2 Absorption: The Death of Photons

If emission may be called, somewhat poetically, the birth of photons, absorption may be called their death, although their spirit (energy) lives on in whatever absorbs them. The previous chapter is devoted mostly but not exclusively to emission because it would be artificial to try to divorce it completely from absorption given that they are inverse processes. In this chapter the emphasis is on absorption.

We purposely used the term absorption by a body in the previous chapter to avoid conveying that radiation is absorbed at or by surfaces. It is not, if only because they do not exist. Surfaces are mathematical idealizations of transitions in material properties over distances smaller than we can perceive with our eyes or fingers or measuring instruments. Perhaps *interface* better conveys this sense of a sharp transition. If radiation incident at the interface between a negligibly absorbing medium and an absorbing medium is attenuated by absorption over a very short distance, we might say that the surface (interface) absorbed the radiation. But it would be more accurate to say that absorption occurred *because* of the interface and over a short distance (on some scale) within the medium. We begin by making no assumption about the characteristic distance over which absorption occurs: it could be meters or kilometers or micrometers or nanometers.

2.1 Exponential Attenuation

Consider attenuation by absorption of a monodirectional beam of monochromatic radiation by an optically homogeneous medium. By *optically homogeneous* is meant homogeneous on the scale of the wavelength; no medium is absolutely homogeneous (continuous). By *monodirectional* is meant that the radiation is confined to a narrow set of directions, and by *monochromatic* is meant that it is confined to a narrow set of wavelengths. Again, a truly monodirectional, monochromatic beam does not exist. Even a laser beam has a finite angular divergence and range of wavelengths.

The beam is directed along the x-axis and its *irradiance*, radiant energy crossing unit area per unit time (Sec. 4.2) at $x = 0$ is F_0. We can imagine the region from 0 to an arbitrary distance x to be subdivided into N identical slices each of thickness $\Delta x = x/N$ (Fig. 2.1). Upon transmission from 0 to Δx the irradiance of the incident beam is reduced (attenuated) from F_0 to F_1 because of absorption within Δx. It is reasonable to assume that if Δx is sufficiently small attenuation is proportional to Δx and to F_0, where the proportionality constant (*absorption coefficient*) is denoted by κ:

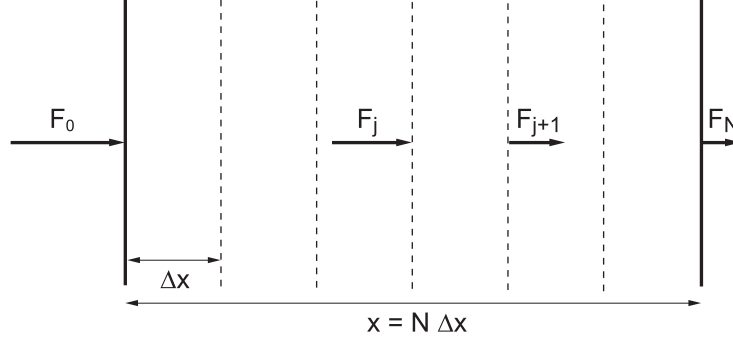$$F_0 - F_1 = F_0 \kappa \Delta x. \tag{2.1}$$

Figure 2.1: An absorbing, plane-parallel medium can be imagined to be made up of N slices, each of thickness Δx. A monodirectional beam with irradiance F_0 is incident on this medium. The transmitted irradiance F_j at a distance $j\Delta x$ into the medium decreases with increasing j because of absorption in all preceding layers.

It is more convenient to write this as

$$F_1 = F_0(1 - \kappa\Delta x). \tag{2.2}$$

This is the irradiance at a distance Δx from the origin. At a distance $2\Delta x$, by the same argument, the irradiance is

$$F_2 = F_1(1 - \kappa\Delta x) = F_0(1 - \kappa\Delta x)^2. \tag{2.3}$$

An implicit assumption in going from Eq. (2.2) to Eq. (2.3) is that transmission by the two slabs is independent, and hence transmission by each can be multiplied to obtain transmission by the two combined (this assumption is scrutinized in Sec. 2.4). The pattern now should be clear: after transmission over a distance $x = N\Delta x$, the irradiance is

$$F_N = F_0(1 - \kappa x/N)^N. \tag{2.4}$$

Now take the limit as N becomes indefinitely large keeping κx constant:

$$F = \lim_{N \to \infty} F_0(1 - \kappa x/N)^N, \tag{2.5}$$

where we omit the subscript on F. Do you recognize this limit? It is one of the many ways of defining the exponential function:

$$e^\xi = \lim_{n \to \infty} (1 + \xi/n)^n. \tag{2.6}$$

From Eqs. (2.5) and (2.6) we obtain the law of exponential attenuation (by absorption)

$$F = F_0 \exp(-\kappa x). \tag{2.7}$$

This law is also valid for attenuation by scattering if multiple scattering is negligible (see Sec. 5.2.3).
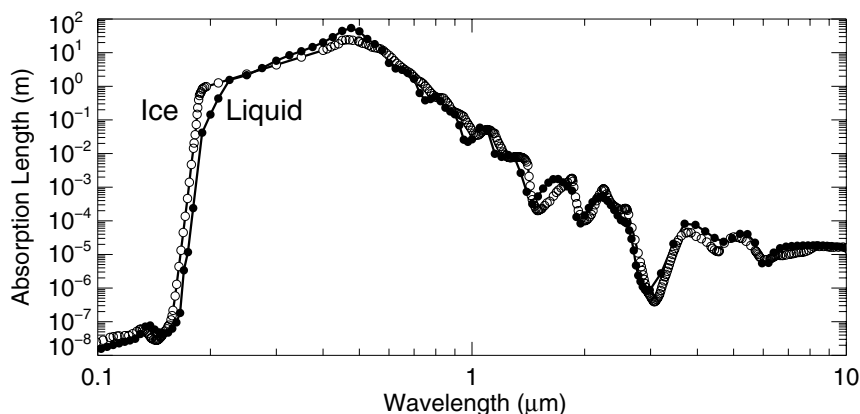
Figure 2.2: Absorption length (inverse absorption coefficient) of pure ice and liquid water from UV to IR. The data for liquid water were taken from Querry *et al.* (1991), those for ice from Warren (1984).

The various names by which this law is called exemplify *Stigler's law of eponymy*: "No scientific discovery is named after its original discoverer." The law of exponential attenuation, often called Lambert's law, was first stated in Pierre Bouguer's *Essay on the Gradation of Light* (1729), although we could find no evidence that he established it experimentally. Most chemists call it Beer's law, which is wide of the mark given that Bouguer preceded Beer by more than 100 years, Beer did not discover an exponential law of attenuation with distance, and, in fact, did not explicitly state *any* exponential law. The most we can say is that by reworking Beer's data one can unearth what he did not: an exponential attenuation law for solutions of fixed thickness but variable concentration of the absorbing solute.

Because κx is dimensionless, κ must have the dimensions of inverse length, and hence $1/\kappa$ must have the dimensions of length. The *e-folding length* is the distance over which a monodirectional beam is attenuated by a factor $1/e$. Because this term is a bit of a mouthful we prefer *absorption length* for $1/\kappa$. As a general rule, whenever any physical quantity can be expressed as a length it is wise to do so. Lengths are easier to get a feel for, more so than time, even mass. We can both touch and see lengths.

The absorption coefficient (absorption length) of a material depends on wavelength, often varying by as much as a factor of 10^{10}, an example of which is as near as a faucet. Figure 2.2 shows the huge range of absorption lengths for liquid water and ice from 0.1 µm to 10 µm. The absorption length for (pure) liquid water and ice over the visible spectrum, shown on a linear scale in Fig. 5.12, is greatest in the blue, least in the red. From this figure it is evident why water in a drinking glass is not noticeably colored: the dimensions of glasses are small compared with the absorption length of water over the visible spectrum. This curve also shows that transmission of white light over several meters through water is sufficient to attenuate much more of the long-wavelength components than the short. Indeed, with increasing distances the only component that survives corresponds to the greatest absorption

length. Water is intrinsically blue and needs no impurities to make it so. In Section 5.3.1 we explore the consequences of this to the observed colors of natural ice bodies such as glaciers, icebergs, ice caves, icefalls, and even holes in snow.

2.1.1   Absorptivity and Absorption Coefficient: A Tenuous Connection

Absorptivity and absorption coefficient are not the same. In the first place, the former is dimensionless whereas the latter has the dimensions of inverse length, which itself ought to signal caution. More to the point, the connection between them is sometimes tenuous at best. Consider, for example, radiation incident on bodies sufficiently thick that transmission by them is negligible. We now can attach a more precise meaning to "sufficiently thick": much thicker than the absorption length at the wavelength of the radiation. With this assumption, the absorptivity of the body is 1 minus its reflectivity. How does the reflectivity of the body depend on its absorption coefficient? For many materials over many wavelength intervals, reflectivity changes hardly at all even with huge increases in absorption coefficient. And if there is a change, it is likely to result in a *decrease* in absorptivity (see Prob. 7.20). For example, the absorption coefficient of metals such as silver and aluminum is usually huge compared with that of insulators such as quartz and salt, a million times or more, especially at visible and near-visible wavelengths. And yet reflectivities of metals are high, and hence their absorptivities are lower than those of insulators. Finally, there is this important distinction to be kept in mind: absorptivity is a property of a *body* whereas absorption coefficient is a property of a *material*.

2.1.2   Absorptance and Absorbance: More Room for Confusion

As if distinguishing between absorptivity and absorption coefficient were not difficult enough, we also have to keep these terms separate from the near homophones *absorptance* and *absorbance*. Although absorptance is sometimes used as a synonym for absorptivity, this is not recommended given that we try to restrict terms ending in "ance" to amounts of radiant power. For example, emittance, which can be looked upon as shorthand for *emitt*ed irradi*ance*, is radiant power per unit area. Similarly, absorptance can be looked upon as shorthand for absorbed irradiance.

Absorbance, a term widely used by chemists, is the negative logarithm (base 10) of the transmissivity of a sample of an absorbing material (usually liquid) in a container (cell). Because a transmissivity less than 1 is a consequence both of reflection by the container and absorption by its contents, the apparent absorbance can be nonzero even with an empty cell or one filled with a negligibly absorbing liquid. To correct for reflection, the absorbance of the cell is subtracted from the apparent absorbance to obtain that of the sample. To good approximation the transmissivity of the sample in the cell often is

$$T = T_0 \exp(-\kappa h), \tag{2.8}$$

where h is the sample thickness, κ its absorption coefficient, and T_0 the transmissivity of the cell without the sample in place. Take the negative logarithm of both sides of Eq. (2.8) to obtain

$$-\log T + \log T_0 = \kappa h \log e = 0.434 \kappa h. \tag{2.9}$$

The left side of this equation is absorbance corrected (approximately) for reflection by the cell. With this correction, absorbance measured by chemists is, except for a constant factor, the absorption optical thickness (κh) of the sample (see Sec. 5.2).

2.1.3 The Sum of Exponentials is not an Exponential

The exponential attenuation law Eq. (2.7) strictly holds only for monochromatic radiation because κ depends on frequency. Any real source is distributed over frequency, and hence the integrated transmitted irradiance is

$$\int F_0(\omega) \exp(-\kappa x)\, d\omega, \tag{2.10}$$

where $F_0(\omega)$ is the *spectral irradiance* (irradiance per unit frequency interval) at $x = 0$. The limits of integration can be anything, and for simplicity we do not express κ as a function of frequency. Although each spectral component of the incident beam is attenuated exponentially with distance, the integrated beam is not. And this is true even if the incident irradiance does not depend on frequency. This basic property of exponential attenuation has sometimes been forgotten, resulting in errors.

To show that the sum (integral) of exponentials is *not*, in general, an exponential, we assume that

$$\exp(-Kx) = \int \exp(-\kappa x)\, d\omega, \tag{2.11}$$

where K is independent of x and ω. Differentiate both sides of this equation with respect to x to obtain

$$K = \frac{\int \kappa \exp(-\kappa x)\, d\omega}{\int \exp(-\kappa x)\, d\omega}. \tag{2.12}$$

The right side of this equation, the average of κ weighted by a normalized exponential, depends on x, in general, and hence we contradict our original assumption that K is independent of x, which therefore must be false. This is a proof by contradiction: assume something is true, explore the consequences, and when a contradiction results, the original assumption must have been false. This, by the way, is a variation on a theme in Section 1.4.2: the average of a function is not necessarily the function of the average.

When κ is independent of frequency over the range of interest, the integrated irradiance does decrease as a simple exponential. And when $\kappa x \ll 1$ for the frequency range and distances of interest we can approximate the exponential in Eq. (2.10) by the first two terms in its Taylor series expansion to obtain the following approximation for the transmitted irradiance:

$$\int F_0(\omega)(1 - \kappa x)\, d\omega. \tag{2.13}$$

This can be written as

$$F_0(1 - \langle \kappa \rangle x), \tag{2.14}$$

where the integrated irradiance at $x = 0$ is

$$F_0 = \int F_0(\omega) \, d\omega \tag{2.15}$$

and the average absorption coefficient is

$$\langle \kappa \rangle = \frac{\int F_0(\omega) \kappa \, d\omega}{\int F_0(\omega) \, d\omega}. \tag{2.16}$$

We obtained Eq. (2.14) by approximating an exponential by the first two terms in a Taylor series, but we can do the reverse, approximate the first two terms in a series by an exponential:

$$1 - \langle \kappa \rangle x \approx \exp(-\langle \kappa \rangle x), \tag{2.17}$$

which yields approximate exponential attenuation for the integrated irradiance:

$$F \approx F_0 \exp(-\langle \kappa \rangle x). \tag{2.18}$$

But in general, this equation is not correct.

2.1.4 Attenuation in a Nonuniform Medium

Nature is not so cooperative as to provide us only with media having uniform properties. The absorption coefficient κ can vary from point to point. The medium depicted in Fig. 2.1 can be nonuniform and subdivided into N equal slices so thin that in each of them the absorption coefficient is nearly constant. Transmission by the jth slice is

$$F_{j+1} \approx F_j \exp(-\kappa_j \Delta x), \tag{2.19}$$

where F_j is the irradiance incident on the jth slice, F_{j+1} is the irradiance transmitted by this slice, κ_j is the absorption coefficient $\kappa(x)$ at some point in the interval (x_j, x_{j+1}), and $\Delta x = x/N$. It follows from this that transmission over the distance x is

$$F = F_0 \exp\left(-\sum_{j=1}^{N} \kappa_j \Delta x\right). \tag{2.20}$$

The limit of the sum in this equation is the integral

$$\int \kappa(x) \, dx = \lim_{N \to \infty} \sum_{j=1}^{N} \kappa_j \Delta x, \tag{2.21}$$

and hence the exponential attenuation law for a nonuniform medium (spatially varying absorption coefficient) is

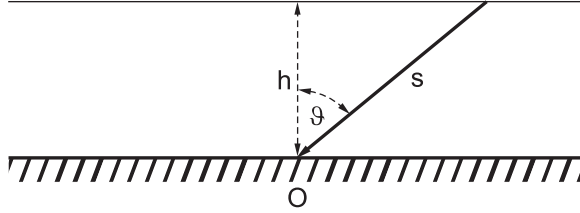$$F = F_0 \exp\left\{-\int \kappa(x) \, dx\right\}. \tag{2.22}$$

Figure 2.3: The path length s through a slab, and hence its absorptivity, is greater the more the path is slanted from the vertical.

2.2 Directional Emissivity of the Atmosphere

We note in Section 1.4 that emissivity depends on direction, and also show calculations of the spectral normal emissivity of a uniform layer of moist air. And in Section 1.5 we show how the emissivity of a layer of water depends on direction. With the law of exponential attenuation in hand we can go a step further and show how the emissivity of the atmosphere depends on direction.

Although Earth is round, its radius is about 1000 times greater than the thickness of that part of the atmosphere containing most of the infrared-active gases. Because of this we can pretend that the atmosphere is a planar slab infinite in lateral extent. For simplicity we assume that the absorption coefficient κ of the atmosphere does not depend on altitude. That this is not true does not affect our general conclusions.

Because the reflectivity of the atmosphere for terrestrial infrared radiation is negligible, the absorptivity of the atmosphere is 1 minus its transmissivity, which over any path is given by the exponential attenuation law:

$$\exp(-\kappa s), \tag{2.23}$$

where s is the path length. If κ varies along the path, Eq. (2.23) is replaced by the path integral of κ as in Eq. (2.22). For a slab atmosphere of thickness h and a path making an angle ϑ with the vertical (zenith), $s = h/\cos\vartheta$ (Fig. 2.3), and hence the emissivity (absorptivity) of the atmosphere in any direction ϑ is

$$\varepsilon = 1 - \exp(-\kappa h/\cos\vartheta). \tag{2.24}$$

Here is an example in which absorptivity (emissivity) does increase with increasing absorption coefficient. Emissivity is least overhead and increases to 1 toward the horizon. But increased emissivity toward the horizon is not the only reason why *emission* increases in this direction. Emission depends on the temperature of the emitting body as well. The temperature of the atmosphere is not uniform. In the troposphere, which contains most of the atmosphere, temperature usually decreases with height.

Radiation from the atmosphere at the surface, in a given direction, comes from all points on an atmospheric path along that direction. But radiation emitted at each point must be transmitted over some distance through the atmosphere to the observation point O (Fig. 2.3). The greater this distance, the less radiation transmitted. This is evident from Eq. (2.24), according to which emissivity does not increase indefinitely with increasing path length but
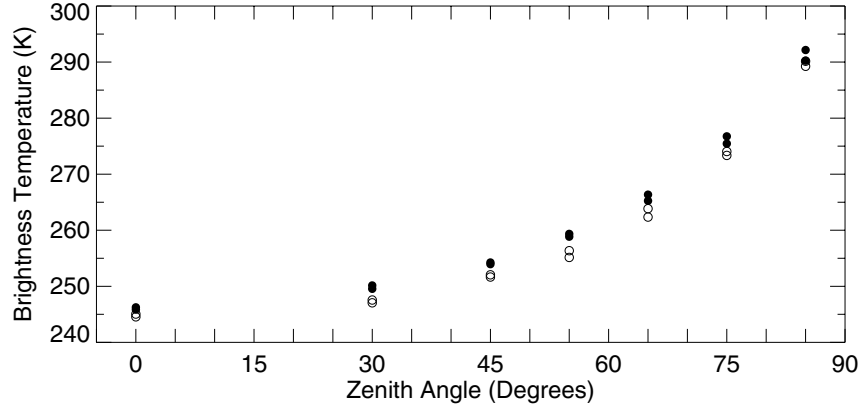
Figure 2.4: Brightness temperature as a function of zenith angle on a clear summer day. The first measurement always was made at zenith ($0°$), then for increasing angles to about $85°$ ($5°$ above the horizon). This was done four times: twice for the radiation thermometer pointed east-northeast (solid circles), twice pointed west-southwest (open circles).

asymptotically approaches 1. All points on the atmospheric path contribute to radiation at the observation point O. But those closest to it contribute most for two reasons: the concentration of emitting gases, and hence κ, decreases with height (distance from the observation point), and the shorter the path from emitting gases to the observation point, the greater the fraction of emitted radiation transmitted to it.

Even if the atmosphere were uniform in temperature, emitted radiation (at the ground) would be least overhead and increase toward the horizon. Because temperature usually decreases with height, this directional variation of emission is even greater. We leave it as a problem (Prob. 5.19) to show that the decrease of temperature with height in Earth's atmosphere is a minor contributor to the total variation of emission with direction.

If emission does indeed increase from zenith to horizon, so should the brightness temperature of the atmosphere. To demonstrate this, we measured brightness temperatures at various zenith angles (Fig. 2.4) from a valley floor near State College on a clear summer day. As predicted by Eq. (2.24) brightness temperature increases with increasing zenith angle.

2.3 Flux Divergence

If more radiant energy enters a region than leaves it, radiant energy must be converted into other forms within the region. This transformation is usually manifested by a temperature increase but could be manifested in other ways (e.g., photosynthesis).

Consider a monochromatic, monodirectional beam propagating along the x-axis in an optically homogeneous medium characterized by an absorption coefficient κ. A region between x and $x + \Delta x$ is bounded by planes of area A. The difference between the radiant energy

entering this region and that leaving it is

$$AF_0 \exp(-\kappa x) - AF_0 \exp\{-\kappa(x + \Delta x)\}, \tag{2.25}$$

where F_0 is the irradiance at $x = 0$. Divide this difference by the volume of the region

$$\frac{AF_0 \exp(-\kappa x) - AF_0 \exp\{-\kappa(x + \Delta x)\}}{A\Delta x} \tag{2.26}$$

and take the limit as $\Delta x \to 0$ to obtain the rate of energy conversion (transformation) per unit volume around a point:

$$-F_0 \frac{d}{dx} \exp(-\kappa x) = -\frac{dF}{dx} = F_0 \kappa \exp(-\kappa x). \tag{2.27}$$

Thus the rate of energy transformation per unit volume is the negative of the spatial derivative of the irradiance, often called the *flux divergence*. The negative flux divergence is proportional to the local rate of temperature change under the assumption that radiant energy transformation results only in temperature increases. Note that the flux divergence is a product of two functions, one of which increases with increasing κ, the other of which decreases; the maximum of Eq. (2.27) occurs for $\kappa x = 1$ for fixed $x > 0$.

Suppose that we want to heat an object by illuminating it with radiation. As a concrete example, take the radiation to lie in the microwave region. How should we choose the frequency? The frequency of maximum absorption coefficient would give the highest heating rate but it might be localized near the surface of the object. For the object to be heated more or less uniformly, the absorption length $(1/\kappa)$ should be comparable with the linear dimensions of the object. The frequency used in microwave ovens is chosen so that the absorption length of water (mostly the water in food enables it to be heated in microwave ovens) is about equal to the linear dimensions of typical objects heated in them (all microwave ovens are about the same size). Many years ago on an examination we gave the absorption properties of water over a wide range of frequencies and asked students to pick the frequency of a microwave oven (we didn't know). Some students fell into the trap of picking the frequency of maximum absorption coefficient, but others recognized that an absorption length of, say, a few mm, would result in braised but not well-cooked food. Following the exam, we consulted the experts in our electronics shop and, to our delight, learned that the frequency chosen on the basis of an absorption length for water of around 10–20 cm is the frequency used in ovens.

2.3.1 The Sum of Exponentials is not an Exponential: Another Example

From now on we use the term *plane-parallel* medium to mean one confined between two parallel planes infinite in lateral extent. The properties of the medium may vary from point to point but only in the direction perpendicular to the planes. We sometimes call this direction the vertical direction even though it need not coincide with gravity. For our purposes, radiation is oblivious to gravity.

Suppose that a plane-parallel, uniform, absorbing medium is illuminated by *isotropic* radiation (Fig. 2.5). By isotropic is meant that the radiation (specifically, the radiance [Sec. 4.1.2]) is the same in all directions in a hemisphere. The total irradiance (Sec. 4.2) is the weighted,
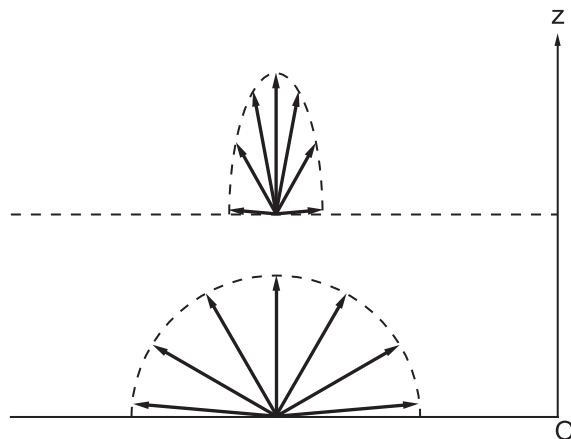
Figure 2.5: An initially isotropic source of radiation, incident at $z = 0$ on an absorbing medium, becomes more sharply peaked toward the z-direction with increasing z.

by cosine and solid angle, sum of contributions from all directions. Any particular direction is specified by ϑ, the angle with the vertical. At any depth z into the medium, radiation along a particular path is attenuated by the factor $\exp(-\kappa z/\cos\vartheta)$. At this depth, the irradiance is again the weighted sum of contributions from radiation in all directions. Although radiation from each direction is the same at the illuminated boundary of the medium ($z = 0$), this is not true at $z > 0$. Thus the initially isotropic radiation field does not remain isotropic. In the limit of indefinitely large κz, the emerging radiation would be monodirectional in the vertical direction, although of vanishingly small irradiance. Moreover, although the irradiance is the sum (integral) of exponentials and depends only on z, it is not an exponential function of z.

2.4   Absorption Cross Section

Determining the absorption coefficient of liquids and solids from the absorption properties of their individual molecules is not an easy task because they are sufficiently close together that they interact strongly. This is evident from Figs. 1.11 and 1.12, which show that the spectral emissivity of liquid water bears little resemblance to that of water vapor. Beginning with the latter it is not an easy step to the former. Interactions between water molecules in the liquid phase all but destroy their individuality. For gases and suspensions of particles, however, we do have a hope of determining absorption coefficients beginning with the properties of a single molecule or particle.

By *particle* we mean a bound collection of molecules sufficient in number that it has macroscopic properties such as temperature and pressure. There is no such thing as the temperature or pressure of a molecule. Even the radius of a molecule is a nebulous quantity: every method for measuring molecular diameters yields a different result. A particle may itself be composed of a material with an absorption coefficient, but there is no such thing as the absorption coefficient of a single molecule. All molecules of the same substance are essentially identical, but every particle is unlike every other particle. Like temperature and

pressure, absorption coefficient is a statistical quantity, an average over an ensemble of many molecules.

We expect the absorption coefficient of a gas to depend on the concentration of its molecules. After all, the inverse of the absorption coefficient is the absorption length, and it would hardly make sense if a gas at one concentration had the same absorption length as the same gas at a higher concentration. The term *concentration* instead of *density* is used here to emphasize that absorption of electromagnetic radiation is not fundamentally dependent on mass. Electromagnetic waves exert forces on charges, not masses, which just go along for the ride. When you use the unqualified term density, make sure that you are clear whether you mean mass density (mass per unit volume) or number density (molecules per unit volume). We use number density and concentration to mean more or less the same thing.

Suppose that an isolated molecule is illuminated by a monodirectional, monochromatic beam of irradiance F. This molecule absorbs energy, by which is meant it transforms radiant energy into other forms, at a rate W_a, which is proportional to F. The dimensions of W_a are power, whereas those of F are power per unit area. Thus the proportionality factor σ_a between the two

$$W_a = \sigma_a F \tag{2.28}$$

must have the dimensions of length squared (area). For this reason σ_a is given the name *absorption cross section*. This is the *effective area* of the molecule for removing energy from the incident beam, and should not be confused with the geometrical cross-sectional area of the molecule even if such an area had a precise meaning, which it does not. The diameters of the kinds of molecules that inhabit the atmosphere are about 3×10^{-8} cm, which corresponds to geometrical cross-sectional areas of about 7×10^{-16} cm^2. But absorption cross sections of molecules are usually much smaller than this value.

Equation (2.28) holds equally well if the illuminated object is a particle rather than a molecule. We have more to say about absorption cross sections of particles in Section 2.9.

Cross sections of various kinds are fundamental in several areas of physics, common to the kinetic theory of gases, neutron physics, high-energy particle physics, and optics.

