with an average IR optical depth at 1 bar within a few units of $\tau_0 \sim 5$, despite the great diversity of compositions. From eq. (2.125), the tropopause pressure is $p_{tp} \approx (0.05/5)^{0.5} = 0.1$ bar. This commonality is surprising because gravity ought to affect optical depth (c.f. eq. (2.98)). However, the radiative properties of gases and the shared τ–p scaling caused by common broadening processes turn out to dominate. Also, the different greenhouse effects of the aforementioned bodies, represented by a gray optical depth τ_0 at 1 bar, ranging ~2–10, have a weak influence on tropopause pressure because the latter decreases only as the inverse square root of τ_0, i.e., $p_{tp} \propto \tau_0^{-1/2}$ from eq. (2.125).

Finally, we note a word of caution: the total gray optical depth is not a real physical property of atmospheres but merely a fitting parameter for a gray model to real observations, such as surface temperature and absorbed solar flux. Broadband optical depths should be viewed as semi-quantitative measures that give us greater understanding of planetary climates. Thus, it is usually never worthwhile laboriously computing a gray optical depth τ_0 from wavelength-dependent optical depths. Instead, τ_0 should be fit.

2.4.5.5 *Runaway Greenhouse: a Limit on Outgoing Longwave Radiation*

The dependence of surface temperature in radiative equilibrium (e.g., eq. (2.110)) on the optical depth can lead to an unstable climate. For wet, Earth-like planets, the optical depth depends on the concentration of water vapor, which is the major greenhouse gas. The connection to water vapor is behind the possibility of a *runaway greenhouse effect*, which is an upper limit on outgoing infrared flux F^+ at the top of the atmosphere that can be developed by a planet with liquid water on its surface.

To understand the runaway greenhouse concept, consider a saturated atmosphere on a warmer Earth. The surface pressure will be proportional to the partial pressure of water, p_{H2O}, which is an exponential function of temperature. In turn, the infrared optical depth, τ_{IR}, is proportional to the column mass or surface pressure via eq. (1.19). Consequently, τ_{IR} is proportional to an exponential function of temperature via the Clausius–Clapeyron equation (1.49). If the solar flux is large enough to produce a sufficiently warm surface, τ_{IR} becomes so high that the atmosphere becomes essentially completely opaque to outgoing thermal-IR radiation across all wavelengths. Earth's 8–12 μm window closes. The atmosphere then radiates to space at a level where $\tau_{IR} \sim 1$, analogous to a star's photosphere except in the thermal-IR. Essentially no thermal-IR fluxes to space directly from the surface. It turns out that if the atmosphere is close to saturation with water vapor, the outgoing flux at the top of the atmosphere F^+ reaches a limit defined by the properties of water and the gravity of the planet, which is $F^+_{limit} \sim 300$ W m^{-2} for Earth. If the solar flux exceeds F^+_{limit} then the excess incoming energy that cannot be radiated away is used to evaporate the oceans inexorably. Once the oceans are all evaporated, the surface melts. Energy balance is achieved when a hot upper atmosphere radiates to space through near-IR windows, which are wavelengths where water vapor is relatively transparent. This *runaway greenhouse effect* is the most widely accepted explanation of the climate history of Venus (see Ch. 13).

2.5 Absorption and Emission of Radiation by Atmospheric Gases

In Sec. 2.4, we described how to calculate the transmission of UV, visible, and IR radiation through atmospheres, given appropriate absorption and scattering coefficients, but we didn't describe *why* certain gases absorb or scatter at particular wavelengths. We now

consider this issue, and how absorption and emission depend upon pressure and temperature. In doing so, we address the question of why gases such as N_2 and O_2 in the Earth's atmosphere are not greenhouse gases, whereas others are, such as H_2O, CO_2, O_3, and CH_4. We also examine how on giant planets, H_2O and CH_4 also absorb in the IR, but there is a predominance of IR absorbers of hydrogen-containing gases such as H_2 itself, NH_3 and C_nH_n.

2.5.1 Overview of Absorption Lines

To understand how gases absorb, we start with the concept that electromagnetic radiation is quantized into elementary particles, *photons*, each with an energy given by Planck's Law, eq. (2.41). A gas molecule can absorb a photon if the energy of the photon is close to a quantized energy jump within the molecule that is needed for vibration, an increase in rotational speed, or the redistribution of charge when an electron leaps to a higher orbit. Consequently, we write the overall internal energy of a gas molecule as:

$$E_{\text{total}} = E_{\text{electronic}} + E_{\text{vibrational}} + E_{\text{rotational}} + E_{\text{translational}} \tag{2.126}$$

Here, $E_{\text{electronic}}$ is the energy associated with electron orbits, and the other terms refer to vibrational energy, rotational kinetic energy, and translational kinetic energy, respectively.

Collisions between molecules, which depend on $E_{\text{translational}}$ and therefore temperature, redistribute energy between the various forms. As discussed in Sec. 2.4.2.4, if collisions are sufficiently rapid compared to the radiative decay time for the molecules, energy exchanges between the forms in eq. (2.126) and excess vibrational or rotational energy is converted into heat. The heating establishes *Local Thermodynamic Equilibrium* (LTE) where a local volume of the atmosphere emits and absorbs like a blackbody with a characteristic brightness temperature. This is generally a good assumption apart from at high altitude (e.g., Earth's mesosphere), where LTE breaks down because of infrequent collisions.

In general, electronic transitions require more photon energy than vibrational ones, which in turn demand more energy than changes in rotational speed. For example, a single quantum of energy between two allowed rotational frequencies of a molecule tends to be relatively small and so would correspond by Planck's Law (eq. (2.41)) to the energy of a low frequency photon in the IR or microwave part of the spectrum. The correspondence between wavelengths and types of transition is as follows.

Dominant transition	Wavelength range	Band
Electronic	< 1 μm	X-ray to UV to visible
Vibration	1–20 μm	Near-IR to far-IR
Rotation	>20 μm	Far-IR to microwave*

* Except H_2 in warm, dense atmospheres, which can have rotational transitions at shorter wavelengths.

The changes in states of molecules that give rise to absorption lines often involve combinations of transitions, such as a simultaneous low-energy rotational transition and a higher energy vibrational or electronic transition. This produces considerable structure in the absorption line spectrum of atmospheres. For quantum mechanical details, the reader is referred to Ch. 13 of McQuarrie and Simon (1997).

When we examine the spectrum of radiation from a planetary atmosphere, we see discrete lines, groups of neighboring lines called *bands*, and smoother varying absorption called the *continuum* (Fig. 2.21), all of which we consider below. Spectral intervals where atmospheric absorption is very weak are termed *windows*.

Various windows exist in Earth's atmospheric spectrum can affect the climate or are useful for astronomical observations. The most well-known window is at 12.5–8 μm where much thermal infrared radiation fluxes to space. Astronomers using ground-based telescopes observe through atmospheric windows to characterize celestial objects with infrared photometry (Bessell, 2005). Their J, H, K, L, and M pass bands are at 1.2, 1.6, 2.2, 3.5, and 4.8 μm, respectively, while N and Q bands are at longer wavelengths of 10.5 and 21 μm, corresponding to relatively transparent atmospheric transmission.

