Notes on Radiative Transfer

G&G 401b

February 7, 2005
# Contents

1 Introduction .......................................................... 7  
1.1 Radiative transfer in the sciences ................................. 7  
1.2 The nature of radiation ............................................ 9  
1.3 Quantifying radiation .............................................. 11  
  1.3.1 Note on units and symbols .................................. 11  
  1.3.2 Quantifying intensity of radiation .......................... 12  
1.4 Interaction between radiation and matter ........................ 15  
  1.4.1 Scattering .................................................... 15  
  1.4.2 Absorption and emission ................................... 16  
1.5 Blackbody radiation ............................................... 17  
1.6 Emissivity ......................................................... 19  

2 The radiative transfer equation ..................................... 23  
2.1 Processes affecting radiative intensity ........................... 23  
  2.1.1 Absorption .................................................. 23  
  2.1.2 Emission .................................................... 24  
  2.1.3 Scattering out of $\Omega$ .................................. 25  
  2.1.4 Scattering into $\Omega$ ..................................... 25  
2.2 Full Equation ...................................................... 25  
  2.2.1 Optical depth ................................................ 27  
2.3 General solution .................................................. 27  
  2.3.1 Polarization ................................................ 28  

3 Scattering within media .............................................. 30  
3.1 Scattering by single atoms or molecules ......................... 31  
  3.1.1 Ideal dipole ................................................ 31  
  3.1.2 Real atoms/molecules ..................................... 34  
3.2 Multiple dipoles .................................................. 34  
3.3 Continuous media ................................................ 38  
  3.3.1 Index of refraction in gases ................................ 40  
  3.3.2 Rayleigh scattering and the blue sky ....................... 43  
  3.3.3 Index of refraction in dense media ......................... 45  
3.4 Scattering in heterogeneous media, and the radiation approximation to Maxwell’s equations ................................................. 47
4 Scattering by finite particles ........................................ 52
4.1 Small spheres \( (x \ll 1) \) ........................................ 52
        4.1.1 Mean field in a medium near a dipole .......................... 55
        4.1.2 Small, non-spherical particles .............................. 56
4.2 The “Mie regime” \( (x) \) ........................................ 56
4.3 Geometrical optics \( (x) \) ........................................ 59
4.4 Surface reflection and albedo ..................................... 59

5 Absorption and Emission ........................................... 63
5.1 Electronic transitions ............................................ 66
5.2 Vibrational transitions .......................................... 67
5.3 Rotational transitions and V-R bands ........................... 70
5.4 Electronic bands .................................................. 73
5.5 Free-free opacity .................................................. 75
5.6 Blackbody radiation and equilibrium, revisited ............... 75
        5.6.1 Non-LTE effects ........................................... 76
        5.6.2 Nonconservative scattering ................................ 78
5.7 Solar emission .................................................... 78
5.8 Planetary atmospheres ........................................... 80
        5.8.1 Earth ....................................................... 81
        5.8.2 The other inner planets .................................. 82
        5.8.3 The gas giants ............................................ 86
5.9 Line shape and broadening ....................................... 86
        5.9.1 Absorption by a single line .............................. 88
5.10 Aborption over finite spectral bands ......................... 90
        5.10.1 Grey model .............................................. 90
        5.10.2 Regular band model .................................... 91
        5.10.3 Random band model .................................... 92
        5.10.4 \( k \)-distributions ...................................... 93
        5.10.5 Observed band characteristics in Earth’s atmosphere . 94

6 Practical Solution of the Radiative Transfer Equation ........ 96
6.1 When is scattering important? .................................. 96
6.2 The plane-parallel approximation ................................ 97
6.3 Diffuse transmission and the diffusivity factor .............. 99
        6.3.1 The planetary greenhouse effect .......................... 100
6.4 Solutions with scattering ........................................ 101
6.5 Separation of direct and diffuse radiation .................... 101
6.6 Approximate approach 1: orders of scattering ............... 103
        6.6.1 Adding and doubling methods ............................ 104
6.7 Approximate approach 2: angular distribution functions .... 105
        6.7.1 Two-stream methods and discrete ordinates ............. 106
        6.7.2 Expansion methods and Eddington’s approximation .... 107
6.8 Solution of two-stream and Eddington models ............... 110
        6.8.1 General solution .......................................... 110
        6.8.2 Solution for diffuse incidence .......................... 110
CONTENTS

6.8.3 A comparison of diffusivity, stream, and polynomial approaches 112
6.8.4 The delta-Eddington method for scattering by large particles . 114
6.9 Three dimensional radiative transfer . . . . . . . . . . . . . . . . . . . 114

7 Radiative heating and cooling ........................................... 117

7.1 General concerns ....................................................... 117
7.2 Heating of a cold layer ............................................... 118
7.3 Net heating in warm layers ........................................... 119
    7.3.1 Heating within thick and thin layers ......................... 119
    7.3.2 Heating near boundaries (“Cooling to space”) ............. 120
    7.3.3 Brightness temperature ....................................... 121
7.4 Heterogeneous media and broadband heating ......................... 122
    7.4.1 The Chapman layer ........................................... 123
    7.4.2 Comments on the broadband heating ......................... 124
    7.4.3 Line broadening effects on heating .......................... 125
7.5 Water vapor on Earth ................................................ 127
    7.5.1 Remote sensing of H$_2$O ..................................... 128
    7.5.2 The water vapor feedback and runaway greenhouse .......... 128
7.6 0-D and 1-D radiative-convective equilibria ......................... 128

8 Radiation and atmosphere-ocean dynamics on Earth ................. 132

8.1 Role of radiation in planetary dynamics .......................... 132
8.2 Radiative damping of fluctuations ................................ 133
    8.2.1 Optically thin medium ....................................... 134
    8.2.2 Optically thick medium ..................................... 134

A Quantum mechanics ................................................... 138
<table>
<thead>
<tr>
<th>symbol</th>
<th>quantity</th>
<th>units</th>
</tr>
</thead>
<tbody>
<tr>
<td>( b )</td>
<td>backscatter coefficient</td>
<td></td>
</tr>
<tr>
<td>( B_{\nu} )</td>
<td>Planck emission function</td>
<td></td>
</tr>
<tr>
<td>( \mathbf{B} )</td>
<td>magnetic induction field</td>
<td></td>
</tr>
<tr>
<td>( c )</td>
<td>speed of light</td>
<td></td>
</tr>
<tr>
<td>( c' )</td>
<td>speed of light in vacuum</td>
<td></td>
</tr>
<tr>
<td>( c_{p} )</td>
<td>specific heat</td>
<td></td>
</tr>
<tr>
<td>( d )</td>
<td>diffusivity factor</td>
<td>[1,2]</td>
</tr>
<tr>
<td>( \mathbf{D} )</td>
<td>electric displacement field</td>
<td></td>
</tr>
<tr>
<td>( e_{\nu} )</td>
<td>mass extinction coefficient</td>
<td></td>
</tr>
<tr>
<td>( E_{\nu}, \text{etc.} )</td>
<td>quantum energies</td>
<td></td>
</tr>
<tr>
<td>( E, \mathbf{E}, \mathbf{E}^{f} )</td>
<td>electric field</td>
<td></td>
</tr>
<tr>
<td>( F, F_{\nu} )</td>
<td>flux (irradiance)</td>
<td></td>
</tr>
<tr>
<td>( g )</td>
<td>asymmetry factor</td>
<td>[-1,1]</td>
</tr>
<tr>
<td>( h )</td>
<td>Planck’s constant (Chapter 5)</td>
<td></td>
</tr>
<tr>
<td>( J_{\nu} )</td>
<td>radiation convergence (Chapter 7-8)</td>
<td></td>
</tr>
<tr>
<td>( I, I_{\nu} )</td>
<td>radianc</td>
<td></td>
</tr>
<tr>
<td>( \overline{I} )</td>
<td>zenith intensity</td>
<td></td>
</tr>
<tr>
<td>( J )</td>
<td>total angular momentum quantum number</td>
<td>[0,( \pm \infty )]</td>
</tr>
<tr>
<td>( J_{\nu} )</td>
<td>source function</td>
<td></td>
</tr>
<tr>
<td>( k )</td>
<td>Boltzmann’s constant</td>
<td></td>
</tr>
<tr>
<td>( k_{\nu} )</td>
<td>mass absorption coefficient</td>
<td></td>
</tr>
<tr>
<td>( \mathbf{k} )</td>
<td>wave vector</td>
<td>( \text{cm}^{-1} )</td>
</tr>
<tr>
<td>( l, m )</td>
<td>rotation quantum numbers</td>
<td>([0,\pm \infty] )</td>
</tr>
<tr>
<td>( m )</td>
<td>index of refraction</td>
<td>([1,\infty],[0,\infty] )</td>
</tr>
<tr>
<td>( m_{e} )</td>
<td>electron mass</td>
<td></td>
</tr>
<tr>
<td>( N )</td>
<td>scatterer concentration</td>
<td>( \text{cm}^{-3} )</td>
</tr>
<tr>
<td>( p )</td>
<td>pressure</td>
<td>( \text{Pa} )</td>
</tr>
<tr>
<td>( \mathbf{P} )</td>
<td>angular momentum</td>
<td></td>
</tr>
<tr>
<td>( \mathbf{P}_{\nu} )</td>
<td>dipole moment</td>
<td></td>
</tr>
<tr>
<td>( P_{l}, P_{l}^{m} )</td>
<td>Legendre functions</td>
<td>[-1,1]</td>
</tr>
<tr>
<td>( \mathbf{P} )</td>
<td>polarization</td>
<td></td>
</tr>
<tr>
<td>( q_{e} )</td>
<td>electron charge</td>
<td></td>
</tr>
<tr>
<td>( Q, Q_{C}, \text{etc.} )</td>
<td>heating rate</td>
<td>( \text{K/day} )</td>
</tr>
<tr>
<td>( \mathcal{R} )</td>
<td>reflectance</td>
<td></td>
</tr>
<tr>
<td>( r )</td>
<td>radial coordinate</td>
<td>( \text{cm} )</td>
</tr>
<tr>
<td>( s )</td>
<td>natural coordinate</td>
<td>( \text{cm} )</td>
</tr>
<tr>
<td>( s_{\nu} )</td>
<td>mass scattering coefficient</td>
<td></td>
</tr>
<tr>
<td>( S )</td>
<td>line strength</td>
<td></td>
</tr>
<tr>
<td>( T )</td>
<td>temperature</td>
<td>( \text{K} )</td>
</tr>
<tr>
<td>( T )</td>
<td>transmittance</td>
<td>([0,1] )</td>
</tr>
<tr>
<td>( u )</td>
<td>mass path</td>
<td>( \text{gm cm}^{-2} )</td>
</tr>
<tr>
<td>( v )</td>
<td>vibrational quantum number</td>
<td>([0,\pm \infty] )</td>
</tr>
<tr>
<td>( x )</td>
<td>size parameter</td>
<td>([0,\pm \infty] )</td>
</tr>
<tr>
<td>( x, \mathbf{x} )</td>
<td>position</td>
<td>( \text{cm} )</td>
</tr>
<tr>
<td>( z )</td>
<td>altitude</td>
<td>( \text{cm} )</td>
</tr>
</tbody>
</table>

Table 1: Symbols used. This list includes all symbols used in more than one chapter or section.
<table>
<thead>
<tr>
<th>Symbol</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\alpha)</td>
<td>dipole polarizability</td>
<td>cm(^3)</td>
</tr>
<tr>
<td>(\alpha_N, \alpha_D, \alpha_L)</td>
<td>line width</td>
<td>(Hertz(^{-1}))</td>
</tr>
<tr>
<td>(\epsilon)</td>
<td>dielectric constant</td>
<td>([1,\infty],[0,\infty])</td>
</tr>
<tr>
<td>(\lambda)</td>
<td>wavelength</td>
<td>cm</td>
</tr>
<tr>
<td>(\mu)</td>
<td>(\cos(\theta))</td>
<td>[-1,1]</td>
</tr>
<tr>
<td>(\nu)</td>
<td>frequency</td>
<td>Hertz</td>
</tr>
<tr>
<td>(\tilde{\omega})</td>
<td>single-scattering albedo</td>
<td>[0,1]</td>
</tr>
<tr>
<td>(\Omega, \Omega')</td>
<td>propagation direction</td>
<td>[sphere]</td>
</tr>
<tr>
<td>(\phi)</td>
<td>azimuth angle</td>
<td>[0,(\pi)]</td>
</tr>
<tr>
<td>(\Phi)</td>
<td>electrostatic potential</td>
<td>statvolt cm(^{-1})</td>
</tr>
<tr>
<td>(\Psi)</td>
<td>quantum wavefunction</td>
<td>([-1,1],[-1,1])</td>
</tr>
<tr>
<td>(\rho)</td>
<td>density</td>
<td>g cm(^3)</td>
</tr>
<tr>
<td>(\sigma_s)</td>
<td>absorption cross section</td>
<td>cm(^{-2})</td>
</tr>
<tr>
<td>(\theta)</td>
<td>zenith angle</td>
<td>[0,(\pi)]</td>
</tr>
<tr>
<td>(\Theta)</td>
<td>scattering angle</td>
<td>[0,(\pi)]</td>
</tr>
<tr>
<td>(\tau, \tau_v)</td>
<td>(extinction) optical depth</td>
<td>[0,(+\infty)]</td>
</tr>
<tr>
<td>(\bar{\tau})</td>
<td>optical path</td>
<td>[0,(+\infty)]</td>
</tr>
</tbody>
</table>

Table 2: Symbols used, cont’d
Chapter 1

Introduction

1.1 Radiative transfer in the sciences

First it is worth trying to explain what “radiative transfer” is, in words, and why anyone should care about it. Electromagnetic radiation is the primary means by which energy is transported from one body to another in the universe. Radiation associated with the gravitational force is very weak by comparison, while the other fundamental (“weak” and “strong”) forces decay too rapidly with distance to produce radiation fields. The field of radiative transfer is the quantitative study of electromagnetic energy transport. The subject becomes nontrivial when we are concerned with substances that are partial transmitters of the radiation, which generally includes not only translucent bodies but the surface layers of opaque bodies (for example, Earth’s atmosphere). Thus, radiative transfer shares with thermodynamics the characteristic that “edge” and non-equilibrium effects are the locus of the most interesting, and difficult to explain, behavior.

Radiative energy transports can crop up in places where one might not expect. I was once related the story of a friend who decided that she would enjoy being able to eat some ice cream after a day of skiing. To keep the pint of ice cream cold during the day, she buried it under a few inches of snow at the base of the mountain. Returning after a day of good skiing, she was dismayed to find that the ice cream had melted, even though it was still buried in snow. Part of this has, of course, to do with the fact that ice cream (a polymer) softens considerably before reaching the freezing point of water. But part also has to do with sunlight, which propagates surprising distances through snow. It is possible to dig a snow cave, deep enough (perhaps a meter) for several people to stand overhead and jump up and down without cave occupants being able to hear them, yet the occupant can easily see by the generous sunshine that seeps in through the snow. How this occurs should become evident through study of Chapter 3.

Radiation takes center stage in understanding planetary climates. On average over time, Earth and the other inner planets radiate to space almost exactly the amount of energy they absorb. Climate (e.g. temperature) affects the amount radiated, while atmospheric composition and surface properties affect both the amount radiated and the amount absorbed. Climate is thus determined, in large measure, by atmospheric
and surface composition. The vastly different surface temperatures of the inner planets are partly due to their equally vast differences in atmospheric composition (mostly radiatively inert gases such as nitrogen on Earth, but mostly carbon dioxide on the much hotter Venus, for example). Changes in the composition of an atmosphere due to volcanic eruptions, bombardment by cosmic debris, or fleets of hydrocarbon-powered automobiles must eventually cause climate change to the degree that they affect the atmosphere’s ability to transmit radiation. These arguments can be applied to the energy budget of Earth’s atmosphere, the upper ~100 km of oceans, or the upper ~10 meters of ice sheets and glaciers, all regions of partial transmission of radiation and significant radiative heating or cooling.

Radiative transfer is obviously a key process in stars too—or any system held at sufficiently high temperatures, high illuminations, and/or low densities so that radiative energy transports may compete with others due to conduction, etc. However, we will not treat certain high-energy radiation and scattering phenomena, such as Bremsstrahlung, which figure prominently in plasmas. Readers interested in those topics should consult an astrophysical text.

In the discussion so far, the word “energy” could have been replaced by information, a property whose transport is generally associated with waves of all kinds. The way to learn about the extraterrestrial universe is by studying the arriving radiation. This, in fact, has provided most of the historical motivation for the development of radiative transfer theory and instrumentation. Even interplanetary spacecraft have collected most of their information radiometrically from orbit or flyby, rather than with in-situ probes.

The business of gleaning properties of an object from a distance using waves is called remote sensing. Remote sensing is an industry of its own, whose techniques ultimately derive from various aspects of radiative transfer. Either emitted or scattered radiation can be exploited, and the technique can either be “passive” (making use of naturally emitted or scattered radiation) or “active” (pinging the object with electromagnetic waves and seeing what comes back, radar for example). Although remote sensing began as a ground-based enterprise, the advent of satellites toward the end of the last century has greatly increased the number of vantage points from which such radiation can be gathered. One improvement, made famous by the Hubble Space Telescope, has been the vast improvement in clarity with which the heavens could be observed from above the distorting effects of Earth’s atmosphere. But brand-new capabilities have also emerged in planetary remote sensing. These include space- and ground-based passive infrared and microwave sounding of planetary atmospheric temperature and humidity structure, tracking of winds, microwave through visible determination of cloud distributions and (to a limited extent) composition, visible and ultraviolet study of suspended aerosol particles, and multi-spectral study of surface properties including vegetation type, surface and subsurface temperatures, surface water and ice cover, topography, ocean currents, and the presence of biological activity in seawater. Recently, active remote sensing has been used from satellite for monitoring precipitation, ice-sheet motions, clouds, and aerosols.

Radiative transfer is also increasingly important for understanding the instrumentation (scatterometers, nephelometers, imagers) used to make modern “in-situ” measurements of atmospheric temperature and composition. We will discuss in this text
the principles and application of radar, and many of the basic principles (but not techniques) that underlie spectroscopy and crystallography. Many observing systems (ordinary and infrared cameras, optical fingerprint and retinal scanners, laboratory equipment, etc.) also rely on the same basic principles.

Other remote sensing strategies rely upon sound waves, detailed analysis of satellite motion, etc. These can be used to do many interesting things ranging from seeing through people’s clothing, to mapping the density anomalies in Earth’s interior, to mapping the velocity field deep within the sun. But this text will remain limited to electromagnetic radiation and its role in climate.

1.2 The nature of radiation

Since the work of Maxwell we know that electromagnetic radiation consists of electric and magnetic fields that are no different, except for their time variation, from those generated by static charges and currents. It is possible to think of these fields as “storing” the energy required to assemble (for electrostatic fields) or move (for dynamic fields) the point charges that create them, at an energy density proportional to the field strengths squared. The electromagnetic energy flux carried by a fluctuating set of electric \( E \) and a magnetic \( H \) fields is quantified by the Poynting flux, which in cgs units is

\[
F = \frac{c^*}{4\pi} \mathbf{E} \times \mathbf{H} \tag{1.1}
\]

(\( c^* \) is the speed of light in vacuum). In general, the spatio-temporal behavior of \( E \) and \( H \) is governed by Maxwell’s equations. Radiative transfer is the calculation of these energy transports in practical circumstances where many radiating sources are present. As such, we will be concerned with transport through media of extent much greater than a light wavelength, but sometimes possessing variability on scales comparable to or smaller than a wavelength.

Far from the charges that generate these fields, the field behavior approaches the radiation limit where it simplifies to a plane wave. Then the \( E \) and \( H \) fields oscillate in phase with one another and have equal magnitude (in cgs units). The Poynting flux associated with such a wave, traveling in direction \( \Omega \), is just

\[
F(\Omega) = \frac{c^*}{4\pi} |E(\Omega)|^2 \tag{1.2}
\]

where \( |E(\Omega)|^2 \) is the mean squared field strength associated with energy travel in the direction \( \Omega \). At large distances from a confined distribution of sources in three dimensions, this flux decreases with the square of the distance. Although we are concerned principally with plane waves (where we don’t have to worry about Maxwell’s equations explicitly), we will have to invoke these equations briefly in Chapter 3 where we examine the scattering (redirection) of waves. Since we will not be concerned with transfer through magnetic media, we will not have to distinguish between the macroscopic field \( H \) and the more familiar microscopic field \( B \). Furthermore, we will be able to ignore both of these fields in favor of \( E \), except during our brief excursion into scattering (at which point we will also have to distinguish between the electric fields \( E \) and \( D \), due to the importance of scattering by dielectric particles).
An electromagnetic wave is characterized by wavelength $\lambda$ and frequency $\nu$ (in Hertz), with $c = \lambda/\nu$ the propagation speed of the wave. In vacuum $c = c^*$. We designate the magnitude of the vector $\mathbf{E}$ by the scalar $E$. In physics and engineering the plane wave is traditionally written something like

$$\mathbf{E} = E_0 e^{-i(kx - \omega t)} \mathbf{e}$$

(1.3)

where $x$ is position, $|k| = 2\pi/\lambda$ is the angular wavenumber with $k$ pointing in direction $\Omega$, and $\omega = 2\pi\nu$ is the angular frequency, and $\mathbf{e}$ is a unit vector indicating the direction of the electric field. In this case, $|E(\Omega)|^2$ averaged over a cycle is $E_0^2/2$.

Electromagnetic radiation is quantized in photons which may be thought of as little wave packets which, when added together, make up the plane wave described above but which, according to the founding principle of quantum physics, are indivisible. Each photon carries energy

$$E_{\text{phot}} = h\nu,$$

(1.4)

and momentum

$$p_{\text{phot}} = h\nu/c^*,$$

(1.5)

where $h$ is Planck’s constant. According to the particle description of radiation, the Poynting flux is simply carried by a stream of photons traveling in a particular direction at velocity $c$. 

Figure 1.1: Schematic depiction of electromagnetic vector relationships for the special case of a propagating plane wave.
1.3. QUANTIFYING RADIATION

In the radiative limit of behavior as addressed in this text, the particle (photon) description of light is especially intuitive. Any location in space can contain photons traveling in all directions and at all frequencies; under normal conditions, these photon streams are like ships passing in the night that have essentially no interaction (except indirectly through their mutual interaction with matter). From the wave point of view, this non-interaction is a straightforward result of the linearity of Maxwell’s equations.

The foregoing introduction is intentionally very brief. Readers who do not have some past experience with oscillators and waves, or are not already familiar with the above material, should consult a basic physics text such as Halliday and Resnick so as to be prepared for Chapters 3 and 4.

1.3 Quantifying radiation

1.3.1 Note on units and symbols

The description of radiative transfer is plagued by an overabundance of defined quantities. Many of these are essentially redundant but have cropped up in different fields (spectroscopy, planetary science, basic physics, etc.). While a subset may be adequate to understand radiation conceptually, some that are not strictly necessary but sometimes encountered will be introduced for the edification of the reader. The overabundance typically exceeds the capacity of the Arabic and Greek alphabets to represent them all without extensive resort to sub- and superscripts; custom also demands at times that the same symbol be used to represent more than one quantity. Here, we have tried to avoid confusing duplication of symbols as much as possible; see the appendix for a list.

Units present another dilemma. We give all equations in cgs (Gaussian) units, the preferred (and generally simpler) system in use by physicists for describing electromagnetic radiation. However, in nearly all practical work in engineering and atmospheric sciences, quantities are expressed in MKS units. Thus, we give all quantities in MKS units. The equations that appear differently in the two systems are those (confined to Chapter 3) involving electromagnetic field strengths or charge densities\(^1\). Otherwise, the equations don’t change, just the numerical values of the physical constants that appear in them. It will turn out that none of the equations that are of practical use to us in calculating radiative transfer will contain field or charge magnitudes explicitly. This enables us to enjoy the best of both worlds, as long as we use the intermediate derivations only to understand what is going on and avoid plugging MKS values into them for results.

Tradition imparts little importance, and no universally accepted symbol, on the wavenumber \(1/\lambda = \nu/c\). However, spectroscopists have found the in-vacuo wavenumber \(\nu/c^*\) to be of great use. They have not played nicely and have stolen the symbol \(\nu\), normally reserved for frequency, to represent in-vacuo wavenumber in their work. These two concepts are equivalent, since they are related by a constant multiple. But since in most commonly used systems of units \(c^* \neq 1\), the notational ambiguity can lead to confusion and forces outsiders to choose sides. In this text we will retain the symbol \(\nu\) to represent frequency in Hertz. By the same token as with MKS units,

\(^1\)For conversion to MKS units see, e.g., the appendix of Jackson (1975).
however, we will quote numerical values of \( \nu \) as in-vacuo wavenumbers in inverse centimeters, e.g., 1000 cm\(^{-1} \) rather than \( 2.998 \times 10^{13} \) Hz, in accordance with widespread custom. Radiation is about equally likely, depending on whom you’re talking to, to be identified by its wavelength or by wavenumber, but not normally by frequency. Wavenumber and wavelength are related by the useful rule of thumb that the product of \( \lambda \) (in microns) and \( \nu \) (in cm\(^{-1} \)) is about 10,000 in vacuum.

While these waffling practices may seem confusing, it is worth exposing the reader to these issues gently here rather than cleaning things up too nicely. We assume most readers cannot be insulated from these issues forever!

### 1.3.2 Quantifying intensity of radiation

As noted earlier, a small volume of space can contain photons traveling in any direction \( \Omega \) and at any frequency \( \nu \). We define the density of power so carried, per unit solid angle, unit wavenumber, and unit area (perpendicular to \( \Omega \)), as the **monochromatic or spectral radiance**, denoted \( I_\nu(x, \Omega) \) (see Table 1.1). This quantity has traditionally been called (specific) **intensity** in the planetary and astrophysical sciences, although the term radiance is less ambiguous and has been widely adopted within the Earth science community. The term “brightness” is sometimes used in astronomy. We will use “radiance” in this text, though still denoting it by \( I \) rather than the \( J \) or \( L \) occasionally encountered in meteorological literature.

The subscript \( \nu \) or \( \lambda \) on radiance and other quantities is important and indicates that they are **densities** rather than proper functions. Specifically, \( I_\nu \) and \( I_\lambda \) have different units and numerical values, since one is power per unit wavenumber while the other is per unit wavelength. In fact it is the integral of either of the quantities from one specific position in the spectrum to another that is physically significant and measurable; it is the same regardless of which one is used.

If either \( I_\nu \) or \( I_\lambda \) is integrated over the whole spectrum, the result is the **total radiance** \( I \):

\[
I = \int_0^\infty I_\nu d\nu = \int_0^\infty I_\lambda d\lambda
\]

In practice, integrals are often taken over a finite range or frequency **band**, rather than the whole spectrum. Usually the band is defined by a limited range over which a particular absorbing constituent is active (e.g., the 15-micron CO\(_2\) band), or a detector is sensitive (e.g., a particular channel on a satellite imager). In the latter case, the integral is sometimes non-uniformly weighted within the band range according to the instrument sensitivity. A subscript can be used to identify the particular band over which a band radiance has been calculated (one example would be the channel number of a detector).

If we instead integrate \( I_\nu \) over a hemisphere of propagation directions \( \Omega \) going one way through a plane surface, weighted by the component of \( \Omega \) perpendicular to that surface, we obtain the one-way **monochromatic or spectral irradiance** through the surface:

\[
F_\nu(x) = \int_{\Omega: \alpha > 0} \Omega \cdot \mathbf{n} I_\nu(x, \Omega) d\Omega
\] (1.6)
1.3. QUANTIFYING RADIATION

In (1.7), the first factor in parentheses is the normal component of a flux contribution from zenith $\theta$ (Figure 1.2), while the second factor is $d\Omega$ expressed in spherical coordinates, i.e., the amount of solid angle per unit zenith and azimuth that exists at zenith $\theta$ (Figure 1.3). The irradiance is more commonly called the flux (and has occasionally been called the exitance if it is the flux leaving an emitting body such as a star). Note that any defined surface will have two fluxes through it, one from each direction, as well as a net flux which is the difference between the two. The net flux is, in reality, a vector. However, opposing streams of radiation typically must be considered separately since they usually have sources of different spectral character. Usually there is a particular orientation of interest (e.g. vertical), dictated by the physical system, such that fluxes in two particular directions (up, down) are of interest. Note that if $I$ is isotropic it can be removed from the integral, with the result that

$$F_\nu = \pi I_\nu \quad \text{(isotropic).} \quad (1.8)$$

A quantity of interest somewhat related to the above is the mean radiance or mean intensity, defined by:

$$I_{0,\nu} = \frac{1}{4\pi} \int I_\nu d\Omega. \quad (1.9)$$

The mean radiance is the zeroth moment of the radiation field; higher moments will be invoked when we discuss solution of the radiative transfer equation. In atmospheric chemistry the mean radiance is referred to as the actinic flux, which is commonly expressed in terms of photon number rather than energy. The actinic flux directly determines the rates of photochemical reactions.

Finally, if $F_\nu$ (or $F_\lambda$) is integrated over the spectrum one obtains the total flux $F$. The quantities and units are shown in Table 1.1. Integration over finite pieces of the spectrum again yields broadband or narrowband fluxes depending on the circumstances. While the total flux determines the transport of energy, the monochromatic and band quantities are more amenable both to theory and observation. Total fluxes must usually be cobbled together from multiple bands calculated or measured individually.

For example, the total flux of solar radiation at a distance of one astronomical unit (AU) from the sun currently averages about 1366 W/m$^2$. This value is called the “solar constant”, although variations in solar output have probably been sufficient to contribute to some of the historical climate changes on Earth over time scales ranging from $10^1$ to $10^6$ years. The flux impinging on (i.e. normal to) the Earth’s surface is equal to the solar constant only at the subsolar location, and is less elsewhere (zero on the night side); when averaged over the planet it is one fourth of the solar constant or 342 W/m$^2$. This illustrates that care must be taken in discussing fluxes as scalars. In discussing planetary or stellar atmospheres, fluxes are most often defined with respect to planes tangent to the planetary surface or stellar boundary rather than perpendicular to the propagation direction of any incident radiation. When measuring fluxes, it can be very important to orient the measuring device appropriately.

Though flux is a simple concept, many students have trouble with the concept of radiance. Since it is defined in inverse steradians, one is tempted to think of $I$ as
quantifying the amount of energy in a narrow cone, but this is not helpful. Where should the vertex of such a cone be located? One should properly think of the cone in the limit where the vertex is infinitely far away, in which case one has a thin cylinder (or as Chandrasekhar appropriately called it a “pencil”) of radiation. This may be seen by defining radiance more formally. First consider flux, an area density that may be formally defined in the following way: divide the flux surface into a grid of \( N \) elements, compute for each element \( i \) the ratio of number of photons per unit time that cross each element to its area, then take limits as the element size goes to zero (\( N \to \infty \)). Radiance may be defined similarly, except that for each element \( i \) one must also consider \( M \) discrete directions \( j \) in which a photon stream could be emitted. A photon stream carries the same amount of power no matter how far away it is intercepted. At large \( M \) the number of streams a receiver of given area will receive, and hence the flux, decreases as the inverse square of distance from the source. Similarly, a bright body

<table>
<thead>
<tr>
<th>Quantity</th>
<th>symbol</th>
<th>units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monochromatic radiance</td>
<td>( I_\nu )</td>
<td>( W \text{ ster}^{-1} \text{ m}^{-2} \text{ cm} ) (( W \text{ ster}^{-1} \text{ m}^{-2} \mu \text{m}^{-1} ))</td>
</tr>
<tr>
<td>Total radiance</td>
<td>( I )</td>
<td>( W \text{ ster}^{-1} \text{ m}^{-2} )</td>
</tr>
<tr>
<td>Monochromatic flux/irradiance</td>
<td>( F_\nu )</td>
<td>( W \text{ m}^{-2} \text{ cm} ) (( W \text{ m}^{-2} \mu \text{m}^{-1} ))</td>
</tr>
<tr>
<td>Total flux/irradiance</td>
<td>( F )</td>
<td>( W \text{ m}^{-2} )</td>
</tr>
</tbody>
</table>

Table 1.1: List of variables quantifying the energy carried by radiation. The variables \( I_\nu \) and \( I_\lambda \) are related by: \( \nu I_\nu = \lambda I_\lambda \), where \( \nu \) is in cm\(^{-1} \).
1.4 Interaction between radiation and matter

Matter is composed naturally of positively and negatively charged particles. Though matter is typically macroscopically neutral, at the subatomic level the charges are separated. Electrons, for example, inhabit a quantum “cloud” of probability density that envelops the nucleus of the nearest atom (classically, they “orbit” the nucleus), or nearest two nuclei if the electron participates in a covalent bond. Ions are present in small quantities in most gases, and are dominant in plasmas such as stellar atmospheres or the inside of a lightning bolt.

1.4.1 Scattering

Scattering of radiation occurs because individual charges are accelerated away from their equilibrium positions by the incident radiation. In so doing, they radiate their own electromagnetic waves in various directions. Any charges, whether bound in atoms, molecules, or free, will participate in scattering. Downstream along the path

Figure 1.3:
of the original wave, the new waves typically interfere destructively with the original, reducing its energy. If this destruction is total then the incident wave has been reflected. Scattering is usually modeled as a classical phenomenon, and calculating it accurately requires a wave-like treatment of the radiation. Far from the scattering molecule we may simply think of a photon that has bounced off of it into a new direction. Due to the wavelike nature of radiation, however, the directions into which scattered energy propagates can be very sensitive to the spatial arrangement of the scattering molecules.

From either the particle or wave perspective, we expect scattered radiation to have the same frequency as the incident radiation. This is called conservative (or elastic) scattering since the photon energy is conserved according to (1.4). Sometimes the term “coherent” is employed, but we will reserve that term for a different concept here. When scattering is conservative, we can write photon conservation (radiative transfer) equations which are completely separable in frequency. In practice, nonconservative scattering can be treated as a combination of absorption at one frequency and simultaneous emission at another frequency. Nonconservative scattering is a quantum phenomenon. It is typically relatively weak, and often can be safely ignored (and will be discussed further in Chapter 5).

Scattered radiation can go in any direction, though usually not all directions will be equally likely. Two limiting cases of interest are specular reflection where all light scatters from a sharp interface between two media such that the angle of incidence equals the angle of reflection, and isotropic scattering where the scattered radiance is equal in all directions regardless of incidence direction.

### 1.4.2 Absorption and emission

In contrast to conservative scattering, absorption of radiation is fundamentally a quantum physics phenomenon. Even the absorption of energy that occurs according to Ohm’s law in a resistive conductor carrying electric current, depends ultimately on the quantum behavior of electrons as they cascade through the substance under an applied electromotive force. Absorption of radiation occurs because atoms and molecules find themselves, at any given time, in one of many quantum states of differing but discrete energies. A transition from one state to another permits the absorption or emission of a photon at precisely the energy corresponding to the difference between that of the initial and final states. Emission and absorption occur due to the same quantum transitions, though generally not at the same rate since most state transitions are non-radiative. Emission rates for a given transition will be proportional to population of the initial state. According to statistical mechanics, when local thermodynamic equilibrium (LTE) holds in the material, the state populations will obey a Boltzmann distribution proportional to

$$e^{-E/kT}$$  

where $E$ is the state energy, $T$ is temperature, and $k$ is Boltzmann’s constant. Absorption rates for a given transition and mass of absorber will, unlike emission, be proportional to the number of photons passing by that have the correct energy. At the risk of stating the obvious, all photons originate by being emitted in a quantum transition, and ultimately perish by being absorbed in one. Scattering, if it occurs, only
1.5 Blackbody radiation

All substances emit radiation. This process, or its result, can generally be called simply emission; if the radiation is in equilibrium with the emitting substance it is called blackbody radiation, although this term is sometimes used loosely (i.e., incorrectly) to refer to any emission. At terrestrial temperatures, the emission is primarily infrared and is often called “thermal emission.” The usual way textbooks introduce blackbody radiation is by invoking a hollow cavity with a tiny hole in it as a perfect black body. Though such a contraption may technically be the closest practical approximation to a true “black body,” it appears to be a rather artificial model for describing a ubiquitous natural process. It also fails to convey some key relationships between blackbody radiation and atomic properties. So here’s another way of viewing blackbody (or equilibrium) radiation.

Suppose we have a large ball of hot gas with no radiation anywhere. Since the gas is hot, many of its atoms (or molecules) must exist in excited states at any given time. Suppose there are many possible states with attainable energies. Excited states have a certain probability per unit time of spontaneous decay to a lower state. When this happens a photon is emitted. The gas will not run out of excited molecules, since collisions are constantly knocking individual molecules into new states. Thus, the initially dark ball of gas will not remain dark but will begin to lighten up. If desired we may think of the gas as being surrounded by perfectly reflecting walls, or we may think of it as being of such great extent that it takes much longer for photons to escape than it does for them to be reabsorbed. In either case the photon density will build up until absorption catches up with emission and an equilibrium is reached.

The emission rate of photons (per unit volume) by a given state transition \( n \rightarrow l \) will be proportional to the absorber density, to the transition likelihood, and to the population of the upper state. For LTE situations, this is just

\[
\text{emission} \sim \rho A_{t, u} e^{-E_u / kT}.
\]

Meanwhile, photons that have already been emitted can be absorbed. The absorption rate by the reverse quantum transition \( l \rightarrow u \) (hence, involving photons of the same energy and frequency) will be proportional to absorber density, to the transition likelihood, to photon density, and to the population of the lower state:

\[
\text{absorption} \sim \rho A_{t, u} U e^{-E_l / kT},
\]

where \( U \) is a measure of the density of photons in the gas (this measure is not necessarily frequency-independent). The ratio of absorption to emission rates for a particular transition is then

\[
\frac{\text{absorption}}{\text{emission}} \sim e^{\Delta E / kT} U \quad (1.11)
\]

\( \Delta E \) is the energy jump of the transition, which determines the photon frequency according to (1.4). For any given transition, \( U \) will continue to rise until this ratio equals...
unity, at which point equilibrium is achieved. For larger values of $T$, this equilibrium requires larger $U$; for higher energy transitions, smaller $U$.

The fully equilibrated photon density is isotropic blackbody radiation. Note the remarkable lack of dependence on the absorber density or, more importantly, on the likelihood $A_{\lambda,u}$ of the state transition—both influences divide out when we compute the absorption to emission ratio. $U$ is determined only by $\Delta E$ and $T$! All that is required from the substance is that, by any means at all, it can emit and absorb at that frequency. Even the feeblest interaction will do, it will just take a long time for the equilibrium to be reached at the corresponding frequencies. The substance does not need to be a gas, of course; anything will eventually form an internal, equilibrated photon field according only to its temperature.

The correct formula for blackbody radiance is

$$B_\nu(T) = \frac{2\nu^3}{c^2} \frac{1}{\exp\left(\frac{h\nu}{kT}\right) - 1}. \quad (1.12)$$

This formula was obtained by Max Planck in a stroke of genius that ushered in the era of quantum physics, and is known as the Planck formula. It differs in detail from that deduced during the discussion (the “$-1$” in the denominator) because we overlooked stimulated emission (see Problem 6); our argument does not predict the leading factor. Our result does possess the correct behavior (at fixed $\nu$) in the low-temperature limit. The $\nu^3$ dependence in the correct formula reflects the classical notion that charges orbiting at higher frequency accelerate more strongly, and therefore radiate energy more efficiently. Prior to Planck’s discovery, a result correct in the low $T$/high $\nu$ limit was obtained by Wein and is known as Wein’s formula. The behavior in the high-$T$, low-$\nu$ limit,

$$B_\nu(T) = \frac{2\nu^2}{c^2} kT \quad (1.13)$$

was also found, from classical arguments, and is now known as the Rayleigh-Jeans law. These limiting expressions are primarily of historical interest but can be useful simplifications for some calculations.

