

Contents

1	Hydrogenic Atoms	<i>page</i> 1
1.1	The Bohr Atom	2
1.1.1	Energy Structure and Transitions	4
1.2	The Schrödinger Atom	8
1.2.1	The Schrödinger Equation	8
1.2.2	Properties of the Wave Function	9
1.2.3	The Schrödinger Wave Function	10
1.2.4	Energy Structure and Transitions	13
1.3	The Dirac Atom	15
1.3.1	Spin	15
1.3.2	Spin-Orbit Coupling	16
1.3.3	Energy Structure and Transitions	19
1.4	Radiative Corrections	23
1.5	Isotope Shifts	25
1.6	Continuum States	26
1.6.1	Energy Structure and Transitions	27
2	Transitions of Hydrogenic Atoms	30
2.1	Transition Probabilities	31
2.1.1	The Dipole Approximation	31
2.1.2	Einstein Coefficients	32
2.2	Schrödinger Oscillator Strengths	34
2.2.1	Averaged Oscillator Strengths	35
2.2.2	The Radial Overlap Integrals	36
2.2.3	Lyman and Balmer Series Oscillator Strengths	37
2.3	Fine-Structure Oscillator Strengths	38
2.3.1	Oscillator Strengths for Multiplets	39
2.3.2	The Spectrum	41

2.4	Bound-Free Oscillator Strengths	43
2.5	Natural Broadening	45
2.6	Emission Power and Absorption Cross Section	51
2.6.1	Bound-Bound Emission	51
2.6.2	Bound-Bound Absorption	52
2.6.3	Bound-Free Absorption	54
3	Multi-Electron Atoms and Transitions	58
3.1	The Many-Electron Problem	59
3.1.1	The Hartree-Fock Method	60
3.1.2	Coupling Schemes	61
3.1.3	The Russell-Saunders Vector Model	62
3.1.4	Russell-Saunders Term and State Symbols	66
3.2	Energy Structure	67
3.2.1	Central Field Approximation	68
3.2.2	L–S Coupling	72
3.2.3	Spin-Orbit Coupling: Fine Structure	74
3.2.4	j – j and Intermediate Coupling Schemes	76
3.3	Bound-Bound Transitions	79
3.3.1	Oscillator Strengths and Line Strengths	80
3.3.2	Emission Power and Absorption Cross Section	83
3.3.3	Selection Rules	84
3.4	Isotope Shifts	87
4	Atomic Properties and Spectra	89
4.1	Ground-State Russell-Saunders Symbols	89
4.2	The Periodic Table	92
4.2.1	Iso-Sequence Ions	93
4.3	Binding Energy and Ionization Potential	96
4.4	Grotrian Diagrams and Spectra	99
4.4.1	Group IA: Alkali Metals	100
4.4.2	Group IIA: Alkali-Earth Metals (and Helium)	105
4.4.3	Group IIIA: Boron Group	111
4.4.4	Iron Family Ions	116
4.5	Common Absorption Lines	123
5	Cosmological Paradigm	129
5.1	Geometry and the Metric	130
5.1.1	Time and Distance	131
5.1.2	Hubble’s Law	133
5.2	The Robertson-Walker Metric	134
5.3	Relativistic Spacetime Dynamics	137

5.4	Parameterized Cosmology	140
5.4.1	Dynamics	140
5.4.2	Friedmann-Einstein Cosmologies	142
5.4.3	The Robertson-Walker Metric	144
5.4.4	The Development Angle	145
5.5	Λ CDM, H_0 , and the Ω s	146
5.5.1	Evolution of the Ω s	149
5.6	The “737 Cosmology”	152
6	Applied Cosmology	154
6.1	Redshift	155
6.2	Observer Time	157
6.2.1	Time Dilation	159
6.3	Expansion Dynamics	160
6.4	Distances, Separations, and Angles in Cosmology	162
6.4.1	Co-Moving and Proper Radial Distances	163
6.4.2	Radial Separations	166
6.4.3	Transverse Separations	167
6.4.4	Angular Diameter “Distance”	169
6.4.5	Angular Separations	171
6.4.6	Proper Separations in Lenses	172
6.4.7	Source Beam Cross Section	176
6.4.8	Luminosity Distance	177
6.4.9	Surface Brightness	179
6.4.10	Absorption “Distance”	180
6.5	Cosmological Volume	183
6.6	Peculiar and Recessional Velocities	185
6.6.1	Peculiar Velocities	186
6.6.2	Recessional Velocities	189
7	Radiative Transfer	192
7.1	The Radiation Field	192
7.1.1	The Geometry	193
7.1.2	The Specific Intensity	195
7.1.3	The mean Intensity	197
7.1.4	The Photon Field	197
7.2	The Flux Vector	198
7.2.1	Defining the Observed Flux	200
7.2.2	Measuring the Observed Flux	204
7.3	Radiative Transfer	205
7.3.1	The Extinction Coefficient and Optical Depth	208

