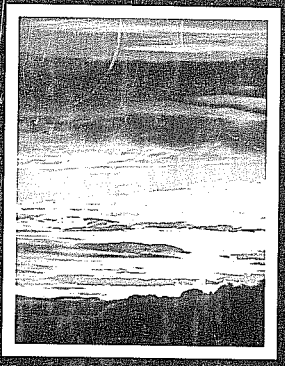
An Introduction to Atmospheric Radiation

SECOND EDITION



K. N. LIOU

INTERNATIONAL GEOPHYSICS SERIES, VOLUME 84

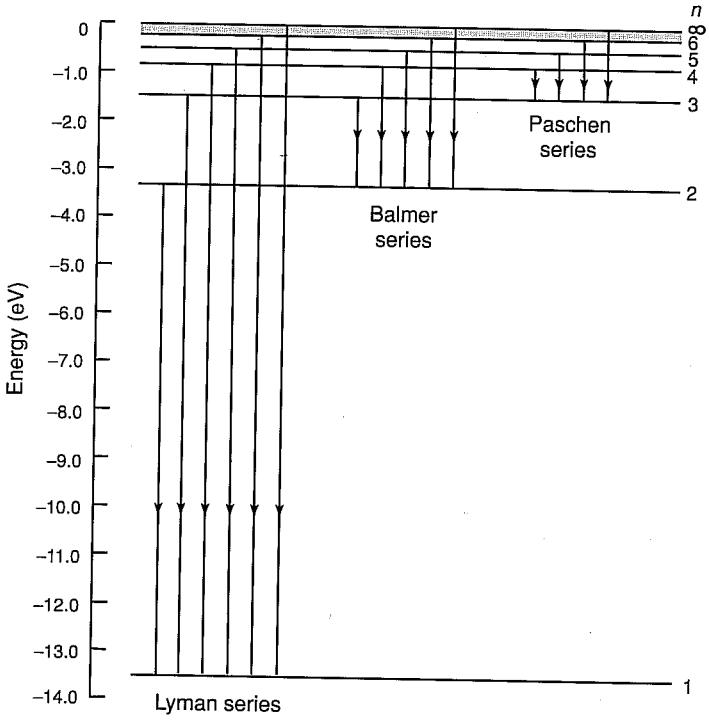


Figure 1.9 Energy level diagram for a hydrogen atom showing the quantum number n for each level and some of the transitions that appear in the spectrum. An infinite number of levels is crowded in between the levels marked $n = 6$ and $n = \infty$.

where j and k are integers defining, respectively, the lower and higher energy states. Figure 1.9 shows the energy diagram for hydrogen. In the field of spectroscopy, energy is usually given in units of electron volts (eV) or in units of wavenumber (cm^{-1}). An electron volt is the energy acquired by an electron accelerated through a potential difference of one volt, and is equivalent to 1.602×10^{-19} J. Exercise 1.11 requires the derivation of Eq. (1.3.3) based on the definitions of kinetic and potential energies of the system.

Each quantum jump between fixed energy levels results in the emission or absorption of a characteristic frequency or wavelength. These quanta appear in the spectrum as emission or absorption lines. For the simple hydrogen atom described previously the line spectrum is relatively simple, whereas the spectra of water vapor, carbon dioxide, and ozone molecules are considerably more complex.

1.3.1.2 Vibrational and Rotational Transitions

In the preceding discussion, we used the electronic transitions of the hydrogen atom to illustrate emission and absorption. It is now helpful to introduce the ways in which a molecule can store various energies. Any moving particle has kinetic

energy as a result of its motion in space. This is known as *translational energy*. The average translational kinetic energy of a single molecule in the x, y, and z directions is found to be equal to $KT/2$, where K is the Boltzmann constant and T is the absolute temperature. A molecule, composed of atoms, can rotate, or revolve, about an axis through its center of gravity and, therefore, has *rotational energy*. The atoms of the molecule are bounded by certain forces like springs such that the individual atoms can vibrate about their equilibrium positions relative to one another. The molecule, therefore, will also have *vibrational energy*. These three molecular energy types are based on a rather mechanical model of the molecule that ignores the detailed structure of the molecule in terms of nuclei and electrons. It is possible, however, for the energy of a molecule to change as a result of a change in the energy state of the electrons of which it is composed, as demonstrated by Bohr's model. Thus, the molecule has *electronic energy*. The last three energy types are quantized and take only discrete values. The absorption and emission of radiation takes place when the atoms or molecules undergo transitions from one energy state to another. In general, these transitions are governed by selection rules.

In radiative transitions, the molecule must couple with an electromagnetic field so that energy exchanges can take place. This coupling is generally provided by the electric dipole moment of the molecule. If the effective centers of the positive and negative charges of the molecule have nonzero separation, then the dipole moment exists. Radiatively active gases in the infrared, such as H_2O and O_3, have permanent electric dipole moments due to their asymmetrical charge distributions. Linear molecules such as N_2 and O_2, however, are inactive in the infrared because of their symmetrical charge distributions. However, they have weak magnetic dipole moments that allow radiative activities to take place in the ultraviolet and, to a lesser extent, in the visible region.

Rotational energy changes are relatively small, with a minimum on the order of $1\ \text{cm}^{-1}$ (see the conversion to energy in Exercise 1.12). For this reason, pure rotational lines occur in the microwave and far-infrared spectra. Many of the rotational energy levels above the lowest level are populated at terrestrial temperatures. Changes in vibrational energy are generally greater than $600\ \text{cm}^{-1}$, which is much larger than the minimum changes in rotational energy. Thus, vibrational transitions never occur alone but are coupled with simultaneous rotational transitions. This coupling gives rise to a group of lines known as the vibrational–rotational band in the intermediate infrared spectrum. An electronic transition typically involves a few electron volts ($\sim 10^4\ \text{cm}^{-1}$) of energy. Because a high-energy photon is required for the transition, absorption and emission usually occur in the ultraviolet or visible spectrum. Atoms can produce line spectra associated with electronic energy. Molecules, however, can have two additional types of energy, leading to complex band systems.

In Subsection 1.3.1.1, we discussed the physical meaning of stationary states for a hydrogen atom. Schrödinger (1926) first introduced the idea of stationary states corresponding to standing matter waves and used this idea as the foundation of *wave mechanics*. In quantum mechanics, to determine the energy states produced by vibrational and rotational transitions, a term referred to as the *Hamiltonian* operator, H,

was introduced as a convenient operator by replacing variables in the classical expression for the energy, E, of a system composed of the atomic nuclei and electrons that form a molecule. Schrödinger's equation can be written in terms of the first-order differential equation involving the wave function and the Hamiltonian operator, as shown in Appendix B. The Hamiltonian may be linearly separated into a time-dependent term and a time-independent term. The stationary states of the molecules can be deduced from the time-independent term, giving discrete eigenvalues (energy levels), E_n, and eigenfunctions, φ_n. Transitions between energy levels result in the absorption and emission of photons with frequency $\tilde{\nu}$ following Planck's relation. The time-dependent term may be treated as a perturbation from which the rate of change of the probability that a stationary state is occupied can be evaluated.

The Hamiltonian operator for the harmonic-oscillator rigid rotator is separable for vibrational and rotational motions so that energies may be added for a combined state. For the rotational states, the kinetic energy of a rigid rotating dipole is equal to one-half the product of angular momentum, L, and angular velocity, ω, i.e., $L\omega/2$, where $L = I\omega$ and I is the moment of inertia. From the solution of the time-independent Schrödinger equation, the quantum restrictions on angular momentum are given by

$$L = \frac{h}{2\pi}[J(J+1)]^{1/2}, \tag{1.3.5}$$

where J is the rotational quantum number (an integer). Thus, the quantized rotational energy can be written as

$$E_J = BhcJ(J+1), \tag{1.3.6}$$

where $B = h/8\pi^2 Ic$ is the rotational constant. This expression is valid for a rigid rotating dipole assuming spherical tops or linear molecules. For asymmetric tops, an additional term is required. The selection rule for radiation transition is governed by $\Delta J = \pm 1$, applicable to the harmonic-oscillator rigid-rotator model. From Planck's relation in Eq. (1.3.1), the spectral line location can be derived and is given by $\nu = 2BJ'$ (cm^{-1}), where J' can be any quantum number. Because of the selection rule, the separation in wavenumber of adjacent lines is simply $2B$ (cm^{-1}), as shown in Fig. 1.10a. As noted above, because of the small energy of a rotational transition, pure rotational spectra occur only in the far infrared and microwave regions.

For vibrational states, the quantized energy levels for a harmonic vibration are given by

$$E_\upsilon = h\tilde{\nu}_k(\upsilon_k + 1/2), \tag{1.3.7}$$

where υ_k is the vibrational quantum number (an integer) and subscript k denotes the normal modes. For triatomic molecules such as H_2O and O_3, there are three normal modes, referred to as *fundamentals*. For linear molecules such as CO_2 and NO_2, there are four fundamentals, but two orthogonal bending modes are degenerate and so only three fundamentals exist (see Fig. 3.3). The term *degenerate* is used to denote states with the same energy but with different sets of quantum numbers.

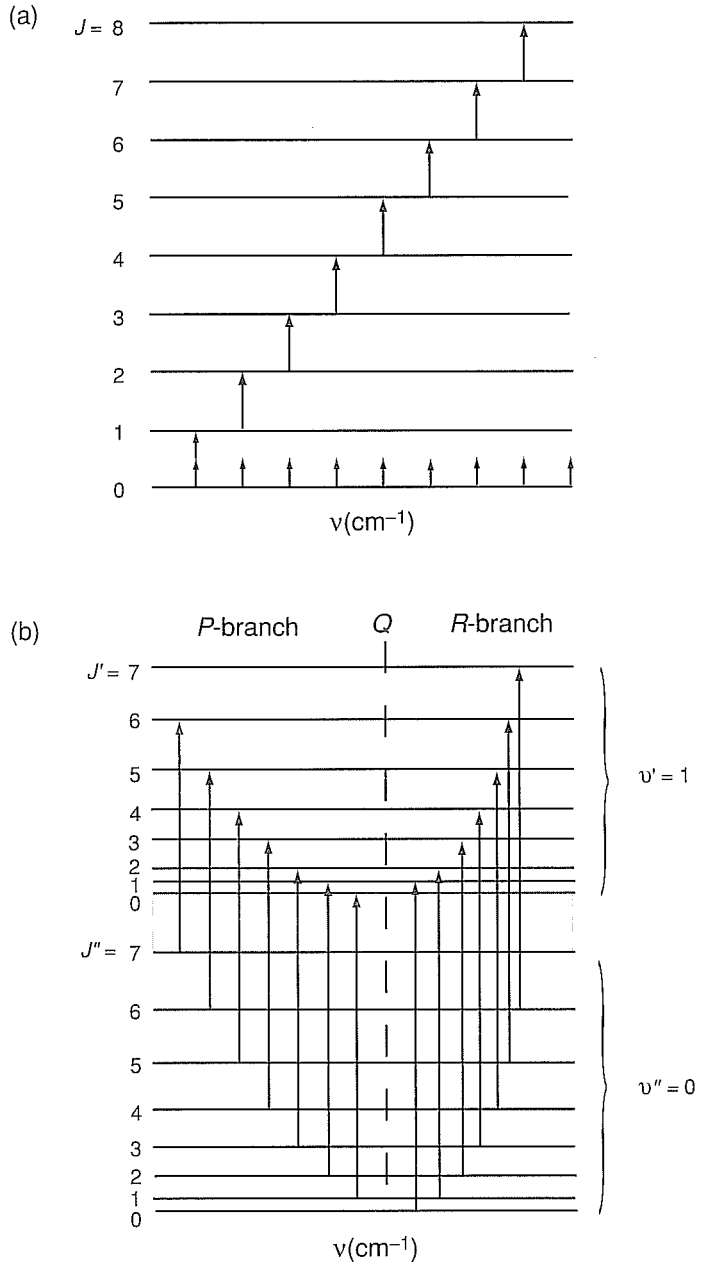


Figure 1.10 (a) Rotational transition following the selection rule $\Delta J = +1$ and equally spaced spectral lines in wavenumber. (b) Simultaneous vibrational and rotational transitions where $\Delta J = -1$ produces the P-branch and $\Delta J = +1$ generates the R-branch. $\Delta J = 0$ shows the Q-branch that overlaps with the vibrational wavenumber, but see text for discussion.

Chapter 3 Absorption and Scattering of Solar Radiation in the Atmosphere

3.1 Composition and Structure of the Earth's Atmosphere

It is now generally accepted that the terrestrial planets were formed by the accretion of solid materials that condensed from the solar nebula about 4.5 billion years ago (BY). The earth's present atmosphere is believed to be a secondary atmosphere that was generated from volatile compounds contained within the solid planetesimals from which the earth formed. Any primary atmosphere that was captured must have been lost because the cosmic abundances, the composition of most stars including the sun, which contains about 90% H and 10% He by mass, had been significantly depleted. The heavy bombardment of the earth ended about 3.8 BY ago and life was probably extant by 3.5 BY, at which time the atmosphere might have contained CH_4 and NH_3. The post-heavy bombardment atmosphere was probably dominated by CO_2 and N_2 with traces of CO and H_2O, but lacking free O_2, referred to as a weakly reduced atmosphere, associated with volcanic activities. Liquid water is believed to have existed on the earth's surface. A large amount of CO_2 or other greenhouse gases is likely to have evolved to compensate for the faint young sun at about 3.5–3.8 BY. It has been suggested that the biota played an integral role in controlling atmospheric CO_2. As a consequence of photosynthesis and organic carbon burial, atmospheric O_2 levels rose naturally. The major increase appears to have occurred between about 1.9 and 2.2 BY. The level of free O_2 is also believed to be associated with the formation of the ozone layer that provided an effective screen for ultraviolet solar radiation. In the following, we define the region of the earth's atmosphere and discuss its present composition.

First

Second

Bit of a co-incidence. Must have been some other stuff around to absorb UV, or early life was UV resistant.

3.1.1 Thermal Structure

To describe the interaction of the earth's atmosphere with solar radiation, we must first understand the atmosphere's composition and structure. For the purpose of defining the region of the atmosphere associated with the absorption and scattering of sunlight, we first present the vertical temperature profile for the standard atmosphere, which is shown in Fig. 3.1. This profile represents typical conditions in the mid-latitude regions. According to the standard nomenclature defined by the International Union of Geodesy and Geophysics (IUGG) in 1960, the vertical profile is divided into

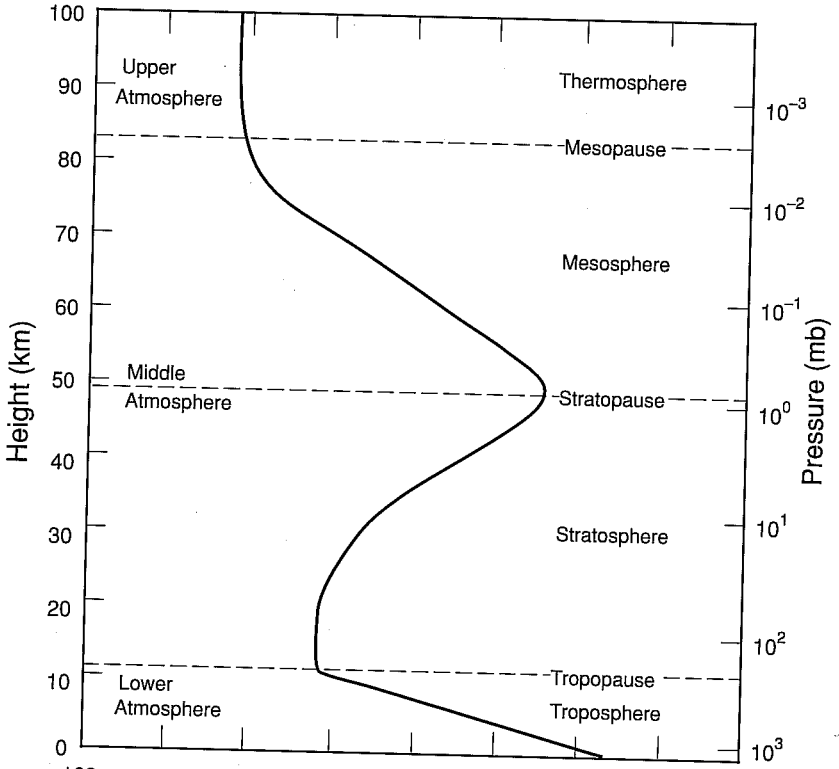
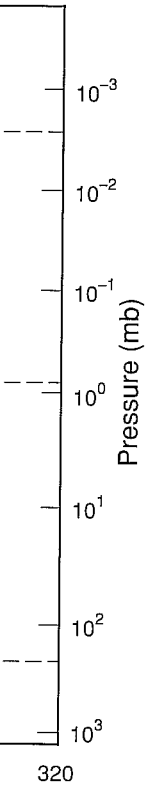


Figure 3.1 Vertical temperature profile after the U.S. Standard Atmosphere and definitions of atmospheric nomenclature.

four distinct layers as displayed in Fig. 3.1. These are the *troposphere, stratosphere, mesophere,* and *thermosphere*. The tops of these layers are called the tropopause, stratopause, mesopause, and thermopause, respectively.