Suppose that a thin slab of gas with area A and thickness Δx, populated by N molecules per unit volume with absorption cross section σ_a, is illuminated by a monodirectional, monochromatic beam directed perpendicular to the faces of the slab (Fig. 2.6). The total number of molecules in the slab is $NA\Delta x$, and hence the total effective area for removing radiant energy from the beam is $\sigma_a NA\Delta x$, where we assume that no molecule overlaps (or shadows) another in the direction of the beam. The decrease in the transmitted power over the distance Δx is

$$A\Delta F = -F\sigma_a NA\Delta x, \tag{2.29}$$

where F is the incident irradiance. From this it follows that

$$\Delta F = -FN\sigma_a \Delta x = -F\kappa\Delta x, \tag{2.30}$$

where

$$\kappa = N\sigma_a \tag{2.31}$$

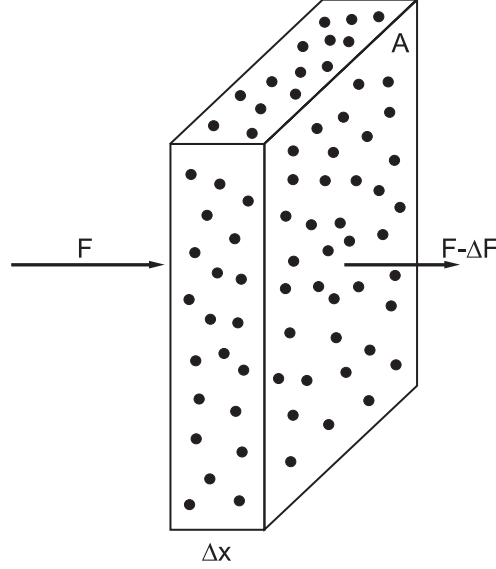is the absorption coefficient from Eq. (2.1).

Figure 2.6: The irradiance F incident on a slab randomly populated by absorbing molecules is attenuated by an amount proportional to the total cross-sectional area (projected effective area) of all the molecules in the slab.

We implicitly assumed that all the molecules are identical. Even if they are, however, they are unlikely to all be oriented identically, and so the absorption cross section in Eq. (2.31) should be interpreted as an average. Because of the additivity of κ:

$$\kappa = \sum_j N_j \sigma_{\mathrm{a}j}, \ N = \sum_j N_j, \tag{2.32}$$

where j indicates a particular molecular orientation and σ_{a} in Eq. (2.31) is the orientational average

$$\frac{1}{N} \sum_j N_j \sigma_{\mathrm{a}j}. \tag{2.33}$$

If all orientations are equally likely (random orientation), all cross sections are equally weighted.

Another assumption underlying Eqs. (2.31)–(2.33) is that the consequences of interference (coherence) are negligible (see Secs. 3.4 and 5.1). If not, to determine the irradiance transmitted at any x, we could not skip the intermediate step of determining the amplitude and phases of the electric and magnetic fields, from which irradiances follow.

The absorption cross section also depends on the polarization state (see Ch. 7) of the beam if the molecule is asymmetric. But if the molecules are randomly oriented, the absorption coefficient of the gas is independent of polarization.
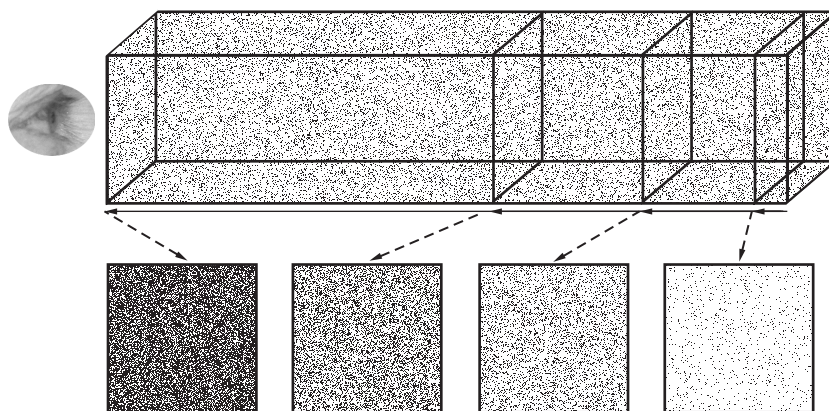
Figure 2.7: The square tube is randomly filled with 25,000 particles (black squares). What an observer would see (looking to the right) at different points along this tube is depicted in the squares below the tube.

Everything we say about absorption by a gas is equally applicable to a suspension of particles. All that need be done is replace the absorption cross section of a molecule with that of a particle (see Sec. 2.9).

Equation (2.7) for exponential attenuation, with the absorption coefficient of a gas (or suspension of particles) given by Eq. (2.31), contains hidden assumptions worth bringing into the open. Even if coherence is negligible, the possibility remains that one molecule or particle can affect absorption by others. Suppose, for example, that a tube contains a fixed number of absorbers each with the same absorption cross section. How they are distributed in space determines attenuation by them. If *all* the absorbers happen to line up one behind the other, their total projected absorption area is just the absorption cross section of a single absorber. But if they are distributed so that *no* absorber is behind another, their total projected absorption area is much greater. Attenuation is vastly different for these two spatial distributions even if the total number of absorbers in the tube is the same. All absorbers lined up in a row is physically unrealistic (except in a crystal), but this extreme distribution signals that how the projected absorption areas of the absorbers overlap determines attenuation by them.

To probe further we did computations for two possible arrangements of absorbers represented by black squares, henceforth called particles. A large square with area 4 units, representing the cross-sectional area of a square tube, was filled with identical particles with cross-sectional area 0.0001 units according to two different prescriptions. In one prescription the positions of the particles within the large square were randomly chosen, added to the large square one after another, and the total cumulative particle area recorded (Fig. 2.7). The total area of overlap of an added particle with all other particles in the tube was subtracted from the cumulative particle area to produce the total projected area. In the other prescription, the particles were not allowed to overlap. If the position of a particle was such that it overlapped any other, it was relocated to remove all overlap, and hence the resulting distribution of particles was not perfectly random.
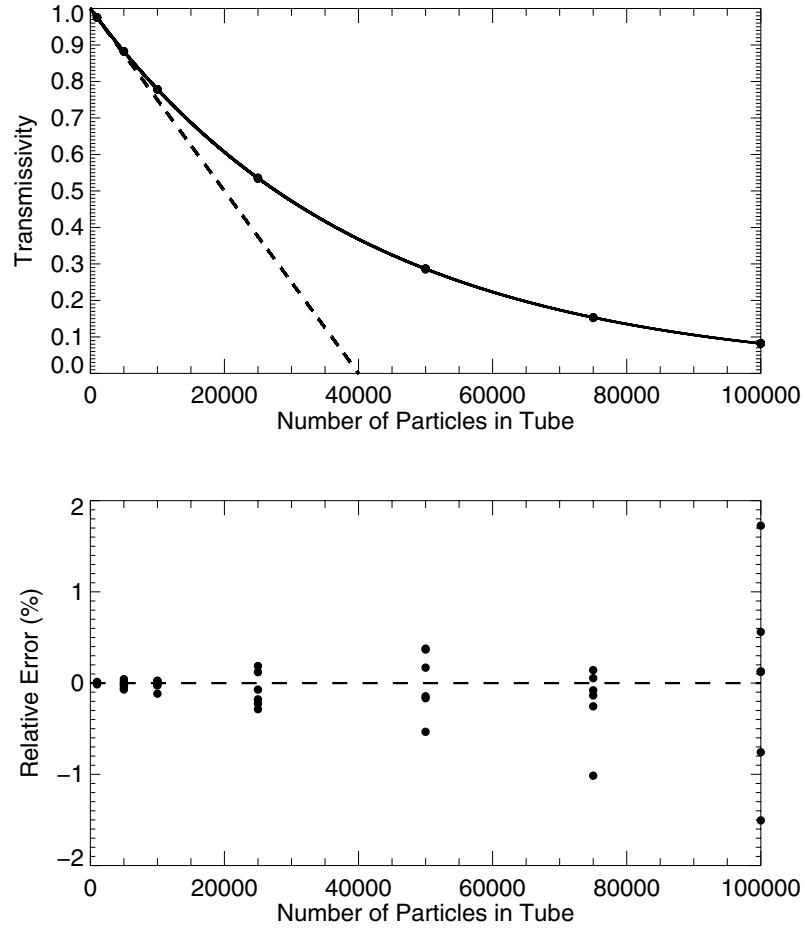
Figure 2.8: The top figure shows transmissivity versus number of particles for two different prescriptions: overlapping particles (solid circles) and non-overlapping particles (dashed curve). Each solid circle on the solid curve (exponential attenuation) is the result of six sets of calculations for 1000, 5000, 10,000, 25,000, 50,000, 75,000, and 100,000 particles. Because the variation over the six sets is not resolvable in the upper part of the figure, the lower part shows the ratio of the difference between each calculated transmissivity and exponential decreasing transmissivity, relative to exponential decreasing transmissivity.

The quantity 1 minus the total projected area of particles divided by the cross-sectional area of the tube is the transmissivity, shown for both overlapping and non-overlapping particles in Fig. 2.8. For a sufficiently small number of particles, the transmissivity is the same, but as this number increases, the two curves diverge. For the overlapping particles, the transmissivity
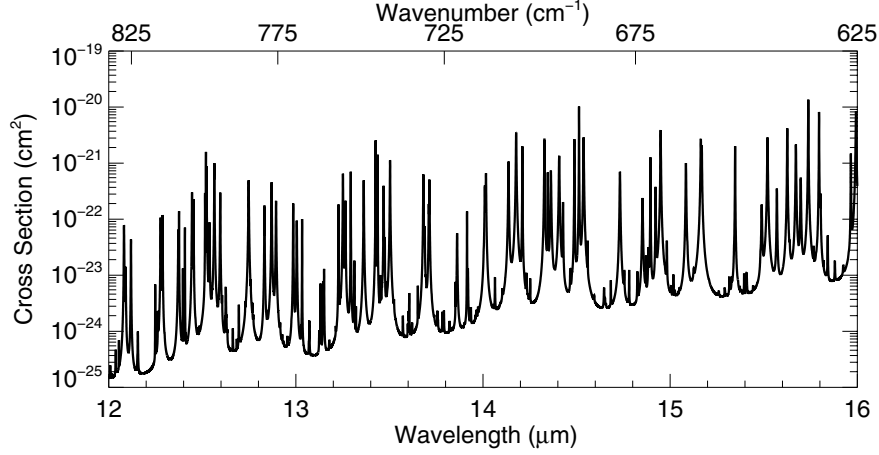
Figure 2.9: Absorption cross section of a water molecule for a temperature of $20°\text{C}$ and a total pressure of 1 atmosphere.

closely follows the exponential law $\exp(-NA_\text{p})$, where N is the total number of particles in the tube and A_p is the cross-sectional area of a particle relative to the cross-sectional area of the tube. Each dot on the solid curve $[\exp(-NA_\text{p})]$ in Fig. 2.8 is the result of six sets of calculations, each of which gave slightly different results, although they are not resolvable. To resolve the differences, the lower part of the figure shows the relative deviation of each calculation, defined as the ratio of the difference between the computed transmissivity and that predicted by the exponential law relative to the exponential law.

For the non-overlapping particles, transmissivity decreases with particle number more rapidly than exponential, a straight line $1 - NA_\text{p}$. And if all the particles were on a line, the transmissivity would be $1 - A_\text{p}$ for all $N > 0$. From this we conclude that Eq. (2.7) is applicable to a perfectly random distribution of absorbers. If the positions of absorbers are correlated, the exponential law is called into question, although we know of no experimental evidence that this ever happens to an appreciable extent with atmospheric molecules or particles.

Now we are better able to understand the origins of the spectral emissivity curve for a uniform layer of moist air (Fig. 1.11). We obtained this curve using Eq. (2.24) for $\vartheta = 0$ with κ given by Eq. (2.31). Thus the frequency dependence of the emissivity originates in that of the absorption cross section of the water molecule (Fig. 2.9). We can write the normal emissivity of a uniform layer of moist air of thickness h as

$$\varepsilon_\text{n} = 1 - \exp(-Nh\sigma_\text{a}), \tag{2.34}$$

with

$$Nh = \frac{\rho_\text{w}d}{m_\text{w}} = 3.3 \times 10^{22}d, \tag{2.35}$$

where ρ_w is the density of liquid water, m_w the mass of a water molecule, and d the depth of a liquid water layer that would result if all the water vapor molecules in the moist air were

condensed to liquid. The quantity $N h \sigma_a$ is the absorption optical depth τ_a (see Sec. 5.2), which for $d = 1\,\mathrm{cm}$ is $3.3 \times 10^{22}\,\sigma_a$, with the cross section in cm^2.

The absorption cross section of water vapor (Fig. 2.9) varies by more than a factor of 1000 over the range of infrared frequencies shown. At the peaks, the cross section is sufficiently large that the corresponding emissivity is almost 1, at the troughs, much less than 1 (Fig. 1.11). The absorption cross section is not strictly a property of a molecule but depends on interactions with neighboring molecules, even in the gas phase. The positions of the peaks do not change so much as their widths, which is not especially difficult to account for (as we do subsequently). What Fig. 2.9 shows is the absorption coefficient divided by N, which defines the effective absorption cross section, a property of the molecule and its environment, which for our calculations is a total pressure of 1 atm and a temperature of $20\,^{\circ}\mathrm{C}$. To interpret the frequency dependence of the cross section we must dig deeper. To do so we need the right shovel, the rudiments of complex variables, which we turn to next.

2.5   The ABCs of Complex Variables

Before proceeding we digress briefly on the rudiments of complex variables, the minimum necessary to understand what follows. For many readers this will be a review.

The history of mathematics is to a large extent the history of the evolution of the concept of number, from the integers, to rational numbers (ratios of integers), to irrational numbers (such as $\sqrt{2}$) and finally to complex numbers. We may define a complex number z as the ordered pair

$$z = (x, y), \tag{2.36}$$

where x and y are just old-fashioned numbers. By ordered pair is meant that (x, y) is not the same as (y, x); x is called the *real* part of z, y the *imaginary* part, sometimes written

$$x = \Re\{z\},\ y = \Im\{z\}. \tag{2.37}$$

The terms "real" and "imaginary" are archaic, relics from a time when our ancestors were nervous about admitting such mathematical unicorns as complex numbers into the mathematical zoo. The rule for addition of complex numbers follows naturally

$$z_1 + z_2 = (x_1 + x_2, y_1 + y_2) \tag{2.38}$$

whereas that for multiplication does not:

$$z_1 z_2 = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1). \tag{2.39}$$

That is, even if you were innocent of any knowledge of complex numbers, if asked to add them, you'd likely come up with Eq. (2.38). Any complex number $(x, 0)$ with zero imaginary part can be written simply as x. We also have from the addition rule [Eq. (2.38)]

$$(x, y) = (x, 0) + (0, y) \tag{2.40}$$

and from the multiplication rule [Eq. (2.39)]

$$(0, y) = (0, 1)(y, 0). \tag{2.41}$$

If the symbol i denotes (0, 1), any complex number can be written simply as

$$z = x + iy. \tag{2.42}$$

The quantity i has the peculiar property that its square is -1,

$$i^2 = -1, \tag{2.43}$$

which is why it was deemed "imaginary." When a proper (real) number, positive or negative, is squared the result is always a positive number.

The ordered pair (x, y) can be represented geometrically as a point in a plane, called the complex plane, with rectangular coordinates x and y or, equivalently, as a displacement vector in this plane spanning the points (0, 0) and (x, y). We can transform rectangular coordinates to plane polar coordinates:

$$x = a\cos\vartheta, \ y = a\sin\vartheta, \tag{2.44}$$

where a is the length of the vector representing the complex number z and ϑ is the angle between this vector and the x-axis in the counterclockwise sense. Thus any complex number can be written

$$z = a\cos\vartheta + ia\sin\vartheta, \tag{2.45}$$

where a is called the *modulus* of z and ϑ its *argument*. The *complex conjugate* of any complex number $z = (x, y) = x + iy$, often written as z^*, is $(x, -y) = x - iy$. The product of a complex number with its complex conjugate is always non-negative:

$$zz^* = x^2 + y^2 = a^2 = |z|^2. \tag{2.46}$$

We may define

$$\exp(i\vartheta) = \cos\vartheta + i\sin\vartheta. \tag{2.47}$$

If we differentiate the right side of Eq. (2.47) with respect to ϑ we get the same result as when we take the derivative of the left side in the usual way treating i as a constant:

$$\frac{d}{d\vartheta}(\cos\vartheta + i\sin\vartheta) = -\sin\vartheta + i\cos\vartheta = i(\cos\vartheta + i\sin\vartheta) = \frac{d}{d\vartheta}\exp(i\vartheta). \tag{2.48}$$

And similarly for derivatives of all orders as well as integrals. Multiplication yields

$$\begin{aligned}
\exp(i\vartheta_1)\exp(i\vartheta_2) &= \cos\vartheta_1\cos\vartheta_2 - \sin\vartheta_1\sin\vartheta_2 + i(\cos\vartheta_1\sin\vartheta_2 + \cos\vartheta_2\sin\vartheta_1) \\
&= \cos(\vartheta_1 + \vartheta_2) + i\sin(\vartheta_1 + \vartheta_2) \\
&= \exp\{i(\vartheta_1 + \vartheta_2)\}.
\end{aligned} \tag{2.49}$$

Equations (2.48) and (2.49) tell us that whenever we encounter the complex number $\cos\vartheta + i\sin\vartheta$ we can write it succinctly as $\exp(i\vartheta)$, and then treat this entity just as we would the exponential function of a real variable. Now we can write an arbitrary complex number in *polar form* as

$$z = a\exp(i\vartheta), \tag{2.50}$$

which for our purposes usually will be the most convenient form. Addition of complex numbers can be interpreted geometrically as the addition of vectors in the complex plane, whereas multiplication of complex numbers

$$z_1 z_2 = a_1 a_2 \exp\{i(\vartheta_1 + \vartheta_2)\} \tag{2.51}$$

can be interpreted as stretching (or shrinking) and rotating vectors. In particular, multiplying any complex number by i doesn't change its length (modulus) but rotates it (counterclockwise) by $\pi/2$ because $\exp(i\pi/2) = i$.

The real part of the sum of complex numbers is the sum of its real parts

$$\Re\{z_1 + z_2\} = \Re\{z_1\} + \Re\{z_2\} \tag{2.52}$$

whereas this is, in general, not true of the product of complex numbers:

$$\Re\{z_1 z_2\} \neq \Re\{z_1\}\Re\{z_2\}. \tag{2.53}$$

Equation (2.52) is the simplest example of a linear operation. Differentiation is another. That is, if f_1 and f_2 are arbitrary differentiable functions of the real variable t then

$$\frac{d}{dt}\{f_1 + f_2\} = \frac{df_1}{dt} + \frac{df_2}{dt}. \tag{2.54}$$

And this is true for all orders of differentiation as well as integration (i.e., the integral of a sum is the sum of integrals). Moreover, we can interchange the order of taking the real part and any other linear operation. For example, if f is a complex-valued function of the real variable t, then

$$\Re\left\{\frac{df}{dt}\right\} = \frac{d}{dt}\Re\{f\}. \tag{2.55}$$

Equation (2.55) is just one example of a general rule:

$$\Re\{Lf\} = L\Re\{f\}, \tag{2.56}$$

where L is *any* linear operator (e.g., differentiation and integration). A consequence of Eq. (2.56) is that if a physical quantity is governed by a linear equation, we can find its complex-valued solution, then take its real part.

Make no mistake about it, all physical quantities are real numbers. So at this point you may be wondering why anyone would be so crazy as to find complex solutions to physical problems. Although lugging around what appears to be excess baggage (i.e., the imaginary

part of a solution), only to toss it on the scrap heap, may seem to be the height of foolishness, the small effort entailed in mastering the rudiments of complex variable theory is an enormous labor-saving device. Although this can be appreciated fully only with experience, a few stories support our claim.

Determining reflection and transmission because of an optically smooth, infinite planar boundary between two arbitrary optically homogeneous and infinite media is a problem in electromagnetic theory, the governing equations for which are linear. For any state of polarization of a plane wave incident on an interface at any angle, we can find the reflected and transmitted electric and magnetic fields relative to the incident fields. The ratios that result from these solutions are complex quantities because we obtained the complex solution to the field equations, called the *Fresnel coefficients* (see Sec. 7.2). They are compact, easy to remember, and take only a line or two to write down. The corresponding real reflectivity and transmissivity, which are ratios of irradiances, are obtained by multiplying these complex ratios by their complex conjugates, which yields real numbers. The Fresnel coefficients are readily programmed using a computer language that allows for operations on complex numbers. The resulting program takes a few lines. We once rewrote the Fresnel coefficients so that the reflectivity and transmissivity were expressed entirely in terms of real variables. The derivation took about 20 pages of close-packed algebra, many hours of work, lots of checking and rechecking. And the result was a cumbersome set of complicated equations, difficult to remember. Moreover, the computer program required to obtain numbers from these equations was at least ten times longer than a program based on complex arithmetic. And we began with the complex solution to the problem. Had we insisted from the outset on not using complex solutions, we might still be trying to obtain the real Fresnel coefficients. Keep in mind that this reflection and transmission problem is perhaps the simplest in electromagnetic theory. A more complicated planar problem is reflection and transmission by a single film and, more complicated yet, by multilayer films. Solving any of these problems without using complex solutions to the field equations would be a nightmare.

A vastly more complicated problem is determining the scattered field when an arbitrary homogeneous sphere is illuminated by a plane wave (discussed in Sec. 3.5). The resulting solution is an infinite series of terms, each of which is a complicated complex-valued function. Many years ago we received a telephone call from a man who proposed to take this complex solution and express it entirely in real terms so that he could then write a computer program using a language that could not do complex arithmetic. We asked him if he had committed some grievous sin for which he felt the need to atone by undertaking this terrible penance. He laughed. He was determined. A week later he called back, chastened: "You were right. It's impossible." And again, he began with a solution that had been obtained by way of complex-valued functions. Had he attempted to solve the scattering problem in a state of willful ignorance of complex numbers, he probably still would be slaving away.

The moral of these stories is that anything you can do with complex solutions to linear problems you can do without them but often you would be crazy to try. The same can be said for vectors. The payoff for learning the rudiments of vector analysis is that you can solve problems that without vectors would be horrible. Vectors are not necessary but they sure are useful. Take a look at old books on hydrodynamics and electromagnetic theory, ones from which vectors are absent, and you'll see what we mean.

Before proceeding, we need to issue just one word of caution. The only time we risk getting into trouble using complex solutions to (real) problems is when we need to perform nonlinear operations (e.g., multiplication) on solutions. To do this we first have to take real parts.

2.6   Interpretation of the Molecular Absorption Coefficient

We could content ourselves with accepting the absorption coefficient as a quantity obtained by transmission measurements and ask no questions about its origin. But such questions arise when we want to know why the absorption coefficient for the same material can be so different at different frequencies or why the absorption coefficient at a given frequency can be so different for different materials. To answer such questions we begin with the simplest possible example, a harmonic oscillator of mass m and positive charge e acted on by a time-harmonic (monochromatic) electromagnetic wave of frequency ω. If $\mathbf{E}$ is the electric field of this wave, the force on the oscillator is $e\mathbf{E}$. You may wonder why, if light is an electro*magnetic* wave, we ignore the force exerted by the magnetic field. The reason is that this force is negligible compared with the electric force if the velocity of the oscillator is small compared with the free-space speed of light.

A time-harmonic electric field has the form $\mathbf{E}_0 \cos \omega t$. According to Eq. (2.47)

$$\exp(-i\omega t) = \cos \omega t - i \sin \omega t, \tag{2.57}$$

and hence the electric field is the real part of the complex electric field:

$$\mathbf{E}_0 \cos \omega t = \Re\{\mathbf{E}_0 \exp(-i\omega t)\}. \tag{2.58}$$

This electric field is the complex representation of the real field of interest. As noted in the previous section, we may deal with complex representations of real quantities as long as we restrict ourselves to linear operations. Note that we could change i to $-i$ in Eq. (2.58) and the result would be the same. There are two conventions for the sign of the complex representation of time-harmonic fields. It doesn't matter which convention you choose as long as you stick with it. If you switch conventions in mid stream you may drown.

The position of the oscillator relative to the origin of some fixed coordinate system is $\boldsymbol{\xi}$. We assume that the oscillator has an equilibrium (no force) position $\boldsymbol{\xi}_0$ when not acted on by an external electric field. We also assume that the oscillator is acted on by a restoring force proportional to the displacement from equilibrium and by a dissipative force proportional to its velocity. Thus the equation of motion for the oscillator is

$$m\frac{d^2\boldsymbol{\xi}}{dt^2} = -K(\boldsymbol{\xi} - \boldsymbol{\xi}_0) - b\frac{d\boldsymbol{\xi}}{dt} + e\mathbf{E}, \tag{2.59}$$

where K and b are constants. If we denote by $\mathbf{x}$ the displacement from equilibrium, Eq. (2.59) can be written

$$\frac{d^2\mathbf{x}}{dt^2} = -\omega_0^2\mathbf{x} - \gamma\frac{d\mathbf{x}}{dt} + \frac{e}{m}\mathbf{E}, \tag{2.60}$$

where $\omega_0^2 = K/m$ and $\gamma = b/m$. Because $\mathbf{x}$ satisfies a linear differential equation we can allow $\mathbf{x}$ to be complex and take its real part if needed.

The time-harmonic electric field is given by Eq. (2.58). After a sufficiently long time $\mathbf{x}$ is also time-harmonic with the same frequency as the electric field: $\mathbf{x} = \mathbf{x}_0 \exp(-i\omega t)$; this is the steady-state solution. When this assumed form of the solution is substituted in Eq. (2.60) we obtain for the complex amplitude

$$\mathbf{x}_0 = \frac{e}{m} \frac{\mathbf{E}_0}{\omega_0^2 - \omega^2 - i\gamma\omega}. \tag{2.61}$$

The instantaneous rate P at which work is done on the oscillator by the external field is

$$P = \mathbf{F} \cdot \frac{d\boldsymbol{\xi}}{dt} = e\mathbf{E} \cdot \frac{d\mathbf{x}}{dt}. \tag{2.62}$$

Here we have to be careful. Taking the dot (or scalar) product of vectors is not a linear operation. So we have to find the real displacement and take its scalar product with the real field. If we multiply and divide Eq. (2.61) by the complex conjugate of the denominator in this equation we obtain

$$\mathbf{x} = \frac{e}{m} \mathbf{E}_0 \frac{\exp\{-i(\omega t - \phi)\}}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}, \tag{2.63}$$

where

$$\sin\phi = \frac{\gamma\omega}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}, \quad \cos\phi = \frac{\omega_0^2 - \omega^2}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}}. \tag{2.64}$$

The quantity ϕ is the *phase difference*, or simply the phase, between the amplitude of the oscillator and the force that drives it. When the phase difference is not zero, the maximum displacement of the oscillator does not occur at the same time as that of the electric field.

Now take the time derivative of the displacement [Eq. (2.63)] and multiply its real part by the real part of the electric field to obtain

$$P = \frac{e^2}{m} E_0^2 \frac{\omega}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}} \{\cos^2\omega t \sin\phi + \sin\omega t \cos\omega t \cos\phi\}. \tag{2.65}$$

This is the *instantaneous* rate at which work is done on the oscillator. But more often we are interested in the *time average* of P, denoted by $\langle P \rangle$, because detectors cannot respond instantaneously to the high frequencies of visible and near-visible radiation. For example, for a detector to respond instantaneously, even to comparatively sluggish $100\,\mu\mathrm{m}$ radiation, would require a response time of about 10^{-11} s.