Two celebrated and useful results may be obtained from Planck’s formula (although both were already known empirically when Planck’s theory came along). The first is Wein’s displacement law, which relates the frequency of peak emission to the temperature. By setting $\partial B_\nu/\partial \nu = 0$, we obtain

$$\nu_{\text{peak}} = 5.88 \times 10^{10} T. \quad (1.14)$$

Figure 1.4 shows how emission varies with temperature. The solar photosphere is about 25 times hotter than Earth’s troposphere, making the peak solar wavenumber higher by an equivalent factor (about 20,000 cm$^{-1}$, or 500 nm for the sun compared to less than 1000 cm$^{-1}$, or $> 10\mu$m for Earth.) Though radiation at all wavelengths increases monotonically with temperature, the peak shifts because radiation increases

---

2 Stimulated emission is the decay to a lower state that is initiated by an incident photon at the same frequency, resulting in two outgoing photons traveling in the same direction. This becomes important at higher temperatures or low transition energies; it is also the key to the functioning of lasers.
1.6. EMISSIVITY

Figure 1.4: Spectral variation of Planck emission at several temperatures.

much faster for photon energies near or above the thermal energy (compare (1.11) and (1.13)) where the initial (upper) states are hard to attain and their populations are very temperature sensitive.

The second important result is the Stefan-Boltzmann law

\[ F_{BB} = \pi \int B_\nu(T) d\nu = \sigma T^4 \]

(1.15)

for the total flux emitted, where \( \sigma \) is the Stefan-Boltzmann constant. In writing the first equality, we invoke the isotropy of blackbody radiance. The Stefan-Boltzmann law is fundamental to the energetics of bodies that cool themselves by radiation, for example, planets.

1.6 Emissivity

What we see as emission in nature is the photons that leak out through the boundary of a material. But to establish the blackbody equilibrium, we had to stipulate that no photons escape. Not surprisingly, then, observed emission spectra never quite match the blackbody ideal. For isothermal bodies, the emittance is always less than the Planck value (1.12); the fraction actually emitted is defined as the (spectral) emissivity,

\[ e(\nu) \equiv \frac{I_\nu}{B_\nu} \leq 1 \]

(1.16)

Spectral emissivity depends on wavelength, and is a proper function (the ratio of two densities). It is an extrinsic property of a body and the direction of the radiation, not an intrinsic property of a substance (for example, a tenuous, elongated blob of material has an emissivity that increases either as the cloud becomes more massive or as the radiating direction becomes closer to that of its long dimension). Radiation emitted by
opaque bodies is usually nearly isotropic, although emissivity can vary with angle in ways determined by the microscopic or macroscopic texture of the body surface.

Bodies may also be characterized by an absorptivity $A(\nu)$, which quantifies the fraction of incident energy that the body absorbs; a reflectance $R(\nu)$; and a transmittance $T(\nu)$, where energy conservation dictates that

$$A + R + T = 1. \quad (1.17)$$

If a body is massive enough to be completely opaque, then $T = 0$ and $A = 1 - R$. For such bodies it is easy to show that,

$$A = \varepsilon, \quad (1.18)$$

known as Kirchoff’s law of radiation.

Kirchoff’s law may be demonstrated by the following thought experiment. Consider two opaque slabs facing each other, separated by vacuum so that they can exchange energy only through radiation. Each slab emits and absorbs radiation as per the above restrictions. For simplicity assume that one of the slabs is a perfect blackbody, the other has wavelength-independent emissivity $\varepsilon$ and absorptivity $a$, and that each slab is initially at the same temperature. The “grey” slab will then emit at the rate $e\sigma T^4$, and all of this emission will be absorbed by the black slab, transferring energy from the grey to the black slab. Meanwhile the black slab will emit $\varepsilon\sigma T^4$, a fraction $a$ of which will be absorbed by the grey slab—the remaining fraction will be reflected (since the slabs are non-transmissive) and reabsorbed by the black slab. Only if $a = \varepsilon$ will the net energy flow be zero; otherwise, one slab will heat up while the other cools off. Since the latter outcome would violate the second law of thermodynamics, which states that energy cannot flow spontaneously from a cold to a hot body, we must have $a = \varepsilon$. We can go farther and show that $a(\nu) = \varepsilon(\nu)$ for spectrally variable materials by repeating the argument with a nonabsorbing filter between the two slabs that transmits only frequency $\nu$ and reflects all other radiation.

Thus, thermodynamics ultimately ensures that objects which are good emitters must also be good absorbers. Imperfect black bodies reflect some portion of incident light but, by the same token, intercept radiation trying to leave the body and scatter it back in, reducing the emission by a like percentage.

If the directional emissivity of an opaque surface varies with emission direction, its directional reflectance with respect to light incident from the corresponding direction will vary oppositely, another manifestation of Kirchoff’s law. This follows from our earlier thought experiment if we insert an imaginary filter that reflects all radiation except that traveling in a particular direction. The important question of what direction(s) radiation is scattered into is not constrained by Kirchoff’s law.

The fact that Kirchoff’s law was derived by invoking a closed, equilibrated system is important not to forget. In this special case, emission will equal absorption at each frequency. However, emission depends on temperature whereas absorption depends on incident flux. If the temperature of the body differs from that of the source that irradiates it, then equality of absorptivity and emissivity does not imply equality of absorption and emission. The body is not in equilibrium with the radiation or its source. The reader may be worried that if absorption and emission are different, this implies more state transitions in one direction than in the other direction. How can this be?
Blackbody equilibrium depends on most transitions being thermal rather than radiative, so that an imbalance in the direction of the radiative transitions is lost in the shuffle. These matters are discussed in more detail in chapter 5.

Exercises

1.1 Starting from $B_\nu$, (a) derive Planck’s formula as a function of wavelength $B_\lambda$, and (b) find the equivalent Wein’s law for the peak in this distribution to three decimal places accuracy. Hint for (b): define a new nondimensional parameter $\alpha$ that includes the product $\lambda T$, and then find an expression of the form $\alpha = F(\alpha)$ where $F(\alpha)$ depends very weakly on $\alpha$, so that inserting an approximate value on the RHS and iterating will very rapidly converge).

1.2 Calculate the sensitivity of emitted power to temperature, in percent per K, of radiation emitted by a 300 K black body in the (a) microwave and (b) far infrared parts of the spectrum (choose a representative wavelength for each to calculate specific values).

1.3 For a full moon, calculate the total (a) radiance observed from Earth as a function of normalized distance $z$ from the center of the lunar disk, and (b) moonlight flux received at a particular location on Earth as a function of local lunar zenith angle $\Theta_L$. The lunar albedo is 0.12. Assume the radiances scattered from a given point on the lunar surface are isotropic, and that the moon occupies a small solid angle viewed from Earth.

1.4 Calculate the downward total flux received by someone standing in the middle of a round, uncovered stadium of radius 100 m and height 70 m at night, if the stadium seats and walls are at temperature $T_S$ and behave as black bodies, and the downward flux far away from the stadium under open sky is given by $\sigma T^4$. Assume the open sky radiation is hemispherically isotropic.

1.5 Determine the emissivity of a material that consists of a perfectly black substrate, covered by a thin non-absorbing layer of dust that reflects 20% of radiation incident upon it, if the dust:

1. is non-absorbing from either direction;

2. also absorbs 20% of incident radiation, and is in good thermal contact with the substrate;

3. absorbs 20% and cannot conduct heat to/from the substrate, so its temperature $T_D$ differs from that, $T_S$, of the substrate. In this case, assume no illumination of the material from outside.

Assume these properties are wavelength-independent.

1.6 Emitted radiation is actually of two types, spontaneous and stimulated. The equi-
librium radiation condition can be written as:

\[ N_l B_{l\mu} I_{0,\mu} = N_u A_{u l} + N_u B_{u l} I_{0,\mu} \]

where \( N_u, N_l \) are the upper and lower state populations, \( B \) is the absorption or stimulated emission probability and \( A \) is the decay probability (the Einstein A and B coefficients). Ignoring degenerate states for simplicity, the Einstein relations are

\[ \begin{align*}
B_{l u} &= B_{u l} \\
A_{u l} &= \frac{2h\nu^3}{c^2} B_{u l}
\end{align*} \]

and state populations are related by

\[ \frac{N_l}{N_u} = \exp\left(\frac{\Delta E}{kT}\right). \]

Using these equations, derive the Planck formula. (Accounting for degenerate states introduces additional weighting factors into these equations, but these weighting factors cancel out so that Planck’s formula is unaffected, just as it is independent of \( B_{u l} \).)

1.7 Calculate the angular distribution of radiance reflecting from a perfect spherical mirror, where all reflection is specular and the sphere is uniformly illuminated by irradiation \( F \) from a distant point source (ignore diffraction).

1.8 During its boost phase, a rocket produces a trailing plume of very hot gases. Suppose that the rocket and its plume each have a projected area \( A \) viewed from the side, and that the plume can be treated as a 1000K black body. A radiation detector of collection area 1 m\(^2\) is aimed at the rocket, and sits at a distance of 1000 km.

1. How much power is received by the detector due to emission from the plume?

2. If a source of radiance \( I_0 \) located next to the detector is aimed at the rocket, how much reflected power (from the rocket) is received? Treat the rocket as an isotropic scatterer of radius \( \sqrt{A/\pi} \).

3. If the source radiates 100 W into one steradian at \( \nu = 10 \) cm\(^{-1}\), \( A = 9 \) m\(^2\), and the distance is 3 km, how many reflected photons per second enter the detector?
Chapter 2

The radiative transfer equation

For radiation impinging on some interface, we wrote conservation of radiant energy (1.17) by noting that incident energy must be equal the total of what is scattered, absorbed, or reflected at the interface (this assumes that photons do not pile up in one location, a safe assumption in practically all situations). The same holds in continuously varying media, provided that we (a) introduce calculus to handle continuously varying radiation fields and (b) take into account the fact that energy can be propagating in all directions at once. This leads us to the radiative transfer equation.

2.1 Processes affecting radiative intensity

Consider a beam of radiation traveling in a direction $\Omega = \{\theta, \phi\}$ through an infinitesimal control volume of length $ds$ in the propagation direction (c.f. Figure 2.1), having area $dA$ normal to this direction. The beam intensity (radiance) $I_\nu$ can be changed by four processes. In this text we will always assume that the radiation field is steady in time, i.e., that photons traverse the medium much faster than the medium’s radiative properties change, and that photon density changes are negligible compared to the product of a typical flux and a time scale of interest. Thus, no distinction need be made between the material derivative $dI/ds$ measured following along with the photons and the instantaneous spatial gradient $\partial I / \partial s$.

2.1.1 Absorption

Lambert’s (a.k.a. Beer’s, Bouguer’s) law states that the probability of absorption of a particular photon by a particular amount of absorber in its path is independent of how many other photons there are or what they are doing. This statement is highly accurate under most circumstances, as discussed further in Chapter 5. The law implies that

$$dI^{(a)}_\nu = -k_\nu \rho I_\nu ds,$$  \hspace{1cm} (2.1)

where $ds$ is an infinitesimal displacement in the direction of travel, $\rho$ is the density of the absorbing substance, and $k_\nu$ is called the mass absorption coefficient or simply
the absorption coefficient. This quantity has units of (area/mass) and depends on the substance, its thermodynamic state, and the wavelength of the radiation. If more than one substance absorbs significantly at $\nu$, then we must sum over substances $i$,

$$dI_{\nu}^{[\text{abs}]} = -\sum_{i} k_{\nu,i} \rho_i I_{\nu} ds,$$  \hspace{1cm} (2.2)

although this complication will not be considered until chapter 7.

A volume absorption coefficient, $k_{\nu} \equiv \bar{k}_{\nu} \rho$, is often defined. A third and quite useful coefficient is the (particle) absorption cross section,

$$\sigma_{\alpha \nu} \equiv k_{\nu} M$$  \hspace{1cm} (2.3)

where $M$ is the mass per particle of the absorber. The cross section has units of area. Lambert’s law written in terms of cross section is

$$dI_{\nu}^{\text{Lam}} = -\sigma_{\alpha \nu} N I_{\nu} ds$$

with $N$ the number of absorbing particles per unit volume. The reader should beware that in spite of the conceptual simplicity here, much semantic confusion surrounds the various expressions of absorption coefficient in the literature, and not all authors mean the same thing by “absorption coefficient.” To avoid confusion we will use only the mass and particle coefficients $k_{\nu}$ and $\sigma_{\alpha \nu}$ until multi-component media are considered in Chapter 7.

### 2.1.2 Emission

The beam is augmented by emission into the propagation direction. We assume here that radiative substances are in LTE. Thus, emission may be stated

$$dI_{\nu}^{\text{emit}} = k_{\nu} \rho B_{\nu}(T) ds,$$  \hspace{1cm} (2.4)

where $B_{\nu}(T)$ is the (isotropic) Planck radiant emission at frequency $\nu$. The reappearance of $k_{\nu}$ here is due to Kirchoff’s law expressed earlier, as applied to an infinitesimal quantity of absorber. Comparing (2.1) and (2.4), we see that if the intensity of a beam of radiation initially matches the Planck intensity of the medium through which it passes,
its subsequent intensity will be unaffected by the net action of absorption and emission, as required by Kirchhoff’s law. Non-LTE effects are discussed in chapter 5, and can (in principle) be handled using appropriately modified source functions in place of the Planck function.

### 2.1.3 Scattering out of $\Omega$

Photons can be scattered out of the beam into other directions. This fate of a photon also obeys Lambert’s law and therefore leads to

$$dI^\text{cat}_\nu = -s_\nu \rho I_\nu ds,$$  \hspace{1cm} (2.5)

where we now define a (mass) scattering coefficient $s_\nu$. A volume scattering coefficient $s_{\nu V}$ and scattering cross section $\sigma_{\nu V}$ may be defined analogous to those for absorption; the latter will be used most often here. Apologies for introducing $s$ and $s_\nu$ for entirely different quantities, but we will not be using either of them for long.

### 2.1.4 Scattering into $\Omega$

Photons traveling in arbitrary directions can be scattered into the direction $\Omega$. This can happen to any photon traveling through the volume $dA ds$ in any previous direction $\Omega'$. We write

$$dI^\text{cat+}_\nu(\Omega) = \frac{s_\nu \rho}{4\pi} \int I_\nu(\Omega') P_\nu(\Omega, \Omega') d\Omega' ds,$$  \hspace{1cm} (2.6)

where the total contribution must be integrated over all previous directions. The factor $P_\nu(\Omega, \Omega')$ under the integral is the probability density that a photon that scatters from direction $\Omega'$ will scatter into the direction $\Omega$ rather than some other direction. This function, which has units of ster$^{-1}$, is called the phase function of the scattering medium. Note that the probability the photon will scatter at all is contained in $\rho$ and is assumed independent of the initial direction of the photon. Thus, the phase function must obey the integral constraint over the sphere

$$\int P(\Omega, \Omega') d\Omega' = 4\pi$$  \hspace{1cm} (2.7)

for any $\Omega$, accounting for the $4\pi$ term in (2.6). Isotropic scattering is represented as $P_\nu(\Omega, \Omega') = 1$. Some authors use a phase function that normalizes to $4\pi$ times the single-scattering albedo (defined below).

### 2.2 Full Equation

Combining these terms yields the radiative transfer equation (RTE) written in natural ($s$) coordinates,

$$\frac{1}{\rho} \frac{dI_\nu}{ds} = -(k_\nu + s_\nu) I_\nu + k_\nu B_\nu(T) + \frac{s_\nu}{4\pi} \int I_\nu(\Omega') P_\nu(\Omega, \Omega') d\Omega'. \hspace{1cm} (2.8)$$
CHAPTER 2. THE RADIATIVE TRANSFER EQUATION

Noting that photons are lost due to absorption and scattering out of the beam, we label the loss due to either of these mechanisms as extinction and define a mass extinction coefficient

\[ e_\nu \equiv k_\nu + s_\nu. \]

Volume extinctions and particle extinction cross sections can be similarly defined by adding those due to absorption and scattering. If we also define the single-scattering albedo

\[ \bar{\omega}_\nu \equiv \frac{s_\nu}{e_\nu}, \]

then the RTE can be transformed into a more convenient form:

\[
\frac{1}{\rho e_\nu} \frac{dI_\nu}{ds} + I_\nu = (1 - \bar{\omega}_\nu)B_\nu(T) + \frac{\bar{\omega}_\nu}{4\pi} \int I_\nu(\Omega')P_\nu(\Omega,\Omega')d\Omega'.
\]

The special case where no absorption occurs and \( \bar{\omega}_\nu = 1 \) is called conservative scattering. Finally, we define the dimensionless optical path

\[ \tau_\nu \equiv \int e_\nu \rho ds \]

which enables us to rewrite the RTE once more as

\[
\frac{dI_\nu}{d\tau_\nu} + I_\nu = (1 - \bar{\omega}_\nu)B_\nu(T) + \frac{\bar{\omega}_\nu}{4\pi} \int I_\nu(\Omega')P_\nu(\Omega,\Omega')d\Omega'.
\]

By writing the RTE in terms of the dimensionless quantity \( \tau \), we avoid confusions concerning the various forms of absorption coefficient. Note that \( \tau \) is technically an extinction optical path although the extinction characteristic is implied (we could also define an absorption optical path, but such a quantity is not particularly useful). Equation (2.12) is often written

\[
\frac{dI_\nu}{d\tau_\nu} + I_\nu = J_\nu
\]

where the total source due to either emission or scattering is called the source function, denoted \( J_\nu \). In general, \( J_\nu \) depends on \( T \) and the distribution of ambient radiation passing through the control volume. This presents one of the fundamental dilemmas of radiative transfer: we want to know the radiance field \( I \), but the source function depends on \( I \) leading to a chicken-and-egg problem. We confront this problem in Chapter 6.

If there is no scattering (\( \bar{\omega}_\nu = 0 \)) then

\[
\frac{dI_\nu}{d\tau_\nu} + I_\nu = B_\nu(T).
\]

The RTE in this case is known as Schwarzschild’s equation. This no-scattering case is much simpler because the radiance field is determined by the temperature field and boundary conditions, avoiding the chicken-egg problem. Situations where this equation is useful include the calculation of infrared radiative transfer in planetary atmospheres, and of stellar radiation.
2.3. **GENERAL SOLUTION**

2.2.1 Optical depth

Typically, an axis exists that is of special importance (normally, the vertical). A measure of opacity changes along this axis known as optical depth \( \tau_\nu \) is fundamentally useful in planetary and astrophysical problems:

\[
\tau_\nu \equiv \pm \tilde{\tau}_\nu(\Omega = 0). \tag{2.15}
\]

We have included the \( \pm \) because optical depth can be chosen to increase or to decrease in the \( \Omega = 0 \) (upward) direction, but have assumed that both quantities are defined to be zero at the same height. Optical depth can be used in two ways. First, it can quantify the opacity of some physical layer in a manner independent of the illumination angle (e.g., “a dust layer of visible optical depth 0.75”); in this capacity, the term *optical thickness* is sometimes used instead. Second, it can serve as an optical coordinate (e.g., “At wavelength \( \lambda \), the tropopause lies at \( \tau = 0.3 \), while the surface is at \( \tau = 2.0 \)”). In this case, one must choose where \( \tau = 0 \) (typically at the top of atmosphere for planets or stars) as well as which direction \( \tau \) increases (typically, downward). We will adopt this coordinate in Chapter 6.

2.3 General solution

If one knows the source function \( J_\nu \), a “general solution” of the RTE is readily obtained. Simply multiply each side of (2.13) by \( \exp(-\tilde{\tau}_\nu) \) to get

\[
\frac{d}{d\tilde{\tau}_\nu}(I_\nu e^{\tilde{\tau}_\nu}) = J_\nu(\tilde{\tau}_\nu)e^{\tilde{\tau}_\nu}, \tag{2.16}
\]

then integrate from the boundary of a medium at \( \tilde{\tau}_\nu \equiv 0 \) to some depth \( \tilde{\tau}_\nu \), obtaining

\[
I_\nu e^{\tilde{\tau}_\nu} \bigg|_{0}^{\tilde{\tau}_\nu} = \int_{0}^{\tilde{\tau}_\nu} J_\nu(x)e^{x-\tilde{\tau}_\nu} dx, \tag{2.17}
\]

\[
I_\nu(\tilde{\tau}_\nu) = I_\nu(0)e^{-\tilde{\tau}_\nu} + \int_{0}^{\tilde{\tau}_\nu} J_\nu(\tilde{\tau}_\nu')e^{\tilde{\tau}_\nu'-\tilde{\tau}_\nu} d\tilde{\tau}_\nu' \tag{2.18}
\]

Thus, the radiance within a medium is the sum of an exponentially decaying contribution from any radiation incident on the surface, plus the accumulated contributions of any sources within the medium itself. Keep in mind that \( \tilde{\tau}_\nu \) as used here specifies a position in the medium, and that the physical position will vary linearly with \( \tilde{\tau}_\nu \) if the medium is homogeneous, but the physical distance implied by a unit change in \( \tilde{\tau}_\nu \) depends on frequency (usually, with enormous variations).

The reader may verify (Problem 4) that in the special case of a nonscattering medium with uniform temperature, the radiation intensity approaches the Planck value exponentially in optical path. In media where \( T \) changes very slowly over unit optical path, \( I_\nu \approx B_\nu \). The interesting (and common) cases are those where \( T \) changes significantly over one optical path, so that the beam remains out of thermal equilibrium with the medium. When calculating heating rates due to radiation, even small deviations from equilibrium can become important.
The solution (2.18) is quite useful for problems with no scattering, since we know \( J_\nu \) if we know the temperature distribution. In the more general case where scattering occurs, we do not normally know \( J_\nu \) \textit{a priori}. In this case, thermal equilibrium may not occur even when temperature varies little over a unit optical path. If only scattering occurs (\( \omega = 1 \)), the situation is not necessarily simplified because (2.12) remains integro-differential and non-separable in \( \Omega \).

Now that we have the RTE and its most basic solution in hand, two tasks remain. First, we must relate the extinction parameters and phase function in the equation to the fundamental physical properties of media. Second, we need methods of approximate solution of the equation in situations where the source function depends on the solution. These two matters will occupy the next three Chapters, after which we move on to applications.

### 2.3.1 Polarization

Some brief discussion of the generalization of this to polarized light.

### Exercises

2.1 Suppose that optical depth is defined as zero at the top of the ocean, increasing downward, and that the volume absorption coefficient of seawater is \( k_{1,\nu} = 10 \text{ m}^{-1} \) near 1 \( \mu \text{m} \) while it is \( k_{0,\nu} = 10000 \text{ m}^{-1} \) near 10 \( \mu \text{m} \). At what depth would you find \( \tau = 1000 \)? \( \tau_{10} = 1000 \)?

2.2 Explain, by invoking the radiative transfer equation, why the blackbody radiation field within an isothermal medium is unchanged if the medium is able to scatter as well as absorb.

2.3 If emission from an infinite plane surface is isotropic (i.e., radiance measured at some height immediately above the surface is independent of direction in the upward hemisphere), find the probability density of the zenith angle at which a particular photon is emitted. Explain why the answer is different if we ask the same question about a photon emitted from a small object measured at large distances from the object.

2.4 Show that any beam of radiation in a isothermal, nonscattering medium approaches the Planck value exponentially as it propagates.

2.5 A small fire is producing a smoke plume. The area-weighted average particle radius is 10 \( \mu \text{m} \), there are \( 10^5 \) particles per cc of air, the density of the particles is 0.1 g/cc, the air density is 1.2 kg/m\(^3\), the air by itself is transparent, and suppose the particles are opaque spheres. Neglecting diffraction, calculate:

1. the smoke particle cross section \( \sigma \);
2. the volume extinction coefficient \( \varepsilon_v \) of smoky air;
3. the mass extinction coefficient \( \varepsilon \) of smoky air;
4. the optical path through a 100 cm segment of the smoke plume;
5. the fraction of incident light that passes through 100 cm of smoke plume unaf-
fected.

2.6 Show by integrating the equation over all directions that if there is no emission, (2.12) can be written

\[- \frac{1}{4\pi e_s \rho} \nabla \cdot \mathbf{F}_\nu = (1 - \tilde{\omega}_\nu) I_{0,\nu}\]

where \( \mathbf{F}_\nu \) is the net flux vector. If, in addition, \( \tilde{\omega}_\nu = 1 \) then we have the flux integral relation \( \nabla \cdot \mathbf{F}_\nu = 0 \) expressing the conservation of electromagnetic energy.

2.7 Show by inserting a linear expansion

\[ B(\tau) = B(T) + \alpha dB/d\tau \]

into the RTE that the radiance emitted by a semi-infinite cloud of top \( (\tau = 0) \) temperature \( T \) differs from the blackbody emission \( B(T) \) by an amount

\[ \delta I = \int_0^\infty \frac{dB}{d\tau} e^{-\tau} d\tau. \]
Chapter 3

Scattering within media

Scattering occurs at boundaries between media of different properties, as is well known. It can also occur, however, within “homogeneous” media whenever the homogeneity does not extend all the way down to scales much smaller than a wavelength. This means scattering occurs to some degree in practically all gaseous, liquid, and solid materials even if they appear macroscopically (or statistically) homogeneous. In this chapter we consider scattering in such media, starting by scattering from single particle and working up. In the following chapter, we will treat scattering from macroscopic objects with well-defined boundaries.

All media are composed of atoms and molecules which, for the purposes of this chapter, may be regarded as point particles. The radiative transfer equation is a continuum equation, with continuum absorption and scattering parameters. The continuum view is a model of reality, required because treatment based on contributions from each particle would be impossible (the same practical requirement exists in fluid dynamics and leads to the development of Navier-Stokes equations rather than kinetic equations for individual molecules). The continuum model works fine as long as important state variables remain smooth on scales large enough to contain plenty of molecules. Thus, our task is to relate the continuum properties of homogeneous materials to the properties of the molecules of which they are composed. Historically such relationships have been used to infer molecular properties, by observing continuum parameters and working backward.

Scattered power always depends on the relative phases of different waves. To discuss phase information quantitatively, we will use complex numbers with the notation

\[ x \equiv (x_R, x_\Im) \equiv x_R + ix_\Im \]

where

\[ x^2 = x_R^2 - x_\Im^2 + 2ix_Rx_\Im. \]

Since all complex quantities we encounter here will also be functions of frequency, throughout the text we will take the real/imaginary subscripts to imply frequency dependence implicitly without the usual \( \nu \) subscript. Further, for clarity’s sake, in this chapter we will not bother with \( \nu \) subscripts on newly introduced quantities (when in...
3.1 Scattering by single atoms or molecules

The behavior of most molecules and atoms is well modeled by treating them as simple dipoles—even though they consist of multiple charges arranged in complicated geometry—as long as $\lambda \gg 1\,\text{Å}$, we allow some flexibility in quantifying the mass and restoring force, and the scattering is primarily from bound (rather than free) electrons.

3.1.1 Ideal dipole

We begin by analyzing the behavior of a single electric dipole of dipole moment

$$\mathbf{p} = q_e \mathbf{x},$$

where $\mathbf{x}$ is the effective separation vector between charges $+q_e$ and $-q_e$. In the absence of external applied field, $\mathbf{p} = 0$ for atoms and linear molecules. Some asymmetric molecules such as H$_2$O have permanent dipole moments; over a large number of molecules, those moments average to a mean value that is nonzero if an electric field is applied, leading to additional radiative interactions beyond those included in our calculation (see Problems).

A dipole moment is induced by an applied electric field $\mathbf{E}_{\text{loc}}$:

$$\mathbf{p} = \alpha \mathbf{E}_{\text{loc}},$$

since opposite charges are forced in opposite directions. In cgs units, the dipole polarizability $\alpha$ has units of volume. To very good approximation we can assume that only one charge, the electron, moves$^1$. In addition to the Lorentz force$^2$

$$m_e \frac{d^2 \mathbf{x}}{dt^2} = q_e \mathbf{E}_{\text{loc}},$$

the electron experiences a linear restoring force $-\xi m_e \mathbf{x}$ and damping force $-\gamma m_e \frac{d\mathbf{x}}{dt}$ (we assume $\xi$ is a constant but $\gamma$ may depend on frequency). Considering only the component in the direction of the applied $\mathbf{E}_{\text{loc}}$, Newton’s law may be written

$$\frac{d^2 x}{dt^2} = \frac{q_e E_{\text{loc}}}{m_e} - \xi x - \gamma \frac{dx}{dt}.$$  

Inserting

$$x = x_0 e^{-\omega t}$$

$^1$In quantum physics, the electrons behave as a probability "clouds" surrounding their center of attraction, rather than as points per se. When an electric field is applied, their average position shifts. Our model here still works though.

$^2$We may take this relation as the definition of electric field.
Figure 3.1: An incident wave travelling in direction $\Omega$ with $E$ pointing instantaneously upward induces a dipole moment $P$ also upward, where the restoring electromotive force (represented by a spring) approximately balances the Lorentz force of the incident field. Scattered power is radiated in directions, e.g., $\Omega'$, with associated (much smaller) fields $E'$. The scattering angle $\Theta$ is that between $\Omega$ and $\Omega'$; we also consider the angle $\theta$ between $\Omega'$ and $P$.

yields a solution where

$$x = \frac{q_e/m_e}{(\xi - \omega^2) - i\omega} E_{\text{loc}}.$$  \hfill (3.5)

The dipole has a natural resonance at $\omega_0 = \sqrt{\xi}$, and an extinction half-width at half-maximum $\delta \omega_N = \gamma/2$. Many authors use the Greek letter $\alpha$ for line width, but it is even more commonly used for polarizability (by most of the same authors!); to avoid confusion, we will simply use $\delta \omega$ for line width. Employing these definitions, (3.1) can be written

$$P = \frac{q_e^2/m_e}{(\omega_0^2 - \omega^2) - 2i\delta \omega N \omega} E_{\text{loc}}.$$  \hfill (3.6)

Notice that complex quantities arise, indicating a phase lag between oscillations of the electron and those of the incident field. Comparing (3.6) with (3.2), we see that dipole polarizability must be

$$\alpha = \frac{q_e^2/m_e}{(\omega_0^2 - \omega^2) - 2i\delta \omega N \omega}$$  \hfill (3.7)
Rationalizing the denominator, we can separate the real and imaginary parts:

\[ \alpha = \frac{q^2/m_e}{(\omega_0^2 - \omega^2)^2 + 4\delta \omega N^2 \omega^2(\omega_0^2 - \omega^2, 2\delta \omega N \omega)}. \tag{3.8} \]

The field \( \mathbf{E}' \) radiated by a classical dipole\(^3 \) at large distances \( r \) is proportional to charge acceleration, specifically, in cgs units:

\[ \mathbf{E}' = \frac{1}{c^2 r} \frac{d^2 \mathbf{p}}{dt^2} \sin \theta \tag{3.9} \]

where \( \theta \) is the angle between the scattering direction and \( \mathbf{p} \). The radiation is polarized so that \( \mathbf{E}' \) lies in the plane containing \( \Omega' \) and \( \mathbf{p} \). Assuming harmonic dependence of the incident field and utilizing (3.2), this becomes

\[ \mathbf{E}' = \frac{\omega^2}{c^2 r} \alpha \sin \theta \mathbf{E}_{0c}. \tag{3.10} \]

Integration of the radiated energy flux about the sphere at radius \( r \) using (1.2) yields the total power

\[ W = \frac{8\pi \omega^4}{3c^4} |\alpha|^2 F_{0c}. \tag{3.11} \]

This is called the Larmor power formula, where \( F_{0c} \) designates the flux incident on the dipole as expressed in equation (1.2). The ratio of \( W \) to \( F_{0c} \) in (3.11) is by definition the particle scattering cross section \( \sigma_s \) introduced in Section 2.1.3, for the special case of a single isolated scattering dipole.

Though damping was included in the electron equation of motion, there is no “friction” in the usual sense acting on an electron. In the absence of absorption events (which will be considered separately), the damping must actually be due to power \( W \) lost in radiation. The resistance felt in trying to accelerate a charge is known as “radiation resistance.” If we demand that the work done against this resistance, associated with the damping term \( \gamma \frac{dx}{dt} \) in (3.4), be equal to the radiated power given by (3.11), we obtain the relationship:

\[ \gamma = \frac{2q^2 \omega^2}{3m_e c^2}. \tag{3.12} \]

From this we may compute the quality factor (ratio of oscillation rate to damping rate \( \omega/\gamma \)) of the oscillator to be of order \( 10^8 \) at visible frequencies. Clearly, the damping of typical dipole oscillators is extremely weak at naturally prevalent wavelengths. The natural line width corresponds to a difference in wavelength

\[ \delta \lambda = \lambda \frac{2\delta \omega N}{\omega} = \frac{2\pi \gamma c \lambda}{\omega^2} = \frac{4\pi q^2}{3m_e c^2} \] \( \tag{3.13} \)

having a constant value of \( 6 \times 10^{-6} \) nm. This indicates a very narrow natural line even for gamma rays (\( \lambda \sim 10^{-2} \) nm), let alone infrared or visible radiation.

\(^3\)The near field of a quasi-static dipole is discussed in Section 4.1.
3.1.2 Real atoms/molecules

Our basic model was a damped oscillator, analogous to a lump of mass on a spring, characterized by a resonant frequency and polarizability determined by fundamental electron properties. Adding more electrons and/or bound atoms to the molecule is like adding more lumps and springs. The resulting system has many “modes” of vibration—one per lump—with each mode possessing a resonant frequency and an excitability per unit forcing analogous to polarizability. These characteristics depend on the spatial arrangement of the lumps and strengths of the various springs. The analogous characteristics for molecules can be calculated accurately only with quantum mechanics, but qualitatively they behave like one might expect from classical reasoning. They will be discussed in detail in Chapter 5.

The implications here are:

- Most substances will have many resonant frequency/polarizability combinations rather than just one. Behavior will be influenced by all resonances, but mainly the nearest one. Liquids and especially solids tend to have resonances at much lower frequencies than gases, due to the strong bonds; these resonances involve coherent motions of many adjacent molecules with large cumulative mass. Gaseous scattering at visible and longer wavelengths may typically be modeled by considering only the lowest resonance.

- Since individual charges will be easier to move in some directions than in others, the polarizability tensor will not generally be diagonal, that is, polarization will not be in exactly the opposite direction as the applied field as in (3.2). This affects light polarization but has a minor impact on scattered power.

- Energy can be absorbed near resonances due to quantum transitions, whose transition energy matches the resonant photon frequency. We haven’t accounted for this yet.

Some of these issues are briefly revisited in Section 3.3.1.

3.2 Multiple dipoles

The next step in building understanding of scattering is to consider what happens when more dipoles are present. Let’s start with two, the second identical to the first, separated by a distance $d$. Away from resonance, the dipole scattering is small enough so that as long as

$$d \ll \lambda$$

(3.14)

where $E$ is the incident field.

First, suppose that $d \ll \lambda$. In this case, each dipole experiences approximately the same Lorentz force and the two dipoles oscillate and radiate in phase with each other. This radiation remains in phase in all propagation directions. The radiated electric
3.2. **MULTIPLE DIPOLES**

Fields add, so according to (1.1) the scattered power is quadrupled. This makes sense according to the Larmor formula (3.11) if we regard the pair of dipoles as being equivalent to a single dipole of twice the polarizability. This result generalizes to \( n \) dipoles inside a region much smaller than a wavelength; the scattered power is proportional to \( n^2 \), unless the number of dipoles becomes large enough so that (3.14) begins to fail.

If \( d \sim \lambda \) interesting things start to happen. We consider the radiation received at a point a long distance \( r \) from the two dipoles. This radiation will be:

\[
E = E_1 e^{i(kx_1 - \omega t)} + E_2 e^{i(kx_2 - \omega t)}
\]  

(3.15)

where \( x_1 \) and \( x_2 \) are the total path lengths traveled by radiation scattering off the respective dipoles, including the path to reach the dipole and the subsequent path from the dipole to the observation point. The complex constants \( E_1 \) and \( E_2 \), which reflect the amplitude and phase of the scattered radiation, will be approximately equal if the dipoles are identical and \( d \ll r \). Then (3.15) describes a single plane wave whose amplitude is

\[
E = E_1 (1 + e^{ik(x_1 - x_2)}).
\]  

(3.16)

Suppose \( d = \lambda/2 \) and the dipoles are located on a \( y \)-axis with the propagation direction \( \Omega \) of the incident wave lying on the \( x \)-axis (Figure 3.2). Since the incident wave arrives simultaneously at both dipoles, the dipoles radiate in phase. Their radiation will remain in phase throughout the \( x - z \) plane since \( x_1 = x_2 \), giving strong scattering, but will become 180° out of phase in the \( y \) direction since \( x_1 - x_2 = -\pi/k \), such that at large distances the two will completely cancel. The result is a dimpled or doughnut-shaped scattering pattern. If the dipoles are located instead on the \( x \)-axis parallel to \( \Omega \), they will radiate 180° out of phase. In the forward and backward (\( x \)-) directions the radiated waves will come back into phase, while in the \( y - z \) plane they will cancel out, leading to a two-lobed scattering pattern. Many other scattering patterns are possible; they can become much more complex if more than two dipoles are present. Such patterns are often exploited in the design of antennas when it is desired to broadcast or receive preferentially in certain directions. Note that the upper bound on radiation in any particular direction—four times that of a single dipole—occurs when \( x_1 = x_2 \), which is approximately true for all directions in the \( d \ll \lambda \) case. Scatterers are most efficient when close together.

In gases and liquids, dipole scatterers are randomly located. It is therefore worth inquiring what happens when we average net results over many possible positions of a second dipole within a “neighborhood,” of small size compared to \( r \) but large compared to \( \lambda \), surrounding the first dipole. Using (3.16), the flux received at \( r \) for a particular location of the second scatterer is equal to

\[
S \propto |E|^2 = EE^* = |E_1^2|[2 + 2\sin(x_1 - x_2)]
\]

where \(^*\) denotes the complex conjugate. The \( x_1 - x_2 \) argument will be constant along a set of surfaces that become regularly spaced parallel planes as \( r \to \infty \), oriented perpendicular to the line that bisects the incoming and outgoing light ray paths (Figure 3.3). Since the neighborhood is large compared to \( \lambda \), when we average over it, the sine term will oscillate many times and average to a very small number, regardless of
Figure 3.2: Scattering pattern from two dipoles. A plane wave is incident from the left, with instantaneous crest and trough positions indicated by solid and dashed straight lines respectively. The scattered wave phase lines (circular in the horizontal plane) are similarly indicated.

\[ \langle S \rangle \approx \langle S_1 \rangle + \langle S_2 \rangle. \]  

(3.17)

Thus, the radiation received on average will almost exactly equal the linear sum of the scattered power from each dipole, ignoring the others:

This principle holds when we add more dipoles, provided that the new dipoles are located randomly, and that we don’t add so many of them that their spacing becomes small compared to \( \lambda \).