2.5.2 Electric and Magnetic Dipole Moments

The ability of a gas molecule to absorb radiation depends on the electric or magnetic field distribution across the molecule. Whether gas molecules have a *dipole moment* determines if they undergo rotational or vibrational transitions, which are important in the greenhouse effect and spectroscopy of planetary atmospheres. A dipole moment comes in two kinds: electric and magnetic. A molecule with a distribution of positive and negative charges $\pm Q_c$ separated by distance l has an electric dipole moment, $p_d = Q_c l$ in units of coulomb·meter. For example, an H_2O vapor molecule has a slight positive charge on its hydrogen atoms and slight negative charge on the oxygen atom, and possesses a permanent electric dipole moment

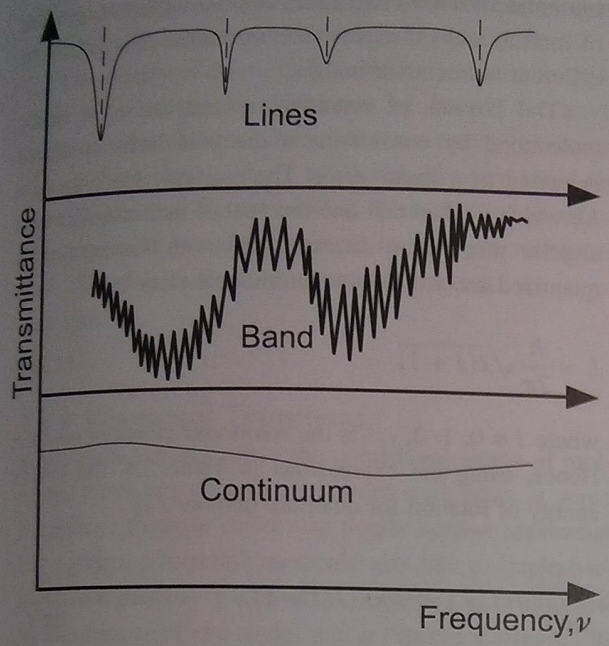


Figure 2.21 Schematic diagram of absorption lines, a band, and a continuum. Note: lines occur at a smaller frequency scale and make up the bands.

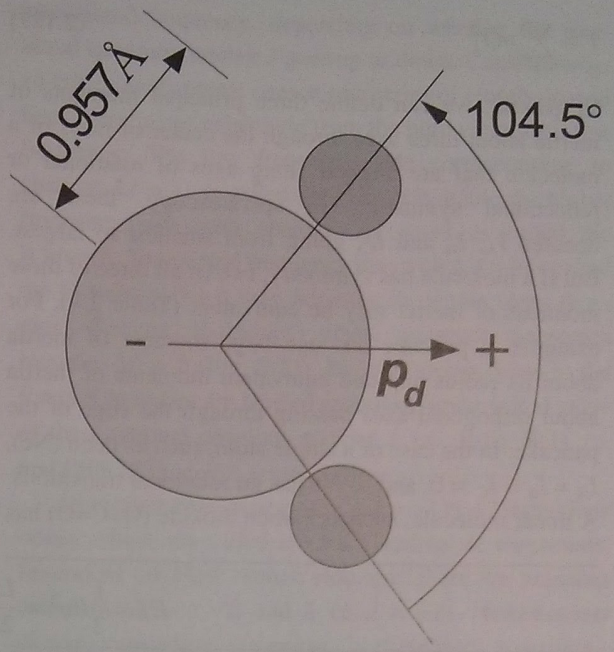


Figure 2.22 The electric dipole moment vector p_d of a water vapor molecule, H_2O.

of 6.1×10^{-30} C·m (Fig. 2.22). An electric field vector $\boldsymbol{E}$ will exert a torque on an electric dipole, given by $\boldsymbol{p_d} \times \boldsymbol{E}$, which tends to align the molecule's dipole with the field. Thus, electromagnetic radiation can induce rotation in molecules with electric dipoles.

Molecular oxygen O_2 possesses a magnetic dipole moment. O_2 does not possess an electric dipole moment because its atoms are identical with no separation of charge, so the electric field in an electromagnetic wave does not interact with an isolated molecule. Homoatomic molecules such as O_2, N_2, and H_2 *can* develop transient induced electric dipoles through collisions that distort their electron clouds, allowing them to interact weakly with the electric field of passing electromagnetic waves, but we consider those later in Sec. 2.5.6. The magnetic field in an electromagnetic wave can torque a molecule with a magnetic dipole moment into rotation. An analogy is the needle of a compass, which possesses a magnetic dipole moment and so points northwards when torqued by the Earth's magnetic field. The magnetic dipole moment is defined by $\boldsymbol{\mu} = I_c \boldsymbol{A}$ in units of A.m^2, where I_c is current and $\boldsymbol{A}$ is the vector area of a loop. A magnetic field vector $\boldsymbol{B}$ will exert a torque on a magnetic dipole, given by $\boldsymbol{\mu} \times \boldsymbol{B}$, which is directly analogous to the electric dipole case.

Why do some molecules have magnetic dipoles? Electrons orbiting an atom classically define a current in a loop. But individual electrons pair up with opposite spins, causing no net magnetic moment for a molecule as a whole. Molecular oxygen O_2, however, has two unpaired electrons. The resulting magnetic dipole allows O_2 to have rotational absorption bands in the microwave at ~60 and 118 GHz. These absorptions are used in satellite observations of temperatures in Earth's atmosphere (Janssen, 1993).

Other common gases may or may not have dipoles. For example, N_2 has neither an electric nor magnetic dipole and so has no pure rotational spectrum, except through collisions (Sec. 2.5.6). Although CH_4 and CO_2 also have no electric or significant magnetic dipoles, their bonds bend, which breaks molecular symmetry and allows a temporary electric dipole with associated rotation–vibrational transitions. We now examine these transitions in more detail.

2.5.3 Rotational Transitions

Rotational transitions occur because molecules spin like tumbling acrobats with a frequency that is the number of revolutions per second. A rotational transition occurs when a photon of appropriate energy is absorbed and the rotational frequency of a molecule is boosted in a quantized way, similar to how a household fan only spins at certain speeds. We describe rotational motion with equations that are analogous to linear motion (Table 2.3).

How quickly a molecule spins depends on its moment of inertia I. This quantity is just the sum of the products of elemental mass m_i and the square of the radial distance r_i from the axis of rotation:

$$I = \sum_i m_i r_i^2 \tag{2.127}$$

In general, we can define three principal moments of inertia about three axes through the center of mass of a molecule that are aligned along axes of rotational or reflectional symmetry. In spectroscopy, these are denoted I_A, I_B and I_C, going from smallest to largest. But if a molecule has symmetry, two or all three of these moments of inertia may be equivalent (Table 2.4). For example, a pancake has one large moment of inertia about its radius and two equivalent moments of inertia about orthogonal axes passing through the edge of the pancake. In the case of a single atom, such as neon (Ne), $I_A = I_B = I_C \approx 0$, and so Ne has no rotational transitions. A linear molecule, such as carbon dioxide (O=C=O) has an axis that passes through the nuclei, where $I_A \approx 0$, while the moments of inertia around the other perpendicular axes are equal and non-zero, $I_B = I_C > 0$. It is an example of a *linear rotor*. For nonlinear polyatomic molecules, three categories occur: (1) *spherical top* molecules with all moments of inertia equal, (2) *symmetric top* molecules with two equal and one different moment of inertia, and (3) *asymmetric top* molecules with three different moments of inertia.