7.3.2	The Emission Coefficient	210
7.3.3	The Transfer Equation and Source Function	210
7.3.4	Solution in One Dimension	212
7.4	Absorption Lines and Optical Depth	215
7.4.1	Transforming Energy Densities	215
7.4.2	Pure Absorption	215
7.4.3	The Astronomical Absorption Spectrum	218
7.4.4	Cosmological Intervening Absorption	219
7.4.5	Associated Absorption and Partial Covering	221
7.4.6	Column Density	224
8	Spectrographs and Spectra	227
8.1	The Spectrograph	227
8.2	The Seeing Disk	228
8.3	Slits	229
8.4	Diffraction Gratings	231
8.4.1	The Grating Equation	232
8.4.2	Free Spectral Range	233
8.4.3	The Interference Function	234
8.4.4	The Blaze Function	235
8.4.5	The Intensity Function	237
8.4.6	Blazing a Grating	237
8.5	Low-Order Spectrographs	240
8.6	Echelle Spectrographs	241
8.7	Resolving Power	244
8.7.1	Diffraction Limited Resolution	245
8.7.2	Seeing Limited Resolution	246
8.8	Instrumental Spread Function	248
8.8.1	Convolution of the ISF	250
8.9	Pixelization	252
8.9.1	Pixel Plate Scale	252
8.9.2	Dispersion per Pixel	253
8.10	Spectroscopic Data	255
8.10.1	Atmospheric Attenuation and Throughput	257
8.10.2	The Recorded Spectrum: Flux to Counts	259
8.10.3	Uncertainties in Recorded Counts	260
8.10.4	Flux Calibration	262
9	Absorption Lines and Ionization Breaks	265
9.1	The Challenge of Spectral Analysis	266
9.2	Absorption Lines	270

9.2.1	The Total Absorption Cross Section	270
9.2.2	The Thermal Broadening Function	271
9.2.3	The Voigt Profile	274
9.2.4	Equivalent Width and the Curve of Growth	278
9.2.5	Curve of Growth Methods	283
9.2.6	Additional Physics and the Curve of Growth	288
9.2.7	Apparent Optical Depth Method	290
9.2.8	Partial Covering from Doublets	292
9.2.9	Observed Redshifted Equivalent Widths	293
9.3	Ionization Breaks	295
10	Spectral Analysis	299
10.1	Continuum Fitting	300
10.2	Combining spectra	304
10.3	Objectively Measuring Absorption Lines	308
10.3.1	Aperture Method	309
10.3.2	An Optimized Method	311
10.3.3	The Line List	313
10.4	Absorption Line Systems	315
10.4.1	Systemic Redshift	323
10.4.2	Equivalent Widths and Mean Wavelengths	324
10.5	Kinematics	327
10.5.1	Rest-Frame Velocities	327
10.5.2	Velocity Moments	327
10.5.3	Fractional Velocity Width	331
10.5.4	Two-Point Velocity Correlation Function	332
10.6	The Apparent Optical Depth Method	334
10.6.1	Apparent Optical Depth Profiles	334
10.6.2	Apparent Optical Depth Column Densities	337
10.7	Deblending Absorption Features	340
10.7.1	Gaussian Decomposition	340
10.7.2	Voigt Profile Decomposition	344
11	Gas Physics and Processes	353
11.1	The Radiation Field	354
11.1.1	In the Intergalactic Medium	354
11.1.2	In the Circumgalactic Medium	356
11.2	The Particle Field	359
11.2.1	Particle and Mass Density Conservation	359
11.2.2	Abundances and Mass Fractions	360
11.2.3	Charge Conservation and Ionization Balance	363

11.2.4	Thermodynamic Equilibrium	364
11.2.5	Thermalized Velocity Distributions	366
11.2.6	Kinetic Energy	369
11.2.7	Pressure	369
11.2.8	The Equation of State	372
11.3	Detailed Balancing	373
11.4	Excitation	376
11.4.1	Excitation Rates and Thermodynamic Equilibrium	377
11.4.2	Excited States and Electron Density	379
11.5	Ionization	381
11.5.1	Photo and Auger Ionization	382
11.5.2	Direct Collisional Ionization	384
11.5.3	Excitation Auto-Ionization	386
11.5.4	Radiative Recombination	387
11.5.5	Dielectronic Recombination	388
11.5.6	Charge Exchange	389
11.5.7	Astrophysical Rates and Rate Coefficients	390
11.6	The Cooling Function	393
11.6.1	Photoionization Heating	396
11.6.2	Collisional Ionization Cooling	398
11.6.3	Recombination Cooling	399
11.6.4	Collisional Excitation Cooling	400
11.6.5	Continuum Cooling	402
11.6.6	Compton Heating and Cooling	403
11.6.7	The Net Cooling Curve	405
12	Ionization Modeling	407
12.1	Rate Equations	408
12.1.1	Hydrogen	408
12.1.2	Helium	409
12.1.3	Metals	410
12.2	Equilibrium Solution	411
12.2.1	Ionization Equilibrium	412
12.2.2	Ionization in Thermodynamic Equilibrium	414
12.3	Applying Ionization Models to Data	415
12.3.1	Cloudy	416
12.3.2	The Ionization Parameter	419
12.3.3	Ionization Structure	420
12.3.4	Ionization Grids	423
12.3.5	Properties of Ionization Grids	426

<i>Contents</i>	7
12.3.6 Constraining Metallicity and Abundances	429
12.3.7 Estimating Cloud Sizes and Masses	433
12.3.8 Caveats and Complications of Ionization Modeling	436
<i>References</i>	438

1

Hydrogenic Atoms

One of the many goals of the astronomical spectroscopist is to interpret the details of spectral features in order to deduce the properties of the radiating or absorbing matter. It is essential to understand the internal working nature of the atom in order to deduce the physical conditions of the matter giving rise to observed spectral features.