Because we were critical in Section 1.4 of how the term "average" is bandied about without qualification, we had better make it clear that by time average here we mean a uniformly weighted average over the averaging interval. That is, if $f(t)$ is a time-dependent function, its average between times t_1 and t_2 is

$$\langle f \rangle = \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} f \, dt. \tag{2.66}$$

The time average of the square of the cosine over a time interval much larger than the period is approximately 1/2 and the average of the product of the sine and cosine is zero. Thus from Eq. (2.65) we have

$$\langle P \rangle = \frac{e^2}{2m} E_0^2 \frac{\omega \sin \phi}{\sqrt{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}}. \tag{2.67}$$

To do work on the oscillator requires energy. Electromagnetic fields are ascribed properties such as energy and momentum (but apparently not mass), which can be thought of as carried by photons. Thus $\langle P \rangle$ may be interpreted as the rate W_a at which energy is absorbed from the incident electromagnetic field by the oscillator. For $\langle P \rangle$ to be nonzero requires a nonzero phase difference ϕ, which in turn requires that dissipation (γ) be nonzero. Because the irradiance of the incident (or exciting) electromagnetic wave is proportional to the square of the amplitude of its electric field, it follows from Eq. (2.28) that $\langle P \rangle / E_0^2$ is, except for a constant factor, the frequency-dependent absorption cross section of the oscillator.

We may combine Eqs. (2.64) and (2.67) to obtain

$$\langle P \rangle = W_a = \frac{e^2}{2m} E_0^2 \frac{\gamma \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}. \tag{2.68}$$

Maximum absorption occurs at ω_0, the frequency at which the derivative of W_a with respect to ω is zero. This frequency is often called the *resonant* or *natural frequency*, the frequency at which the oscillator, subject to no external forces, would oscillate forever in the absence of dissipation if displaced from its equilibrium position and released. W_a and hence the absorption cross section at the resonant frequency are proportional to $1/\gamma$. To remind ourselves that we are interested in the frequency dependence of absorption cross sections we can rewrite Eq. (2.68) as

$$\sigma_a = \gamma \sigma_{am} \frac{\gamma \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2 \omega^2}, \tag{2.69}$$

where σ_{am} is the maximum absorption cross section. In the high-frequency limit ($\omega \gg \omega_0$) the absorption cross section decreases as $1/\omega^2$; in the low-frequency limit ($\omega \ll \omega_0$) it increases as ω^2 provided that for both limits $\gamma \ll \omega_0$. The frequencies at which absorption falls to one-half its maximum value are obtained from Eq. (2.69):

$$\omega_{1/2} = \sqrt{\omega_0^2 + \gamma^2/4} \pm \gamma/2. \tag{2.70}$$

Equation (2.69) is the simplest example of an *absorption line*, a range of frequencies over which absorption is comparatively high. The *width* γ of the line, the range of frequencies for which absorption is at least one-half the maximum, follows from Eq. (2.70).

We are almost always interested in lines that are narrow in the sense that the width is much less than the resonant frequency ($\gamma \ll \omega_0$), in which instance Eq. (2.69) is approximately

$$\sigma_a = \frac{\gamma \sigma_{am}}{4} \frac{\gamma}{(\omega_0 - \omega)^2 + \gamma^2/4} \tag{2.71}$$

and Eq. (2.70) is approximately

$$\omega_{1/2} = \omega_0 \pm \gamma/2. \tag{2.72}$$

The frequency-dependent function multiplying the invariant quantity $\gamma\sigma_{am}/4$ is sometimes called the *Lorentz line shape*. Because its width at half maximum is γ, whereas its peak is $4/\gamma$, this suggests that the area

$$\gamma \int_0^\infty \frac{d\omega}{(\omega_0 - \omega)^2 + \gamma^2/4} \tag{2.73}$$

underneath the profile is independent of γ. By a transformation of variable the integral becomes

$$\gamma \int_{-\omega_0}^\infty \frac{d\mu}{\mu^2 + \gamma^2/4}. \tag{2.74}$$

The antiderivative of the integrand is

$$\frac{2}{\gamma} \tan^{-1} \frac{2\mu}{\gamma}, \tag{2.75}$$

and hence Eq. (2.74) is approximately 2π provided that $\gamma \ll \omega_0$, which confirms our expectation that the integral of the absorption cross section is independent of γ. As the width of the line decreases its peak increases in such a way that the area underneath the line is constant.

What does all this have to do with molecules? By definition a molecule is a *bound* object: its constituent electrons and nuclei hang together and their range of motion is limited. Similarly, the harmonic oscillator is bound by the spring constant (restoring force) K. Although molecules are electrically neutral (unless ionized), they are composed of equal numbers of oppositely charged electrons and protons. Our molecular oscillator has charge e, so lurking in the background is an equal and opposite charge $-e$. We can take the motion of e to be relative to $-e$. Two equal and opposite charges separated by a distance d is called a *dipole*; its *dipole moment* $\mathbf{p}$ is e times the position vector of the positive relative to the negative charge. Thus our oscillator is a dipole with an oscillating (or vibrating) dipole moment $\mathbf{p} = e\mathbf{x}$. Similarly, molecules have dipole moments, either *induced* by the exciting field or *permanent* as a consequence of their structure or both. Although molecules are electrically neutral, their centers of positive and negative charge may not coincide, either because an external electric field pushes the centers of positive and negative charge in opposite directions or because this charge separation is built into the molecule. Here we consider only induced dipole moments; that is, $\mathbf{x}$ vanishes when the external electric field vanishes. Moreover, $\mathbf{p}$ lies along the direction of the field, which therefore exerts no torque on the dipole. But a permanent dipole moment in an electric field experiences a torque and hence rotates. In general, a dipole both vibrates and rotates.

A harmonic oscillator driven by an external electric field leads to a simple classical model of the frequency dependence of absorption of electromagnetic radiation by individual molecules. To good approximation a molecule can be considered to be a dipole, and this is true even for sufficiently small particles (in Sec. 3.4 we elaborate on what is meant by "sufficiently").

Before proceeding we pause to contemplate the form of the resonant frequency

$$\omega_0^2 = \frac{K}{m}, \tag{2.76}$$

which tells us something about the origin of all oscillations. This resonant frequency is the ratio of a restoring force (K) to inertia (m). This is a general result. Oscillations of any mechanical system usually can be traced to a struggle between a restoring force and inertia. For example, a simple pendulum oscillates because of the opposition between the restoring force of gravity and the inertia of the pendulum. Oscillations of a plucked string (see Sec. 3.3) occur because of opposition between the restoring force provided by the tension in the string and its inertia. Whenever you encounter an expression for the resonant frequency of a mechanical system, examine it carefully and try to interpret it as a consequence of a struggle between a restoring force and inertia. For an atom the restoring force is the attraction between the opposite charges of the induced dipole and the inertia is from the mass of the electrons associated with the dipole (Sec. 3.4.9).

2.6.1   Why the Obsession with Harmonic Oscillators?

At first glance, the amount of space devoted in textbooks to harmonic oscillators seems out of proportion to their importance. Yet the reason for this emphasis is that under fairly general conditions *any* mechanical system can be considered to be a collection of independent harmonic oscillators.

Consider N interacting point masses. For simplicity, we specify the positions of the point masses by their rectangular cartesian coordinates, denoted by ξ_i ($i = 1, 2, \ldots, 3N$). Suppose that the point masses exert position-dependent forces on each other that can be derived from a potential V, and suppose further that the system has an equilibrium configuration ($\xi_1^0, \xi_2^0, \ldots$). When the coordinates take these values the net force on any point mass is zero. We expand the potential in a Taylor series about the equilibrium position:

$$V = V_0 + \sum_i \left(\frac{\partial V}{\partial \xi_i}\right)_0 (\xi_i - \xi_i^0) + \frac{1}{2} \sum_i \sum_j \left(\frac{\partial^2 V}{\partial \xi_i \partial \xi_j}\right)_0 (\xi_i - \xi_i^0)(\xi_j - \xi_j^0) + \ldots \tag{2.77}$$

By definition the equilibrium position is that for which all interaction forces vanish:

$$\left(\frac{\partial V}{\partial \xi_i}\right)_0 = 0. \tag{2.78}$$

V_0 is a constant, so can be omitted because its derivatives do not contribute to the forces. If we truncate the series in Eq. (2.77) at the quadratic term, we obtain the potential in the *harmonic approximation*:

$$V \approx \frac{1}{2} \sum_i \sum_j K_{ij}(\xi_i - \xi_i^0)(\xi_j - \xi_j^0), \tag{2.79}$$

where

$$K_{ij} = K_{ji} = \left(\frac{\partial^2 V}{\partial \xi_i \partial \xi_j}\right)_0. \tag{2.80}$$

This approximation, also called the small-amplitude approximation, is better the smaller the departure from the equilibrium position.

Denote by x_i the displacement $\xi_i - \xi_i^0$. By the definition of a potential, the force component F_j is

$$F_j = -\frac{\partial V}{\partial \xi_j}, \tag{2.81}$$

which from Eq. (2.79) is

$$F_j = -\sum_i K_{ji}(\xi_i - \xi_i^0) = -\sum_i K_{ji}x_i. \tag{2.82}$$

With this equation and the assumption of a damping force

$$-b_i\frac{d\xi_i}{dt} = -b_i\frac{dx_i}{dt} \tag{2.83}$$

associated with each velocity component, the equations of motion for the system of point masses are

$$m_j\frac{d^2\xi_j}{dt^2} = m_j\frac{d^2x_j}{dt^2} = -\sum_i K_{ji}x_i - b_j\frac{dx_j}{dt} \; (j = 1, 2, \ldots, 3N). \tag{2.84}$$

Although there are $3N$ distinct coordinates, there are only, at most, N distinct masses. Because these equations are linear with constant coefficients we can express them in compact form by defining a column matrix $\mathbf{x}$, the $3N$ elements of which are x_j:

$$\mathbf{M}\frac{d^2\mathbf{x}}{dt^2} = -\mathbf{K}\mathbf{x} - \mathbf{B}\frac{d\mathbf{x}}{dt}, \tag{2.85}$$

where $\mathbf{K}$ is a symmetric $3N \times 3N$ matrix with elements K_{ij}, $\mathbf{M}$ is a diagonal matrix with elements m_j, and $\mathbf{B}$ is a diagonal matrix with elements b_j. Multiply both sides of Eq. (2.85) by the inverse matrix $\mathbf{M}^{-1}$, a diagonal matrix the elements of which are the inverse masses:

$$\frac{d^2\mathbf{x}}{dt^2} = -\mathbf{M}^{-1}\mathbf{K}\mathbf{x} - \mathbf{M}^{-1}\mathbf{B}\frac{d\mathbf{x}}{dt} = \mathbf{K}'\mathbf{x} - \mathbf{B}'\frac{d\mathbf{x}}{dt}. \tag{2.86}$$

Now assume that we can find a linear transformation

$$\mathbf{x} = \mathbf{A}\mathbf{u} \tag{2.87}$$

such that both $\mathbf{A}^{-1}\mathbf{K}'\mathbf{A}$ and $\mathbf{A}^{-1}\mathbf{B}'\mathbf{A}$ are diagonal, where $\mathbf{A}^{-1}$ denotes the inverse transformation. With the transformation Eq. (2.87), Eq. (2.86) becomes

$$\frac{d^2\mathbf{u}}{dt^2} = -\mathbf{\Omega}\mathbf{u} - \mathbf{\Lambda}\frac{d\mathbf{u}}{dt}, \tag{2.88}$$

with

$$\mathbf{A}^{-1}\mathbf{K}'\mathbf{A} = \mathbf{\Omega}, \; \mathbf{A}^{-1}\mathbf{B}'\mathbf{A} = \mathbf{\Lambda}, \tag{2.89}$$

where $\mathbf{\Omega}$ and $\mathbf{\Lambda}$ are diagonal matrices. Thus the $3N$ coupled equations of motion [Eq. (2.85)] become $3N$ uncoupled equations

$$\frac{d^2 u_j}{dt^2} = -\omega_j^2 u_j - \Gamma_j \frac{du_j}{dt} \ (j = 1, 2, \ldots, 3N), \tag{2.90}$$

where ω_j^2 are the elements of $\mathbf{\Omega}$ and Γ_j are the elements of $\mathbf{\Lambda}$.

Do these equations of motion look familiar? They should. They are the equations of motion of $3N$ *independent* harmonic oscillators. The frequencies ω_j are called the *normal* (or *characteristic*) *frequencies* of the system and the coordinates u_j are called the *normal coordinates*. Each independent oscillation is called a *normal mode*, and the general motion of the entire system is, from Eq. (2.87), a superposition of normal modes.

By purely mathematical trickery we reduced a system of interacting point masses to a system of non-interacting harmonic oscillators. This reduction requires for its validity that the system have a position of equilibrium the departures from which are sufficiently small. With these assumptions, *every* mechanical system behaves as if it were a set of independent harmonic oscillators of different frequencies. The system can range from a single water molecule composed of two hydrogen atoms and an oxygen atom to an entire building.

We once attended a colloquium given by a civil engineer interested in the forces that turbulent winds exert on buildings. He claimed that changes in materials and designs had altered the natural frequencies of buildings. We don't know much about turbulence, even less about structural engineering, but all we needed to understand the gist of his talk was Eqs. (2.76) and (2.90).

Reduced to its simplest elements, a building (or any structure) is a system of independent harmonic oscillators (e.g., weights and springs). The natural frequencies of this system depend on the stiffness of the structure (K) and the density of materials (m). If either or both are changed, the natural frequencies change. The time-varying turbulent wind field can be decomposed into components with different frequencies, and the aim of the designer is to be sure that these frequencies, especially the dominant ones, do not coincide with the natural frequencies of the structure. As we have seen, the amplitude of an oscillation depends on the frequency of the driving force.

2.7 Classical versus Quantum-Mechanical Interpretation of Absorption

What once was, and perhaps still is, the standard treatise on atomic spectra, by Condon and Shortley, fills 432 pages of text. Herzberg's treatises fill 581 pages for diatomic molecules, 538 pages for polyatomic molecules, and 670 pages for the electronic spectra of polyatomic molecules. Townes and Schawlow devote 648 pages to microwave spectroscopy. And the physical strain of just lifting these nearly 3000 pages is as nothing compared to the mental strain of absorbing them. Add to these tomes the dozens if not hundreds of books on quantum mechanics and modern physics. Spectroscopy is the science of details. Indeed, this is the source of the almost religious awe in which quantum mechanics is held by physicists. A theory capable of ordering – if not explaining – so many details is akin to magic.

This somewhat depressing inventory of details, systematized by a theory that, as Richard Feynman has so often been quoted as saying, "nobody understands", signals that what follows can only scratch the surface of a vast subject. Given the few pages we can devote to it, our aim must be the modest one of conveying the bare bones of what students of atmospheric science should know about the infrared spectra of molecules that inhabit the atmosphere.

According to classical mechanics, the mechanics of Newton, a mechanical system can have a continuous set of energies (kinetic plus potential). This seems so self-evidently true that it is rarely stated explicitly and even more rarely questioned. But in reality, energy is like dollar bills: you can have one bill in your wallet or two or three, but never 1.523. Dollar bills are *quantized*. And so is energy, but this is not evident until we consider mechanical systems on a scale not directly accessible to our senses. We do not live at this microscopic (or atomic) scale, so we have no right to expect that the physics of ordinary macroscopic objects is valid there. This is much like moving from one country to another. In one society, certain rules of behavior and customs are taken for granted. But when you enter a different society, you sometimes discover that many of the familiar rules no longer apply. In India, Pakistan, and Sri Lanka, people eat with their fingers. To do otherwise would seem unnatural to the inhabitants of these countries. But in Western countries, eating with your fingers is considered to be extremely impolite. Children who do this are told in no uncertain terms to desist.

The differences between the rules of behavior of macroscopic and microscopic objects are considerably greater than the differences between those of the inhabitants of New York penthouses and the inhabitants of huts in the Amazon jungle.

If the laws of quantum mechanics must be taken on faith, consider that so must the laws of classical mechanics. Why $F = ma$? You learn this first as an axiom, and so what follows it may seem strange. But if you had learned quantum mechanics first, $F = ma$ might seem strange. Whatever you learn first sets the standard for what is normal.

Consider first a harmonic oscillator with natural frequency ω_0. According to classical mechanics this oscillator can have a continuous range of energies depending on its initial amplitude. But according to quantum mechanics the energies of this harmonic oscillator are quantized, having only the discrete set of energies

$$E_n = \hbar\omega_0 \left(n + \frac{1}{2} \right),$$ (2.91)

where $n = 0, 1, 2, \ldots$. This result seems to contradict common sense, but it really doesn't because common sense is based on our experience with macroscopic objects. The difference in energy between physically realizable (allowed) adjacent energy states of the harmonic oscillator is

$$\Delta E = \hbar\omega_0.$$ (2.92)

The oscillators we encounter at the macroscopic level might have natural frequencies as high as 100 Hz. This corresponds to an energy difference of 6.63×10^{-32} J. To get a feeling for how much energy this is, consider how much energy the eyebrow of a flea has when it falls from a flea on the back of a dog. A flea, which can be seen without a microscope, has dimensions of a millimeter or so, and hence the total volume of the flea is of order $10^{-9}\,\mathrm{m}^3$. The characteristic linear dimension of a flea's eyebrow is at least 100 times smaller than the

overall dimensions of the flea. Thus the volume of the flea's eyebrow is of order $10^{-15}\,\mathrm{m}^3$. We estimate the density of flea flesh to be that of water, $1000\,\mathrm{kg\,m}^{-3}$, which gives a mass of $10^{-12}\,\mathrm{kg}$ for the flea's eyebrow. Suppose that the flea resides on the back of a Great Dane, about 1 m high at the shoulders. The potential energy of the flea's eyebrow, relative to that at the ground, is about $10^{-11}\,\mathrm{J}$. As small as this potential energy is, it is still 10^{21} times greater than the difference between adjacent energy levels of a harmonic oscillator with natural frequency 100 Hz. Although the energies of macroscopic oscillators are quantized in principle, the spacing of energy *levels*, as they are called, is so small macroscopically that in practice the levels are continuous. We are forced to come to grips with the discreteness of energy levels only when we consider systems with very high natural frequencies. All else being equal, natural frequencies increase with decreasing mass [Eq. (2.76)], and hence the consequences of discrete energies are not negligible at the atomic scale.

With this preamble, consider absorption of electromagnetic energy by a single isolated oscillator from the classical and quantum-mechanical points of view. According to the classical analysis in Section 2.6, the rate at which power is absorbed by an oscillator from a time-harmonic electromagnetic wave of given amplitude depends on its frequency ω. Absorption is sharply peaked in a narrow range of frequencies, called an absorption line (or band), centered on the natural frequency of the oscillator. The width of the line is a consequence of damping of the oscillator.

Now consider the same process from a quantum-mechanical point of view. The incident monochromatic electromagnetic wave is considered to be a stream of photons, each with energy $\hbar\omega$. Absorption of electromagnetic energy is a consequence of absorption of photons. If the oscillator absorbs a photon, the energy of the oscillator must increase. But this increase can be only one of a set of discrete values. Unless the energy of the photon is equal to the difference between two energy levels of the oscillator, it cannot absorb the photon. This accounts for the narrowness of absorption lines.

2.7.1 Molecular Energy Levels

An electrically neutral atom consists of a nucleus composed of uncharged neutrons and positively charged protons bound together and surrounded by negatively charged electrons equal in number to the protons. Almost the entire mass of an atom resides in its nucleus, which is smaller by a factor of about 1000 than the atom as a whole. That is, the electrons, on average, are at distances from the nucleus large compared with its size.

A molecule is a collection of two or more atoms bound together. They can vibrate about their equilibrium positions and the molecule as a whole can rotate. A vibrating object has vibrational kinetic and potential energy; a rotating object has rotational kinetic energy. The center of mass of the molecule can move and it can interact with neighboring molecules. And electrons have potential and kinetic energies. These various modes of motion of a molecule can be decomposed *approximately* into *translational* (position and velocity of the center of mass of the molecule), *rotational*, *vibrational*, and *electronic* modes, and hence the total energy of the molecule is

$$E = E_{\mathrm{trans}} + E_{\mathrm{rot}} + E_{\mathrm{vib}} + E_{\mathrm{elec}}. \tag{2.93}$$

That this decomposition is only approximate is evident from the rotations and vibrations of two mass points connected by a spring. The rotational kinetic energy of this system depends on its moment of inertia, which in turn depends on the separation between the mass points. But if they vibrate, their separation changes, and hence so does the moment of inertia. Thus the two types of motion, rotation and vibration, are not completely decoupled. Although Eq. (2.93) often is a good approximation, don't be surprised when it fails to explain what is observed.

Each of the separate contributions to the total energy E in Eq. (2.93) is quantized except the translational kinetic energy of the center of mass of a molecule. A truly isolated molecule is an idealization. We are almost always faced with an ensemble of many molecules. Even in a low density gas, interactions between molecules are not completely negligible. Strictly speaking, translational energy is also quantized but the level separation is so small that it cannot be observed. To the extent that Eq. (2.93) is a good approximation we can therefore speak of the electronic energy levels of a molecule, its vibrational energy levels, and so on. The lowest allowed energy of any kind is called the *ground state*. All higher energy levels, or states, are called *excited states*. When a molecule absorbs a photon it is said to be excited into a higher-energy state or to undergo a transition from one energy state to another of higher energy. A molecule in an excited state can then spontaneously drop to a lower energy state accompanied by the emission of a photon equal in energy to the difference in energy levels, which underscores our previous assertions about absorption and emission being inverse processes. The Lorentz line shape [Eq. (2.71)] is just as valid quantum-mechanically as it is classically, but the terms in it are interpreted quite differently. Classically, ω_0 is the natural frequency of an oscillator and γ is a factor in a viscous damping term; quantum-mechanically, $\hbar\omega_0$ is the difference between two energy levels and $1/\gamma$ is the *lifetime* of the transition between them (i.e., the average time the molecule exists in the higher energy state).

The energy of a photon is

$$\hbar\omega = h\nu = \frac{hc}{\lambda} = hc\tilde{\nu}. \tag{2.94}$$

Except for the factor hc, a universal constant, the energy of a photon is inversely proportional to the wavelength of the associated wave. Thus we may express photon energies as inverse wavelengths, which are called *wavenumbers*, usually with units cm^{-1}. Why wave*number*? Because it is the number of waves in unit length. Spectroscopists are careless in the use of the symbol ν: sometimes it denotes frequency, sometimes wavenumber (inverse wavelength). And to make matters worse, the units of wavenumber are usually cm^{-1}, and so if hc is in SI units and $\tilde{\nu}$ in cm^{-1}, photon energy is $100hc\tilde{\nu}$. We try to be consistent and use ν for frequency, $\tilde{\nu}$ for wavenumber. A word of caution: $2\pi/\lambda$, often written as k or q, is sometimes called wavenumber.

A wavelength of $10\,\mu\mathrm{m}$ corresponds to a wavenumber of $1000\,\mathrm{cm}^{-1}$. Wavelengths in the middle of the visible spectrum correspond to wavenumbers of around $20,000\,\mathrm{cm}^{-1}$. Sometimes you will encounter statements about a photon energy being so many wavenumbers, 2000, say. Unless stated otherwise, this probably means $2000\,\mathrm{cm}^{-1}$.

Another quantity with dimensions of energy that keeps cropping up in all kinds of problems is $k_\mathrm{B}T$, which we can divide by hc so as to express it in the same units as wavenumbers. At typical terrestrial temperatures (300 K say), $k_\mathrm{B}T/hc$ is around $200\,\mathrm{cm}^{-1}$.

Consider now a gas of molecules in thermal equilibrium at temperature T. These molecules can collide and exchange energy. Indeed, a molecular collision is *defined* as an interaction between two or more molecules in which the energy of each molecule after the interaction is different from that before (total energy being conserved). The most probable kinetic energy of a gas molecule is $k_B T/2$ (see Sec. 1.2), so this is how much energy can be exchanged in a typical collision. Some molecules have greater energies than the most probable, some less, the probability of energy E being proportional to $\exp(-E/k_B T)$. (We omit the other factor dependent on energy because it does not vary so strongly). This quantity, called the *Boltzmann factor*, crops up in all kinds of problems. Indeed, it would not be an exaggeration to say that it, not love, is what makes the world go round. Rates of chemical reactions are determined by the Boltzmann factor. Our bodies are complicated, finely-tuned engines in which many chemical reactions are continuously taking place, their rates determined by the Boltzmann factor. A small change in temperature changes not only the absolute rates of these reactions but their relative rates as well. This is why our bodies constantly attempt to keep our deep core temperature at around $37\,°\text{C}$. An increase in this temperature of only 1–$2\,°\text{C}$ is sufficient to send us to bed wracked with pain and fever. An increase of a few more degrees might be fatal. When you cook food you are exploiting the Boltzmann factor. An egg will rot before it hardens at room temperature, but immerse it in water at $100\,°\text{C}$ and it is hard-boiled in minutes. A rough rule of thumb used by chemists is that chemical reaction rates double with every $10\,°\text{C}$ increase in temperature. A temperature increase from $20\,°\text{C}$ to $100\,°\text{C}$ corresponds to 8 doubling times, a factor of 256.

If a gas is in thermal equilibrium its molecules are distributed in their energy states. The ratio, on average, of the number N_j of molecules having energy E_j and the number of molecules N_i having energy E_i is approximately

$$\frac{N_j}{N_i} \approx \frac{\exp(-E_j/k_B T)}{\exp(-E_i/k_B T)} = \exp\{-(E_j - E_i)/k_B T\}. \tag{2.95}$$

Typically, the separation between electronic energy levels is around $10{,}000\,\text{cm}^{-1}$. Suppose that E_i corresponds to the ground electronic state and that E_j corresponds to the first excited electronic state. Because the difference ΔE between these two energy states is around $10{,}000\,\text{cm}^{-1}$, the ratio of the number of molecules with electrons in the first excited state to the number in the ground electronic state at 300 K is of order $\exp(-10{,}000/200) = \exp(-50) \approx 10^{-25}$. At typical terrestrial temperatures almost all molecules in a gas are in their ground electronic state. To populate excited electronic states would require a ten-fold or more increase in absolute temperature. Typical separations between adjacent vibrational energy levels are about $1000\,\text{cm}^{-1}$. Again, this is large compared with $k_B T$ for ordinary temperatures, and hence most gas molecules at these temperatures are in their vibrational ground states. Typical separations between rotational energy states are 10–$100\,\text{cm}^{-1}$, which are comparable with $k_B T$. Thus at ordinary temperatures many molecules are in excited rotational states.

Merely because molecules vibrate and rotate does not necessarily mean that they radiate (emit). Absorption, as we have seen, is the inverse of emission. It is perhaps easier to discuss spontaneous emission because it occurs in the absence of an external exciting field. One of the results of classical electromagnetic theory is that accelerated charges radiate electromagnetic

waves. A corollary of this is that for an electrically neutral charge distribution (e.g., molecule) to radiate, its dipole moment must change with time, either in magnitude or direction or both. A molecule with a permanent dipole moment is said to be *polar*. A common example is the water molecule (H_2O), which owes its permanent dipole moment to its asymmetry: it is composed of two atoms of one kind (hydrogen) and one atom of another (oxygen), which all do not lie on a line. As a consequence of its permanent dipole moment, a rotating water molecule radiates (emits).