The physics of rotational spectra are most easily understood by considering a diatomic linear molecule behaving as a *linear rotor*. The molecule has only $I_B = I_C$, which we can call one moment of inertia I. The total angular momentum L, associated with the rotation is quantized according to quantum mechanics by

$$L = \frac{h}{2\pi}\sqrt{J(J+1)} \tag{2.128}$$

where $J = 0, 1, 2, \ldots$ is the *rotational quantum number*. Hence, using the expressions in Table 2.4, the kinetic energy of rotation for quantum number J is:

$$E_J = \frac{1}{2}I\omega^2 = \frac{L^2}{2I} = \frac{h^2}{8\pi^2 I}(J(J+1)) = hB(J(J+1)), \qquad \text{where } B = \frac{h}{8\pi^2 I}$$

Quantity B is called the *rotational constant*. (Sometimes the rotational constant is given as $h/(8\pi^2 Ic)$ when working in wavenumbers, i.e. units of cm^{-1}). Only rotational transitions between adjacent quantum states J and $J + 1$ are allowed, so the energy change in a transition is

$$\Delta E = hB[(J+1)(J+2) - J(J+1)] = hB\left[J^2 + 3J + 2 - J^2 - J\right] = 2hB(J+1)$$

Table 2.3 Equations of linear and rotational motion. Here m is mass, F is force, a is acceleration, v is velocity, T is torque, ω is angular velocity, and I is moment of inertia.

	Linear	Rotational or Angular
Force and acceleration	Force, $F = ma$	Torque, $T = I\, d\omega/dt$
Kinetic energy	$\frac{1}{2}\, mv^2$	$\frac{1}{2}\, I\omega^2$
Momentum	mv	$I\omega$

Consequently, the frequency of a photon for a rotational transition in a linear rotor is

$$\nu = \frac{\Delta E}{h} = 2B(J+1) \tag{2.129}$$

The pure rotational absorption spectrum is therefore a series of lines uniformly separated by frequency differences of $\Delta\nu = 2B$. However, in reality, as a molecule rotates faster with increasing J, centrifugal force causes the bond to stretch slightly and the frequency spacing to

Table 2.4 Rotational symmetry and moments of inertia in common molecules present in atmospheres of giant and terrestrial planets.

Molecular type	Moment of inertia	Example gases
Monoatomic	$I_A = I_B = I_C \approx 0$	Ar, He, Ne
Linear (or linear rotor)	$I_A \approx 0$, $I_B = I_C > 0$	O_2, N_2, CO_2, CO, N_2O, C_2H_2
Spherical top (or spherical rotor)	$I_A = I_B = I_C > 0$	CH_4, GeH_4
Symmetric top (or symmetric rotor)	$I_A \neq 0 < I_B = I_C$ Or $I_C > I_A = I_B$	C_2H_6, C_2H_4
Asymmetric top	$I_A \neq I_B \neq I_C$	NH_3, PH_3, CH_3Cl H_2O, O_3, H_2S, SO_2, C_3H_8

change. A higher-order "centrifugal distortion" correction to eq. (2.129) can account for this behavior.

Our treatment above only considers a diatomic molecule but similar expressions can be derived for more complicated molecules. These have n moments of inertia I_n each associated with J_n rotational quantum numbers (McQuarrie and Simon, 1997). The absorption of a photon can also cause simultaneous changes rotational and vibrational quantum numbers, which we now consider.

2.5.4 Vibrational Transitions

A molecule vibrates when the constituent atoms of molecules move toward and away from each other at some frequency. Because molecular bonds behave somewhat like springs, a diatomic molecule acts like a simple harmonic oscillator with a resonant frequency ν_0 determined by the masses of the atoms and a "spring constant." In quantum mechanics, however, the actual oscillation frequency is quantized according to

$$\nu = \left(\mathrm{v} + \frac{1}{2}\right)\nu_0$$

where "v" (the letter vee) is a *vibrational quantum number*, v = 0, 1, 2,. . .. Consequently, the allowed energy levels are separated by an integer times the energy at the classical resonant frequency:

$$E_\mathrm{V} = h\nu_0\left(\mathrm{v} + \frac{1}{2}\right)$$

The lowest vibrational state (v = 0) has a finite energy of E= $h\nu_0/2$ rather than zero. This *zero point energy* is consistent with Heisenberg's uncertainty principle because if the oscillating particle were stationary the uncertainty on its position would be zero, giving it infinitely uncertain momentum, which cannot be so. The transition from the lowest energy state to the next highest, or vice versa, is known as the *fundamental*. This typically produces the strongest line for statistical reasons: at typical planetary temperatures and pressures near STP, most of the molecules of a gas are in the ground state. Usually, the quantum mechanical rule is that only transitions of $\Delta\mathrm{v} = \pm 1$ are allowed. However, large displacements of atoms in molecules in so-called anharmonic oscillations give rise to cases where $\Delta\mathrm{v} = \pm 2, \pm 3$, etc., which produce *overtone* bands that are less intense than the fundamental absorption.

Because vibrational and rotational transitions often occur together, the combined transition frequency can be slightly less or greater than the pure vibrational transitional frequency, depending on whether the rotational quantum number J goes up or down. Consequently, *ro-vibrational* spectra consist of a series of closely spaced lines fanning out either side from the pure vibrational line frequency. The low frequency side corresponding to changes $\Delta J = -1$ is called the P branch, while the higher frequency side with changes $\Delta J = +1$ is called the R branch. The central frequency line with $\Delta J = 0$ is called the Q branch, but it may not exist in certain cases (e.g., most diatomic molecules) where quantum mechanics requires $\Delta J \neq 0$ (Fig. 2.23). By convention, lines in the P and R branches are labeled according to the initial value of the rotational quantum number J, i.e., R(0), R(1),. . ., and P(0), P(1),. . ..

In another case, superposed $\Delta J = -2$ or +2 transitions occur, which are called *electric quadrupole transitions*. Instead of the PQR branch structure, there are branches called O ($\Delta J = -2$) and S ($\Delta J = +2$). Homonuclear diatomic molecules can engage in these weak transitions, despite their lack of a permanent electric dipole. Giant planet atmospheres have electric quadrupole absorption lines of H_2 in the far IR, for example.

Molecules with more than two atoms have a large variety of vibrations, and the theory for their vibrations concerns the superposition of normal modes. These

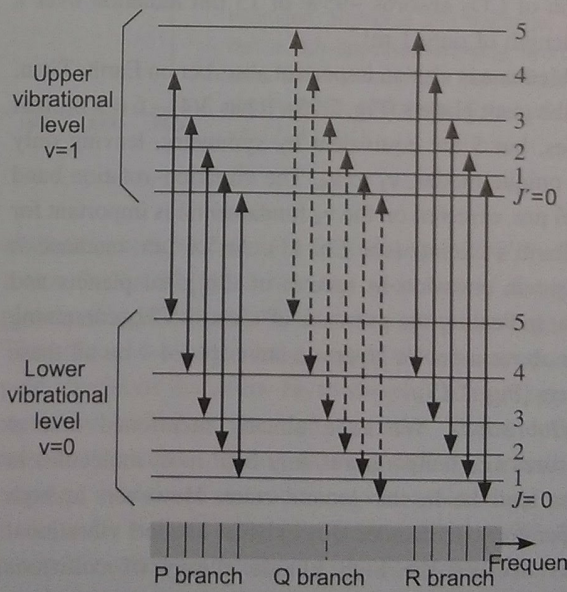


Figure 2.23 The meaning of the P, Q, and R branches in vibrational–rotational transitions, showing a vibrational transition $\Delta\mathrm{v}$ and superposed rotational transitions, ΔJ. The P branch corresponds to $\Delta\mathrm{v} = 1$ with rotational transitions $\Delta J = -1$. The Q branch corresponds to $\Delta\mathrm{v} = 1$ and $\Delta J = 0$ and the R branch corresponds to $\Delta\mathrm{v} = 1$ and $\Delta J = +1$. The lower shaded panel shows the appearance of the lines in the spectrum schematically, noting that the Q branch is offset from P and R branches for clarity in order to show the Q-branch ΔJ transitions.