This result is important. In a gas, molecular positions obey Boltzmann statistics to a high degree of accuracy; in other words, each molecule’s position is random and independent of that of the other molecules. Similar behavior is expected for macroscopic particles (e.g., dust) embedded in another medium. This allows us to conclude that the power scattered by many such particles will just be the amount we calculated for an individual dipole, multiplied by the number of dipoles. Also, the directional distribution will match that of an individual scatterer. This conclusion comes with important
3.2. MULTIPLE DIPOLES

Figure 3.3: Scattering from a distant source, angle $\Omega$, into $\Omega'$ toward a distant target. Long dashed lines show surfaces of constant total ray distance at constant phase. One scattering particle is labeled 1; we consider others added either individually at random, or in a close-packed linear arrangement denoted by the dotted line.

provisos, and is reliable only for gases as will be shown in the next few sections.

Before getting to that, however, we discuss an alternative case that concerns a point made at the beginning of this chapter regarding the meaning of “homogeneity” in a scattering medium. Consider a plane filled with evenly spaced dipoles, oriented at a nonzero angle to the planes of constant $x_1 - x_2$ noted above. In the limit as $d/\lambda \to 0$, the radiance scattered toward $r$ (in direction $\Omega'$) becomes the sum of an infinite number of scattered waves originating at all points along the plane of dipoles:

$$S \propto \left( E_{inc} \int e^{ikx(s)} \, ds \right)^2,$$

where $x(s)$ is the total path traveled by the radiation, which depends on the position $s$ along the intersection of the dipole and scattering planes (the “scattering plane” is defined as the plane containing the incident and scattered propagation directions). Since the harmonic function will again average (integrate) to zero over a large scattering plane where its argument oscillates ad infinitum, we are left with the result that the scattered
energy is zero. This result holds for any scattering direction \( \Omega' \) except \( \Omega' = \Omega \); in this special case, the path \( x \) is independent of position and the integrand is a constant. This means that light propagates through truly homogeneous media with no scattering\(^4\). This explains why many liquids and a few solids are transparent.

If \( d \) is finite, the above integral is replaced by a finite sum. For \( d < \lambda/2 \), the same result will occur, due to destructive interference; for smaller \( \lambda \), resonances will occur that cause large scattering in certain directions (the basis of crystallography). So we see that gases scatter occur only because their molecules are randomly, hence irregularly, distributed. We will return to this point at the end of the chapter.

### 3.3 Continuous media

In this section we begin the transition from dipole-oriented to continuum description of a medium, now considering a field of dipoles having some density that is (statistically) homogeneous over a large volume. We begin by considering a volume element that is large enough to contain many dipole scatterers, but has linear dimensions small compared to \( \lambda \). Due to the small size, we may treat the macroscopic field \( E \) as being spatially uniform with harmonic time variation throughout the volume element. It will turn out that after this we will essentially understand scattering in gases, though condensed matter will require further development.

First we define the polarization \( P \) of a dielectric medium composed of \( N \) identical electric dipole scatterers per unit volume,

\[
P \equiv N(p).
\]

(3.18)

In the cgs system, \( P \) and \( E \) have equivalent units. Note from this definition that in multicomponent gas the total polarization will be given by the sum of the contributions from individual components.

At moderate field strengths, polarization turns out to be proportional to the ambient electric field in a way analogous to the dipole moment of a single scatterer:

\[
P = \chi_E E.
\]

(3.19)

This relationship, called a constitutive relation, does not derive from electromagnetic theory but from models and/or observations of the material in question. It is a necessary closure relation analogous to the equation of state in thermodynamics. In general the electric susceptibility \( \chi_E \) is a nondimensional, complex, frequency-dependent tensor—but in gases and liquids we can safely assume the medium is isotropic so that \( \chi_E \) is scalar. The field \( E \) is the macroscopically smooth average electric field within the medium, where the average is taken over all (equally likely) possible arrangements of the individual molecules that are consistent with the macroscopic properties of the substance. By contrast, the instantaneous field \( E' \) (if it could be measured) would obviously vary on tiny scales depending on where exactly the nearby charges were at that instant.

\(^4\)We have in effect obtained the celebrated Huygens’s principle which states that the subsequent development of an electromagnetic wave can be thought of as being caused by a continuum of sources aligned with the present wave fronts.
3.3. CONTINUOUS MEDIA

The macroscopic field $\mathbf{E}$ in the medium is equal to the incident field plus the mean field generated by induced dipole moments:

$$\mathbf{E} = \mathbf{D} + \mathbf{E}'$$  \hspace{1cm} (3.20)

where $\mathbf{D}$ (known as the electric displacement$^5$) is the field that would have prevailed in the absence of a dielectric response. The field $\mathbf{E}'$ generated by the medium opposes the incident field, making $\mathbf{E} < \mathbf{D}$. It is easily shown that a homogeneous distribution of dipoles obeying (3.1) and (3.18) will generate a mean field

$$\mathbf{E}' = -4\pi \mathbf{P}.$$  

Then (3.19) and (3.20) can be combined to give

$$\mathbf{E} = \frac{\mathbf{D}}{\epsilon}$$  \hspace{1cm} (3.21)

where we have defined the relative permittivity or dielectric constant

$$\epsilon \equiv 1 + 4\pi\chi_E.$$  \hspace{1cm} (3.22)

Maxwell’s equations describe the behavior of $\mathcal{E}$ but would be impossible to handle in the presence of many moving charges. Fortunately, as long as (3.19) holds, a set of equations (given in Chapter 4) can be derived that have the same form as Maxwell’s equations but govern the macroscopic mean fields in continuous dielectric media. These equations support waves just like the “vacuum” ones, except they propagate at velocity $c$ related to that $c^*$ of light in a vacuum by

$$c^*/c = m = \sqrt{\epsilon/\mu}$$  \hspace{1cm} (3.23)

where $m$ is defined as the index of refraction of the medium. All quantities save $c^*$ can have imaginary as well as real parts. The magnetic permeability $\mu$ does not significantly differ from unity in any materials of likely interest here, so we’ll assume it is equal to one.

Real parts of quantities in (3.23) are associated with the propagation rate of radiation and, of equal importance, the ratio of the $\mathbf{D}$ and $\mathbf{H}$ field strengths in an electromagnetic wave. In the continuum approximation it is not individual scatterers but spatial variations in $m$ that produce scattering, since waves then cannot proceed undisrupted. We return to this matter in Section 3.4.

Imaginary parts of quantities in (3.23) are associated with damping or absorption of radiation. Consider a plane wave traveling in the $x$ direction, described in vacuum by

$$E = E_0 e^{i(kx-\omega t)} = E_0 e^{2\pi i(x/\lambda - \nu t)}.$$  

$^5$In the CGS system of units $\mathbf{D}$, $\mathbf{E}$, $\mathbf{H}$, and $\mathbf{B}$ all have the same units, but this is not the case in other systems like MKS where all are different. The physical interpretation of $\mathbf{D}$ given here is somewhat of a simplification; it is formally defined as $\mathbf{E} + 4\pi \mathbf{P}$. 

If the wave enters a dielectric material, the wavelength $\lambda = c/\nu$ must decrease since $c = c^*/m$ is lower while $\nu$ must remain the same everywhere. With

$$m \equiv m_R + im_\Theta,$$  \hspace{1cm} (3.24)

a distance $ds$ into the dielectric we have

$$\frac{|E(ds)|}{|E(0)|} = \exp \left(-2\pi \frac{m_\Theta ds}{\lambda^*}\right),$$

where $\lambda^*$ denotes the wavelength in vacuum. As $ds$ goes to zero,

$$\frac{d|E^2|}{|E^2|} = -4\pi \frac{m_\Theta ds}{\lambda^*}.$$  

Thus, most of the energy is absorbed within a distance comparable to $\lambda^*$ if $m_\Theta$ is about $10^{-1}$. Comparing with (2.1) gives a key relation between the absorption coefficient $k_\nu$ employed in the radiative transfer equation and the index of refraction:

$$\rho k_\nu = \frac{4\pi \nu m_\Theta(\nu)}{c^*}. \hspace{1cm} (3.25)$$

In order to make use of this result, however, we must be able to calculate $m$ from the molecular properties of a substance.

### 3.3.1 Index of refraction in gases

In some materials, the scattering may be sufficiently weak for the field experienced by a molecule to be little affected by its neighbors. In this case we can continue to assume (3.14) in our continuum treatment, and (3.6) may be combined directly with (3.18), (3.19), (3.22), and (3.23) to obtain

$$m^2 = \epsilon = 1 + 4\pi \chi_E \approx 1 + 4\pi N\alpha \hspace{1cm} (3.26)$$

which we call the *incoherent or noninteracting scatterer* approximation.

This approximation should hold if two conditions, which may be anticipated from the discussion of Section 3.2, are met:

1. The total field scattered from dipoles within a distance $d < \lambda$, which is always coherent in phase with $E$, must be weak compared to $E$;
2. Scatterers at $d > \lambda$ must be distributed randomly so that their waves do not superpose coherently with $E$, so that the average work done by the incident field in accelerating the dipole moment is unaltered; or else, their total scattered field contribution must also be weak compared to $E$.

The first condition requires a low scatterer density and is typically satisfied only by gases. The second condition is typically not met by crystalline solids but, as noted in Section 3.2, is satisfied by gases and often by liquids. Thus, the incoherent scatterer approximation should be good for gases (notice that we didn’t require the scattered field
3.3. CONTINUOUS MEDIA

to be weak at a distance; if it is, however, this simplifies things further as discussed in Section 3.4. In Section 3.3.3 and 3.4, by doing additional work, we will respectively relax the first, and then both, of these conditions. This will lead to results that are useful respectively for liquids and for arbitrary media.

In neutral molecules at terrestrial temperatures, \( n > 100,000 \text{ cm}^{-1} \), so that for most incoming solar radiation \( \nu \ll n \) and both \( \alpha \) and \( \chi_E \) are small. Also, we have seen that the damping of the harmonic oscillator \( m_\alpha \) is tiny. These facts imply that \( m \) is very close to unity. This may be taken as an \( a \ posteriori \) justification of the first assumption above for gases. With \( m \) close to unity, we can write

\[
m \approx 1 + \frac{m^2 - 1}{2}
\]

which when combined with (3.26) gives

\[
2(m - 1) \approx e - 1
\]

Further, we have

\[
m^2 \approx (m^2_R, 2m_Rm_\alpha)
\]

since \( m^2_R \gg m^2_\alpha \).

As discussed earlier, \( m_\alpha \) is associated with absorption. We can quantify the relationship between \( m_\alpha \) and dipole properties in a tenuous medium by comparing (3.26) with (3.29) and recalling the dipole polarizability result (3.8). Equating real terms in the respective equations and solving for \( m_R \) using (3.29) first yields the expression

\[
m_R - 1 = \frac{2\pi N q_e^2}{m_e} \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + 4\delta \omega_N^2 \omega^2}.
\]

For frequencies near resonance, this expression may be simplified by noting that \( \omega^2 - \omega_0^2 = (\omega + \omega_0)(\omega - \omega_0) \approx 2\omega_0(\omega - \omega_0) \), and that the second term in the denominator need only be correct to zeroth order \( (\omega^2 \approx \omega_0^2) \). Then

\[
m_R - 1 \approx \frac{\pi N q_e^2}{m_e \omega_0} \frac{\omega_0 - \omega}{(\omega_0 - \omega)^2 + 2\delta \omega_N^2} \quad (\omega \approx \omega_0).
\]

If we may equate the imaginary terms in the earlier equations, we may follow a similar procedure; making the same approximations and using (3.31) to eliminate \( m_R \) yields

\[
m_\alpha \approx \frac{\pi N q_e^2}{m_e \omega_0} \frac{\delta \omega_N}{(\omega_0 - \omega)^2 + 2\delta \omega_N^2} \quad (\omega \approx \omega_0)
\]

The real and imaginary indices of refraction near a resonance are depicted in Figure 3.4. The (relatively narrow) region where \( m_R \) decreases with \( \omega \) is known as the region of anomalous dispersion, since the spreading of colors (given by \( \partial m_R / \partial \omega \) by a prism will be in the opposite direction to that encountered elsewhere in the spectrum.

Note that significantly nonzero \( m_\alpha \) is concentrated very close to the resonant frequency \( \omega_0 \). Away from this frequency, the real index of refraction changes gradually while the imaginary part quickly drops to zero. Combining (3.32) with (3.25) yields
the (continuous) absorption coefficient in tenuous media, in terms of the parameters describing the bound electrons in the individual dipoles. Thus, absorption is predicted to be large near resonances. This is a key milestone in our goal to see how the parameters governing radiative transfer arise from the physical nature of the medium.

**Limitations of the classical model**

The above result implies a paradox. We noted before that the power dissipated in our classical dipole model, which led to the natural line width $\delta \omega_N$, corresponded to a transfer of power from the incident wave into a scattered wave. But (3.32) establishes a quantitative relation between this and absorption! This paradox arises because of shortcomings in our classical and continuum models. The classical model cannot really predict absorption: the parameter $\gamma$ was essentially a “fudge factor” that we ini-
tially set to account for the scattered energy, but we could easily boost it to represent internal absorption too, if this absorption acted like friction. Meanwhile, our derivation of (3.25) was for a uniform material, which we know from earlier arguments does not actually scatter at all if it is truly uniform (in which case, due to interaction between scatterers, the “radiation resistance” and $\gamma$ vanish). These issues are subtle, and to resolve them is beyond our scope here. Suffice it to say that in real media, (i) the parameters $\omega_0$, $\gamma$ (or the line width $\delta \omega$), and effective $q^2/m_e$ are determined by quantum physics as discussed in Chapter 5, with line widths usually being significantly greater than the “natural” line width $\delta \omega_N$; (ii) given these parameters, our classical arguments are quite good at predicting scattering and absorption.

3.3. Rayleigh scattering and the blue sky

Rayleigh’s explanation for the blue sky followed from arguments based on incoherent scattering. Since $\omega_0^2$ in gas-phase particles is much larger than the other terms in the denominator in (3.6), $\alpha$ and $\chi_E$ do not vary greatly with wavelength for molecular scattering of solar radiation. This leaves the $\omega^4$ factor in the Larmor power formula (3.11) as the dominant source of wavelength dependence in the scattered radiation. The result is a blue sky on earth (when one looks away from the sun at radiation that has been scattered out of the direct beam), since blue wavelengths are scattered so much more than red wavelengths. The direct beam itself is correspondingly reddened, as becomes apparent at sunset.

The total scattered power per molecule is given in terms of $\alpha$ by (3.11). Using (3.26) we may write the scattering cross section as

$$\sigma_s \approx \frac{\omega^4}{6\pi c^4 N \hbar} |e - 1|^2, \quad (3.33)$$

or using (3.28),

$$\sigma_s \approx \frac{2\omega^4}{3\pi c^4 N \hbar} |m - 1|^2. \quad (3.34)$$

The scattered power per unit volume is then

$$s_v = N \sigma_s \approx \frac{2\omega^4}{3\pi c^4 \hbar} |m - 1|^2. \quad (3.35)$$

The strange thing about this equation is that the scattering coefficient ($s_v$) is related to another continuum quantity, $m$, in a way that depends on the atomicity $N$ of the medium! If we divided the gas’s dipoles into smaller and smaller “pieces”, conserving $N\alpha$, we would not change $m$ but would keep reducing $s_v$ and eventually obtain a non-scattering gas in the limit $N \to \infty$. This result is consistent with the discussion in Section 3.2, and follows from the quadratic dependence of scattered power per dipole on $\alpha$. Equation (3.35) is useful for deducing $N$ from observations of a gas of known composition since $m$ and $s_v$ are easily observable in the laboratory.

We also need to know the angular distribution of the scattered radiation, or phase function. In general, this depends on four angles, two characterizing the incident direction $\Omega$ and two the scattered direction $\Omega'$. However, if the scattering particles are not
CHAPTER 3. SCATTERING WITHIN MEDIA

Figure 3.5: Angular distribution of Rayleigh scattering, incident from left. For unpolarized incidence, the pattern is symmetric about the axis of incident propagation. The dashed line indicates scattered energy polarized perpendicular to the scattering plane (the page), while the solid line includes also the energy polarized parallel to the scattering plane. Arrows show field strength in the scattering plane, with dot-filled circles indicating perpendicular fields.

oriented in any preferential direction, then by symmetry, the scattered intensity from a small, homogeneous volume can depend only on the angle between $\Omega$ and $\Omega'$, the scattering angle $\Theta$ (see Figure 3.1). This is true for molecular scattering under ordinary circumstances.

We have already argued that swarms of incoherent scatterers will produce the same angular distribution as a single constituent dipole. From (3.9) we may write

$$S \propto \sin^2 \theta$$

(3.36)

where $\theta$ is the angle between $\mathbf{p}$—which aligns with the incident field—and the scattering direction. This is a doughnut-shaped radiation pattern, with maximum radiation in directions perpendicular to the dipole, going to zero in the direction parallel to the dipole.

Light arriving from the sun is unpolarized (so-called natural light), which means that instantaneous $\mathbf{E}$ points randomly in the plane perpendicular to the propagation direction (the “field plane”). The principle of superposition enables us to calculate the net scattered energy by adding the contributions from any two perpendicular polarized components of the incident light (and therefore perpendicular component directions of $\mathbf{p}$ in the field plane). This means adding two doughnut-shaped patterns in equal measure, with one rotated $90^\circ$ on the propagation axis relative to the other, or

$$S \propto \sin^2 \theta_1 + \sin^2 \theta_2$$

with $\theta_1$ and $\theta_2$ the angles between the scattering direction and either of the two perpen-
3.3. CONTINUOUS MEDIA

dicular basis vectors in the field plane. From geometry one may obtain

\[ \sin^2 \theta_1 + \sin^2 \theta_2 = 1 + \cos^2(\Theta). \]

This watermelon-shaped distribution is shown by the solid curve in Figure 3.5. Normalization of this yields the *Rayleigh phase function*

\[ P(\Theta) = \frac{3}{4} \left[ 1 + \cos^2(\Theta) \right]. \]  

(3.37)

Note that if the incident radiation is polarized, the two “doughnut” patterns will add in unequal measures. Then the scattered distribution will become dependent on azimuth (angle between the scattering plane and the major axis of polarization), resembling a watermelon flattened so that its girth is greater perpendicular to the polarization direction.

A key feature of Rayleigh scattering is that natural light becomes polarized upon scattering, in all directions except \( \Theta = 0 \) and \( \Theta = \pi \). The scattered component polarized perpendicular to the scattering plane (as it happens, the first term in (3.37)) is isotropic, but the parallel component (the second term) decreases to zero at \( \Theta = \pi/2 \), accounting for all of the decrease in overall scattered power perpendicular to \( \Omega \) (see Exercise 10). This may be observed using polarizing sunglasses, by looking at the sky in various directions. Looking in a direction 90° away from the sun, one finds a significant difference in sky brightness depending on the orientation of the polarizing lenses.

The observed polarization is weaker than we calculated, however. One reason is scattering in the real atmosphere by particulates, but the main reason has to do with our assumptions regarding dipole polarizability. As discussed in Section 3.1.2, we assumed that \( \alpha \) would be a scalar, independent of the orientation of the incident field, and that \( p \) would always lie in the field plane. Most molecules, due to their lack of spherical symmetry, are easier to polarize in some directions than others. An individual dipole moment \( p \) will not be exactly parallel to the incident field, as it skews toward the direction(s) of easiest polarizability in that particular molecule. Since molecules have no preferred orientation, this does not shift or rotate the overall scattering pattern, but it does cause smoothing of the pattern due to all the random deviations of \( p \) out of the field plane. It also slightly alters the scattering cross section (Rayleigh 1918; Cabanes 1929).

### 3.3.3 Index of refraction in dense media

Within liquids and solids, the independent scatterer approximation fails and we must calculate the difference \( \mathbf{E}_{\text{loc}} - \mathbf{E} \) between the local field (that is, the Lorentz force per unit charge) truly experienced by a given molecule \( j \) and the field estimated in the continuum approximation. We still consider the volume element described in Section 3.3. In this section, we relax the first assumption on page 40 but must retain the second of these assumptions for the time being.

The difference between \( \mathbf{E}_{\text{loc}} \) and \( \mathbf{E} \) arises from two key facts. First, the Lorentz forces experienced by \( j \) are governed by the field generated at \( j \)'s location by *other*
CHAPTER 3. SCATTERING WITHIN MEDIA

charges, not counting that of \( j \) itself (this is an inherent feature of the field concept for understanding forces). But \( j \)'s own field contribution was implicitly included along with all the others in calculating \( \mathbf{E} \). We must subtract this contribution back out:

\[
\mathbf{E}_{\text{loc}} = \mathbf{E} - \mathbf{E}_{\text{self}}. \tag{3.38}
\]

Now one might ask how removal of a single dipole could make any difference, especially since we have argued that gaseous molecules obey Boltzmann statistics implying that the average spatial distribution of the remaining molecules will be unaffected by the removal, so that the typical field at \( j \) should be no different from that anywhere else. However, molecular positions are not quite independent since it is impossible for any other molecule to occupy exactly the same location as \( j \), which rules out certain (very unlikely) configurations. For almost all purposes, this restriction is unimportant. However (our second key fact), these ruled-out configurations turn out to have contributed substantially to \( \mathbf{E} \) at the specific location of \( j \), due to the extreme strength of the electric field very close to a dipole. From this point of view it is the contribution of these ruled-out configurations that is being removed in (3.38).

To calculate \( \mathbf{E}_{\text{self}} \) we calculate the change in average field inside a spherical region surrounding \( j \) that would attend the hypothetical removal of \( j \), in other words the mean field generated by \( j \). A spherical geometry is the only choice for a fully isotropic medium that possesses the proper symmetry. This calculation is performed in Section 4.1; the result may be expressed as

\[
\mathbf{E}_{\text{loc}} = \mathbf{E} + \frac{4\pi \mathbf{P}}{3}. \tag{3.39}
\]

From (3.2) and (3.18) we obtain

\[
\mathbf{P} = N\alpha \left( \mathbf{E} + \frac{4\pi}{3} \mathbf{P} \right). \tag{3.40}
\]

Solving this for \( \mathbf{P} \) and employing (3.19) yields

\[
\chi_{E} = \frac{N\alpha}{1 - \frac{4\pi}{3} N\alpha},
\]

which using (3.23) and (3.22) can be rearranged into the Lorentz-Lorenz equation\(^6\)

\[
\frac{m^2 - 1}{m^2 + 2} = \frac{4\pi N}{3} \alpha. \tag{3.41}
\]

Note that for \( m \approx 1 \), (3.41) reduces to the approximate result (3.26). What we have here is an improvement on (3.26) that accounts for the crowding of dipoles in the medium, be it a gas, liquid, or solid.

We still haven’t relaxed the second assumption on page 40 that scattering from distant charges be incoherent, so (3.41) is unreliable for solids. Other simplifications mean

---

\(^6\)If \( \varepsilon \) is retained in this equation in place of \( m^2 \), as is more appropriate for electrostatic rather than optical applications, the equation is known as the Clausius-Mossotti relation.
3.4. SCATTERING IN HETEROGENEOUS MEDIA, AND THE RADIATION APPROXIMATION TO MAXWELL’S EQUATIONS

Figure 3.6: The real and imaginary parts of $n$ for liquid and ice water. Note the presence of a number of resonances where increases in absorption and regions of anomalous dispersion are indicated.

it is not always accurate for liquids either. For many substances, however (notably including H$_2$O), (3.41) works well over a huge range of $N$ at constant $\alpha$, even through the liquid-gas phase change. This permits calculation of changes of $n$ with changes of state. Figure 3.6 shows how water’s index of refraction varies with frequency in its condensed forms.

3.4 Scattering in heterogeneous media, and the radiation approximation to Maxwell’s equations

In the previous sections we considered the behavior of a small parcel of air bathed in an effectively uniform electric field. We now outline the nature of the scattering problem in its full generality. In the continuum treatment of radiation we are interested in scales large compared to $\lambda$. This means adding up the contributions from many parcels such as the one considered earlier, located at all positions within the medium. We will characterize the behavior of each parcel in terms of macroscopic quantities. This material is designed purely to acquaint the reader with the issues and make contact with basic electromagnetic theory, rather than for practical use, so it is sketched only very roughly. (This section may be omitted without loss of continuity).

A wave equation in a dielectric medium with dielectric constant $\epsilon_0$ can be written
using the macroscopic Maxwell’s equations that reads (Jackson, p. 419),
\[
\nabla^2 \mathbf{D} - \frac{\varepsilon_0}{c^2} \frac{\partial^2 \mathbf{D}}{\partial t^2} = -\nabla \times \nabla \times (\mathbf{D} - \varepsilon_0 \mathbf{E}) + S(\mathbf{H}).
\] (3.42)

The source term \(S(\mathbf{H})\) is zero for nonmagnetic media and will be ignored further. Substituting a wave solution where
\[
k = \sqrt{\frac{1}{\varepsilon_0}} \frac{\omega}{c}
\]

enables the wave equation to be written
\[
(\nabla^2 + k^2) \mathbf{D} = -\nabla \times \nabla \times (\mathbf{D} - \varepsilon_0 \mathbf{E}).
\] (3.43)

In this equation, the right hand side appears as a source of waves whose propagation is governed by the left-hand side.

If the R.H.S. “source term” is fully known, the solution of (3.43) is straightforward. The homogeneous part of the solution is just the incident wave:
\[
\mathbf{D}^{(0)} = D_0 e^{ik \cdot \mathbf{x}}
\]

which is determined by the boundary conditions. Here, \(\mathbf{e}\) is the (complex) polarization vector of the wave. There is also an inhomogeneous part of the solution obtained in the usual way at a position \(\mathbf{x}\) by integrating the Green’s function of the \(\nabla^2 + k^2\) operator,
\[
G(\mathbf{x}, \mathbf{x}') = \frac{e^{ik|\mathbf{x} - \mathbf{x}'|}}{4\pi |\mathbf{x} - \mathbf{x}'|},
\]
over all positions \(\mathbf{x}'\) weighted by the local source due to the deviations.

We now consider specifically a medium whose properties vary from one parcel to the next, but not greatly. Then \(\varepsilon \approx \varepsilon_0\) at each location. We write
\[
\varepsilon(\mathbf{x}) = \varepsilon_0 + \delta \varepsilon(\mathbf{x})
\] (3.44)

With this substitution, using (3.21), the combined solution to (3.43) is
\[
\mathbf{D} = \mathbf{D}^{(0)} + \frac{1}{4\pi} \int \int \frac{e^{ik|\mathbf{x} - \mathbf{x}'|}}{|\mathbf{x} - \mathbf{x}'|} \nabla \times \frac{\delta \varepsilon(\mathbf{x}')}{\varepsilon_0} \mathbf{D}(\mathbf{x}') d^3 \mathbf{x}'
\] (3.45)

The reader may observe (perhaps with a bit of squinting and imagination) that (3.45) resembles an integral form of the RTE (2.8) with emission and absorption left out. It may be regarded as the “missing link” between Maxwell’s equations (describing the full behavior of the fields) and our simplified RTE treatment of their behavior in the radiative limit. The most important difference is that (3.45) governs the field \(\mathbf{D}\), while the RTE governs radiance (proportional to \(D^2\)). As with (2.12), we do not actually know the R.H.S. without having already obtained the full solution!

The scattered field is given by linear contributions from all the locations in the medium. If we discretized the triple integral into a sum over volume elements, each element would behave as a discrete dipole (we would have to make sure that the volume
elements were small compared to $\lambda$ or the discretization would fail). Thus, we can reduce the problem of radiative transfer in a heterogeneous continuum to the problem we previously faced, of scattering by a large number of dipoles. This can be exploited to treat scattering in situations of arbitrary medium boundaries and heterogeneous properties by specifying an appropriate array of dipoles and performing a brute-force computation, in what is known as the discrete dipole approximation.

Aside from such numerical methods, solution of (3.45) usually requires the approximate approach of perturbation theory. If $\varepsilon$ is sufficiently close to unity or $\delta \varepsilon$ sufficiently small, scattering will be sufficiently weak that the field experienced by each parcel is approximately equal to the incident field. Going ahead and setting $\mathbf{D}$ on the R.H.S. equal to the incident amplitude, we make what is known as the single-scattering or Born approximation (most of this theory was first worked out in quantum mechanics). The total solution is then

$$\mathbf{D} \approx \mathbf{D}^{(0)} + \frac{k^2 D_0}{4\pi} \int \int \int \frac{e^{i(k-k') \cdot \mathbf{x}}}{|\mathbf{x} - \mathbf{x}'|} (\mathbf{e} \cdot \mathbf{e}') \frac{\delta \varepsilon(\mathbf{x}')}{\varepsilon_0} d^3 \mathbf{x}'$$

(3.46)

Note that if the medium is truly uniform (fluctuations $\delta \varepsilon$ are zero) there is no scattering, even though each of those volumes contains plenty of scatterers! This is not paradoxical from the continuum viewpoint, since we simply have a wave propagating undisturbed in a uniform medium. The explanation of this result from the viewpoint of contributions by individual scatterers is essentially that given at the end of Section 3.2. The scattering observed in “homogeneous” gases results from the Maxwellian distribution of random fluctuations in $N$ throughout the gas, which leads to fluctuations $\delta \varepsilon$ that produce results corresponding to those derived by assuming independent scatterers.

A final note concerns the second requirement noted on page 40, relating to the coherency of scattered power arriving from various points throughout the medium. If we calculate the scattered power leaving any part of the medium using (3.45) or (3.46), we must compute $|\mathbf{D}^2|$ which requires squaring a sum (integral) taken over all space. This yields a doubly infinite number of cross product terms. If the fields associated with contributions from different $\mathbf{x}'$ are randomly phased, the cross product terms will vanish on average and the total scattered power will equal that scattered from each region of the medium, just as we found in obtaining (3.17). This is the essence of our previous approximations (at the microscopic level), and the RTE approach (at the macroscopic level). If at either level scattering is not incoherent, then the discrete dipole or Born approximations or other methods of solving (3.45) must be followed in order to properly account for the contributions of all scattered waves in a heterogeneous medium. The field of radiative transfer may be defined as the study of those electromagnetic problems where distant scattering is incoherent and 2.12 is sufficient.

**Exercises**

3.1 Derive (3.11) from (3.10).

3.2 What can you say about $\omega_0$ in a metal?
3.3 Derive (3.12) from (3.4) and (3.11) by computing the work done during an oscillation.

3.4 Scattering from unbound electrons (those with no restoring force) is called Thompson scattering. This scattering is the main source of radiation from the solar corona, visible during an eclipse. Using the approach of Section 3.1, calculate the scattering cross section of a free electron. What “color” is Thompson-scattered light, if the incident light is white?

3.5 The dipole polarizability (3.8) included only the polarization induced by charge separation, working against the Lorentz restoring force. However, substances such as H₂O whose molecules have permanent dipole moments \( p_0 \) are also polarized by rotating the molecules into preferred orientations, working against the thermal tendency to homogenize the orientations. It can be shown from statistical mechanics that for the electrostatic (\( \omega \to 0 \)) case the additional polarization is

\[ \alpha_{\text{rot}} = \frac{1}{3} \frac{p_0^2}{kT} \]

Calculate the temperature at which \( \alpha_{\text{rot}} = \alpha \) at \( \omega = 0 \), assuming the dipole moment of H₂O is equivalent to that of opposite unit charges separated by 10⁻⁸ cm. How would you expect the resonant frequency of this type of polarization to compare with that for charge separation polarization?

3.6 Devise an arrangement of two dipoles that will scatter strongly in the forward direction, but not backward toward the source (given a known incident direction). How could you increase the sharpness of the forward peak by (a) adding more dipoles, or (b) changing their distance?

3.7 How would you expect the scattering by a dense monatomic gas to change if all of the atomic nuclei somehow acquired an extra proton? Why?

3.8 At \( \lambda = 12.0 \mu m \), \( m \) of liquid water is (1.1170.196). Calculate the “skin depth” of penetration of this radiation into water (the distance at which \( \tau = 1 \)), neglecting scattering. Compare this for \( \lambda = 0.65 \mu m \), where \( m = (1.331, 1.6 \times 10^{-8}) \).

3.9 Calculate the real and imaginary parts of the index of refraction of air of molecular weight 29 and density 1.2 kg m⁻³, as a function of \( \omega \) for \( \omega \ll \omega_0 \). Use a value for \( \omega_0 \) equivalent to \( \nu = 10^9 \) cm⁻¹; explain why your \( m_R \) differs from the actual value of 1.00023 at visible wavelengths.

3.10 By taking one of the arbitrary basis vectors for polarization to be in the scattering plane and the other perpendicular to this plane, demonstrate from (3.36) that the Rayleigh phase function must be as given by (3.37), and show that the component of scattered radiation with polarization perpendicular to the scattering plane is isotropic.

3.11 Find \( m \) of a two-component gas mixture as a function of \( N_i \) and \( \alpha_i \) of the individual components.

3.12 Use the Lorentz-Lorenz equation to estimate the index of refraction for water in its vapor and solid phases at \( \lambda = 10.7 \mu m \), given its liquid-phase value of \( m_{10.7} = \)
Assume the densities of vapor, liquid, and solid water are $10 \text{ g m}^{-3}$, $10^6 \text{ gm}^{-3}$, and $9.5 \times 10^5 \text{ gm}^{-3}$, respectively.
Chapter 4

Scattering by finite particles

We now consider scattering and absorption by an aggregate of molecules large enough to be treated as a chunk of dielectric medium. Applications include scattering from cloud and aerosol particles in planetary atmospheres or the interstellar medium, scattering in turbid surface waters. For liquid water, the most common dielectric material, \( \varepsilon_0 \approx 1.78 \) at visible frequencies. For simplicity, we will consider spherical particles. The insight gained from this carries over fairly well to a variety of realistic situations, and results for spherical particles are still widely used in many cases.

One discusses scattering of a spherical particle in terms of the particle’s size parameter

\[
x \equiv \frac{2\pi a}{\lambda} = ka
\]

where \( a \) is the particle radius. If the particle is not spherical, one can still define a size parameter using, for example, a radius corresponding to a sphere of equal surface area. Scattering changes qualitatively in the small, medium, and large size parameter regimes, which will be discussed in turn.

Not only is a sphere the simplest shape for mathematical treatment, but happens to be a very good approximation for liquid cloud droplets, one of the most important geophysical applications of scattering theory. Even for ice crystals that are far from spherical in shape, calculations based on spherical particles can produce radiative transfer results that are quite accurate as long as one is interested only in the aggregate transmission and albedo characteristics of a large number of crystals (Grenfell and Warren, 1999). Similar arguments may be expected to hold for small aerosol or sediment particles of arbitrary shape suspended in the atmosphere or ocean, or small impurities in any scattering medium.

4.1 Small spheres \((x \ll 1)\)

(The following derivation is adapted from Jackson 4.4.) We assume that a plane wave (1.3) is incident on a small sphere of radius \( a \). Since \( a \ll \lambda \), the problem reduces to the classical problem of calculating the static electric field in the presence of a dielectric
4.1. SMALL SPHERES ($X \ll 1$)

A sphere, subject to boundary conditions. The first step in doing this is to obtain general solutions to Maxwell's equations, written inside and outside the sphere (where the solutions will differ owing to the different dielectric constants). Then the solutions must be matched at the boundary, for example, fields parallel to the spherical surface must be continuous across the boundary. Also, the solutions must decay to just the incident wave itself far away from the sphere. The field solution that satisfies all these constraints must be the correct one. Implied in the solution is a charge redistribution that occurs within the sphere, which we won't explicitly be concerned with.

Maxwell's macroscopic equations are:

\begin{align*}
\nabla \cdot \mathbf{D} &= 4\pi \rho \\
\nabla \times \mathbf{H} - \frac{1}{c^*} \frac{\partial \mathbf{D}}{\partial t} &= 4\pi \mathbf{J} \\
\nabla \times \mathbf{E} - \frac{1}{c^*} \frac{\partial \mathbf{B}}{\partial t} &= 0 \\
\n\nabla \cdot \mathbf{B} &= 0
\end{align*}

The field $\mathbf{H}$ and, as mentioned previously, the electric displacement $\mathbf{D}$, are fields that include the response of the medium of propagation and essentially behave in the medium as the fundamental $\mathbf{B}$ and $\mathbf{E}$ fields do in vacuum (where, in fact, the fields become identical to their more fundamental counterparts). The charge and current densities are denoted $\rho$ and $\mathbf{J}$, respectively. In the $a \ll \lambda$ limit, the magnetic fields and currents need not be considered explicitly (we must also assume that the dielectric material is nonmagnetic).

The scalar potential $\Phi$ is defined such that

$$\mathbf{E} = -\nabla \Phi;$$

equation (4.3) guarantees that an electrostatic field can be written this way. By combining this with (4.2) we obtain Poisson's equation,

$$\nabla^2 \Phi = -4\pi \rho,$$

Which reduces to Laplace's equation where $\rho = 0$. In practice, aside from polarization which is already implicit in $\mathbf{D}$, the only charge density that will arise is what collects on the surface of the sphere as a response to the applied field. Thus, the potentials governing the electric fields inside ($\mathbf{E}_{\text{in}}$) and outside ($\mathbf{E}_{\text{out}}$) the sphere must each obey Laplace's equation. The fields must also blend at the spherical boundary according to the following conditions:

\begin{align*}
(\mathbf{D}_{\text{out}} - \mathbf{D}_{\text{in}}) \cdot \mathbf{n} &= 4\pi \rho \\
(\mathbf{E}_{\text{out}} - \mathbf{E}_{\text{in}}) \times \mathbf{n} &= 0
\end{align*}

where $\mathbf{n}$ is the unit normal to the local surface, and $\rho$ is the surface charge density. The condition (4.5) may be obtained by integrating (4.2) through a small pancake-shaped volume that surrounds a tiny patch of the surface of the sphere, while (4.6) is obtained from (4.3) by integrating around a small loop on the surface that encloses such a patch.
(if you’ve seen this in an E/M physics course, great; if not, understanding the basis of the final result is more important than understanding all the steps to get to it). The other boundary conditions of this problem are that at large distances from the sphere, the solution must just equal the incident plane wave; and that the solution must remain finite everywhere.

This problem is most easily solved in spherical coordinates aligned with the reference direction parallel to the \( \mathbf{E} \) field polarization direction in the incident wave. If Laplace’s equation is written in this coordinate system, orthogonal candidate solutions can be written in product form and superposed to match any boundary conditions (see Appendix?). For the present, axially symmetric case, the solutions reduce to

\[
\Phi_{\text{in}} = \sum_{l=0}^{\infty} \left[ A_l r^l + B_l r^{-(l+1)} \right] P_l(\cos \theta)
\]

\[
\Phi_{\text{out}} = \sum_{l=0}^{\infty} \left[ C_l r^l + D_l r^{-(l+1)} \right] P_l(\cos \theta).
\]

(4.7)

for \( r < a \) and \( r > a \) respectively. \( P_l(\cos \theta) \) are the Legendre polynomials. Also, writing (1.3) in spherical coordinates and using (4.4) we may establish the boundary condition at infinity as

\[
\Phi_{\text{out}}(r \to \infty) = -E_0 r \cos \theta.
\]

(4.8)

Since the field must be finite at the center of the sphere, we immediately have from (4.7) that \( B_l = 0 \) for all \( l \). From (4.8), we find that \( C_1 = -E_0 \) and all other \( C_l \) are zero thanks to our advantageous orientation of the coordinates. Substituting (4.7) into (4.5-4.6) leads to the result that \( A_l = D_l = 0 \) for \( l \neq 1 \); computation of the \( l = 1 \) terms leads to the solution

\[
\Phi_{\text{in}} = -\left( \frac{3}{2 + \epsilon} \right) E_0 r \cos \theta
\]

(4.9)

\[
\Phi_{\text{out}} = -E_0 r \cos \theta + \left( \frac{\epsilon - 1}{\epsilon + 2} \right) E_0 \frac{a^3}{r^2} \cos \theta
\]

(4.10)

Using (4.4), we find that inside the sphere the field is

\[
E_{\text{in}} = \frac{3}{\epsilon + 2} E_0,
\]

a constant field parallel to the applied field but weaker if \( \epsilon > 1 \).