The two most abundant molecules in the atmosphere are the diatomic molecules nitrogen (N_2) and oxygen (O_2). Both of these molecules are *homonuclear*, which is just a fancy way of saying that they are composed of identical atoms. Such molecules cannot have a permanent dipole moment. To have such a moment would require the center of positive charge to be associated with one atom and the center of negative charge to be associated with the other. But symmetry rules this out: both atoms are identical, and hence one cannot be positively and the other negatively charged. Although nitrogen and oxygen molecules can rotate, in so doing they do not radiate (much). These molecules also can vibrate. But again symmetry rules out the possibility of a changing dipole moment during this vibration. The two atoms must enter into the vibration in a symmetric way, and hence the dipole moment associated with one atom is equal and opposite to that associated with the other.

We have seen that the motions of a molecule can be expressed as a sum of normal modes, each with a characteristic frequency. These frequencies lie in the infrared and a radiating mode is called *infrared active*. A mode that does not radiate is called *infrared inactive*. The terms infrared active and inactive, which are familiar to infrared spectroscopists, are preferable to the popular but misleading term "greenhouse gas." Water vapor is infrared active; nitrogen and oxygen, for the most part, are infrared inactive. Greenhouse gases are produced by resident cats with digestive problems.

The normal modes of vibration of the water molecule are shown schematically in Fig. 2.17. There are three modes: the O-H (symmetric) stretching mode, the H-O-H bending mode, and the O-H (asymmetric) stretching mode. The corresponding normal frequencies are $\tilde{\nu}_1 = 3657.1 \, \text{cm}^{-1}$ ($2.73 \, \mu\text{m}$), $\tilde{\nu}_2 = 1594.8 \, \text{cm}^{-1}$ ($6.27 \, \mu\text{m}$), and $\tilde{\nu}_3 = 3755.9 \, \text{cm}^{-1}$ ($2.66 \, \mu\text{m}$). The O-H stretching modes are so named because the vibration occurs approximately along the O-H bond.

Keep in mind that these normal frequencies are for the isolated water molecule, water in the gaseous phase. When water vapor condenses to form liquid water, the positions of the absorption bands shift. And when liquid water freezes to form ice, there is yet another shift in the absorption bands. Nevertheless, the normal frequencies of the isolated water molecule are good guides to the approximate positions of absorption bands in the condensed phases. This is evident in Fig. 2.2. Note the large dips in the absorption length around $3 \, \mu\text{m}$ and $6 \, \mu\text{m}$. These are the consequence, even in the condensed phases, of vibrations of the isolated water molecule, which does not lose its identity completely when it condenses.

According to quantum mechanics, the vibrational energy levels of the water molecule are quantized:

$$E_n = h\nu_i \left(n + \frac{1}{2} \right), \tag{2.96}$$

where ν_i is any one of the three normal mode frequencies of the water molecule. On the basis of this equation alone we would expect absorption bands not only at the fundamental frequency ν_i but at overtones as well, integral multiples of the fundamental frequency: $2\nu_i, 3\nu_i, \ldots$. But this expectation is not quite borne out by experience. Equation (2.96) does not tell the entire story. We have to account for *selection rules*, rules that tell us which energy transitions are allowed. The different energy states described by Eq. (2.96) are also states of different angular momentum. The photon carries quantized angular momentum. Thus when a photon is absorbed by a molecule, and it undergoes a transition to a higher energy level, the transition must be such that angular momentum is conserved. Thus if one unit of angular momentum is annihilated (so to speak) when a photon is absorbed, the molecule must increase its angular momentum, and hence energy state, by the same amount. And similarly for emission: when a photon is emitted, the one unit of angular momentum created upon the birth of the photon must be compensated for by the same decrease in the angular momentum of the molecule. Remember that angular momentum is a vector, and hence the angular momentum of the emitted or absorbed photon and the change in angular momentum of the molecule must add vectorially to zero. This requirement of angular momentum conservation leads to the selection rule $\Delta n = \pm 1$. That is, transitions are allowed only for integral changes in the quantum number n in Eq. (2.96).

This selection rule is not absolute because it is based on the harmonic approximation (Sec. 2.6.1). Because the forces between atoms in the water molecule are not exactly harmonic, the previous section rule can be violated. That is, transitions corresponding to *overtones* of the fundamental frequencies and even *combinations* of different frequencies (e.g., $\nu_1 + 2\nu_2$, $\nu_1 - \nu_2$, etc.) are possible. These transitions are very weak, but not so weak as to be unobservable.

Although absorption drops precipitously from infrared to visible, it does not reach zero. Moreover, absorption over the visible spectrum is least in the blue–green and rises toward the red in both the liquid and solid phases. (This is evident in Fig. 2.2 but even more so in Fig. 5.12.) Water is a weak blue dye, and this is intrinsic, not the result of some vague impurities. This intrinsic selective absorption by water leads to observable consequences: the blue of the sea, of crevasses in glaciers, ice caves, and frozen waterfalls (see Sec. 5.3.1). What is the molecular mechanism for this blueness?

Overtones of fundamental vibration frequencies in the infrared and combinations of these frequencies make their presence felt, weakly, but observably in the visible. The experimental proof of the vibrational origin of the visible absorption spectrum of water was obtained several years ago by Chuck Braun and Sergei Smirnov, who published their results in a delightful paper, "Why is water blue?", one of those papers you must read before going to your grave. Braun and Smirnov measured the visible and near-visible absorption spectra of ordinary water (H_2O) and heavy water (D_2O). The vibration spectrum of a molecule depends on the masses of its constituent atoms [see Eq. (2.76)]. Because the nuclear mass of heavy water is greater than that of ordinary water we expect a shift (*isotope shift*) toward lower frequencies in the vibration spectrum of D_2O relative to that of H_2O. And this shift was indeed observed. In ordinary water there is a peak at about 750 nm, which shifts to about 1000 nm for heavy water. Braun and Smirnov did not calculate the magnitude of the isotope shift because of the greater mass of deuterium relative to that of hydrogen, so we did using Eq. (2.76) and found good agreement with the measured shift.

Rising absorption in the red gives ordinary water its bluish color, whereas absorption by heavy water is flat throughout the visible spectrum and begins to rise only well into the infrared.

When questions about visible absorption by water arise, chemists seem to be divided into two groups: those who adamantly deny that water has a visible absorption spectrum and those who admit that it does but are just as adamant that it is a consequence of hydrogen bonding.

You may think that we are joking about the first group, yet we could tell you many stories. One will suffice. Several years ago one of the authors gave a talk in Italy on colors in nature for an audience mostly of biological scientists. Afterwards, a photochemist accused him of being a criminal and pervert for telling innocent biologists that water has a visible absorption spectrum.

In chemistry, hydrogen bonding is the universal solvent. . . for ignorance. Whenever a puzzle arises and there is any liquid water in the neighborhood, hydrogen bonding dissolves the puzzle in the same way that Alexander untied the Gordian knot with his sword. Hydrogen bonding plays the same role that friction does in the undergraduate physics laboratory. Liquid water, however, is blue not *because* of hydrogen bonding but *despite* it. Hydrogen bonding in liquid water shifts all the infrared absorption bands, including their overtones and combinations, to lower frequencies compared with the gas phase.

Braun and Smirnov are of the opinion that the color of water is the only example of a color in nature that results from vibrational rather than electronic transitions. It is commonly believed that all colors in nature can be traced to electronic excitations. Water provides an example to the contrary.

Carbon dioxide is a linear, symmetric molecule. By linear is meant that its bonds lie on a straight line. By symmetric is meant that it is composed of a carbon atom flanked on either side by oxygen atoms. This symmetry implies that the carbon dioxide molecule does not have a permanent dipole moment, which precludes this molecule from having what is called a *rotational* absorption band: an absorption band associated with rotation of the molecule unaccompanied by vibrations. But the carbon dioxide molecule can vibrate in such a way that it has a changing dipole moment. This in turn implies that this molecule, when vibrating in one of its vibrational modes, also can rotate to give a changing dipole moment. So carbon dioxide has *vibration-rotational* bands.

Carbon dioxide has a bending mode with frequency $\tilde{\nu}_2 = 667.4\,\mathrm{cm}^{-1}$ (15 μm) and an asymmetric stretching mode with frequency $\tilde{\nu}_2 = 2349.2\,\mathrm{cm}^{-1}$ (4.25 μm). The symmetric stretching mode is infrared inactive. It is the vibration-rotational band near 15 μm that is the major player in the global warming scenario associated with increased carbon dioxide in the atmosphere.

Rotational energies also are quantized. For example, those of the *rigid rotator*, a massless rigid rod of length r with one end fixed, the other end attached to a mass m, and free to rotate are given by

$$E_J = \frac{\hbar^2 J(J+1)}{2I}, \; J = 0, 1, 2, 3, \ldots, \tag{2.97}$$

where $I = mr^2$ is the moment of inertia about the rotation axis. As with transitions of the harmonic oscillator, those of the rigid rotator are restricted by a selection rule: $\Delta J = \pm 1$.

Although molecules can rotate, and hence have rotational energy levels, they are never exactly describable as rigid rotators. Like the harmonic oscillator, the rigid rotator is an idealization, and hence this selection rule is not absolute. When we say with seeming absoluteness that a transition is forbidden, we mean that the transition probability is very small compared with that for an allowed transition but not identically zero. Quantum mechanics is a most permissive theory, allowing improbable events that according to classical mechanics are impossible.

You may wonder why we jumped directly to the quantum-mechanical rigid rotator instead of making a more gradual transition by way of the classical rotator. This is what we did for the harmonic oscillator, beginning with its classical equation of motion and then stating, but not proving, how the discrete energy levels of the quantum-mechanical oscillator are related to the classical natural frequency of the oscillator. Why didn't we follow the same approach for the rigid rotator? And why were we unable to find any authors who do or even bring up the subject?

As it happens, analysis of the classical harmonic oscillator is considerably less difficult than analysis of the quantum-mechanical oscillator. When we turn to the rigid rotator, this happy state of affairs is reversed: the quantum-mechanical rotator is more amenable to analysis. The classical rigid rotator has no natural frequency. It can spin at any angular speed. A simple pendulum does have a natural frequency for small-amplitude oscillations, but this frequency depends on the magnitude of an *external* force: gravity. Thus even if we could solve the classical equation of motion for a driven rigid rotator, we wouldn't acquire any fruitful analogies or equations we could carry with us into the quantum domain. And similarly for electronic energy levels. But with only a few exceptions, these energy levels are not of great relevance to the atmosphere. One exception is the Chappuis bands of ozone, which play an important role in the color of the zenith twilight sky (see Sec. 8.1.3).

2.8   Absorption by Molecules: The Details

The previous section laid out the bare bones of molecular absorption and its physical interpretation. In this section we put some flesh on these bones, a meal that requires considerable chewing and digestion. At the very least students of atmospheric science should be familiar with Figs. 2.12 and 2.13 and understand their implications. Everything in this section is aimed at explaining the strange shapes of the curves in these figures.

A beam of radiation transmitted by a gas undergoes attenuation by absorption if photons in the beam have energies equal to differences in two molecular energy levels for which the selection rules are obeyed. Molecules make transitions from lower to upper levels. But the reverse is also possible. Some molecules can be *induced* or *stimulated* to make transitions to a lower level, energy conservation therefore requiring that photons be emitted. These photons have the same energy and direction as those that induced them. Although this process may seem strange at first glance, remember that any process in which energy is conserved is possible. *Spontaneous emission* of photons by gas molecules, in contrast to this *induced emission*, is, of course, also possible. Indeed, by emission without qualification, we usually mean spontaneous emission.

Induced emission and absorption are opposite processes, the former increasing the radiant energy in a beam, the latter decreasing it, but all we can observe is the net result. Thus the

apparent (observed) absorption coefficient κ is less than it would be in the absence of induced emission. And we also might have to correct for spontaneous emission because a detector downstream of a source necessarily receives induced and spontaneous radiation from the gas as well as source radiation not absorbed by it.

2.8.1 Absorption versus Spontaneous and Induced Emission

Estimating κ in the presence of induced emission is a difficult but worthwhile task. Let E_ℓ and E_u ($E_u > E_\ell$) be two molecular energy levels, for which radiative transitions between levels are allowed. The energy levels of a molecule depend on its structure and on the forces exerted on it by its neighbors (no molecule lives alone) and these forces depend on the ever-changing distances between molecules. These interaction forces, often called collisions, result in perturbations of the energy levels of the isolated molecule, and hence in their difference, often expressed as a frequency $(E_u - E_\ell)/h$. This frequency for the *isolated* molecule is denoted by ν_0. The distribution of frequencies about ν_0, called a *line shape*, is narrow but its width is never identically zero (see Fig. 2.18). Because they are in motion, molecules are illuminated by radiation of frequency slighted shifted from that of the source, and emit radiation the frequency of which is also shifted (see Sec. 3.4.6). This *Doppler broadening* of the spectral line (centered on ν_0) is equivalent to perturbing energy level differences but is a consequence of relative motion rather than inter-molecular forces. In the lower atmosphere Doppler broadening of spectral lines is usually less than that resulting from interactions, often called *collisional broadening* or *pressure broadening*, a misleading term because the broadening has nothing fundamental to do with pressure. Inter-molecular force broadening is a more descriptive, but also more awkward, term.

Keep in mind that treating a gas as a collection of identical molecules each with the same but fuzzy or perturbed energy levels is an approximation. The only rigorous way to determine the energy levels of a gas would be to treat it as a system of molecules each interacting with all the others. This system as a whole has a set of energy levels. Alas, this is an insoluble many-body problem. So we settle for less and adopt the fiction that each molecule has the same set of fuzzy energy levels, the fuzziness specified by the line width. This is a good approximation for a gas because its molecules interact weakly, but fails completely when the gas is compressed into a liquid or a solid, for which the inter-molecular forces are dominant rather than perturbations. The absorption spectrum of sodium gas, for example, bears no resemblance to that of sodium metal.

N is the total number of molecules per unit volume, N_u in an upper state, N_ℓ in a lower state. Divide frequency space into M bins centered on frequency ν_0. Denote by n_{um} the number density of molecules with an upper energy level such that the difference between this level and a lower level lies in the mth bin. The rate of change of n_{um} because of spontaneous emission is assumed to be given by

$$\left(\frac{dn_{um}}{dt}\right)_s = -R_{s,um}n_{um},\tag{2.98}$$

where $R_{s,um}$ is the rate constant for spontaneous emission, for generality assumed to be possibly different for every transition, and s denotes spontaneous. Equation (2.98) has the form

of a first-order chemical kinetics equation based on the plausible assumption that the instantaneous rate at which something changes is proportional to the amount of that something at any instant. This is also the equation of radioactive decay. The frequency of the photon emitted is that associated with the mth bin. If $f_{\text{u}m}$ is the probability that a molecule has upper and lower states such that the energy difference lies in the mth bin, we can write Eq. (2.98) as

$$\left(\frac{dn_{\text{u}m}}{dt}\right)_{\text{s}} = -R_{\text{s},\text{u}m} f_{\text{u}m} N_{\text{u}}. \tag{2.99}$$

The rate of change of $n_{\text{u}m}$ because of induced emission must depend on the radiation environment of the molecule, and hence the rate equation must contain the number density $n_{\text{p},m}$ of photons with energies (frequencies) in the mth bin:

$$\left(\frac{dn_{\text{u}m}}{dt}\right)_{\text{e}} = -R_{\text{e},\text{u}m} f_{\text{u}m} N_{\text{u}} \frac{n_{\text{p},m}}{\Delta\nu}, \tag{2.100}$$

where e denotes induced emission and $\Delta\nu$ is the bin width. Note that the dimensions of the rate constant in this equation are different from those in Eq. (2.99).

The number density of lower states $n_{\ell m}$ decreases because of absorption of photons and, like induced emission, depends on the radiation environment. Thus the rate equation analogous to Eq. (2.100) is

$$\left(\frac{dn_{\ell m}}{dt}\right)_{\text{a}} = -R_{\text{a},\ell m} f_{\ell m} N_{\ell} \frac{n_{\text{p},m}}{\Delta\nu}, \tag{2.101}$$

where a denotes absorption and $f_{\ell m}$ is the probability that a molecule has lower and upper states such that the energy difference lies in the mth bin. The quantity $R_{\text{a},\ell m}$ is essentially the *Einstein absorption coefficient*, $R_{\text{e},\text{u}m}$ the *Einstein induced emission coefficient*, and $R_{\text{s},\text{u}m}$ the *Einstein spontaneous emission coefficient*. Einstein assumed in 1916 that the transition rates Eq. (2.100) and Eq. (2.101) are proportional to the spectral density $n_{\text{p},m}/\Delta\nu$, later proven by quantum mechanics when it had developed into a sufficiently detailed theory. We are content to accept Einstein's assumption. Ultimately, its justification is agreement with measurements.

Because only the product of a rate constant and an occupation probability occurs in Eqs. (2.99)–(2.101), we can fold any dependence of the rate constants on bin into the occupation probabilities and write

$$\left(\frac{dn_{\text{u}m}}{dt}\right)_{\text{s}} = -R_{\text{s}} f_{\text{s},\text{u}m} N_{\text{u}}, \tag{2.102}$$

$$\left(\frac{dn_{\text{u}m}}{dt}\right)_{\text{e}} = -R_{\text{e}} f_{\text{e},\text{u}m} N_{\text{u}} \frac{n_{\text{p},m}}{\Delta\nu}, \tag{2.103}$$

$$\left(\frac{dn_{\ell m}}{dt}\right)_{\text{a}} = -R_{\text{a}} f_{\text{a},\ell m} N_{\ell} \frac{n_{\text{p},m}}{\Delta\nu}. \tag{2.104}$$

The sum of Eqs. (2.102) and (2.103), times a constant factor, is the rate at which molecules lose energy by emission, whereas Eq. (2.104), times the negative of that same constant factor,

is the rate at which molecules gain energy by absorption. In equilibrium losses must balance gains and hence

$$\frac{\overline{N}_u}{\overline{N}_\ell} = \frac{R_a f_{a,\ell m}(n_{p,m}/\Delta\nu)}{R_s f_{s,um} + R_e f_{e,um}(n_{p,m}/\Delta\nu)}, \tag{2.105}$$

where the bars indicate equilibrium values. We assume that this energy balance holds bin by bin, that is, we invoke the principle of detailed balance (see Sec. 1.3). From this equation follows

$$\frac{n_{p,m}}{\Delta\nu} = \frac{(R_s f_{s,um}/R_e f_{e,um})}{(\overline{N}_\ell R_a f_{a,\ell m}/\overline{N}_u R_e f_{e,um}) - 1}. \tag{2.106}$$

In equilibrium the ratio of number densities is given approximately by

$$\frac{\overline{N}_u}{\overline{N}_\ell} = \frac{g_u \exp(-E_u/k_B T)}{g_\ell \exp(-E_\ell/k_B T)} = \frac{g_u}{g_\ell} \exp(-h\nu_0/k_B T), \tag{2.107}$$

where g_u and g_ℓ are the multiplicities of states, the number of different ways a molecule can have the same energy. We say approximately because we approximate a system of weakly interacting molecules with definite, but unknown, energy levels by a collection of non-interacting molecules with fuzzy energy levels. You have seen a version of Eq. (2.107) before in Section 1.2, the Maxwell–Boltzmann distribution of molecular kinetic energies, which are not quantized. The same kind of probability distribution also holds for the quantized (internal) molecular energy levels.

In equilibrium the number density of photons in the mth bin is proportional to the Planck irradiance $P_e(\nu_m)$ at frequency ν_m (see Prob. 5.41):

$$\frac{n_{p,m}}{\Delta\nu} = \frac{4P_e(\nu_m)}{ch\nu_m}. \tag{2.108}$$

If we set Eq. (2.108) equal to Eq. (2.106) and use Eq. (2.107) we obtain

$$\begin{aligned} P_e(\nu_m) &= \frac{(ch\nu_m R_s f_{s,um}/4R_e f_{e,um})}{(g_\ell R_a f_{a,\ell m}/g_u R_e f_{e,um})\exp(h\nu_0/k_B T) - 1} \\ &= \frac{b_m}{c_m \exp(h\nu_0/k_B T) - 1}. \end{aligned} \tag{2.109}$$

This must be equal to the Planck function [Eq. (1.11)], with, of course, $\omega = 2\pi\nu$ transformed to ν:

$$\begin{aligned} P_e(\nu_m) &= \frac{2\pi h\nu_m^3}{c^2}\frac{1}{\exp(h\nu_m/k_B T) - 1} \\ &= \frac{a_m}{\exp(h\nu_m/k_B T) - 1}, \end{aligned} \tag{2.110}$$

from which it follows that

$$\frac{a_m}{\exp(h\nu_m/k_B T) - 1} = \frac{b_m}{c_m \exp(h\nu_0/k_B T) - 1}. \tag{2.111}$$

If we assume that a_m, b_m and c_m are independent of temperature (they contain only ratios of quantities that depend on temperature in approximately the same way), then the only way Eq. (2.111) can be satisfied for arbitrary T is if

$$a_m = b_m = \frac{ch\nu_m R_s f_{\mathrm{s,u}m}}{4 R_{\mathrm{e}} f_{\mathrm{e,u}m}} = \frac{2\pi h\nu_m^3}{c^2}, \tag{2.112}$$

and

$$c_m \exp(h\nu_0/k_{\mathrm{B}}T) = \frac{g_\ell R_{\mathrm{a}} f_{\mathrm{a},\ell m}}{g_{\mathrm{u}} R_{\mathrm{e}} f_{\mathrm{e,u}m}} \exp(h\nu_0/k_{\mathrm{B}}T) = \exp(h\nu_m/k_{\mathrm{B}}T). \tag{2.113}$$

We can solve these equations to obtain

$$R_{\mathrm{s}} f_{\mathrm{s,u}m} = \frac{g_\ell}{g_{\mathrm{u}}} R_{\mathrm{a}} f_{\mathrm{a},\ell m} \frac{8\pi\nu_m^2}{c^3} \exp[-h(\nu_m - \nu_0)/k_{\mathrm{B}}T], \tag{2.114}$$

$$R_{\mathrm{e}} f_{\mathrm{e,u}m} = \frac{g_\ell}{g_{\mathrm{u}}} R_{\mathrm{a}} f_{\mathrm{a},\ell m} \exp[-h(\nu_m - \nu_0)/k_{\mathrm{B}}T]. \tag{2.115}$$

Thus all three rate coefficients are related and depend only on molecular properties and temperature. Although we obtained Eqs. (2.114) and (2.115) by considering a gas in equilibrium with radiation, we assume that these relations hold even when Eq.(2.105) does not, which we turn to next.

Consider the radiance L in a gas. We define radiance much more carefully and thoroughly in Section 4.1. For the moment we define $L_m \Delta\nu \Delta\Omega$ as the amount of radiant energy in the mth bin crossing unit area normal to the beam per unit time and confined to a set of directions with solid angle $\Delta\Omega$. Over a distance Δz the radiance increases because of spontaneous and induced emission and decreases because of absorption. We assume that scattering (see Ch. 3) is negligible. Let A be the area of a rectangle to which the z-axis is perpendicular. The number of molecules in the volume $A\Delta z$ is the number density times this volume. The rate of energy decrease depends on the various transition rates in Eqs. (2.102)–(2.104):

$$\Delta L_m \, \Delta\nu \, A \, \Delta\Omega = A\Delta z h\nu_m \left[-\left(\frac{dn_{\mathrm{u}m}}{dt}\right)_{\mathrm{s}} \frac{\Delta\Omega}{4\pi} - \left(\frac{dn_{\mathrm{u}m}}{dt}\right)_{\mathrm{e}} + \left(\frac{dn_{\ell m}}{dt}\right)_{\mathrm{a}} \right]. \tag{2.116}$$

The first two terms on the right side are gains, the third a loss. We assume that photons are spontaneously emitted with equal probability in all directions, which accounts for the factor $\Delta\Omega/4\pi$. For the general radiation field, $n_{\mathrm{p},m}$ in Eqs. (2.103) and (2.104) is *not* that given by Eq. (2.108) but instead is obtained from

$$ch\nu_m n_{\mathrm{p},m} = L_m \, \Delta\Omega \, \Delta\nu. \tag{2.117}$$

If we substitute Eqs. (2.102)–(2.104) in Eq. (2.116) and use Eqs. (2.107), (2.110), (2.114), (2.115) and (2.117) we obtain, in the limit $\Delta z \to 0$,

$$\frac{dL_m}{dz} = -\kappa_{\ell\mathrm{u},m} L_m + \frac{N_{\mathrm{u}}}{\overline{N}_{\mathrm{u}}} \overline{\kappa}_{\ell\mathrm{u},m} \frac{P_{\mathrm{e}}}{\pi}, \tag{2.118}$$

where

$$\kappa_{\ell u,m} = \frac{N_\ell R_a}{c} \frac{f_{a,m}}{\Delta\nu} \left[1 - \frac{\overline{N}_\ell}{N_\ell} \frac{N_u}{\overline{N}_u} \exp(-h\nu_m/k_B T) \right].$$ (2.119)

Now take the limit as the bin width goes to zero:

$$\kappa_{\ell u} = \frac{N_\ell R_a f}{c} \left[1 - \frac{\overline{N}_\ell}{N_\ell} \frac{N_u}{\overline{N}_u} \exp(-h\nu/k_B T) \right],$$ (2.120)

where

$$f = \lim_{\Delta\nu \to 0} \frac{f_{a,m}}{\Delta\nu}$$ (2.121)

is a probability density. We suppress the bin index and treat frequency as a continuous variable. We also omit the subscript ℓ on f_a, it being understood that absorption is always from a lower to a higher energy state.

According to Eq. (2.31) an absorption coefficient is a number density times an absorption cross section. If we divide Eq. (2.120) by N_ℓ we obtain the absorption cross section for those molecules with energy level E_ℓ that make a transition to E_u:

$$\sigma_{\ell u} = \frac{R_a f}{c} \left[1 - \frac{\overline{N}_\ell}{N_\ell} \frac{N_u}{\overline{N}_u} \exp(-h\nu/k_B T) \right].$$ (2.122)

If the number densities of molecules with upper and lower energy levels are those corresponding to thermal equilibrium [i.e., those given by Eq. (2.107)] at temperature T, the equation of transfer Eq. (2.118) becomes

$$\frac{dL}{dz} = -\overline{\kappa}_{\ell u} L + \overline{\kappa}_{\ell u} \frac{P_e}{\pi},$$ (2.123)

where

$$\overline{\kappa}_{\ell u} = \frac{\overline{N}_\ell R_a f}{c} \left[1 - \exp(-h\nu/k_B T) \right].$$ (2.124)

The multiplier of L is the absorption coefficient, that of P_e/π the emission coefficient, and in Eq. (2.123) the two are equal, which is a form of Kirchhoff's law (see Sec. 1.4). The validity of this law does not require a gas at temperature T to be bathed only in blackbody radiation corresponding to this temperature but rather that the number of molecules in the upper and lower energy states be the thermal equilibrium values.