multiple modes require more vibrational quantum numbers to describe all the energy levels than in a diatomic molecule, but the number of modes is given by some simple rules. A linear molecule consisting of N atoms has $3N - 5$ normal modes (or degrees of freedom), whereas a nonlinear molecule of N atoms has $3N - 6$ normal modes. Examples of normal modes of vibration of some common gas molecules are shown in Fig. 2.24. In order for the absorption of a photon to excite a mode, the dipole moment must vary during the normal mode motion. When this occurs for IR photons, the mode is said to be *infrared active*. Otherwise, the mode is *infrared inactive*.

The vibrational absorption bands of CO_2 are of great importance for the atmospheres and climates of Venus, Earth, and Mars. CO_2 has $3 \times 3 - 5 = 4$ normal vibration modes shown in Fig. 2.24. There is no change in the dipole moment during the symmetric stretch ν_1, so this mode is infrared inactive. Two bending modes are equivalent and are treated as a single mode denoted ν_2. In fact, the 15 μm rotation–vibration band is centered on the ν_2 fundamental. This is actually a PQR branch structure but individual lines are indistinct in the spectra of Venus, Earth, or Mars because of pressure broadening (see below). The center of the 15 μm band is remarkably absorbing. On Earth at 1 bar, air with a typical concentration of CO_2 absorbs ~95% of 15 μm radiation over a pathlength of only 1 m!

Methane is also an important absorber on Earth, Titan, and the giant planets (Fig. 2.25). It has $3(4) - 6 = 9$ normal modes, but 5 are equivalent by symmetry, leaving only four unique modes, ν_1 to ν_4. The vibration–rotation band at 7.6 μm, centered on the ν_4 fundamental is important for the Earth's climate (see Ch. 11). At 7.6 μm, methane is evident in emission in spectra of the giant planets and Titan, indicating the presence of a warm CH_4-containing layer above a cooler layer – a stratosphere – on all these planets (Fig. 2.25).

Hot bands. We have already mentioned that at pressures and temperatures near STP most molecules in a gas tend be in the ground state. However, at high temperatures molecules can exist in excited vibrational states because the high kinetic energy of collisions excites vibrations. "Hot bands" refer to transitions where the vibrational state does not involve the ground state. Hot bands are also important in gases at low pressure, where collisions are infrequent and there is a departure from Local Thermodynamic Equilibrium, or non-LTE, which allows higher vibrational bands to become densely populated, as reviewed by López-Puertas and Taylor (2001).

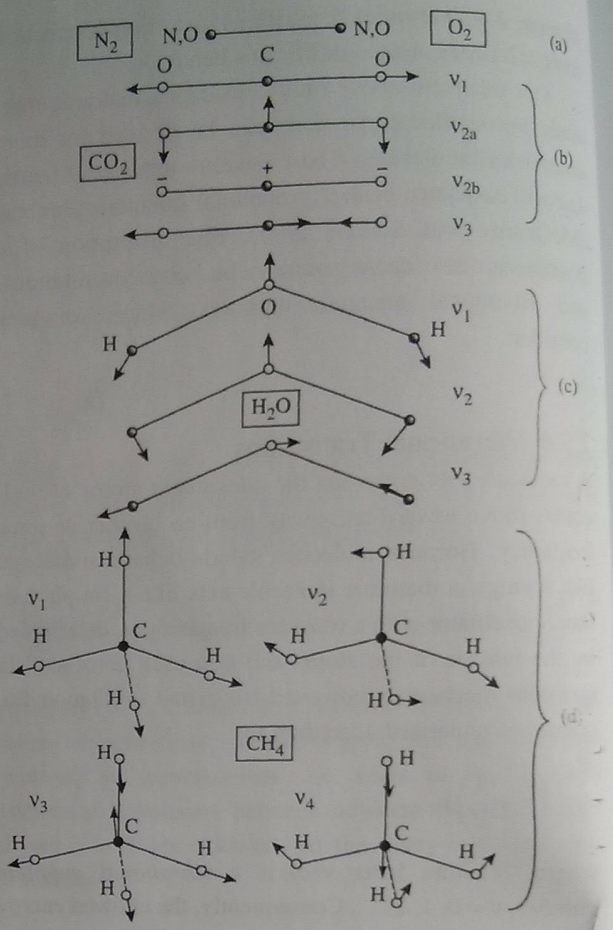


Figure 2.24 Normal vibrational modes of some common gas molecules. The arrows indicate a half-cycle direction of motion, which is followed by movement in the opposite direction to complete the full cycle. Signs +, – indicate motion out of and into the page. (a) Homonuclear molecules N_2 and O_2 have a single stretching mode of vibration. (b) CO_2 has a symmetric stretch (ν_1), which is optically inactive, and an asymmetric stretch ν_3 with a fundamental at 4.3 μm. The bending mode ν_2, with a fundamental at 15 μm, has two equivalent (or *degenerate*) modes 2a and 2b through a 90° rotation. N_2O, which is "NNO" structurally (not shown), behaves similarly to CO_2 with ν_1 to ν_3 modes at fundamentals of 7.8, 17.0, and 4.5 μm, respectively. (c) H_2O has a bending mode ν_2, with a fundamental at 6.3 μm. Water vapor's symmetric stretch ν_1 (fundamental at 2.74 μm) and asymmetric stretch ν_3 (fundamental at 2.66 μm) contribute to near infrared absorption. O_3 (not shown) behaves similarly to H_2O with ν_1 to ν_3 modes at fundamentals of 9.01, 14.3, and 9.6 μm, respectively. (d) CH_4 has four distinct bending modes, ν_1 to ν_4, but only ν_3 and ν_4 are infrared active. Modes ν_3 (with a fundamental at 3.3 μm) and ν_4 (with a fundamental at 7.6 μm) are important bands in spectra of planetary atmospheres (Modified from Thomas and Stamnes (1999).)