The electrostatic potential created by a dipole \( \mathbf{p} \) is:

\[
\Phi_d(\mathbf{x}) = \frac{\mathbf{p} \cdot \mathbf{x}}{r^2} \cos \theta
\]

(4.11)

if \( \theta \) is again the angle between \( \mathbf{p} \) and \( \mathbf{x} \). Comparing this with (4.10), we see that the field or potential around the small sphere can be written as the superposition of that associated with the incident wave plus a contribution from a single electric dipole whose strength is

\[
\mathbf{p} = \left( \frac{\epsilon - 1}{\epsilon + 2} \right) a^3 E_0.
\]

(4.12)
In other words, the spherical particle behaves exactly like a dipole as discussed in the previous chapter, and a (sufficiently sparse) collection of them will cause Rayleigh scattering just like a gas! This is true only for small size parameters $x$. In light of (3.2) we see that the particle has a polarizability

$$\alpha_{\text{sphere}} = \left( \frac{\epsilon - 1}{\epsilon + 2} \right) a^3$$

(4.13)

Comparison of this polarizability with the Lorentz-Lorenz formula (3.41) also yields the interesting result that the earlier formula, though derived for extensive media, gives the polarizability of a sphere if one sets $N$ equal to the reciprocal of the volume of the sphere. Although (4.13) was derived using electrostatic theory, the use of complex values for $m$ and $\epsilon$ yields correct results for harmonically varying fields since the sphere is small compared to $\lambda$.

According to (4.13), particle polarizability is proportional to volume and mass of the particle. This agrees with the conclusions from Section 3.2, where polarizability was found proportional to the number of dipoles when the dipoles were clustered closely together compared to $\lambda$. It also means that particle scattering is proportional to radius to the sixth power. This very sharp dependence is the basis for rain radar, which uses sufficiently large $\lambda$ (in the range of a few mm to $\sim 10$ cm) to penetrate small particles with no difficulty, scattering only from larger precipitation particles. This is only one example of the ubiquitous importance of the great size sensitivity in the Rayleigh regime.

### 4.1.1 Mean field in a medium near a dipole

In Section 3.3.3 a formula was quoted for the average field within a spherical shell surrounding a particular dipole. Now that we have the results of the previous section in hand, that formula is not difficult to obtain. (This section may be omitted without loss of continuity.)

From (4.4) we have the volume integral within the sphere of radius $a$

$$\int_V \mathbf{E}(x)d^3x = - \int_V \nabla \Phi(x)d^3x$$

(4.14)

which can be converted to a surface integral

$$\int_V \nabla \Phi(x)d^3x = \int_S a^2 \Phi(x) d\Omega$$

(4.15)

where $n$ is the outward normal vector to the surface. Substitution of (4.11) and integrating yields

$$\int_V \mathbf{E}(x)d^3x = -\frac{4\pi}{3} \mathbf{p}.$$  

(4.16)

Notably, the integral is independent of $a$. Thus, all the contribution comes from a vanishingly small neighborhood around the dipole (hence the importance of the fact
that two dipoles cannot occupy the same position!) The sphere will contain on average $NV$ dipoles, so the mean field due to all dipoles in the sphere is

$$E_{\text{self}} = \frac{4\pi}{3} \mathbf{p} (NV)/V \quad (4.17)$$

or, using (3.18),

$$E_{\text{self}} = -\frac{4\pi}{3} \mathbf{P} \quad (4.18)$$

the desired result (which can also be obtained by setting $V$ in (4.16) equal to $1/N$). This line of reasoning involves some further subtleties which are discussed in Jackson 4.5.

### 4.1.2 Small, non-spherical particles

Shapes other than spheres have more complicated solutions for two reasons. First, the higher-order coefficients in (4.7) will not, in general, vanish; also, the solutions must be further expanded azimuthally if the scattering particle is not azimuthally symmetric with respect to the polarization direction (a circumstance that clearly will not hold for arbitrary polarization and particle orientation unless the particles are spheres).

Despite these problems, however, the behavior of small particles ($x \ll 1$) is for most shapes still dominated by the first dipole moment of the scattering particle, and this moment has the same scaling with radius as that of a sphere. The primary problem is to compute the polarizability for the shape of interest, and multipole moments if these are important. For small particles this is an electrostatics problem; dipole and multipole solutions for various shapes have been computed. Generally, one does not go far wrong by approximating small particles as spheres of equivalent surface area.

### 4.2 The “Mie regime” (medium $x$)

Our derivation of the scattered field of a very small sphere can be repeated without making the quasistatic approximation, and can then be applied to any size sphere. This was done by Lorenz and Mie around the turn of the 20th century, and in many respects remains the most useful calculation of scattering behavior at moderate size parameters. Consequently scattering at moderate $x$ is often dubbed “Mie scattering” or scattering in the “Mie regime.” Unfortunately the full solution is involved, and is not presented here. It is obtained as a series of Legendre polynomials that does not converge very fast for large size parameters, and becomes computationally burdensome. The first term in the series corresponds to Rayleigh scattering, and series convergence is very fast for small size parameters. Higher terms have no particular intuitive significance. The full solution is presented in Liou (2002), Hansen and Travis (1974), and others.

For scattering in this regime, one typically expresses the particle cross section in terms of a geometric cross section multiplied by an efficiency factor $Q$:

$$\sigma_s = Q(\pi r^2) \quad (4.19)$$
4.2. THE ‘MIE REGIME’ (MEDIUM X)

This efficiency depends only on the real and imaginary parts of $m$, and on $x$. It is shown for a few values of $m$ in Figure 4.1. For small $x$ the efficiency increases as $x^4$ because the amount of scattering varies as the square of the particle mass ($r^6$). As $x$ approaches one, $Q$ peaks near $x = 3-5$ at a value of up to 4-5 if $m_3 \approx 0$. Then a series of oscillations occur in $Q(x)$, which are due to a sequence of increasingly weak resonant modes on the surface of the particle. The values of $x$ at which these occur vary as $m_8$ is varied, since the modes are affected, but they are little affected by changes in $m_3$. On the other hand, the oscillations decrease in intensity as $m_3$ increases, damping all oscillations that involve field penetration into the particle. In most real situations (for example, in almost any cloud on Earth), one has a distribution of differently sized particles so that in the average behavior within a cloud element (a region small compared to the overall cloud size, but large enough to contain many droplets) the oscillations are smoothed away, leaving just a single broad peak in $Q$ near $x = 4$ with nearly monotonic decay toward $Q = 2$ at high $x$.

Corresponding to the resonant mode excitations evident in the $Q$ variations is a lobed behavior to the scattering pattern, with the number of lobes increasing as one traverses the oscillations in $Q(x)$.

At large size parameters, $Q$ begins to level off at a value of two. At first, it seems strange that this value should be other than unity. The reason for this is that diffraction (sometimes called Fraunhofer diffraction) of light rays around the particle contribute to the scattered energy. The integrated amount of scattered energy turns out to be equal, in the limit of small $\lambda$, to that intercepted by the particle itself, resulting in $Q = 2$. For this reason, single-scatter albedos are generally no less than 0.5 for large particles no matter what kind of material they are made of.

In the limit of large $x$, a very large amount of energy begins to go into an increasingly narrow, forward-scattering lobe. This lobe is due mainly to the diffraction around the particle, and as a result, contains about half the total scattered power at large $x$. The angular width of this lobe is determined by interference effects from the waves that diffract around either side of the obstacle: these will be in phase and will constructively interfere when $\Theta = 0$, but will become out of phase when

$$\sin(\Theta) \approx \Theta = \frac{0.5\lambda}{2a} = \frac{1}{4x}.$$  

The solid angle of this peak therefore decreases as $x^{-2}$. Since the total diffracted energy is independent of $x$, it must be true that the radiance scattered in the direction $\Theta = 0$ increases as $x^2$. This is observed in detailed Mie calculations carried out to large $x$. With a bit more work, we could obtain the shape of the forward scattering peak to very good approximation using diffraction theory.

All of the basic points made in this section apply to non-spherical particles to some degree. Analytical solution of Maxwell’s equations for arbitrary particle geometries has, in general, not been achieved, so a number of approximate methods must be used if the spherical assumption is deemed inadequate (though it is often adequate even for highly nonspherical particles). The most popular is anomalous diffraction theory, popularized by Van de Hulst (1963), which does a very good job of reproducing the gross behavior of Mie scattering from spheres, and can be applied easily to other shapes. This theory is essentially what we just used in the previous paragraph to predict the shape
and intensity of the forward scattering peak, except that away from this strong diffraction peak, energy that passes through the particle contributes significantly. To account for this, one can treat the passage of photons through the particle using geometric optics, but in doing so, keep track of their phases rather than completely ignoring the wave nature of the light. These rays are then blended with those that diffract around the particle, to produce the overall scattering pattern. Another approximation that has become more popular with the advent of computers is the discrete dipole approximation, briefly introduced in Section 3.4, in which the scattering particle is represented by a compact lattice of dipoles and field strengths are computed taking full account of dipole-dipole interactions.

This is an intentionally very brief discussion of scattering in the resonant regime. For further information, the reader is urged to consult the quite readable paper by Hansen and Travis (1974).
4.3 Geometrical optics (large $x$)

It may be anticipated from Section 3.2 that large agglomerations of scatterers will have an increasing tendency to behave like an infinite distribution in which all energy keeps going in the forward direction. Recall that when integrating over an infinite volume of regularly spaced scatterers (such as would occur in a solid or, to some extent, a liquid), the contributions from all scatterers canceled in every direction except the direction of the incident radiation, leading to transparency. The implication of this is that inside a reasonably homogeneous particle (e.g., one where there is relatively little scattering inside the particle compared to what happens upon entering or leaving it), whose dimensions are large compared to $\lambda$, photons travel in straight lines as light rays with relatively little dispersion. Then the rules of geometric optics—refraction and reflection from boundaries where there are discontinuities in $m$—apply.

The topic of geometric optics, and its application to rainbows, haloes, and other optical phenomena is well documented in introductory textbooks. It is assumed that the reader is familiar with the basics of refraction and reflection from dielectric interfaces. The reader is then referred to other texts (Liou 2002, Hansen and Travis 1974) for further discussions.

One noteworthy point will be made here that emerges from geometric optics and anomalous diffraction theory, concerning absorption and scattering in clouds of many dielectric particles. For $x > 2$ or so, the scattering cross section of particles varies roughly in proportion to the geometric cross section, that is, as $a^2$. However, the absorption by any given particle will be proportional to the product of the amount of intercepted light and the typical path length through the absorbing material, which implies an $a^3$ dependence. The result is that, at a given wavelength, the ratio of scattering to absorption, or single-scatter albedo $\tilde{\omega}_s$, increases as scattering particles become smaller. This trend doesn’t continue all the way to very small particles, however. Once we have $x < 1$ and begin to enter the Rayleigh regime, scattering falls off as $a^6$, while absorption continues to do so as $a^3$. Thus, at very small sizes the single-scatter albedo can approach zero.

4.4 Surface reflection and albedo

Finally we consider a chunk of dielectric that is large compared not only to $\lambda$ but to the medium itself; that is, reflection or absorption by a semi-infinite macroscopic solid or liquid surface such as a planetary surface. With some exceptions, most macroscopic surfaces scatter radiation roughly isotropically regardless of its incident direction. Surface roughness at the scale of $\lambda$ tends to produce all possible scattering angles; the probability of a photon escaping the surface without further scattering is roughly proportional to $\cos(\theta)$ for $\theta < 90^\circ$, leading to outgoing radiances that are isotropic in the backscattering hemisphere. Radiation scattered deeper into the surface topography by its first reflection ($\theta > 90^\circ$) must of course be scattered again (or absorbed). Of course, many surfaces include a specular reflection component; this includes surfaces that are smooth at scale $\lambda$ (e.g. wet or highly polished) and surfaces whose microscopic orientations have a preferred value.
If a surface scatters isotropically into the backward hemisphere, we speak of it as a Lambertian surface. This approximation is especially common when it is only desired to compute the net fluxes of radiation in the directions perpendicular to the surface. The ratio of reflected to incident flux of any surface is called its (hemispherical) albedo. The term bidirectional reflectance is used when discussing reflection into a specific angle of radiation coming from another specific angle; it is usually expressed as the ratio

\[ R(\Omega, \Omega') \equiv \frac{2\pi I(\Omega')}{F_0} \]

where \( F_0 \) is the flux into the reflecting surface carried by an incident beam from direction \( \Omega \), and \( I(\Omega') \) is the backscattered radiance. By this definition, \( R(\Omega, \Omega') = 1 \) for a Lambertian surface that reflects 100% of incident radiation.

One can define Lambertian scatterers and albedos for objects other than planar surfaces. An occasionally-used alternative is to define properties a propos of a spherical target object. A spherical Lambertian object scatters radiation isotropically into all 4\( \pi \) steradians; its spherical albedo is the total power scattered divided by the total incident power intercepted. Thus the bidirectional reflectance in backward directions is larger when calculated relative to spherical Lambert scatterer than when calculated relative to a Lambertial planar surface, by a factor of \( 2 \cos \theta_0 \) (where \( \theta_0 \) is the zenith angle of the incident light relative to vertical).

Example: wetness and darkness of surfaces

Have you ever noticed that wet objects become darker than when they were dry? This is quite a general property that is true for all soil types and most everyday surfaces. The effect is of recognized importance to the melting of open ice on planetary surfaces since meltwater ponds lower the albedo of the underlying snow or ice, increasing sunlight absorption and accelerating melting.

We can understand this property by considering a solid Lambertian surface covered with a coating of liquid water that is thin compared to the surface’s macroscopic curvature but thick compared to \( \lambda \) of the incident light. In this case, geometric optics applies. Due to water’s high \( n \approx 1.33 \), light is refracted toward the normal direction as it enters the water, according to Snell’s law (Figure 4.2). If the reflections from the solid surface were specular rather than isotropic, all light that reflected from the solid surface would encounter the liquid surface at the entry zenith and upon leaving the water would refract back to its original zenith direction. However, a Lambertian surface scatters in all directions, so that some of the reflected light will be scattered into angles too shallow to leave the water (below the critical refraction angle) and will experience total internal reflection back toward the solid. These rays must scatter a second time from the solid surface in order to be reflected, reducing the overall probability of reflection and therefore the albedo of the wet surface.

Exercises

4.1 Show that the radar backscattering coefficient (ratio of backscattered radiance to...
Figure 4.2: Refraction of light rays in the geometrical optics limit obeys Snell’s law: \( m_1 \sin \theta_1 = m_2 \sin \theta_2 \). Diagram shows \( m_2 > m_1 \). Lambertian scattering at lower boundary produces scattering in all directions, some of which cannot refract back into medium #1.

Incident flux at the target, in units of \( m^{-1} \text{ster}^{-1} \), can be written

\[
\beta = \frac{64 \pi^5}{\lambda^4} \frac{m^2 - 1}{m^2 + 2}^2 Z
\]

where \( m \) is the refractive index of the scatterer material and the “reflectivity factor” \( Z \) is defined

\[
Z \equiv \int a^6 n(a) da.
\]

\( n(a) \) is the scatterer number per unit volume per unit scatterer radius \( a \).

4.2 Derive (4.10).

4.3 Derive (4.16).

4.4 According to the boundary conditions (4.5, 4.6), one should be able to obtain the electrostatic field solutions for the situation of an empty spherical cavity embedded in an infinite dielectric easily from those of a dielectric sphere embedded in vacuum by substituting \( \epsilon \rightarrow 1/\epsilon \). Do this, and show that the field obtained in the cavity is closer to that in the dielectric than was the case in (3.39). Explain the difference physically (consider the implicit behavior of neighboring charges in the two calculations). (answer: \( \mathbf{E}_{in} - \mathbf{E} = \frac{4\pi}{2\gamma+1} \mathbf{P} \).

4.5 A corona is a diffraction pattern that appears around the “direct” solar beam when viewed through a cloud of small droplets all of similar size. Calculate (a) the angle to either side of the main beam where one will see the first maximum in intensity, as a
function of $x$ assuming $x \gg 1$; (b) the color of the light on the inside and outside of this maximum.

4.6 Rayleigh scattering produces blue scattered light. On some occasions (“once in a blue moon”), scattered light can be red and the direct beam can be left blue. What two conditions would have to be met for this to occur, and why is it so rare in Earth’s atmosphere? (Hint: see the numbers given in the next problem.)

4.7 Aerosol particles range in size from a few nm to a few microns; cloud and precipitation particles, from a few microns to a few mm. Identify what size range (Rayleigh, Mie, or geometric optics) corresponds to scattering from (a) aerosol particles and (b) cloud and precipitation particles at (i) visible, (ii) infrared, and (iii) radar (10 cm) wavelengths.

4.8 Write an expression for the albedo $\alpha_w$ of a wet surface whose dry albedo is $\alpha$, if the probability of incident light penetrating the liquid-air interface is $p_i$ for incoming light and $p_s$ for light that has scattered upward from the solid surface (note that these probabilities depend only on the index of refraction and the angular distribution of radiation, but require some work to calculate). Assume the water is nonabsorbent and the surface is Lambertian.
Chapter 5

Absorption and Emission

Absorption and emission are quantized processes usually involving jumps between states in atoms or molecules. Quantum theories are based on the wave nature of particles (photons and electrons, for example). The DeBroglie wave (designated $\Psi$) of any particle has a wavelength inversely proportional to the particle’s momentum $p$:

$$\lambda_\Psi = \frac{\hbar}{p}$$

(the photon is a special case with zero mass but nonzero momentum (1.5) due to its relativistic velocity.) The wave $\Psi$ for an electron or other nonrelativistic particle obeys an equation called Schroedinger’s equation (see Appendix A). Stationary solutions to this equation in finite systems occur only for certain DeBroglie wavelengths, resulting in what we call “quantum behavior” where energy and momentum are limited to discrete values associated with a set of states. Thus, the radiation by an atom or molecule is not continuous, as would be predicted by classical theory of a negatively charged electron accelerating centripetally as it orbits the positive nucleus; instead it is zero most of the time, except during a transition where a pulse of energy (photon) suddenly issues forth or disappears. An atom in its lowest-energy, or ground state can emit nothing.

The key constraint on emission is that the total energy of the initial state, plus that of the photon, must equal the energy of the final state. The relation (1.4) then ties this pair of states to a particular light frequency. We will see that the frequencies predicted by this constraint match the resonant frequencies predicted by classical theory, but include other frequencies besides those. Spectroscopic observations not only confirm quantum theory but form the basis for observational determination of many properties of molecules such as the distance between atoms, strength of forces connecting the atoms, etc. At the level of treatment here, classical and quantum theories are both helpful, although at a more advanced level quantum theory can stand alone.

For our purposes, the total energy of an atom includes the electronic binding energy and the translational energy. For molecules, we can add vibration and rotation energies. We do not consider the binding energy of the nucleus or nuclear reactions, which require “hard” radiation that is not naturally encountered in significant amounts. Of the four relevant energy reservoirs for a molecule, all except translational energy
are quantized. Photon absorption by a neutral particle (e.g. most molecules and atoms) requires that at least one of the three quantized states change. The requirement that the energy differences of all the transitions must exactly equal that of the photon may be written:

$$\Delta E_{\text{Elec}} + \Delta E_{\text{Vib}} + \Delta E_{\text{Rot}} + \Delta E_{\text{Trans}} = h\nu.$$  

The energy $\Delta E_{\text{Trans}}$ that goes into particle translation (kinetic energy) is not quantized, so if it comes into play it can enable continuum absorption\(^1\) anywhere within a spread of frequencies. Usually, this spread is tiny in a gas since particles interact weakly. Only if a moving particle carries a net charge can its translational energy absorb a photon by itself without the particle having to involved in a collision at the time. Continua are therefore important in ionized gases; they are also important in condensed matter since molecules are in constant proximity (water vapor produces some continuum absorption for more subtle reasons discussed in section 5.8). Otherwise, absorption is dominated by discrete lines in which discrete energy level changes balance almost exactly. The effect of the small $\Delta E_{\text{Trans}}$ values is to give the lines a small but finite width, in effect making each line a tiny continuum. Readers who want to know more about quantum mechanics may benefit from a (very) brief review in Appendix A.

The three relevant quantum transitions involve different typical energy ranges, although these depend on the temperature. Going from highest to lowest, at terrestrial temperatures we have

- electronic (ultraviolet, some visible)
- vibrational (infrared)
- rotational (far infrared and microwave)

Figure 5.1 shows schematically the energy levels associated with each of these; a transition moves the system from one level to another. These transitions can either be “bound-bound” transitions between two bound states or “bound-free” transitions where the bond is broken entirely. The latter is called ionization for an atomic (electronic) transition, where the electron is stripped from the atom, or photodissociation for the breaking of bonds in a polyatomic molecule (there is no bound-free rotational transition). Bound-free transitions produce continua, since the particles can carry an arbitrary amount of kinetic energy in addition to that required to break the bond. But this requires higher photon energies than the corresponding bound-bound transitions. Ionization and photodissociation are important to heating in Earth’s stratosphere and above, and are clearly important for atmospheric chemistry.

It is common for two quantum transitions to occur simultaneously. Because of the wide variation between typical energies associated with the three transition types, absorption bands result that are comprised of clusters of many closely-spaced lines. The center frequency of the band is determined by the more energetic transition, while distance from the center is determined by the energy of the less energetic transition.

---

\(^1\)“Continuum” absorption is not to be confused with the continuum approximation discussed in the scattering context, but refers instead to continuous behavior in the spectral dimension.
Figure 5.1: This figure will be a standard energy level diagram, but showing the widely-spaced electronic levels, more narrow vibrational levels broken out, and another breakout showing rotational levels, for a symmetric diatomic molecule where the "electronic levels" assume only one atom is excited.

The lower transition can either be upward or downward, leading to lines on either side of the center. Clearly, this greatly increases the number of lines possible in a given species: if there were three electronic transitions and 20 vibrational transitions possible, we would now have 120 possible dual transitions in addition to the individual ones. “Electronic” bands can be quite complicated since they can be coupled with both vibrational and rotational transitions at once. “Vibration-rotation” bands are critical in providing the infrared opacity of planetary atmospheres, and are somewhat simpler to understand.

Atoms and molecules are constantly jumping from one state to another due a variety of excitations. The most common excitation is by collision with another particle, or (in an ionized gas or plasma) acceleration by a nearby charge. Other possibilities are spontaneous emission of a photon (only works downward in energy), or absorption of a photon passing by (only works upward). The mean lifetime of a state until spontaneous emission occurs, given by quantum theory, is typically of order:

- electronic: $10^{-8}$ sec
- vibrations: $10^{-1}$ sec
- rotations: $10^0$ sec

By contrast, the typical lifetime due to collisional transitions is only $10^{-9}$ sec in the Earth’s troposphere (this becomes much longer toward Earth’s upper atmosphere or in more tenuous gases). Thus, in the lower and middle atmosphere, excited states—regardless of how they got that way—usually relax “thermally” rather than by spontaneous emission. This keeps the population of states close to that predicted by kinetic gas theory. Otherwise, Planck’s function would not accurately describe the emission,
because the state populations would be distorted by incoming radiation so as to cause anomalous emission resembling the absorbed component of the incident radiation. This problem does emerge in the Earth’s ionosphere, but is beyond our scope here. The time required for a quantum transition to actually take place is short compared to the lifetime of the state (except for electronic transitions where the two can be comparable), but long compared to the vibration or orbit period characterizing the state.

One of the fascinating features of quantum physics is that many of the quantum transitions that one might imagine are “forbidden” by so-called selection rules, and are observed to be quite infrequent (sometimes, totally absent) compared to allowed transitions. Such selection rules will be introduced in the following sections; they can usually be “explained” (to the extent that this is ever possible in quantum mechanics!) by invoking a fundamental physical principle.

5.1 Electronic transitions

Newcomers to this subject are probably most familiar with the quantized electronic structure of atoms (a mainstay of basic chemistry as well as quantum physics). Further, electronic transitions are the most energetic transitions of importance, and occur whether matter is organized into molecules or not. They are therefore a logical place to begin. Though electronic transitions occur in an individual atom, this includes atoms bound up in molecules! The presence of the other atoms does produce complications, which we defer for the moment.

For a single electron in stationary orbital state about a nucleus, the Hamiltonian is determined by a central electrostatic potential ($1/r^2$) of the positively charged nucleus. Solutions to Schroedinger’s equation are only possible when a discrete number of wavelengths fit around the orbit, i.e. for discrete energy values (eigenvalues of $H$) and their associated electron momentum values. The energy levels for the simplest case of the hydrogen atom are:

$$E_n = -\left(\frac{2\pi^2 m_e q_e^4}{\hbar^2}\right) \frac{1}{n^2}$$

where $n$ is the number of azimuthal nodes in $\Psi$. The energy absorbed in a transition from $n_1$ to $n_2$ is consequently

$$\Delta E = \left(\frac{2\pi^2 m_e q_e^4}{\hbar^2}\right) \left(\frac{1}{n_1^2} - \frac{1}{n_2^2}\right).$$

There are no selection rules governing these transitions, so any combination of initial and final $n$ can occur, though with more than one electron the initial and final states must obey the Pauli exclusion principle.

In heavier elements, electronic transitions become more complicated because of level splitting due to the electron orbital angular momenta. Transition energies for the outermost electron tend to be similar to, or somewhat greater than, those for hydrogen, since the extra nuclear charge is mostly shielded by the remaining electrons. Shielding is much less efficient for the interior electrons, typically preventing any except the outermost electron from seeing any action.
At terrestrial temperatures nearly all atoms are in their ground state, so the only transitions likely to be observed are occasional ones between the ground state and lowest excited state(s). These typically lie in the ultraviolet range for common atoms. At higher temperatures, particles have enough kinetic energy to access all the energy levels, so one sees a series of many lines corresponding to different combinations, known as “Rydberg” series. These include the well-known Lyman \((n_1 = 1)\) and Balmer \((n_1 = 2)\) series of hydrogen, and typically reach down through the visible parts of the spectrum. As temperatures rise further, the more closely-spaced, higher levels become populated causing the spectra to reach down to very long wavelengths. Also, many atoms become ionized and the Rydberg series starts filling in to become an ionization continuum.

This discussion implies that although absorption can occur at any temperature, electronic emission only becomes important at temperatures well above those in Earth’s atmosphere. As temperature increases, atoms can begin to attain states that have closer neighbors, enabling lower-energy photonic transitions. Meanwhile, the Planck radiation is shifting to higher frequencies and energies. Eventually overlap occurs, and electronic emission becomes significant. The particular temperature where this happens depends very much on the atomic number, since the outer electron of a heavier element starts out in the higher, more closely-spaced shells. Molecular bonds can also lead to closer level spacings when electrons are “shared.” An example is oxygen in Earth’s atmosphere, which possesses a near-infrared “A-band” at 0.76 \(\mu\)m. This transition between the ground state of O\(_2\) and one of its lowest two excited states is sufficiently strong to absorb a significant fraction of incoming solar radiation near the band center. Even this band is not, however, of sufficiently low energy to produce much emission in Earth’s atmosphere.

Because electronic transitions are weak at terrestrial temperatures, and atoms are typically found in molecules at these temperatures, lower-energy molecular quantum transitions (vibration and rotation) are the key to understanding absorption and emission in planetary atmospheres at infrared wavelengths. Further, when electronic transitions do occur, it can be concurrently with these others, forming electronic bands. We will return to this after discussing the lower-energy, molecular transitions.

### 5.2 Vibrational transitions

Diatomic molecules have one vibrational mode, while larger molecules have three or more. Each mode is independent and is most simply approximated by a harmonic oscillator, in which the restoring force is assumed to vary linearly with deviation from equilibrium, leading to a quadratic potential. The true potential for bound atoms deviates significantly from a pure quadratic at higher energies (Figure 5.2), even for the simplest case of a diatomic molecule. More accurate computations are possible by adding higher-order terms to the potential.

The harmonic oscillator is associated with a potential function \(V = k(r - r_0)^2\), where \(r\) is the distance from the center of mass, \(r_0\) the equilibrium value of \(r\), and \(k\) a
force constant. The classical frequency of vibration with two bodies is

$$\nu_0 = \sqrt{\frac{k}{\mu}},$$

(5.3)

with $\mu$ the reduced mass of the system ($\frac{M_1 M_2}{M_1 + M_2}$). The Schroedinger equation with a harmonic potential function again has a discrete series of solutions assigned with integer quantum number $v = 0, 1, \ldots$ where $v$ equals the number of radial nodes in $\Psi$. The vibrational state energies turn out to be

$$E_v = h\nu_0(v + \frac{1}{2}),$$

(5.4)

so that the spacing between adjacent states is just $h\nu$. In accord with classical theory, this implies that an absorbed photon’s frequency $\nu$ must equal the resonant frequency $\nu_0$ for transitions between adjacent levels. In real systems, the level spacing decreases somewhat with $v$ (Figure 5.2).

Note that isotopes of a given molecule will have the same force constant, but different reduced mass. Their resonant frequencies will therefore differ according to an isotope shift, $(\nu' / \nu = (\mu' / \mu)^{1/2})$, making possible their spectroscopic detection.
5.2. VIBRATIONAL TRANSITIONS

The selection rule

\[ \Delta v = \pm 1 \]

disallows the absorption of photons at higher harmonics of the resonant frequency. Classically, such transitions would not have been expected either, since the light frequency must match the resonant frequency. From the quantum viewpoint, the level-skipping transitions are ruled out only because while in the superposition state \( \Psi = a_1\Psi_{\text{initial}} + (1 - a)\Psi_{\text{final}} \), the electron probability distribution \( \Psi\Psi^* \) is symmetrical if \( \Delta v \neq \pm 1 \) thereby failing to produce the transient electric dipole needed to radiate or absorb.

Deviations of the true potential from a pure quadratic cause certain effects in the spectra of real molecules. For one thing, the so-called “hot” transitions between two excited states occur at somewhat lower frequencies than the “fundamental” transitions to/from the ground state, since the higher energy levels are spaced somewhat closer together. At terrestrial temperatures, “hot” transitions are much less common than fundamentals since most molecules are in the ground state. Another effect of anharmonicity is that the selection rule based on a harmonic potential does not hold absolutely, since asymmetry in the net attractive force (derivative of \( H \) with respect to \( r \)) allows nonzero (though small) transition probabilities for “forbidden” transitions with \( \Delta v \neq \pm 1 \). Transitions with \( \Delta v > 1 \) are called overtone transitions, since they typically lead to frequencies that are nearly integral multiples of the fundamental. Observed forbidden transitions include a weak CO\(_2\) line at 2.0 \( \mu m \).

The above points apply independently to each vibrational mode of a molecule. Multiple modes, present in polyatomic molecules, are designated by the quantum numbers \( v_1, v_2, \) etc. These can change individually or in combination yielding “combination bands,” but each mode is subject to the selection rule. An important note is that, in many molecules (including all diatomic ones), any vibrational transition is forbidden unless accompanied by a second, simultaneous transition of another type (see the next section). Though to a good approximation the modes are independent, this is not exactly true. Thus there is a slight energy difference between, say, the \( v_1 = 0 \rightarrow 1 \) transition when \( v_2 = 1 \) as opposed to the more usual case of \( v_2 = 0 \). Transitions of one mode where another mode is in an excited state are also called “hot” transitions.

Triatomic molecules have three vibrational modes named “(symmetric) stretching”, “bending” and “asymmetric stretching” (Figure 5.3) which are numbered \( v_1, v_2, \) and \( v_3 \) respectively. Linear triatomic molecules have two indistinguishable bending modes (for a total of four); this degeneracy gives rise to a fourth quantum number \( l \) (\( l \leq v_2 \)) which quantifies the angular momentum about the symmetry axis due to excitation of the two perpendicular bending modes in tandem (imagine a figure skater going into a spin while bent at the waist). \( l \) must change during an allowed bending transition in such molecules, so that these are called “perpendicular” transitions. The quantum number \( l \) cannot change during stretching mode changes, which are termed “parallel” transitions. This distinction between parallel and perpendicular transitions is important in its effects on the selection rules for rotational transitions, described next. Molecules with greater numbers of atoms have many vibrational modes, still designated \( v_i \).
A photon carries not only energy but also angular momentum $\hbar/2\pi$. Both of these quantities must be conserved among the photon and molecule during absorption or emission. Unlike energy, angular momentum has a direction (sign) to it which must be taken into account in the conservation. Since individual vibration modes store energy but not angular momentum, pure vibration transitions usually cannot conserve momentum and are forbidden. The exception is perpendicular transitions described above, where the angular momentum stored in $l$ can come into play. In other cases, angular momentum conservation demands that a vibrational transition always be accompanied by simultaneous rotational transition in what is called a “vibration-rotation” (V-R) transition. Since rotational energies are much lower than vibration energies, the V-R transitions occur at frequencies near, but not equal to, the vibration resonance $\nu_0$.

The rigid rotator assumption is the usual starting point for understanding quantum rotation behavior of molecules. The assumed independence between vibration and
5.3. **ROTATIONAL TRANSITIONS AND V-R BANDS**

rotation modes is not perfect but a good first approximation. For the rigid rotator,

\[ E_r = \frac{p^2}{2\mu r_0^2}, \]

with the angular momentum \( p = 2\pi \mu r_0^2 \nu_r \), \( \nu_r \) the rotation frequency. Here, Schroedinger’s equation has solutions with energy eigenvalues

\[ E = \frac{\hbar^2}{2\mu r_0^2} J(J + 1), \]

with \( J = 0, 1, \ldots \), the total angular momentum quantum number. This is usually written

\[ E = cBJ(J + 1) \tag{5.5} \]

with \( B \equiv \hbar^2/2\mu r_0^2 c \), so that \( B \) has wavenumber units and is typically quoted in cm\(^{-1}\).

The selection rule for rotational transitions is

\[ \Delta J = \pm 1. \]

In quantum theory this is again derived from overlap integrals of \( \Psi \), but physically it can be thought of as resulting from the conservation requirement for angular momentum. Note that \( J \) can increase or decrease by a single unit regardless of whether the event is an emission or absorption, since a photon’s angular momentum can be aligned either positively or negatively with respect to that of the molecule. Again, \( \Delta J = 0 \) is possible only for perpendicular V-R transitions, where the photon’s angular momentum appears in the quantum number \( l \) rather than in \( J \).

According to (5.5), the energy spacing between levels \( J \) and \( J + 1 \) is \( 2BJ \). The selection rule leads to transition energies proportional to the lower of the two quantum numbers involved. All the transitions together produce a band structure (Figure 5.4). The band has two wings or *branches*, labelled P and R, for rotation transitions that lower or raise the net transition energy, respectively. If the band corresponds to a “perpendicular” \( v_2 \) transition in a linear molecule, there is a very strong absorption line right at \( \nu = \nu_0 \) (i.e., at band center) labelled the Q branch; otherwise, there is a gap at the band center (reminiscent of a missing tooth) with no absorption. At terrestrial temperatures, the wings typically spread out for quite some number of lines away from the center (i.e., to high values of \( J \)), and in fact usually peak at some distance away from the Q branch. This is because the state populations peak at \( J \) values significantly greater than one, in contrast to the vibrational situation.

Pure rotation transitions (not involving changes in vibration) are not possible unless the molecule has a permanent electric dipole moment, for example \( \text{H}_2\text{O} \). A pure rotation spectrum, when it exists, may be thought of as a P branch centered on zero frequency. The pure rotation spectrum of \( \text{H}_2\text{O} \) extends into the infrared at terrestrial temperatures and is quite important to the Earth’s greenhouse effect. The V-R and R spectrum of water vapor is so complex, due to the asymmetric nature of the molecule, that it is often approximated theoretically as a random distribution of lines (Section 5.10.3).
Figure 5.4: Modeled parallel V-R band from Goody and Yung. The missing Q-brand corresponds to the fundamental vibration frequency at 1285 cm\(^{-1}\). Smaller contributions to the band from isotopes are also visible.

The idealized rotation band described above is in reality distorted by a number of complicating factors. First, since real molecules are not rigid but are more accurately modeled as masses connected by springs, they will “stretch out” as they rotate faster, increasing their moment of inertia. The result is greater level spacings as \(J\) increases, which appears as a slight widening of the line spacing away from the band center. Second, there is interaction between vibration and rotation motions, leading to slight anomalies in rotation spacing that increase in hot and overtone V-R bands. Third, different isotopes possess slightly different line spacings due to their differing moments of inertia. The distance between the atoms is essentially unchanged, so the line spacing is inversely proportional to the reduced mass of the isotope (closer spacing for heavier isotopes). Finally, rotational energies are slightly affected by contributions to the total system rotation from the electrons themselves, causing each “line” in a V-R band to appear as a multiplet of lines, whose number increases with molecular weight.

Remember also that a fundamental V-R band will exist for each vibrational mode or combination, and that each will potentially produce any number of weaker “hot” and/or overtone bands in addition to the fundamental. Hot bands appear as “ghost” bands shifted somewhat downward (usually) in energy from their fundamental counterparts. An illustrative, if somewhat extreme, example of a single fundamental V-R band (mode #4 for this particular molecular structure) in an octahedral molecule (SF\(_6\)) is presented in Figure 5.5. The band shows the expected structure of a P, Q, and R branch. The presence of a strong Q branch indicates that this vibrational mode, like the bending mode in a triatomic molecule, breaks molecular symmetry in a way that allows the photon angular momentum to be stored without a simultaneous rotational transition. In addition to the main Q-branch a number of “hot” Q-branches appear
toward lower wavenumbers; these blend together at the FTIR resolution. The individual V-R lines are labelled according to $\Delta J$. It is clear that in between each of these lines falls a weaker line belonging to a “ghost” band; this is one of the hot bands in which the sixth vibrational mode is in its first excited state during the transition of the fourth mode. This is only the strongest of many hot bands, others contributing to the remaining “clutter” between the lines of the fundamental band. Interestingly, a heavy sulphur isotope also produces as a weak band centered near 612 cm$^{-1}$. Finally, Figure 5.6 shows a higher-resolution spectrum obtained with a tunable laser, in which a single “line” in the fundamental band is expanded to reveal that it actually consists of some 12 components. A given transition in a single molecule can fall into any of these components, depending on what combination of angular momentum states are occupied by the electrons in the atoms of that particular molecule. The structure shown in Figures 5.5-5.6 may seem overwhelming and, indeed, we will leave further study of such details up to the spectroscopists.

### 5.4 Electronic bands

As complex as V-R spectra can be with all the variants on the fundamental transition, electronic bands are even richer in complexity, and will not be discussed in as much detail. The first added complication is that the electronic transitions are coupled with
Figure 5.6: The same spectrum as in Fig. 5.5, taken with a tunable diode laser at $10^{-5}$ cm$^{-1}$ resolution.
both vibrational and rotational transitions, leading to three-way combinations. Moreover, any vibrational transition is allowed if it is coupled to an electronic transition. This allows all the overtone transitions to come to full fruition, as well as allowing the symmetric stretching mode to participate in linear molecules, both forbidden in ordinary V-R transitions.