The simplest atom is hydrogen, since it comprises a single proton and a single electron. In general, hydrogenic atoms (those with a single electron) serve to clearly illustrate atomic structure while avoiding the complication of multiple interacting electrons. In addition to neutral hydrogen, H^0 , hydrogenic atoms include the ions He^+ , Li^{+2} , Be^{+3} , etc.

We begin with the semi-classic Bohr model. We then briefly address the wave nature of matter and some basic quantum mechanics, including wave functions. The Schrödinger equation and resulting stationary states of hydrogen are presented. We then incorporate electron spin, spin-orbit coupling, and relativistic energies, all of which lead to the Dirac model and fine structure. We complete our discussion of bound states with radiative corrections (vacuum polarization) and isotope shifts. Finally, we present a brief formalism and wave function for unbound states.

Three excellent additional recourses include the books *Quantum Mechanics of One and Two Electron Atoms* by Bethe & Salpeter (1957), *Physics of Atoms and Molecules* by Bransden & Joachian (2003), *Atomic Astrophysics and Spectroscopy* by Pradhan & Nahar (2011), and *Astronomical Spectroscopy* by Tennyson (2011).

1.1 The Bohr Atom

The hydrogenic atomic model proposed by Neils Bohr (1913) is a semi-quantum mechanical approach to what is essentially classical physics. The Bohr model is based upon two postulates:

1. The electron moves in circular planer orbits about the nucleus such that the total angular momentum of the atom (electron plus nucleus) must be an integer number of Planck's constant, h .
2. The orbit of an electron is a stationary state; an atom emits or absorbs electromagnetic radiation only when its electron changes states; the energy of the photon equals the energy difference of the initial and final states.

Bohr's first postulate is that the total angular momentum is a constant of motion. It is expressed

$$\oint L d\phi = \oint (m_N \omega r_N^2 + m_e \omega r_e^2) d\phi = nh, \quad (1.1)$$

where L is the total angular momentum of the atom, which is purely azimuthal (the orbit is confined to a plane). The geometric configuration is illustrated in Figure 1.1(a), where $r_n = r_N + r_e$ for a nucleus of mass m_N and electron of mass m_e orbiting about a common center of mass with angular frequency $\omega = d\phi/dt$.

Since the orbits are postulated to be circular, ω , r_e , and r_N are constants. Applying $m_N r_N = m_e r_e$, we have

$$L = \hat{m} \omega r_n^2 = n\hbar, \quad (1.2)$$

where $\hbar = h/2\pi$, and where $\hat{m} = m_e/(1 + m_e/m_N)$, is the reduced mass of the electron. Assuming a Coulomb potential, $V(r) = -Ze^2/r_n$, where Ze is the charge of the nucleus, the total energy of the atom in state n is

$$E_n = -\frac{Ze^2}{2r_n}, \quad r_n = \frac{\hbar^2}{\hat{m}e^2} \frac{n^2}{Z} = a_0 \frac{n^2}{Z} \left(\frac{m_e}{\hat{m}} \right), \quad (1.3)$$

where

$$a_0 = \frac{\hbar^2}{m_e e^2} = 5.29177 \times 10^{-9} \text{ cm} = 0.529177 \text{ \AA}, \quad (1.4)$$

is the Bohr radius for an infinite mass nucleus, and where the ratio m_e/\hat{m} explicitly accounts for the reduced electron mass. For hydrogen $\hat{m}/m_e = 0.99946$. For helium $\hat{m}/m_e = 0.99986$. We also can define the reduced Bohr radius $\hat{a}_0 = (m_e/\hat{m})a_0$.