2.5.5 Electronic Transitions

Electronic transitions concern the energy levels of electron orbits around the nucleus of an atom. There are larger energy differences between these levels than in vibrational transitions, so electronic transitions of outer

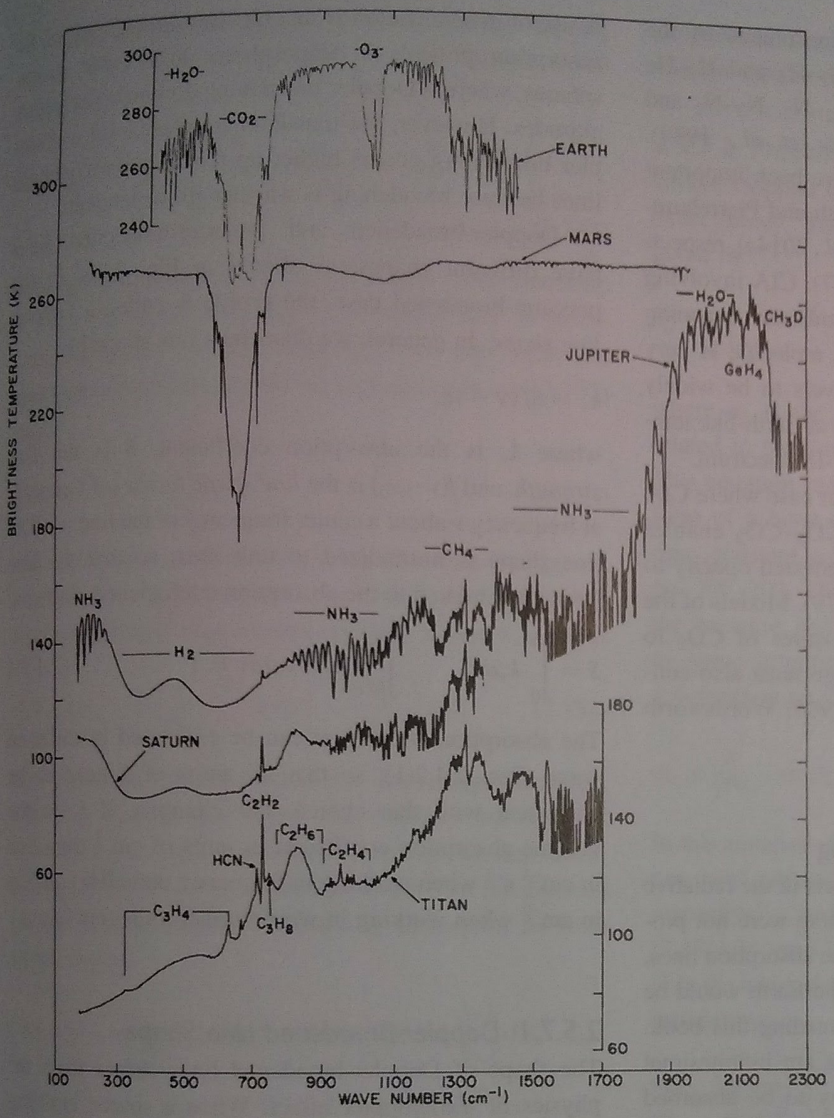


Figure 2.25 Spectra of the Earth, Mars, Jupiter, Saturn, and Titan from the Infrared Interferometer Spectrometer (IRIS) instruments on Nimbus 4, Mariner 9, and Voyager 1 and 2 spacecraft. (From Hanel (1981), Courtesy of NASA.) Carbon dioxide at 15 μm (667 cm^{-1}) is a common feature of the atmospheres of Earth, Mars, and Venus (not shown), but only the Earth has prominent ozone absorption at 9.6 μm (1043 cm^{-1}) and abundant water bands. The giant planets, including Uranus and Neptune, which are not shown, have absorption bands due to hydrogen, ammonia and methane. Titan has a rich spectrum showing hydrocarbons and nitriles. Note the emission of stratospheric acetylene, C_2H_2, at 13.7 μm (729 cm^{-1}).

electrons are associated with absorption and emission at shorter wavelengths, from the near-IR to the UV. For inner electrons, which are tightly bound to the nucleus, absorption and emission occur in the x-rays. Superposed upon the electronic energy levels are the smaller vibrational transitional levels, which in turn have superposed rotational energy levels, as discussed earlier. Consequently, each electronic transition appears as a series of closely spaced lines. All molecules, including those lacking a permanent dipole, such as N_2, have electronic spectra because a dipole moment change always accompanies a change in charge distribution.

2.5.6 Collision-Induced Absorption: Giant Planets, Titan, Early Earth, and Venus

Homonuclear molecules such as H_2 or N_2 lack a permanent electric dipole moment but they still interact with electromagnetic radiation. In fact, absorption by H_2 dominates the far IR (10–1000 μm) spectrum of the giant planets (Fig. 2.25) and H_2 is a key greenhouse gas on Titan (Sec. 14.4.2). We have already met electric quadrupole transitions that allow H_2 to absorb. Of more importance are collisions at high pressure that induce a temporary electric dipole, allowing radiative transitions amongst vibrational and rotational states that would never occur in the isolated molecule (Trafton, 1966, 1998). This process is called *collision-induced absorption* (CIA). It is also called pressure-induced absorption, but this term has the potential for confusion with pressure broadening, so we avoid it. Although the absorption is weak, broad absorption features occur when the gas is abundant and there is a long pathlength. Minor CIA of N_2 and O_2 occurs in the Earth's present atmosphere with radiative effect of little importance (Farmer and Houghton, 1966), but in other atmospheres CIA dominates absorption.

CIA is important in reducing atmospheres of the present and past. CIA results from H_2–H_2 and H_2–He collisions on the giant planets and H_2–N_2, N_2–N_2 and N_2–CH_4 collisions on Titan (McKay *et al.*, 1989). H_2–N_2 and H_2–CO_2 collisions may have been important in warming prebiotic Earth (Wordsworth and Pierrehumbert, 2013) and early Mars (Ramirez *et al.*, 2014a), respectively, as first suggested by Sagan (1977). CIA involving H_2 is particularly effective for greenhouse warming because the moment of inertia of this molecule is very small, causing its rotational energy levels to be widely spaced (see eq. (2.129)). Consequently, at Earth-like temperatures H_2 absorbs across the thermal-IR spectrum.

Dense CO_2 atmospheres are another case where CIA is important. On Venus, CIA from CO_2–CO_2 enables transitions that contribute significant infrared opacity to the atmosphere (Moskalenko *et al.*, 1979). Models of the early Earth that use high partial pressures of CO_2 to offset a fainter Sun ~4 billion years ago must also consider CIA (Kasting and Ackerman, 1986; Wordsworth *et al.*, 2010).

2.5.7 Line Shapes and Broadening

Broadening processes are essential effects in the radiative transfer of planetary atmospheres. If there were not processes in atmospheres that broadened the absorption lines, there would be no greenhouse effect, the Earth would be uninhabitable, and you would not be reading this book. Absorption lines at precise frequencies are infinitesimal delta functions but in order for energy to be absorbed there must be absorption over a wavelength range. The three line-broadening processes are as follows.

- *Natural broadening*: Heisenberg's uncertainty principle expressed in terms of energy (E) and time (t), states that $\Delta E\, \Delta t \geq \hbar/2$ where $\hbar = h/2\pi$. Excited states have a natural lifetime τ_D with respect to decay or τ_C between collisions, so that energy levels must be spread out in energy by at least $\Delta E \sim (\hbar/2\tau_D)$ or $(\hbar/2\tau_C)$. Using $\Delta E = h\Delta\nu$, this translates to a spread of frequency of $\Delta\nu \sim 1/4\pi\tau_D$ or $\sim 1/4\pi\tau_C$. However, apart from in the UV, this broadening is miniscule compared to the other two processes described below so we will not discuss it further.
- *Doppler broadening:* Gas molecules move randomly relative to the source of radiation. Such motion produces absorption and emission at frequencies that are Doppler-shifted compared with the rest position of a spectral line.
- *Pressure* (or *collisional) broadening:* Collisions between molecules provide or remove energy during radiative transitions so that emission and absorption occur over a broader wavelength range.

A useful generalization is that pressure broadening of IR absorption prevails in tropospheres and lower stratospheres, whereas Doppler broadening dominates at higher altitudes. However, the transition from pressure to Doppler broadening occurs higher up for longer wavelength lines because broadening is wavelength-dependent.