If a substance made of diatomic or polyatomic molecules is at sufficiently high temperature that excited electronic states are significantly populated, it is typically hot enough to populate many excited vibrational states too, leading to a great multiplicity of possible energy combinations. These are generally described by spectroscopists as progressions, identified with either the vibrational quantum number of the upper (\(v'\)) or lower (\(v''\)) state. The set of bands involving a given \(v'\) (say, 2) in the higher-\(n\) state, and any possible \(v''\), is designated the \(v' = 2\) progression for that pair of \(n\); conversely, the set of transitions involving a given \(v''\) in the lower electronic state is referred to as the (e.g.) \(v'' = 1\) progression.

Although selection rules are absent for electronic-vibration transitions, not all possible combinations are equally probable. Typically, certain combinations of \(v_1, v_2\) and \(n_1, n_2\) will allow \(\Psi_1, \Psi_2\) to overlap particularly well (for example, the equilibrium interatomic distance in the old and new configurations happens to be very nearly the same). Such “path of least resistance” combinations will often dominate the spectrum, concentrating most of a given electronic transition into a single V-R band. This is why classical series like the Rydberg series are often clearly visible in an observed spectrum as a sequence of relatively narrow features.

### 5.5 Free-free opacity

The interactions above, which all involve transitions to or from a bound state in a molecule or atom, are those of dominant importance in planetary atmospheres. In ionized gases, another form of emission is possible: free-free or Bremsstrahlung emission, due to interactions between charged, free particles. This is essentially a classical effect, in which the charged particles are accelerated by applied fields and interact with neighboring charges (resulting in absorption) or are accelerated by neighboring charges (resulting in emission). The energy gained or lost from the radiation field is stored in the kinetic energy of translation of the particles. The charged particles may be ions or free electrons. Free-free emission contributes significantly to the solar continuum, and to absorption in Earth’s ionosphere.

### 5.6 Blackbody radiation and equilibrium, revisited

It may confuse the beginning student that the complicated and spiky absorption spectra due to quantum transitions can be consistent with the smooth spectrum of radiation expected from a so-called black body. The solution to this puzzle lies in distinguishing between the nature of an equilibrium state and the path toward that state.

From the discussion in Chapter 1, we know that an equilibrated field of radiation develops in a mass of absorber and that this radiation field possesses a smooth spec-
tral distribution. The surprising thing is that the Planck distribution is independent of the properties of the material. One way to see that this is so follows from Kirchoff’s law. Imagine inserting a different kind of material into a previously single-species gas. Whatever the new material’s absorption properties are, its emission properties must match them wavelength by wavelength. When equilibrium is achieved, the invading material must emit photons at each \( \lambda \) at the same rate it absorbs them, which means the (isotropic) radiation field will be undisturbed by the new material.

So what about the complicated spectra? The spectral features of a particular substance determine how fast the equilibrium is reached, as a function of frequency. If the substance cannot absorb at a particular frequency, then radiances at that frequency will never reach local equilibrium, and the photon density will always be determined by fluxes passing through the material from outside. At highly absorbing wavelengths, a small amount of absorber will reach equilibrium very quickly. When peering into a body from the outside, one sees photons that were emitted from greater depths as the opacity at the particular wavelength decreases. Even inbetween the farthest-spaced absorption lines, opacity is not absolutely zero, so in a sufficiently massive body the Planck intensity (subject to emissivity, scattering, and temperature heterogeneity effects) will always be seen.

One can therefore observe the absorption spectrum of a material only in non-equilibrium situations. As it happens, these are the rule rather than the exception and species-specific features appear in almost any measured spectrum. If sunlight is observed passing through a thin substance, the substance’s absorption features appear; if a thin substance is emitting in front of a darker background then emission lines appear. Similarly, spectral features appear whenever the emitting medium’s temperature changes with depth into the medium, with absorption features appearing if temperature increases with depth.

Statistical mechanics tells us that the population of states of molecules in LTE falls off exponentially with ratio of state energy to temperature (1.10), tailing off rapidly at energies greater than \( kT \). This is \( 2.35 \times 10^{-2} \) electron volts (eV) at 273 K, for example, corresponding to a photon wavelength of 53 \( \mu \)m. The peak in the 273 K Planck spectrum actually occurs at a wavelength shorter than this (near 10 \( \mu \)m) due to the influence of the \( v^2 \) factor in the Planck formula—in other words, the Planck spectrum wants to maximize at energies where the medium’s ability to radiate or absorb is already in steep decline. This causes continuum opacity within an absorber to decrease toward the shortwave end of its equilibrium emission spectrum. The effect is noticeable in hot stars, whose continuum opacity decreases at higher wavelengths, allowing one to see deeper into the star. Absorption in cooler stars and planetary atmospheres is dominated by bands of lines, however. Wherever these band happen to fall in the thermal spectrum is where opacity will be greatest.

### 5.6.1 Non-LTE effects

In formulating the radiative transfer equation, we assumed that photons could be added to a radiation stream only through Planck emission or conservative scattering from other directions. In some cases, sources exist that cannot be represented by either of these processes. Two well-known examples are flourescence, and Earth’s aurorae. A
less well-known example is Raman scattering. We avoid the complication of adding the physics to our equations necessary to represent these phenomena, because in planetary applications they are generally not important energetically. However, nonequilibrium radiative and nonconservative scattering effects deserve a brief discussion, since they may come up in other contexts, and in fact may be useful in remote sensing applications (plus they are interesting!).

Florescence occurs when atoms are excited by incident radiation of sufficient energy to induce an electronic jump of more than one level. Some of the atoms will decay to an intermediate state inbetween the original and excited states, and some of these decays will be through photon emission. The emitted photon will be at a wavelength longer than the photon that stimulated the atom to begin with (a well-known example is the emission of visible light by fluorescent paint illuminated in the ultraviolet). Aurorae occur by a closely related process; the difference is atoms are “pumped up” into excited states by collisions with highly energetic particles from the solar wind, rather than by photonic excitation. Since the emitted wavelength(s) are not necessarily present at all in the incident radiation, it is obvious that neither process can be treated as conservative scattering; and since the emission is not directly related to the temperature of the medium, it is not ordinary thermal emission either.

Both of these phenomena occur because state populations are distorted away from the Boltzmann distribution (1.10) by some outside influence. This influence is, in effect, radiation or matter that is much “hotter” than the medium. As the extra excited atoms/molecules decay back to lower states, they emit extra radiation compared to that predicted in (1.12) on the basis of Boltzmann state populations. This is what is referred to as “non-LTE” emission. LTE holds only when all constituents are at the same temperature, which is true in Earth’s atmosphere up to altitudes of about 70 km, and is true in most of the sun.

LTE can also fail if the ambient radiation field differs sufficiently (in an absolute sense) from the Planck spectrum to affect state populations significantly. Based on the mean lifetimes quoted earlier in this chapter, it would appear that in parts of the spectrum dominated by vibrational and rotational transitions, this would be nearly impossible. For electronic transitions, however, the emission probabilities are only about an order of magnitude smaller than those for collisions even at one bar of pressure. Thus, at low densities and high temperatures, it is possible for electronic state populations to become rather sensitive to the radiation field.

Note that the presence of strongly non-equilibrium photon distributions does not automatically mean that LTE does not hold. This is only the case if absorption of the anomalous radiation causes state transition rates significant compared to those arising from other excitations. For example, sunlight traveling through Earth’s atmosphere presents no problem below the mesosphere even though solar radiation is spectrally almost totally different from Earth’s own emission spectrum. The intensity of sunlight is simply too weak to compete with collisions as a source of excitation, until low pressures are reached. Scattering, of course, is very important anytime radiation differs significantly from black-body radiation at the local temperature, either through its spectral distribution or anisotropy. This is a far more common situation than non-LTE, occurring anytime a cold object is illuminated by a hot one.

When non-LTE effects are important, accurate calculations require that one ex-
plicitly model all important transition probabilities and state populations, a very challenging enterprise. For this reason there has been a tendency, even in astrophysical problems where non-LTE effects are significant, to give up and just use the Planck function.

5.6.2 Nonconservative scattering

Though most scattering is conservative, as we have so far assumed, an interesting non-conservative scattering process called Raman scattering can occur. It is somewhat like fluorescence, except that it is a one-step rather than a two-step process with no interim, quasi-steady state. A photon is scattered in a way that leaves the molecule in a different V-R state than that in which it started out, such that the photon frequency is raised or lowered accordingly. Lines appearing at frequencies lower than the incident photon are called Stokes lines, with the higher frequency ones called anti-Stokes (this terminology arises from the so-called Stokes’ rule of fluorescence dictating that fluorescent radiation must be at lower frequency than the stimulant radiation). Raman scattering produces yet another type of “ghost” spectrum, and was first observed through its production of such a ghost of the solar absorption spectrum, through Raman scattering from O₂. Vibrational Raman scattering produces very weak anti-Stokes lines at terrestrial temperatures compared to the Stokes ones, since most molecules start out in the ground state and cannot provide extra energy to the photon.

From a classical standpoint, Raman scattering is not a resonant phenomenon (indeed, the photon must have a much higher frequency than the particle resonance). Instead, the process is more like a Doppler shift. Recall that the selection rules for absorption/emission in a vibrational transition required \( \Delta v = \pm 1 \), due to the need for the transition state to have a dipole moment. Raman scattering does not require a transient dipole moment, but instead requires that the particle polarizability must fluctuate as the oscillation takes place. This condition yields the same \( \Delta v = \pm 1 \) selection rule; but more importantly, it means that diatomic molecules, which are inactive in the infrared but dominate Earth’s atmosphere, can Raman scatter. For rotational changes, consideration of the transition wavefunction leads to the selection rule \( \Delta J = \pm 2 \), different from that for absorption/emission.

The angular distribution of Raman-scattered radiation is the same as that for Rayleigh scattering, though at much reduced amplitude. This makes Raman scattering an interesting hybrid of classical and quantum-type phenomena. Raman scattering is useful in some remote sensing applications as a measure of the mass of diatomic gases, and causes the gaps in the solar spectrum to be partially “filled in” by the time radiation transits Earth’s atmosphere.

5.7 Solar emission

The “effective radiating temperature” of the sun is about 6,000K. Its spectrum resembles that of a blackbody at this temperature, except that emission at the highest wavelengths is significantly reduced, and a few other distortions are visible (figure 5.7). The spectrum deviates from a blackbody partly due the lowering of emissivity by scattering
5.7. SOLAR EMISSION

Figure 5.7: The measured flux distribution arriving at the Earth’s surface (lower solid curve) and top of atmosphere (upper solid). A black body spectrum at 6,000K is also shown (dashed line). The black areas show radiation removed during transit through the atmosphere by molecular absorption; the remaining difference between the upper and lower curves is due mainly to atmospheric Rayleigh scattering. (From Petit 1951.)

in the outer layers (especially at shorter wavelengths), and partly due to nonuniform solar temperature. The thin outer layer responsible for most of the emission that escapes the sun is called the photosphere. Temperatures within the sun vary considerably, from about 30,000K deep in the interior, to a minimum near 4,000K in the outer photosphere, rising again in the (even thinner) chromosphere and attaining over a million degrees in the very tenuous corona. The optical depth of the corona is about $10^{-10}$, so its radiation is undetectable against the background, except during a total solar eclipse. LTE holds well in the sun except in the chromosphere and corona regions, where scattering also becomes important.

Not surprisingly, electronic transitions are more important in stars than in terrestrial atmospheres. Stars such as the sun are composed almost entirely of hydrogen and helium. The 13.6 eV necessary to ionize hydrogen becomes attainable at temperatures of 5,000-6,000 K and above, which includes most of the sun. Consequently, the sun is a partly ionized plasma. Also, all the excited states are attainable in the photosphere and the full hydrogen series is visible in most stars. They appear as absorption lines, because the photospheric temperature decreases outward. The excitation and ionization potentials for helium are higher, so that temperatures must be at least about 10,000K to ionize this element. Because of this, helium lines are only visible during solar eclipse (which observation led to the naming of this element). In hotter stars, helium lines are present in the photospheric (main) emission. Temperatures in the corona are so high that coronal emission spectra indicate that heavier elements (called “metals” by astrophysicists) have lost most of their electrons, a key clue in revealing the extreme

---

2This cannot be due to radiative heating, but is believed to be due to upward energy transport in dynamical waves.
temperatures there.

Continuum emission is dominant in the sun and most stars. What may be surprising is that the most important continuum-generating mechanisms in the solar photosphere is not due to the ionization of neutral atoms or the free-free scattering of positive ions or free electrons, though these each occur. The main source is actually due to formation and destruction of the negative hydrogen ion \( \text{H}^- \). This ion forms when a neutral hydrogen atom loosely attaches a free electron, a union made possible by the relatively strong instantaneous dipole moment of the neutral hydrogen atom, plus the healthy supply of free electrons (this is reminiscent of water vapor dimer absorption, also a surprise player in determining the infrared opacity of Earth’s atmosphere). The energy of an \( \text{H}^- \) is substantially closer to that of neutral hydrogen than is that of the ordinary ion \( \text{H}^+ \), causing \( \text{H}^- \) to contribute far more to solar opacity. In stars much hotter than the sun, the ionization potential of neutral H is more easily reached and free-free continuum effects begin to dominate owing to the greater ionization. In cooler stars, the scarcity of electrons and ions causes ordinary atomic transitions to dominate. Bound-bound transitions are also clearly visible as absorption features in solar output, revealing much about stellar composition, but are of secondary importance to the overall opacity of most stars.

The binding energies for most molecules common to stars (the most familiar example is \( \text{H}_2\text{O} \)) are several eV. These energies are easily attainable at temperatures of \( \sim 5,000K \) and above, so that molecules break up easily and are rare in stars. The sun contains molecules in its cooler, outer photosphere, but only in low concentrations. Molecular spectra are useful for constraining stellar temperatures, since molecular concentrations are quite temperature-sensitive.

### 5.8 Planetary atmospheres

As discussed in Chapter 3, gases at terrestrial temperatures are effectively transparent to visible radiation. Thus, the molecular composition of atmospheres is relevant mainly to their internal transport and emission of infrared radiative energy. The planets have quite different atmospheres as will be seen below, although visible opacity is always dominated by clouds (or other aerosols), while infrared opacity is dominated by carbon compounds and/or oxides, and clouds. Some of the planets (Mercury and Pluto), and most moons, do not have real atmospheres although molecules will always be floating about due to sputtering of the surface.

In analyzing atmospheric spectra, the concept of brightness temperature \( T_b \) is often employed. This is simply the temperature of a blackbody that would account, via (1.12), for the observed radiance \( I \). In other words,

\[
B_\nu(T_b) \equiv I_\nu.
\]

Obviously, \( T_b \) is a function of \( \nu \) (unless the object being viewed really is a black body!) that increases monotonically with \( I_\nu \). We will see later that an observed value of \( T_b \) roughly conveys the local temperature at a position within the medium where the optical path to the observation point is unity. Typically (though not always) temperature
increases downward into planetary and stellar atmospheres, so that opacity is low at frequencies where $T_g$ is high and vice-versa.

### 5.8.1 Earth

Earth’s atmosphere is 78% N\textsubscript{2} by volume, followed by O\textsubscript{2} (21%) and a series of trace gases. The two main species are essentially inactive in the infrared, being symmetric, diatomic molecules. Oxygen, however, dissociates in the atmosphere and recombines to form ozone (O\textsubscript{3}), a trace gas that absorbs ultraviolet and infrared radiation. Water vapor is the next most abundant constituent after O\textsubscript{2}, and is the most important infrared absorber in Earth’s atmosphere. H\textsubscript{2}O and O\textsubscript{3} share the characteristic of being polar molecules with permanent dipole moments, endowing them with pure rotation spectra plus a plethora of V-R bands. Their polar character also probably facilitates the water-vapor continuum, discussed below.

After water vapor, the important greenhouse gases in Earth’s atmosphere are CO\textsubscript{2}, CH\textsubscript{4} (methane), O\textsubscript{3}, NO, CO, and CFC’s. Clouds are also very effective infrared absorbers, and are essentially black to all infrared radiation if they contain any more than about 50 g/m\textsuperscript{2} of water (equivalent to a layer of water 50 m thick).

Visible opacity arises almost entirely from aerosols and clouds, with essentially no absorption and only a small amount of Rayleigh scattering from gases. Clouds absorb very little, but scatter light effectively, elevating the planetary albedo to about 30%. Aerosol particles, where present, may absorb significantly at a variety of wavelengths depending on their composition; dust and smoke are good absorbers, while sulphate and sea-salt aerosols (which absorb water vapor from the air like sponges, forming haze) tend to behave more like cloud droplets and have a predominantly scattering effect.

Though very little visible absorption occurs, much of the invisible solar output is absorbed by the atmosphere. Ultraviolet light is absorbed (primarily in the stratosphere and above) by electronic transitions of various species, and by photolysis of molecules (mainly O\textsubscript{2} and O\textsubscript{3}). Many of the constituents active in the terrestrial spectrum, particularly water vapor, also absorb near-infrared sunlight (significant energy arrives from 0.7 out to about 3 microns). Hence it is important to distinguish between “visible” and “solar” radiation.

Water vapor and CO\textsubscript{2} are sufficient to make Earth’s atmosphere opaque at most energetically important wavelengths outside the visible range; exceptions, where the atmosphere is reasonably transparent, are critically important for remote sensing and for the planetary energy budget and are called spectral windows. The most prominent window is one from about 8-12 microns, near the peak of Earth’s emission spectrum. This window is not entirely transparent due to the narrow ozone feature at 9.6 μm; also, water vapor produces weak absorption throughout the window at low latitudes where vapor concentrations are highest, the so-called water vapor continuum. This continuum is not fully understood, but it is somehow caused by relatively strong and/or long-range interactions between an H\textsubscript{2}O molecule and its neighbors, discussed further in Section 5.9. Chloro-flourocarbons also absorb in (or “dirty”) this window, making it possible for them to contribute significantly to Earth’s greenhouse effect despite their tiny concentrations.
CHAPTER 5. ABSORPTION AND EMISSION

H2O
pure rotation 0-1000 cm\(^{-1}\)
\(v_2\) 6.3 \(\mu\) m 900-2400 cm\(^{-1}\)
\(v_1, v_3, h\alpha v_2\) 2.7 \(\mu\) m 2800-4400 cm\(^{-1}\)
over tone/combination 4500-11000 cm\(^{-1}\)

CO2
\(v_2\) 15 \(\mu\) m perpendicular
\(v_3\) 4.3 \(\mu\) m parallel

CH\(_4\)
\(v_4\) 7.7 \(\mu\) m 1306 cm\(^{-1}\)
O\(_3\)
\(v_1, v_3\) 9.6 \(\mu\) m

Table 5.1: Short list of key infrared absorption features in Earth’s atmosphere. All vibration bands listed are fundamentals of the indicated mode. Each band absorbs over a range of frequencies (right column), but is commonly named according to a central wavelength (middle column).

For a more thorough description of the gases and state transitions involved in Earth’s atmosphere, see Sections 3.2.1, 3.2.3 and 4.2.1 of Liou.

5.8.2 The other inner planets

Venus

Venus is often referred to as Earth’s “sister planet” due to its similar size and nearest-neighbor status in the solar system. Reasons for why its atmosphere turned out so different from Earth’s continue to be a subject of active research and interest, especially in the context of the “runaway greenhouse” idea discussed later.

The Venusian atmosphere consists almost entirely (97.5%) of CO\(_2\), with the only other significant component being N\(_2\). Consequently the dominant spectral feature is the 15 \(\mu\)m band (Figure 5.8). The surface atmospheric pressure on Venus is 92 bars, making concentrations of CO\(_2\) near the surface more than \(10^6\) times as high as on Earth. At these concentrations, every conceivable transition involving the CO\(_2\) molecule and its isotopes becomes active (in fact, there is no way to approach in laboratory settings the optical paths of CO\(_2\) that occur naturally on Venus, making Venus our best laboratory for understanding weak transitions in this gas). Venus’s CO\(_2\) is nearly sufficient to make its atmosphere totally opaque throughout the infrared (seeFigure 5.9). Those few gaps that remain in the spectrum are mostly filled in by the trace species H\(_2\)O in the mid-infrared and and SO\(_2\) in the mid and far infrared; CO, HCl and COS may also contribute to opacity here and there.

Venus has an overcast cloud layer at the top of its troposphere that is also opaque in the infrared. The clouds are composed mainly of sulfuric acid droplets that condense out at temperatures of around 200 K. Due to the enormous greenhouse effect of all these
Figure 5.8: Observed brightness-temperature spectra of three inner planets (From de Pater and Lissauer, fig. 4.6)
Figure 5.9: Variation of infrared spectrum of CO$_2$ with increasing optical path. $x$ axis goes from 2350 to 550 cm$^{-1}$, $y$ axis shows transmittance. Path length through the gas was 3.3 m (except 5th fig., 31 m), with CO$_2$ gas at pressures of 13 mb, 130 mb, 5 bar, 10 bar and 20 bar, at standard temperature. Hot bands near 1000 cm$^{-1}$ and combination bands near 1900 cm$^{-1}$ emerge at sufficient optical path, while a Fermi resonance near 1350 cm$^{-1}$ emerges at high pressure. From Chackerian and Boese, *Workshop on V-R Spectroscopy for planetary atmospheres*, 1980.
infrared absorbers, the surface temperature is over 700 K, and the cloud-forming temperatures do not occur until altitudes of ~60 km with a dry-adiabatic temperature lapse rate. Even though less than 2% of Venus’ atmosphere lies above its cloud deck, that portion still contains orders of magnitude more CO₂ than Earth’s entire atmosphere, and obscures the cloud deck at most infrared wavelengths.

The cloud layer is thick by terrestrial standards, similar in optical depth to that of the most severe storms on Earth. Its opacity to solar radiation gives Venus a Bond albedo of 75% and allows only about 2.5% of incident sunlight to reach the Venusian surface on average, according to measurements made by the Soviet Venera landers 8-12 deployed during the 1970’s (most of which lasted less than an hour in the heat and pressure). The only radiation that can penetrate these clouds and the gaseous atmosphere is at wavelengths near 10 cm, where emitted radiances roughly correspond to 700 K Planck values and provided our first clue as to surface temperature.

**Mars**

The Martian atmosphere is also composed mostly of CO₂ (about 95%), followed by the inert gases N₂ and Argon. The main difference from Venus is that the surface pressure is only about six millibars. This nonetheless adds up to a CO₂ “column mass” (mass per unit area of planetary surface) roughly 50 times that of preindustrial Earth. Other radiatively active gases are CO (of which Mars also has about 40 times the column mass as Earth), and small amounts of H₂O.

Mars lacks significant cloud cover, but does experience occasional, intense dust storms. The dust absorbs sunlight, heating the tenuous atmosphere and driving further
dust lofting, in a positive feedback that makes the storms long-lived and severe. Another interesting feature of the Martian atmosphere is that significant amounts of the main constituent, CO$_2$, condense out during the winter at high latitudes, causing large seasonal fluctuations in total atmospheric mass (this is why the above numbers were approximate).

5.8.3 The gas giants

The gas giant planets (Jupiter, Saturn, Uranus, and Neptune) have compositions similar to the sun. The main constituent, hydrogen, accounts for between 85% (Uranus and Neptune) and 96% (Saturn) of the total gas volume. The remained is each case is mostly Helium.

The radiative transfer in the gas giants is relatively complicated. As on Earth, the dominant species contribute little to opacity, at least in the outer part of the planets. Opacity in the infrared is dominated by methane (CH$_4$); ammonia (NH$_3$) makes some contributions here, but is more important at microwave and radio wavelengths. Deeper within the gas giants, high pressures (~20-80 Mbar) and temperatures (~10,000 K) cause the gases to adopt new thermodynamic phases. For example, H$_2$ dissociates at pressures greater than 1 Mbar into singlet H, and when compressed a bit further becomes “fluid metallic hydrogen” in which free electrons can flow as in a metal. This leads to continuum absorption throughout all but the outermost layer of the planets, much as one finds in the Sun.

Water vapor is present in significant concentrations deep within the gas giants. Though a dominant greenhouse absorber on Earth, H$_2$O makes little contribution to the greenhouse effect on gas giants because it condenses out at lower altitudes than the other gases. This condensation forms clouds in regions of ascent, as on Earth. Clouds are in fact the main thing we see when we examine the planets with visible telescopes, most famously on Jupiter where the banding structure associated with multiple atmospheric jets was easily discernible in early ground-based observations. In addition to H$_2$O, the clouds are composed of CH$_4$, NH$_3$, and H$_2$S, which condense at temperatures as low (50-100 K) as those found at the tropopause of the giants (near 100 mb, as on Earth). The highest clouds are probably made of condensed ammonia on Saturn and Jupiter, and methane on Uranus and Neptune.

5.9 Line shape and broadening

As one might expect, there are limits to how exactly the photon energy must match that of the transition. Recall that for an individual scatterer we calculated a classical, wavelength-dependent scattering coefficient (3.8), which had a natural line width $\delta \omega_N$ about the classical resonant frequency where excursions of the oscillating dipole

---

3 The information on gas giants, and much of that on the inner planets, comes from De Pater and Lissauer, *Planetary Sciences*, which the reader is urged to consult for further information on all aspects of the solar planets.
became very large. Since then we have found that quantum theory also predicts absorp-
tion at these resonant frequencies. Here we discuss the absorption line shapes from a
quantum viewpoint.

According to Heisenberg’s uncertainty principle, the uncertainty in the energy of a
state must be at least $\frac{\hbar}{\Delta t}$ where $\Delta t$ is the occupation time of the state. This implies
a similar uncertainty in the transition energy difference, implying a finite width to the
absorption line. In the Earth’s atmosphere, decay of an excited state due to collision
with another molecule is generally far more likely than spontaneous decay by emitting
a photon. For this reason, collisions make the lines much broader than their “natural”
line widths $\delta \omega_N$ (the latter set by the rate of spontaneous decay).

Both photon emission and molecular collisions in a gas can be modeled success-
fully as Poisson processes, where the probability per unit time of an event (emis-
sion/collision) is independent of how long it has been since the last event. We then
deduce that both state-altering processes produce the same line shape. It turns out that
this is just the shape already derived classically (3.32) for natural lines, known as the
Lorentzian profile:

$$f(\nu) = \frac{\delta \nu_L / \pi}{(\nu_0 - \nu)^2 + \delta \nu_L^2}. \quad (5.6)$$

This profile is accurate to a few percent for collisional broadening of most species.
Though the collision-broadened line has the same profile as the unbroadened natural
line, a new Lorentzian width $\delta \nu_L$ is specified, where $\delta \nu_L \gg \delta \nu_N$. Kinetic theory
predicts that the time between collisions in a gas varies as $\sqrt{T/p}$, which implies that
for a given absorption line,

$$\delta \nu_L \propto pT^{1/2}. \quad (5.7)$$

Pressure is the more important variable since it varies more than temperature through
an atmosphere and it appears at a higher power. This dependence leads to use of the
terms “pressure broadening” and, less often, “temperature broadening.”

When collisions are sufficiently rare, another source of line broadening emerges,
the Doppler effect associated with the different velocities of the individual molecules.
If each molecule is emitting or absorbing at the same frequency $\nu_0$ in its own frame of
reference, photons observed by a stationary detector will have a spread of frequencies.
Molecular velocities are distributed $\sim \exp(-au^2)$, which with the Doppler relation-
ship $\nu - \nu' = \frac{\nu_0}{c}$ gives lines a Gaussian shape

$$f(\nu) = \sqrt{\frac{\delta \nu_D}{\pi}} \exp \left( -\frac{(\nu - \nu_0)^2}{\delta \nu_D} \right). \quad (5.8)$$

if they are broadened purely by Doppler effects. This is approximately true in the
Earth’s ionosphere, for example, where density is low but molecular velocities can be
high. If neither collisional nor Doppler broadening is dominant by itself, more elabo-
rate line width calculations may be required (the most popular is the “Voigt” approxi-
mation) that possess characteristics of both line shapes. In the Earth’s atmosphere this
hybrid situation usually occurs in or above the stratosphere, but for some absorption
lines it can occur in the upper troposphere. Below this, Doppler broadening is relatively
unimportant.
It is common to describe an absorption line as the product of two terms, the “strength” $S$ and the “shape” $f(\nu)$, defined by

$$k_\nu = Sf(\nu) \quad (5.9)$$
$$S = \int k_\nu d\nu \quad (5.10)$$
$$\int f(\nu) d\nu = 1 \quad (5.11)$$

The expressions (5.6) and (5.8) obey the normalization condition (5.11). A pure Lorentzian or Doppler line can be fully characterized by quoting $S$, $\nu_0$, and $\delta\nu_L$ or $\delta\nu_D$. Note that $\delta\nu_L$ depends on temperature and pressure, and is determined by molecular dynamics of the medium; $S$ depends on temperature, but not on pressure, and is determined by the quantum physics of the absorbing molecule and the molecular quantum state populations; $\nu_0$ is a quantum characteristic of the molecule that is independent of the state (temperature and pressure) of the medium.

The strength $S$ of a transition is determined by two factors: the population of the initial state, and the degree of overlap of the quantum wavefunctions $\Psi_i$ and $\Psi_f$ describing the initial and final states. The degree of overlap is a fundamental property of the molecule that depends only very weakly on the state of the gas due to effects such as (for example) centrifugal forces associated with rotation that can slightly affect vibrational states. The initial state population, however, is strongly dependent on $T$ according to (1.10). Thus, $S$ of a given line is temperature dependent and is greatest when $T \sim E_i/k$ ($E_i$ the energy of the initial state).

### 5.9.1 Absorption by a single line

The falloff of absorptivity with distance from the line center is rather slow for a Lorentzian shape compared with, say, the Doppler (Gaussian) shape which decays exponentially away from the center. For this reason the absorption in the “wings” of Lorentzian lines plays an important role in radiative transfer. These wings will still be present even if densities are low enough for Doppler effects to determine the line shape closer to its center. Here we consider the properties of a single Lorentzian line in isolation, with Doppler effects ignored.

We begin by defining the mass path $u$ as

$$u \equiv \int \rho ds \quad (5.12)$$

which has units of mass per area and may be thought of as the total mass encountered by a beam of radiation having unit cross section. Note that $u$ is independent of $\nu$. In this discussion we will also neglect scattering ($k_\nu = e_\nu$). Then from (2.11) and (2.15), we have

$$\tau_\nu = \int_0^u k_\nu du' \quad (5.13)$$

which is just $k_\nu u$ if the medium is homogeneous.
5.9. LINE SHAPE AND BROADENING

Setting \( \nu = \nu_0 \) in (5.6), we see that at the line center, the absorption is proportional to \( S/\delta \nu_L \). In the wings \( ((\nu - \nu_0)^2 \gg \delta \nu) \), however, the absorption is given by

\[
k_\nu = \frac{S \delta \nu_L}{\pi (\nu - \nu_0)^2}.
\]

The heavy wings can add up to substantial absorption away from any line centers, especially at higher pressures (note that \( \delta \nu_L \) and \( u \) are each proportional to \( p \) for isothermal gas variations, at constant \( ds \)).

The cumulative effect of the wings of many lines can lead to a type of continuum absorption. A good example is water vapor, which produces the dominant absorption in the 8-12 micron “window” region of Earth’s atmosphere, even though no strong \( H_2O \) lines exist there. The strength of this continuum is still uncertain, since the interactions that give rise to it are not well understood. The behavior far from line centers is determined by very high-energy, long-range or long-lasting collisions. High energy collisions involve such close encounters between molecules that the multiatomic particles can no longer accurately be treated as point masses and subtle quantum effects become important. Coulomb interactions between polarized \( H_2O \) molecules may be important too, and it is known that water vapor “dimers” often form where two \( H_2O \) molecules loosely attach to one another for a while before parting ways again. This greatly increases the types of absorption possible during the time the dimer exists. The formation of dimers is expected to scale with the square of water vapor concentration.

The total absorption by a single line (including its entire tail according to the Lorentzian profile) is given by

\[
A_{\text{line}} = \int_{-\infty}^{\infty} [1 - e^{-\tau_\nu}]d(\nu - \nu_0).
\]

We assume for the time being that no other lines are nearby in the spectrum. The dimensionless parameter

\[
x \equiv \frac{S u}{2 \pi \delta \nu_L}
\]

may be defined, which is equal to one-half the optical depth at the center of the line as may be found from (5.6) and the definition of \( S \). Note that the quantity \( A_{\text{line}} \) actually has the dimensions of frequency or wavenumber, and represents the width of a spectral interval containing the amount of energy that the line absorbs. For this reason, the quantity \( A_{\text{line}} \) defined in (5.14) is called the equivalent width of the line.

We may consider two cases, the optically thin \( (x \ll 1) \) and optically thick \( (x \gg 1) \) limits. In the thin limit, the factor in brackets in (5.14) may be approximated by \( \tau_\nu \). This, using (5.9) and (5.13), renders the integral trivial with the result

\[
A_{\text{line}} =Su.
\]

Thus, for optically thin paths, the total absorptance is linearly proportional to the path length and to the line strength, but is not changed by line broadening due to pressure changes. This is known as the “linear limit.”

In the thick (saturated line) limit, we recognize that the quantity in brackets will assume values significantly different from unity only for \( |\nu - \nu_0| \gg \delta \nu_L \). In parts
of the range of integration where this is so, we have \( k_\nu \approx \frac{S \delta_\nu}{\ln(1 + \nu u)} \). Closer to the line center this approximation becomes poor, but then the integrand in (5.14) becomes insensitive to \( k_\nu \) so this does not significantly affect the result. Thus (5.14) may be integrated to yield

\[
A_{\text{line}} \approx 2 \sqrt{S \delta_\nu u}. \tag{5.17}
\]

We see that greater broadening of lines leads to more absorption in the optically thick case, though it did not do so in the optically thin case. Conversely, greater optical paths or line strengths have less effect on total absorption when the line becomes saturated. This is known as the “square-root” limit.

When one is between these limits, the solution of 5.14 can be expressed more generally in terms of a superposition of Bessel functions (see Liou). The absorption by a line always increases as it is broadened (for example, by increasing pressure), and as the line strength and \( u \) increase. The variation of \( A_{\text{line}} \) with \( u \) is customarily called the “curve of growth” of the line.

### 5.10 Absorption over finite spectral bands

Actual substances typically possess thousands of absorption lines within any broad part of the spectrum (e.g., over all wavelengths at which terrestrial or solar blackbody radiation is significant). Furthermore, there are many gases to contend with in a typical atmosphere or other emitting substance. As we saw earlier, many of the lines occur in V-R band systems. Each substance will typically have only a few important band systems. Since Lorentzian lines have broad wings, lines overlap so that the total absorption will be less than one would obtain by adding up the total absorptions calculated for each line in isolation (this is obvious if the equivalent width happens to be greater than the spacing between lines, which would have led to absorbance of over 100%!).

Several “band models” will be considered here for estimating transmission in bands of many lines. We simplify the problem by assuming that we deal with a spectral interval small enough so that both the Planck emission (and/or incident radiation) is nearly constant across it, and that the properties of the lines are (statistically at least) homogeneous across the interval. However, the interval must be large enough to contain many lines. In practice, this is typically called a narrowband calculation. Entire V-R bands are often too wide for these conditions to be adequately satisfied, however, in which case a wideband calculation must be performed, either by stitching together narrowband calculations or by attempting an approximate solution using average conditions within the band.

#### 5.10.1 Grey model

The simplest treatment of a spectral region is to ignore the presence of lines and compute bulk gas properties that represent averages over the spectral band of interest. Then the monochromatic RTE may be used by substituting these properties for the usual monochromatic ones. This procedure is known as the grey medium approximation,
since in effect a medium of complicated spectral properties is replaced by an “equivalent” greybody over the spectral interval. The greybody transmittance is simply

$$\mathcal{T} = \exp(-u\bar{k})$$

where $\bar{k}$ is an appropriately chosen bulk absorption coefficient over all frequencies. The grey emissivity $\varepsilon = 1 - \mathcal{T}$ fully describes the absorption and emission of the layer.

This approximation is not a real band model per se but a way of avoiding a band model. It is typically applied as a crude wide- or broadband approximation for primitive or pedagogical calculations. Though often useful in this regard, it fails to capture a key feature of absorption in most real situations: that great variations in $k_\nu$ with frequency allow some photons to penetrate a layer much more readily than others. This produces heating and transmission properties that no grey model can accurately reproduce except in special cases, as will be seen below.

### 5.10.2 Regular band model

The first of two band models to be considered here is the regular band model. This was proposed by Elsasser to calculate the total absorption in the 15 micron CO$_2$ band, which possesses the evenly spaced lines typical of V-R bands. In this model one assumes an infinite sequence of identical, evenly spaced lines separated by a frequency spacing $D = \Delta \nu$. The absorption coefficient at a particular frequency is then given (dropping the “Lorentzian” subscript $L$ on line widths, though still assuming shapes to be Lorentzian) by

$$k_a(\nu) = \frac{S}{\pi} \sum_{j=-\infty}^{\infty} \frac{\delta \nu}{(\nu - jD)^2 + \delta \nu^2}. \quad (5.18)$$

Each term in this series is the contribution from one line. The sum can be expressed in terms of hyperbolic functions; the transmittance turns out to have the form of a special function, named after Elsasser, whose exact form may be found in many texts.

It simplifies in the limits of widely-spaced and crowded lines:

$$\mathcal{T} \approx \exp\left(-\frac{Su}{D}\right) \quad (D \ll \delta \nu) \quad (5.19)$$

$$\mathcal{T} \approx 1 - \text{erf}\left(\frac{\sqrt{\pi u} \delta \nu}{D}\right) \quad (D \gg \delta \nu) \quad (5.20)$$

For lines that are crowded together ($D \ll \delta \nu$), optical properties become smooth in the frequency domain and we recover the grey approximation with $\bar{k} = S/D$. The case of widely spaced lines ($D \gg \delta \nu$) is more relevant at terrestrial pressures. Here we may consider further the weak line limit $Su\delta \nu \ll D^2$, in which (5.20) reduces to

$$\mathcal{T}_{\text{band}} \approx 1 - \frac{\sqrt{\pi Su} \delta \nu}{D} \quad (5.21)$$

In this limit, the band transmittance behaves like that of noninteracting, saturated lines, with square-root dependence on all three parameters (although quantititively slightly more transmissive per line).
It is at first surprising that even in this limit, one does not obtain an overall transmission approaching the linear limit shown by single weak lines. This perplexing result comes from the assumption of an infinite band, and from the Lorentzian profile with its heavy tails. Bands with a finite number $N$ of lines will approach the linear limit \( T = 1 - NSu \) as $S \to 0$, but will depart from the infinite-band theory only for extremely weak lines. The fact that the asymptotic results differ illustrates the importance of the line wing overlap and continuum absorption in real band systems. Since we are presently unable to characterize this continuum accurately from theory, what is usually done in practice is that empirical continuum corrections are applied to lift or depress the continuum absorption in the model until it agrees with data. These corrections are all the more necessary when the finiteness or heterogeneity of real bands (in contrast to the Elsasser theory) is considered.