The troposphere is characterized by a decrease in temperature with respect to height from a mean surface temperature of about 288 K to a temperature of about 220 K with a typical lapse rate of 6.5 K km^{-1}. The temperature structure in this layer is a consequence of the radiative balance and the convective transport of energy from the surface to the atmosphere. Virtually all water vapor, clouds, and precipitation are confined in this layer. The stratosphere is characterized by an isothermal layer from the tropopause to about 20 km, above which the temperature increases to the stratopause with a temperature of about 270 K. Ozone occurs chiefly in the stratosphere. In addition, thin layers of aerosol are observed to persist for a long period of time within certain altitude ranges of the stratosphere. The state of the stratosphere is primarily determined by the absorption of solar fluxes by ozone and through the emission



definitions of atmo-

re, stratosphere,
the tropopause,

with respect to
erature of about
ture in this layer
t of energy from
precipitation are
al layer from the
the stratopause
stratosphere. In
d of time within
here is primarily
gh the emission

of infrared fluxes by carbon dioxide. Like the troposphere, the temperatures in the mesosphere decrease with height from about 50 to 85 km. Above this height and extending upward to an altitude of several hundred kilometers lies the thermosphere where temperatures range from 500 K to as high as 2000 K, depending on the level of solar activity. The outermost region of the atmosphere above the thermosphere is called the *exosphere*.

Atmospheric nomenclature is customarily derived from its thermal state, as just defined. However, there is practically no physical distinction between the stratosphere and the mesosphere. In recent years, the atmosphere has been defined by the following terms: *lower atmosphere* (troposphere), *middle atmosphere* (stratosphere and mesosphere), and *upper atmosphere* (above 80 km). Moreover, the lowest 1 km or so of the atmosphere differs from the remaining troposphere in that interactions with the surface are strong and significant. This is referred to as the *planetary boundary layer*.

3.1.2 Chemical Composition

The earth's atmosphere is presently composed of two groups of gases, one with nearly permanent concentrations and another with variable concentrations. The atmosphere also contains various kinds of aerosols, clouds and precipitation, which are highly variable in space and time. Table 3.1 lists the chemical formulas and volume ratio for the concentrations of permanent and variable gases. Nitrogen, oxygen, and argon account for more than 99.96% of the atmosphere by volume. The permanent gases have virtually constant volume ratios up to an altitude of about 60 km.

Although listed as a permanent constituent, the carbon dioxide concentration has been increasing by about 0.4% per year as a result of the combustion of fossil fuels,

Table 3.1
The Composition of the Atmosphere[a]

Permanent constituents		Variable constituents	
Constituent	% by volume	Constituent	% by volume
Nitrogen (N_2)	78.084	Water vapor (H_2O)	0–0.04
Oxygen (O_2)	20.948	Ozone (O_3)	0–12 × 10^{-4}
Argon (Ar)	0.934	Sulfur dioxide (SO_2)[b]	0.001 × 10^{-4}
Carbon dioxide (CO_2)	0.036	Nitrogen dioxide (NO_2)[b]	0.001 × 10^{-4}
Neon (Ne)	18.18 × 10^{-4}	Ammonia (NH_3)[b]	0.004 × 10^{-4}
Helium (He)	5.24 × 10^{-4}	Nitric oxide (NO)[b]	0.0005 × 10^{-4}
Krypton (Kr)	1.14 × 10^{-4}	Hydrogen sulfide (H_2S)[b]	0.00005 × 10^{-4}
Xenon (Xe)	0.089 × 10^{-4}	Nitric acid vapor (HNO_3)	Trace
Hydrogen (H_2)	0.5 × 10^{-4}	Chlorofluorocarbons	Trace
Methane (CH_4)	1.7 × 10^{-4}	($CFCl_3$, CF_2Cl_2	
Nitrous oxide (N_2O)[b]	0.3 × 10^{-4}	CH_3CCl_3, CCl_4, etc.)	
Carbon monoxide (CO)[b]	0.08 × 10^{-4}		

[a] After the U.S. Standard Atmosphere (1976) with modifications.
[b] Concentration near the earth's surface.

absorption and release by the oceans, and photosynthesis. In addition, a number of measurement series indicate that the atmospheric methane concentration, with a present value of ~1.7 parts per million by volume (ppmv), has increased by 1–2% per year and that it may have been increasing for a long period of time. The most likely cause of the increase in the CH_4 concentration is the greater biogenic emissions associated with a rising human population. Rice paddies seem to be another prime source of CH_4. There is no direct evidence of an increase in carbon monoxide concentration. However, deforestation, biomass burning, and modification of CH_4 sources could lead to changes in the atmospheric CO concentration. There is also some evidence of an increase in nitrous oxide. A possible global increase of ~0.2% per year in N_2O has been suggested. This increase is attributed to the combustion of fossil fuels and, in part, to fertilizer denitrification.

The amounts of variable gases listed in Table 3.1 are small, but they are extremely important in the radiation budget of the atmosphere. Water vapor is the major radiative and dynamic element in the earth's atmosphere. The H_2O concentration varies significantly with both space and time. The spatial distribution of tropospheric H_2O is determined by the local hydrological cycle via evaporation, condensation, and precipitation, and by large-scale transport processes. Specific humidity decreases rapidly with pressure, almost following an exponential function. Specific humidity also decreases with latitude. More than 50% of water vapor is concentrated below ~850 mb, while more than 90% is confined to the layers below ~500 mb. The variability of the H_2O concentration shows a bimodal distribution with a maximum in the subtropics of both hemispheres below ~700 mb. The variability is very small in the equatorial region and poleward of ~60°. The stratospheric H_2O concentration is relatively small, with a value of ~3–4 ppmv in the lower stratosphere. It has been suggested that H_2O in the lower stratosphere is controlled by the temperature of the tropical tropopause, and by the formation and dissipation of cirrus anvils due to outflow from cumulonimbus.

The ozone concentration also varies significantly with space and time, but ozone occurs principally at altitudes from ~15 to 30 km, an area referred to as the *ozone layer*. The maximum ozone concentration occurs at ~20–25 km, depending on latitude and season. Atmospheric ozone is continually created and destroyed by photochemical processes associated with solar ultraviolet radiation. The absorption of deadly solar ultraviolet radiation by the ozone layer is essential to life on earth. Many photochemical reactions associated with O_3 involve H_2O, CH_4, and CO. Total ozone varies significantly in terms of latitude and season, with the maximum occurring during the polar night.

Nitrogen oxides ($NO_x = NO, NO_2$) appear to be important in the determination of both tropospheric and stratospheric O_3 concentrations. Atmospheric NO_x are emitted by transportation and combustion processes at the surface and by high-flying aircraft in the upper troposphere and lower stratosphere. In the stratosphere, the major source of NO_x is the dissociation of N_2O by excited oxygen atoms. In the lower atmosphere, the major source of NO_x appears to be the anthropogenic combustion of fossil fuels and biomass burning.

Chlorofluorocarbons
ozone layer. Large amo
used in solvents, refriger
$CFCl_3$ (CFC-11), CF_2Cl
(CFC-114), CF_2ClCF_3 (

Sulfur dioxide in the
and other sulfur-based g
aerosols. Emissions of
tropospheric aerosols as
cloud and precipitation
the gases listed in Table

The atmosphere conti
to ~20 μm. These aero
as by human activity. Na

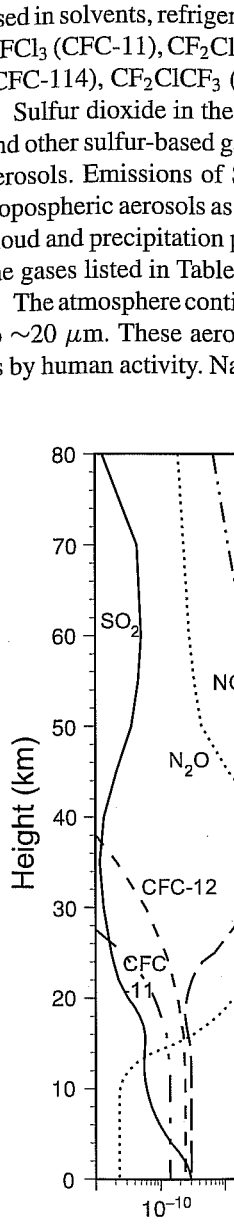


Figure 3.2 Representat
conditions.

Chlorofluorocarbons are also recognized as presenting a potential threat to the ozone layer. Large amounts of these chemicals are produced by industry and are used in solvents, refrigerants, and spray-can propellants. Chlorofluorocarbons include $CFCl_3$ (CFC-11), CF_2Cl_2 (CFC-12), CF_3Cl (CFC-13), CF_3CCl_3 (CFC-113), CF_4CCl_2 (CFC-114), CF_2ClCF_3 (CFC-115), CHF_2Cl (CFC-22), CH_3Cl, and CCl_4.

Sulfur dioxide in the stratosphere is largely produced by volcanic eruptions. SO_2 and other sulfur-based gases are believed to be the primary precursors of stratospheric aerosols. Emissions of SO_2 from the surface may be important in the formation of tropospheric aerosols as well and, hence, related to the production of acid rain through cloud and precipitation processes. Figure 3.2 shows representative vertical profiles of the gases listed in Table 3.1 for mean midlatitude conditions.

The atmosphere continuously contains aerosol particles ranging in size from $\sim 10^{-3}$ to ~ 20 μm. These aerosols are known to be produced by natural processes as well as by human activity. Natural aerosols include volcanic dust, smoke from forest fires,

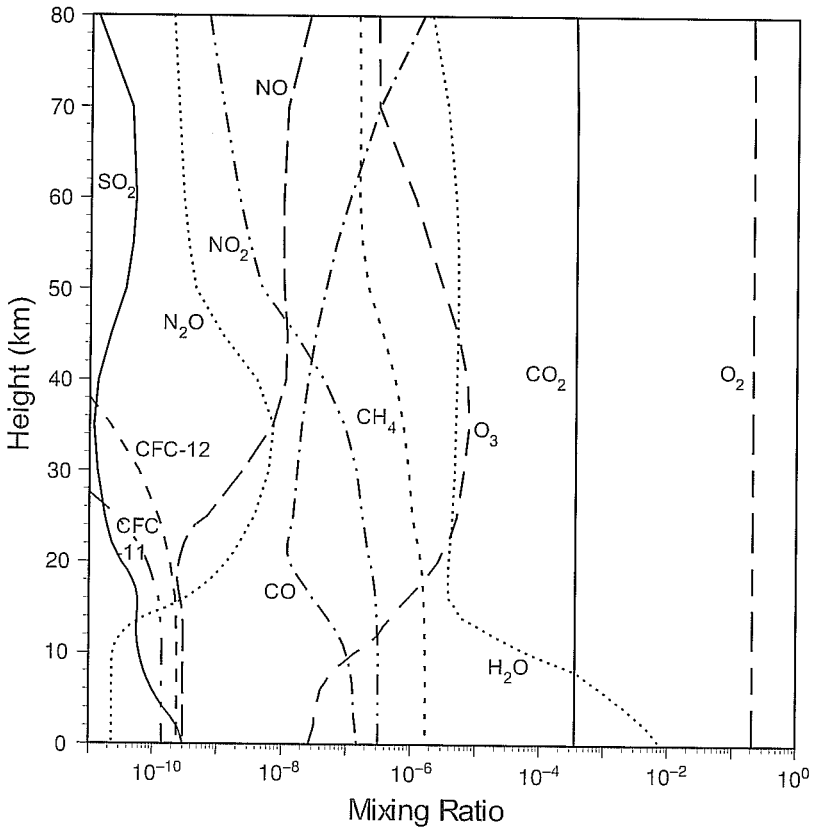


Figure 3.2 Representative vertical profiles of mixing ratios of selected species for midlatitude conditions.

particles from sea spray, windblown dust, and small particles produced by the chemical reactions of natural gases. Primary man-made aerosols include particles directly emitted during combustion processes and particles formed from gases emitted during combustion. The atmospheric aerosol concentration varies with locality; the largest concentrations generally occur in urban and desert areas. In normal conditions, the background aerosol concentration has a visibility of ~20–50 km. Aerosol concentrations generally decrease rapidly with height in the troposphere. Some aerosols are effective condensation and ice nuclei upon which cloud particles may form. For the hygroscopic type, the size of the aerosol depends on relative humidity. Thin layers of aerosols are observed to persist for a long period of time in some altitudes of the stratosphere, as noted previously.

Clouds are global in nature and regularly cover more than 50% of the sky. There are various types of clouds. Cirrus in the tropics and stratus in the Arctic and near the coastal areas are climatologically persistent. The microphysical composition of clouds in terms of particle size distribution and cloud thickness varies significantly with cloud type. Clouds can also generate precipitation, a discrete event generally associated with midlatitude weather disturbances and tropical cumulus convection.

3.2 Atmospheric Absorption

Understanding the atmospheric absorption produced by the various molecules listed in Table 3.1 requires an understanding of their molecular structure. Consider first the diatomic structure (N_2, O_2, CO). Its two nuclei can only move toward and away from each other during vibration. Thus, diatomic molecules have but one vibrational mode, known as symmetric stretch and denoted by ν_1. Because of symmetrical charge distributions, N_2 and O_2 molecules lack a permanent dipole moment that could acquire oscillating momentum during vibration. As a result, they have little radiative activity in the visible and infrared regions (see also Subsection 1.3.1.2).

For triatomic molecules with a linear symmetrical configuration (CO_2, N_2O), there are three vibrational modes: ν_1 for symmetric stretch, ν_2 for bending motion, and ν_3 for antisymmetric stretch, as shown in Fig. 3.3. As a result of their linear symmetry, the CO_2 molecules have no permanent electric dipole moment. Because of vibrational symmetry, the symmetric stretch mode, ν_1, is radiatively inactive at its fundamental, although it has been identified in the Raman spectrum near 7.5 μm. The bending mode, ν_2, consists of ν_{2a} and ν_{2b} vibrations at the same frequency, referred to as degenerate, as noted in Subsection 1.3.1.2. The triatomic structure for H_2O and O_3 molecules forms an isosceles triangle that is obtuse, referred to as the asymmetric top (bent triatomic) configuration. This molecular shape has three fundamental vibration modes, as shown in Fig. 3.3 (Herzberg, 1945). Not shown in the figure is the CH_4 molecule, which has a spherical top configuration, no permanent electric dipole moment, and four fundamental vibration modes.

The rotation of a rigid body is defined by its center of gravity with respect to three orthogonal axes. The axes of rotational freedom of linear and asymmetric top

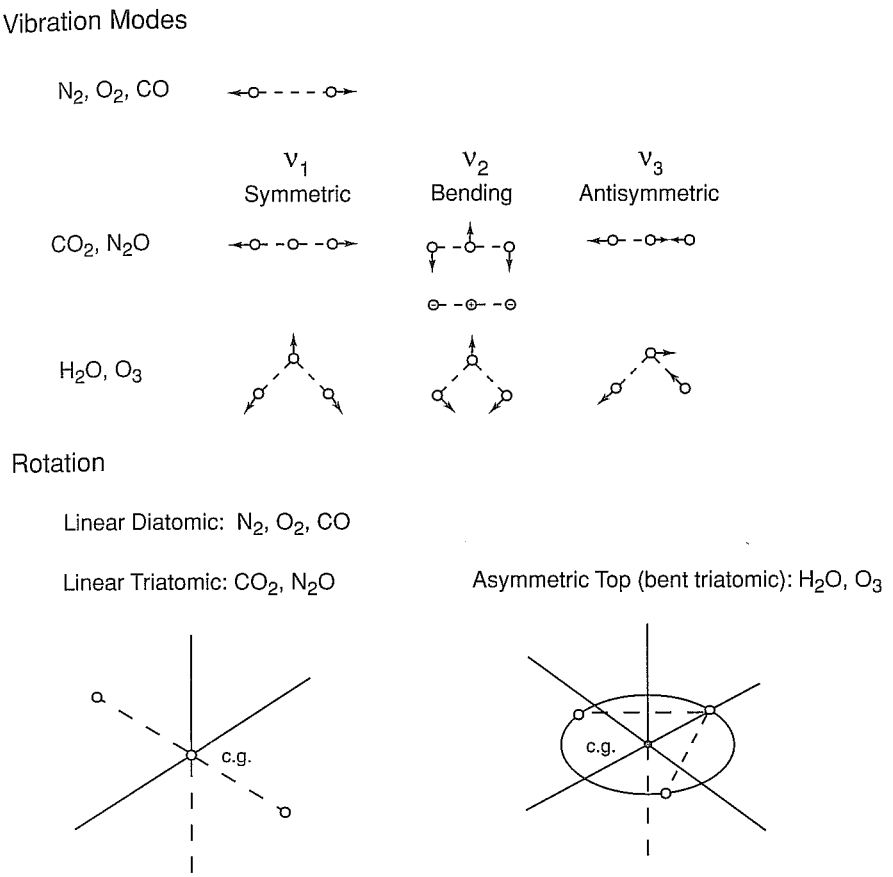


Figure 3.3 Vibrational modes of diatomic and triatomic atmospheric molecules and the axes of rotational freedom for linear and asymmetric top molecules.

molecules are displayed in Fig. 3.3. A diatomic and a linear triatomic molecule have two equal moments of inertia and two degrees of rotational freedom. Asymmetric top molecules have three unequal moments and three degrees of rotational freedom. Molecular structures of greater complexity have additional degrees of freedom.