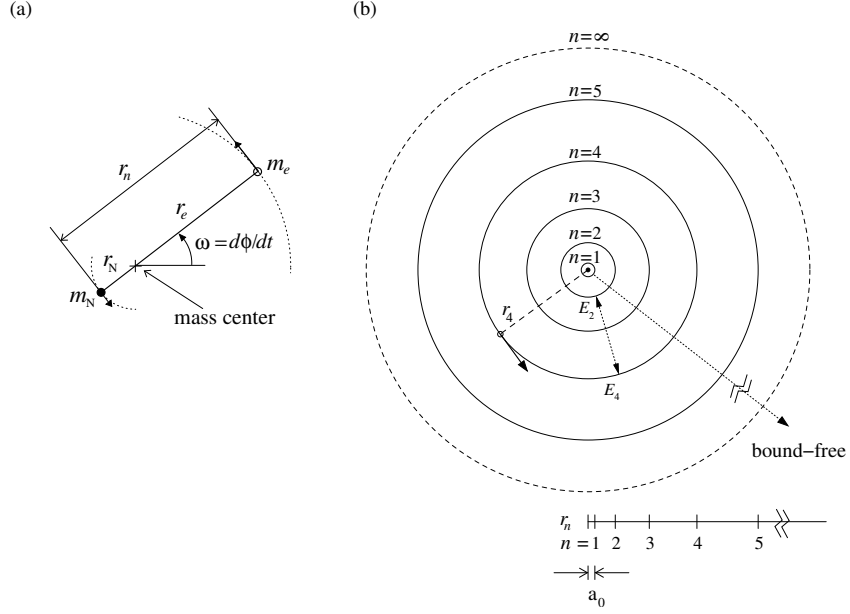


Figure 1.1 (a) The geometric configuration of the Bohr model in arbitrary state n illustrating the relationships between r_n , r_N , r_e , and the center of mass (offset from the nucleus is exaggerated). (b) A schematic of the Bohr model with the relative sizes of the first five orbits in units of the Bohr radius, a_0 . The limit $n = \infty$ is not to scale. When electrons move from one bound orbit to another, photons with energies given by the energy difference of the orbits are absorbed (upward transition) or emitted (downward transition); a bound-bound transition between $n = 4$ and $n = 2$ is shown, for which the photon has energy $E_\gamma = E_4 - E_2$. Ionization from $n = 1$ is also shown as a bound-free transition.

Substituting r_n into E_n (Eq. 1.3), we obtain the binding energy for orbit n ,

$$E_n = -\frac{\hat{m}e^4 Z^2}{2\hbar^2 n^2} = -\frac{e^2 Z^2}{2a_0 n^2} \frac{\hat{m}}{m_e} = -\frac{m_e c^2}{2} \frac{(Z\alpha)^2}{n^2} \left(\frac{\hat{m}}{m_e} \right), \quad (1.5)$$

where the last form provides the energy in terms of the rest energy of the electron, $m_e c^2$, and the fine structure constant $\alpha = e^2/\hbar c$. We define the Rydberg constant,

$$R = \frac{m_e e^4}{2\hbar^2} = \frac{e^2}{2a_0} = \frac{m_e c^2}{2} \alpha^2 = 2.17987 \times 10^{-11} \text{ erg} = 13.60570 \text{ eV}, \quad (1.6)$$

for an infinite mass nucleus, i.e., $\hat{m} = m_e$. In general, for a hydrogenic atom with nuclear charge Ze ,

$$R_z = Z^2 \frac{\hat{m}}{m_e} R. \quad (1.7)$$

We can thus rewrite Eq. 1.5 in the simplified form

$$E_n = -\frac{R_z}{n^2}. \quad (1.8)$$

For hydrogen, $Z = 1$ and $\hat{m}/m_e = 0.99946$, we obtain $R_H = 13.59843$ eV.

1.1.1 Energy Structure and Transitions

In general, atomic energy structure is elucidated using various quantities, including the transition energy, the excitation energy and potential, the ionization energy, and the ground state ionization potential. Here, we briefly define each of these terms.

Transition Energy

According to Bohr's second postulate, the electron is induced to move from a lower bound orbit to a higher orbit via the absorption of a photon. Alternatively, if the electron spontaneously transitions from a higher orbit to a lower orbit, a photon is emitted. When discussing transitions, we adopt the general convention of denoting an orbit lower than orbit n as n' , though contextual exceptions to the convention will arise. An example bound-bound transition between $n' = 2$ and $n = 4$, is illustrated in Figure 1.1. The photon energy, E_γ , is the difference between the upper level and lower level binding energies¹,

$$\begin{aligned} E_\gamma &= E_n^{n'} = E_{n'}^n \equiv E_n - E_{n'} \\ &= h\nu_n^{n'} = h\nu_{n'}^n = hc/\lambda_n^{n'} = hc/\lambda_{n'}^n \end{aligned} \quad (1.9)$$

where $\nu_n^{n'}$ and $\nu_{n'}^n$ are the frequencies [Hz] of the emitted and absorbed photons, which correspond to the wavelengths $\lambda_n^{n'}$ and $\lambda_{n'}^n$, respectively. Note that the transition energy for a upward transition, $E_n^{n'}$ is equivalent to the transition energy for a downward transition, $E_{n'}^n$. Invoking Eq. 1.8, the transition energy between states n and n' is written

$$E_n^{n'} = E_{n'}^n = R_z \left[\frac{1}{n'^2} - \frac{1}{n^2} \right]. \quad (1.10)$$

¹ For quantities such as the transition energy, etc., we adopt the convention that the subscript denotes the initial state and the superscript denotes the final state. For example, $E_n^{n'} \equiv E_n^{n'}(\uparrow)$, where the arrow indicates the direction of the transition, in this case emission.