### 5.10.3 Random band model

Another interesting model for bands of lines is the random model of Goody (1952). He noticed that the spectra of many gases (particularly $H_2O$) are so complicated that they appear random. It turns out that with advantageous randomness assumptions, a model is produced that is both accurate in calculations and computationally tractable. With an average of $n$ lines per unit frequency interval, the assumptions are: (1) lines are positioned randomly and independently within any spectral interval with uniform probability density

\[
P(\nu_0) = \frac{n}{\Delta \nu}
\]

with $n$ constant; (2) line strengths are drawn randomly from a Poisson probability distribution,

\[
P(S) = S^{-1} \exp(-S/S)
\]

that is independent of $\nu_0$. The mean transmittance of radiation within a spectral interval $\Delta \nu$ due to line $i$ is

\[
T_i = \frac{1}{\Delta \nu} \int_{\Delta \nu} d\nu \int_0^\infty P(S)e^{-ku}dS
\]

(5.22)

where $k_\nu$ depends on the difference $\nu - \nu_0$. Since the random variable $\nu_0$ is uniformly distributed over all frequencies, the integral is independent of $\nu$. Also, since the line positions and strengths are assumed independent, the transmissivities are independent for each line and the total transmittance is simply $T = \prod T_i$ or

\[
T = \left[ \frac{1}{\Delta \nu} \int_{\Delta \nu} d\nu \int_0^\infty P(S)e^{-ku}dS \right]^n.
\]

(5.23)

This can be integrated to give (e.g., Liou, 4.5.3)

\[
T = \exp \left[ \frac{Su}{D} \left( 1 + \frac{Su}{\pi \delta \nu} \right)^{-1/2} \right]
\]

(5.24)

where we have defined a mean line spacing $D \equiv \Delta \nu/n$ analogous to that for the regular model.
While in the regular band model the key parameter was line spacing, here the key parameter is line breadth. If \( \delta \nu \gg S_u \) (broad weak lines), the transmittance reduces again to that of the grey approximation with \( \bar{k} = S/D \) which was also approached by crowded bands of regular lines. In the opposite limit \( \delta \nu \ll S_u \),

\[
T \approx \exp\left(-\frac{\sqrt{\pi} S_u \delta \nu}{D}\right) \tag{5.25}
\]

Considering further the limit of weak or widely spaced lines \( S_u \delta \nu \ll D^2 \) yields the same limiting transmittance obtained for the regular band model in this limit. However, the error function decreases much faster even than an exponential, indicating that with large paths and widely-spaced lines, the transmission becomes much smaller in a regular band than in a random band with the same lines. This is physically reasonable since randomly spaced lines will tend to leave more gaps in the spectrum that radiation can sneak through, as compared with regularly spaced lines. If the lines are so thick that spaces are crowded out even in the random model, then the two models give equivalent results.

Figure 5.11 shows the transmission vs. the equivalent greybody optical depth \( S_u/D \) for each of the three aforementioned calculations. In the nongrey models the flux into the layer decreases more rapidly at first, then more slowly at significant optical depths. Even a much thinner grey layer (dotted line in the figure) does not possess the same qualitative behavior and allows significantly less transmission at greater depths than the other models. The band model behavior is recognizable as the superposition of decaying exponential solutions (2.18) with a broad range of decay rates. More heating is concentrated very close to the surface, but also at significant depths, than in the grey case where heating occurs mostly near unit optical depth. These characteristics may be physically interpreted as rapid absorption of all photons near line centers, followed by the relatively unfettered propagation through the layer of photons at frequencies far from line centers (especially in the random model). This behavior is responsible for Earth’s thermosphere, a very warm outer layer above 50 km that is also found in other planetary atmospheres but would not be predicted in the grey approximation. Nongrey effects are also responsible for many other phenomena. Heating rates will be discussed in greater detail in Section 7.3.

### 5.10.4 \( k \)-distributions

The main differences between the three models just considered can be ascribed to differences in the breadth and shape of the opacity distribution. In the grey model, \( k_\nu \) always has the same value and therefore obeys a delta distribution. The distribution becomes much broader in the Elsasser model, and develops longer tails in the random model due to the appearance of large gaps between lines (thus lower values of \( k \) than previously occurred) as well as the appearance of very strong lines leading to a tail at the high end. The extreme values of \( k \) are key in producing that model’s results. As long as the medium is homogeneous, it makes no difference which frequencies match up with high or low values of \( k \); the only thing that matters is how often each value of \( k \) occurs within the band, and how much incident radiation (or Planck emission) there
Figure 5.11: Transmittance vs. equivalent grey optical depth $S_{\nu}/D$ of three band models. The regular and random model results are shown for two different ratios of $\delta \nu$ to $D$ (the grey result is independent of this ratio). The first grey result assumes $\bar{k} = S/D$; a second grey calculation is shown (dotted line) with $\bar{k}$ one fifth of this.

is integrated over all pieces of the spectrum corresponding to any given $k$ value. This leads naturally to the k-distribution approach to band calculations, in which a distribution of $k$ is specified without regard for photon frequencies per se except to compute the total energy associated with each $k$ value. This method is simple, but must be modified for heterogeneous media leading to the more complicated correlated $k$ methods. See Liou 2002 for further details.

5.10.5 Observed band characteristics in Earth’s atmosphere

Observations of the spectrally-averaged transmittance $\tau^f$ over a broad swath of the spectrum (say the Planck curve for a 300 K black body) usually do not show a particularly rapid approach toward 100% absorption (H$_2$O is notably somewhat of an exception and, at path lengths perhaps an order of magnitude higher than those found in Earth’s atmosphere, could serve as a good black body due to continuum absorption). This is because lines are not distributed randomly over the entire spectrum, but are clustered in bands with large windows left in between that remain open even at very large $u$. Despite this, it is possible to achieve accurate computation of radiation by breaking the spectrum into intervals and treating these independently. The random model is widely useful in this regard. In practice, model parameters (e.g., $S$ and $\delta$) can be tuned to agree with observed dependence of band absorption on path length, which is easier and (in some cases) more accurate than trying to specify parameters from first principles based on spectroscopy.
In most gases (notably CO$_2$ at terrestrial quantities), the band structure of lines produces a logarithmic behavior of $\mathcal{F}(u)$ over many orders of magnitude of $u$. Thus, each doubling of such a gas in the atmosphere produces an approximately constant addition of infrared opacity regardless of the original concentration. In the past this logarithmic dependence was sometimes used by “climate change skeptics” to discount the importance of increasing CO$_2$ in the atmosphere, but this casual dismissal of CO$_2$ change is not born out by quantitative calculation. The main uncertainty in estimating terrestrial climate sensitivity to CO$_2$ comes not from poor knowledge of the radiative behavior of the gas, but from our inability to predict confidently how the concentrations of other radiatively active constituents may respond as climate changes.

**Further reading:** An old but very good reference for the theory and history of molecular spectroscopy is the classic work of Herzberg (1939).

### Exercises

**5.1** Given that the ionization potential $E_{\text{ion}}$ of hydrogen is 13.6 eV, what would the (theoretical) $E_{\text{ion}}$ of helium be if the other electron (a) fully shielded the second positive nuclear charge, or (b) were absent. Explain why the actual $E_{\text{ion}}$ for helium is the highest of all elements.

**5.2** Explain why (a) N$_2$ and O$_2$ are inactive in the infrared at terrestrial temperatures, and (b) why H$_2$O has a more complicated spectrum than CO$_2$.

**5.3** Describe how a typical V-R band will change as temperature is increased.

**5.4** For each of these light sources, indicate whether the observed spectrum should be smooth Planck emission, a smooth spectrum with absorption features, or a spectrum with emission features: (a) incandescent light bulb (if the glass is transparent), (b) fluorescent bulb (ditto), (c) moonlight (observed above Earth’s atmosphere), (d) sodium lamp, (e) aurorae.

**5.5** Roughly how does the equivalent width of a CO$_2$ line in the 15 $\mu$m band change with pressure in Earth’s atmosphere?

**5.6** CO$_2$ is uniformly mixed through most of inner planets’ atmospheres, and its lines are saturated in the 15 $\mu$m band. Will there be any difference in how far (in mass path $u$) thermal radiation emitted from the surface penetrates upward, in contrast to similar downward emission from the top of atmosphere (say, from a thick cirrus cloud)? Why?

**5.7** Show that $x$ given by (5.15) is equal to one-half the optical depth at the center of a line.

**5.8** Show that (5.17) holds in the thick-line limit. (Hint: you should obtain an approximate integral that can be done by letting $x \equiv \nu - \nu_0$, integrating by parts, and then making the substitution $y = 1/x$.)

**5.9** Derive (5.19) and (5.20).
Chapter 6

Practical Solution of the Radiative Transfer Equation

At this point we suppose that the parameters appearing in equation (2.12) have been calculated or measured according to the physics described in Chapters 3-5. Now our task is to solve this difficult equation, which will require a variety of approximate techniques chosen for the particular situation. First, one must recognize what processes need to be included in the calculation.

6.1 When is scattering important?

Absorption is usually important in radiative transfer problems, but scattering or thermal emission can often be neglected in making quantitative calculations. Emission, of course, is negligible at sufficiently short wavelengths when the medium is cold enough compared to the temperature of incoming radiation; in Earth’s atmosphere this is the case below \(\sim 3\mu\text{m}\), which includes nearly all of the incoming energy. Gaseous absorption is negligible far from line centers, though it will usually be significant somewhere in a solar or surface emission spectrum.

Scattering is the main concern since it significantly complicates radiative transfer calculations. It can be neglected when \(\tilde{\omega}\) is low enough compared to unity. Absorption tends to scale with the mass of the absorber; as seen in Chapter 4, scattering per unit mass (hence \(\tilde{\omega}\)) tends to be greatest when the scattering particle radii \(a \sim \lambda\), diminishing especially rapidly when \(a \ll \lambda\). In Earth’s atmosphere this makes clouds the main scatterers. Clouds are especially good scatterers at visible wavelengths, where \(\tilde{\omega} > 0.99999!\) At longer wavelengths water becomes more absorptive, so that by the time one reaches the far infrared, scattering is usually ignored even though \(a \sim \lambda\) for many cloud particles. In gases, the importance of scattering decreases rapidly with wavelength due to the \(\nu^4\) dependence. Earth’s atmosphere is too thin to scatter much infrared radiation, though enough visible light scatters to make the sky blue. Even in an atmosphere thick enough to scatter some infrared light, scattering becomes moot in the presence of even the slightest traces of absorbing material.
6.2 The plane-parallel approximation

Having decided what terms to keep in the radiative transfer equation, the next question is how to simplify the geometry of the physical problem. Often, the medium and the radiation incident on it vary mainly along one (z or “vertical”) axis. Solutions to the RTE where quantities vary only in this direction are called plane-parallel solutions. We do not assume that the photons all move straight up and down, only that properties of the medium and the radiation are the same at all horizontal positions. We must still reckon with the full variation of $I_\nu(\Omega)$ in angle, frequency, and altitude; this is sufficiently challenging that in the context of Earth’s atmosphere, attempts to perform realistic radiative transfer calculations outside of the plane-parallel framework (i.e., “3-D radiative transfer”) have been attempted only recently and have not (for example) been implemented into any climate models as of this writing other than by (in a few cases) making fudge-factor type adjustments to the input or output of plane-parallel calculations.

In a plane-parallel situation we may write (Figure 6.1)

$$d_s = \frac{1}{\mu} dz$$

where $\mu \equiv \cos \theta$, and we choose $\theta$ as the zenith angle between $\Omega$ and the upward direction. Thus, for beams traveling upward $\mu = 1$, while for downward beams $\mu = -1$. Recalling (2.15) and choosing $\tau$ to increase downward,

$$\tilde{\tau}_\nu = -\frac{1}{\mu} \tau_\nu$$

under plane-parallel assumptions. Then (2.13) becomes

$$-\mu \frac{dI_\nu}{d\tau_\nu} + I_\nu = J_\nu.$$  \hspace{1cm} (6.2)

Though it may not look very different, this is a key improvement because the equation now applies in a fixed (Eulerian) rather than natural coordinate system. Given $c_\nu(z)$, $\tau_\nu$ is a fixed function of $z$. One may take $\tau_\nu = 0$ wherever is convenient; we will take this to lie at the top of the medium. The minus sign in (6.2) arises because beams traveling downward have $\mu < 0$ but travel toward increasing $\tau$ (sorry, but no matter what conventions you adopt, nonintuitive minus signs will crop up somewhere!).

The RTE solution (2.18) applies to the plane-parallel RTE (6.2) with the simple substitution (6.1). Thus, the fraction of incident radiance that penetrates through a layer of optical depth $\tau_\nu$ is

$$T_{\nu,\Omega}(\tau_\nu) = \exp(- \left| \frac{\tau_\nu}{\mu} \right|),$$ \hspace{1cm} (6.3)

which we designate as the directional or beam transmittance. Note that this transmittance does not include any scattered light, some of which may also penetrate through a given layer resulting in greater total transmittance than given by (6.3) if $\Omega > 0$. The beam transmittance is greatest at normal incidence ($\mu = -1$), and approaches zero at glancing incidence, due to the increasing path lengths that a beam must traverse to achieve the same degree of vertical penetration.
Figure 6.1: Plane parallel geometry. For problems of incident or outgoing radiation to/from above, we define optical depth to increase downward from zero at the top of the medium (atmosphere), and define $\theta$ such that $\mu > 0$ for upwelling radiation.
6.3 Diffuse transmission and the diffusivity factor

Thermal radiation (radiation at wavelengths where emission is important) is inherently diffuse, since emission is isotropic. Scattering also tends to smooth out any anisotropy in the radiation field, so that (for example) sunlight passing through a cloud quickly becomes diffuse. We must account for radiation traveling in all directions.

In many cases, however, one is interested only in hemispherically-averaged energy fluxes $F_{\nu}$ in the $+z$ and $-z$ (upward and downward) directions. It is useful to define the monochromatic diffuse transmittance or transmission function $\mathcal{T}_\nu$ as the fraction of an incident flux that penetrates through a layer of optical thickness $\tau_\nu$:

$$\mathcal{T}_\nu(\tau_\nu) = \frac{F_{\nu(\text{transmitted})}}{F_{\nu(\text{incident})}}.$$  \hspace{1cm} (6.4)

In writing this we again ignore contributions from emission or scattering, so that transmittance may be thought of as the average probability of an incident photon at one depth penetrating to another without being absorbed or scattered.

The flux definition (1.7) may be rewritten as

$$F_{\nu} = \int_0^{2\pi} d\phi \int_0^1 \tilde{\mu} I_{\nu}(\tilde{\mu}, \phi) d\tilde{\mu} = \pi \langle I_{\nu} \rangle,$$  \hspace{1cm} (6.5)

where $\tilde{\mu}$ is the (positive definite) cosine of the zenith angle between the flux and radiance directions, and $\langle \rangle$ denotes a “flux-weighted” hemispheric mean. We may then invoke the simple concept of the diffusivity factor $\tilde{d}$, defined so that

$$\exp(-\tau_\nu \tilde{d}) \approx \frac{\langle I_{\nu} \exp(-\frac{\tau_\nu}{\tilde{d}}) \rangle}{\langle I_{\nu} \rangle}.$$  \hspace{1cm} (6.6)

as closely as possible for realistic radiance distributions. This establishes a rescaled, flux optical depth $\tau_\nu \tilde{d}$ that accounts for the fact that path lengths in general are longer than in the vertical direction. Accordingly, we have

$$\mathcal{T}_\nu(\tau_\nu) = \exp(-\tau_\nu \tilde{d}).$$  \hspace{1cm} (6.7)

Most importantly, we can integrate (6.2) over either the upper or lower hemispheres using (6.5) to obtain the flux transfer equation

$$\pm \frac{1}{\tilde{d}} \frac{dF_{\nu}}{d\tau_\nu} + F_{\nu} = \pi \langle J_{\nu} \rangle$$  \hspace{1cm} (6.8)

where plus (minus) applies to downward (upward) fluxes. The utility of the diffusivity approximation is clear in the resemblance of this flux equation to the natural (2.12) and plane-parallel (6.2) radiance RTEs. In particular, we can use the same solution (2.18) simply by making the associations

$$\tilde{\tau}_\nu = \pm \tau_\nu \tilde{d}$$  \hspace{1cm} (6.9)

$$F_{\nu} = \pi I_{\nu}.$$  \hspace{1cm} (6.10)
giving
\[ F_\nu(\tau_\nu) = F_\nu(0) e^{-|\tau_\nu|d} + \pi \int_0^{\tau_\nu} \langle J_\nu(x) \rangle e^{-|\tau_\nu-x|d} dx , \] (6.11)
again assuming for simplicity that the boundary where incoming flux is known is located at \( \tau_\nu = 0 \).

Though more accurate approaches exist, the diffusivity approximation is simple and yields qualitatively meaningful results whenever the radiation field may be assumed to be reasonably diffuse. The “correct” value of \( \tilde{d} \) depends on the angular distribution of radiance according to (6.6), which will obviously vary with circumstances. Clearly, \( \tilde{d} > 1 \) unless all radiation is exactly vertical; if the incident radiation is hemispherically isotropic, then \( \tilde{d} \) cannot exceed 2. Investigators have found that a value of \( \tilde{d} = 1.66 \) provides fluxes to the best overall accuracy (a few percent or better) for a wide range of \( \tau_\nu \), so this value is often used. The reason a single value works across many situations is subtle, and will be revisited after we consider solutions in scattering atmospheres.

Finally, it is worth obtaining an alternative form of (6.11). We can use any convenient variable of integration in the last term; instead of optical depth \( x \), for example, we can use transmission from \( x \) to \( \tau_\nu \), since this is monotonic in \( x \). Let \( y \) be this new integration variable; according to (6.7), \( y = \exp(-\tilde{d}(\tau_\nu - x)) \), and \( dy = y \tilde{d} dx \). Substitution gives
\[ F_\nu(\tau_\nu) = F_\nu(0) T(\tau_\nu) + \pi \int_{T(\tau_\nu)}^{1} \langle J_\nu(y) \rangle dy . \] (6.12)
Note that the \( \tilde{d} \) no longer appears; in fact, by a roundabout way, we have obtained a more general formula that does not depend on this particular approximation. However, to use this formula one needs a transmission function, if not (6.7) then something else.

### 6.3.1 The planetary greenhouse effect

Planetary thermal radiation can be calculated fairly accurately without worrying about scattering. Consequently the source function is just Planck emission, and (6.12) gives for the upward flux \( F^\dagger \) within the atmosphere,
\[ \frac{F^\dagger_\nu(\tau_\nu)}{\pi} = B_\nu(T_0) T_\nu(\tau_\nu) + \int_{T(\tau_\nu)}^{1} B_\nu(T(y)) dy , \] (6.13)
where \( T_0 \) and \( \tau_{\nu 0} \) are the surface values and we assume a black surface. By defining the diffuse absorption function
\[ A_\nu \equiv 1 - T_\nu , \]
and performing a few manipulations (see Problem 2) we obtain
\[ \frac{F^\dagger_\nu(z)}{\pi} = B_\nu(T_0) + \int_0^z A_\nu(z, z') dB'_\nu \frac{dT}{dz'} dz' , \] (6.14)
where we have also changed the vertical coordinate to height \( z \) to help make the result more intuitive. We are using the notation \( A_\nu(z, z') \) to designate the absorption between
two specific heights. Equation (6.14) may be integrated over frequency to get

\[ F^1(z) = \sigma T^1_0 + \pi \int_0^z \frac{dT}{dz'} \left[ \int_0^\infty A(\nu, z') \frac{dB(\nu)}{dT} d\nu \right] dz', \quad (6.15) \]

The second term on the right hand side is the greenhouse effect of the atmosphere, quantified by its alteration of the outgoing flux from what would have been emitted by a bare surface. Clearly \( A(\nu) \) is always positive, as is the derivative of \( B(\nu) \) with respect to \( T \), so that the quantity in brackets in (6.15) is always positive. This means that the sign of the greenhouse effect depends on the atmospheric lapse rate. Since temperature typically decreases with height, the greenhouse effect decreases the outgoing flux and warms the planet. The greenhouse effect maximizes at a tropopause where \( \frac{dT}{dz} \) is maximized, and is sometimes defined at this level for Earth climate sensitivity analyses, though it is the greenhouse effect defined by taking \( z \to \infty \) that is relevant for a planetary energy budget.

### 6.4 Solutions with scattering

When scattering cannot be neglected, the problem of solving the RTE becomes much more challenging. Evaluation of the scattering integral requires knowledge of the entire angular distribution of radiation at a given point in order to establish its spatial derivatives; this couples all the directions and creates a daunting closure problem. Another difficulty sometimes associated with scattering cases is that, if absorption and emission are weak, the results become more sensitive to the angular distribution of incoming radiation at the boundary. We first address the treatment of anisotropic incoming radiation, then discuss two approaches to solving the RTE in scattering situations. In addition to the plane-parallel assumption, we assume here that \( \bar{\omega} \) is constant with depth.

### 6.5 Separation of direct and diffuse radiation

If scattering is important, it usually means there is intense illumination from outside. For planetary problems, solar radiation arrives as almost a unidirectional beam, while emitted or scattered radiation tends to have a smooth distribution over all angles. It is a practical necessity to calculate separately the “direct” or “solar” beam \( I_S \) (consisting of whatever remains of the original source beam), and the “diffuse” radiation field \( I_D \) consisting of all other radiation in the medium:

\[ I = I_S + I_D, \quad (6.16) \]

temporarily dropping the \( \nu \) subscripts. The solar beam is the easy part. It diminishes in the atmosphere according to the RTE with \( J = 0 \):

\[ -\mu \frac{dI_S}{dT} + I_S = 0 \quad (6.17) \]

since by definition it has no sources. Since the radiance of this beam is infinite in the \( \Omega_0 \) (solar propagation) direction and zero in other directions, it is best quantified initially by its monochromatic flux density \( \pi S \) in the downward direction.
change this to just $S$? I actually like seeing "$4\pi$" rather than just "$4$" in the equations, and find it weird to introduce a new quantity like this. Also, should we define this as the downward flux or the flux in the propagation direction? If the latter, then subsequent equations need a factor of $\mu_0$ in them. Then the solution to (6.17) appears as

$$I_S = \frac{\pi S}{\mu_0} \exp(-\frac{\tau}{\mu_0}) \delta(\Omega, \Omega_0)$$  \hspace{1cm} (6.18)$$

where $\mu_0$ is the cosine of the solar zenith angle and $\delta(\Omega, \Omega_0)$ is the Dirac delta function. The solar beam travels in a direction where $\mu = -\mu_0$, because solar zenith is the angle from the overhead sun position while $\mu$ was defined as positive for upwelling radiances.

If our decomposition (6.16) and the solar beam solution (6.18) are substituted into the RTE (6.2), and the solar beam terms eliminated using (6.17), we obtain an equation for the diffuse part of the field:

$$-\mu \frac{dI_D(\tau, \Omega)}{d\tau} + I_D(\tau, \Omega) = (1 - \bar{\omega}) B(T) + \frac{\bar{\omega}}{4\pi} \int I_D(\tau, \Omega') P(\Omega, \Omega') d\Omega' + \frac{\bar{\omega}}{4\mu_0} S P(\Omega_0, \Omega) \exp(-\frac{\tau}{\mu_0}).$$ \hspace{1cm} (6.19)$$

This equation looks like our original RTE (6.2) except there is now an additional “source” of diffuse radiation due to scattering out of the solar beam. If we write (6.19) out explicitly in angular variables using

$$d\Omega = \sin \theta d\theta d\phi = d\mu d\phi,$$

and rearrange terms a bit, we obtain

$$\mu \frac{dI_D(\tau, \mu, \phi)}{d\tau} = I_D(\tau, \mu, \phi) - \frac{\bar{\omega}}{4\pi} \int_{-1}^{1} \int_{0}^{2\pi} I_D(\tau, \mu', \phi') P(\mu, \phi, \mu', \phi') d\phi' d\mu' - J_S.$$ \hspace{1cm} (6.20)$$

This form of the equation neatly combines all the source of diffuse radiation other than diffuse radiation itself into an “external” source

$$J_S \equiv \frac{\bar{\omega}}{4\mu_0} S P(-\mu_0, \phi_0, \mu, \phi) \exp(-\frac{\tau}{\mu_0}) + (1 - \bar{\omega}) B(T)$$ \hspace{1cm} (6.21)$$

which includes emission and scattering out of the solar beam. Note that (6.20) is identical to the original RTE (2.12) with $\tau/\mu$ substituted for $\tau$ (6.1), except that $J_S$ replaces the emission term. The formal solution of (6.20) is therefore given by (2.18), with the same substitution for $\tau$, with $I$ in that equation understood to be the diffuse portion of the field, and with the source function

$$J = \frac{\bar{\omega}}{4\pi} \int_{-1}^{1} \int_{0}^{2\pi} I(\tau, \mu', \phi') P(\mu, \phi, \mu', \phi') d\phi' d\mu' + J_S,$$ \hspace{1cm} (6.22)$$

that includes all sources of diffuse radiation from emission and from scattering of either the solar beam or the diffuse field.
The above decomposition is typically necessary for solar heating of atmospheres, but not infrared cooling, where incoming fluxes at boundaries are diffuse. Regardless, we end up with an equation that includes a scattering integral, and the field to be solved for (either the total field or the diffuse part) appears both outside and inside this integral. The remainder of this chapter discusses approaches for obtaining approximate solutions of this difficult type of equation.

6.6 Approximate approach 1: orders of scattering

We will discuss two broad approaches that have been developed for solving the RTE with scattering. The first is a specific method known as successive orders of scattering. The idea is simple: begin by evaluating the scattering integral in the RTE with \( I \) set equal to the unscattered incident field (minus attenuation due to absorption or scattering) \( I_0 \). This yields a first guess \( I' \) as to the scattered radiation field. Then, reevaluate the scattering integral with \( I \) set to \( I_0 + I' \), to get a better estimate of \( I' \). This may be repeated, using the improved estimate in place of the previous one, until the result converges to sufficient accuracy or one runs out of patience.

The first iteration of this procedure is equivalent to assuming that each photon can scatter no more than once; the second, that each photon scatters no more than twice; and so on. The successive orders of scattering approach is obviously well suited to thin atmospheres where higher orders of scattering rapidly decay in importance. It is not well suited to cumulus clouds for example, where a photon can scatter hundreds of times.

The first order of scattering, called the single-scatter approximation, is the simplest to calculate, especially for the case of beam incidence discussed in the previous section. Assuming there are no other, diffuse sources, we have already calculated the unscattered (decaying) incident beam \( I_s \). We would simply use (6.20) with its integral term set to zero (since there is no diffuse illumination). A small amount of algebra yields the backscattered radiation field at the top of the layer,

\[
I(\tau = 0, \mu, \phi) = \frac{\varphi S}{4} \frac{1}{\mu + \mu_0} P(\mu_0, \phi_0, \mu, \phi) \left\{ 1 - \exp \left( -\tau_0 \left( \frac{1}{\mu} + \frac{1}{\mu_0} \right) \right) \right\}. \tag{6.23}
\]

The single-scatter approximation is reasonable only for optically thin layers. If, in addition, \( \mu/\tau \ll 1 \) and \( \mu_0/\tau \ll 1 \) for the angles of interest, the above solution can be further approximated by

\[
I(\tau = 0, \mu, \phi) \approx \frac{\varphi S \tau_0}{4 \mu \mu_0} P(\mu_0, \phi_0, \mu, \phi). \tag{6.24}
\]

The transmitted field is left as an exercise. The bidirectional reflectance may be found by dividing the backscattered radiances by \( S \).

In the case of a diffuse incident field, we would no longer have any external source \( J_s \), but we would instead have to evaluate the scattering integral based on the diffuse incident field (minus extinction) analogous to \( I_s \) in (6.18). This clears the way for solving (approximately) the RTE. This is also the procedure to continue to higher orders...
of accuracy, except that the sum of the most recently obtained scattered field plus the unscattered incident field must always be inserted into the scattering integral the next time around.

### 6.6.1 Adding and doubling methods

An important strategy for obtaining solutions to the RTE, known as the *adding method*, is similar in spirit to the orders of scattering method described above. Suppose one is interested only in transmission and reflection, and has solutions $T_1$, $R_1$, $T_2$, $R_2$ already in hand for diffuse behavior through two layers (typically, two very optically thin layers where the single-scatter approximation has been used). Then one can calculate the total transmission $T$ and reflection $R$ of the layers in combination by considering successive orders of reflection and transmission between them. The combined transmission $T$ given diffuse incidence will be

$$
T = T_1 T_2 + T_1 R_2 R_1 T_2 + T_1 R_2 T_1 R_1 R_2 T_2 + \ldots
$$

$$
= T_1 T_2 \sum_{i=0}^{\infty} R_1^i R_2^i
$$

$$
= T_1 T_2 (1 - R_1 R_2)^{-1}
$$

One can then repeat the calculation, combining one new layer at a time with the previous result to build up to arbitrarily large optical depth solutions. Not only overall transmission and reflection, but also net fluxes inbetween each layer, are obtained.

In this simple description of the method, we have not considered the angular distribution of the radiation. In practice the results are not very accurate unless $R$ and $T$ are computed as functions of angle (i.e., $R(\mu, \phi, \mu', \phi')$ etc.), since these functions will typically be sensitive to the angle of incidence and the distribution of this will change deeper into the medium (see Section 6.8.3). This is especially so if there is collimated incidence. If angular information is retained, then at each stage of the adding process, one must integrate over all angles in performing the noted multiplications. The benefit of this, besides better accuracy, is that one ends up with a calculation of the full radiance field rather than just the fluxes. This method is commonly used for calculation of backscattered sunlight for use in planetary remote sensing applications.

Large optical depths can be attained much more rapidly using the adding method for homogeneous layers, since one can use the previous solution as both the top and bottom layer in the next iteration; then $\tau$ increases as $2^n$ rather than $n$. This is called the *doubling method*. *Adding-doubling* methods are widely used to obtain accurate simulations of fluxes in plane-parallel situations with large degrees of scattering and vertical variation of scattering properties; one can divide the atmosphere into homogeneous layers, compute the properties of each layer by the doubling method, then combine the layers using the adding method.
6.7 Approximate approach 2: angular distribution functions

The other fundamental approach to solving the RTE is to somehow parameterize the angular distribution of radiance by representing it by a finite number of basis function amplitudes or values at quadrature points. This enables the continuum of RTE’s at different angles to be approximated by a closed, linear system of finite size. The approach is not limited to small optical depth, and can usually be applied to a medium with vertically varying scattering properties. Here we will consider homogeneous layers only.

This general approach comes in many forms which have reappeared in the radiative transfer literature under many names. A word of caution concerning these methods is the loose terminology and use of symbols that appear; “intensity” is often used to mean radiance or any quantity derived from radiance, including the flux density, and many of these quantities are typically denoted by the same symbol as radiance. One source of trouble is the frequent assumption that all photons travel in certain directions (c.f. (6.18)), which leads to infinite radiances in those directions and requires the use of other quantities to describe the quantity of energy conveyed. Here, overbars will be used to designate averaged quantities and tildes to designate parameters, but this is not standard notation.

A nearly universal assumption made prior to embarking on this approach is that the phase function depends only on the scattering angle \( \Theta \), i.e.,

\[
P(\mu, \phi, \mu', \phi') = P(\cos \Theta).
\] (6.25)

This holds if the scattering medium is isotropic, which is true for gases and small aerosol particles but not always strictly true in clouds since crystals or droplet shapes can have a preferred orientation relative to gravity\(^1\). Decomposition of \( P \) into Legendre polynomials will be discussed in Section 6.7.2.

For low-order approximations, the problem is usually simplified by solving only for the azimuthally averaged radiance field, defining

\[
\bar{I}(\tau, \mu) \equiv \int_0^{2\pi} I(\tau, \mu, \phi) d\phi,
\] (6.26)

which is \( 2\pi \) times the mean radiance in a cone of constant \( \mu \). In this case there is no point in retaining azimuthal scattering information so we also define the probability of scattering from one such cone into another by integrating over all relative azimuths at which this could happen,

\[
P(\mu, \mu') \equiv \frac{1}{2\pi} \int_0^{2\pi} P(\mu, \phi, \mu', \phi') d(\phi - \phi').
\] (6.27)

The zenith phase function \( P \) is independent of \( \phi \) by symmetry. Note that the normalization condition (2.7) on \( P \) becomes

\[
\int_{-1}^{1} P(\mu, \mu') d\mu = 2
\]

\(^1\)Flying over stratiform ice clouds, one can occasionally observe a specular reflection of the sun due to the crystalline shapes’ orientation with flat faces facing upward, an effect that contradicts (6.25).
when applied to $P$. Integrating (6.20) over all $\phi$ and employing (6.26-6.27), we obtain

$$
\frac{dI(\tau, \mu)}{d\tau} = I(\tau, \mu) - \frac{\bar{\omega}}{2} \int_{-1}^{1} I(\tau, \mu') P(\mu, \mu') d\mu' - (1 - \bar{\omega}) B(T) - \frac{\bar{\omega}}{2} S P(\mu_0, \mu) \exp\left(-\frac{\tau}{\mu_0}\right).
$$

### 6.7.1 Two-stream methods and discrete ordinates

The simplest way of approaching the scattering problem is to idealize the full radiance field as two streams of radiation, one moving up and one down. This two-stream approach is still widely used in numerical weather and climate models of Earth’s atmosphere, where radiation calculations must be performed many times and consume a significant fraction of the total CPU effort, and where only vertical fluxes and their gradients are needed. Though crude, the ability of this approach to capture the gist of flux behavior is well recognized. Many variations of the approach have been proposed and compared in the literature. Here we present a particularly simple version whose value is mainly pedagogical: the approximation (originally employed by Coakley and Chylek, 1975) that the radiance takes on the constant value

$$
I(\tau, \mu) = \begin{cases} 
\frac{\bar{\omega}}{2} & \text{for } \mu > 0 \\
-\bar{\omega} & \text{for } \mu < 0
\end{cases}
$$

Thus, in this particular case the two “streams” are hemispherically isotropic sprays.

Since angular information is sacrificed in any two-stream approach, one usually begins by reexpressing the RTE in terms of the two fluxes

$$
F^\uparrow = \int_0^1 \mu I(\tau, \mu) d\mu, \\
F^\downarrow = -\int_{-1}^0 -\mu I(\tau, \mu) d\mu.
$$

Both fluxes are defined so as to be positive. Integrating (6.28) over the individual hemispheres and neglecting emission yields the flux equations

$$
\frac{dF^\uparrow(\tau)}{d\tau} = \int_0^1 I(\tau, \mu) d\mu + \frac{\bar{\omega}}{2} \int_{-1}^1 I(\tau, \mu') P(\mu, \mu') d\mu' d\mu - \pi \bar{\omega} S \beta_0 \exp\left(-\frac{\tau}{\mu_0}\right),
$$

$$
\frac{dF^\downarrow(\tau)}{d\tau} = -\int_0^1 I(\tau, -\mu) d\mu + \frac{\bar{\omega}}{2} \int_{-1}^1 I(\tau, -\mu') P(\mu, -\mu') d\mu' d\mu + \pi \bar{\omega} S (1 - \beta_0) \exp\left(-\frac{\tau}{\mu_0}\right),
$$

where we have defined the \textit{backscattering function}

$$
\beta(\mu) \equiv \frac{1}{2} \int_0^1 P(\mu, -\mu') d\mu',
$$

which is the fraction of scattered radiation that scatters back into the same hemisphere it came from, as a function of incident $\mu$, and $\beta_0$ is the value of this function at the solar incidence angle. For now, we simply take $P$ (hence $P$) to be known and $\beta(\mu)$...
6.7. APPROXIMATE APPROACH 2: ANGULAR DISTRIBUTION FUNCTIONS

6.7.1 Approximate Approach 2: Angular Distribution Functions

To be calculable in a straightforward way. Making the assumption of hemispherically constant radiances yields

\[ \int_0^1 I(\tau, \left( \frac{\mu}{\sqrt{1-\mu^2}} \right)) d\mu = 2 \left( \frac{F^\dagger}{F^\perp} \right). \]

Defining also the mean backscatter over all incident angles allows (6.31-6.32) to be rewritten under the hemispheric constant assumption, in matrix form, as

\[ \frac{d}{d\tau} \left( \begin{array}{c} F^\dagger \\ F^\perp \end{array} \right) = 2\omega \left( \begin{array}{c} 1 - b \\ b \end{array} \right) F^\dagger + \left( \begin{array}{c} b \\ 1 - b \end{array} \right) F^\perp + \pi\omega S \left( \begin{array}{c} -\beta_0 \\ 1 - \beta_0 \end{array} \right) \exp\left( -\frac{\tau}{\mu_0} \right). \]

The solution of this system of equations will be addressed in Section 6.8.

The most popular “stream” method is the discrete ordinates method, introduced by Chandrasekhar and made popular by Liou (1973), in which Gaussian quadrature is used to evaluate the scattering integral. At the lowest (two-stream) order of approximation, this is equivalent to assuming that streams of radiation travel in the specific quadrature directions. This two-stream scheme leads to a system of equations similar to (6.34) but with different coefficients. No single two-stream method is superior to others, but the discrete ordinates approach has a distinct virtue: it can be increased to any desired accuracy using higher orders of quadrature. Four-stream calculations have become quite popular as simple methods for calculating fluxes with good accuracy—the addition of two more streams allows for the tendency of radiances distributions to narrow toward the vertical to be represented. Many-stream discrete ordinates approximations are regarded as a rigorous standard that is comparable to fully angular adding-doubling methods for obtaining the directionally-resolved \( I(\Omega) \).

Though we will not present the discrete ordinates method here, we will mention some of its characteristics in passing. A detailed presentation can be found in Liou.

6.7.2 Expansion methods and Eddington’s approximation

Yet another approach to parameterizing the angular distribution of radiation is the method of moments or the closely-related method of polynomial expansion. In the former, (6.28) is multiplied by successively higher powers of \( \mu \) to produce a sequence of coupled equations for the moments of the radiation field. In the latter, the radiance field is expanded in Legendre polynomials and a system of equations obtained for the coefficients. At the lowest useful order (two equations), the methods of moments and polynomial expansion are identical and are known as Eddington’s approximation, one of many approximations with Eddington’s name attached to it. Here we present Eddington’s approximation followed by a general presentation of the polynomial expansion method.
Eddington assumed

\[
\bar{I}(\tau, \mu) = \sum_{l=0}^{N} \bar{I}_l P_l(\mu)
\]

(6.35)

\[
\bar{I}_0(\tau) + \bar{I}_1(\tau) \mu, \quad (N = 1)
\]

(6.36)

neglecting azimuthal dependences according to (6.26) and (6.25). Insertion into (6.30) gives \(I\) in terms of the fluxes as

\[
I(\tau, \mu) = (1 + \frac{3\mu}{2}) P^\uparrow + (1 - \frac{3\mu}{2}) P^\downarrow
\]

(6.37)

for the \(N = 1\) case. Things get messy when \(N > 1\), and the procedure described at the end of this section should be followed instead.