The absorption coefficient Eq. (2.124) decreases with increasing temperature because $\overline{N}_\ell$ decreases whereas $\overline{N}_u$ increases: there are fewer molecules with the lower energy state to absorb photons and more molecules with the upper energy to be induced to emit photons (equivalent to negative absorption). Although we usually think of emission, the Planck function times an emission coefficient, or emissivity, as being a consequence of spontaneous emission, the absorption coefficient, and hence emission coefficient, is affected by induced emission.
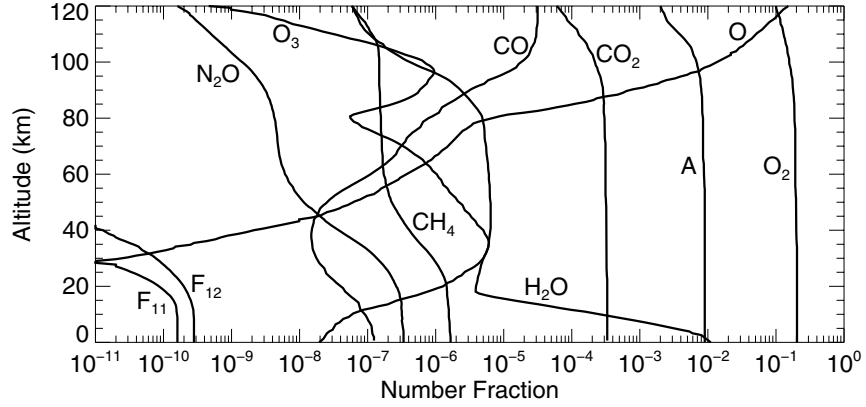
Figure 2.10: Average concentration profiles (expressed as a fraction of the total number density) of various atmospheric molecules. F11 and F12 denote the chlorinated fluorocarbons Freon-11 and Freon-12.

We implicitly assumed molecules with only two energy levels between which transitions are allowed. But there are many such pairs of levels, and hence the absorption coefficient is the sum of many transitions

$$\kappa = \sum_{\ell u} \overline{\kappa}_{\ell u} \tag{2.125}$$

with the corresponding absorption cross section defined by

$$\sigma_a = \frac{\kappa}{N} = \sum_{\ell u} \frac{N_\ell}{N} \sigma_{\ell u}, \tag{2.126}$$

where N is the total molecular number density. All the different lines are separate but not completely: their tails overlap, and the sum of all this overlap gives a continuum background on which is superimposed a series of sharp lines.

We may write Eq. (2.120) so as to emphasize that the absorption coefficient is the product of the line shape f and the *line strength*

$$S_{\ell u} = \frac{N_\ell R_a}{c} \left[1 - \frac{\overline{N}_\ell}{N_\ell} \frac{N_u}{\overline{N}_u} \exp(-h\nu/k_B T) \right]. \tag{2.127}$$

At normal terrestrial temperatures $\exp(-h\nu/k_B T)$ is usually $\ll 1$ for vibrational and, especially, electronic transitions, but not necessarily for rotational transitions. At such temperatures an appreciable fraction of molecules may be in excited (higher than the ground state) rotational energy states.
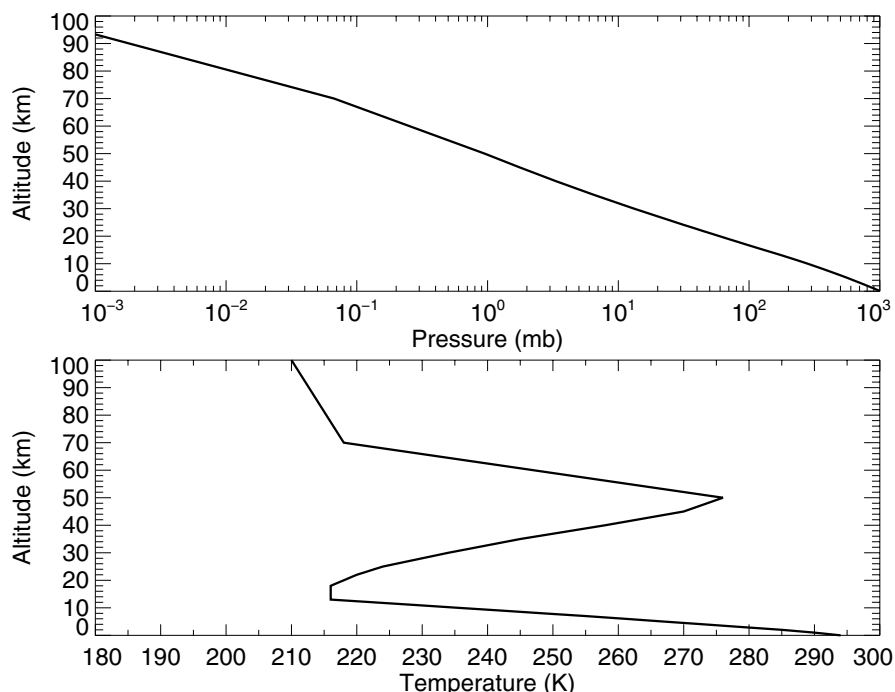
Figure 2.11: Typical mid-latitude summer pressure and temperature profiles. After McClatchey *et al.* (1972).

2.8.2 Absorption by Atmospheric Molecules

The basic determinant of transmission by Earth's atmosphere is the absorption coefficient Eq. (2.125), which depends on concentration profiles of atmospheric gases, their electronic, vibrational, and rotational energy levels, and the variation of temperature and pressure with height.

Several years ago students in a graduate seminar course surveyed the scientific literature for ranges of concentrations of atmospheric gases. The average values they obtained are shown in Fig. 2.10. A more recent perusal of the literature indicates that these values are typical. Although molecular nitrogen (not shown) is the most abundant gas, it is not a major player in absorption by the lower atmosphere except in an indirect way by virtue of interaction (collisional broadening) with its less abundant but radiatively active neighbors.

The absorption coefficient is proportional to molecular concentration, which varies with height. Less evident is an additional variation with height because the line shape depends on total pressure and, to a lesser extent, temperature. Vertical pressure profiles are monotonously similar whereas temperature profiles exhibit more latitudinal and seasonal variation. For the calculations shown here we used typical mid-latitude, summer pressure and temperature profiles (Fig. 2.11).

Cross Sections and Transmissivities

Figure 2.12 shows the absorption cross section at sea level (1013 mb and 294 K) over six wavelength decades for the dominant radiatively active atmospheric molecules. The absorption cross section of a molecule is defined as the ratio of its absorption coefficient to its number density [Eq. (2.126)]. We computed the absorption coefficient with a *line-by-line* radiative transfer code. By this is meant that the spectrum is divided into such small intervals (e.g., wavenumber intervals $0.001\ \mathrm{cm}^{-1}$ or less) that the finest spectroscopic details are captured. Absorption lines are narrowest for the lowest pressures (see Fig. 2.19), so these lines dictate the interval width; the necessity of line-by-line calculations for accurate results is evident from the huge variation in absorption, as much as four decades, over a small range of wavenumbers. Although wavenumber is proportional to energy level differences, to show how cross sections vary from the ultraviolet (0.1–0.4 μm) through the visible (0.4–0.7 μm) into the infrared (0.7 μm to 1 mm) and out to the microwave (1 mm to 10 cm) spectral regions, wavelength is easier to visualize.

Three fundamental lessons follow from Fig. 2.12: the magnitudes of absorption cross sections vary over a huge range, more than 10 decades; the spectra of the molecules are quite different; and the spectra are extraordinarily complicated. To determine how these spectra determine transmission by the atmosphere we need absorption coefficients for each molecular species at many altitudes, the total absorption coefficient at a height being the sum over species. The exponential of the negative of this absorption coefficient integrated along a vertical path from sea level to high in the atmosphere is the transmissivity [see Eq. (2.22)]. The result is shown in Fig. 2.13, in six separate panels so as to clearly display the marked differences between spectral regions. We also include the consequences of scattering by atmospheric molecules, the total transmissivity being the product of the absorption and scattering transmissivities (assuming single scattering; see Ch. 5).

Ozone absorbs almost all solar radiation shortward of 0.3 μm. With the exception of ozone and, to a lesser extent oxygen, atmospheric molecules are negligibly absorbing over the visible, and hence the smooth decrease in transmissivity from about 0.6 μm to 0.3 μm is a consequence of scattering by atmospheric molecules (mostly nitrogen because it is the dominant species). Absorption from 0.6 μm to 0.8 μm is dominated by oxygen, and from 0.8 to 5 μm by water vapor helped by two appreciable peaks attributable to carbon dioxide. Within the spectral region 0.3–3 μm there are several broad regions over which the transmissivity exceeds 0.7. Most of the solar spectrum lies in this region, and hence much solar radiation penetrates clear skies and is transmitted to Earth's surface.

From 4 μm to 1 mm absorption by water vapor, with contributions from carbon dioxide and ozone, is often so strong that the transmissivity is nearly zero over broad ranges within this region. One important exception is 8–12 μm, where transmissivity often exceeds 0.6. Emission by Earth peaks in this region, sometimes called the *window region* because transmission of radiation from the surface is high. From 0.1 mm to 1 mm, transmission is essentially zero, and beyond about 1 mm solar and terrestrial radiation do not contribute much to Earth's radiation budget. But from 1 mm to 10 cm atmospheric transmissivity generally increases, approaching 1 at 10 cm. No wonder weather radars operate at 10 cm!
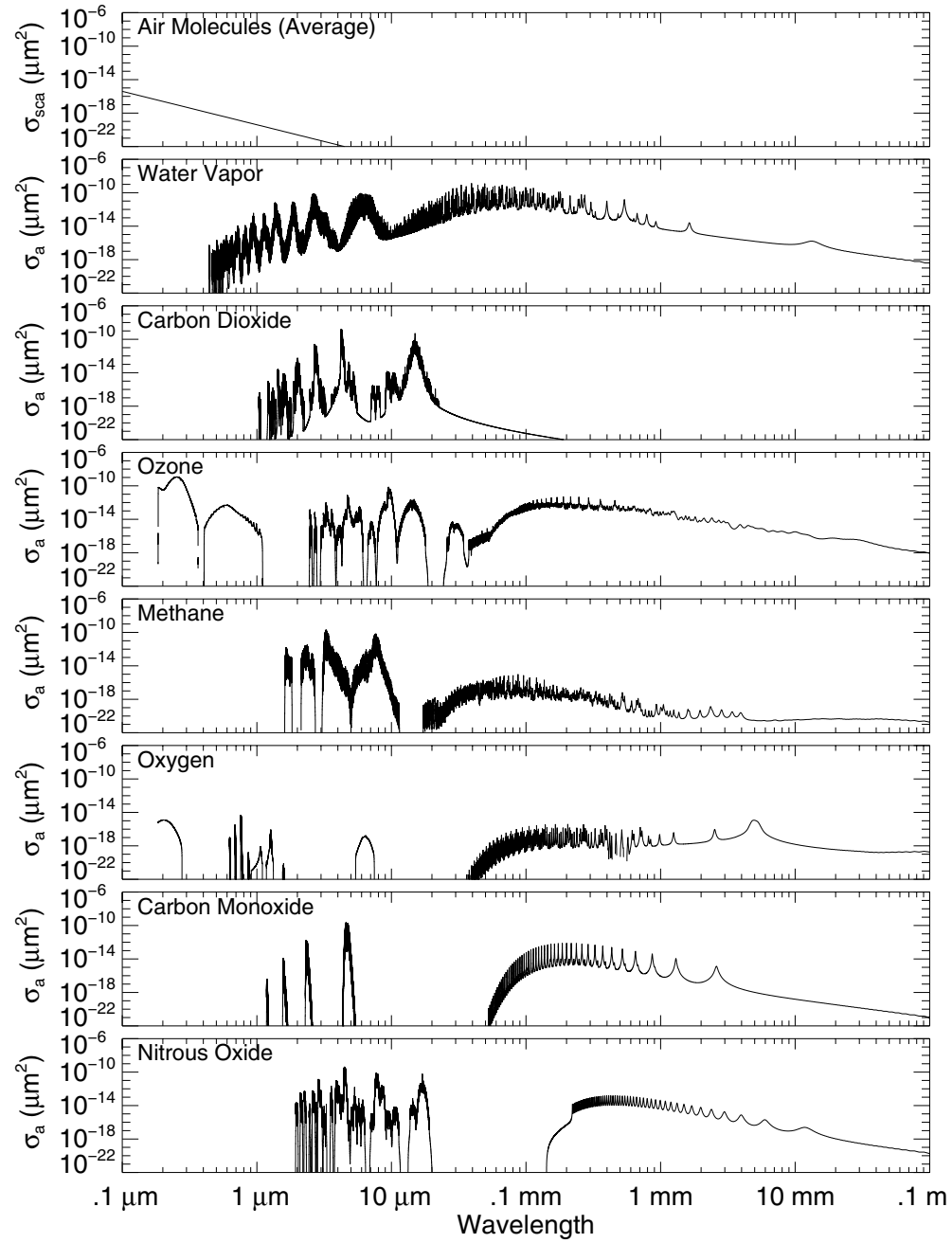
Figure 2.12: Absorption cross sections for most of the strongly absorbing atmospheric gases (1013 mb and 294 K) and average scattering cross section for air molecules (top panel).
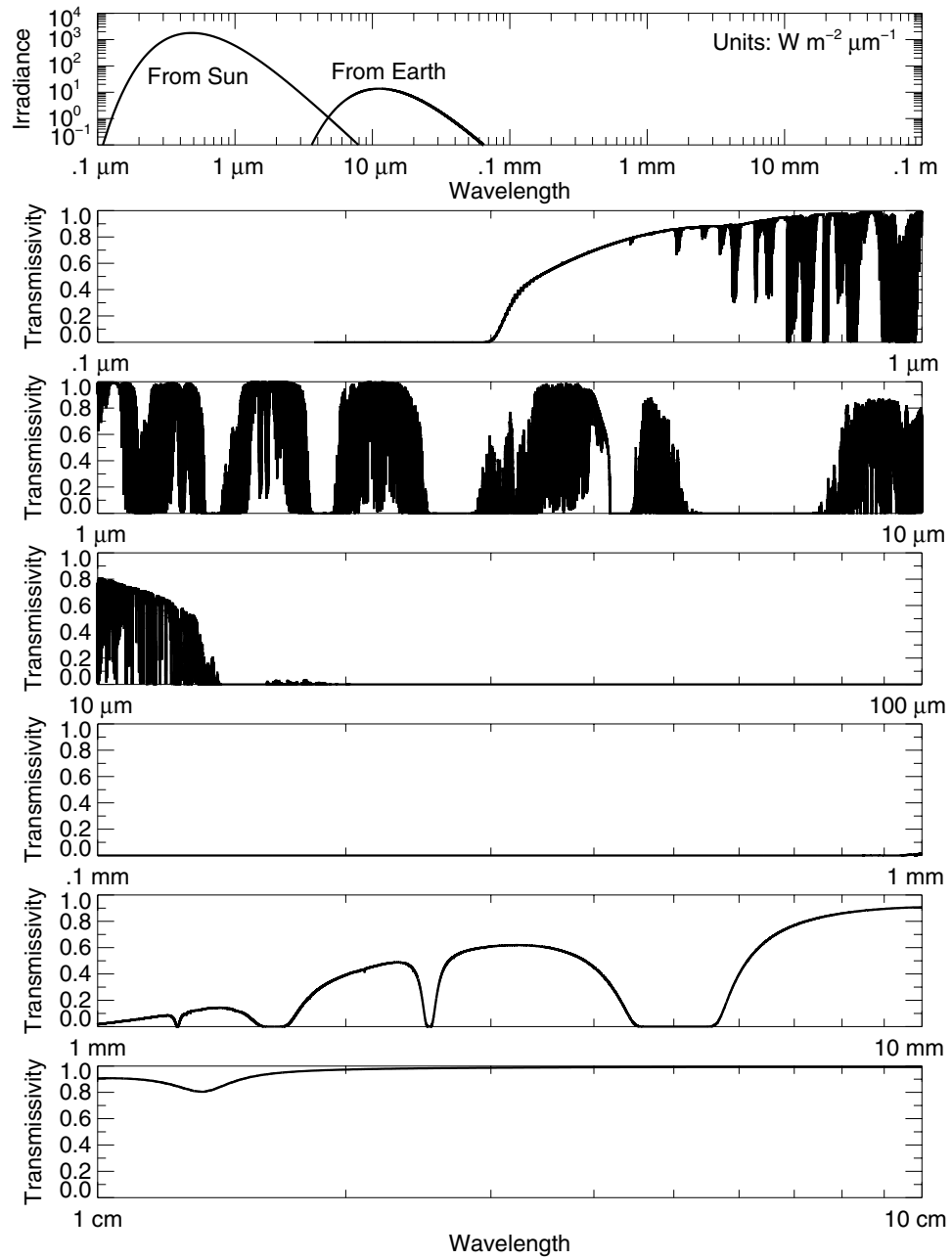
Figure 2.13: Transmissivity along a vertical path through the entire atmosphere. Each panel is one wavelength decade. The top panel shows approximate solar and terrestrial irradiances.

State Populations and the Structure of Absorption

To understand the detailed structure of the absorption cross sections in Fig. 2.12 we have to probe deeper into the energy states of molecules, the statistical distribution of these states, and line broadening.

The probability that a molecule of a gas in equilibrium at temperature T has a quantized energy E_s is

$$p(E_s) = \frac{g_s \exp\left(-E_s/k_B T\right)}{\displaystyle\sum_{s=0}^{\infty} g_s \exp\left(-E_s/k_B T\right)}, \tag{2.128}$$

where g_s, as we note following Eq. (2.107), is the multiplicity (or density or degeneracy, an unfortunate term give its unsavory connotations, or number) of states with energy E_s. You have seen a density of states before perhaps without realizing it. For example, in the continuous distribution function for molecular kinetic energies [Eq. (1.8)] $\sqrt{E}$ multiplying the Boltzmann factor $\exp(-E/k_B T)$ is the density of states. This becomes more evident if we transform from E to speed v, in which instance the density of states is proportional to v^2. If molecular translational states are uniformly distributed in a three-dimensional velocity space, the number of states between v and $v + \Delta v$ is proportional to v^2, and all these states correspond to approximately the same energy. And similarly for the Planck function Eq. (1.11). If we divide it by $\hbar\omega$, we obtain a quantity proportional to the spectral distribution of the number density of photons, which divided by the integral over all frequencies is the probability distribution. Here the density of states is proportional to ω^2. The factor multiplying this density of states is not the Boltzmann factor, except for photon energies $\hbar\omega \gg k_B T$, because photons don't obey the same statistical law as molecules.

The simplest molecule for which absorption is shown in Fig. 2.12 is carbon monoxide (CO), and so we would be wise to attempt to understand its spectrum before considering more complicated molecules. Carbon monoxide has a single vibrational mode with frequency ν_1; the corresponding vibrational energy levels are

$$E_n = h\nu_1 \left(n + \frac{1}{2}\right) = hc\tilde{\nu}_1 \left(n + \frac{1}{2}\right), \ n = 0, 1, \ldots, \tag{2.129}$$

where $\tilde{\nu}_1 = 2143.3\,\text{cm}^{-1}$. If we treat the carbon monoxide molecule as a rigid rotator its rotational energy levels are approximately

$$E_J = BhcJ\left(J + 1\right), \ J = 0, 1, \ldots, \tag{2.130}$$

where we take $B = 1.9563\,\text{cm}^{-1}$ for the molecule in its lowest (ground) vibrational state. Although a truly rigid rotator cannot vibrate, the simple energy levels in Eqs. (2.129) and (2.130) result in a cross section spectrum that agrees with much more detailed line-by-line calculations.
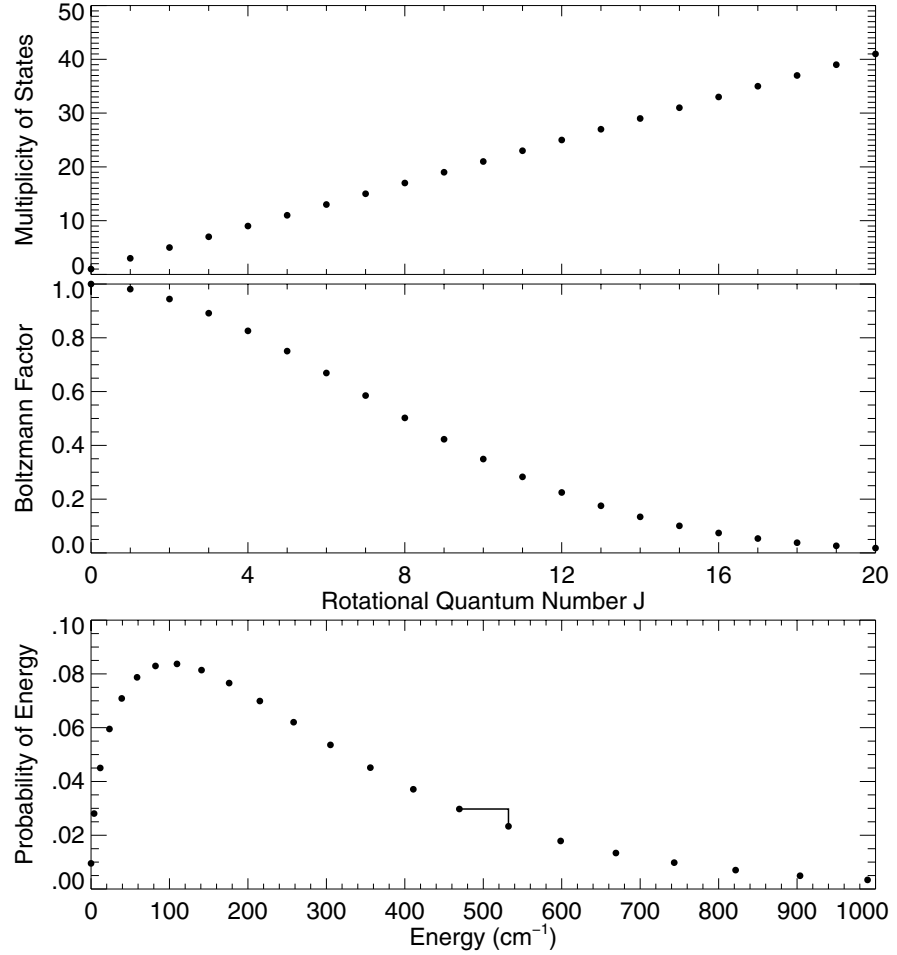
Figure 2.14: Multiplicity of states of a rigid rotator for different rotational quantum numbers J (top). Boltzmann factor for the carbon monoxide molecule at $T = 294$ K (middle). Probability that this molecule has a particular energy (bottom); the energy (wavenumber) difference indicated by the two connected dots corresponds to a transition from $J = 15$ to $J = 16$.

The multiplicity of vibrational states is $g_n = 1$, whereas that of the rotational states is $g_J = 2J + 1$. Thus the probability that a CO molecule has the vibrational energy E_n is

$$p(E_n) = \frac{\exp\{-hc\tilde{\nu}_1(n + 1/2)/k_{\mathrm{B}}T\}}{\sum_{n=0}^{\infty} \exp\{-hc\tilde{\nu}_1(n + 1/2)/k_{\mathrm{B}}T\}} \tag{2.131}$$

and the probability that it has the rotational energy E_J is

$$p(E_J) = \frac{(2J+1)\exp\{-B\,h\,c\,J(J+1)/k_{\mathrm{B}}T\}}{\displaystyle\sum_{J=0}^{\infty}(2J+1)\exp\{-B\,h\,c\,J(J+1)/k_{\mathrm{B}}T\}}. \tag{2.132}$$

At normal terrestrial temperatures, almost all CO molecules are in their ground $(n = 0)$ vibrational state (see the top panel in Fig. 2.16), which is why the simple rotational energy level structure [Eq. (2.130)] yields good results.

Figure 2.14 shows the multiplicity of states g_J as a function of rotational quantum number J, the corresponding Boltzmann factor $\exp(-E_J/k_{\mathrm{B}}T)$ for $T = 294\,\mathrm{K}$, and the probability $p(E_J)$. Because g_J increases monotonically with increasing J whereas the Boltzmann factor decreases, the maximum of $p(E_J)$, unlike that of $p(E_n)$, does not correspond to the ground state.

Energy level transitions of a linear harmonic oscillator with rigid rotations, resulting in emission or absorption of radiation, must obey specific selection rules: the rotational quantum number J must change by 1 and the vibrational quantum number n can change by *no more* than 1. Every allowed energy level difference corresponds to a photon energy. For example, the bottom panel of Fig. 2.14 shows an energy level difference of about 60 cm^{-1} for an allowed transition from $J = 15$ to $J = 16$, which corresponds to a photon of this energy. The top panel in Fig. 2.15 shows the absorption cross section $\overline{\kappa}_{\ell u}/N$ for each rotational transition. These cross sections are a discrete set for frequencies at the peak of each absorption line. In reality, lines are broadened and partly overlap. The cross section [Eq. (2.126)] taking this into account is shown in the middle panel in Fig. 2.15. These calculations require the line shape (discussed in the following subsection) and the rate constant R_{a} for rotational transitions of carbon dioxide. For a rigid rotator theoretical estimates are available and we used those in Eqs. 3.9 and 3.38 (converted to SI units) of Goody and Yung (1989). These relatively simple calculations are in remarkably good agreement with line-by-line calculations (bottom panel) that include all whistles and bells. But the results are not identical. The very fine features in the line-by-line calculation are a consequence of the different isotopes of carbon, which result in CO molecules with different masses. The simple theory also neglects anharmonic terms, which is a fancy way of saying that there is no such thing as an exactly harmonic oscillator. Note that in Section 2.6.1 we referred to the harmonic approximation. Although we can ignore terms higher than quadratic in the expansion of the potential, Nature cannot.

The Sierra Nevada mountain range in California (also in Spain) owes its name to its long, jagged profile seen at a distance against the horizon sky: *sierra nevada* is literally "snow-covered saw." Sierras are abundant in Fig. 2.12, for example, the *Sierra Rotatoria* of carbon monoxide stretching eastward across the spectrum from about 50 µm. Far to the west lie four *picachos* (to borrow another geographical term from the Spanish), isolated peaks jutting from the surrounding plane, to which we now turn our attention, especially the *Gran Picacho Vibratorio* at about 4.67 µm. These picachos, less imaginatively called bands with their associated lines, are a consequence of transitions of the carbon monoxide molecule from its $n = 0$ to $n = 1$ vibrational state accompanied by a change $\Delta J = \pm 1$ of its rotational state.
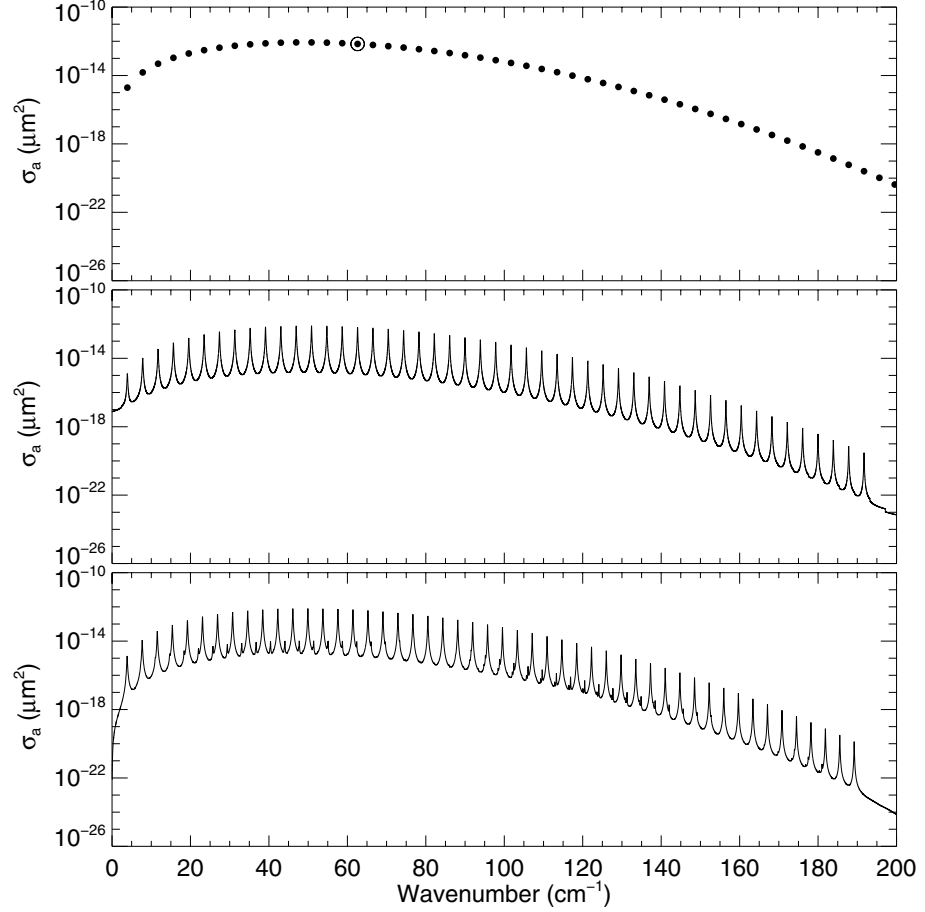
Figure 2.15: Absorption cross sections of carbon dioxide for the transitions shown in Fig. 2.14. The black dot enclosed by a circle (top panel) corresponds to a photon energy (wavenumber) equal to the difference between the $J = 15$ and $J = 16$ energy levels. These discrete cross sections are calculated at line centers and do not include overlap. The middle panel illustrates all of each broadened line and overlapping of lines. The bottom panel shows detailed line-by-line calculations. All calculations are for $T = 294$ K at a pressure of 1013 mb.