Excitation Energy and Potential

The excitation energy is the excess stored internal energy of the atom when the electron orbits in some state n above the ground state ($n_0 = 1$). The energy required to raise the internal energy of the atom to some state n above the ground state is called the excitation potential. Both these quantities are obtained from Eq. 1.10 according to

$$\chi_n = E_n - E_1 = R_z \left[1 - \frac{1}{n^2} \right]. \quad (1.11)$$

The excitation energy and excitation potential is commonly expressed in units of electron volts [eV]. Note that the excitation energy of the ground state is always 0 eV and that χ_n increases with increasing n . An alternative method for computing the energy of emitted or absorbed photons and/or the transition energy is to employ the difference of the excitation energies,

$$h\nu_n^{n'} = h\nu_{n'}^n = E_n^{n'} = E_{n'}^n \equiv \chi_n - \chi_{n'}. \quad (1.12)$$

Ionization Energy

The minimum energy required to liberate, or free, an electron originally bound in orbit n is called the ionization energy and is expressed

$$I_n = E_\infty - E_n = -E_n = \frac{R_z}{n^2}, \quad (1.13)$$

where the superscript “i” denotes ionization.

Incident photons with energies $E_\gamma = h\nu = hc/\lambda \geq I_n$ can ionize the atom from state n . The photon frequencies and wavelengths must satisfy

$$\nu \geq \frac{R_z}{hn^2}, \quad \text{or} \quad \lambda \leq \frac{hcn^2}{R_z}. \quad (1.14)$$

Photons with $E_\gamma = I_n$ will liberate the electron from the atom, but the electron will have no kinetic energy in the center of mass frame of the atom. For $E_\gamma > E_n^i$, the additional energy above and beyond the ionization energy imparts kinetic energy $(1/2)m_e v^2$ to the free electron.

Ground State Ionization Potential

The ground-state ionization potential is the minimum energy required to ionize the atom from the ground state, $n_0 = 1$. As such, it is equal to the

ionization energy of the ground state. For hydrogenic atoms, the ground-state ionization potential is

$$\chi_1 = E_\infty - E_1 = -E_1 = R_z. \quad (1.15)$$

Grotrian Diagrams and Spectra

In Table 1.1, expressions for the orbital energy, E_n , transition energy, $E_n^{n'}$, excitation potential, χ_n , ionization energy, I_n , and the ground state ionization potential, χ_1 , are listed for hydrogenic atoms. Note that the ionization energy from level n is simply the negative of the orbital binding energy, and that the ionization potential from the ground state is simply the Rydberg constant.

Table 1.1 *Hydrogenic Atom: Energies and Potentials*

Energy/Potential	Symbol	Expression	Equation
Energy	E_n	$= -R_z/n^2$	1.8
Transition Energies	$E_n^{n'}, E_{n'}^n$	$= R_z(1/n'^2 - 1/n^2)$	1.10
Excitation Potential	χ_n	$= R_z(1 - 1/n^2)$	1.11
Ionization Energy	I_n	$= R_z/n^2$	1.13
Ionization Potential	χ_1	$= R_z$	1.15

Energy diagrams called Grotrian diagrams (after Walter Grotrian, who first introduced them in 1928) are a useful visual aid for elucidating the energy structure and transition energies of the atom. Grotrian diagrams of neutral hydrogen (H^0) and the singly ionized helium ion (He^+) are shown in Figures 1.2(a) and 1.2(b). The left axes are the excitation energies (Eq. 1.11) and the right axes are the binding energies (Eq. 1.8). Note that the relative energy structures for the hydrogenic atoms are in direct proportion by the scaling of $Z^2(\hat{m}/m_e)$.

The excitation energy for the ground state is $\chi_1 = 0$, which corresponds to the binding energy $E_1 = -R_z$. As n increases above the ground state, the excitation potential increases (always positive), and reaches a maximum $\chi_\infty = -E_1 = R_z$. The binding energy of the electron is always negative; as n increases, the binding energy decreases in magnitude (increases, getting less negative) until it reaches a maximum value of zero at $n = \infty$.

Selected transitions are shown as vertical lines showing the electron transition; each is labeled with the photon wavelength corresponding to the