Now it behoves anyone following this method (for reasons to become apparent below) to expand the phase function \(P(\cos \Theta)\) in the same set of functions:

\[
P(\cos \Theta) = \sum_{l=0}^{N} \alpha_l P_l(\cos \Theta)
\]

(6.38)

\[
= \alpha_0 + \alpha_1 \cos \Theta \quad (N = 1)
\]

(6.39)

The expansion coefficients are

\[
\alpha_l = \frac{2l + 1}{2} \int_{-1}^{1} P(x) P_l(x) dx
\]

(6.40)

as dictated by the orthogonality relations obeyed by these polynomials:

\[
\int_{-1}^{1} P_l P_k = \begin{cases} 
0 & (l \neq k) \\
\frac{2}{2l + 1} & (l = k)
\end{cases}
\]

From geometry, the cosine of the scattering angle is

\[
\cos \Theta = \mu \mu' + (1 - \mu^2) \frac{1}{2} (1 - \mu'^2) \frac{1}{2} \cos(\phi - \phi'),
\]

(6.41)

where as before the unprimed and primed coordinates refer to the incident and scattering directions. If this expression is substituted into (6.39) using (6.40), and the resulting polynomials substituted into (6.27) we obtain the simple result

\[
P(\mu, \mu') = 1 + 3g \mu \mu'
\]

(6.42)

where the \textit{asymmetry factor} \(g\) (first moment of the phase function) has been defined:

\[
g \equiv \frac{\alpha_1}{3} = \frac{1}{2} \int_{-1}^{1} xP(x) dx
\]

with \(x = \cos \Theta\). The asymmetry factor ranges from \(-1\) to \(+1\) depending on how much energy scatters forward vs. backward, and is important because it controls how deeply scattered light can penetrate into optically thick media.
6.7. APPROXIMATE APPROACH 2: ANGULAR DISTRIBUTION FUNCTIONS

Since \( P(\mu, \mu') \) and \( I \) are expressed in the orthonormal Legendre basis set, evaluation of the scattering integrals is not difficult and yields

\[
- \frac{\mathcal{Q}}{2} \int_0^1 \int_{-1}^1 \tau(\tau, \mu') P(\mu, \mu') \, d\mu' \, d\mu = - \frac{\mathcal{Q}}{4} [(4 \pm 3g) F^\uparrow + (4 \mp 3g) F^\downarrow].
\]

Substitution of these results into (6.28), and again neglecting emission for simplicity, yields

\[
\frac{d}{d\tau} \left( \begin{array}{c} F^\uparrow \\ F^\downarrow \end{array} \right) = \frac{1}{4} \left( \begin{array}{c} 7 - \mathcal{Q}(4 + 3g) \\ \mathcal{Q}(4 - 3g) - 1 \end{array} \right) F^\uparrow + \frac{1}{4} \left( \begin{array}{c} 1 - \mathcal{Q}(4 - 3g) \\ \mathcal{Q}(4 + 3g) - 7 \end{array} \right) F^\downarrow + \frac{\mathcal{Q}}{4} S \left( \begin{array}{c} -2 + 3g \mu_0 \\ 2 + 3g \mu_0 \end{array} \right) \exp(-\frac{\tau}{\mu_0}).
\]

\subsection*{Legendre expansion with \( N > 1 \)}

Any function spanning the interval \([-1, 1]\) can be represented by a Legendre series with sufficiently large \( N \). The general approach to (6.38) is to use the addition theorem,

\[
P_l(\cos \Theta) = \sum_{m=0}^{l} \frac{(2 - \delta_{m0})(l - m)!}{(l + m)!} P^m_l(\mu) P^m_l(\mu') \cos m(\phi - \phi'),
\]

(6.44)
to express the Legendre polynomials in terms of spherical harmonics. The associated Legendre functions can be found by

\[
P_l^m(x) = \sqrt{1 - x^2} \frac{d^m P_l(x)}{dx^m}
\]

(6.45)

Introducing (6.44) into (6.38), and interchanging the order of summation, leads to

\[
P(\mu, \phi, \mu', \phi') = \sum_{m=0}^{N} (2 - \delta_{m0}) \cos m(\phi - \phi') \left[ \sum_{l=m}^{N} \alpha_l \frac{(l - m)!}{(l + m)!} P^m_l(\mu) P^m_l(\mu') \right].
\]

(6.46)

Inserting this into (6.27) gives

\[
P(\mu, \mu') = \sum_{i=0}^{N} \alpha_i P_i(\mu) P_i(\mu'),
\]

(6.47)

which may be used with (6.35) to expand the RTE.

\subsection*{Full expansion with azimuthal variation}

Those who are really ambitious may want to dispense with (6.26) and fully calculate \( \tau(\tau, \mu, \phi) \). This requires expanding the radiance azimuthal dependence as

\[
I(\mu, \phi) = \sum_{m=0}^{N} (2 - \delta_{m0}) \cos m(\phi - \phi') \left[ \sum_{l=m}^{N} \frac{(l - m)!}{(l + m)!} P^m_l(\mu) P^m_l(\mu') \right],
\]

(6.48)
The quantity in brackets in (6.48) is to be solved for. Substituting (6.46) and (6.48) into (6.28) will, by virtue of the orthogonality relation of the polynomials and straightforward integration over $\phi'$, lead to a set of $N+1$ equations

$$
\frac{d\tilde{I}}{dT} = I - (1 - \tilde{\omega})B(T) - \tilde{\omega} \int_{-1}^{1} I(\tau, \mu') P(\mu, \mu') d\mu' - \tilde{\omega} S P(\mu_0, \mu) \exp\left(-\frac{\tau}{\mu_0}\right).
$$

(6.49)

### 6.8 Solution of two-stream and Eddington models

#### 6.8.1 General solution

Now we return to the two-stream and Eddington methods, using them to explore the behavior of scattering layers. Each system (6.34) and (6.43) of equations has a similar form; in fact, they can both be rearranged into the general matrix representation

$$
\begin{pmatrix}
D_\tau - \gamma_1 & \gamma_2 \\
-\gamma_2 & D_\tau + \gamma_1
\end{pmatrix}
\begin{pmatrix}
F^\uparrow(\tau) \\
F^\downarrow(\tau)
\end{pmatrix}
= 
\begin{pmatrix}
S^\uparrow(\tau) \\
S^\downarrow(\tau)
\end{pmatrix}
$$

(6.50)

where $D_\tau$ is the derivative operator with respect to $\tau$, and $\gamma_1$ and $\gamma_2$ are constants that depend on the method used. The specified forcings $S^\uparrow, S^\downarrow$ represent the “external” sources of flux from the collimated beam and/or Planck emission. These equations are rearranged by the standard elimination technique of linear algebra to yield

$$
\begin{pmatrix}
D_\tau^2 - \gamma_1^2 + \gamma_2^2 & 0 \\
0 & D_\tau^2 - \gamma_1^2 + \gamma_2^2
\end{pmatrix}
\begin{pmatrix}
F^\uparrow(\tau) \\
F^\downarrow(\tau)
\end{pmatrix}
= 
\begin{pmatrix}
(D_\tau + \gamma_1)S^\uparrow(\tau) - \gamma_2 S^\downarrow \\
(D_\tau - \gamma_1)S^\downarrow(\tau) + \gamma_2 S^\uparrow
\end{pmatrix}
$$

(6.51)

from which we obtain the solution

$$
\begin{pmatrix}
F^\uparrow(\tau) \\
F^\downarrow(\tau)
\end{pmatrix}
= 
\begin{pmatrix}
A & B \\
Aw & Bu
\end{pmatrix}
\begin{pmatrix}
\xi^\tau \\
e^{-\xi^\tau}
\end{pmatrix}
+ \text{Particular sol'n}
$$

(6.52)

where

$$
\xi^2 \equiv \gamma_1^2 - \gamma_2^2, \quad w = \frac{\gamma_1 - \xi}{\gamma_2}, \quad v = \frac{\gamma_1 + \xi}{\gamma_2}
$$

(6.53)

The factors $w$ and $v$ result from insertion of the general solution for $F^\uparrow$ into (6.51). The form of the particular solution will depend on the nature of $S$. The two integration constants $A$ and $B$ must be determined from the boundary conditions (specified diffuse flux into the layer top, and at bottom either another specified flux or a reflectivity condition). If the incident radiation is collimated and the layer extends to the top of atmosphere, for example, the downward diffuse flux at top would be zero yielding $Aw + Bu = 0$.

#### 6.8.2 Solution for diffuse incidence

As an example, we solve (6.52) for the situation of diffuse incoming flux $F_0$ at the top ($\tau = 0$) of the layer with no collimated incidence, no emission, and no upwelling or
reflection from below the layer bottom at \( \tau = \tau_0 \). Thus \( F^\dagger(0) = F_0 \) and \( F^\dagger(\tau_0) = 0 \) yeilding, respectively,

\[
Aw + Bv = F_0 \quad Ae^{\xi\tau_0} + Be^{-\xi\tau_0} = 0 \tag{6.54}
\]

and a particular solution of zero. Inserting these constraints into (6.51) and rearranging yields

\[
F^\dagger(\tau) = F_0 \frac{e^{-\xi(\tau_0 - \tau)} - e^{\xi(\tau_0 - \tau)}}{we^{-\xi\tau_0} - ve^{\xi\tau_0}} \quad (6.55)
\]

\[
F^\dagger(\tau) = F_0 \frac{we^{-\xi(\tau_0 - \tau)} - ve^{\xi(\tau_0 - \tau)}}{we^{-\xi\tau} - ve^{\xi\tau}}. \tag{6.56}
\]

Of particular interest in this case is the layer albedo

\[
R \equiv \frac{F^\dagger(0)}{F_0} = \frac{e^{-\xi\tau_0} - e^{\xi\tau_0}}{we^{-\xi\tau} - ve^{\xi\tau}}. \tag{6.57}
\]

In optically thin layers (\( \tau_0 \xi \ll 1 \)),

\[
R \approx \frac{2 \tau_0 \gamma_2}{2 + \gamma_1 \tau_0}
\]

which, if \( \tau_0 \gamma_1 \ll 1 \) also, may be further simplified to

\[
R \approx \tau_0 \gamma_2.
\]

Referring back to our two-stream solution (6.34), we see that for the hemispheric constant method, \( \gamma_2 = 2\omega b \). Thus, the albedo of a thin layer depends linearly on the optical depth, the single-scattering albedo and the backscattering coefficient, a result consistent with the single-scatter result (6.24). However, referring back to (6.43) we see that Eddington’s method implies \( \gamma_2 = -[1 - \tilde{\omega}(4 \tilde{g})] / 4 \). For isotropic scattering (\( g = 0 \)), this gives an albedo too low by more than a factor of two. Also, the albedo is nonzero when \( \tilde{\omega} = 0 \) (no scattering). These results demonstrate that the \( N = 1 \) Eddington truncation of the radiance expansion can lead to unphysical results in low-albedo cases.

In optically thick layers (\( \tau_0 \rightarrow \infty \)),

\[
R \rightarrow \frac{1}{v} = \frac{\gamma_2}{\gamma_1 + \xi}.
\]

Thus the albedo of very optically thick layers becomes independent of \( \tau \) and depends only on the single-scattering parameters and the method. This follows from physical considerations, since the transmittance decreases to zero. This result, therefore, does not depend on our choice of diffuse vs. collimated incidence. The albedo concept, however, inherently assumes there is no emission (or, if there is, that the emitted photons are not counted as part of the reflected radiation).

Note that these albedo results hold only for diffuse incident radiation, with no backscattering from optical depths greater than \( \tau_0 \). If the incident radiation is collimated, a different result occurs (see Problem 9). If the layer is instead backed by a reflective boundary of known albedo, then the albedo of the layer and its lower boundary together can be found by the adding method (Section 6.6.1).
Note on cloud and aerosol layers

These results have practical implications for planetary albedo, to the extent that it depends on atmospheric scattering. On many planets including Earth, a dominant contribution to the albedo is made by cloud layers, whereas the atmosphere is optically thin when clouds are absent. Albedo is always proportional to optical depth for small $\tau$, until transmission through the cloud drops well below unity, whereupon albedo begins to plateau. The final value at $\tau = \infty$ depends on $\tilde{\omega}$ and the phase function, and approaches unity for conservative scattering. Since optical depth depends on cloud water content and is inversely proportional to droplet size, clouds with more water or smaller droplets will reflect more energy to space. A significant issue in climate modeling is capturing the variations in these cloud parameters so that their impact on radiation can be appropriately represented.

Albedo also shows a general tendency to increase at low $\mu$ and $\mu_0$. This is because the optical path through the layer becomes large at glancing incidence, allowing more scattering opportunities. On the other hand, low angles of incidence imply low incident vertical fluxes, so the backscattered radiation itself generally does not increase with decreasing $\mu_0$.

Aerosol layers can also contribute to planetary albedo. However, the dominant impact of aerosol layers is on transmission, due to their lower $\tilde{\omega}$ values ($\sim 0.9$ as opposed to nearly unity for clouds at visible wavelengths). Thick aerosol layers can substantially reduce surface insolation. Whether an aerosol layer increases or decreases planetary albedo depends on how the albedo of the layer compares to that of the underlying surface and depends on the aerosol type.

6.8.3 A comparison of diffusivity, stream, and polynomial approaches

It is useful to examine a bit further the behavior of these methods in the nonscattering limit, even though they are not usually used in this limit\(^2\). Recall that in discussing non-scattering solutions we employed a diffusivity factor $\tilde{d}$ to rescale optical depth. Examining (6.52) when $\tilde{\omega} = 0$ and $S = 0$, we see that incoming fluxes decay at a rate $\xi_\tau$; thus, we may identify $\xi$ with $\tilde{d}$. When $\tilde{\omega} = 0$, $\xi = \gamma_1$ for the hemispheric constant method, we then have $\xi = 2[1-\tilde{\omega}(1-b)] = 2$, compared to the “optimal” $\tilde{d}$ of $1.66$. Other two-stream methods produce diffusivity factors closer to $1.66$ (two-stream quadrature gives $1.73$), and can be doctored to give exactly $1.66$ if desired.

It is illustrative to consider why the hemispheric-constant extinction rate is too rapid in nonscattering media. The method assumes that the downward radiances remain isotropic as they penetrate into the layer. What would really happen, with no sources, is that the radiances at $\mu \sim 0$ would diminish very rapidly, causing the radiance distribution to be increasingly weighted toward more vertical, less attenuated angles as shown in Figure 6.2. One may observe this by swimming well below the surface of an ocean or lake and looking up to see the most brightness in the overhead direction. The hemispheric-constant method, however, essentially keeps redistributing the radiance into the most attenuated directions, exaggerating the extinction rate if the more horizontal radiances are not being replenished by actual sources. Of course, if nearly

---

\(^2\)Much of the material here comes from Obrien et al, 1997.
isotropic sources were important (hot media or large $\omega$) then the field really could remain isotropic and this approximation should work much better. Note also that $\xi$ decreases as $\omega$ increases, and may approach values much lower than 1.66, so this problem decreases when scattering re-enters the picture.

The important point to be made here is not the minutia of various approximate methods, but the tendency of radiation to become more concentrated toward the vertical directions as it propagates; the dependence of this process on the nature of the source $J_\nu$; and its impact on quantitative calculation of radiative transfer. Diffusivity and two-stream approximations may be expected to have difficulty in cases where $J_\nu$ is small or highly anisotropic. Nonetheless, experimentation shows that they can obtain surprisingly reasonable fluxes and heating rates. The basic reason for this will be investigated further in Section 7.4, in the context of heating of Earth’s atmosphere.

The nonzero albedo predicted by the Eddington method for $\omega = 0$ recurs at all optical depths because the truncated polynomial representation of $I$ does not allow downward and upward fluxes to be independent, as they must be in reality if $\omega = 0$. Since Eddington’s method is normally used for nearly conservative scattering, this problem may not be critical. In general, “stream” or discrete ordinates methods are more accurate for cases with weak scattering than are polynomial-expansion-type methods of similar complexity, while the latter are sometimes more useful in cases of scattering problems with complicated phase functions.
6.8.4 The delta-Eddington method for scattering by large particles

One particular pathology that develops in the phase function is that for scattering off of macroscopic particles having large size parameters \( x \) (for example, cloud droplets illuminated by visible radiation), the scattering pattern develops a huge forward-scattering lobe (cf. Chapter 4). The fraction \( f \) of scattered power contained in the lobe can approach unity if \( x \) is large and \( m \) is not too large. Since accurate polynomial expansion of the phase function requires \( N \sim 3x \) or better, solutions based on low-order expansions such as the Eddington approximation will perform poorly.

A common trick in treating this situation is to approximate \( P(\cos \Theta) \) as the sum of a smooth function \( P'(\cos \Theta) \) and a delta function in \( \cos \Theta \sim 1 \), with weights of \( 1 - f \) and \( f \), respectively. Then photons that scatter into the forward peak (fraction \( f \) of total scattering events) are in effect treated as unscattering photons. The same radiative transfer equations are used except that the appropriately reduced optical depth and increased single-scattering albedos

\[
\omega' = \frac{\bar{\omega}(1 - f)}{1 - \bar{\omega}f} \quad \tau' = (1 - \bar{\omega}f)\tau
\]

are substituted for the original ones. This strategy, known as the delta scattering method, can be combined with any radiative transfer method (e.g., “delta-two stream”) but is most associated with the Eddington approximation (“delta-Eddington”).

It is also common in treating cloud and aerosol problems to employ the Henyey-Greenstein phase function

\[
P(\cos \Theta) = \frac{1 - g^2}{(1 + g^2 - 2g\cos \Theta)^{3/2}}
\]

which is found to represent smoothed Mie phase functions fairly well and can be conveniently represented by the Legendre expansion

\[
P(\cos \Theta) = \sum_{l=0}^{\infty} (2l + 1)g^l P_l(\cos \Theta).
\]

6.9 Three dimensional radiative transfer

Recently, there has been a lot of work on fully three-dimensional radiative transfer calculations. Atmospheric applications have focused almost entirely on the problem of solar radiation propagating through water clouds, where scattering is dominant and the scattering medium (cloud) has complex structure. Remote sensing of cloud properties using backscattered sunlight can be sensitive to the three-dimensional structure of clouds, especially cumulus clouds. Also, it has been found that absorption within clouds is systematically underestimated by plane-parallel calculations (by something like 10%, although absorption is small in any case) even when the albedo is correct.

\(^3\)There is some confusion in the literature surrounding the “delta” method nomenclature, since it is sometimes used to represent simply the direct-diffuse decomposition of \( I \) even when the adjustments in this section are not needed, i.e., for solar radiative transfer through gases.
Cloud albedos and transmittances can be substantially over- and underestimated respectively by plane-parallel calculations based on observed horizontal average water content, due to horizontal heterogeneity in real clouds and the nonlinear dependence of $R$ and $\mathcal{T}$ on $\tau$.

Specifying the three-dimensional structure of a cloud is itself a difficult problem, since no observing system is currently able to measure condensed water contents in three dimensions. A popular method is to model the cloud statistically—often as a fractal—and then compute a (random) synthetic 3-D distribution of cloud water from the statistical model using a computer.

The most popular way of modeling radiative transfer in 3-D is the so-called Monty Carlo method, in which calculations are performed for individual photons. The trajectory of each photon is calculated probabilistically according to the RTE. Fluxes are then computed by brute force methods. Radiances can also be computed (for example, to simulate the observation by a satellite); in this case, photons are usually run backward through the RTE (which is fully reversible) starting from the eventual point and direction of observation.

**Exercises**

6.1  a) Calculate the flux transmission $\tau_\nu$ of a layer of thickness $(\tau_\nu)$, for the special case where the incident radiation is isotropic and where $J_\nu = 0$. b) Use this to calculate an effective diffusivity factor as a function of $\tau_\nu$. Your answer should approach 2 or 1 as $\tau_\nu \to 0$ or $\infty$, respectively, and it should equal 1.66 when $\tau_\nu \approx 0.40$.

6.2  By integrating by parts the last term in (6.13), and noting that

$$dB_\nu = \frac{dB_\nu}{d\tau'} d\tau' = \frac{dB_\nu}{dz'} dz',$$

derive (6.14).

6.3  Find the expression analogous to (6.23) for the single-scattered, transmitted radiance at the bottom of a thin layer.

6.4  Using the single-scatter approximation, calculate the backscattered radiance distribution for $\tau_0 = 0.05$ and $\bar{\omega} = 0.5$, when the incident radiation is isotropic. Assume a Rayleigh scattering phase function with no emission.

6.5  Find an expression for the reflectance $R$ of two layers in the adding method, analogous to that given for $\mathcal{T}$. If a layer 125 m thick has a transmittance of 0.95 and a reflectance of 0.03 for isotropic incident radiation, calculate $R$ and $\mathcal{T}$ of a 1 km layer of the same medium using this method. Neglect changes in the angular distribution of the radiation. *Answer: $\mathcal{T} = 0.677$, $R = 0.174$.*

6.6  Explain under what conditions, if any, an isotropic radiation field incident on a medium will remain isotropic at all depths (consider the phase function and single-scatter albedo).

6.7  Show that the method of moments yields the same result (6.43) for $N = 1$ as was
obtained using polynomial expansion.

6.8 Calculate the near-infrared albedo of a cloud of very large water content given diffuse incidence, if $\tilde{\omega} = 0.8$, following (a) the Eddington and (b) two-stream approximations with isotropic scattering, and (c) two-stream with $b = 0.25$ (a more realistic value for Mie scattering at moderate $x$). Roughly what optical depth is necessary to approach the maximum albedo in case (c)?

6.9 Obtain a two-stream (hemispheric constant) solution for the albedo of a layer backed by a dark surface, starting from (6.52), for the case of a collimated incident beam from zenith $\mu_0$. Carefully articulate the boundary conditions before proceeding. Your albedo should reduce to

$$R = \frac{\tilde{\omega} [\gamma_1 \beta_0 + \gamma_2 (1 - \beta_0) + \beta_0 \xi]}{(1 + \xi \mu_0)(\xi + \gamma_1)}$$

for infinite optical depth.
Chapter 7

Radiative heating and cooling

7.1 General concerns

The convergence of the net radiative energy flux

\[ h = -\nabla \cdot \mathbf{F} \]  

(7.1)

will lead to a heating rate (rate of temperature change per unit time) of

\[ Q = \frac{1}{\rho c_p} h. \]  

(7.2)

In general, for quantitative calculations we must consider the behavior as a function of wavelength. The “monochromatic” heating per unit frequency will be denoted \( Q_\nu \) so that

\[ Q = \int_0^\infty Q_\nu d\nu. \]

For a fluid layer in a horizontally-invariant (plane parallel) radiation field, we may write

\[ \mathbf{F}_\nu = (F'_\nu - F''_\nu) \mathbf{k}, \]

where the field is now characterized by separate upward and downward vertical fluxes. Here, \( \mathbf{k} \) refers to the upward unit vector. Then \( Q_\nu \) is

\[ Q_\nu = \frac{1}{\rho c_p} \left( \frac{dF'_{\nu}}{dz} - \frac{dF''_{\nu}}{dz} \right), \]  

(7.3)

This can be written in optical depth coordinates using (2.2), (2.11) and (2.15) as

\[ Q_\nu = \frac{1}{\rho c_p} \left( \frac{dF'_\nu}{d\tau_\nu} - \frac{dF''_\nu}{d\tau_\nu} \right) e_{\nu,\nu'}, \]  

(7.4)

where we assume \( \tau \) decreases with height (otherwise a minus sign is needed), and define

\[ e_{\nu,\nu'} = \sum_i (k_{\nu,i} + s_{\nu,i}) \rho_i. \]  

(7.5)
as the multicomponent volume extinction coefficient. We now recognize that multiple constituents of total density \( \rho \) may be present, but only specific constituents \( i \) contribute to the optical depth at a given wavenumber \( \nu \). The single coefficient \( e_{\nu,i} \) includes all such constituents.

If the mass mixing ratios\(^1\)
\[
q_i \equiv \frac{\rho_i}{\rho}
\]
of all absorbers are constant throughout the medium, (7.4) indicates that the rate of temperature change is independent of density, since density affects heat capacity and volume absorptance equally. Further, if we are concerned with transfer through a planetary atmosphere or ocean, the hydrostatic relation
\[
dp = -\rho g dz
\]
(with \( g \) the gravitational constant) enables us to write
\[
\tau_{\nu} = \left( \frac{e_{\nu,i}}{\rho g} \right) \rho.
\]
For simplicity we will assume that \( g \) is constant with height (a quite adequate approximation for the inner planets). Then, if the atmosphere is also well mixed, the factor in parentheses is constant throughout the medium. Thus, when dealing with well-mixed gases such as \( \text{CO}_2 \), our \( \tau_{\nu} \) coordinate system is equivalent to the pressure coordinate system familiar to meteorologists, but with a scaling constant that varies with \( \nu \). The convergence of radiation in \( \tau_{\nu} \)-coordinates is also then proportional to \( Q_{\nu} \). Our recourse (7.5) to volume extinction quantities is symptomatic of having to consider multicomponent media that are not well mixed; but we will put off this concern until after considering some simple cases of absorption by monolithic media. When we do consider heterogeneous media, we will assume that absorption at a particular frequency can be ascribed to a single absorber of mixing ratio \( q \) and absorption coefficient \( k_{\nu} \).

7.2 Heating of a cold layer

The simplest case of heating is that due to hot incident radiation on a cold layer so that emission is negligible. For simplicity, we will also neglect scattering so that \( e_{\nu} = k_{\nu} \). This provides qualitatively instructive results, and is a good approximation for, e.g., heating of clean lakes and oceans (where scattering is small). We assume that \( q \) is constant throughout the medium.

With these simplifications there is only a downward radiation stream. Assuming collimated incidence of flux \( \pi F_{\nu,0} \), we may apply (7.4) to the general solution (2.18) using also (6.1) to obtain
\[
Q_{\nu} = \left[ \frac{q k_{\nu}}{\mu c_p} \right] \pi F_{\nu,0} \exp\left(-\frac{\tau_{\nu}}{\mu}\right),
\]
The heat deposited simply decreases exponentially with optical path. The factor in brackets is just the derivative of optical path with respect to heat capacity.

\(^1\)Mixing ratio is usually defined as \( q_i \equiv \rho_i / (\rho - \rho_i) \) although the distinction is not important here.
If we have a layer of finite optical thickness backed by a surface of reflectivity \( R_{g,\nu} \), the heating including one surface reflection will be

\[
Q_\nu = \frac{q k_\nu}{\mu c_p} \pi F_{\nu,0} \left[ \exp\left(\frac{-\tau_\nu}{\mu}\right) + T_\nu^\mu R_{g,\nu} \exp\left(-d(\tau_{g,\nu} - \tau_\nu)\right) \right] \tag{7.9}
\]

where we employ the layer beam transmittance (6.3), assume diffuse reflection from the surface, and conveniently treat this reflection using the diffusivity approximation.

### 7.3 Net heating in warm layers

Net heating is more complicated when the layer can emit as well as absorb. The results will obviously depend on the temperature profile of the layer as well as the boundary conditions. To obtain the heating (or cooling) in this case, we begin with the diffusivity flux equation (6.8), employing the effective path length (6.9) and neglecting scattering:

\[
\pm \frac{dF_\nu}{d(\tilde{\tau}_\nu)} + F_\nu = \pi B_\nu(T) \tag{7.10}
\]

Recall that plus/minus in (7.10) refers to the flux toward/away from increasing \( \tau_\nu \). Noting that \( k_\nu \) is independent of depth, we can combine this with (7.3) and (2.15) to obtain

\[
Q_\nu = \frac{q k_\nu d}{c_p} [F_\nu^+ + F_\nu^- - 2\pi B_\nu(T)]. \tag{7.11}
\]

Thus the net heating is determined by the difference between the intensity of the radiation streams arriving at a given position, and the intensity of blackbody emission from that position.

Now consider two layers with their interface at \( \tau_\nu = 0 \), so that the lower layer contains positive values of \( \tau_\nu \) up to a total thickness \( \tilde{\tau}_{\nu(L)} \) while the upper layer contains negative values down to \( \tilde{\tau}_{\nu(U)} < 0 \). Temperature is continuous across the interface and through the two layers. The layers are bounded by surfaces at temperatures \( T(U) \) and \( T(L) \) which do not necessarily match the adjacent layer temperatures. The upward flux from the lower layer is given by (2.18), if we employ (6.9) and note the change in flux direction:

\[
F_\nu^+(0) = F_\nu^+(\tilde{\tau}_{\nu(L)}(L)) e^{-\tilde{\tau}_{\nu(L)}} + \pi \int_0^{\tilde{\tau}_{\nu(L)}} B_\nu(\tilde{\tau}_{\nu(L)}) e^{\tilde{\tau}_{\nu(L)} - \tilde{\tau}_{\nu(U)}} d\tilde{\tau}_{\nu(U)} \tag{7.12}
\]

An identical equation holds for \( F_\nu^+(0) \) with \( \tilde{\tau}_{\nu(L)} \rightarrow -\tilde{\tau}_{\nu(U)} \) everywhere except in the argument of \( B_\nu \).

#### 7.3.1 Heating within thick and thin layers

We first consider the thin limit. Where \( \tilde{\tau}_{\nu(L)} \) and \( \tilde{\tau}_{\nu(U)} \) both approach zero, we have fluxes that approximately equal the boundary values, therefore

\[
Q_\nu \approx \frac{\pi q k_\nu d}{c_p} [B_\nu(T(U)) + B_\nu(T(L)) - 2B_\nu(T)]. \tag{7.13}
\]
In this case, heating will be zero for $T$ such that the Planck emission from the layer equals the mean of that at the two boundaries. In the Rayleigh-Jeans limit (for emission at terrestrial temperatures, toward the far infrared) we have $B_\nu \propto T$, so that the temperature will tend toward the mean of $T(U)$ and $T(L)$; in the near-IR, $B_\nu$ is highly nonlinear in $T$ so the profile will tend toward a value close to the warmer of the two boundaries. Thus, we see that at two wavelengths having the same $k_\nu$, $Q_\nu$ will not necessarily be the same even though the layer is homogeneous.

Now consider the thick limit, where $\tau_\nu(L)$ and $\tau_\nu(U)$ both approach infinity. In this case the boundary contributions vanish and we have, from (7.11) dropping primes on the variable of integration,

$$Q_\nu = \frac{\pi q k_\nu \tilde{d}}{c_p} \left[ \int_0^{\infty} B_\nu(\tilde{\tau}_\nu) e^{-\tilde{\tau}_\nu} d\tilde{\tau}_\nu + \int_0^{\infty} B_\nu(-\tilde{\tau}_\nu) e^{-\tilde{\tau}_\nu} d\tilde{\tau}_\nu - 2B_\nu(0) \right] \tag{7.14}$$

It is now useful to exploit the Taylor expansion

$$B_\nu(\tilde{\tau}_\nu) = B_\nu(0) + \tilde{\tau}_\nu \left( \frac{d B_\nu}{d \tilde{\tau}_\nu} \right)_0 + \frac{1}{2!} \tilde{\tau}_\nu^2 \left( \frac{d^2 B_\nu}{d \tilde{\tau}_\nu^2} \right)_0 + \ldots \tag{7.15}$$

Inserting this into (7.14) and collecting terms we find that the zero, first, and all higher odd order terms cancel, leaving

$$Q_\nu = \frac{\pi q k_\nu \tilde{d}}{c_p} \int_0^{\infty} (\tilde{\tau}^2 \left( \frac{d^2 B_\nu}{d \tilde{\tau}_\nu^2} \right)_0 + O(4)) e^{-\tilde{\tau}} d\tilde{\tau}_\nu. \tag{7.16}$$

If temperature is not changing too rapidly with optical depth then we may ignore the fourth-order and higher terms in this expansion. Performing the integral with only the leading term yields

$$Q_\nu \approx \frac{2\pi q k_\nu \tilde{d}}{c_p} \left( \frac{d^2 B_\nu}{d \tilde{\tau}_\nu^2} \right)_0. \tag{7.17}$$

Equation (7.17) has the form of a diffusion equation, indicating that with smoothly varying temperature, thermal radiative transport is approximately diffusive over distances longer than one photon mean free path. The assumption of equality in (7.17) is known as the *diffusion approximation*, valid for example within the Sun out to about 0.95 times the solar radius. However, in view of (7.14), we note that over distances short compared to a photon path the behavior deviates strongly from that of photon diffusion. This has important consequences for planetary radiative equilibria, as discussed later, since most planetary atmospheres are thin compared to a photon mean free path over significant parts of the spectrum.

### 7.3.2 Heating near boundaries (“Cooling to space”)

The last case we consider is the semi-thick limit where only $\tilde{\tau}_\nu(L)$ approaches infinity. We will denote $\tilde{\tau}_\nu(U)$ more succinctly as $\tilde{\tau}_1$ and assume that $T(U) = 0$. The downward flux at $\tilde{\tau}_\nu = 0$ is

$$F_\nu^+(0) = \pi \int_0^{\tilde{\tau}_1} B_\nu(\tilde{\tau}') e^{-\tilde{\tau}'} d\tilde{\tau}' \tag{7.18}$$

\[\text{This is different from the *diffusivity approximation* of pg. 99}\]
where we have taken \( \bar{\tau}' = -\bar{\tau}_u \) to be positive. Once again expanding \( B_\nu \) about \( \bar{\tau}_u = 0 \) and performing the integrations by parts, we have

\[
F^\downarrow_\nu(0) = B_0(1 - e^{-\bar{\tau}_1}) - B_1[1 - (1 + \bar{\tau}_1)e^{-\bar{\tau}_1}] + B_2[1 - (1 + \bar{\tau}_1 + \frac{\bar{\tau}_2}{21})e^{-\bar{\tau}_1}] - \ldots \quad (7.19)
\]

where we have denoted the \( n \)th derivative of \( \pi B_\nu \) evaluated at \( \bar{\tau}_n u = 0 \) by \( B_n \). Following a similar procedure for the upward flux from the semi-infinite lower layer, but defining \( \bar{\tau}' = +\bar{\tau}_\nu \), yields

\[
F^\uparrow_\nu(0) = B_0 + B_1 + B_2 + \ldots \quad (7.20)
\]

Inserting these two expressions into (7.11), noting that \( B_\nu = B_0/\pi \) at \( \bar{\tau}_\nu = 0 \), and canceling terms, we have

\[
Q_\nu = \frac{k_s q d}{c_p} \left\{ B_0(-e^{-\bar{\tau}_1}) + B_1[e^{-\bar{\tau}_1}(1 + \bar{\tau}_1)] + B_2[2 - e^{-\bar{\tau}_1}(1 + \bar{\tau}_1 + \frac{\bar{\tau}_2}{21})] + \ldots \right\}.
\]

This expression gives the cooling rate at a point \( \tau = 0 \), at a specified temperature, as a function of how much absorber exists above that point. If the entire layer is isothermal \((B_1 = B_2 = \ldots = 0)\), then we simply have “cooling to space” that decreases exponentially away from the upper boundary. If the amount of absorber above the point is infinite, the zeroth and first order terms approach zero and we obtain (7.17). The diffusive behavior noted for the optically thick case continues for all values of \( \bar{\tau}_1 \), but with a diffusion rate constant that decreases by a factor of two going from the thick limit to \( \bar{\tau}_1 \to 0 \).

For finite \( \bar{\tau}_1 \) and nonuniform \( T \), we have exponentially decaying heating that is modified by terms due to the slope and curvature of \( B_\nu(\bar{\tau}_u) \). However, inspection of these two terms reveals that their parenthesized factors are the first- and second-order Taylor expansions, respectively, of \( \exp(\bar{\tau}_1) \). Thus, for values of \( \bar{\tau}_1 \) between zero and order unity, both factors in brackets will be nearly constant; as long as temperature profiles are not changing drastically, they will not vary much over several optical depths. These extra terms arise from the changes in emission far from \( \bar{\tau} = 0 \) implied by extrapolation of the second-order fit to \( B_\nu(\bar{\tau}_u) \). In actual atmospheres in which radiative, convective, and diffusive processes all act to limit the curvature of the temperature profile over large vertical distances, it is typically the case that the net heating due to these higher-order terms is quite small. Thus, the simple, exponential “cooling to space” behavior of isothermal layers is typically close to the actual \( Q_\nu \). Also, although the atmosphere is not semi-infinite, the surface emissivity is typically high and its temperature is constrained by the same diffusive mechanisms, so it does not radically alter the picture. These matters are taken up further in Chapter 8.

### 7.3.3 Brightness temperature

One of the most important concepts in applied radiative transfer arises from the above discussion. Comparing (7.15) and (7.20) we see that, to first order in \( \bar{\tau}_\nu \), the emitted radiation has a value equal to the blackbody value \( B_\nu(\bar{\tau}_\nu = 1) \) one optical path into
CHAPTER 7. RADIATIVE HEATING AND COOLING

the medium. The higher order terms differ by a factor of $n!$, indicating that the outgoing flux will be somewhat greater or less than $\pi B_\nu(\bar{\tau}_\nu = 1)$ if temperatures at greater depths are greater or less than those obtained by a linear extrapolation. However, for the same reasons as given above, such curvatures of the emission profile are typically modest at most wavelengths in planetary atmospheres. Thus, the emitted monochromatic radiance (or flux) from a layer can usefully be interpreted in terms of an effective brightness temperature defined by

$$F'^\downarrow_\nu = \pi B_\nu(T_b),$$  \hspace{1cm} (7.22)

where $T_b$ approximately equals the actual temperature one optical depth into the medium.

This is invaluable in interpreting satellite observations, and is the basis for remote sensing of cloud heights and atmospheric temperature profiles. Though accurate retrieval of vertical atmospheric temperature profiles from space requires more careful inversion techniques to account for the higher-order terms above, the process can be understood qualitatively as the observation of $T_b$ at different wavelengths $\nu$—chosen such that $\bar{\tau}_\nu = 1$ falls at different altitudes in the atmosphere—and consequent determination of $T$ at each of those levels. Not surprisingly, determination of $T(z)$ from observed emission becomes inaccurate when there are strong inversions or sudden jumps in $T$, unless these jumps are known and modeled as part of a careful inversion process.

Note that the optical path $\bar{\tau}_\nu$ is equivalent to optical depth $\tau_\nu$ only for radiances at $\mu = 1$, e.g. direct overhead observation. The “viewing depth” will be shallower as $\mu$ of the observed radiance decreases. Since $T$ typically decreases with height, this usually leads to reduced emitted radiance from the same atmosphere. This effect is known as “limb darkening” and must be taken into account in interpreting any satellite observation where $B_\nu$ changes significantly over unit optical depth. “Limb brightening” can occur in the stratosphere or when viewing backscattered radiation.

7.4 Heterogeneous media and broadband heating

So far we have found that in many situations $Q_\nu$ decreases approximately exponentially away from the exposed boundary (e.g., top of atmosphere). This behavior is exact for both self-cooling and exogenous heating of isothermal, non-scattering layers. The $Q_\nu$ decay rate (in optical depth $\tau_\nu$) is $\tilde{d}$ or $1/\mu$ for diffuse or collimated radiation respectively. In the case of a grey gas, the same behavior holds for $Q$, since by construction $\tau$ is independent of wavelength. We will now explore how exponentially decaying $Q_\nu$ plays out in more realistic media.

In a non-grey treatment of a homogeneous medium, things are still fairly simple, since $\tau_\nu \propto \nu$ for all $\nu$. Now $Q$ will simply be a sum of many decaying exponentials with different decay constants

$$Q(p) \propto \int_0^\infty F_{\nu,0} \exp\left(\frac{k_0 q p}{g}\right) d\nu.$$  

The final shape of heating will be determined by the distribution of $\tau_\nu$ through the range of incident and radiated frequencies. The fluxes predicted by the random band model of
section 5.10.3 yield the profile shown in Figure 5.11, which resembles the superposition so described. In the random band model, the heating rate actually becomes infinite at the top of the layer.