As discussed in Section 2.3.1, electrons can be removed from an orbit if a characteristic amount of energy is available. This amount with respect to the ground level is called the *ionization potential*. For atoms and molecules with more than one electron, the ionization potential usually refers to the most loosely bound electron, the one that requires the least energy for removal. Likewise, a certain amount of energy can dissociate or separate the atoms and is referred to as the *dissociation potential*.

Electronic energy is closely related to vibrational energy because both derive from the elastic valence bonds that bind the atoms into a molecular entity. The sign and magnitude of the force between two atoms in a molecule depend on two factors: the

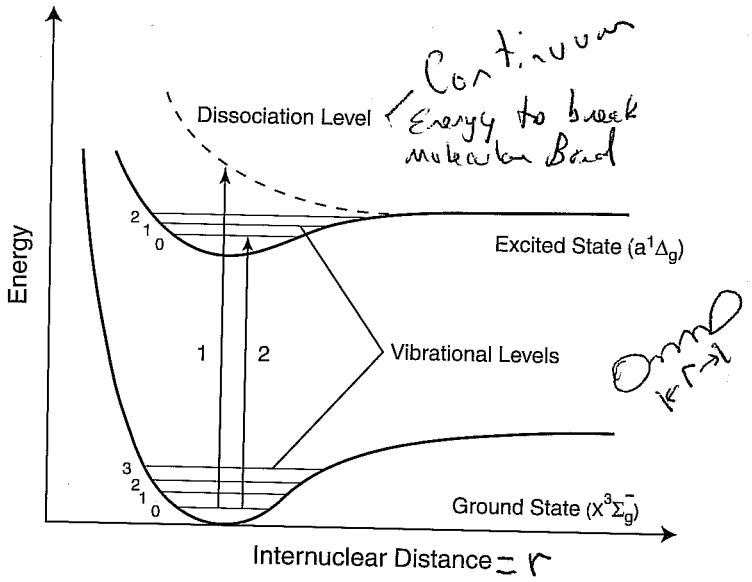


Figure 3.4 Illustrative potential energy curves for two electronic states of a diatomic molecule. The horizontal lines in the potential well represent vibrational energy levels.

distance between the two nuclei and their electronic configuration. It is customary to represent the dependence of this force on the internuclear distance as a *potential curve,* with potential energy as the ordinate and distance as the abscissa, as shown in Fig. 3.4 in which two possible potential curves are displayed schematically. These two curves are representative of the ground and excited states of a diatomic molecule. At large distances the two atoms of a molecule exert no significant attractive or repulsive force on one another. As the atoms approach one another in the ground state of the molecule, they begin to exert attractive force, as shown by the decrease in potential energy. The ground electronic state of the molecule exhibits maximum stability at the minimum in the potential curve that represents a stable bound configuration of the molecule. As the distance becomes smaller, the atoms begin to repel one another so that potential energy increases.

When a high-energy photon is absorbed, the electron configuration changes to one that has potential energy even when the nuclei are at equilibrium distance. This is a transition to an excited state whose potential energy is represented by the upper curve. The energy gain due to the transition is given by the difference between the two minima. Also shown are the vibrational levels denoted by the horizontal lines in the potential well.

As noted earlier and as shown in Fig. 3.4, the absorption of a photon of appropriate energy can lead to the excitation of a molecule from one electronic state to another. If the excited state is unstable, two possibilities may take place: the molecule may decompose into its constituent atoms and photoionization may also occur. At

atmospheric temperatures, most molecules are in the ground vibrational state. Possible transitions are denoted by the vertical lines labeled 1 and 2 in Fig. 3.4. Transition 1 leads from the ground state to a state that is not quantized, denoted by the dissociation level, which has more energy than the quantized level. Thus, a continuum of wavelengths near the energy associated with transition 1 are possible. The spectrum should then consist of a smooth continuum at short wavelengths above the dissociation limit. Transition 2, on the other hand, requires absorption into a specific vibrational level of the upper electronic state. It is quantized at the energy level defined by transition 2. The spectra corresponding to these transitions are then shown at discrete wavelengths. Many absorption processes associated with diatomic molecules are largely a combination of electronic and vibrational transitions that occur in the ultraviolet (UV), visible, and near infrared (IR) regions. As noted in Subsection 1.3.1.2, it is customary to refer to transitions as forbidden and allowed. However, even highly forbidden transitions occur under certain conditions.

As discussed in Section 1.1.4, the strength of absorption is customarily represented by cross section σ_a in units of cm^2. Also frequently used is the absorption coefficient k in units of $(cm\ atm)^{-1}$. The two are related through the Loschmidt's number, $N_0 (= 2.687 \times 10^{19}$ particles cm^{-3} at standard temperature of 273 K and standard pressure of 1013 mb), and the fact that $k = \sigma_a \cdot N_0$.

3.2.1 Absorption in the Ultraviolet

3.2.1.1 Molecular Nitrogen

The UV absorption spectrum of the most abundant gas in the atmosphere consists of a band system from about 1450 to 1120 Å, referred to as the *Lyman–Birge–Hopfield bands,* associated with the forbidden ground-state transition. Dissociation and ionization continua have been observed below 800 Å. From about 1000 to 850 Å, the nitrogen spectrum consists of various sharp bands whose absorption coefficients are highly variable. Absorption of N_2 in the solar spectrum is generally considered to be insignificant. The photodissociation of N_2 in the atmosphere plays a minor role in atmospheric chemistry below 100 km.

3.2.1.2 Molecular Oxygen

The UV absorption spectrum of O_2 begins with the weak *Herzberg band* between 2600 and 2000 Å, due primarily to the forbidden ground-state transition and dissociation continuum, which lead to the formation of two oxygen atoms in the ground ^3P state. Absorption by this band system is weak and of little importance in the absorption of solar radiation because of overlap with the much stronger O_3 bands in this spectral region. It is, however, considered to be of significance in the formation of ozone. Adjacent to the Herzberg continuum are the *Schumann–Runge bands* produced by ground-state transitions that occupy the spectral region from 2000 to 1750 Å. At 1750 Å, the bands converge to a stronger dissociation continuum in which one of the oxygen atoms is formed in the excited ^1D state. Referred to as the *Schumann–Runge continuum,* this extends to about 1300 Å and represents the most important

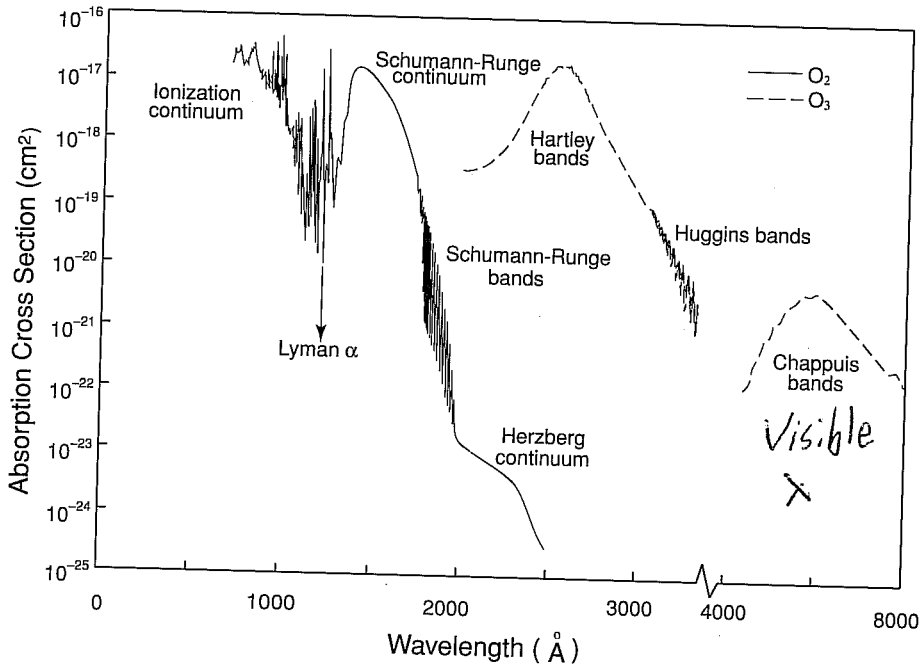
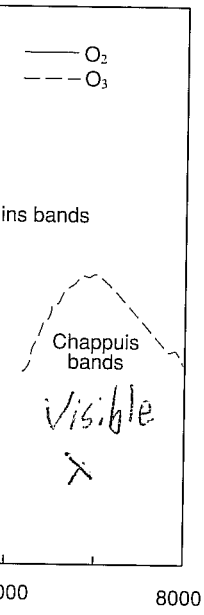


Figure 3.5 Absorption cross section of ozone and molecular oxygen in the ultraviolet spectral region. Data taken from Brasseur and Solomon (1986), Vigroux (1953), and Griggs (1968).

absorption spectrum of O_2. At shorter wavelengths some diffuse bands occur, which have not yet been identified. These are interspersed with a series of windows allowing some wavelengths to penetrate deeply into the atmosphere. Of particular interest is the Lyman α line located at 1216 Å (see the solar spectrum in Fig. 2.9), which happens to lie in one of the windows. Between 850 and 1100 Å are a series of distinct Rydberg bands, associated with transitions between excited states, known as the *Hopfield bands*. Below about 1026 Å, O_2 absorption is in the form of an ionization continuum. The spectral distribution of the absorption cross section of O_2 is shown in Fig. 3.5.

Because of the absorption of solar UV radiation, some of the oxygen and nitrogen molecules in the upper atmosphere undergo photochemical dissociation into atomic oxygen and nitrogen. Atomic nitrogen exhibits an absorption spectrum from about 10 to about 1000 Å. Although atomic nitrogen probably is not abundant enough to be a significant absorber in the upper atmosphere, it may play an important role in the absorption of UV radiation in the thermosphere. Atomic oxygen also shows an absorption continuum in the region of 10 to 1000 Å. Because of the absorption of solar UV radiation, a portion of molecular and atomic oxygen and nitrogen becomes ionized. The ionized layers in the upper atmosphere are formed mainly as a result of these processes.



violet spectral region.

nds occur, which
f windows allow-
particular interest
Fig. 2.9), which
re a series of dis-
ites, known as the
n of an ionization
of O_2 is shown in

gen and nitrogen
ation into atomic
trum from about
ndant enough to
important role in
en also shows an
he absorption of
itrogen becomes
nly as a result of

3.2.1.3 Ozone

The absorption of ozone in the solar spectral region is due to electronic transitions. The strongest ozone bands are the *Hartley bands,* which cover the region from 2000 to 3000 Å and are centered at 2553 Å. The absorption of solar flux in these ozone bands takes place primarily in the upper stratosphere and in the mesosphere. The weak bands between 3000 and 3600 Å have more structure and are called the *Huggins bands.* Ozone also shows weak absorption bands in the visible and near-IR regions from about 4400 to 11,800 Å. These bands are referred to as *Chappuis bands.* The absorption coefficients in these bands are slightly dependent upon temperature. They have been measured by a number of earlier researchers (Inn and Tanaka, 1953; Vigroux, 1953) and more recently by Molina and Molina (1986) and Anderson and Mauersberger (1992). Absorption cross sections of the preceding three O_3 bands are also displayed in Fig. 3.5.

3.2.1.4 Other Minor Gases

The NO_2 molecule absorbs solar flux in the UV as well as in the visible section of the wavelength range between 0.2 and 0.7 μm. Accurate absorption cross sections are required for the retrieval of aerosols and ozone using ground-based sunphotometer measurements (see Section 7.2.1). The UV absorption cross sections of NO, N_2O, H_2O, CO_2, and other trace gases listed in Table 3.1 have been measured by numerous researchers and they are important in the discussion of atmospheric chemistry and ionization in the middle and upper atmospheres. However, either because they occur in too small quantities or because they are dissociated at high levels, these gases absorb relatively little energy in the UV and are overshadowed by the absorption of O_2 and O_3. The most important absorption bands in the UV particularly associated with photochemical processes are listed in Table 3.2.

3.2.1.5 Absorption of Solar Radiation

In reference to Section 1.4.2, consider an atmosphere that is plane-parallel and nonscattering. We may define a normal absorption optical depth for a monochromatic

Table 3.2

Important Absorption Spectral Regions Associated with Photochemistry in the Atmosphere

Wavelength range (Å)	Absorber	Principal location
1000–1750	O_2 Schumann–Runge continuum	Thermosphere
	O_2 1216 Lyman α line	Mesosphere
1750–2000	O_2 Schumann–Runge bands	Mesosphere
2000–2420	O_2 Herzberg continuum; O_3 Hartley band	Stratosphere
2420–3100	O_3 Hartley band; $O(^1D)$ formation	Stratosphere
3100–4000	O_3 Huggins bands; $O(^3P)$ formation	Stratosphere/ troposphere
4000–8500	O_3 Chappuis bands	Troposphere

wavelength and for a given absorber as follows:

$$\tau(z) = \int_z^\infty n(z)\sigma_a dz, \tag{3.2.1}$$

where σ_a is the absorption cross section (cm^2), and n is the number density (cm^{-3}) for a specific absorber. The attenuation of solar flux that enters the atmosphere at a solar zenith angle of θ_0 is given by

$$F_{act}(z) = F_\odot e^{-\tau(z)/\mu_0}, \tag{3.2.2a}$$

where $\mu_0 = \cos\theta_0$ and $F_\odot$ is the solar irradiance at the top of the atmosphere. The flux here is in reference to the direct solar beam without accounting for the cosine dependence [see Eq. (1.1.8)] and is referred to as *actinic flux* associated with photodissociation, defined by

$$F_{act}(z) = \int_0^{2\pi} \int_{-1}^{1} I(z;\mu,\phi)\, d\mu\, d\phi, \tag{3.2.2b}$$

with units of photons $cm^{-2}\,sec^{-1}$. This definition can be applied to both direct and diffuse beams.