Doppler-broadened and pressure-broadened lines have different shapes, as shown in Fig. 2.26. In the pressure-broadened case, the profile is called a *Lorentz line shape*. In general, we describe a line shape by

$$k_\nu = Sf(\nu - \nu_0) \tag{2.130}$$

where k_ν is the absorption coefficient, S is the *line strength*, and $f(\nu - \nu_0)$ is the *line shape factor* (or *function*) at frequency ν about a center frequency of the line, ν_0. The line shape is normalized to unit area, so that the line strength is related to the absorption coefficient as follows:

$$S = \int_0^\infty k_\nu d\nu \qquad \int_0^\infty f(\nu - \nu_0)d\nu = 1 \tag{2.131}$$

The absorption coefficient can be expressed in different ways (Sec. 2.4.2.1), so that the units of S need to be consistent with that choice. For example, if k_ν is the volume absorption coefficient in units of cm^{-1} then S is in $cm^{-1}\ s^{-1}$ when working in frequency units (Hz), or S is in cm^{-2} when working in wavenumber units (cm^{-1}).

2.5.7.1 *Doppler-Broadened Line Shape*

The shape of Doppler-broadened lines comes from the physics of the Doppler effect. When a source emitting radiation moves toward an observer at a relative speed V_x, which is the component of the velocity along the line of sight to the observer, the frequency ν of a detected photon

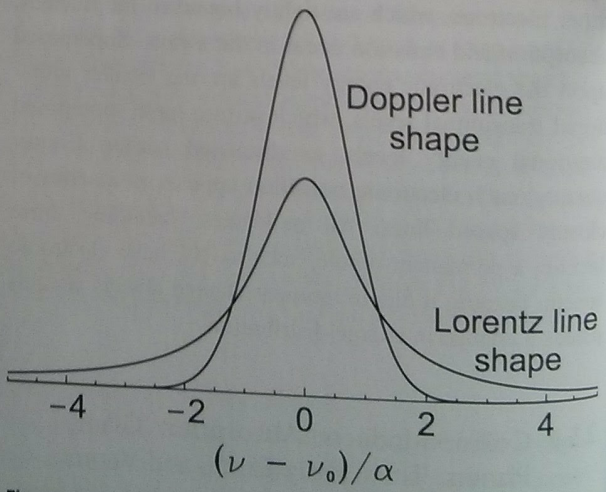


Figure 2.26 Lorentz and Doppler line shapes for the same line widths α and strengths.

is higher than the frequency ν_0 emitted in the moving frame. The difference is given by the Doppler shift formula:

$$\nu = \nu_0\left(\frac{1+V_x/c}{1-V_x/c}\right)^{1/2} \approx \frac{\nu_0}{1-V_x/c} \approx \nu_0\left(1+\frac{V_x}{c}\right) \tag{2.132}$$

The approximations in eq. (2.132) arise from assuming $V_x << c$, the speed of light. For a Doppler shift in recession, we would swap all minus signs for plus signs and vice versa. Rearranging the last expression in eq. (2.132) gives

$$\Delta\nu = \nu - \nu_0 = \nu_0\frac{V_x}{c} \tag{2.133}$$

The one-dimensional version of the Maxwell–Boltzmann distribution of gas molecule velocities gives the probability that a molecule has a speed between V_x and $V_x + dV_x$ as

$$P(V_x)dV_x = \left(\frac{m}{2\pi kT}\right)^{1/2}\exp\left(-\frac{mV_x{}^2}{2kT}\right)dV_x \tag{2.134}$$

where m is the molecular mass, k is Boltzmann's constant, and T is temperature. From eq. (2.133), we see that $V_x = c\Delta\nu/\nu_0$ and $dV_x = (c/\nu_0)d(\Delta\nu)$, which we can substitute in eq. (2.134) to give the probability that the detected photon is at frequency ν shifted by an amount $\Delta\nu$ from frequency ν_0:

$$\begin{aligned}\text{Prob}(\nu-\nu_0=\Delta\nu)d(\Delta\nu) &= \text{Prob}\left(V_x = c\frac{\Delta\nu}{\nu_0}\right)dV_x \\ &= \left(\frac{m}{2\pi kT}\right)^{1/2}\exp\left[-\frac{mc^2}{2kT}\left(\frac{\Delta\nu}{\nu_0}\right)^2\right]\frac{c}{\nu_0}d(\Delta\nu)\end{aligned} \tag{2.135}$$

If we substitute

$$\alpha_D = \frac{\nu_0}{c}\sqrt{\frac{2kT}{m}} \tag{2.136}$$

in eq. (2.135), we get the Doppler line profile:

$$f_D = \frac{1}{\alpha_D\sqrt{\pi}}\exp\left(-\left[\frac{\nu-\nu_0}{\alpha_D}\right]^2\right) \tag{2.137}$$

We describe the width of the line by the *half-width at half-maximum*, $\alpha_{1/2}$, which is the value of $\nu - \nu_0$ at which the absorption falls by half. Solving $f_D(\alpha_{1/2})/f_D(0) = ½$ gives:

$$\alpha_{1/2} = \alpha_D\sqrt{\ln 2} \tag{2.138}$$

Doppler broadening depends on temperature as $T^{1/2}$ but not on pressure. This makes intuitive sense: the higher the temperature, the larger the line width because molecules acquire a wider range of speeds.

2.5.7.2 *Pressure-Broadened Line Shape*

Unlike Doppler broadening, pressure broadening has no exact theory. The Lorentz line shape approximation is given by

$$f_L = \frac{\alpha_L}{\pi\left[(\nu-\nu_0)^2+\alpha_L{}^2\right]} \tag{2.139}$$

where α_L is the half-width at half-maximum, which is related to the time between collisions by $\alpha_L = (2\pi\tau_C)^{-1}$. This equation can be derived by taking the Fourier transform of a burst of light that decays exponentially with time constant τ_C, e.g., see Zdunkowski *et al.* (2007), pp. 205–209. The Lorentz half-width is proportional to the pressure and inversely proportional to temperature to some empirical exponent N_L (where N_L = ½ to 1 depending on the gas),

$$\alpha_L \propto \frac{p}{T^{N_L}} \quad \text{or} \quad \alpha_L = \alpha_{\text{ref}}\left(\frac{p}{p_{\text{STP}}}\right)\left(\frac{T_{\text{STP}}}{T}\right)^{N_L} \tag{2.140}$$

In this equation, T_{STP} and p_{STP} and are standard temperature and pressure values 273.15 K and 1013 mbar, respectively, and α_{ref} is empirical. If we divide eq. (2.136) by eq. (2.140), and put a typical exponent N_L = ½ in the latter, the result is the ratio of Doppler to pressure broadening:

$$\frac{\alpha_D}{\alpha_L} = \frac{\nu_0 p_{\text{STP}}T}{\alpha_{\text{ref}}cP}\left(\frac{2k}{mT_{\text{STP}}}\right)^{1/2} \simeq \left(5\times10^{-11}\ \text{Pa Hz}^{-1}\right)\left(\frac{\nu_0}{P}\right) \tag{2.141}$$

For the last approximation, we have used terrestrial values of m = 30 a.m.u. and a mid-tropospheric temperature T ~ 225 K. Values of α_{ref} range from 0.1 to 1 cm^{-1} in wavenumbers and we have used α_{ref} = 0.07 cm^{-1} (close to the value for CO_2), which in frequency units is α_{ref} = 0.07 cm^{-1} × c × 100 cm/m = 2.1 × 10^9 Hz. Thus, for the prominent 15 μm or 2 × 10^{13} Hz band of CO_2, Doppler and pressure broadened line widths become equal at ~10 hPa, which is in the stratosphere at ~31 km altitude. For the 2.5 mm or 118 GHz microwave line for O_2, eq. (2.141) indicates equal line widths at ~6 Pa, which is ~68 km altitude (Elachi and Van Zyl, 2006). See Fig. 2.27.