The probability that a molecule has a vibrational energy E_n *and* a rotational energy E_J is

$$p(E_n, E_J) = p(E_n)p(E_J), \tag{2.133}$$

with $p(E_n)$ given by Eq. (2.131), $p(E_J)$ by Eq. (2.132). The probability $p(E_n)p(E_J)$ is shown in the top panel of Fig. 2.16 for $n = 0$ and $n = 1$, each for a range of J. The horizontal and vertical lines indicate energy differences corresponding to the $n = 0$ to $n = 1$ transition with J changing from 15 to 14 and to 16. The set of higher probabilities ($n = 0$) shown in this figure duplicates that shown in the lower panel of Fig. 2.14. Added to Fig. 2.16 is a set of probabilities ($n = 1$) that are much lower (because $E_{n=1} > E_{n=0}$).
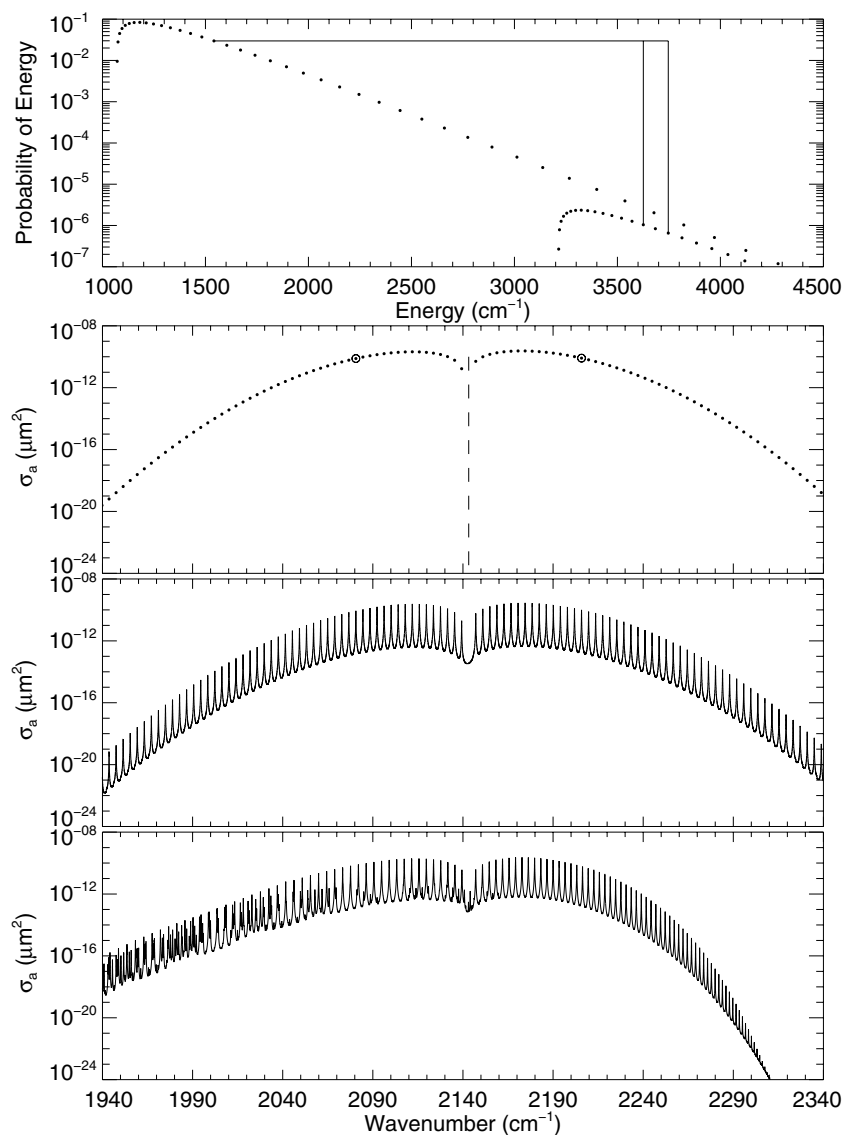
Figure 2.16: Top panel: Probability of a vibration-rotational energy level of carbon monoxide for vibrational quantum numbers $n = 0$ and $n = 1$ and a range of rotational quantum numbers J. Horizontal and vertical lines indicate energy level differences corresponding to the $n = 0$ to $n = 1$ transition with J changing from 15 to 14 and to 16. Second panel from top: Absorption cross sections with no overlapping or broadening of lines for a set of allowed energy differences for two branches: one for an increase in J by 1, one for a decrease by 1. Third panel from top: Absorption cross sections with overlapping and broadening of lines according to simple theory. Bottom panel: Absorption cross sections from line-by-line calculations.

The second panel from the top in Fig. 2.16 shows the absorption cross section with no overlapping or broadening of lines for a set of allowed energy differences, and hence photon energies, for two branches: $\Delta J = 1$ and $\Delta J = -1$. This set of points is similar to that shown in Fig. 2.15 with its mirror image added. That is, the points in Fig. 2.16 are nearly symmetric about the dashed vertical line, which corresponds to the forbidden transition $\Delta J = 0$. The third panel in Fig. 2.16 is the counterpart of the second panel in Fig. 2.15. Because there is no simple theory of the rate constant for $n = 0$ to $n = 1$ transitions, we used the same line strengths as for the rotational band, then multiplied the resulting cross sections by a scale factor so that their peak value matched that obtained from the line-by-line calculation (bottom panel in Fig. 2.16). The lack of obvious resemblance between the cross section spectrum in Fig. 2.16 and the *Gran Picacho Vibratorio* in Fig. 2.12 is a result of the quite different scales.

As evidenced by comparing the third and fourth panels in Fig. 2.16, the relatively simple theory is adequate to explain the general form of the lowest energy vibration-rotational band of carbon monoxide (and other atmospheric gases). That agreement between simple theory and detailed line-by-line calculations is not as good for vibration-rotational bands as for rotational bands (with a molecule in its lowest vibrational state) is to be expected because the higher the vibrational state, the less valid the rigid rotator model of a molecule.

What about the lower picachos to the west of *Gran Picacho Vibratorio*? These are inexplicable by theory based on a strictly harmonic oscillator. Anharmonic terms (i.e., what was discarded in the expansion of the potential) give rise to transitions between $n = 0$ and $n = 2, 3, 4$ states, and hence frequencies that are integral multiples (overtones) of the fundamental frequency. The picachos for carbon monoxide are equally spaced in wavenumber, each about 100 times lower than its neighbor to the east. We briefly discuss overtones and combination bands of the water molecule in Section 2.7.1. Carbon monoxide has only overtones because of its single normal mode frequency. In Fig. 2.15 and 2.16 (bottom panels) absorption cross sections at the peaks are perhaps a thousand times greater than in the troughs. But the cross section never drops to zero.

With one exception, the absorption spectra in Fig. 2.12 are, for the most part, attributable to rotational and vibration-rotational transitions. The fundamental normal mode frequencies for all molecules (except oxygen) in Fig. 2.12 are shown in Fig. 2.17. By taking combinations and overtones of these fundamental frequencies, one can estimate the positions of the vibrational peaks shortward in wavelength of pure rotational lines. The exception is ozone at wavelengths less than about 1 μm. For this molecule electronic transitions are responsible for absorption from approximately 0.4 μm to 1 μm (*Chappuis* bands), from 0.3 μm to 0.35 μm (*Huggins* bands), from 0.2 μm to 0.3 μm (*Hartley* bands), and at wavelengths less than 0.2 μm. The relevance of the Chappuis bands to the color of the twilight sky is discussed in Section 8.1.3. At sufficiently short wavelengths (high energies) photons destroy molecules rather than merely promote them into higher energy states. A molecule is an association of atoms, so dissolution of the association is called dissociation, or more precisely, photo-dissociation to indicate that the breakup is mediated by photons.

Oxygen is different from the other six gases of Fig. 2.12 in that its relatively weak rotational structure is a consequence of its permanent *magnetic* dipole moment (see Sec. 7.3). The vibration-rotational band centered on 6.43 μm is a result of electric quadrupole transitions, and the sets of bands between 0.9 μm and 1.4 μm and between 0.6 μm and 0.9 μm are vibration-rotational bands superimposed on electronic transitions, which also are also re-
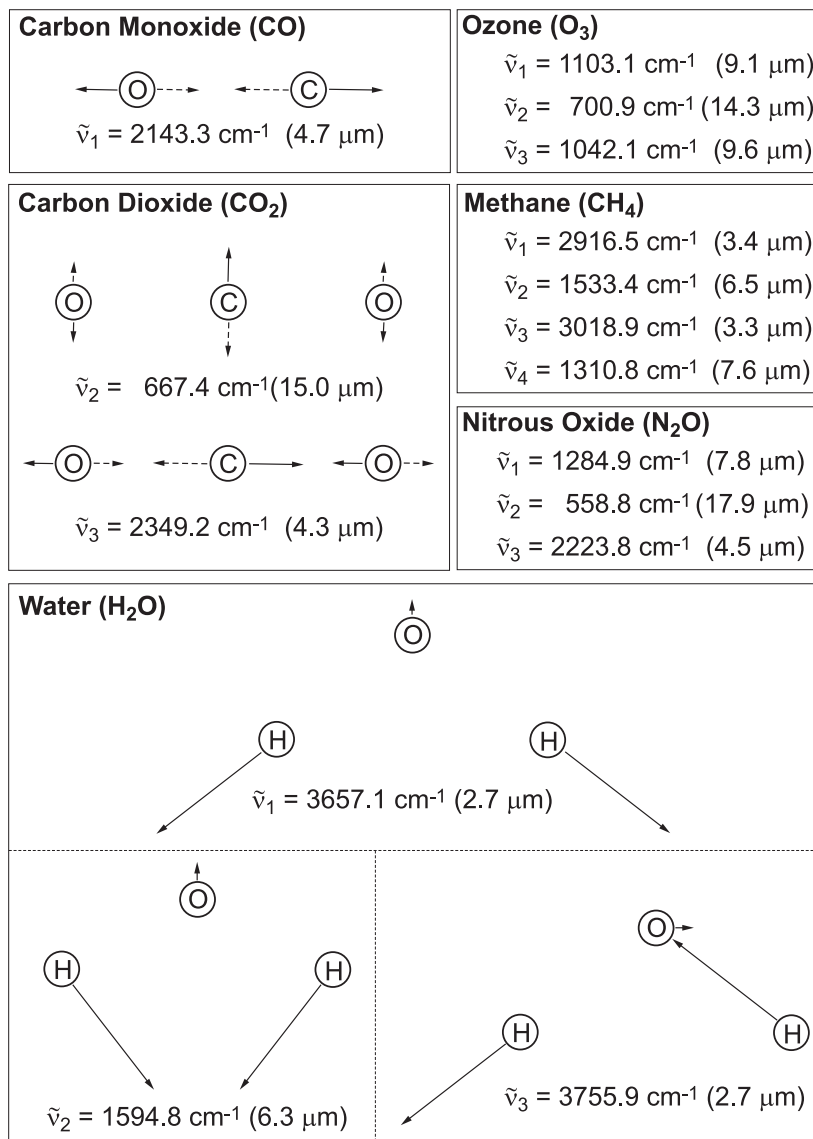
Carbon Monoxide (CO)

$\tilde{v}_1$ = 2143.3 cm-1 (4.7 μm)

Carbon Dioxide (CO$_2$)

$\tilde{v}_2$ =   667.4 cm-1(15.0 μm)

$\tilde{v}_3$ = 2349.2 cm-1 (4.3 μm)

Ozone (O$_3$)

$\tilde{v}_1$ = 1103.1 cm-1   (9.1 μm)

$\tilde{v}_2$ =   700.9 cm-1 (14.3 μm)

$\tilde{v}_3$ = 1042.1 cm-1   (9.6 μm)

Methane (CH$_4$)

$\tilde{v}_1$ = 2916.5 cm-1   (3.4 μm)

$\tilde{v}_2$ = 1533.4 cm-1   (6.5 μm)

$\tilde{v}_3$ = 3018.9 cm-1   (3.3 μm)

$\tilde{v}_4$ = 1310.8 cm-1   (7.6 μm)

Nitrous Oxide (N$_2$O)

$\tilde{v}_1$ = 1284.9 cm-1   (7.8 μm)

$\tilde{v}_2$ =   558.8 cm-1 (17.9 μm)

$\tilde{v}_3$ = 2223.8 cm-1   (4.5 μm)

Water (H$_2$O)

$\tilde{v}_1$ = 3657.1 cm-1 (2.7 μm)

$\tilde{v}_2$ = 1594.8 cm-1 (6.3 μm)

$\tilde{v}_3$ = 3755.9 cm-1 (2.7 μm)

Figure 2.17: Schematic normal mode vibrations of carbon monoxide, carbon dioxide, and water vapor and normal mode frequencies for ozone, methane, and nitrous oxide. The frequencies for carbon monoxide are from Goody and Yung (1989, pp. 210–11), for carbon dioxide from Goody and Yung (1989, pp. 204–07) and Herzberg(1945, pp. 66, 272–6), for water from Goody and Yung (pp. 198–201) and Herzberg (pp. 171, 280–93), for ozone from Goody and Yung (pp. 207–10) and Herzberg (pp. 285–7), for methane from Goody and Yung (p. 211) and Herzberg (pp. 306–9) and nitrous oxide from Goody and Yung (p. 210) and Herzberg (pp. 277–8).

sponsible for absorption between 0.2 μm and 0.26 μm (*Herzberg* band and continuum) and between 0.13 μm and 0.2 μm (*Schumann–Runge* bands and continuum).

Line Shapes

Absorption lines are broadened by the line shape function. We derive one line shape in Section 2.6, the Lorentz line shape, from the classical equation of motion of a harmonic oscillator. The form of this line shape carries over into the quantum domain but the terms in it are interpreted differently. For example, the frequency ω_0 in Eq. (2.60) is a resonant frequency whereas its quantum-mechanical interpretation is an energy level difference divided by $\hbar$. If we transform from circular frequency ω to wavenumber $\tilde{\nu}$ in Eq. (2.73) and multiply by a factor so that the integrated line shape is normalized, we obtain

$$f_c(\tilde{\nu}) = \frac{\alpha_c}{(\tilde{\nu} - \tilde{\nu}_0)^2 + \alpha_c^2}, \tag{2.134}$$

where the half-width $\alpha_c = \gamma/4\pi c$ and $\tilde{\nu}_0$ is the wavenumber at the center of the line. The subscript c denotes collisional broadening, a consequence of inter-molecular, in contrast with intra-molecular, forces. Equation (2.134) also applies to *natural* line broadening, the irreducible line broadening of absorption lines of an isolated molecule at rest. The natural line broadening half-width is so small for the electronic, vibrational, and rotational transitions of interest in atmospheric applications that we need not consider it further.

At 1013 mb and 294 K the collision-broadened half-width α_c is typically 0.03–0.10 cm^{-1} for vibration-rotational transitions. For example, a water vapor line at 1085.436 cm^{-1} has a *foreign-broadened* half-width $\alpha_f = 0.0755$ cm^{-1} and a *self-broadened* half-width $\alpha_s = 0.3580$ cm^{-1}. These terms are self-explanatory: a molecule interacts with its fellow countrymen and with foreigners. The total half-width is the number-weighted average of the two half-widths. High in the atmosphere collisional broadening of the water vapor line is dominated by foreign broadening because water vapor decreases with height more rapidly than do nitrogen and oxygen (Fig. 2.10). Foreign broadening is (approximately) proportional to pressure.

The Lorentz line shape for this transition for the mid-latitude summer profile (Fig. 2.11) is shown in Fig. 2.18 at the surface (1013 mb, 294 K) and high in the atmosphere (10 mb, 238 K). At high altitudes the concentration of water vapor is so low (Fig. 2.10) that broadening of this line is dominated by foreign broadening. If collisional broadening is a consequence of molecular interactions, we expect broadening to decrease with increasing altitude, and the line shapes in Fig. 2.18 support this expectation. (Note the great difference in wavenumber scales between low and high altitudes.) This figure also shows two other line shapes, which we turn to next.

Doppler broadening of lines is a result of relative motion between a source of photons and molecules that absorb them. We discuss the Doppler shift in detail in Section 3.4.6 and derive an expression for the distribution of the frequencies of scattered radiation as a consequence of random molecular motions of an ideal gas [Eq. (3.80)]. Given this expression, and taking into account that the Doppler shift for absorption is half that for scattering, we obtain the Doppler
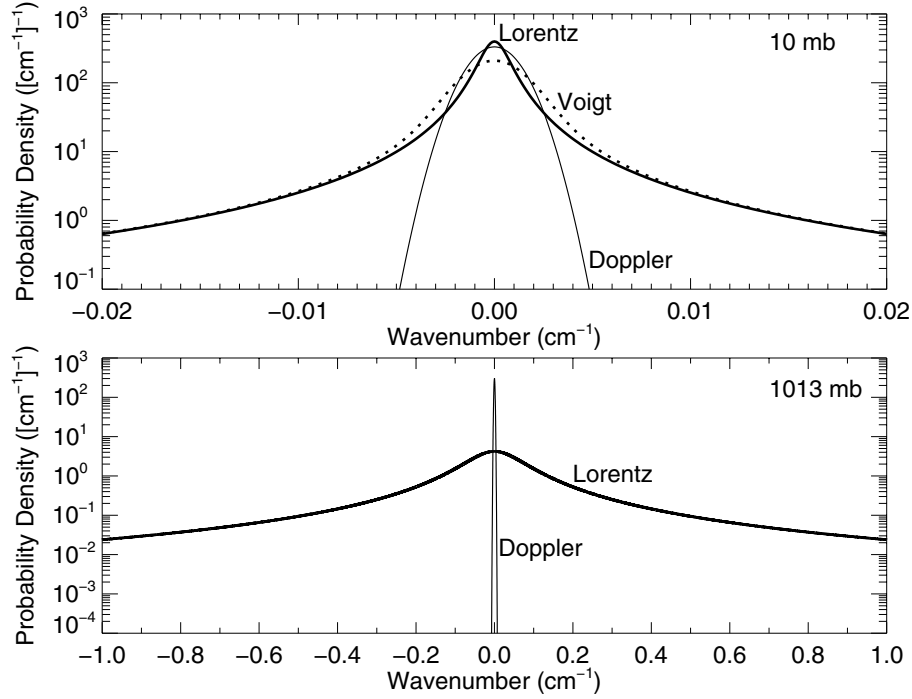
Figure 2.18: Line shapes for a water vapor line high in the atmosphere (10 mb, 238 K) and at the surface (1013 mb, 294 K). Wavenumbers are relative to that at the peak. The Voigt line shape is a convolution of the Lorentz and Doppler line shapes.

line shape

$$f_D(\tilde{\nu}) = \frac{1}{\alpha_D \sqrt{\pi}} \exp\left\{ -\left(\frac{(\tilde{\nu} - \tilde{\nu}_0)}{\alpha_D} \right)^2 \right\}, \tag{2.135}$$

where the Doppler half-width $\alpha_D = (\tilde{\nu}_0/c)\sqrt{2k_B T/m}$. Doppler line shapes for the water vapor line also are shown in Fig. 2.18.

Determining the line shape because of both collisional and Doppler broadening is not trivial. To understand why, consider that we approximate absorption by gas molecules as transitions between two sets of energy levels, where the energy level difference (frequency) is distributed according to the Lorentz line shape. Each one of these energy level differences is Doppler shifted by molecular motion. Because of this, the total line shape is a *convolution* of the two line shapes, called the *Voigt* line shape:

$$f(\tilde{\nu}) = \int_{-\infty}^{\infty} f_c(\tilde{\nu}') f_D(\tilde{\nu} - \tilde{\nu}') \, d\tilde{\nu}' = \int_{-\infty}^{\infty} f_c(\tilde{\nu} - \tilde{\nu}') f_D(\tilde{\nu}') \, d\tilde{\nu}', \tag{2.136}$$
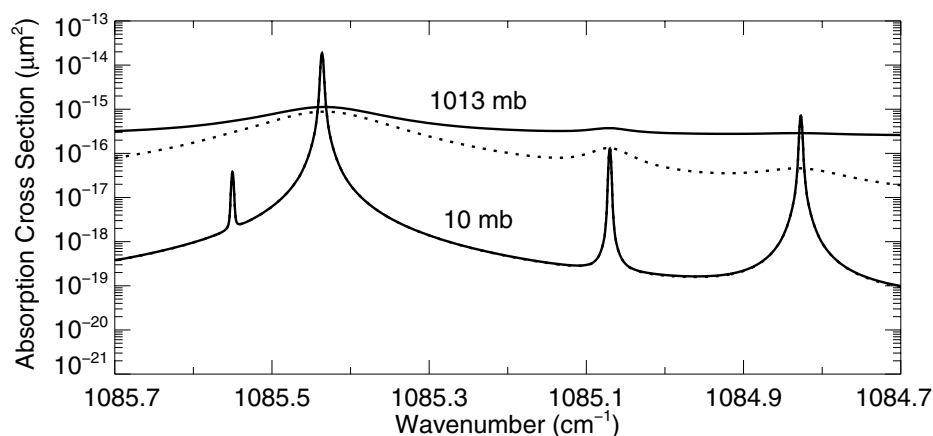
Figure 2.19: Line-by-line calculations of water vapor absorption at two pressure altitudes. The dashed curve shows a calculation from which continuum absorption is omitted. At 10 mb, omission of the continuum results in a curve indistinguishable from that shown.

where the integration limits are symbolic not literal. Equation (2.136) is the line shape for *most* temperatures and pressures, but in certain limits reduces approximately to the Doppler or Lorentz line shape. For example, low in the atmosphere (Fig. 2.18) the Doppler line shape is nearly a spike compared with the much broader Lorentz line shape. When the two are convolved, the result is almost the Lorentz line shape. The other extreme is high in the atmosphere, where the Lorentz line shape is a spike compared with the much broader Doppler line shape. The two profiles at 10 mb in Fig. 2.18 convolve to a line shape that is neither one nor the other.

Figure 2.19 shows four water vapor lines for two pressures. At the lower pressure (10 mb) all lines are distinguishable, but at the higher pressure (1013 mb) collisional broadening is so great that two lines are completely lost and two are demoted from proud peaks to humble dunghills. This figure also shows the results of line-by-line calculations with and without *continuum* absorption; the difference is, however, imperceptible at the lower pressure. Continuum absorption is that added to bring theory and measurements into agreement, usually by theoretically or empirically adjusting lines so that they are accurate far from their centers. By positing stable aggregates of two (dimers) or three (trimers) or more water molecules, additional, but weak, absorption is expected (and has been found experimentally) in infrared regions of the spectrum.

Line shapes (e.g., Lorentz and Voigt) obtained by modeling the weak interaction between a single infrared-active molecule and only one other molecule are accurate near line centers. But line shapes far from centers are much more difficult to model. In spectral regions far from all line centers, in which only the wings of lines overlap, accurate treatment of these wings is essential for estimates of absorption.
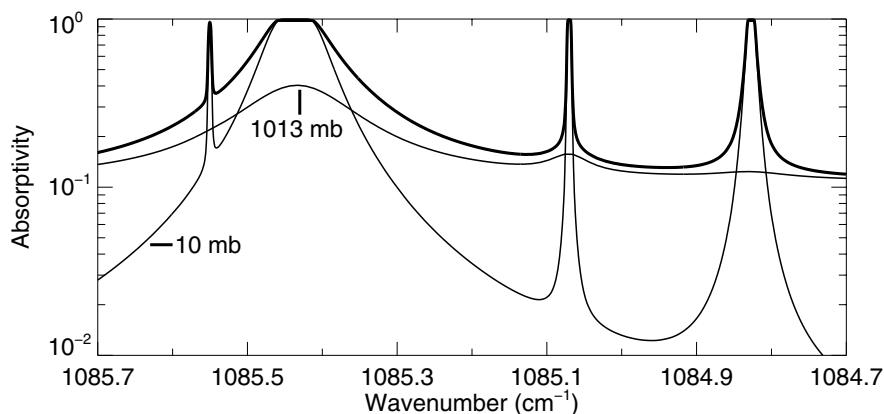
Figure 2.20: Water vapor absorptivities of a two-layer model atmosphere, one layer at 1013 mb and 294 K, one layer at 10 mb and 238 K. The thicker dark curve is the absorptivity of the two layers combined.

Implications for Remote Sensing of Water Vapor

The dependence of collisional broadening on pressure (altitude) and line strength on concentration provides a way to remotely determine vertical water vapor profiles using transmission measurements. To show this we consider a simple two-layer atmosphere. Each layer has a uniform pressure and temperature, the lower one at 1013 mb and 294 K, the upper one at 10 mb and 238 K. The physical thicknesses are chosen to make the absorptivity, averaged over the spectral region shown in Fig. 2.19, of the layers the same.

The absorptivity of each layer is quite different because of collisional broadening (Fig. 2.20). Although all that can be measured from the ground is the total absorptivity of the two layers together, absorptivity at the peaks is dominated by the upper layer, between peaks by the lower layer. Thus a measurement of total absorptivity (transmission) at sufficiently high resolution (< 0.01 cm^{-1}) yields information about both layers, in particular the amount of water vapor in each because of the dependence of the line strength [Eq. (2.127)] on concentration.

The mesosphere, that part of Earth's atmosphere between about 50 km and 80 km, is a difficult place to visit: too high for balloons, too low for satellites. And yet for many years water vapor profiles in the mesosphere have been inferred from ground measurements of high-resolution radiance spectra around a 22.2 GHz water vapor line. Near this frequency even a cloudy atmosphere is not opaque, pressure broadening occurs to high altitudes, and the line shape can be resolved.