### 7.4.1 The Chapman layer

In heterogeneous atmospheres, $\tau_\nu$ and $p$ are no longer proportional. Going from $Q_\nu$ to $Q$ requires conversion from $\tau_\nu$ into a mass- or position-based coordinate. Customarily, height $z$ is chosen.

In planetary atmospheres where absolute temperature does not deviate greatly as a percentage of its mean value, the concentration $\rho$ is approximately

$$\rho(z) = \rho(0)e^{-z/H} \quad (7.23)$$

where $H$ is a the atmospheric scale height ($\sim 8 \text{ km on Earth}$). Obviously this relation also holds for the density $\rho_i$ of any individual well-mixed species $i$. A similar relation holds—to varying degrees of approximation—for certain non-well-mixed trace constituents on Earth such as water vapor and aerosols, but with much shallower scale heights $H_i$ (about $2 \text{ km for these two examples}$). In the upper atmosphere, molecular diffusion dominates turbulent mixing so that gas profiles tend toward their individual natural scale heights (which decrease with increasing atomic weight) rather than the lower atmospheric $H$. Some gases, such as ozone, have profiles completely unlike (7.23). However, near absorption lines even these gases will usually be approximately exponential in $z$ for the topmost few optical depths, in which case the discussion below will still apply.

Assuming for simplicity that there is only one absorbing constituent $i(\nu)$ at $\nu$, neglecting scattering, we have

$$d\tau_\nu = \rho_i k_\nu dH. \quad (7.24)$$

Inserting (7.23) written for the absorbing species and integrating this from $z$ to infinity gives the optical depth at height $z$,

$$\tau_\nu = \rho_i(0) k_\nu H_i \exp \left(-\frac{z}{H_i}\right). \quad (7.25)$$

From the flux solution and (7.1), we have

$$h_\nu = \pi F_{\nu,0} \rho_i k_\nu d \exp(-\tilde{\alpha}_\nu), \quad (7.26)$$

where we have defined a monochromatic radiation convergence $h_\nu$, and have set $\tau_\nu = d\tilde{\alpha}_\nu$ although for solar incidence (6.1) can be used instead. Insertion of the previous expressions for $\tau_\nu$ and $\rho_i$ gives

$$h_\nu = \pi F_{\nu,0} \rho_i(0) k_\nu d \exp\left(-\frac{z}{H_i} - \tilde{\alpha}H_i \rho_i(0) k_\nu \exp\left(-\frac{z}{H_i}\right)\right). \quad (7.27)$$

We now define

$$z' = \ln(\tilde{\alpha}H_i \rho_i(0) k_\nu)H_i. \quad (7.28)$$
Note that \( \bar{\tau}_\nu(z') = 1 \), and that
\[
\frac{z'}{H_t} \approx \ln(\bar{\tau}_\nu(z = 0)).
\]
Inserting this definition into (7.27) yields
\[
h_\nu = \frac{\pi F_{\nu,0}}{H_t} \exp \left( \frac{z' - z}{H_t} - \exp \left( \frac{z' - z}{H_t} \right) \right). \tag{7.29}
\]
This expression for the convergence of energy in \( z \) defines what is known as a Chapman layer. We see that \( h_\nu \) approaches zero at large \( z \) due to vanishing absorber density, and approaches small values at the surface if \( z' \sim H_t \) or greater due to the insulating effect of the atmosphere above. The maximum of \( h_\nu \) occurs exactly at \( z = z' \), as can be seen by setting
\[
\frac{\partial h}{\partial z} = 0.
\]
Thus, \( h_\nu \) peaks at one optical depth into the medium, the same level as the effective radiating level for thermal emission. Significant values of \( h_\nu \) occur in a layer confined mainly to \( z' \pm H_t \); this means there will be significant energy convergence all the way to the ground unless the total optical path is greater than one.

If the absorber scale height \( H_t = H \), then from (7.2) the heating rate \( Q_\nu \) will decay exponentially in \( \tau_\nu \) or \( \rho \) as found previously. However, if \( H_t \ll H \) (as in the case of water vapor in Earth’s atmosphere) then \( \rho \) will not vary greatly over the scale height \( H_t \), and the heating \( Q_\nu \) will also develop a local maximum within the medium, near that of \( h_\nu \). The Chapman layer concept is quite important since it quantifies the localization of radiative effects about \( z' \sim z \) for a variety of situations, with notable Earthly examples including solar heating of ozone and thermal emission from water vapor, aerosols, and clouds at various wavelengths.

### 7.4.2 Comments on the broadband heating

We are now ready to consider \( Q \), the heating integrated over all frequencies. At each \( z \), one must simply sum the \( h_\nu \) over all \( \nu \) and then divide by \( \rho c_p \). In practice, this is done on a computer using band models to represent the effects of different absorbers, and using the adding method and/or multistream methods to determine the fluxes and their convergence in each band over a discrete set of vertical layers. However, certain qualitative insights are possible here.

If we look at (7.29) not as a function of \( z \) but as a function of the absorber characteristics, we may ask: what absorber parameters yield the greatest heating or cooling at a given level of interest \( z' \)? The key parameters are the scale height \( H_t \) and volume absorption coefficient \( h_\nu \rho \), though in place of the latter it is convenient here to consider the difference \( z' - z \) which carries the same information. The local maximum of \( h_\nu \) obtained by setting
\[
\frac{\partial h_\nu}{\partial (z' - z)} = 0
\]
clearly occurs again where \( z' = z \), i.e., for gases and at frequencies where the densities and mass absorption coefficients are just right so that unit optical depth coincides with the chosen level \( z \). This is true regardless of \( H_i \). If we examine the dependence of \( h_\nu \) on \( H_i \), we find that at the peak \( z' = z \), the heating is inversely proportional to the absorber scale height. This is because larger scale heights cause the radiative effect to be spread out more.

If we neglect variations of \( F_{\nu,0} \) over the spectrum, we are led to the conclusion that \( Q \) at a given level \( z \) is caused largely by species whose optical depth from \( z \) to the boundary is near unity (to within a factor of \( \epsilon \)) at that level. Recalling the line structures described in Section 5.9, it is apparent that any sufficiently strong line will meet this criterion at some distance \( \delta \nu \) from the line center. Furthermore, the distribution of \( k_\nu \), that occurs over a single strong line is huge, to say nothing of the range over many lines or over many gases. Typical \( k \)-distribution models (c.f. Section 5.10) have 10th to 90th percentile intervals that span one or two orders of magnitude in \( k \). These ranges of absorption coefficient, hence total optical depth, are of similar or greater breadth as the range of \( \tilde{\tau}_\nu \) that contributes significantly in (7.29), so that at most values of \( \nu, \tau_\nu \) is too far away from unity for a significant contribution to \( Q \). So, in practice, the influence of parts of the spectrum where \( z' = z \) is not only important but actually tends to dominate \( Q \). The heating or cooling at a given level can often be ascribed to a particular band of a particular absorber, particularly for absorbers of smaller scale heights.

Figure 7.1, which shows the net thermal cooling vs. \( \nu \) and \( p \), illustrates this characteristic. Note that at most wavelengths, most of the cooling is concentrated within a narrow range of altitudes. Though cooling prevails overall, heating occurs near the tropopause due to the 15-\( \mu \)m band of CO\(_2\), and throughout the lower stratosphere due to the 9.6 \( \mu \)m O\(_3\) band. Water vapor dominates the cooling in the lower troposphere, acting broadly across the spectrum. The net cooling turns out to be a nearly constant value of about 2 K day\(^{-1} \) through most of the troposphere, dropping to near zero through the tropopause and lower stratosphere, then increasing again to large values that balances the intense heating by solar absorption in the upper atmosphere.

### 7.4.3 Line broadening effects on heating

So far, we have assumed that \( k_\nu \) is spatially invariant, i.e., that volume absorption is proportional to the concentration of the absorber. This is not strictly true, because line breadth can vary particularly due to pressure. Pressure and temperature broadening have modest effects for sufficiently small \( H_i \) since all the action is confined to a narrow range of \( p \) and \( T \), but the broadening becomes quite a problem when \( H_i \sim H \).

We may discuss the effects qualitatively in the following way. Recall that near the center of a Lorentzian line, at a given temperature,

\[
k_\nu \propto \frac{S}{p}
\]

while in the wings

\[
k_\nu \propto Sp.
\]
In the wings, we see that the variation of volume absorption coefficient with height is identical to that in a gas of constant $k_0$ if the latter’s density is taken to be

$$\rho_i = \frac{p}{p(0)} \rho_i(0) \exp \left(-\frac{z}{H_i}\right)$$

(7.30)

and the mass absorption coefficient is taken as its true value at the surface. Since pressure decreases exponentially with the scale height $H$, we have

$$\rho_i = \rho_i(0) \exp \left(-\frac{z}{H_i}\right)$$

where the “reduced scale height” is

$$H' = \frac{H_i H}{H_i + H}.$$ 

In other words, the line wings behave as though they were non-broadened wings of a gas with a shallower scale height than the actual scale height of the gas (half for well-mixed gases where $H_i = H$). This, incidentally, helps wing absorption exhibit the Chapman behavior described above. Temperature broadening has a qualitatively similar (since $T$ increases with $p$ where most of the atmospheric mass is) though much smaller effect.

Near line centers a similar argument holds, except the reduced scale height is

$$H' = \frac{H_i H}{H_i - H},$$
which becomes infinite if the gas is well mixed. This perplexing result occurs because
we have neglected Doppler broadening, without which the line becomes infinitely nar-
row as $z \to \infty$. Treatment of line centers requires consideration of both Doppler and
collision effects and is beyond our scope here. The qualitative result is that at line cen-
ters, absorption occurs higher in a hydrostatic atmosphere than it would in an isobaric
medium of otherwise equivalent properties.

### 7.5 Water vapor on Earth

Due to its unique properties, $\text{H}_2\text{O}$ deserves special consideration. First, as we have
observed, water vapor obeys a shallow scale height on Earth, and its V-R and pure
rotation bands fill much of Earth’s infrared spectrum. More importantly, water vapor
concentrations are strongly constrained by thermodynamics. Gaseous mixing ratios $q$
rarely exceed the saturation mixing ratio with respect to a plane liquid surface $q_s$; the
temperature dependence of the latter is given by the Clausius-Clapyron equation

$$dq_s = \frac{L}{R_e T_e^2}dT$$

where $L$ is the latent heat of fusion and $R_e$ is the gas constant for $\text{H}_2\text{O}$. Integration of
this relation, combined with tropospheric temperature profiles, yields a vertical profile
of $q_s$ that is very nearly exponential with a scale height of 2 to 2.5 km depending on
latitude.

Though $q$ almost never exceeds $q_s$, it can be up to two orders of magnitude lower
than $q_s$ in the middle troposphere and even drier in the stratosphere. The ratio of $q$ to $q_s$
is called relative humidity, here denoted RH. The huge variations of relative humidity
are important to weather, and to climate as well since the lack of a commonly accepted
theory to explain climatological RH statistics creates some uncertainty in our ability
to predict climate variations. However, it is widely expected that the dynamics and
thermodynamics of the atmosphere, while producing stark horizontal contrasts in RH,
should also tend to keep climatological means and variances of RH relatively invariant
as Earth’s climate changes.

The temperature dependence of water vapor has interesting consequences for ra-
diation. Noting the exponential behavior and shallow scale height of $\text{H}_2\text{O}$, we may
for the moment suppose RH is a constant with height of order unity. Since the (nadir)
brightness temperature in an $\text{H}_2\text{O}$ band is determined approximately by $T'(z')$ where
$\tau_p(z') = 1$, and $\tau_p(z)$ is proportional to $\rho(z)$ by virtue of (7.23) and (7.25), it follows
that to leading approximation the observed $T_b$ at a given $\nu$ and for a given RH will
be independent of the atmospheric temperature profile! That is, the radiation will (on
average) originate at that altitude where $T$ has the value required for the water loading
above to add up to unit optical depth. The special value of $T_b$ will vary with frequency,
and will increase if RH $< 1$ since warmer temperatures will have to be reached to attain
the same vapor; the sensitivity of $T_b$ to RH is given approximately by the Clausius-Clapyron
equation again. If curvature effects of $B(T(\tau))$ are taken into account, some
sensitivity of $T_b$ to $T(z)$ does emerge though it is much less than the sensitivity to RH.
7.5.1 Remote sensing of H\textsubscript{2}O

This characteristic enables the remote sensing of relative humidity (but not \( q \), unless \( T(z) \) is known a priori) using passive monochromatic/narrowband emission, as is currently performed by both geostationary and polar-orbiting weather satellites. The shallow scale height enables RH retrievals with some degree of vertical resolution, by choosing wavelengths of different \( k \). However, detailed retrievals of water vapor are made tricky by the influence of clouds on the retrievals and by the large (sometimes sudden) changes in RH with height at specifically sampled locations and times. Use of the microwave portion of the spectrum can reduce the severity of the cloud problem, but only techniques that make use of auxiliary information on vertical structure can succeed in the presence of discontinuities in RH with height.

7.5.2 The water vapor feedback and runaway greenhouse

The bad news side of this story is that emitted flux in water vapor bands is insensitive to Earth’s temperature. If water vapor were the dominant absorber at all wavelengths, this would lead to a runaway greenhouse warming on Earth, since no matter how hot Earth got it would radiate the same amount of energy to space. This is known as Simpson’s paradox. In reality, only about half of the infrared spectrum is strongly affected by H\textsubscript{2}O, with the other half of the spectrum escaping in the so-called “window” regions or blocked only by other greenhouse gases such as CO\textsubscript{2}. Thanks to the window regions, Earth’s climate is stable after all, albeit about twice as sensitive to changes in energy input as it would be with fixed \( q \) at all levels. This increased sensitivity is one of the strongest known positive climate feedbacks, dubbed the water vapor feedback. The strength of this feedback could differ from the doubling of sensitivity currently anticipated, if it turned out that RH itself were sensitive to Earth’s mean climate (a sensitivity currently not shown in global climate models). Also, the water vapor feedback could become stronger and the climate much more unstable if the critical window regions became dirtied by the addition of sufficient quantities of new anthropogenic gases such as CFC’s that, unlike any naturally occurring substances, are good absorbers in this region.

7.6 0-D and 1-D radiative-convective equilibria

We can take the plane-parallel approach to extremes and represent the entire planet as a single, horizontally-uniform column with radiation coming in the top and heating the ground. In Chapter 8, we will consider variations of an atmosphere in space and time, and the role played by radiation in damping those. But the 1-D picture is actually quite useful in obtaining a first-order understanding of the time-averaged vertical structure of Earth’s atmosphere.

An even simpler, zero-dimensional equilibrium model of the climate system is:

\[
\frac{(1 - a)S}{4} = \sigma T_E^4,
\]
where $a$ is the planetary albedo, $S$ is the solar constant, the factor 4 is the ratio of a sphere’s surface area to its extinction cross section, and $T_E$ is called the planetary effective temperature. The effective temperature characterizes the average temperature at optical depths near unity into the atmosphere, and is therefore related to the surface temperature in a way that depends on atmospheric opacity (with no atmosphere, $T_E$ is just the surface $T$.)

If radiation alone acts on the atmosphere and underlying surface, we say that the fluid is in radiative equilibrium. Any fluid with no internal energy transports other than radiation will relax toward an equilibrium where temperatures are just right so that net thermal emission balances solar absorption at each altitude. This equilibrium temperature distribution $T_E(z)$ will be radiatively stable to any temperature perturbation since $\frac{\partial B(T)}{\partial T} > 0$ always. It cannot be calculated accurately using a grey model, but can be calculated fairly well by dividing the infrared spectrum into window and non-window regions (see Problem 5), or better yet using a model with multiple bands for different gases.

Though radiatively stable, the radiative equilibrium distribution $T_E(z)$ is gravitationally unstable through most of Earth’s atmosphere and those of the other inner planets, as well as through the outer part of the sun. This is because $T_E(z)$ is high near the planetary surface (or conversely, is low in the solar photosphere) compared to deeper within the fluid since so much solar radiation is received from (lost to) space. Heated air near the ground is lighter due to the ideal gas law,

$$\rho = \frac{p}{RT}$$

where $\rho$ is density and $R$ is the gas constant, so it rises and causes convection or overturning of the fluid. Fluid will approximately conserve its equivalent potential temperature

$$\theta_e \approx T \left( \frac{p_0}{p} \right) \frac{\partial}{\partial z} \exp \left( \frac{Lq}{c_pT_0} \right)$$

as it rises or sinks. Note that in saturated air at a given pressure, $\theta_e$ increases monotonically with $T$.

The result will be a tendency for convection to adjust $T(z)$ toward a “convectively neutral” profile in which there is no longer any gravitiational instability and any parcel of fluid that is displaced vertically will always attain the same density as other fluid already at that level. This occurs only in the unstable portions of the profile. At any level $z_S$ such that no parcel lifted to $z_S$ from below can be warmer than $T_E(z_S)$, while a parcel lifted from $z_0$ to any higher level will be colder than $T_E(z_0)$, convection will not occur and we have $T(z_S) = T_E(z_S)$. On the inner planets this holds in the stratosphere and above. Convection is confined to the troposphere, which on Earth includes about the lowermost 90% of its atmosphere.

Unfortunately, no firm theory exists to predict the action of convection against the destabilizing influence of radiation. One generally assumes that convection will act as a relaxation, as just described, but even this simplification leaves open the question of how rapid the adjustment will be as a function of altitude. The simplest procedure, widely used in 1-D models, is to assume an infinite relaxation rate at all levels where
adiabatically displaced parcels can mix (conditions above on $z_S$ not satisfied), and a zero rate elsewhere. The problem with that is that observed temperature profiles show that convective relaxation is very rapid compared to radiative relaxation only in the lower part of the convective layer, with the two processes gradually shifting in relative importance toward the stratosphere. Furthermore, due to the inertia of convective eddies, it is possible for convective mixing to extend somewhat outside the boundaries established by the above reasoning. These issues do not go away when we start considering horizontally heterogeneous atmospheres, since convection is a very small scale process that must be parameterized in even the most finely resolved numerical models.

**Exercises**

7.1 Using the hemispheric constant two-stream method, calculate the heating profile in a semi-infinite layer, with isotropic phase function and no emission, heated by diffuse incident radiation, for the cases (a) $\omega=0$, and (b) $\omega=0.5$. Comment on the net effect of scattering on the heating profile.

7.2 Compute the integrated net heating of Earth’s troposphere with overhead sun, using the single-scattering and grey approximations with a clear-sky optical depth of 0.1, under conditions of (a) clear skies over oceans, (b) clear skies over snow, and (c) cloudy skies over oceans. Assume you have the following observations: greybody $R_g = 0.1$ for ocean and $R_g = 0.8$ for snow; clear-sky scene albedo $\mathcal{R} = 0.1$ and cloudy scene albedo $\mathcal{R} = 0.5$. Assume the clouds lie at the top of the troposphere. (hints: i. you must know the backscattering coefficient, ii. the measured albedos include all reflections).

7.3 Calculate as a function of relative humidity $\text{RH}$ the $T_b$ observed by looking down into an atmosphere near the 6-7 micron $\text{H}_2\text{O}$ band, given $T_b(T_\text{sat})$ for a saturated atmosphere. Assume RH is vertically uniform, the Planck function $B(z)$ is linear with height, and that $T_b(T_\text{sat})$.

7.4 Calculate the radiative equilibrium temperature profile of an atmosphere and its underlying surface, if the atmosphere is grey in the emission spectrum and optically thin at incoming solar wavelengths, and the surface is black. Neglect scattering. Represent the atmosphere as $N$ layers (numbered from $i = 1$ for the top layer) each having thermal emissivity $\varepsilon$ ($\varepsilon = 1 - T$), each absorbing an equal fraction $A/N$ of net incoming solar flux $S_0 = \sigma T_\text{E}^4$, and each experiencing zero net energy gain. Designate the upward and downward thermal fluxes at the top and bottom of layer $i$ as $F_\text{up}^i$ and $F_\text{down}^i$, respectively, and the layer temperatures as $T_i$. When $A = 0$, your result should reduce to:

$$T_G = T_E \sqrt[4]{2 + (N - 1)\varepsilon \over 2 - \varepsilon},$$

$$T_i = T_E \sqrt[4]{1 + (i - 1)\varepsilon \over 2 - \varepsilon}.$$

7.5 Consider a three-layer representation of Earth’s atmosphere, where layer 1 represents the stratosphere and layers 2-3 are the upper and lower troposphere. Suppose the
EXERCISES

optical properties and fluxes can be treated separately in three spectral bands: solar, IR window (8-12 \( \mu \text{m} \)), and IR non-window (all other terrestrial emission). Assume that the fraction \( f \) of thermal radiation falling outside the window is fixed, estimating \( f \) based on a thermal spectrum at \( T = 300 \text{K} \). Suppose the optical depths are:

<table>
<thead>
<tr>
<th>Solar</th>
<th>Window</th>
<th>IR non-window</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.03</td>
<td>0.1</td>
</tr>
<tr>
<td>2</td>
<td>0.10</td>
<td>1.0</td>
</tr>
<tr>
<td>3</td>
<td>0.10</td>
<td>1.0</td>
</tr>
</tbody>
</table>

(The opacities in the IR outside the window come \( \text{CO}_2 \) and \( \text{H}_2\text{O} \), while the solar and window opacities are due to ozone in the stratosphere and water vapor in the troposphere.) For solar radiation, make the approximation that absorbed energy equals \( \tau \sigma T_E^4 \) in each layer.

1. Obtain a linear system of four equations \( \mathbf{AT} = \mathbf{b} \mathbf{T}_E^4 \), where \( \mathbf{A} \) is a 4 by 4 matrix of constant coefficients, \( \mathbf{T} \) is the vector of temperatures, and \( \mathbf{b} \) is a constant vector.

2. Evaluate the solution (hopefully using a computer) for the values given above, and setting an observed value for \( T_E^4 \).

3. Reevaluate the solution for an ozone-depleted atmosphere (50\% less ozone).

4. Reevaluate the solution for an atmosphere with CFC’s boosting the window optical depth by 0.01 in each layer.

5. Reevaluate the solution for non-window IR \( \tau \) increased by 20\% due to \( \text{CO}_2 \) increase.

6. Reevaluate the solution with volcanic aerosol boosting all stratospheric optical depths by 0.02 and decreasing \( T_E^4 \) by 1\%.

Comment on any results that seem nonintuitive.
Chapter 8

Radiation and atmosphere-ocean dynamics on Earth

This chapter is currently incomplete, but here is some material on radiative damping.

8.1 Role of radiation in planetary dynamics

Radiative transports through Earth’s atmosphere are of critical importance since they ultimately drive the circulation of the atmosphere and oceans. In an overall sense, the strong heating of low latitudes requires the existence of net poleward heat transports in the two fluids. Furthermore, the net cooling of the atmosphere requires heating by fluxes from the surface. In addition to these climatological roles, radiation can be important as a damping agent on some large-scale variations in the climate system since net radiative transports are sensitive to temperature.

We may understand the latitudinal distribution of Earth’s atmosphere and surface temperatures by an argument analogous to that used in the previous chapter for the vertical structure. In the absence of fluid energy transport, there is a radiative equilibrium temperature distribution $T_E(x)$, where $x$ is a position vector. Though we think of $T_E$ as depending primarily on latitude and height, it can also depend on longitude as long as there are continents that cause variations in albedo and orography. Earth’s atmosphere is in constant motion because of the instability of the $T(x) = T_E(x)$ state. The state is not unstable to changes in $T$ but is, in several ways, unstable to changes in wind $u$.

The first source of instability, convective instability, was discussed in Chapter 7 and leads to overturning eddies of horizontal length scale comparable to the thickness of the troposphere (with possible organization to larger scales by a variety of nonlinear mechanisms). This type of instability by itself does not, however, lead to significant horizontal winds.
8.2. RADIATIVE DAMPING OF FLUCTUATIONS

The distribution $T_E(x)$ also possesses a strong equator-to-pole gradient due to strong solar heating of the tropics and negligible heating of the poles. Variations in thermal expansion produce horizontal pressure gradients that tend to drive flow toward the equator at low levels and away at high levels. While the Earth’s rotation strongly modifies this flow, the final flow field that results must, for sufficiently low atmospheric viscosity, include a so-called Hadley cell at low latitudes. This cell modifies $T$ away from $T_E$ at these latitudes by very efficiently transporting energy, and in fact makes $T$ nearly constant with latitude out to about 23° on Earth. Such “axisymmetric” (longitudinally-averaged) circulations occur on all the solar planets with atmospheres, but with varying character and extent.

Another important instability is the baroclinic instability of the midlatitude jet stream which occurs poleward of Earth’s Hadley Cell. This instability leads to increased poleward transport of heat in longitudinally asymmetric midlatitude cyclones. It is also believed to occur in the Martian and Jovian atmospheres. This instability is fundamentally due to the excessive latitudinal gradient in $T_E(x)$.

Motions associated with each of these instabilities produce net enthalpy fluxes from regions of radiative heat buildup ($Q > 0$ near the surface and at low latitudes) toward regions of net loss (aloft and high latitudes), such that the atmosphere functions as a heat engine. Locally, there are also energy fluxes into and out of the surface which may reinforce or reduce the atmospheric radiative flux convergence. Underneath solid surfaces, diffusive horizontal heat transport will be negligible and long-term (e.g. more than a year) averages of the surface net vertical energy flux must approach zero. Oceans however can redistribute heat horizontally; on Earth, poleward transports occur that are qualitatively similar to, while quantitatively smaller than, those in the atmosphere. The net effect of all energy sources is to produce a new mean state in which net energy sources add up to zero on average at each location. This state may be regarded (i.e., modeled) as a radiative-dynamical equilibrium $T_0$.

Steady atmospheric circulations cannot be properly understood without acknowledging the role of radiation, since the circulations are driven (or at least energetically balanced) by radiation. However, dynamical analyses of geophysical flows often downplay the role of radiative heating while emphasizing instead the roles of surface energy fluxes or the conversion of latent heat in the atmosphere. The following sections discuss various aspects of the radiative interaction with circulations.

8.2 Radiative damping of fluctuations

The previous discussion was focused on gaining an understanding for what determines heating rates. Here, we consider a problem of practical interest within the atmospheric sciences: the radiative damping of large-scale disturbances involving deviations from an equilibrium temperature structure $T_E$.

It was noted earlier that radiative heating anomalies will work against such temperature deviations. We may define the radiative damping time constant $\tau_R$ by

$$\frac{1}{\tau_R} \equiv -\frac{\partial Q}{\partial T}$$
where $\tau_R$, $Q$ and $T$ are associated with a specific small perturbation of the system away from $T_0$. If $\tau_R$ is a constant, independent of position or type of perturbation in the system, the radiative damping is known as Newtonian cooling. The Newtonian cooling assumption is ubiquitous in the meteorological literature and will merit some attention. Its effect is to produce linear damping of all thermal disturbances at a constant rate without changing the structure of any disturbance, an important simplification.

### 8.2.1 Optically thin medium

We now consider the heating rate at the location $\tau_\nu = 0$ within a nonscattering medium. In optically thin cases, each flux in (7.11) differs little from what comes in at the respective upstream boundaries. In the atmosphere as a whole the upward and downward fluxes would approximately equal the flux $\pi B_\nu(T_g)$ emitted from the surface and zero respectively, yielding

$$Q_\nu \approx \frac{\pi k_\nu d}{c_p} \left[ B_\nu(T_g) - 2B_\nu(T) \right] \quad \text{(IR).} \quad (8.1)$$

Thus the longwave heating would become zero at a temperature equal to $2^{-1/4} T_g$. For small departures from this temperature, we would have

$$\frac{\partial Q_\nu}{\partial T} \approx -\frac{2\pi k_\nu d}{c_p} \frac{\partial B_\nu}{\partial T}. \quad (8.2)$$

This sensitivity applies to the whole spectrum if $k_\nu$ and the solar heating can be assumed independent of $T$. Equation (8.2) indicates that radiative heating of a thin layer will relax exponentially toward an equilibrium temperature, at a constant Newtonian cooling rate. This rate is determined by the properties of the gas and of the Planck function.

The net radiative cooling of the atmosphere is often called “cooling to space,” a term we applied earlier to the cooling by emission that occurs within a few photon mean free paths of the upper boundary of a deep medium. Examination of (8.1) reveals, however, that the equilibrium temperature reached by an optically thin layer is determined equally by the upper and lower boundary conditions; fluxes into the ground adjust to layer temperature as readily as do fluxes to space. But it is also true that subsequent warming of the surface might in some situations reduce the effective sensitivity of the net flux at the surface to $T$. Over a surface with zero effective heat capacity, the “cooling to space” terminology becomes appropriate for describing the aggregate behavior.

### 8.2.2 Optically thick medium

Does the Newtonian cooling approximation hold for optically thick cases as it did for the thin case? It is easy to see that the general answer is no. Consider a layer of $\tau_\nu \gg 1$. Deep in the interior of such a layer, $F^\uparrow_\nu$ and $F^\downarrow_\nu$ come almost entirely from emission from the layer itself. If the whole layer warms uniformly then both fluxes will
8.2. RADIATIVE DAMPING OF FLUCTUATIONS

increase as the same rate as $B_{x}(T)$, rendering the net heating in (7.11) independent of temperature. The relaxation rate approaches zero as $\tau_{\nu} \rightarrow \infty$.

Consider the average heating through the layer when perturbed uniformly by a small amount $dT$. We assume the layer has optical thickness $\tau$, geometrical thickness $H$, is bounded by surfaces of unit emissivity and fixed temperature $T_0$, and use the diffusivity approximation. Such a perturbation may also be thought as a top-hat shaped perturbation within a thick medium of greater extent. In either case, (7.3) may be integrated through the perturbed layer, and (6.7) used to obtain the outward transmitted fluxes, to obtain the mean heating rate

$$\langle Q \rangle_{\text{layer}} = \frac{2\pi\epsilon}{\rho H \epsilon_{p}} [B(T) - B(T + dT)],$$

(8.3)

where the greybody emissivity $\epsilon$ is given by (6.7). In the optically thin limit $\epsilon \propto \tau$ and we recover (8.2). At the thick limit, however, $\epsilon \approx 1$ so that the relaxation rate loses its dependence on gas properties and becomes inversely proportional to layer heat capacity $\rho H \epsilon_{p}$ instead. Thus, the linearity of the damping rate will depend on the scale of the perturbation. Note also that perturbations are most rapidly damped in the optically thin limit.

**Form of the perturbation**

The astute reader will also realize that the actual heating rate must vary greatly within the layer considered above, maximizing at the layer boundaries and minimizing in the interior. If the layer optical thickness is much greater than one, the diffusion approximation (7.17) holds so that radiative energy diffuses slowly into the boundaries of the layer. The form of the initial perturbation will obviously not be preserved as it decays; in mathematical language, the uniform perturbation is not an eigenfunction of the system. Naturally arising fluctuations in which radiation plays a key role will be eigenfunctions of the combined system equations (including dynamical and radiative) and should be calculated from them. Note that sine and cosine functions are eigenfunctions of the diffusion operator, so these perturbations will decay without changing form in a very optically thick medium. The damping rate of a sinusoidal disturbance falls as $L^{-2}$ ($L$ the wavelength) in contrast to the $H^{-1}$ dependence noted for a uniformly heated layer. The difference arises because the temperature gradients are in the sinusoidal case are proportional to $L^{-1}$, while they are scale-independent in the uniform perturbation.

In many cases, the form of temperature perturbations is dictated by factors that can be considered external and is not expected to be greatly affected by radiation. One example is upward propagating waves. If the waves are launched in the troposphere and when reaching the upper atmosphere are weakly damped by radiation, this damping can be calculated by considering perturbations of the form dictated by the wave equations whether or not they are eigenfunctions of the combined radiative-dynamical system. The damping time of such waves propagating into Earth’s stratosphere is typically a few weeks or less. As long as this damping time is longer than the natural period of the (undamped) wave, radiation is not expected to substantially alter the form of the wave.
Another interesting case is planetary boundary layers. To the extent that an adiabatic vertical temperature profile is maintained by vertical mixing from the surface, perturbations within such layers will also be uniform in potential temperature; if radiative damping is nonuniform, the heat delivered by convection will simply readjust vertically to compensate. The condition for this is that vertical variations in a disturbance be damped by convective heat transports much more rapidly than the radiative damping rate. The Earth typically has a boundary layer a couple of km thick which has this character; in the tropics and summer midlatitudes, it may be appropriate to regard the entire troposphere as such a layer, as long as convection occurs throughout the domain of interest and small details of the vertical temperature structure are not important. In these cases, the uniform layer analysis presented earlier is valid.

If and when convective or other processes can be relied upon to enforce vertical uniformity upon (potential) temperature perturbations, the plane-parallel approximation produces automatically the Newtonian cooling approximation. If more than one vertical perturbation structure is possible (i.e. the first and second harmonics in the vertical direction which characterize the convective and stratiform heating of convective systems), then each will be damped at a different rate and Newtonian cooling cannot be applied locally.

\section*{Exercises}

8.1 Assuming the troposphere is nonabsorbing inside a “window” region containing a fraction $f$ of the Planck emission, but optically thick outside the window, and treating the troposphere as an isothermal layer of temperature $T^*$, calculate the relaxation time $\tau_R$ of vertically uniform temperature perturbations with lower boundary conditions of: (a) fixed surface temperature, (b) surface temperature that tracks air temperature. Evaluate the expressions for $T^* = 255$K and $f = 0.5$. 

Bibliography

Appendix A

Quantum mechanics

For readers who haven’t studied quantum mechanics—or for whom memories of it are not crystal clear—a brief review is provided here of basic concepts invoked in the main text. These concepts are not really essential to understand, but should help make the behavior described in Chapter 5 (e.g., selection rules) more intuitive, especially for readers who are somewhat familiar with the concepts already. For further information, readers should consult a quantum physics textbook; I have found French and Taylor’s book to be a very readable starting point, and Gasiorowicz’s book to be a useful (though terse) overview of more advanced topics.

Quantum mechanics is concerned with the behavior of (DeBroglie) waves describing matter. In an atom, for example, the electron may be thought of as a wave; the nucleus, another wave; and so on, with the whole system being described by the product of the wavefunctions $\Psi$ of all the constituent parts. The wavelength is related to the momentum of the particle and, for a particle of given mass, decreases with increasing particle energy. A particle traveling in free space can have any momentum, and therefore any DeBroglie wavelength. The wavefunction behaves according to the Schroedinger equation, a partial differential wave equation quite similar to that for describing a sound wave (the quantum wavefunction is a complex scalar). The physical interpretation of this wave first given by Born is now widely accepted: the probability of finding a particle in a particular location, if one goes to the trouble of exactly locating it by means of an observation, is proportional to $|\Psi|^2 = \Psi^*\Psi$ at that location (where $*$ denotes complex conjugate). Otherwise, the behavior of the physical system behaves as the average of the ways that would have been expected, averaged over the potential positions, without you ever knowing the “true” position (whether there is any true position is still a matter of debate). As the wavefunction evolves or propagates and its (complex) amplitude grows in certain places and shrinks in others, the particle has (probabilistically) moved. We think of the particle’s (potential) position as being spread out according to $\Psi$.

Quantized behavior of importance to us arises because this wave must obey boundary conditions. An everyday analogy is that although a steel wire or string can physically support any frequency of oscillation, if it is strung onto a guitar—held fixed at either end—a wave must “fit” into its length exactly or it will rapidly destroy itself as it
reflects incoherently, over and over, from the fixed ends of the string (the time it takes a wave to move from one end to the other is very short compared to the duration of a played note). This boundary condition establishes a set of possible wavelengths, and therefore frequencies, for a sustained note audible after initial transients have died out (which takes only a small fraction of a second). The longest wave ($\lambda = 2l$) that fits is called the fundamental, but others ($n\lambda = 2l$, $n > 1$), called harmonics, will also fit. These patterns that fit are called resonant modes. How strongly each mode occurs when the guitar is played depends on where and how rapidly the string is plucked.

Similarly, the DeBroglie wave of a trapped particle must “fit” into the containing vessel (which is described more accurately as a potential well) except for transient fluctuations. A steady state (mode) in which this “fit” is achieved is called a stationary state and must be a solution of the simpler time-independent version of the Schroedinger equation (obtained by setting $\partial \Psi / \partial t = 0$ in the full equation) which is most simply written

$$H \Psi_n = E_n \Psi_n.$$  \hspace{1cm} (A.1)

The Hamiltonian function $H$ describes the “vessel” (boundary conditions and/or force potentials) and the eigenvalue $E_n$ equals the energy of the mode $n$, whose wavefunction is $\Psi_n$.

When a state transition from, say, $n = 1$ to $n = 2$ occurs, the system must pass through a superposition state whose wavefunction is

$$\Psi = A \Psi_1 + B \Psi_2.$$  \hspace{1cm} (A.2)

The coefficients $A$ and $B$ are 1 and 0 respectively before the transition, then change eventually to 0 and 1 by the time the transition is completed. The key to understanding whether a transition is photometrically allowed or forbidden is to examine the superposition state. If, in the superposition state, positive and negative charge are separated into a dipole, then it will be possible for a photon to drive the system from 1 to 2 (and, by extension, a spontaneous transition from 2 to 1 will radiate a photon). The charge separation during the superposition state can be thought of as a temporary antenna.

Since an electron and its nucleus always represent a dipole, the above requirement does not restrict electronic transitions. Classically, an electron orbiting a positive charge would always be expected to radiate at the orbital frequency. However, the requirement poses restrictions on vibrational and rotational transitions. Linear, asymmetric molecules like CO possess a transient dipole moment during extrema of vibration, but not on average. As discussed in the main text, other molecules may never develop such moments, may develop them only for certain vibration modes, or may have them all the time.

Here is presented an idealized example. The wavefunction describing the orientation of a molecule must obey a periodic boundary condition in polar coordinates (its value and all derivatives must repeat every $2\pi$ radians). The resulting modes obey

$$\Psi \propto e^{iJ\phi},$$

where $\phi$ is the molecular angle and $J$ is the quantum number. Note that in any such state, the probability $\Psi \Psi^*$ is uniform on $\phi$, so there is no orientation preference. For
Figure A.1: The angular probability distribution for superposition of two stationary states under polar symmetry, labelled by the difference in quantum number, for equal mixture of states \((A=B)\).

the superposition state between \(J_1\) and \(J_2\), this is not true. We can combine the above equations to get a probability distribution

\[ \Psi(t)\Phi^*(t) \propto AA^* + BB^* + 2AB|e^{i(J_1 - J_2)\phi}. \]

The relevant dipole moment will be proportional to the product of the instantaneous dipole moment of the molecule (if any) and the first moment of the above distribution. This is nonzero only for \(|J_1 - J_2|=1\) (Fig. A.1), hence the selection rule for rotational transitions.

A mathematically identical argument yields the \(\Delta v = \pm 1\) rule for vibrational transitions. In this case we consider solutions with a quadratic Hamiltonian. Though this yields modes of nonuniform amplitude in space, they still possess the same harmonic-series character. Superpositions of two immediate neighbor states will produce flexure with a directional preference (e.g. stretched as opposed to compressed) while others will not. Thus, if we have \(\Delta v = \pm 1\), then a dipole moment will persist over the duration of the transition. If, further, we have \(\Delta J = \pm 1\), then this dipole moment will tend to orient in a persistent direction as well, allowing the photon interaction.