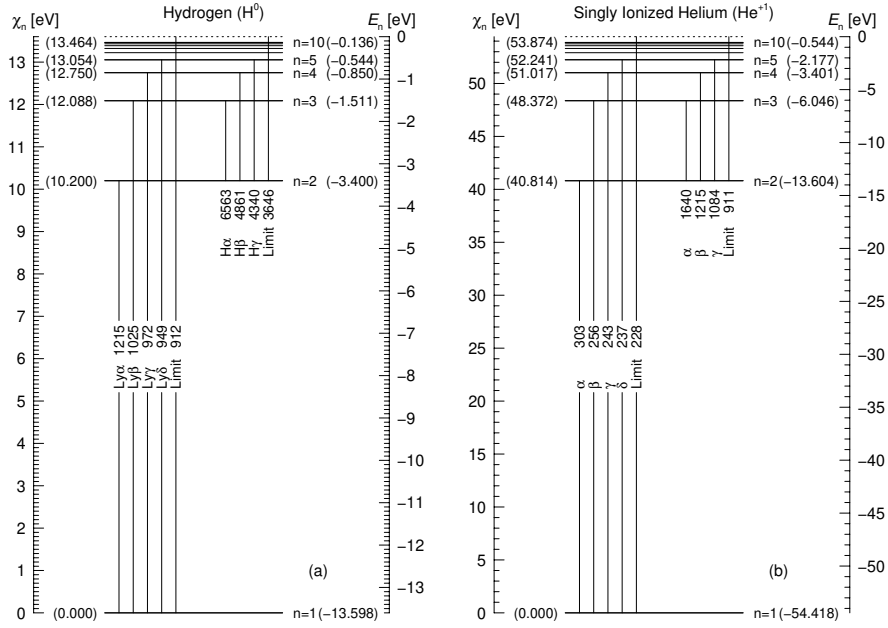


Figure 1.2 Grotrian energy diagrams for neutral hydrogen (panel a) and singly ionized helium (panel b). The left axes gives the excitation energy, χ_n [eV], and the right axes give the electron energy, E_n [eV]. The horizontal lines represent the respective binding energies for a given n . (a) For H^0 , the first four transitions are shown as vertical lines for the $n = 1$, or Lyman series, as is the Lyman limit. Also shown are the first three transitions of the $n = 2$, or Balmer series, as is the Balmer limit. The corresponding wavelengths of the photons are provided (in angströms) for each transition. (b) For He^+ ; the same diagram as for H^0 , except that the energies are all scaled by $Z^2(\hat{m}/m_e)$. (The energies in parenthesis are computed from the expressions listed in Table 1.1; they do not include relativistic and higher-order corrections and are not in full agreement with observations.)

transition energy. For each lower state, n , there is a series of energy differences between E_n and the successive higher states. Consider H^0 ; for all transitions with lower state $n = 1$, the series is called the Lyman series. For lower state $n = 2$, it is called the Balmer series. For the Lyman series, it is customary to denote the first transition in the series ($\Delta n = 1$) as Ly α , the second transition in the series ($\Delta n = 2$) as Ly β , etc., in order of the Greek alphabet. For large Δn , the transitions are simply numbered by Δn , i.e., Ly8, etc. The notation for the Balmer series follows the same convention as the Lyman series except that “H” is the prefix, i.e., H α ($\Delta n = 1$), H β ($\Delta n = 2$), H γ ($\Delta n = 3$), etc. The ionization threshold ($\Delta n = 1 \rightarrow \infty$), is known as the Lyman limit, whereas the series limit $\Delta n = 2 \leftrightarrow \infty$ is called

the Balmer limit. The Lyman series gives rise to ultraviolet spectral lines (912–1216 Å), whereas the Balmer series gives rise to optical spectral lines (3646–6563 Å).

1.2 The Schrödinger Atom

The key to further understanding the atom derives from the particle-wave duality of matter. The fundamental principle of the wave nature of matter, as proposed by Louis De Broglie in 1924, is that the motion of a particle with mass m and velocity v is equivalent to a propagating wave with wavelength

$$\lambda = h/mv. \quad (1.16)$$

De Broglie's hypothesis suggested that the motion of an electron in an atom be investigated using the formalism of wave mechanics. One characteristic of waves is that they constructively or destructively interfere with themselves. If a wave is spatially bound in the interval Δx , it interferes with itself until it sets up a resonance standing wave pattern with $n\lambda = \Delta x$, where n is an integer. These considerations lead directly to the Bohr orbits. A wave confined to a circular path is periodically bound to a path length $2\pi r$. This path length must equal an integer number of wavelengths, $2\pi r = n\lambda = n(h/mv)$, yielding $2\pi r = nh/m\omega r$, or the electron wave will destructively interfere with itself. We immediately obtain $m\omega r^2 = n\hbar$, the consequence of Bohr's first postulate (Eq. 1.2). However, there is no *a priori* reason to assume circular orbits confined to a plane.

1.2.1 The Schrödinger Equation

With three degrees of freedom for the electron, the wave model yields a wide variety of three dimensional orbital configurations that satisfy the condition of constructive interference. The effect is that the radius and angular momentum of bound electrons are quantized into discrete allowed stationary states. The single assumption of the wave model of the atom is that the particles obey the laws of wave mechanics, i.e., the particles are described by a wave function, $\Psi(\mathbf{r}, t) = \Psi(r, \theta, \phi, t)$, which is the solution to a wave equation. In 1925, Erwin Schrödinger derived what is now commonly referred to as Schrödinger's equation,

$$i\hbar \frac{\partial}{\partial t} \Psi(\mathbf{r}, t) = H\Psi(\mathbf{r}, t), \quad (1.17)$$

where the classical, non-relativistic Hamiltonian is the energy operator

$$H = -\frac{\hbar^2}{2\hat{m}}\nabla^2 + V(\mathbf{r}, t), \quad (1.18)$$

and where $-(\hbar^2/2\hat{m})\nabla^2$ is the kinetic energy operator, $V(\mathbf{r}, t)$ is the potential specific to the system under consideration, and \hat{m} is the reduced electron mass.

1.2.2 Properties of the Wave Function

The resulting form of the wave function depends on the exact form of the potential, $V(\mathbf{r}, t)$. However, the wave function must be single valued, piecewise smooth, and bounded. For atomic bound states, the wave function is time varying complex eigenfunction of three eigenvalues, n , l , and m , which are a consequence of the periodic boundary conditions in the azimuthal (l ; $0 \leq \phi \leq 2\pi$) and polar (m ; $0 \leq \theta \leq \pi$) spatial coordinates. The boundary condition on the radial spatial coordinate is that the wave function must vanish as $\mathbf{r} \rightarrow \infty$.