The solar energy absorbed per time and volume along the direction of the solar beam, referred to as the *volume absorption rate* (photons $cm^{-3}\,sec^{-1}$), can be expressed by

$$q(z) = \frac{dF_{act}(z)}{dz/\mu_0} = F_\odot n(z)\sigma_a e^{-\tau(z)/\mu_0}$$
$$= J(z)n(z), \tag{3.2.3a}$$

where the monochromatic *photodissociation coefficient*, $J(z)(sec^{-1})$, is defined by

$$J(z) = \sigma_a F_{act}(z). \tag{3.2.3b}$$

To gain a physical insight into solar attenuation, consider an atmosphere containing a concentration of an absorber that varies exponentially with altitude according to a scale height, H (Exercise 3.1), in the form

$$n(z) = n_0 e^{-z/H}, \tag{3.2.4}$$

where n_0 is the value of n at some arbitrary level at which z is taken to be zero. On substituting Eq. (3.2.4) into Eqs. (3.2.3a) and (3.2.1), we obtain

$$q(z) = F_\odot \sigma_a n_0 \exp\left(-\frac{z}{H} - \frac{1}{\mu_0}\sigma_a n_0 H e^{-z/H}\right). \tag{3.2.5a}$$

We may define a level, z_0, at which the normal absorption optical depth is equal to 1 such that

$$\tau(z_0) = 1, \text{ i.e., } \sigma_a H n_0 = e^{z_0/H}. \tag{3.2.5b}$$

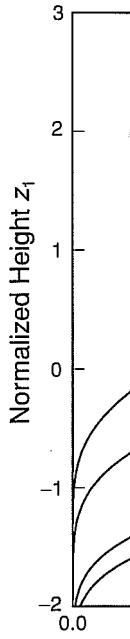


Figure 3.6 Norma
solar zenith angles (see

By carrying out so

$r(z$

where $z_1 = (z - z_0$

The function $r($
This function displ
layer (Chapman, 1
the extremely smal
at the top of the at
declines more rapi
The rate of energy
is low, the maximu
effective optical de
the level at which

The preceding
centration. In pract

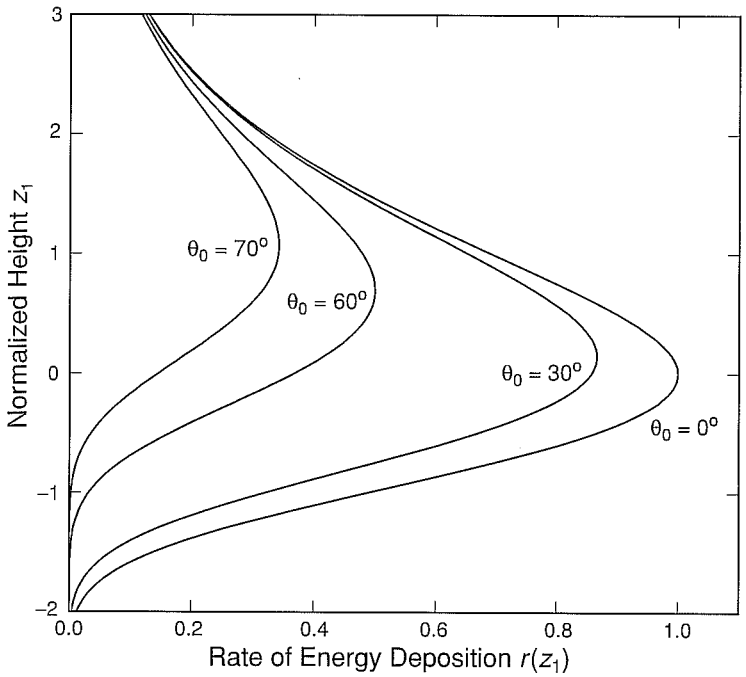


Figure 3.6 Normalized rate of energy deposition as a function of normalized height for a number of solar zenith angles (see text for the definition of r and z_1).

By carrying out some straightforward analyses, we find

$$r(z_1) = \frac{q(z)}{q[z_0(\mu_0 = 1)]} = \exp\left(1 - z_1 - \frac{1}{\mu_0} e^{-z_1}\right), \tag{3.2.6}$$

where $z_1 = (z - z_0)/H$ (Exercise 3.2).

The function $r(z_1)$ is illustrated in Fig. 3.6 for a number of solar zenith angles θ_0. This function displays a well-defined layered structure, referred to as the *Chapman layer* (Chapman, 1931). The small value of $q(z)$ at the top level is associated with the extremely small value of $n(z)$ due to exponential falloff and a constant $J(z)$ value at the top of the atmosphere. Deep down in the atmosphere $n(z)$ increases but $J(z)$ declines more rapidly, again leading to a small value of $q(z)$, as shown in Eq. (3.2.3a). The rate of energy deposition is also related to the solar zenith angle. When the sun is low, the maximum deposition level moves higher because of the increase in the effective optical depth. As demonstrated in Eq. (3.2.5b), the optical depth of unity is the level at which maximum energy deposition occurs.

The preceding discussion of the Chapman layer is limited to one particular concentration. In practice, however, the combined effects of several absorbing gases must

be accounted for. Thus, the total absorption optical depth for a given wavelength is

$$\tau(z, \lambda) = \sum_i \int_z^\infty n_i(z)\sigma_a(n_i, \lambda)\, dz, \tag{3.2.7a}$$

and the photodissociation coefficient defined in Eq. (3.2.3b) for a spectral interval can be written as

$$J_i(z) = \int_{\Delta\lambda} \sigma_a(n_i, \lambda)F_\odot(\lambda)e^{-\tau(z,\lambda)/\mu_0}\, d\lambda. \tag{3.2.7b}$$

When the solar zenith angle is larger than about 75°, as in the case of sunrise or sunset, the effect of the earth's curvature must be accounted for in the calculation of the effective optical depth. In this case a more complicated function should replace $1/\mu_0$ (Exercise 3.3).

To illustrate the relative absorption effects of oxygen and nitrogen molecules, Fig. 3.7 shows the reduction of solar flux when it penetrates the atmosphere. The curve represents the altitude where the optical depth is unity. Alternately, we can express $F_{\text{act}} = F_\odot e^{-1}$, referred to as the *e-folding transfer*. The EUV (extreme ultraviolet) fluxes are absorbed at high altitudes, resulting in the dissociation and ionization of the major constituents in the thermosphere and leading to the formation of the layers of the ionosphere. At longer wavelengths, from 1750 to about 2400 Å, the solar flux penetrates deeper into the atmosphere and is chiefly absorbed by O_2 in the Schumann–Runge band and Herzberg continuum (Table 3.2), leading to the production of O and

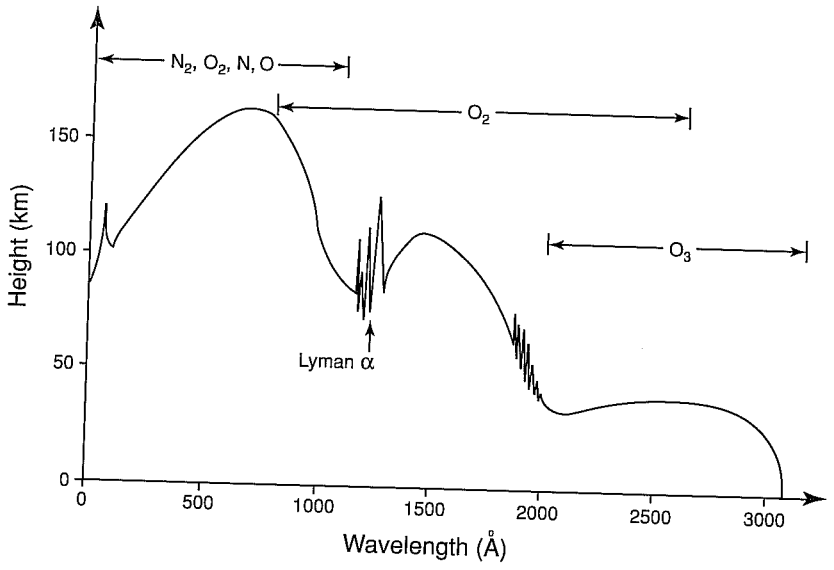


Figure 3.7 Depth of penetration of solar UV radiation in the earth's atmosphere for $\mu_0 = 1$ and an averaged ozone profile. The line indicates the height where the optical depth is equal to 1 (data taken from Herzberg, 1965).

O_3. Maximum ozone absorption occurs at about 50 km, associated with absorption in the Hartley band in the wavelength range 2400–3100 Å. Above 3100 Å, the atmosphere is relatively transparent except for Rayleigh scattering and scattering produced by aerosols and clouds.

3.2.2 Photochemical Processes and the Formation of Ozone Layers

Because of the absorption spectrum of various molecules and atoms in the solar UV region, a great variety of photochemical processes take place in the upper and middle atmospheres. Those involving various forms of oxygen are critical in determining the amount of ozone in the stratosphere. The classical photochemistry of the middle atmosphere concerning the ozone problem was first postulated by Chapman (1930), who described the following five basic reactions.

Ozone is basically formed by the three-body collision

$$O + O_2 + M \xrightarrow{K_{12}} O_3 + M, \tag{3.2.8a}$$

where M is any third atom or molecule, and K_{12} is the rate coefficient involving O and O_2. Atomic oxygen is produced when the oxygen molecule is dissociated by a quantum of solar energy

$$O_2 + h\tilde{\nu}(\lambda < 2420\,\text{Å}) \xrightarrow{J_2} O + O, \tag{3.2.8b}$$

where J_2 is the dissociating quanta per molecule absorbed by O_2. Ozone is destroyed by photodissociation,

$$O_3 + h\tilde{\nu}(\lambda < 11,000\,\text{Å}) \xrightarrow{J_3} O + O_2, \tag{3.2.8c}$$

as well as by collision with oxygen atoms,

$$O_3 + O \xrightarrow{K_{13}} 2O_2, \tag{3.2.8d}$$

where J_3 is the dissociating quanta per molecule absorbed by O_3, and K_{13} denotes the rate coefficient involving O_3 and O. At the same time, oxygen atoms generated by reactions (3.2.8b) and (3.2.8c) may undergo a three-body collision,

$$O + O + M \xrightarrow{K_{11}} O_2 + M, \tag{3.2.8e}$$

with K_{11} denoting the rate coefficient involving O and O. Normally, the reaction denoted in (3.2.8e) may be neglected below 50 to 60 km.

The preceding five reactions take place simultaneously. If an equilibrium state is reached, the number of ozone molecules formed exactly equals the number destroyed in unit volume and time. To evaluate the equilibrium amount of ozone, let [O], [O_2], [O_3], and [M] be the number densities, respectively, for O, O_2, and O_3, and air molecules. Based on the observed concentrations of oxygen species, almost all oxygen below 70 km is in the form of O_2, which is fairly constant with a value of about 21%. Thus, below 70 km, [O_2] may be treated as an independent variable. It follows that the

photochemical processes given by Eqs. (3.2.8a)–(3.2.8c) may be expressed in terms of the rate of change of the number density of O and O_3 in the forms:

$$\frac{\partial[O]}{\partial t} = -K_{12}[O][O_2][M] + 2J_2[O_2] - K_{13}[O][O_3] + J_3[O_3] - 2K_{11}[O][O][M], \tag{3.2.9a}$$

$$\frac{\partial[O_3]}{\partial t} = K_{12}[O][O_2][M] - K_{13}[O][O_3] - J_3[O_3], \tag{3.2.9b}$$

with the photodissociation coefficients given by

$$J_2(z) = \int_0^{\lambda_1} \sigma_a(O_2, \lambda) F_\odot(\lambda) e^{-\tau(z,\lambda)/\mu_0} d\lambda, \tag{3.2.10a}$$

$$J_3(z) = \int_0^{\lambda_2} \sigma_a(O_3, \lambda) F_\odot(\lambda) e^{-\tau(z,\lambda)/\mu_0} d\lambda, \tag{3.2.10b}$$

where $\lambda_1 = 0.2420\,\mu$m and $\lambda_2 = 1.1\,\mu$m.

On adding Eqs. (3.2.9a) and (3.2.9b), we find

$$\frac{\partial}{\partial t}([O] + [O_3]) = 2J_2[O_2] - 2K_{13}[O][O_3]. \tag{3.2.11a}$$

Below 75 km, $K_{13}[O][O_3] \ll K_{12}[O][O_2][M]$, so that Eq. (3.2.9b) may be simplified to give

$$\frac{\partial[O_3]}{\partial t} \cong -J_3[O_3] + K_{12}[O][O_2][M]. \tag{3.2.11b}$$

Under the assumption of photochemical equilibrium, $\partial[O]/\partial t = \partial[O_3]/\partial t = 0$. Thus, from Eqs. (3.2.11a) and (3.2.11b), the solution for ozone is given by

$$[O_3]_{eqb} \cong \left(\frac{J_2 K_{12}}{J_3 K_{13}}[M]\right)^{1/2} [O_2]. \tag{3.2.12}$$

At low levels, $[O_3]$ is small since J_2 is small. At high levels, $[O_3]$ is also small because $[O_2]$ and $[M]$ are both small. $[O_3]$ has a maximum in the stratosphere. Exercise 3.4 requires the calculation of $[O_3]$ using Eq. (3.2.12).

Figure 3.8 depicts the equilibrium ozone concentration determined from the classical theory. In the same diagram, the observational range of ozone number densities (shaded area) is also shown. It is evident that the classical theory overestimates the ozone number densities at almost all heights. The total ozone in an atmospheric column based on theoretical calculations exceeds the observed values by as much as a factor of three or four. Obviously, additional loss mechanisms are required to explain the observed data.

In the early 1970s, two independent theoretical analyses suggested that ozone can be destroyed by minute concentrations of chemical species associated with human activities. These species are chlorine atoms in the chlorine molecule and nitrogen atoms in the nitrogen molecule. The importance of chlorine was established by Molina

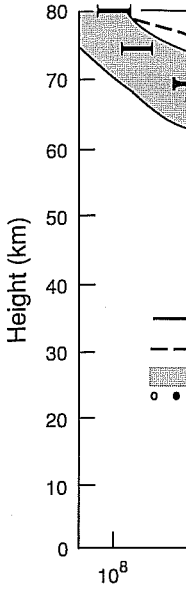


Figure 3.8 Observ
ozone number densities

and Rowland (1974
refrigerants and spr
oxides was first rep
supersonic transpor
The catalytic de

where X may be ni
hydrogen (H). The

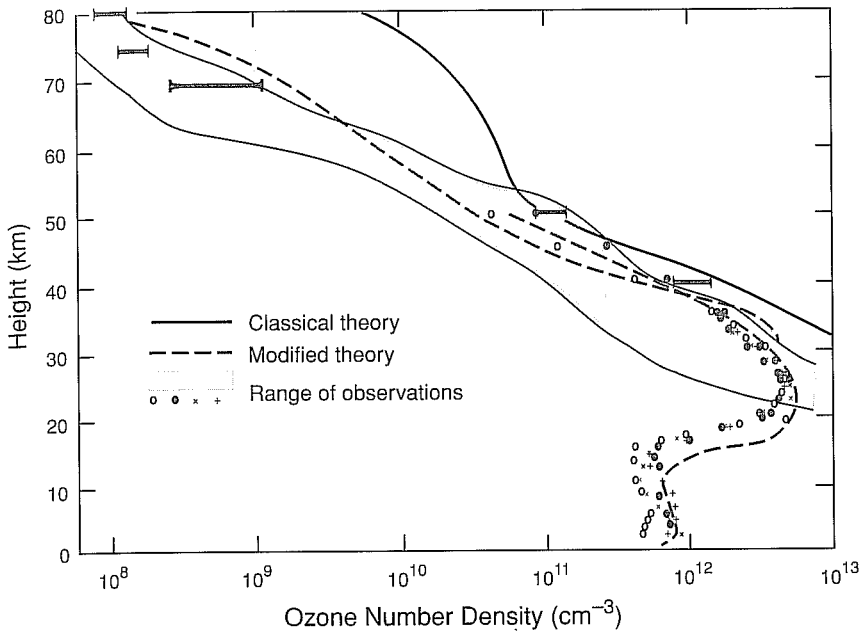


Figure 3.8 Observational range in ozone number densities and theoretical calculations for equilibrium ozone number densities (data from Leovy, 1969, 20–80 km; Logan *et al.*, 1978, 0–50 km).

and Rowland (1974) from their study of the chlorofluorocarbons (CFCs) produced by refrigerants and spray-can propellants. Ozone reduction from the effect of nitrogen oxides was first reported by Crutzen (1970) and that from the exhaust of a fleet of supersonic transports was documented by Johnston (1971).

The catalytic destruction reactions of ozone have been found to be

$$O_3 + h\tilde{\nu} \longrightarrow O + O_2, \tag{3.2.13a}$$

$$O + XO \longrightarrow X + O_2, \tag{3.2.13b}$$

$$X + O_3 \longrightarrow XO + O_2, \tag{3.2.13c}$$

$$\text{net } 2O_3 + h\tilde{\nu} \longrightarrow 3O_2,$$

where X may be nitric oxide (NO), chlorine (Cl), hydroxyl radical (OH), or atomic hydrogen (H). The possible sources of NO and OH are the following reactions:

$$O_3 + h\tilde{\nu}(\lambda < 3100\,\text{Å}) \longrightarrow O(^1D) + O_2, \tag{3.2.14a}$$

$$O(^1D) + M \longrightarrow O + M, \tag{3.2.14b}$$

$$O(^1D) + N_2O \longrightarrow 2NO, \tag{3.2.14c}$$

$$O(^1D) + H_2O \longrightarrow 2OH, \tag{3.2.14d}$$

where $O(^1D)$ denotes the excited atomic oxygen in the ^1D state, which is essential to these reactions. It is clear that the high concentration of ozone from Chapman's theoretical prediction is due to the neglect of these additional loss mechanisms and possibly others. If we introduce reactions involving (3.2.13a)–(3.2.14d), the calculated equilibrium ozone concentration labeled as the modified theory appears to closely match the observed values (Logan *et al.*, 1978).