Before leaving line shapes, we note that the Lorentz and Doppler line shapes can be convolved into a *Voigt profile* line shape, which is a useful representation when both broadening processes contribute. The Voigt profile equation is more complicated and requires numerical solution.

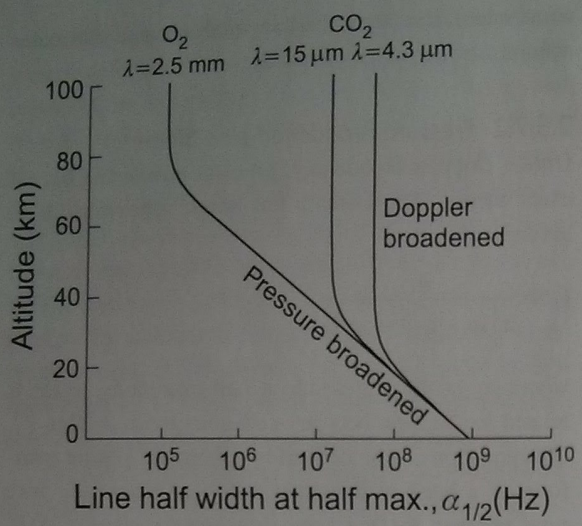


Figure 2.27 Although wavelength-dependent, in the thermal-infrared *pressure broadening* generally dominates in tropospheres, whereas *Doppler broadening* dominates above middle stratospheres. This graph shows the broadening of CO_2 infrared lines and an O_2 microwave line as a function of altitude on Earth. (After a similar diagram by Elachi and van Zyl (2006), p. 453, originally by J. Waters of JPL.)

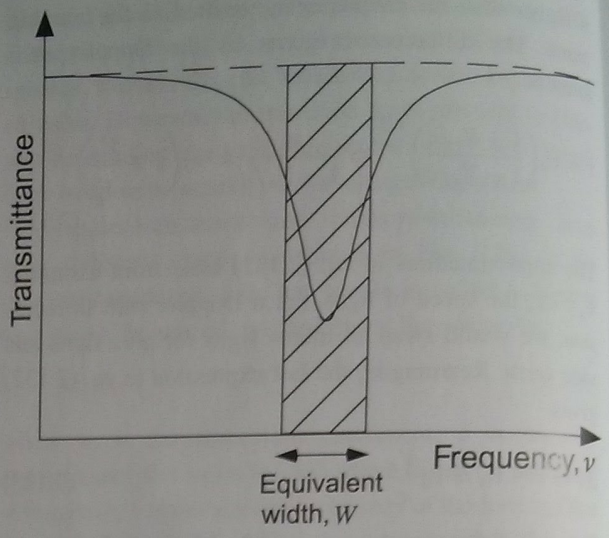


Figure 2.28 Equivalent width, *W*, is the width of a fully absorbing rectangle with the same area as the dip in the signal caused by the absorption line.

2.5.8 Continuum Absorption

In addition to spectral lines and bands, the continuum is smoothly varying absorption over a broad range of wavelengths. In the infrared window of the Earth's atmosphere from 800 to 1200 cm^{-1} (12.5 to 8 μm), weak continuum absorption is mainly due to water vapor (Shine *et al.*, 2012). However, the absorption coefficient is proportional to the *square* of the water vapor density, which is unusual. Two ideas are proposed for this continuum absorption. The first is that the continuum comes from the distant wings of very many lines in the far IR. The second idea is that continuum absorption arises from clusters of H_2O molecules, including dimers ($H_2O{\cdot}H_2O$), trimers, and polymers.

In contrast, in the shortwave, continuum absorption arises from well-known processes, including *photoionization*, when photons are energetic enough to create ions, and *photodissociation* when photons break or dissociate molecules into atoms. Of course, photodissociation is very important for atmospheric chemistry, as we describe in Sec. 3.1.

2.5.9 Band Transmission and Weak and Strong Absorption

Rather than deal with the varying depth of broadened lines, it is sometimes convenient to have a fundamental integrated measure of the extent to which a spectral line can reduce radiation. Such a quantity is the *equivalent width*, which is the width of a completely absorbed rectangle having the height of the continuum and the same absorption area as the line (Fig. 2.28). In frequency units, the equivalent width is defined as

$$W = \int_{\Delta\nu} (1 - e^{-\tau_\nu}) d\nu = \int_{\Delta\nu} (1 - \mathcal{T}_\nu) d\nu = \Delta\nu (1 - \overline{\mathcal{T}}_{\text{band}}) = \Delta\nu \mathcal{A}_{\text{band}} \tag{2.142}$$

where τ_ν is the optical path at frequency ν (eq. (2.56)), $\mathcal{T}_\nu$ is the transmission, and $\Delta\nu$ is the spectral interval. The other quantities are the average *band transmission* and *band absorption* over the spectral band, which are defined as follows:

$$\overline{\mathcal{T}}_{\text{band}} = \frac{1}{\Delta\nu} \int_{\Delta\nu} \mathcal{T}_\nu d\nu \qquad \mathcal{A}_{\text{band}} = 1 - \overline{\mathcal{T}}_{\text{band}} \tag{2.143}$$

If the mass absorption coefficient is $k_{m,\nu}$, we define a *mass path* u_a as a function of absorber density, ρ_a, over an arbitrary pathlength l, from point s_1 to s_2,

$$u_a = \int_{s_1}^{s_2} \rho_a(s) ds, \qquad = \rho_a l \text{ for constant } \rho_a \tag{2.144}$$

Then, the transmission is $\mathcal{T}_\nu = \exp(-k_{m,\nu} u_a)$. There are two limiting cases that help us understand how increases of certain gases will affect greenhouse warming in atmospheres or how to use measured absorption in spectral lines to derive gas abundances.

2.5.9.1 The Weak Line Limit (for an Optically Thin Gas)

If absorption is weak, the transmission and equivalent width are

$$\mathcal{T}_\nu = \exp(-k_{\nu,m}u_a) \approx 1 - (k_{\nu,m}u_a) \qquad W = \int_{\Delta\nu} (1 - \mathcal{T}_\nu)d\nu = \int_{\Delta\nu} k_{\nu,m}u_a \, d\nu = Su_a \tag{2.145}$$

Here, we have used the definition of the line strength S (eq. (2.131)) in the last step. This linear law between W and u_a is called the *weak approximation*. It tells us that if the gas is optically thin (i.e., $k_{m,\nu}u_a << 1$), doubling the amount of absorbing gas or pathlength will double the absorption, which is shown as the linear section of Fig. 2.29(b).