Denoting the eigenstate nlm as a generic state n , the time-dependent wave equation can be written as a spatial part and a time-varying part,

$$\Psi_n(\mathbf{r}, t) = \psi_n(\mathbf{r}) \exp \{-i(E_n/\hbar) t\}, \quad (1.19)$$

where $\psi_n(\mathbf{r})$ satisfies the time-independent Schrödinger equation,

$$H\psi_n(\mathbf{r}) = \left[-\frac{\hbar^2}{2\hat{m}}\nabla^2 + V(\mathbf{r}) \right] \psi_n = E_n\psi_n(\mathbf{r}). \quad (1.20)$$

Eq. 1.20 shows that when the Hamiltonian operates on a wave function having eigenstate n , the result is the product of the energy for that eigenstate, E_n , and the wave function, i.e., $H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r})$, where E_n is called the eigenenergy.

For bound states, i.e., $E < 0$, only certain value of E_n are allowed. Being periodically bounded, the $\psi_n(\mathbf{r})$ are complex functions. The E_n are real and appear in the time-varying part of the wave functions, which reflect the oscillation of the amplitude of $\psi_n(\mathbf{r})$ at arbitrary position $\mathbf{r} = (r, \theta, \phi)$ with frequency $\omega_n = E_n/\hbar$. That is, $\psi_n(\mathbf{r})$ is the wave function amplitude of the eigenstate at \mathbf{r} and this amplitude oscillates at a frequency in proportion to the eigenenergy of the state. The wave functions are orthonormal and obey the condition

$$\langle \psi_{n'}(\mathbf{r}) | \psi_n(\mathbf{r}) \rangle \equiv \int_0^\infty \oint \psi_{n'}^*(\mathbf{r}) \psi_n(\mathbf{r}) dV = \delta_{n'n}, \quad (1.21)$$

where $\delta_{n'n}$ is the Dirac δ -function, which evaluates to $\delta_{n'n} = 1$ when $n' = n$ and $\delta_{n'n} = 0$ when $n' \neq n$. The expression $\langle \psi_{n'}(\mathbf{r}) | \psi_n(\mathbf{r}) \rangle$ is the Dirac notation for the integral over all space, where $\psi_n^*(\mathbf{r}) = \langle \psi_n(\mathbf{r}) |$ and $\psi_n(\mathbf{r}) = | \psi_n(\mathbf{r}) \rangle$.

In 1926, Max Born postulated that the square of the wave function provides the probability of finding the electron at \mathbf{r} at time t within the volume element dV centered on the nucleus, where the volume element is $dV = r^2 dr d\Omega = r^2 dr \sin \theta d\theta d\phi$. Thus, the probability density of the particle position is

$$P_n(\mathbf{r}) dV = \psi_n^*(\mathbf{r}) \psi_n(\mathbf{r}) dV. \quad (1.22)$$

1.2.3 The Schrödinger Wave Function

For hydrogenic atoms, a Coulomb potential, $V(r) = -Ze^2/r$, is applied, where Ze is the charge of the nucleus. For $E \geq 0$ (unbound states), the effective potential yields a continuum of acceptable eigenfunctions. For $E < 0$ (bound states), only certain acceptable eigenfunctions and eigenvalues solve the Schrödinger equation. As such, the wave function is forced to have an integer number of nodes in the radial, and in both the azimuthal and polar coordinates. The bound-state eigenfunctions are of the form

$$\psi_{nlm}(r, \theta, \phi) = R_{nl}(r) Y_{lm}(\theta, \phi), \quad (1.23)$$

where $R_{nl}(r)$ is the radial component, and where $Y_{lm}(\theta, \phi)$ is the spherical harmonic function, which governs the azimuthal and polar modulations of $R_{nl}(r)$. The integers n , l , and m are the eigenvalues, called quantum numbers.

The principle quantum number is

$$n = n_r + l + 1, \quad (1.24)$$

where n_r is the radial quantum number with $n_r = 0, 1, 2, \dots, \infty$. The value of n_r gives the number of radial nodes of $R_{nl}(r)$. The eigenvalue l is the angular momentum quantum number; it provides the number of nodes in the azimuthal direction of the $Y_{lm}(\theta, \phi)$ functions. Examining Eq. 1.24, we see that for a given n , l is a maximum for $n_r = 0$ and is equal to $n - 1$. Since the minimum value of n is $n = 1$ and $n_r \leq n$, the minimum value is $l = 0$. We have the condition

$$l = 0, 1, 2, \dots, n - 1. \quad (1.25)$$

The eigenvalue m is known as the magnetic quantum number; its absolute value provides the number of nodes in the polar direction of the $Y_{lm}(\theta, \phi)$

functions. It is interpreted as the z-axis projection, L_z of the orbital angular momentum, L for state nlm . These projections are limited by the value of l , so that

$$m = -l, -l+1, \dots, 0, \dots, l-1, l. \quad (1.26)$$

A schematic of the allowed quantum numbers for $n = 1, 2$, and 3 is presented in Figure 1.3. In Table 1.2, we summarize the quantum numbers.