Low in the atmosphere (e.g., the troposphere) collisional broadening flattens the 22.2 GHz line (Fig. 2.21). At these altitudes emission does not vary appreciably near line center, and hence if all the water vapor in a vertical column above 40 km were to vanish, the radiance spectrum would be flat (Fig. 2.21). But from 40 km to 80 km collisional broadening decreases,
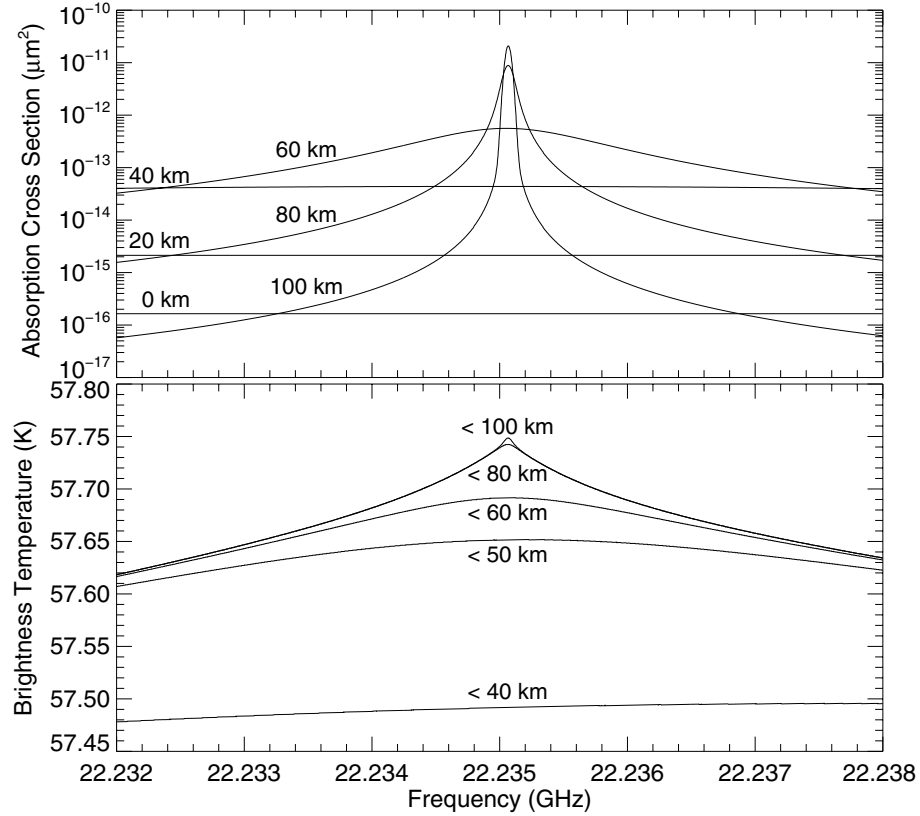
Figure 2.21: Water vapor absorption cross section spectra every 20 km from the surface to 100 km (top panel). Brightness temperatures that would be measured at the surface if all the water vapor above the indicated altitude were not present (bottom panel). Note the switch from wavenumber to frequency, and radiance expressed as a brightness temperature (Sec. 1.4.4), the variables favored by those who measure microwave radiance.

the absorption cross section near line center increases, and hence so does emission relative to that in the wings. By 80 km, Doppler broadening dominates, and the line shape no longer changes apart from small changes with temperature. As a consequence, surface radiances are not sensitive to the location of water vapor above 80 km. Only below this altitude, but above the troposphere, are surface radiances across the line sensitive to the location of water vapor. The radiance near line center increases relative to that in the wings as the water vapor is concentrated at higher altitudes within this range. Spatial resolution of retrieved water vapor profiles is about a pressure scale height.

Water vapor retrievals are not limited to the mesosphere, and have been obtained for the (upper) stratosphere as well. Any middle-atmosphere constituent with sufficient concentration and a rotational microwave line can be probed by this approach, for example, CO, ClO, HNO_3, and N_2O.
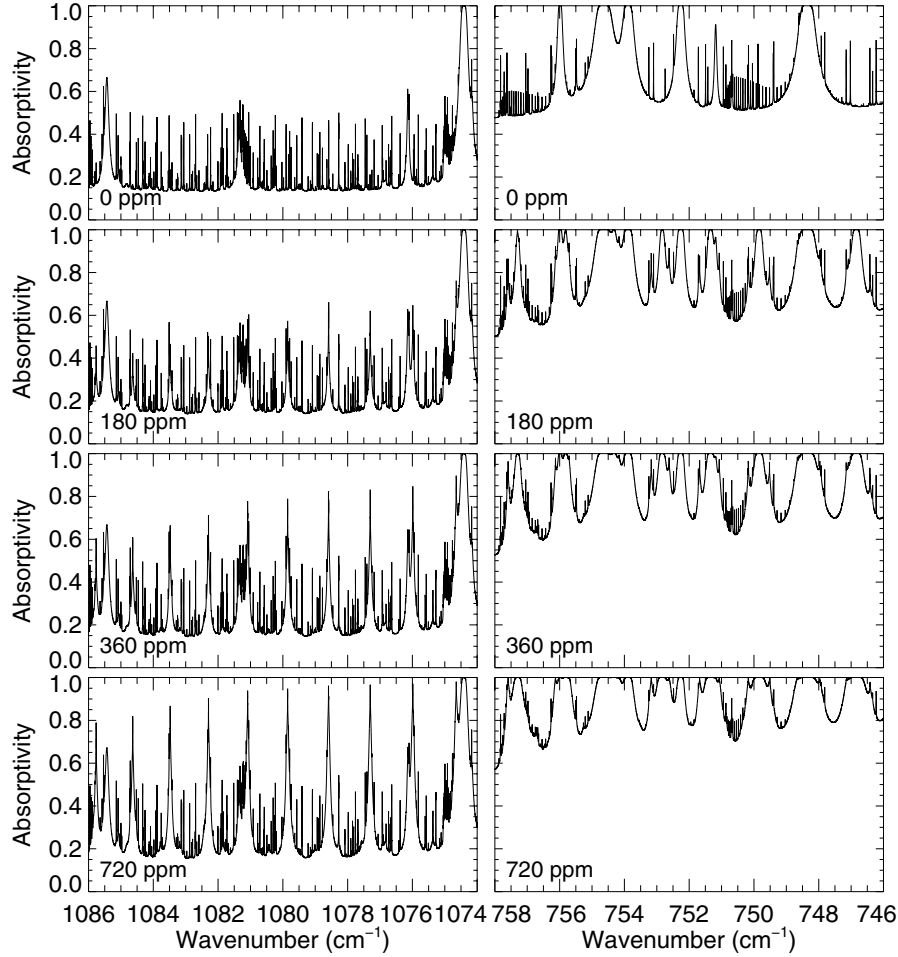
Figure 2.22: Absorptivity along a vertical path, for two spectral regions, for a mid-latitude summer temperature and pressure profile. The concentration of carbon dioxide (ppm = parts per million) for each absorptivity is indicated on the panels.

Line Absorptivity and Saturation

We now turn to absorption by a uniform layer of thickness h of a gas with a single, isolated absorption line. As previously, we divide the line into M equally spaced (in wavenumber) bins. Suppose that n_m photons in the mth bin are incident on the layer. The total line transmissivity is the ratio of the total transmitted to the total incident photons (reflection is assumed

negligible):

$$T = \frac{\displaystyle\sum_{m=1}^{M} n_m \exp(-\kappa_m h)}{\displaystyle\sum_{m=1}^{M} n_m}, \tag{2.137}$$

where κ_m is the absorption coefficient for the mth bin. If n_m is uniform

$$T = \frac{1}{M} \sum_{m=1}^{M} \exp(-\kappa_m h). \tag{2.138}$$

The *line absorptivity* is

$$A = 1 - T = 1 - \frac{1}{M} \sum_{m=1}^{M} \exp(-\kappa_m h) = \frac{1}{M} \sum_{m=1}^{M} A_m, \tag{2.139}$$

where $A_m = 1 - \exp(-\kappa_m h)$ is the absorptivity for the mth bin. Thus A is the average absorptivity of the line.

Consider first the limit of a weak line ($\kappa_m h \ll 1$ for all m):

$$A \approx \frac{1}{M} \sum_{m=1}^{M} \kappa_m h. \tag{2.140}$$

If we assume that the exponential in Eq. (2.127) is negligible, Eq. (2.140) becomes

$$A \approx \frac{hS}{M} \sum_{m=1}^{M} f_m = \frac{hS}{\tilde{\nu}_2 - \tilde{\nu}_1} \sum_{m=1}^{M} f_m \Delta\tilde{\nu} = \frac{hS}{\tilde{\nu}_2 - \tilde{\nu}_1}, \tag{2.141}$$

where $\Delta\tilde{\nu}$ is the bin width and $(\tilde{\nu}_1, \tilde{\nu}_2)$ are what we agree to call the (somewhat arbitrary) limits of the line. Because the line strength S is proportional to the number density of molecules, the line absorptivity in the weak line limit increases linearly with number density. But this can't go on indefinitely. In the strong line limit ($\kappa_m h \gg 1$ for all m), $A \to 1$. The line is saturated in the sense that a further increase in number density does not change the line absorptivity.

The spectral absorptivity, over two fairly broad spectral regions, corresponding to a mid-latitude summer profile (Figs. 2.10 and 2.11) but with varying amounts of uniformly mixed carbon dioxide, is shown in Fig. 2.22. The consequences of line saturation are evident for the two spectral regions. Figure 2.23 is a closer look at a narrow spectral region within each of the broad regions in Fig. 2.22. To show the behavior of carbon dioxide itself we use the absorption optical thickness with *no* carbon dioxide to determine what the absorptivity would be if carbon dioxide were the *only* absorbing gas. The line absorptivity for the two spectral regions of Fig. 2.23 is shown as a function of carbon dioxide concentration in Fig. 2.24.
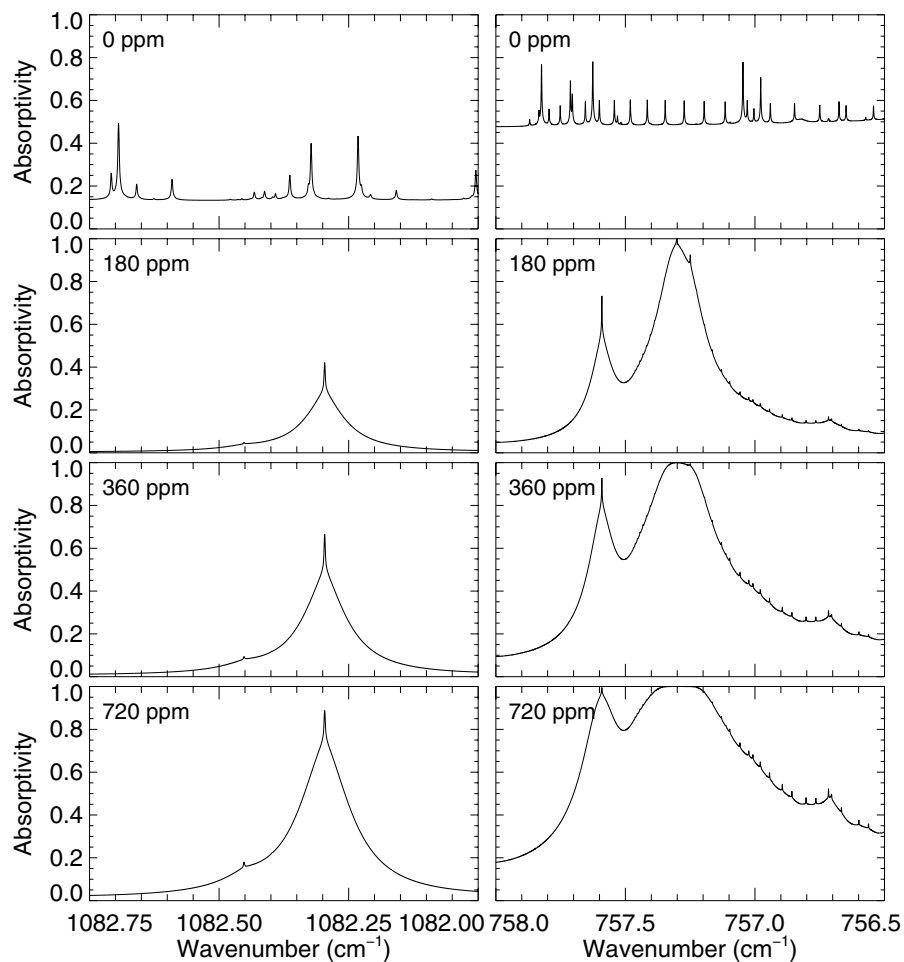
Figure 2.23: Absorptivity along a vertical path for two narrow regions within the spectral regions in Fig. 2.22. The concentration of carbon dioxide is indicated on the panels. Each curve is what the absorptivity would be if carbon dioxide were the only absorbing gas.

For the spectral region 1082–1082.75 cm^{-1} the line absorptivity increases almost linearly with concentration (weak line limit), whereas for the spectral region 756.5–758 cm^{-1}, the line absorptivity increases linearly with concentration at first but then increases with ever decreasing slope, gradually reaching the asymptotic value (strong line limit) 1.

If a carbon dioxide line occurs in a spectral region where lines from other molecular species are already saturated, carbon dioxide has little effect on absorptivity (Fig. 2.22). The converse of this is that carbon dioxide, indeed any infrared-active gas, has its greatest effect in regions where lines from other molecules are not saturated. Saturation of absorptivity does
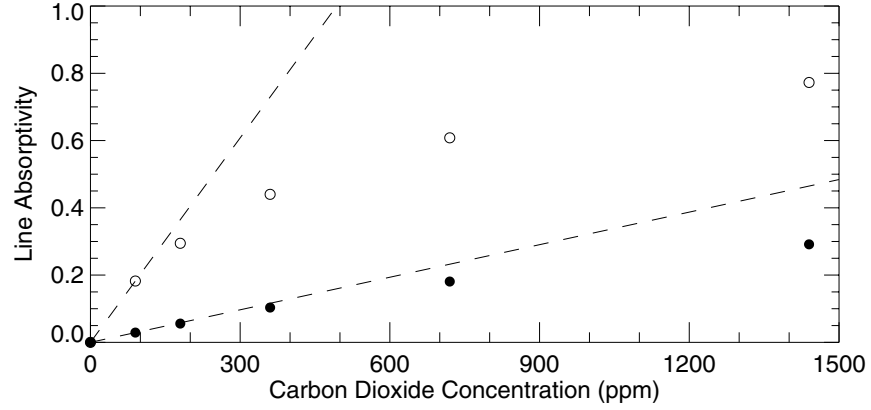
Figure 2.24: Line absorptivity of an atmosphere with carbon dioxide the only absorbing gas for the spectral regions 1082–1082.75 cm^{-1} (closed circles) and 756.5–758 cm^{-1} (open circles) as a function of carbon dioxide concentration (parts per million). Dashed lines represent linear fits to the results at the two lowest carbon dioxide concentrations.

not necessarily imply saturation of consequences. Emission from the atmosphere to Earth depends on emissivity (absorptivity). But it also depends on the atmospheric temperature profile. Although one might be tempted to think that increasing carbon dioxide can only eventually saturate all lines with no other effect, resulting in an upper limit on infrared radiation from the atmosphere, this assumes (probably incorrectly) that the temperature profile does not change. It is as unreasonable to expect an increase in carbon dioxide in the atmosphere to eventually result in some asymptotic value for infrared atmospheric irradiance [Eq. (1.74)] as to expect it to increase indefinitely [Eq. (1.79)].

2.9   Absorption by Particles

Although the form of Eq. (2.28) is indifferent to whether the illuminated object is a molecule or a particle, we write the symbol C_{abs} for the absorption cross section of a particle to signal that particles and molecules are different. Implicit in the definition of cross sections is that the irradiance of the incident illumination be constant over lateral dimensions large compared with the size of the particle. As with molecules the absorption cross section of a particle depends on its orientation (unless it is an isotropic sphere) and the state of polarization of the illumination (again, unless it is an isotropic sphere). The absorption cross section of a particle often is normalized by its geometrical cross sectional area G projected onto a plane perpendicular to the illumination. The resulting dimensionless quantity

$$Q_{\mathrm{abs}} = \frac{C_{\mathrm{abs}}}{G} \tag{2.142}$$

is called an *efficiency* or *efficiency factor* for absorption. The advantage of this normalization, namely, not having to fret about units, is outweighed by several disadvantages. In the first place, we expect quantities called efficiency factors to be less than or at most equal to 1 (100%), whereas it is not unusual to encounter overachieving particles with Q_{abs} *greater* than 1, sometimes appreciably greater, at some wavelengths. Also, some particles don't have well-defined or easily determined geometrical cross sections (e.g., soot aggregates, the kinds of particles found in smokes), and molecules most definitely do not. And the normalization in Eq. (2.142) is arbitrary. Why not normalize by total surface area or the square of mean chord length or . . . ? To be fair Q_{abs}, at least for a sphere, does have a physical meaning. At a given wavelength it is the emissivity of the particle, which underlies the assertion in Section 1.3 about emissivities greater than 1. But the efficiency factors for scattering and extinction (Sec. 3.5) have no physical meaning.

Both molecules and particles have well-defined absorption cross sections, in principle measurable; both molecules and particles have masses, which again are measurable. So the only really meaningful normalized cross section, by means of which we can compare molecules and particles, is cross section per unit mass. But if a particle has a definite volume v (a molecule does not), another normalized cross section that does have a physical meaning is the *volumetric* cross section, the cross section per unit particle volume. The absorption coefficient of a dilute suspension of N identical particles is

$$\kappa = NC_{\mathrm{abs}} = NvC_{\mathrm{abs}}/v = fC_{\mathrm{abs}}/v, \tag{2.143}$$

where $f = Nv$ is the volume fraction of particles in the suspension (i.e., the fraction of the total volume containing something you can kick). If any quantity deserves to be called an efficiency it is the volumetric absorption cross section C_{abs}/v. Equation (2.143) can be generalized to a suspension of non-identical particles:

$$\kappa = f \left\langle \frac{C_{\mathrm{abs}}}{v} \right\rangle, \tag{2.144}$$

where

$$\left\langle \frac{C_{\mathrm{abs}}}{v} \right\rangle = \frac{\displaystyle\sum_j f_j \left(C_{\mathrm{abs},j}/v_j \right)}{\displaystyle\sum_j f_j} = \frac{1}{f} \sum_j f_j \left(C_{\mathrm{abs},j}/v_j \right) \tag{2.145}$$

and the subscript j denotes anything that makes one particle different from another – shape, size, orientation, or composition.

We now turn to another advantage of expressing cross sections of particles as volumetric cross sections. Consider the simplest imaginable particle, an optically homogeneous slab of uniform thickness d and surface area A with dimensions much larger than λ illuminated at normal incidence by a beam with irradiance F. The law of exponential attenuation [Eq. (2.8)] strictly applies only to a medium without boundaries. Boundaries add the complication of reflections, with the attendant possibility of interference (see Sec. 3.4). We assume that the optical properties of the slab are sufficiently similar to those of the surrounding medium that

reflections are negligible and that the slab is sufficiently thick compared with the wavelength that interference is negligible. With these assumptions the rate W_a at which energy is absorbed by the slab is

$$W_a = FA\{1 - \exp(-\kappa_b d)\},\tag{2.146}$$

where we write the bulk absorption coefficient of the slab material as κ_b to distinguish it from the absorption coefficient [Eq. (2.144)] of a dilute suspension of particles composed of that material. If we further assume that $\kappa_b d \ll 1$ (weak absorption), Eq. (2.146) becomes

$$W_a \approx FAd\kappa_b.\tag{2.147}$$

From Eq. (2.28) the absorption cross section of the slab is therefore $Ad\kappa_b$ and its cross section per unit volume is

$$C_{abs}/v \approx \kappa_b.\tag{2.148}$$

At the other extreme $\kappa_b d \gg 1$ (strong absorption) and Eq. (2.146) yields

$$C_{abs}/v \approx 1/d.\tag{2.149}$$

For some materials Equations (2.148) and (2.149) for a slab particle are good estimates for the volumetric absorption cross section of some particles in the limits of weak and strong absorption. A further advantage of considering volumetric absorption by a particle is that it can be compared with the bulk absorption coefficient of its parent material. Sometimes the absorption spectrum of a particle is similar to that of its parent material, but sometimes exhibits not even a trace of a family resemblance.

If the distribution of sizes and shapes of particles in a suspension, all of identical composition, is such that they can fill all space (suspensions of identical spheres cannot), κ given by Eq. (2.144) should approach κ_b as f approaches 1. And indeed it does *if* the volumetric absorption cross section is given by Eq. (2.148), but definitely does not *if* it is given by Eq. (2.149). For strong absorption the volumetric absorption cross section of particles is dominated by their size and not the bulk absorption coefficient of their parent material. This signals caution in applying Eq. (2.144) to dense suspensions of particles. Unfortunately, we can't give precise criteria for what is meant by "dense", although we touch on this subject in Sections 5.1 and 5.3. Depending on the size of the particles relative to the wavelength, when a suspension of them becomes sufficiently dense, it is probably more realistic to look upon it as a slightly porous medium ($f \approx 1$), the absorption coefficient of which is approximately

$$\kappa = f\kappa_b.\tag{2.150}$$

This expression does have the correct asymptotic behavior as $f \to 1$. Thus Eq. (2.144) can be used with confidence for $f \ll 1$, and Eq. (2.150) for $f \approx 1$, but the values of f at which the transition from the one to the other begins and ends cannot be specified precisely.

For what materials and particles is Eq. (2.148) not a good approximation? Suspensions of small gold and silver particles (colloidal gold and silver) when illuminated by white light display upon transmission vivid colors completely unrelated to the appearance of the bulk
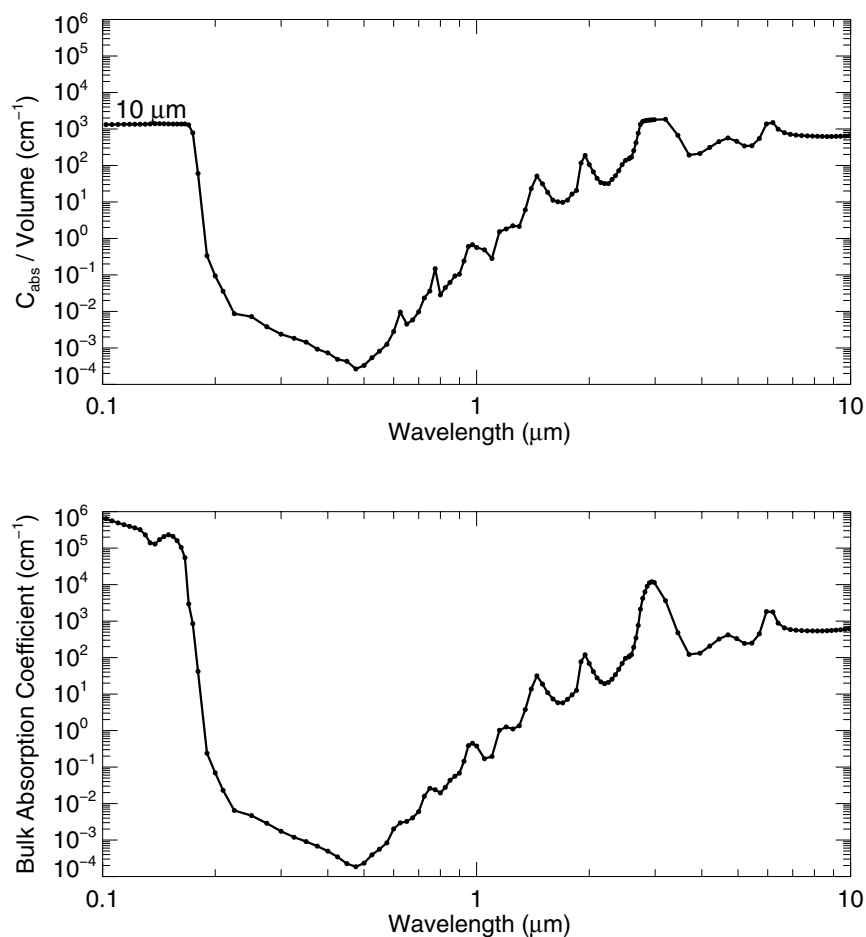
Figure 2.25: Volumetric absorption cross section spectrum of a 10 μm-diameter water droplet (top) compared with the bulk absorption spectrum of liquid water (bottom). For smaller diameters volumetric absorption in the UV more closely follows the bulk absorption coefficient; for larger diameters, peaks in the volumetric absorption spectrum beyond about 2 μm are flattened.

metals. The absorption spectrum of silver, in particular, is as nearly featureless as a dry lake bed, and although that of gold is more interesting, which is why gold is golden, absorption spectra of particles of these metals bear no resemblance whatsoever to those of their parent materials.

Alas, gold and silver particles are not commonly found in the atmosphere but particles of insulating crystalline materials (e.g., quartz, ammonium sulfate) are, and at infrared wavelengths the absorption spectra of such particles do not always dutifully follow those of their
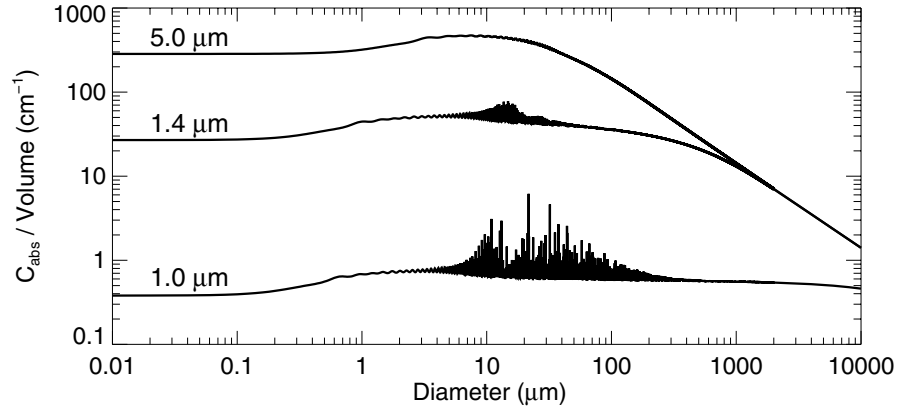
Figure 2.26: Volumetric absorption cross section for water droplets of varying diameter at the three wavelengths shown.

parent materials. But the differences are not so striking as they are for metallic particles, a matter of shifts of absorption peaks rather than their appearance seemingly from nowhere.

Water, however, is an example of a material for which Eqs. (2.148) and (2.149) often are good approximations. The volumetric absorption cross section of a $10\,\mu\mathrm{m}$-diameter water droplet (Fig. 2.25), calculated using the exact theory for a sphere (Sec. 3.5), is similar to the bulk absorption coefficient of water from visible to infrared. For droplets smaller than about $10\,\mu\mathrm{m}$ (not shown), the volumetric absorption cross section approaches more closely the bulk absorption coefficient in the ultraviolet ($< 0.2\,\mu\mathrm{m}$), whereas for droplets larger than about $10\,\mu\mathrm{m}$ the peaks in the infrared become more and more flattened. This flattening of the absorption spectrum of a particle is expected on physical grounds. As the bulk absorption coefficient of a particle increases, absorption is concentrated more and more in the outer layers of the illuminated particle, and what changes is not *how much* energy is absorbed but *where*.

Figure 2.26 shows the volumetric absorption cross section as a function of droplet diameter at three different wavelengths. For sufficiently small diameters, the cross section is approximately the bulk absorption coefficient and independent of diameter, as predicted by Eq. (2.148). With increasing diameter, the cross section decreases as the inverse of diameter and is independent of the bulk absorption coefficient. Moreover, the diameter at the transition from a constant to a decreasing value is smaller the greater the bulk absorption coefficient, which increases with increasing wavelength. All this is consistent with Eq. (2.149). The exceedingly fine structure in the absorption cross section at $\lambda = 1.0\,\mu\mathrm{m}$ is a consequence of interference (see Sec. 3.4). Note that this interference structure does not occur until the diameter is greater than the wavelength, and is increasingly suppressed with increasing bulk absorption coefficient. Equations (2.148) and (2.149) for a slab particle were obtained under the assumption of negligible interference. If we had included interference we would have obtained oscillations in the absorption cross section of a slab, interpreted as a consequence of interference between multiply-reflected waves. The greater the value of $\kappa_{\mathrm{b}} d$, the more these waves are damped (attenuated), and the more the interference pattern disappears.