Ozone is a natural trace ingredient of the atmosphere that occurs at an average concentration of about 3 parts per million by volume. Its concentration varies with season and latitude and is modulated by radiative and dynamic processes in the middle atmosphere. High intensities of UV radiation shorter than 3200 Å, which are harmful to nearly all forms of life, are largely (~99%) screened out by ozone. It has been postulated that surface life on earth did not evolve until after the ozone layer was formed. The effect of small increases in the intensity of ultraviolet radiation due to the reduction of ozone by human activities has been a subject of extensive scientific research.

More recently, there has been continuous concern with the question of whether the ozone layer in the stratosphere is actually thinning globally because of the increased use of CFCs. A decrease in Antarctic ozone during the austral spring, just after the end of the Antarctic polar night, has been discovered. Data from the British Antarctic Survey collected at the Halley Bay station indicate a 40% total ozone decrease during the month of October between 1977 and 1984 (Farman *et al.*, 1985). This decrease has been confirmed by data derived from the Total Ozone Mapping Spectrometer and the Solar Backscatter Ultraviolet Spectrometer on board Nimbus 7. The largest percentage decrease corresponds roughly to the polar vortex, covering approximately the entire area of the Antarctic continent. It has been suggested that the so-called *ozone hole* is the precursor of a major decrease in the ozone layer worldwide caused by the addition of CFCs to the atmosphere. It has also been noted that the ozone hole appears to have a dynamic origin and is related to the rather special meteorological conditions prevalent over Antarctica each October. The cold polar vortex traps the air parcels for weeks, during which time *polar stratospheric clouds* are formed. These clouds may facilitate certain chemical reactions that favor the destruction of ozone (Toon and Turco, 1991).

3.2.3 Absorption in the Visible and Near Infrared

3.2.3.1 Molecular Oxygen and Ozone

Molecular oxygen has absorption bands in the visible and near-IR. The ground electronic state of O_2 is designated by X and two excited states by a and b. The a←X and b←X transitions coupled with vibrational–rotational transitions produce weak absorption lines in the near-IR and visible, respectively. These are referred to as *infrared bands* and *red bands*. The most important red bands are the A band centered at 0.762 μm (0←0), the B band at 0.688 μm (1←0), and the γ band at 0.628 μm (2←0). See Fig. 3.4 for the meaning of b←X (ground state electronic transition)

and 0←0 (vibrational transition). Because the positions of these three bands are near the peak of the solar spectrum, absorption of the solar flux due to O_2 in the visible is important in the middle and upper atmospheres and could affect the solar flux available at the surface and in the troposphere. Line-by-line information for O_2 is available and a precise computation of the band absorption can be performed. The oxygen A band has been utilized for the remote sensing of the level of cloud pressure from space. Absorption of O_3 in the visible and near-IR, the Chappuis band, was discussed in Subsection 3.2.1.3.

3.2.3.2 Water Vapor

The water vapor molecule has an asymmetric top (bent triatomic) configuration with the oxygen atom in the middle and a bond angle of 104.45°, as shown in Fig. 3.3. The distance between the oxygen and hydrogen atoms is 0.958 Å. It has three rotational constants. The hydrogen atom has two isotopic forms, ^1H and ^2D, with relative abundances of 99.9851 and 0.0149%, respectively. The isotopic forms of water vapor that are important in infrared radiative transfer are $HH^{16}O$, $HH^{18}O$, $HD^{16}O$, and $HD^{18}O$. Each of these molecules has a different vapor pressure, and their abundances are a function of the hydrological cycle. The H_2O molecule has three fundamental vibration modes (Fig. 3.3). The bending vibration, ν_2, has the lowest wavenumber; both ν_1 and ν_3 have wavenumbers about twice the wavenumber for ν_2. The ν_2 fundamental band of H_2O is centered at 6.25 μm and is important in thermal infrared radiation transfer and remote sensing applications (Section 4.2.1). The ν_1 and ν_3 fundamentals of H_2O produce bands centered at 3657.05 cm^{-1} (2.74 μm, 100←000) and 3755.93 cm^{-1} (2.66 μm, 001←000), respectively. These two bands are close to one another and combine to form a strong band in the solar spectrum, referred to as the 2.7 μm band. The three integers represent the three vibrational quantum numbers $\upsilon_1\upsilon_2\upsilon_3$. The ground state is designated by 000. As described in Subsection 1.3.1.2, the transition from the ground state to the first excited state 001 or 100 is referred to as the fundamental. The $2\nu_2$ band is centered at 3161.60 cm^{-1} (3.2 μm band, 020←000) in the tail of the solar spectrum. In addition, the solar spectrum contains a large number of overtone and combination bands, which arise from ground state transitions. In the near-IR region, these bands absorb a significant amount of solar flux in the lower atmosphere. They are centered at 0.94, 1.1, 1.38, and 1.87 μm, and are commonly identified in groups by the Greek letters (ρ, σ, τ), ϕ, Ψ, and ω, respectively. Although the overtone and combination bands centered at 0.72 and 0.82 μm are relatively weak, their contribution to the solar heating of the atmosphere appears to be not insignificant. There are also a number of H_2O lines in the visible region associated with overtone and combination transitions. However, their contributions to absorption and the heating rate appear to be quite small.

3.2.3.3 Carbon Dioxide

The CO_2 molecule has a linear symmetrical configuration, with the carbon atom in the middle and an oxygen atom on each side, as shown in Fig. 3.3. The length of the C–O bond in the fundamental vibration is 1.1632 Å and it has one rotational constant.

The natural isotopes of the carbon atom are ^{12}C and ^{13}C with relative abundances of 98.892 and 1.108%, respectively. For the oxygen atom, the isotopes are ^{16}O (99.758%), ^{17}O (0.0373%), and ^{18}O (0.2039%). Thus, several CO_2 isotopes are present in the atmosphere. The most significant of these for the radiation problem are $^{12}C^{16}O^{16}O$, $^{13}C^{16}O^{16}O$, and $^{12}C^{16}O^{18}O$.

As a result of its symmetrical arrangement, the CO_2 molecule has no permanent dipole moment and no permitted rotation band. The symmetrical stretch mode, ν_1, is radiatively inactive at its fundamental. The bending mode, ν_2, produces the most important band in the thermal infrared: the 15 μm CO_2 band. This subject will be discussed further in Section 4.2.1. Because of asymmetric stretching vibration, the ν_3 fundamental transition has a wavenumber centered at 2349.16 cm^{-1} (the 4.3 μm band) for $^{12}C^{16}O^{16}O$. In addition, the combination band $\nu_1 + \nu_3 - 2\nu_2$ is centered at 2429.37 cm^{-1} and occurs when the ν_3 transition originates at the vibrational level $\nu = 2$ or higher. Because of parallel vibration, the Q branch corresponding to $\Delta J = 0$ does not appear. There are also a number of bands for the isotopes. The 4.3 μm CO_2 band is in the tail of the solar spectrum and its effect on solar absorption is insignificant. This band has been used in connection with the 15 μm band for the nighttime detection of atmospheric temperatures (see Section 7.4.3). Carbon dioxide exhibits a number of rather weak overtone and combination bands in the solar region: 2.0, 1.6, and 1.4 μm. The stronger 2.7 μm band of CO_2 overlaps with the 2.7 μm band of H_2O and contributes to the absorption of solar flux in the lower stratosphere.

3.2.3.4 Other Minor Gases

The N_2O molecule is linear and asymmetric (NNO). The three fundamentals are ν_1 (1284.91 cm^{-1}), ν_2 (558.77 cm^{-1}), and ν_3 (2223.76 cm^{-1}). The ν_3 band (4.5 μm) is in the solar region, but is insignificant in solar absorption. There are also four overtone and combination bands: 4.06, 3.90, 2.97, and 2.87 μm. The CO molecule has fundamental (2.34 μm) and first-overtone (4.67 μm) bands in the solar spectrum. The CH_4 molecule is a spherical top with four independent (ν_1, ν_2, ν_3, ν_4) fundamentals. The ν_3 fundamental occurs at 3018.92 cm^{-1} (3.31 μm). Other overtone and combination bands are located at 3.83, 3.53, 3.26, 2.37, 2.30, 2.20, and 1.66 μm. The NO_2 molecule absorbs solar flux in the wavelength range 0.2–0.7 μm, as pointed out in Subsection 3.2.1.4. Absorption of solar radiation by the transitional molecules, $O_2 \cdot O_2$ and $O_2 \cdot N_2$, has also been reported (Solomon *et al.*, 1998). A summary of the preceding absorption bands is given in Table 3.3.

3.2.3.5 Transfer of Direct Solar Flux in the Atmosphere

The monochromatic downward solar flux density, following the discussion in Section 1.4.2, may be written as

$$F_\lambda(z) = \mu_0 F_\odot(\lambda) \exp\left(-\frac{k_\lambda u(z)}{\mu_0}\right), \tag{3.2.15}$$

Solar

Species	Band (μm
	3.2, 2.7
	1.87
	1.38
	1.1
H_2O	0.94
	0.82
	0.72
	Visible
	4.3
	2.7
CO_2	2.0
	1.6
	1.4
	4.74
	3.3
O_3	Visible
	UV
	1.58
	1.27
	1.06
O_2	0.76(A)
	0.69(B)
	0.63(γ)
$O_2 \cdot O_2$	Visible
$O_2 \cdot N_2$	1.26
	4.5
N_2O	4.06, 3.9
	2.97, 2.8
	3.83, 3.5
	3.31, 3.2
CH_4	2.37, 2.3
	2.20
	1.66
CO	4.67
	2.34
NO_2	Visible

where $k_\lambda u$ represents the opt
It is customary to use the p
particularly for water vapor,

where ρ_a denotes the density

Table 3.3
Solar Absorption Bands of Atmospheric Gases

Species	Band (μm)	$\Delta\nu$ (cm^{-1})	Major transitions
	3.2, 2.7	2500–4500	020, 001, 100
	1.87	4800–6200	110, 011
	1.38	6400–7600	021, 200, 101
	1.1	8200–9400	111
H_2O	0.94	10,100–11,300	121, 201, 003
	0.82	11,700–12,700	211
	0.72	13,400–14,600	221, 202, 301
	Visible	15,000–22,600	500, 203
	4.3	2000–2400	00^01
	2.7	3400–3850	10^01
CO_2	2.0	4700–5200	20^01
	1.6	6100–6450	30^01
	1.4	6850–7000	00^03
	4.74	2000–2300	101
	3.3	3000–3100	003
O_3	Visible	10,600–22,600	—
	UV	29,000–50,000	—
	1.58	6300–6350	a←X (0←1)
	1.27	7700–8050	a←X (0←0)
	1.06	9350–9400	a←X (1←0)
O_2	0.76(A)	12,850–13,200	b←X (0←0)
	0.69(B)	14,300–14,600	b←X (1←0)
	0.63(γ)	14,750–15,900	b←X (2←0)
$O_2 \cdot O_2$	Visible	7600–30,000	—
$O_2 \cdot N_2$	1.26	7600–8300	—
	4.5	2100–2300	00^01
N_2O	4.06, 3.9	2100–2800	12^00, 20^00
	2.97, 2.87	3300–3500	02^01, 10^01
	3.83, 3.53		
	3.31, 3.26	2500–3200	0002, 0101, 0200
CH_4	2.37, 2.30		
	2.20	4000–4600	1001, 0011, 0110
	1.66	5850–6100	0020
CO	4.67	2000–2300	1
	2.34	4150–4350	2
NO_2	Visible	14,400–50,000	—

where $k_\lambda u$ represents the optical depth with k_λ the absorption coefficient (cm^2 g^{-1}). It is customary to use the path length (g cm^{-2}) to denote the amount of absorber, particularly for water vapor, in the form

$$u(z) = \int_z^{z_\infty} \rho_a(z')\, dz', \tag{3.2.16}$$

where ρ_a denotes the density of the absorbing gas. The flux density here is a hemispheric

quantity weighted by the cosine of the solar zenith angle. It is this flux density that is related to the solar heating rate defined in Section 3.5.

Because of the structure of absorption lines, it is important to define the spectral absorptance, also referred to as absorptivity or absorption, in the form

$$A_{\bar{\lambda}}(u/\mu_0) = \int_{\Delta\lambda} \left(1 - e^{-k_\lambda u/\mu_0}\right) \frac{d\lambda}{\Delta\lambda}, \tag{3.2.17}$$

where $\Delta\lambda$ denotes the spectral interval. If the solar flux variation is small in this interval, the downward spectral flux density may be expressed by

$$F_{\bar{\lambda}}(z) \cong \mu_0 F_\odot(\bar{\lambda})[1 - A_{\bar{\lambda}}(u/\mu_0)]. \tag{3.2.18}$$

Computation of the spectral absorptance is the key to the evaluation of the downward solar flux in the atmosphere, which can be done by line-by-line techniques or simplified methodologies such as the correlated k-distribution method to be introduced

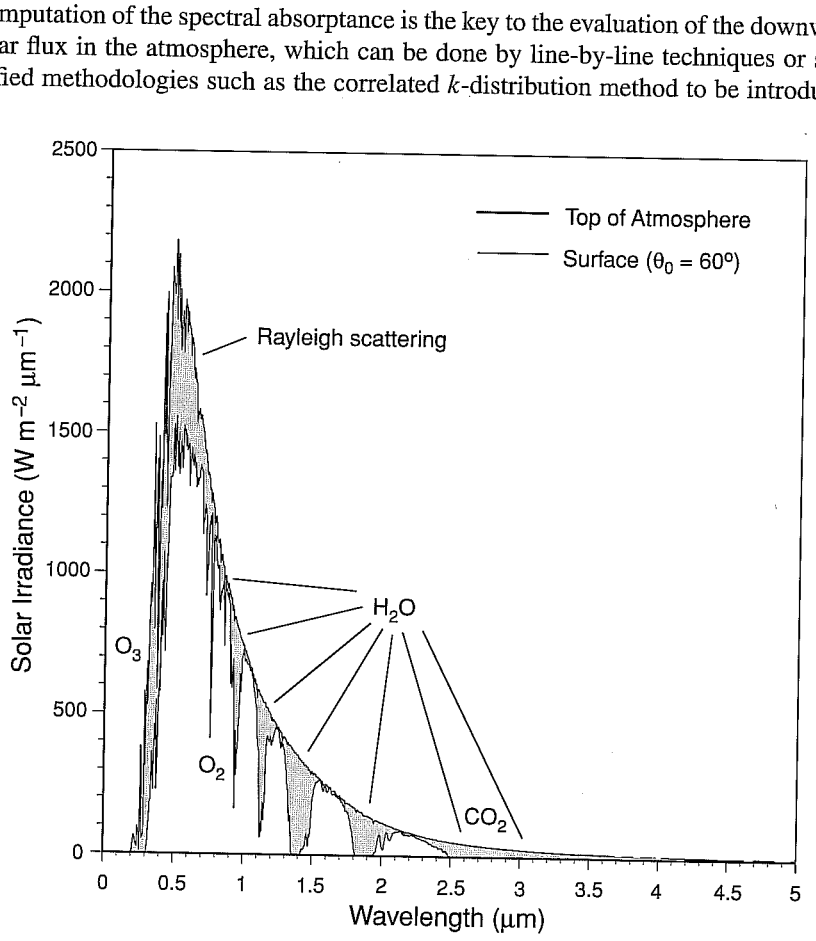


Figure 3.9 Solar irradiance curve for a 50 cm^{-1} spectral interval at the top of the atmosphere (see Fig. 2.9) and at the surface for a solar zenith angle of 60° in an atmosphere without aerosols or clouds. Absorption and scattering regions are indicated. See also Table 3.3 for the absorption of N_2O, CH_4, CO, and NO_2.

in Section 4.5. Exercise 3.6 requires the evaluation of spectral absorptance in the limit of strong line approximation using the Lorentz line shape. The total amount of water vapor can be estimated by this approximation using the 0.94 μm band, which is the foundation for the determination of precipitable water from sunphotometers (see Section 7.2).

Figure 3.9 shows the depletion of solar flux in a clear atmosphere. The top curve is an observed solar irradiance with a spectral resolution of 50 cm^{-1} at the top of the earth's atmosphere as depicted in Fig 2.9. The depletion of solar irradiance in the UV region (<0.4 μm) is chiefly due to the absorption of molecular oxygen and ozone discussed previously. In the visible, the depletion of solar flux is caused by the absorption produced by oxygen red bands, the ozone Chappuis band, and some water vapor weak bands; but the chief attenuation is associated with Rayleigh scattering to be discussed in Section 3.3. In the near-IR, the prime absorber is water vapor with contributions from carbon dioxide in the 2.7 μm band. Other minor absorbers such as N_2O, CO, and CH_4 also contribute to the depletion of solar flux but are less significant. It is evident that water vapor is the most important absorber in the solar near-IR spectrum, which contains about 50% of solar energy.