2.5.9.2 Strong Line Limit (for an Optically Thick Gas)

The strong line limit applies when there is complete absorption in the line center, shown on the right-hand side of Fig. 2.29(a). When this happens, we say that the line center has *saturated*. In this case, adding extra absorbing gas can only increase absorption in the wings of the line, which has a nonlinear relationship between absorber amount and absorption. We can derive this relationship using the Lorentz profile (eq. (2.139)). The equivalent width, using eqs. (2.130), (2.139), and (2.142), will be given by

$$W = \int_{\nu-\nu_0=-\infty}^{\infty} [1 - \exp(-Sf_L u_a)]d(\nu - \nu_0) = \int_{-\infty}^{\infty} \left[1 - \exp\left(\frac{-S\alpha_L u_a}{\pi\left[(\nu-\nu_0)^2 + {\alpha_L}^2\right]}\right)\right] d(\nu - \nu_0) \tag{2.146}$$

When the line center is completely absorbing, we need only consider the wings of the line. Consequently, $(\nu - \nu_0)^2 >> {\alpha_L}^2$, so ${\alpha_L}^2$ can be omitted from the denominator. Noting that the line is symmetrical, we can use substitutions to compute the integral as follows:

$$W \approx 2\int_0^{\infty} \left[1 - \exp\left(\frac{-S\alpha_L u_a}{\pi x^2}\right)\right] dx = \sqrt{\frac{S\alpha_L u_a}{\pi}} \int_{\beta=0}^{\infty} \frac{[1 - \exp(-\beta)]}{\beta^{3/2}} d\beta \tag{2.147}$$

$$\text{Using standard integral } \int_0^{\infty} \frac{[1 - \exp(-\beta)]}{\beta^{3/2}} d\beta = 2\sqrt{\pi} \quad \Rightarrow W \approx 2\sqrt{S\alpha_L u_a} \tag{2.148}$$

To solve the integral in eq. (2.147), we made the following substitutions in eq. (2.147),

$$x = \nu - \nu_0, \ \beta = \frac{S\alpha_L u_a}{\pi x^2}, \ \text{so} \ \frac{d\beta}{dx} = -2\frac{S\alpha_L u_a}{\pi x^3} \ \text{and} \ dx = \frac{-1}{2}\sqrt{\frac{S\alpha_L u_a}{\pi}}\frac{d\beta}{\beta^{3/2}}$$

We also absorbed the minus sign in substituting for dx in eq. (2.147), given that the limits on the integral in eq. (2.147) should otherwise reverse in changing from x to β. The result that $W \approx 2(S\alpha_L u_a)^{1/2}$ is known as the *square root law*.

A log–log plot of the equivalent width W versus the mass path of the absorber u_a, is called the *curve of growth* (Fig. 2.29(b)), in which we can see the progression from the "weak" linear relationship to the "strong" square root relationship.

What does the curve of growth mean? If a trace greenhouse gas, such as CCl_2F_2 in the Earth's atmosphere, doubles in abundance, its band-averaged absorption will double along a path. On the other hand, a gas such as CO_2, which is saturated in strongest bands such 15 μm, needs to quadruple in abundance in order to double its band-averaged absorption. In fact, the radiative forcing change due to perturbations in the abundances of these particular gases in the current Earth's climate can each be represented by simplified analytical expressions that reflect the line broadening. There is a linear relationship of forcing to CCl_2F_2 concentration and a logarithmic one for CO_2 concentration (e.g., Sec 6.3.5 of Ramaswamy *et al.* (2001), noting that hot climates with much higher CO_2 levels deviate from purely logarithmic (Caballero and Huber, 2013)).

Curves of growth are also useful in deriving the abundance of gases from planetary spectra. For example, one of the authors (DCC) when an undergraduate, measured reflected sunlight from Jupiter's cloud-tops using a small optical telescope, grating spectrometer, and

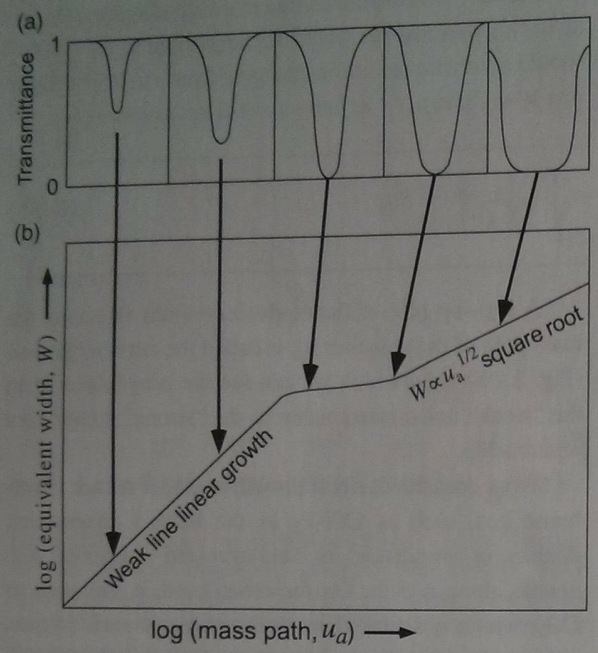


Figure 2.29 How the *curve of growth* of equivalent width W, is related to changes in line shape. (a) The schematic change in the line shape as the absorber density and/or pathlength increases. (b) The related change in the log–log plot of equivalent width versus the mass path u_a, i.e., the product of absorber density and pathlength.

charged-couple device detector, which revealed absorptions due to methane vibration–rotation bands at wavelengths centered at 543 nm and 619 nm. From the measured absorption, one can calculate equivalent widths. Then one can use laboratory data of the curve of growth for methane in the linear weak limit to calculate the column abundance of methane above Jupiter's cloud-tops.

2.6 Calculating Atmospheric Absorption in Climate Calculations

Above, we have considered single lines, but for calculating the climate of a planet, we have to deal with thousands to millions of lines with a great range of line widths rather than isolated lines, and often these lines overlap. Consequently, sophisticated methods have been devised to quantify atmospheric absorption as a function of wavelength.

Line-by-line calculations add up the effects of all the contributions of known individual lines. This method is as accurate as possible (and line lists are continually updated to provide even higher accuracy), but computers take a very long time to calculate broadband fluxes at every level in an atmosphere, which proves impractical for repeated climate calculations. Consequently, two other approaches have been developed to reduce the computational time by orders of magnitude.

- *Band models* rely on a simple analytical form for the positions and strengths of lines, which can be solved much faster than line-by-line calculations. Lines are assigned mean line strengths and are spaced randomly or in some other arrangement, which allows analytical calculation over some spectral interval.
- The *correlated k-distribution method* recasts the very complicated variation of absorption coefficients k_ν over a spectral interval $\Delta\nu$ into a smooth, monotonic, cumulative probability function of the fractional occurrence of k_ν within $\Delta\nu$. This approach is a statistical technique that speeds up calculations. Each spectral interval $\Delta\nu$ is sufficiently narrow that the Planck function within the interval is treated as a function of temperature only, without frequency dependence. Because the cumulative probability function is smooth and monotonic, the k_ν values can be summed with relatively large bins, which allows for fast computation of transmissivity. See Appendix A for a description of this method in a one-dimensional (1-D) radiative–convective climate model.

Validation against line-by-line calculations shows that band models and k-distribution methods can produce realistic results if applied correctly. Petty (2006) gives a very readable introduction to such methods, so we will not describe these models further. More detailed treatments are given in lengthier books devoted to atmospheric radiation by Liou (2002), Thomas and Stamnes (1999), Goody and Yung (1989) and Zdunkowski (2007). Pierrehumbert (2010) also includes detailed discussion.