2.9.1 Molecules and Particles: Similarities and Differences

As far as absorption (or scattering) is concerned a molecule is a particle of zero dimensions. Although molecules do indeed have extension in space they are fuzzy. In any interaction of electromagnetic radiation with matter, the relevant measuring stick is the wavelength, against which molecules are quite small, even for wavelengths well into the ultraviolet. The separate parts of molecules therefore radiate in unison. A corollary of this is that absorption by molecules and by small (compared with the wavelength) particles ought to be similar. And indeed they are, with some notable exceptions. The absorption spectrum of water vapor at infrared frequencies exhibits many narrow, closely spaced rotational lines (Fig. 2.12), whereas these lines vanish completely in bulk water, which we interpreted as collisional broadening taken to its extreme, and hence vanish from the absorption spectra of water droplets of all sizes. Although vibrational bands are broadened and shifted in going from vapor to the condensed phases (liquid and ice), they still are prominent in the infrared absorption spectrum of water (Fig. 2.25) and hence in the absorption spectra of small water droplets. Where molecules and particles go their separate ways is when particles become comparable with or greater than the wavelength. This additional degree of freedom (size) results in particle absorption spectra that may bear no resemblance to that of individual molecules or even the condensed phases of these molecules. A single particle larger than the wavelength and weakly absorbing can exhibit a series of narrow, closely spaced lines in its absorption spectrum, reminiscent of rotational bands in molecular spectra but arising from a completely different cause: interference.

Another difference between molecules and particles is that we have a hope of calculating the energy levels of molecules, and hence the frequency dependence of molecular cross sections, by quantum mechanics. Difficult, yes, but not impossible. Calculating the cross sections of particles, aggregations of many closely-packed interacting molecules, by quantum mechanics is essentially impossible. This would be like trying to forecast the weather using quantum mechanics – in principle not impossible, but in practice not advisable. To calculate absorption cross sections of particles we must use classical electromagnetic theory and obtain the electromagnetic properties of the materials of which they are composed from measurements. This is how the volumetric absorption cross sections for water droplets shown in Figs. 2.25 and 2.26 were obtained.

References and Suggestions for Further Reading

Examples of Stigler's law of eponymy, drawn mostly from mathematics, are in Stephen M. Stigler, 1999: *Statistics on the Table*, Harvard University Press, Ch. 14. Another name for this "supremely important law of the history of science" is "the Infinite Chain of Priority: Somebody Else Always Did It First" (Tony Rothman, 2003: *Everything's Relative and Other Fables from Science and Technology*, John Wiley & Sons, p. xiii.).

Pierre Bouguer's *Optical Treatise on the Gradation of Light* (1760) was translated, with an introduction and notes, by William Edgar Knowles Middleton, 1961, University of Toronto Press. For a sketch of Bouguer's life and work see Middleton's entry in *Dictionary of Scientific Biography*.

For Lambert's magnum opus see Johann Heinrich Lambert, 2001: *Photometria*, Illuminating Engineering Society of North America, translated from the Latin by David DiLaura, who, in the lengthy historical introduction (p. lxxxi), acknowledges that "Bouguer... was the first to recognize exponential decay in absorbing media."

For what Beer did (and did not) do see Heinz G. Pfeiffer and Herman A. Liebhafsky, 1951: The origins of Beer's law. *Journal of Chemical Education*, Vol. 28, pp. 123–5.

For a history of the exponential attenuation law see Dorothy R. Malinin and John H. Yoe, 1961: Development of the laws of colorimetry: a historical sketch. *Journal of Chemical Education*, Vol. 38, pp. 129–31. Despite its title, this paper is *not* about colorimetry (see Sec. 4.3). Its authors add yet another claimant to the law of exponential attenuation, that of F. Bernard, who in the same year as Beer (1852) "arrived at a similar relationship between absorptive capacity and concentration."

The absorption length for liquid water in Fig. 2.2 is from Table 1 in Marvin R. Querry, David M. Wieliczka, and David J. Segelstein, 1991: Water (H_2O), pp. 1059–77 in Edward D. Palik, Ed., 1991: *Handbook of Optical Constants of Solids*, Vol. II, Academic Press. That for ice is from the compilation by Stephen G. Warren, 1984: Optical constants of ice from the ultraviolet to the microwave. *Applied Optics*, Vol. 23, pp. 1206–25.

We have to be clear as to exactly what nonexponential attenuation, a topic of current interest, means. To us Eq. (2.22) is not an example of nonexponential attenuation but rather of nonuniform attenuation; the underlying law is still exponential. Nor is integrated attenuation over an absorption line (Sec. 2.1.3) nonexponential attenuation. Exponential attenuation may hold locally (point by point) but not globally (integrated). To us nonexponential attenuation (at a single wavelength) is a consequence of spatial correlations among absorbers (or scatterers) in a statistically homogeneous medium. This is the sense in which nonexponential attenuation is meant by Alexander B. Kostinski, 2001: On the extinction of radiation by a homogeneous but spatially correlated random medium. *Journal of the Optical Society of America A*, Vol. 18, pp. 1929–33, Alexander B. Kostinski, 2002: On the extinction of radiation by a homogeneous but spatially correlated random medium: reply to comment. *Journal of the Optical Society of America A*, Vol. 19, pp. 2521–25, and Raymond A. Shaw, Alexander B. Kostinski, and Daniel D. Lanterman, 2002: Super-exponential extinction of radiation in a negatively-correlated random medium. *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 75, pp. 13–20 in which one can find citations of relevant papers. Note in particular that Fig. 3 (dot diagrams similar to our Fig. 2.7) in the second paper by Kostinski shows at a glance the differences between positive correlation (sub-exponential), no correlation (exponential) and negative correlation (super-exponential), a graphic realization of the old adage that one picture is worth a thousand words.

Careful examination of Fig. 2.2 (see also Prob. 5.31) reveals small differences between absorption by liquid and solid water at wavelengths around 1 μm. These differences are the basis of methods for remotely determining cloud composition (ice or liquid water). See Peter

Pilewskie, 1989: Cloud Phase Discrimination by Near-Infrared Remote Sensing. Ph.D. dissertation, University of Arizona. This is a model of what a dissertation should be (graduate students take note). It is short (107 pages), blessedly free of acronyms and non-nutritious fluff, the figures are few, uncluttered and comprehensible at a glance, and, as evidenced by the following paper, it has not grown stale with age: Wouter H. Knap, Piet Stammes, and Robert B.A. Koelemeijer, 2002: Cloud thermodynamic-phase determination from near-infrared spectra of reflected sunlight. *Journal of the Atmospheric Sciences*, Vol. 59, pp. 83–96.

For journal articles on this method see Peter Pilewskie and Sean Twomey, 1987: Discrimination of ice from water in clouds by optical remote sensing. *Atmospheric Research*, Vol. 21, pp. 113–22; Peter Pilewskie and Sean Twomey, 1987: Cloud phase discrimination by reflectance measurements near 1.6 and 2.2 µm. *Journal of the Atmospheric Sciences*, Vol. 44, pp. 3419–21; and Peter Pilewskie and Sean Twomey, 1992: Optical remote sensing of ice in clouds. *Journal of Weather Modification*, Vol. 24, pp. 80–82. In this paper the authors present spectral measurements of transmission as well as reflection by clouds, and point out that transmission samples an entire cloud whereas reflection samples only its outer layers.

For a clear treatise on complex variables see Ruel V. Churchill, 1960: *Complex Variables and Applications*. McGraw-Hill.

The following are the bibliographic details for the formidable reading list at the beginning of Section 2.7: Edward U. Condon and George H. Shortley, 1967: *The Theory of Atomic Spectra*. Cambridge University Press; Gerhard Herzberg, 1950: *Molecular Spectra and Molecular Structure*: I. *Spectra of Diatomic Molecules*, 2nd ed. D. Van Nostrand; Gerhard Herzberg, 1945: *Molecular Spectra and Molecular Structure*: II. *Infrared and Raman Spectra of Polyatomic Molecules*. D. van Nostrand; Gerhard Herzberg, 1966: *Molecular Spectra and Molecular Structure*: III. *Electronic Spectra and Electronic Structure of Polyatomic Molecules*. D. Van Nostrand; Charles H. Townes and Arthur L. Schawlow, 1975: *Microwave Spectroscopy*. Dover.

The paper by Charles L. Braun and Sergei N. Smirnov, 1993: Why is water blue? *Journal of Chemical Education*, Vol. 70, pp. 612–14, could have been published in an archival journal (loosely translated from the Latin, archival means "no one reads this stuff") but instead its authors chose a journal devoted to exposition. The result is a paper that is a pleasure to read. When you finish it you only wish it were longer so that the pleasure will linger.

The temperature and pressure profiles in Fig. 2.11 were taken from R. A. McClatchey, R. W. Fenn, J. E. A. Selby, F. E. Volz, and J. S. Garing, 1972: Optical Properties of the Atmosphere, 3rd ed. AFCRL-72-0479, Air Force Cambridge Research Laboratories.

The spectral calculations in this chapter were done using the latest version of LBLRTM, a line-by-line code developed by Tony Clough and his collaborators. See, for example, Shepard A. Clough, Michael J. Iacono, and Jean-Luc Moncet, 1992: Line-by-line calculations of atmospheric fluxes and cooling rates: application to water vapor. *Journal of Geophysical Research – Atmospheres*, Vol. 97, pp. 15761–85. The spectroscopic parameters in this code come

from HITRAN, an acronym for high-resolution transmission molecular absorption database. The version most recently published (in a journal) is Laurence S. Rothman *et al.*, 2003: The HITRAN 2000 molecular spectroscopic database: edition of 2000 including updates through 2001. *Journal of Quantitative Spectroscopy and Radiative Transfer*, Vol. 82, pp. 5–44. The 2004 database eventually will be published in the same journal. We depart from our usual practice of giving (when known) the full names of all authors because this paper has 31 of them.

For an English translation of Einstein's 1916 paper Emission and Absorption of Radiation in Quantum Theory see Alfred Engel's 1997 translation *The Collected Papers of Albert Einstein, Vol. 6, The Berlin Years: Writings, 1914–1917*, pp. 212–16. Einstein did not use the terms spontaneous and induced emission, absorption, or photons in this paper. What we call induced emission and absorption Einstein associated with the "work done by the electric field on the resonator", the "energy change due to incident radiation."

For a discussion of the Einstein coefficients see, for example, Robert Martin Eisberg, 1961: *Fundamentals of Modern Physics*. John Wiley & Sons, p. 458–60.

The treatise by Richard M. Goody and Yuk L. Yung, 1989: *Atmospheric Radiation: Theoretical Basis*. Oxford University Press has been invaluable to us in writing this chapter. Textbooks can serve up nonsense when their authors leave their areas of expertise and carelessly discuss subjects with which they are not familiar. But they can also serve up gems when their authors concentrate on clear exposition of subjects with which they are familiar. Goody and Yung's discussions of absorption by atmospheric molecules falls into the latter category. Their sections on thermal emission, vibration-rotational spectra, and line shapes served as roadmaps that we came back to time and again as we wrote our sections on absorption, induced emission and spontaneous emission.

Schematic diagrams of the broadening of spectral lines because of interactions are standard fare in solid state physics treatises. See, for example, Figs. 1 and 2 on page 225 of Frederick Seitz, 1987: *The Modern Theory of Solids*. Dover. Seitz notes that "If we have a large number of free atoms that are infinitely separated from one another and are stationary, their electronic levels are discrete so that the energy levels of the entire system are discrete. Thus the width of the emission lines is determined entirely by natural broadening. The atoms interact, however, if they are brought within a finite distance of one another, and this interaction broadens these levels of the entire system..."

The history of the development of line shapes with far wings that accurately account for continuum water vapor absorption is an interesting one. To learn more about this see Shepard A. Clough, Frank X. Kneizys and R. W. Davies, 1989: Line shape and the water vapor continuum. *Atmospheric Research*, Vol. 23, pp. 229–41, and Qiancheng Ma and Richard H. Tipping, 1991: A far wing line shape theory and its application to the water continuum absorption in the infrared region. 1. *J. Chem. Phys*, Vol. 95, pp. 6290–6301. For measurements of water dimer concentrations in the atmosphere and estimates of the weak absorption by them see

Klaus Pfeilsticker, Andreas Lotter, Christine Peters, and Hartmut Bösch, 2003: Atmospheric detection of water dimers via near-infrared absorption. *Science*, Vol. 300, pp. 2078–80.

For a discussion of retrieving water vapor profiles in the mesosphere see Richard M. Bevilacqua, 1982: An Observational Study of Water Vapor in the Mid-Latitude Mesosphere. Doctoral Thesis, The Pennsylvania State University. Parts of this thesis are more readily available in Richard M. Bevilacqua, John J. Olivero, Philip R. Schwartz, Christopher J. Gibbins, Joseph M. Bologna, and Dorsey J. Thacker, 1983: An observational study of water vapor in the mid-latitude mesosphere using ground-based microwave techniques. *Journal of Geophysical Research*, Vol. 88, pp. 8523–34.

For more on absorption by particles see the treatises on scattering by particles cited in the references at the end of Chapter 3. Scattering and absorption by particles go hand in hand.

Problems

2.1. Given the assumption that the frequency of a microwave oven is chosen so that the absorption length in water is comparable with the size of objects to be heated in it, you should be able to estimate how long it takes to heat a cup of coffee from room temperature to what we call the *McDonald's point*, the temperature (85 °C) to which McDonald's *used* to heat its coffee. The power of a typical microwave oven is 1000–1500 W. The specific heat capacity of liquid water is about $4200 \, \text{J} \, \text{kg}^{-1} \, \text{K}^{-1}$, its density $1000 \, \text{kg} \, \text{m}^{-3}$.

HINT: There is no need to make this problem complicated. All that is wanted is an estimate. The first time we did this problem we made crude approximations, guessed at some of the relevant physical parameters, and yet still came up with a number that made sense.

2.2. Morning fog often disappears soon after sunrise, and as a consequence you often hear it said that the sun "burned off the fog". Because this explanation seems plausible, its correctness is rarely questioned. But can sunlight really vaporize fog droplets in less than an hour? You should be able to answer this question quantitatively. State all assumptions.

The density of liquid water is about $1000 \, \text{kg} \, \text{m}^{-3}$ and about 2.5×10^6 J is required to evaporate 1 kg of liquid water. Over the solar spectrum the absorption cross section per unit volume of a cloud droplet is approximately equal to the bulk absorption coefficient of pure water. At Earth's surface, most of the solar irradiance lies at wavelengths less than about $2.5 \, \mu\text{m}$.

HINT: To determine the exact rate at which a cloud droplet absorbs solar radiation would require considerable computation. But this isn't necessary if all that is wanted is to determine if this rate is sufficient to evaporate a droplet in an hour or less. First set up the problem as exactly as you can. Then make approximations, always erring toward overestimating the rate of absorption. If after making such approximations you obtain an evaporation time much greater than an hour, more detailed calculations will not change your conclusion.

2.3. Physical quantities can be represented as complex variables provided they are subjected to linear operations only. Nevertheless, for some purposes one may multiply the complex representations of real quantities without introducing any error. Show that the time average of

the product of two time-harmonic quantities with complex representations $A \exp(-i\omega t)$ and $B \exp(-i\omega t)$, where A and B are complex numbers, is

$$\frac{1}{2}\Re\{AB^*\},$$

where the asterisk denotes the complex conjugate.

2.4. A radiation thermometer can be used to measure the temperature of the ocean (from above its surface, of course). But temperature varies with depth in the ocean, so what temperature does the radiation thermometer measure? Figure 2.2 will help you answer this question.

2.5. Estimate the highest natural frequency of a (macroscopic) mechanical oscillator that you are likely to encounter in your everyday lives.

2.6. Find an expression for the resonant frequency of two one-dimensional oscillators of different masses m_1 and m_2 connected by a spring with constant K. Find the normal modes of motion of this system and give them a simple physical description.

2.7. With Eq. (2.47) you can express $\cos \vartheta$ and $\sin \vartheta$ in terms of $\exp(i\vartheta)$ and $\exp(-i\vartheta)$, then derive all those trigonometric identities you may have slaved over in high-school trigonometry. Begin by showing that $\sin^2 \vartheta + \cos^2 \vartheta = 1$, then move on to expressions for the cosine and sine of the sum of angles, and so on.

2.8. Define the transmissivity of a slab as the ratio of the transmitted to the incident radiant energy. Is the transmissivity of a given slab for a diffuse source greater than, equal to, or less than that for a monodirectional source? The spectra of the two sources are identical. In answering this question, first ignore reflection. How does your answer depend on the angle of incidence of the monodirectional source?

Suppose that you do not ignore reflection. In which way, if any, does this change your conclusions about the relative transmissivities of a slab for monodirectional and diffuse radiation?

2.9. Almost everyone these days knows that a layer of ozone in the stratosphere absorbs ultraviolet (UV) radiation from the sun. Suppose that above this ozone layer there is a layer of scattering particles as a consequence of volcanic eruptions. One might expect this particle layer together with the ozone layer to reduce the solar UV irradiance at the surface even more than does the ozone layer alone. It has been suggested, however, that the particle layer could result in a greater UV irradiance at the surface. Where (latitude) and when (season) might this be true? Explain your answer. You may take the particle layer to be of uniform thickness and infinite in lateral extent. The trivial solution to this problem is obtained by taking the sun to be below the horizon.

HINT: You do not need to know anything more about scattering by particles than that it exists. This problem, which is related to the previous one, should be easier if you draw a diagram.

2.10. We stated without proof in Section 1.5 that the infrared emissivity of a layer of water is similar to that of a cloud of water droplets. Now you have the tools necessary to show this. A typical liquid water path value for a thick cloud is of order 1 cm. That is, if all the water in the cloud were compressed into a uniform layer it would be about 1 cm thick. The diameter of a typical cloud droplet is about 10 μm. The absorption cross section per unit volume of such a

droplet (Fig. 2.25) will help you to show that the infrared emissivity of a cloud, like that of a layer of water, is close to 1 over a broad range of wavelengths.

2.11. The quantity κh in Section 2.2 is the (normal) absorption optical thickness of the (slab) atmosphere (see Sec. 5.2 for a discussion of optical thickness). Strictly, the normal optical thickness is the integral of κ along a radial path through the entire atmosphere. Estimate the absorption optical thickness of the clear atmosphere for the wavelength range (10 μm) of the infrared thermometer used to make the brightness temperature measurements shown in Fig. 2.4. What is the corresponding emissivity? Assume that the (absolute) temperature of the atmosphere is constant. This is not a bad assumption given that water vapor is the major contributor to the absorption optical thickness, the scale height (e-folding distance) for the decrease of water vapor concentration is 1–2 km, and over this distance the absolute temperature decreases by only about 5%. At 10 μm and for temperatures around 300 K, the exponential in the Planck function is much greater than 1.

2.12. In Section 2.9 we derive approximate expressions for the absorption cross section per unit volume of a normally illuminated slab particle in the limits of weak ($\kappa_b d \ll 1$) and strong absorption ($\kappa_b d \gg 1$), where d is the thickness of the slab. Show that, subject to the same underlying assumptions, these expressions are valid for a particle of more general shape. Take the particle to be homogeneous (no holes) and bounded by a surface such that no line in the direction of the beam intersects it in more than two points. For such a particle, d in Eq. (2.148) is the average distance through the particle along the direction of the incident beam; d in Eq. (2.149) is A/v, where A is the cross-sectional area of the particle projected onto the beam and v is its volume.

HINTS: A sketch is essential. Consider a thin tube, its axis parallel to the beam, intersecting the particle. Determine the net amount of incident radiant energy absorbed by this tube (energy in minus energy out). Then add up (integrate) the total for all the tubes into which the particle can be considered to be composed.

2.13. This is Problem 2.2 stated a different way. Show that the radiative equilibrium temperature of a typical cloud droplet exposed to direct sunlight cannot be more than a small fraction of a degree greater than what its temperature would be in the shade (all else being equal). Figure 2.25 is essential for this problem.

2.14. Our local newspaper published a front-page article about auroras that were on seen two successive nights in our county. According to the article, "Auroras...occur when charged particles from the sun interact with atoms of nitrogen and oxygen in the upper atmosphere, giving off particles called photons that appear red or green.... Light is created when the atoms in the atmosphere attempt to cool down." Please discuss.

2.15. We state in Section 2.7.1 that we calculated the measured spectral shift in the visible absorption spectrum of water vapor as a consequence of the different masses of heavy water (D_2O) and ordinary water (H_2O). With a few hints and having done Problem 2.6 you should be able to duplicate our calculation. Braun and Smirnov attribute the absorption peak in the visible at 700 nm and in the infrared at 1000 nm to the combination-overtone $\nu_1 + 3\nu_3$, where ν_1 is the symmetric OH vibrational frequency and ν_3 is the antisymmetric OH vibrational frequency. Both frequencies are nearly equal (within a few percent).

2.16. Suppose that we have a gas in equilibrium at temperature T of N molecules that can exist in only one of two states, upper with energy E_u and lower with energy E_ℓ. Show that at sufficiently low temperature essentially all of the molecules are in the lower state and that the number of molecules in this state decreases with increasing temperature. Also show that at sufficiently high temperature *not* all of the molecules are in the upper state.

HINT: You need Eq. (2.95).

2.17. In 1852 Beer published an attenuation law for absorbing aqueous solutions according to which the diminution λ in *amplitude* of a beam upon transmission through a solution over a path of length D (in decimeters) is given by $\lambda = \mu^D$. Show that this leads to an exponential law for attenuation of irradiance. What we call transmissivity is λ^2. For more on this see Heinz G. Pfeiffer and Herman A. Liebhansky, 1951: The origins of Beer's law. *Journal of Chemical Education*, Vol. 28, pp. 123–125. These authors argue that Beer did not recognize that concentration and length are symmetric variables, that is, the absorption coefficient κ (in modern notation) of an absorbing solution is proportional to the concentration of the absorbing solute, and hence it is the product of concentration and path length that determines the transmissivity. What would have been the form of Beer's law if Beer had recognized this symmetry?

2.18. We are often told in textbooks on classical mechanics that the energy of a body is defined only to within an additive constant. In fact, similar statements are made even in treatises on quantum mechanics. For example, on page 5 of Herzberg's *Molecular Spectra and Molecular Structure I. Spectra of Diatomic Molecules* he states "The energy E of the atom contains an arbitrary additive constant." Yet such statements are never, to our knowledge, made about the energy of a photon. It is $h\nu$. Period. Why the difference?

2.19. The Maxwell–Boltzmann distribution [Eq. (2.8)] for molecular kinetic energies E contains the factor $\sqrt{E}$, which makes it difficult to determine analytically the fraction of molecules with kinetic energies greater than a particular value. But there is a ploy we can resort to if we are willing to settle for an exactly integrable inexact theory. We note in Section 2.8 that when the M–B distribution for energy E is transformed to a distribution for speed v [see also Prob. 3.13] the factor v^2 multiplying the exponential reflects the density of states. The number of points (states) in velocity space corresponding to (approximately) the same kinetic energy is proportional to v^2 because the total volume in velocity space between v and $v + \Delta \nu$ is proportional to v^2. Knowing this, you should be able to determine the distribution function for the kinetic energies of molecules constrained to lie in a *plane*. With this result you can then determine, analytically, the fraction of these molecules with kinetic energies greater than a given value. Then answer the following question. For a given energy $\overline{E}$ appreciably greater than $k_\mathrm{B}T$, how does the fraction of molecules with energies greater than $\overline{E}$ change if the *mean* kinetic energy is changed by only a factor of, say, 2? To answer this question is to obtain insight into why rates of chemical reactions depend so strongly on temperature. This problem was inspired by Cyril N. Hinshelwood, 1940: *The Kinetics of Chemical Change*, Oxford University Press, p. 5–6, an uncommonly good writer (and eminent scientist), whose (1951) *The Structure of Physical Chemistry*, Oxford University Press is, in places, lyrical (see, e.g., the opening section of the first chapter).

2.20. A chemical reaction occurs when two (or more) stable molecules unite to form another stable molecule. If the reacting molecules are stable then their potential energies must be

minima. And if the product is stable, it must have a potential energy minimum. The system, so to speak, makes a transition from one valley to another, and in so doing crosses a mountain. The energy required to cross this mountain, which comes from collisions, is the activation energy. Typical activation energies are around 30 times k_BT at room temperature. With this result and the previous problem you should be able to obtain a rule of thumb of chemists we mention in Section 2.7.1: increasing temperature by $10\,°C$ doubles the rate of chemical reactions. As with the previous problem, this one was inspired by Cyril Hinshelwood's book on chemical kinetics (p. 41).

2.21. How does the absorption cross section of atmospheric molecules compare with their geometrical cross section? Keep in mind that the diameter of a molecule is not a precisely defined quantity.

2.22. We note that some of the differences in Fig. 2.15 between the the results of simple and line-by-line calculations for the rotational levels, and hence absorption cross section, of carbon monoxide are a consequence of different isotopes of carbon. What about the isotopes of oxygen? Estimate the difference in the energy levels, and hence photon energies, for the most abundant isotopic forms of CO. Compare your results with Fig. 2.15. Also, estimate the relative contributions (at the peak of a line) of these isotopes to the average absorption cross section for naturally occurring carbon monoxide. Compare your result with the line-by-line calculations.

HINT: To do this problem you will have to dig up atomic masses and abundances of carbon and oxygen isotopes.

2.23. What is the fundamental difference between stimulated emission and scattering (in Sec. 1.2 we describe scattered radiation as being "stimulated" by an external source)?

2.24. At approximately what pressure altitude would the line shape for the water vapor line shown in Fig. 2.18 be dominated by Doppler broadening?

HINT: No calculations are needed.

2.25. In the discussion of Fig. 2.18, which shows line shapes of a water vapor band, we note that at the surface the foreign-broadened half-width is 0.0755 cm^{-1} whereas the self-broadened half-width is 0.3580 cm^{-1} (for the same temperature and pressure), a factor of almost 5 greater. Why the big difference?

2.26. Show that if the Doppler and Lorentz line shape functions [Eqs. (2.135) and (2.134), respectively] are normalized, so is their convolution [Eq. (2.136)].

2.27. Figure 2.23 shows absorptivity over narrow spectral regions by an atmosphere containing different concentrations of carbon dioxide. A line-by-line computer program contains all absorbing gases, most notably water vapor. How does one correct the total absorptivity in order to obtain that for an atmosphere with only carbon dioxide as the absorbing gas? What assumption has to be made?

2.28. Show that the effect of an increase in the concentration of an infrared-active molecule on emission from the atmosphere (i.e., on emissivity) is least, all else being equal, in spectral regions where the transmissivity is least. For simplicity, consider only the normal emissivity and assume a uniform atmosphere of finite thickness.

2.29. James Howard Kunstler, on page 139 of *The Long Emergency: Surviving the Converging Catastrophes of the Twenty-First Century*, asserts that "methane freed into the atmosphere is a ten times more effective greenhouse gas than carbon dioxide." Discuss this assertion.

2.30. Problem 1.5 asks you to estimate plausible maximum and minimum values at Earth's surface of radiation emitted downward from the atmosphere. Why is it easier to estimate an upper limit than a lower limit?