3.3 Atmospheric Scattering

3.3.1 Rayleigh Scattering

The simplest and in some ways the most important example of a physical law of light scattering with various applications is that discovered by Rayleigh (1871). His findings led to the explanation of the blue color of the sky. In this section we formulate the scattering of unpolarized sunlight by air molecules and describe its important application to the atmosphere.

3.3.1.1 Theoretical Development

Consider a small homogeneous, isotropic, spherical particle whose radius is much smaller than the wavelength of the incident radiation. The incident radiation produces a homogeneous electric field $\mathbf{E}_0$, called the *applied field.* Since the particle is very small, the applied field generates a dipole configuration on it. The electric field of the particle, caused by the electric dipole, modifies the applied field inside and near the particle. Let $\mathbf{E}$ be the combined field, i.e., the applied field plus the particle's own field, and further, let $\mathbf{p}_0$ be the induced dipole moment. Then we apply the electrostatic formula to give

$$\mathbf{p}_0 = \alpha \mathbf{E}_0. \tag{3.3.1}$$

This equation defines the polarizability α of a small particle. The dimensions of $\mathbf{E}_0$ and $\mathbf{p}_0$ are in units of charge per area and charge times length, respectively, and α has the dimension of volume. In general, α is a tensor, because the vectors $\mathbf{p}_0$ and $\mathbf{E}_0$ may not align along the three mutually perpendicular directions. In the very common case where these two vectors coincide, α is a scalar.

The applied field $\mathbf{E}_0$ generates oscillation of an electric dipole in a fixed direction. The oscillating dipole, in turn, produces a plane-polarized electromagnetic wave, the scattered wave. To evaluate the scattered electric field in regions that are far away from the dipole, we let r denote the distance between the dipole and the observation point, γ the angle between the scattered dipole moment $\mathbf{p}$ and the direction of observation, and c the velocity of light. According to the classical electromagnetic solution given by Hertz (1889), the scattered electric field is proportional to the acceleration of the scattered dipole moment and $\sin\gamma$, but is inversely proportional to the distance r. Thus, we have

$$\mathbf{E} = \frac{1}{c^2}\frac{1}{r}\frac{\partial^2 \mathbf{p}}{\partial t^2}\sin\gamma. \tag{3.3.2}$$

In an oscillating periodic field, the scattered dipole moment may be written in terms of the induced dipole moment as

$$\mathbf{p} = \mathbf{p}_0 e^{-ik(r-ct)}. \tag{3.3.3}$$

Note that k is the wavenumber, and $kc = \omega$ is the circular frequency. By combining Eqs. (3.3.1) and (3.3.3), Eq. (3.3.2) yields

$$\mathbf{E} = -\mathbf{E}_0 \frac{e^{-ik(r-ct)}}{r} k^2 \alpha \sin\gamma. \tag{3.3.4}$$

Now we consider the scattering of sunlight by air molecules. Let the plane defined by the directions of incident and scattered waves be the reference plane (plane of scattering). Since any electric vector may be arbitrarily decomposed into orthogonal components, we may choose the two components perpendicular (E_r) and parallel (E_l) to the plane of scattering. The sunlight is characterized by the same electric field in the r and l directions and by a random phase relation between these two components, and is referred to as *natural* or *unpolarized* light (see Section 6.6 for a more advanced discussion of the representation of polarized light). In this case, we may consider separately the scattering of the two electric field components E_{0r} and E_{0l} by molecules assumed to be homogeneous, isotropic, spherical particles. Based on Eq. (3.3.4), we have

$$E_r = -E_{0r}\frac{e^{-ik(r-ct)}}{r} k^2 \alpha \sin\gamma_1, \tag{3.3.5a}$$

$$E_l = -E_{0l}\frac{e^{-ik(r-ct)}}{r} k^2 \alpha \sin\gamma_2. \tag{3.3.5b}$$

Referring to Fig. 3.10, we see that $\gamma_1 = \pi/2$ and $\gamma_2 = \pi/2 - \Theta$, where Θ is defined as the scattering angle, which is the angle between the incident and scattered waves. Note that γ_1 is always equal to 90° because the scattered dipole moment (or the scattered electric field) in the r direction is normal to the scattering plane defined previously.

In matrix form, we may write

$$\begin{bmatrix} E_r \\ E_l \end{bmatrix} = -\frac{e^{-ik(r-ct)}}{r} k^2 \alpha \begin{bmatrix} 1 & 0 \\ 0 & \cos\Theta \end{bmatrix} \begin{bmatrix} E_{0r} \\ E_{0l} \end{bmatrix}. \tag{3.3.6}$$

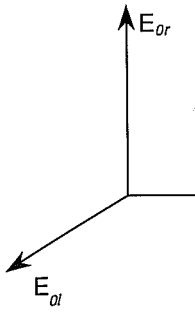


Figure 3.10 Scatteri
into a parallel (l) and a pe
We may select the compo
incident and scattering be

A complete descript
be given in Section
of the continuity of
components (per so
$I_0 = C|E_0|^2$ and I =
C/r^2 implies a solid
in the form of intens

where I_r and I_l are
the plane containing
The total scattered i
direction of Θ is the

But for unpolarized
obtain

This is the original
light by molecules
sunlight scattered b
inversely proportio

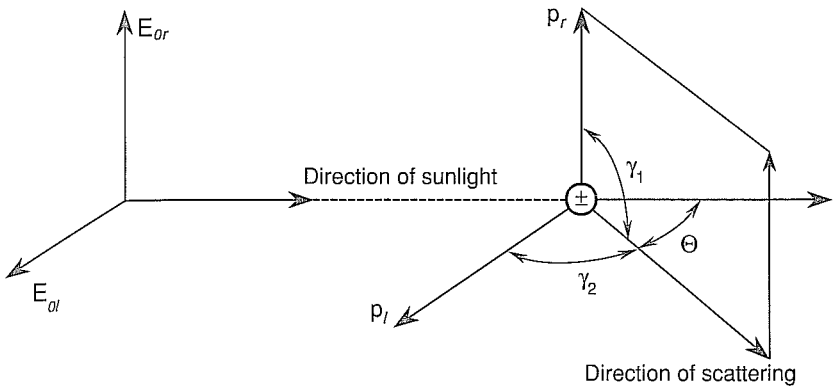


Figure 3.10 Scattering by a dipole. The incident electric field, a vector, can be arbitrarily decomposed into a parallel (l) and a perpendicular (r) component, where each undergoes the scattering by the dipole. We may select the component that is always perpendicular to the scattering plane that is defined by the incident and scattering beams (i.e., $\gamma_1 = 90°$). All the notations are defined in the text.

A complete description of the intensity of a light beam and its polarized state will be given in Section 6.6 in which the Stokes parameters are introduced. For the sake of the continuity of the present discussion, however, we may define the intensity components (per solid angle) of the incident and scattered radiation in the forms $I_0 = C|E_0|^2$ and $I = C|E|^2$, where C is a certain proportionality factor such that C/r^2 implies a solid angle. It follows that Eqs. (3.3.5) and (3.3.6) can be expressed in the form of intensities as

$$I_r = I_{0r} k^4 \alpha^2 / r^2, \tag{3.3.7a}$$

$$I_l = I_{0l} k^4 \alpha^2 \cos^2 \Theta / r^2, \tag{3.3.7b}$$

where I_r and I_l are polarized intensity components perpendicular and parallel to the plane containing the incident and scattered waves, i.e., the plane of scattering. The total scattered intensity of the unpolarized sunlight incident on a molecule in the direction of Θ is then

$$I = I_r + I_l = (I_{0r} + I_{0l} \cos^2 \Theta) k^4 \alpha^2 / r^2. \tag{3.3.8}$$

But for unpolarized sunlight, $I_{0r} = I_{0l} = I_0/2$, and by noting that $k = 2\pi/\lambda$, we obtain

$$I = \frac{I_0}{r^2} \alpha^2 \left(\frac{2\pi}{\lambda} \right)^4 \frac{1 + \cos^2 \Theta}{2}. \tag{3.3.9}$$

This is the original formula derived by Rayleigh, and we call the scattering of sunlight by molecules *Rayleigh scattering*. By this formula, the intensity of unpolarized sunlight scattered by a molecule is proportional to the incident intensity I_0 and is inversely proportional to the square of the distance between the molecule and the

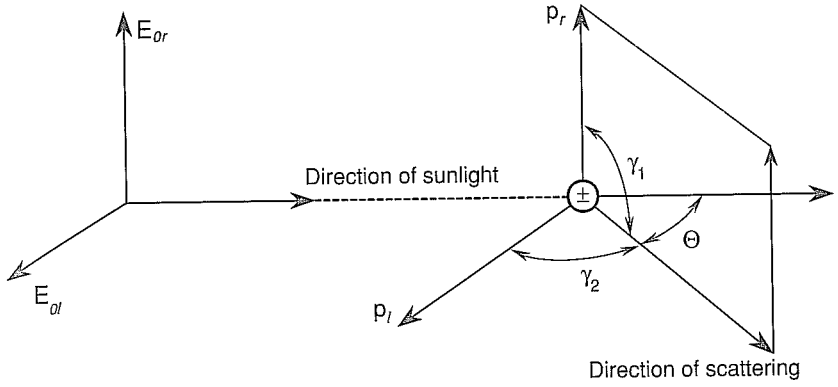


Figure 3.10 Scattering by a dipole. The incident electric field, a vector, can be arbitrarily decomposed into a parallel (l) and a perpendicular (r) component, where each undergoes the scattering by the dipole. We may select the component that is always perpendicular to the scattering plane that is defined by the incident and scattering beams (i.e., $\gamma_1 = 90°$). All the notations are defined in the text.

A complete description of the intensity of a light beam and its polarized state will be given in Section 6.6 in which the Stokes parameters are introduced. For the sake of the continuity of the present discussion, however, we may define the intensity components (per solid angle) of the incident and scattered radiation in the forms $I_0 = C|E_0|^2$ and $I = C|E|^2$, where C is a certain proportionality factor such that C/r^2 implies a solid angle. It follows that Eqs. (3.3.5) and (3.3.6) can be expressed in the form of intensities as

$$I_r = I_{0r} k^4 \alpha^2 / r^2, \tag{3.3.7a}$$

$$I_l = I_{0l} k^4 \alpha^2 \cos^2 \Theta / r^2, \tag{3.3.7b}$$

where I_r and I_l are polarized intensity components perpendicular and parallel to the plane containing the incident and scattered waves, i.e., the plane of scattering. The total scattered intensity of the unpolarized sunlight incident on a molecule in the direction of Θ is then

$$I = I_r + I_l = (I_{0r} + I_{0l} \cos^2 \Theta) k^4 \alpha^2 / r^2. \tag{3.3.8}$$

But for unpolarized sunlight, $I_{0r} = I_{0l} = I_0/2$, and by noting that $k = 2\pi/\lambda$, we obtain

$$I = \frac{I_0}{r^2} \alpha^2 \left(\frac{2\pi}{\lambda} \right)^4 \frac{1 + \cos^2 \Theta}{2}. \tag{3.3.9}$$

This is the original formula derived by Rayleigh, and we call the scattering of sunlight by molecules *Rayleigh scattering*. By this formula, the intensity of unpolarized sunlight scattered by a molecule is proportional to the incident intensity I_0 and is inversely proportional to the square of the distance between the molecule and the

point of observation. In addition to these two factors, the scattered intensity also depends on the polarizability, the wavelength of the incident wave, and the scattering angle. The dependence of these three parameters on the scattering of sunlight by molecules introduces a number of significant physical features.

3.3.1.2 PHASE FUNCTION, SCATTERING CROSS SECTION, AND POLARIZABILITY

On the basis of Eqs. (3.3.7) and (3.3.9), the intensity scattered by a molecule depends on the polarization characteristics of the incident light. For vertically (r) polarized incident light, the scattered intensity is independent of the direction of the scattering plane. In this case then, the scattering is isotropic. On the other hand, for horizontally (l) polarized incident light, the scattered intensity is a function of $\cos^2\Theta$. When the incident light is unpolarized, such as sunlight, the scattered intensity depends on $(1+\cos^2\Theta)$. The angular scattering patterns in space for the three types of incident polarization are illustrated in Fig. 3.11. We see that the scattering of unpolarized sunlight by molecules (Rayleigh scattering) has maxima in the forward (0°) and backward (180°) directions, whereas it shows minima in the side directions (90° and 270°). Light scattered by particles or molecules is not confined only to the plane of incidence, but is visible in all azimuthal directions. Because of the spherical symmetry assumed for molecules, scattering patterns are symmetrical in three-dimensional space, as demonstrated in Fig. 3.11.

To describe the angular distribution of scattered energy in conjunction with multiple scattering and radiative transfer analyses and applications for planetary atmospheres, we find it necessary to define a nondimensional parameter called the *phase function,* $P(\cos\Theta)$, such that

$$\int_0^{2\pi}\int_0^{\pi}\frac{P(\cos\Theta)}{4\pi}\sin\Theta\,d\Theta\,d\phi=1. \tag{3.3.10}$$

By this definition, the phase function is said to be normalized to unity. Upon performing simple integrations, the phase function of Rayleigh scattering for incident unpolarized sunlight is given by

$$P(\cos\Theta)=\frac{3}{4}(1+\cos^2\Theta). \tag{3.3.11}$$

Employing the definition of the phase function, Eq. (3.3.9) may be rewritten in the form

$$I(\Theta)=\frac{I_0}{r^2}\alpha^2\frac{128\pi^5}{3\lambda^4}\frac{P(\Theta)}{4\pi}. \tag{3.3.12}$$

It follows that the angular distribution of the scattered intensity is directly proportional to the phase function.

The scattered flux f (or power, in units of energy per time) can be evaluated by integrating the scattered flux density ($I\Delta\Omega$) over the appropriate area a distance r

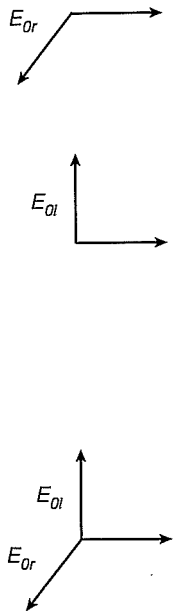


Figure 3.11 Polar
light with the electric v
electric vector on the sc

away from the scat

where $r^2 d\Omega$ repres
ing the expressions
Eqs. (3.3.12) and (
over the solid angl

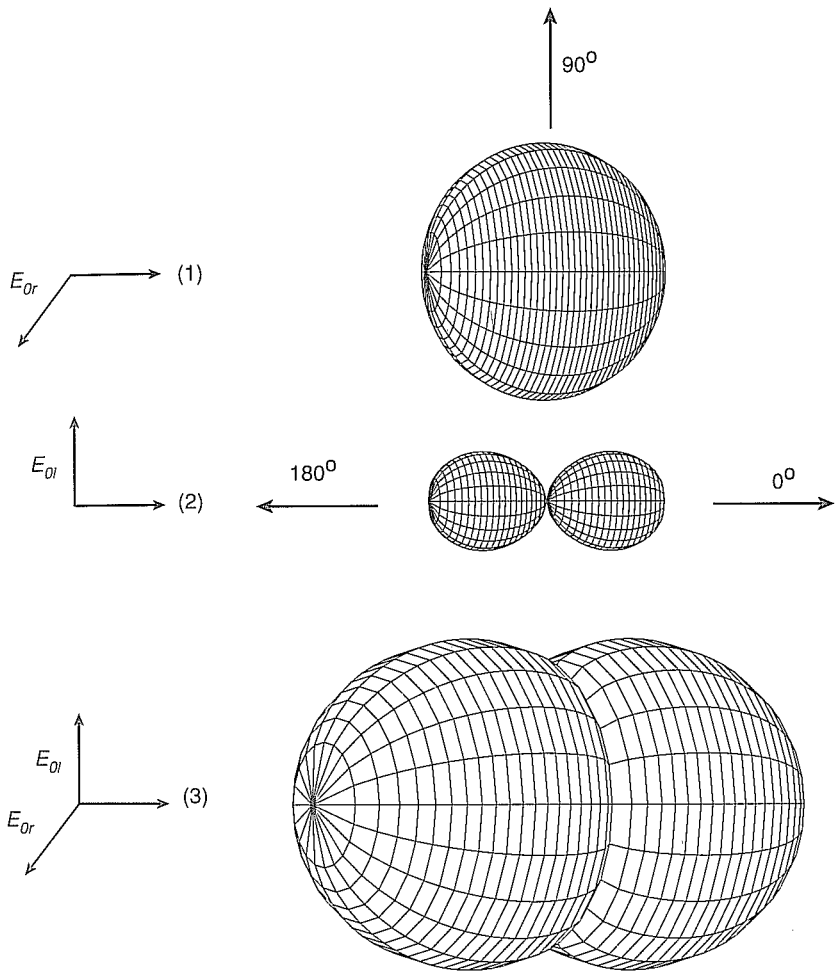


Figure 3.11 Polar diagram of the scattered intensity for Rayleigh molecules: (1) polarized incident light with the electric vector perpendicular to the scattering plane, (2) polarized incident light with the electric vector on the scattering plane, and (3) unpolarized incident light.

away from the scatterer. Thus,

$$f = \int_{\Omega} (I \Delta\Omega) r^2 \, d\Omega, \tag{3.3.13a}$$

where $r^2 d\Omega$ represents the area according to the definition of the solid angle. Inserting the expressions for scattered intensity and the differential solid angle defined in Eqs. (3.3.12) and (1.1.5), respectively, into Eq. (3.3.13a) and carrying out integrations over the solid angle of a sphere, we obtain the equivalent isotropically scattered flux

in the form

$$f = F_0\alpha^2 128\pi^5/(3\lambda^4), \tag{3.3.13b}$$

where the incident flux density F_0 is equal to $I_0\Delta\Omega$. Moreover, we may define the scattering cross section per one molecule as

$$\sigma_s = f/F_0 = \alpha^2 128\pi^5/(3\lambda^4). \tag{3.3.14}$$

The scattering cross section (in units of area) represents the amount of incident energy that is removed from the original direction because of a single scattering event such that the energy is redistributed isotropically on the area of a sphere whose center is the scatterer and whose radius is r. In terms of the scattering cross section, the scattered intensity can be expressed by

$$I(\Theta) = I_0 \frac{\sigma_s}{r^2} \frac{P(\Theta)}{4\pi}. \tag{3.3.15}$$

This is the general expression for scattered intensity, which is valid not only for molecules but also for particles whose size is larger than the incident wavelength, as will be discussed in Section 5.2.

The *polarizability* α, which was used in the preceding equations, can be derived from the principle of the dispersion of electromagnetic waves and is given by

$$\alpha = \frac{3}{4\pi N_s}\left(\frac{m^2-1}{m^2+2}\right), \tag{3.3.16}$$

where N_s is the total number of molecules per unit volume and m is the nondimensional refractive index of molecules. This equation is called the *Lorentz–Lorenz formula,* and its derivation is given in Appendix D. The refractive index is an optical parameter associated with the velocity change of electromagnetic waves in a medium with respect to a vacuum. Its definition and physical meanings are also given in Appendix D. Normally, the refractive indices of atmospheric particles and molecules are composed of a real part m_r and an imaginary part m_i corresponding, respectively, to the scattering and absorption properties of particles and molecules. In the solar visible spectrum, the imaginary parts of the refractive indices of air molecules are so insignificantly small that absorption of solar radiation by air molecules may be neglected in the scattering discussion. The real parts of the refractive indices of air molecules in the solar spectrum are very close to 1, but they depend on the wavelength (or frequency) of the incident radiation as illustrated in Appendix D. Because of this dependence, white light may be *dispersed* into component colors by molecules that function like prisms. The real part of the refractive index derived in Appendix D [(Eq. D.17)] may be approximately fitted by

$$(m_r - 1)\times 10^8 = 6432.8 + \frac{2{,}949{,}810}{146-\lambda^{-2}} + \frac{25{,}540}{41-\lambda^{-2}}, \tag{3.3.17}$$

where λ is in units of micrometers. Since m_r is close to 1, for all practical purposes, Eq. (3.3.16) may be approximated by

$$\alpha \approx \frac{1}{4\pi N_s}\left(m_r^2 - 1\right). \tag{3.3.18}$$

Thus, the scattering cross section defined in Eq. (3.3.14) becomes

$$\sigma_s = \frac{8\pi^3\left(m_r^2 - 1\right)^2}{3\lambda^4 N_s^2} f(\delta). \tag{3.3.19}$$

A correction factor $f(\delta)$ is added in Eq. (3.3.19) to take into consideration the anisotropic property of molecules, where $f(\delta) = (6 + 3\delta)/(6 - 7\delta)$ with the anisotropic factor δ of 0.035. Anisotropy implies that the refractive index of molecules varies along the x, y, and z directions, and thus is a vector, not a scalar. Hence, the polarizability α is a tensor, as noted previously.

The optical depth of the entire molecular atmosphere at a given wavelength may be calculated from the scattering cross section in the form

$$\tau(\lambda) = \sigma_s(\lambda) \int_0^{z_\infty} N(z)\, dz, \tag{3.3.20}$$

where $N(z)$ denotes the number density of molecules as a function of height, and z_∞ is the top of the atmosphere. The optical depth represents the attenuation power of molecules with respect to a specific wavelength of the incident light. Exercises 3.7–3.11 require the calculation of a number of parameters based on Rayleigh scattering results.

3.3.1.3 Blue Sky and Sky Polarization

Returning to Eq. (3.3.12), we see that the scattered intensity depends on the wavelength of incident light and the index of refraction of air molecules contained in the polarizability term. According to the analyses given in Appendix D and Eq. (3.3.17), the index of refraction also depends slightly on the wavelength. However, the dependence of the refractive index on the wavelength is relatively insignificant in calculating the scattered intensity as compared to the explicit wavelength term. Thus, the intensity scattered by air molecules in a specific direction may be symbolically expressed in the form

$$I_\lambda \sim 1/\lambda^4. \tag{3.3.21}$$

The inverse dependence of the scattered intensity on the wavelength to the fourth power is a direct consequence of the theory of Rayleigh scattering and is the foundation for the explanation of blue sky.

In reference to the observed solar energy spectrum displayed in Fig. 3.9, a large portion of solar energy is contained between the blue and red regions of the visible spectrum. Blue light ($\lambda \approx 0.425\ \mu$m) has a shorter wavelength than red light ($\lambda \approx 0.650\ \mu$m). Consequently, according to Eq. (3.3.21) blue light scatters about 5.5 times

more intensity than red light. It is apparent that the λ^{-4} law causes more blue light to be scattered than red, green, and yellow, and so the sky, when viewed away from the sun's disk, appears blue. Moreover, since molecular density decreases drastically with height, it is anticipated that the sky should gradually darken to become completely black in outer space in directions away from the sun. And the sun itself should appear whiter and brighter with increasing height. As the sun approaches the horizon (at sunset or sunrise), sunlight travels through more air molecules, and therefore more and more blue light and light with shorter wavelengths are scattered out of the beam of light, and the luminous sun shows a deeper red color than at its zenith. However, since violet light ($\sim$0.405 μm) has a shorter wavelength than blue, a reasonable question is, why doesn't the sky appear violet? This is because the energy contained in the violet spectrum is much less than that contained in the blue spectrum, and also because the human eye has a much lower response to the violet color.

Another important phenomenon explained by the Rayleigh scattering theory is sky polarization. For many atmospheric remote sensing applications utilizing polarization, a parameter called the *degree of linear polarization* has been used (Subsection 7.3.5.2). In the case of Rayleigh scattering it is given by

$$LP(\Theta) = -\frac{I_l - I_r}{I_l + I_r} = -\frac{\cos^2\Theta - 1}{\cos^2\Theta + 1} = \frac{\sin^2\Theta}{\cos^2\Theta + 1}. \tag{3.3.22}$$

In the forward and backward directions the scattered light remains completely unpolarized, whereas at the 90° scattering angle, the scattered light becomes completely polarized. In other directions, the scattered light is partially polarized with the percentage of polarization ranging from 0 to 100%. Interested readers may wish to refer to Section 6.6 for further details on this subject.

The theory of Rayleigh scattering developed in Section 3.3.1 is based on the assumption that molecules are homogeneous and isotropic spheres. However, molecules are in general anisotropic, whereby their polarizability, as defined in Eq. (3.3.16), varies along three axes and, hence, is a tensor instead of a scalar. The anisotropic effect of molecules reduces the degree of linear polarization defined in Eq. (3.3.22) by only a small percentage. At the 90° scattering angle, the degree of linear polarization for dry air is about 0.94. Further, the theory of Rayleigh scattering developed previously considers only single (or primary) scattering, i.e., where scattering occurs only once. But in the earth's atmosphere, which contains a large number of molecules and aerosol particles, light may undergo an infinite number of scattering events. In addition, the earth's surface also reflects light that reaches it. Multiple scattering processes involving the atmosphere and the surface become complicated and require a more advanced treatment of radiative transfer theory, which will be discussed in Chapter 6.

The theory of Rayleigh scattering predicts *neutral points,* i.e., points of zero polarization, only at the exact forward and backward directions. However, owing to multiple scattering of molecules and particulates, and reflection of the surface, there normally exist a number of neutral points in cloudless atmospheres. The first observations of neutral points and partially polarized sky light were made by Arago in

1809. He discovere
25° above the antis
other two neutral p
below the sun, were
tively. These three
neutral points in the
of aerosol loadings
characteristics of th

Figure 3.12 illu
points for a pristin
condition under the
Loa Observatory fr

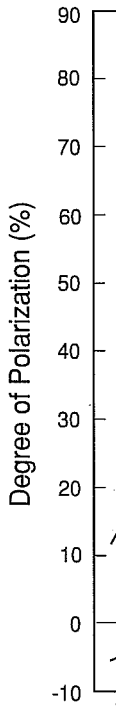


Figure 3.12 Illust
plane of the sun's vertic
atmospheric condition
on July 27, 1982 (dashe
The sun's elevation ang
and Babinet (data taken

1809. He discovered the existence of a neutral point at a position in the sky about 25° above the antisolar direction (the direction exactly opposite that of the sun). The other two neutral points, which normally occur in the sunlit sky 25° above and 20° below the sun, were discovered by Babinet in 1840 and by Brewster in 1842, respectively. These three neutral points were named to honor these three discoverers. The neutral points in the sky vary and depend on the turbidity (an indication of the amount of aerosol loadings in the atmosphere), the sun's elevation angle, and the reflection characteristics of the surface at which observations are made.

Figure 3.12 illustrates the distribution of the degree of polarization and neutral points for a pristine, clear atmosphere (January 20, 1977) and for an atmospheric condition under the El Chichon volcanic cloud (July 27, 1982) observed at the Mauna Loa Observatory from a polarimeter developed by Coulson (1983). The observations

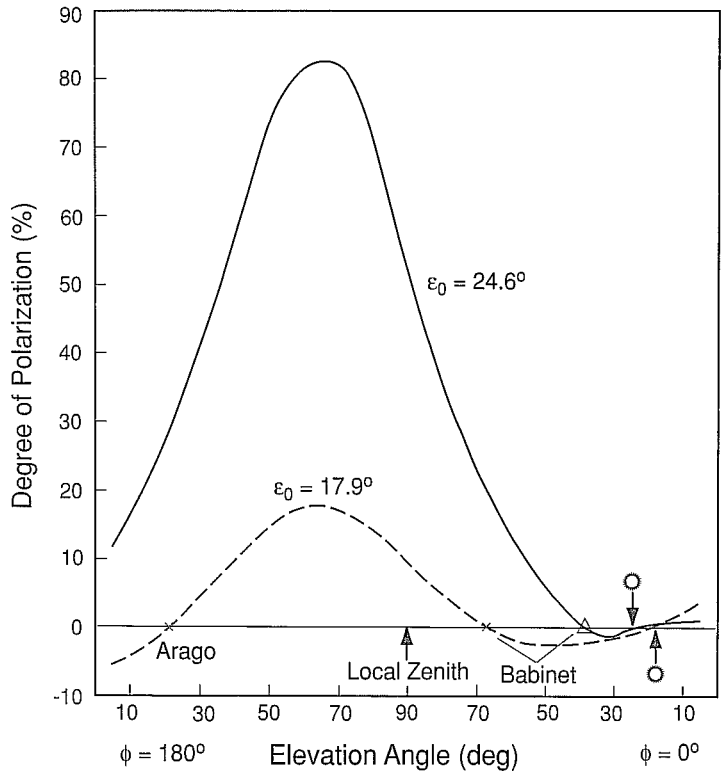


Figure 3.12 Illustration of neutral points in the distribution of the degree of polarization through the plane of the sun's vertical at a wavelength of 0.7 μm observed at the Mauna Loa Observatory for a clear atmospheric condition on January 20, 1977 (solid line), and for an atmosphere under the volcanic cloud on July 27, 1982 (dashed line). The azimuthal angles $\phi = 0°$ and $\phi = 180°$ are on the sun's vertical plane. The sun's elevation angles ε_0 for these two cases are indicated in the graph, as are the positions of Arago and Babinet (data taken from Coulson, 1983).

were made on the sun's vertical plane, referred to as the principal plane in radiative transfer, using a wavelength of 0.7 μm. The solar elevation angle, ε_0 (90° – solar zenith angle θ_0), differed slightly on these two dates, but the observed polarization patterns suffice to demonstrate their substantial variabilities in clear and turbid atmospheres. The clear Rayleigh atmosphere produced a maximum polarization of about 80%, 60% more than that generated in the volcanic cloud condition. The neutral points in the Rayleigh scattering atmosphere occurred at the positions close to the sun (forward direction) and about 20° above the sun, the Babinet point, which was about 50° above the sun when a significant aerosol loading was present. In this case, the Arago point was also shown at about 20° above the horizon at the opposite position of the sun. Because of the sun's position, the Brewster point was not observed. The neutral points' positions are dependent on the aerosol optical depth and composition. Thus, a systematic observation of these points could be a valuable approach for inferring aerosol optical properties and perhaps composition information.

3.3.2 Light Scattering by Particulates: Approximations

In Section 1.1.4, we defined the size parameter, $x = 2\pi a/\lambda$, where a is the particle radius. Rayleigh scattering is concerned with scattering events when $x \ll 1$. When $x \gtrsim 1$, scattering events are often called *Lorenz–Mie scattering*. Lorenz (1890) and Mie (1908) independently derived the solution for the interaction of a plane wave with an isotropic homogenous sphere. The mathematical theory of Lorenz–Mie scattering begins with Maxwell's equations and will be detailed in Chapter 5, along with some new developments in research on light scattering by nonspherical ice crystals and aerosols. In this section, however, we shall present a brief discussion of Lorenz–Mie scattering and two elementary approximations: geometric optics and anomalous diffraction.

3.3.2.1 Lorenz–Mie Scattering

The intensity scattered by a particle as a function of direction, as presented in Eq. (3.3.15), is given by

$$I(\Theta) = I_0 \Omega_{\text{eff}} \frac{P(\Theta)}{4\pi} = I_0 \left(\frac{\sigma_s}{r^2}\right) \frac{P(\Theta)}{4\pi}, \tag{3.3.23}$$

where I_0 is the incident intensity, P is the phase function normalized according to Eq. (3.3.10), Ω_{eff} is the effective solid angle upon which scattering occurs, r is the distance between the particle and the observer, σ_s is the scattering cross section, and 4π is the solid angle for the entire spherical space. The scattering cross section can be derived from the Lorenz–Mie theory of light scattering by spheres and is given by the following expansion:

$$\sigma_s/\pi a^2 = Q_s = c_1 x^4 \left(1 + c_2 x^2 + c_3 x^4 + \cdots\right), \tag{3.3.24}$$

where a is the radius, $x = 2\pi a/\lambda$, Q_s is referred to as the *scattering efficiency,* and the coefficients in the case of nonabsorbing particles are given by

$$c_1 = \frac{8}{3}\left(\frac{m^2-1}{m^2+2}\right)^2, \quad c_2 = \frac{6}{5}\left(\frac{m^2-1}{m^2+2}\right),$$

$$c_3 = \frac{3}{175}\frac{m^6+41m^4-28m^2+284}{(m^2+2)^2} + \frac{1}{900}\left(\frac{m^2+2}{2m^2+2}\right)^2\left[15+(2m^2+3)^2\right].$$

The leading term is the dipole mode contribution associated with Rayleigh scattering. Note that for light scattering by spheres, we may replace the total number of molecules per volume N_s by $1/V$ where $V = 4\pi a^3/3$. For molecules, $a \sim 10^{-4}$ μm, so that $x \sim 10^{-3}$ in the visible. Thus, the higher order terms can be neglected and the scattered intensity is proportional to λ^{-4}. For aerosols and cloud particles, $a \gtrsim 10^{-1}$ μm, and $x \gtrsim 1$ in the visible. In this case, the scattered intensity is less wavelength dependent and is primarily dependent on particle size. As a result, clouds and nonabsorbing aerosols in the atmosphere generally appear white. In a cloudy atmosphere, the sky appears blue diluted with white scattered light, resulting in a less pure blue sky than would have been expected from pure Rayleigh scattering.

On the basis of Eq. (3.3.23), the scattered intensity is dependent on the phase function, which can be computed from the Lorenz–Mie theory for spheres. Figure 3.13 shows typical examples of the phase function for polydispersed cloud droplets ($\sim$10 μm) and aerosols ($\sim$1 μm) illuminated by a visible light. Also shown is the phase function for Rayleigh scattering. The mean size parameters in these cases are about 100, 10, and 10^{-3}, respectively. The scattering by cloud droplets is characterized by a strong forward diffraction; a minimum at $\sim$100° scattering angle; a peak at $\sim$138° scattering angle, the well-known rainbow feature; and a peak in the backscattering direction associated with the glory pattern. The diffraction pattern and the rainbow feature will be discussed further later; the explanation of the glory pattern requires more advanced discussion and will be presented in Chapter 5. The scattering of typical aerosols also displays a forward diffraction maximum and a maximum pattern in the 150°–170° scattering region (see also Fig. 1.4).

3.3.2.2 Geometric Optics

The principles of geometric optics are the asymptotic approximations of the fundamental electromagnetic theory and are valid for light-scattering computations involving a particle whose dimension is much larger than the wavelength, i.e., $x \gg 1$. In this case, a light beam can be thought of as consisting of a bundle of separate parallel rays that hit the particles, which is referred to as the *localization principle*. Each ray will then undergo reflection and refraction and will pursue its own path along a straight line outside and inside the scatterer with propagation directions determined by the *Snell law,* as shown in Fig. 3.14a. In the context of geometric optics, the total electric field is assumed to consist of the diffracted rays and the reflected and refracted rays, as illustrated in Fig. 3.14b, using a sphere as an example. The diffracted rays

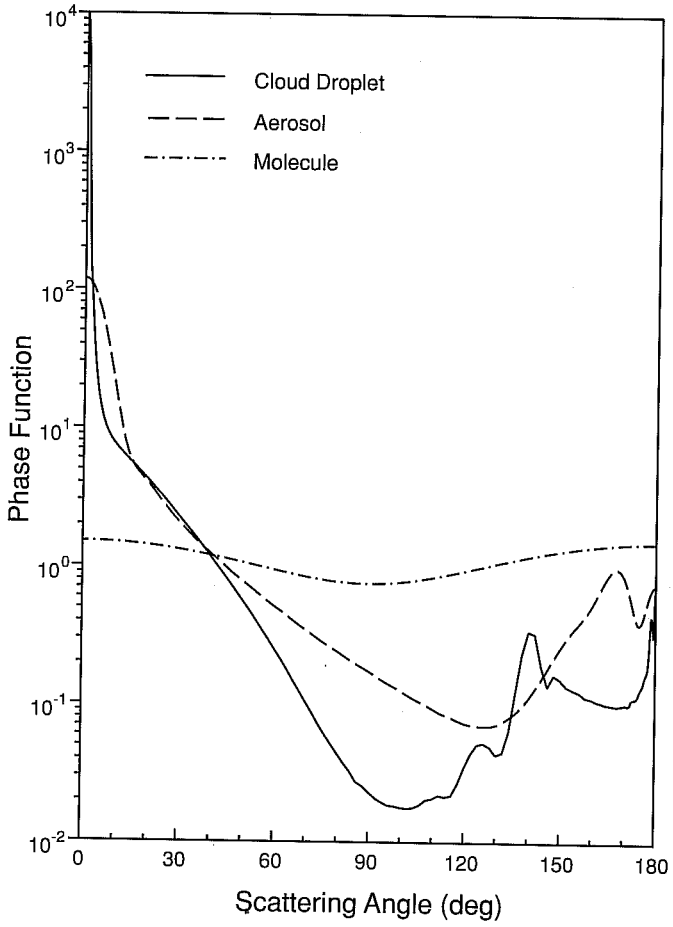


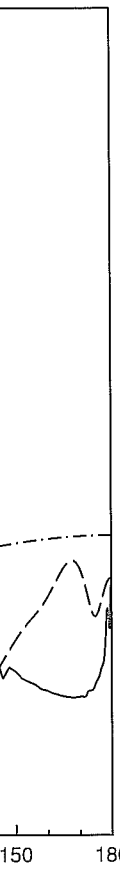
Figure 3.13 Normalized phase functions for cloud droplets (~10 μm), aerosols (~1 μm), and molecules (~10^{-4} μm) illuminated by a visible wavelength of 0.5 μm, computed from the Lorenz–Mie theory.

pass around the scatterer. The rays impinging on the scatterer undergo local reflection and refraction, referred to as *Fresnelian interaction*. The energy that is carried by the diffracted and the Fresnelian rays is assumed to be the same as the energy that is intercepted by the particle cross section projected along the incident direction.

In reference to Fig. 3.14a, let v_1 and v_2 be the velocities of propagation of plane waves in the two media such that $v_1 > v_2$. Also, let θ_i and θ_t be the angles corresponding to the incident and refracted waves. Thus, we have

$$\sin\theta_i / \sin\theta_t = v_1/v_2 = m, \tag{3.3.25}$$

where m is the index of refraction for the second medium with respect to the first.



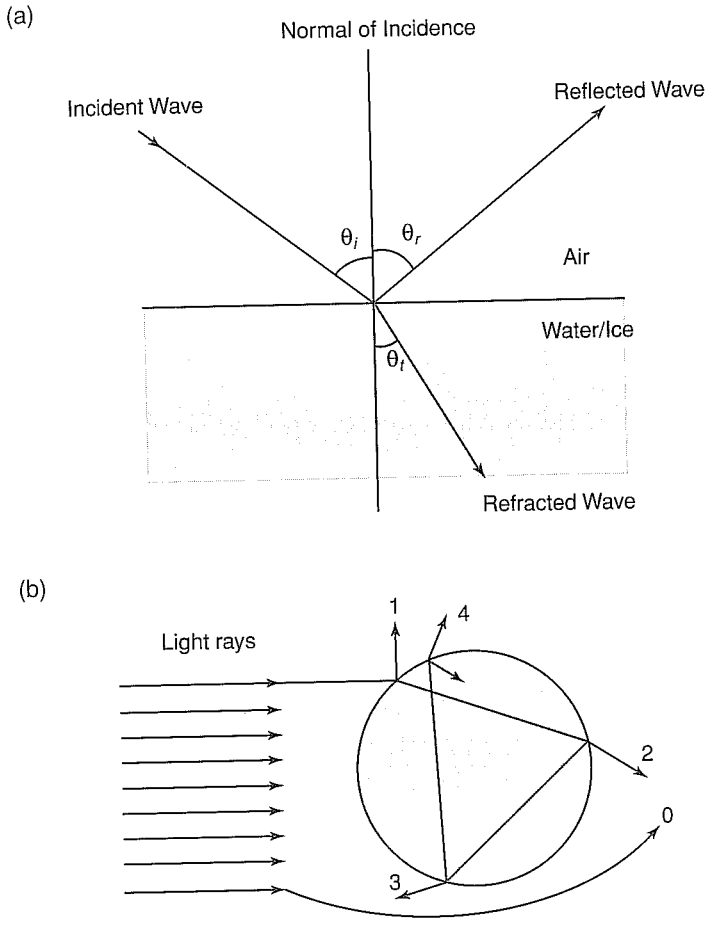


Figure 3.14 (a) Reflection and refraction of a plane wave from air to water/ice surface. (b) Representation of light rays scattered by a sphere based on the geometric optics principle: 0, exterior diffraction; 1, external reflection; 2, two refractions; 3, one internal reflection; and 4, two internal reflections.

For the purpose of this discussion, we shall assume that there is no absorption in the medium. This is the *Snell law* relating the incident and refracted angles through the index of refraction. Exercises 3.12 and 3.13 require the derivation of the minimum deviations of light rays that produce *rainbows* from spherical water droplets and *halos* from hexagonal ice crystals. Moreover, white sunlight is decomposed into component colors after the rays undergo geometric reflection and refraction through water droplets and ice crystals.

The diffraction component in geometric optics can be determined from *Babinet's principle*. This principle states that the diffraction pattern in the far field, referred to as *Fraunhofer diffraction,* from a circular aperture is the same as that from an opaque

disk or sphere of the same radius. Based on this principle and geometric consideration, the scattered intensity is proportional to

$$I_p = \frac{x^4}{4}\left[\frac{2J_1(x\sin\Theta)}{x\sin\Theta}\right]^2, \tag{3.3.26}$$

where J_1 is the first-order *Bessel function* and Θ is the scattering angle. Exercise 3.14 requires the calculation of maxima and minima of the diffraction pattern that can be used to explain an optical phenomenon known as the *corona*.

One final note is in order here. If a particle of any shape is much larger than the incident wavelength, the total energy removed is based on geometric reflection and refraction, giving an effective cross-section area equal to the geometric area A. In addition, according to Babinet's principle, diffraction takes place through a hole in this area, giving a cross-section area also equal to A. The total removal of incident energy is therefore twice the geometric area. Thus, the extinction cross section is given by

$$\sigma_e = 2A, \text{ or } Q_e = \sigma_e/A = 2, \tag{3.3.27}$$

where Q_e is called the *extinction efficiency*. This is referred to as the *optical theorem of extinction*. If a particle is nonabsorbing, then we have $Q_e = Q_s$, where the extinction and scattering efficiencies are the same.

3.3.2.3 Anomalous Diffraction Theory

Consider large optically soft particles such that $x \gg 1$ and $|m-1| \ll 1$. The second condition implies that rays are negligibly deviated as they cross the soft particle boundary and are negligibly reflected because the refractive indices inside and outside the particle are similar. In this case, the extinction is largely caused by absorption of the light beam passing through the particle, as well as by the interference of light passing through the particle and passing around the particle. This is the physical foundation for the anomalous diffraction theory originally developed by van de Hulst (1957). In reference to Fig. 3.15, let the plane wave be incident on a spherical particle with a radius a and a refractive index $m \to 1$. The wave front on the forward side of the particle can be divided into two types: one within the geometric shadow area denoted by $A = \pi a^2$, and one outside this area denoted by B. The incident rays can undergo diffraction and pass around the particle. The rays can also hit the particle and undergo reflection and refraction. Since $m \to 1$, we may assume that the rays enter into the particle and pass through it, as illustrated in Fig. 3.15. However, these rays will have phase lags due to the presence of the particle. The phase lag for the ray indicated in the figure is $2a\sin\alpha(m-1)\cdot 2\pi/\lambda$. If we define the phase shift parameter

$$\rho = 2x(m-1), \tag{3.3.28}$$

the phase lag can then be expressed by $\rho\sin\alpha$.

Consider a screen that collects the field. The resultant wave on the screen is the sum of the incident and scattered fields. If the incident field is assumed to be unity,

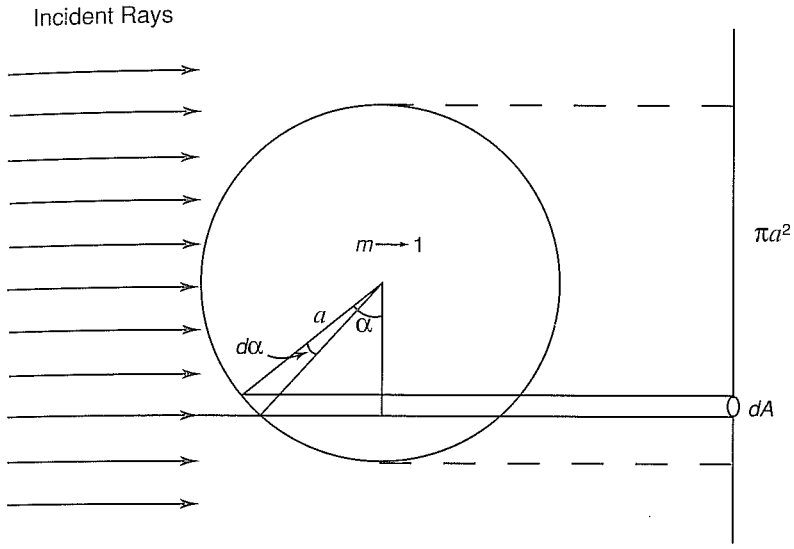


Figure 3.15 Geometry of anomalous diffraction through a sphere with a radius a and an index of refraction $m \to 1$. πa^2 denotes the geometric cross-section area of the sphere and dA denotes the differential cross-section area.

then in the forward direction ($\Theta = 0$), the change in the electric field is proportional to

$$A = \int\int \left(1 - e^{-i\rho \sin\alpha}\right) dx\, dy. \tag{3.3.29a}$$

The differential area can be replaced by an area in the polar coordinate such that $dx\, dy = a\cos\alpha\, d(a\cos\alpha)\, d\phi$. Thus, we have

$$A = \int_0^{2\pi}\int_0^{\pi/2} \left(1 - e^{-i\rho\sin\alpha}\right) a^2 \sin\alpha\, d\sin\alpha\, d\phi = 2\pi a^2 K(i\rho), \tag{3.3.29b}$$

where

$$K(i\rho) = \frac{1}{2} + \frac{e^{-i\rho}}{i\rho} + \frac{e^{-i\rho} - 1}{(i\rho)^2}. \tag{3.3.30}$$

The extinction cross section σ_e is proportional to the differential change in the scattered intensity I. Since $I \sim |E|^2$, as shown in Eq. (3.3.6), $dI \sim 2d|E|$. Thus, we have $\sigma_e = 2\mathrm{Re}(A)$. It follows that the extinction efficiency is given by

$$Q_e = \sigma_e/\pi a^2 = 4\mathrm{Re}[K(i\rho)] = 2 - \frac{4}{\rho}\sin\rho + \frac{4}{\rho^2}(1 - \cos\rho), \tag{3.3.31}$$

where Re denotes the real part of the function. Exercise 3.15 requires calculations of Q_e.

We may also determine the absorption efficiency by the following procedure. The ray path as shown in Fig. 3.15 is $l = 2a\sin\alpha$. The absorption coefficient $k_i = m_i 2\pi/\lambda$,

where m_i is the imaginary part of the refractive index. Thus, the absorption path length associated with the electric field is lk_i. The attenuation of the intensity of the ray is then $\exp(-2lk_i)$ and the absorption cross section for all possible rays is

$$\sigma_a = \int\int \left(1 - e^{-2lk_i}\right) dx\, dy. \tag{3.3.32}$$

Following the procedure just illustrated, the absorption efficiency is given by

$$Q_a = \sigma_a/\pi a^2 = 1 + \frac{2}{b}e^{-b} + \frac{2}{b^2}(e^{-b} - 1), \tag{3.3.33}$$

where $b = 4xm_i$ and $x = 2\pi a/\lambda$. The approximation based on the anomalous diffraction theory (ADT) is useful for the calculation of the extinction and absorption coefficients when $m \to 1$. It can also be applied to nonspherical particles such as spheroids and hexagons. Since refractions and reflections of rays are neglected in this approximation, its accuracy must be examined carefully when applied to the scattering of ice crystals ($m \sim 1.31$) and aerosols ($m \sim 1.5$). Finally, it should be noted that the ADT approximation cannot produce the phase function pattern.

3.4 Multiple Scattering and Absorption in Planetary Atmospheres

3.4.1 Fundamentals of Radiative Transfer

In Section 1.1.4, we pointed out that scattering is often coupled with absorption. In the following we formulate the fundamental equation governing the transfer of diffuse solar radiation in plane-parallel atmospheres. The term *diffuse* is associated with multiple scattering processes and is differentiated from *direct* solar radiation. In reference to Fig. 3.16 and considering a differential thickness Δz, the differential change of diffuse intensity emergent from below the layer is due to the following processes: (1) reduction from the extinction attenuation; (2) increase from the single scattering of the unscattered direct solar flux from the direction $(-\mu_0, \phi_0)$ to (μ, ϕ); (3) increase from multiple scattering of the diffuse intensity from directions (μ', ϕ') to (μ, ϕ); and (4) increase from emission within the layer in the direction (μ, ϕ). Consider a small volume containing a spectrum of molecules and/or particulates and denote the extinction, scattering, and absorption coefficients (in units of per length) as β_e, β_s, and β_a, respectively, defined by

$$\beta_{e,s,a} = \int_{\Delta z} \sigma_{e,s,a}(z) n(z)\, dz/\Delta z, \tag{3.4.1}$$

where the symbol σ denotes the cross section and n is the number density. Moreover, let the phase function corresponding to a volume of particulates be P. Thus, $P(\mu, \phi; \mu', \phi')$ denotes the redirection of the incoming intensity defined by (μ', ϕ') to the outgoing intensity defined by (μ, ϕ). Also note that the differential length