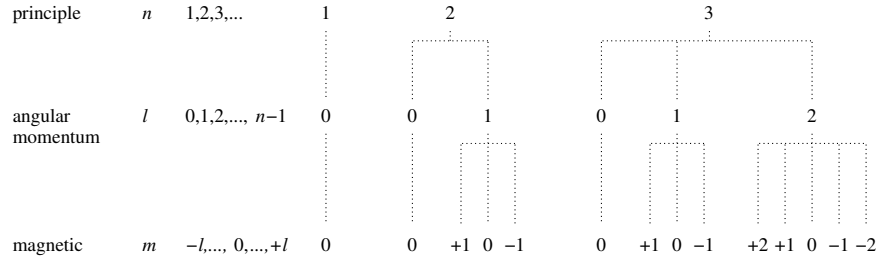


Figure 1.3 A schematic chart of the allowed quantum numbers, nlm , based upon the spatial boundary conditions and wave interference properties of an orbiting electron.

Table 1.2 *Schrödinger model bound state quantum numbers*

	Description	Allowed ranges	Number of states
n	principle	$1 \leq n \leq \infty$	∞
l	angular momentum	$0 \leq l \leq n-1$	n
m	magnetic	$-l \leq m \leq l$	$2l+1$

Angular Momentum and Multiplicity of States

Whereas the constant of motion for the Bohr model is the azimuthal orbital angular momentum, the constants of motion for the Schrödinger wave model are the total angular momentum vector, \mathbf{L} , and the azimuthal component (z axis projection) of the angular momentum, L_z , where,

$$|\mathbf{L}| = L = (L_x^2 + L_y^2 + L_z^2)^{1/2} = \sqrt{l(l+1)} \hbar, \quad (1.27)$$

$$L_z = m\hbar,$$

where L_z is the azimuthal or z component in Cartesian coordinates.

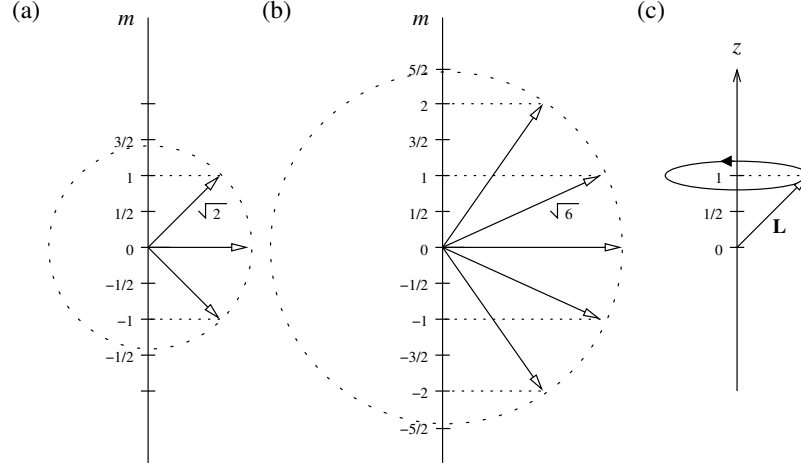


Figure 1.4 A schematic of the angular momentum vector, \mathbf{L} , and the z axis projection, L_z . These two quantities are the constants of motion for the Schrödinger model of the atom. (a) The $l = 1$ case for which $L = \sqrt{2}\hbar$, showing the three possible allowed L_z projections $m = -1, 0, +1$ in units of \hbar . (b) The $l = 2$ case, for which $L = \sqrt{6}\hbar$, showing the five possible allowed L_z projection $m = -2, -1, 0, +1, +2$ in units of \hbar . (c) The precession of \mathbf{L} around the z axis is illustrated for $l = 1$ with $m = +1\hbar$.

The angular momentum vector and its z projection are illustrated in Figure 1.4(a) for $l = 1$ and in Figure 1.4(b) for $l = 2$. The vector \mathbf{L} precesses about the z axis as illustrated in Figure 1.4(c). Its projection on the z axis can be any one of $2l + 1$ allowed projections. For $l = 1$, the magnitude of L is $\sqrt{2}\hbar$ and the possible projections are $-\hbar, 0, +\hbar$; for $l = 2$, the magnitude of L is $\sqrt{6}\hbar$ and the possible projections are $-2\hbar, -\hbar, 0, +\hbar, +2\hbar$.

The number of m stationary states that an electron can occupy for state nl is $g_{nl} = 2l + 1$, which is called the multiplicity of states. The total multiplicity of states for level n is

$$g_n = \sum_{l=0}^{n-1} g_{nl} = \sum_{l=0}^{n-1} (2l + 1) = n^2. \quad (1.28)$$

1.2.4 Energy Structure and Transitions

In general, the expectation value, $\langle y \rangle$ of a given operator, Y , is given by the integral over all space

$$\langle y \rangle = \langle \Psi(\mathbf{r}, t) | Y | \Psi(\mathbf{r}, t) \rangle \equiv \int \oint \Psi^*(\mathbf{r}, t) Y \Psi(\mathbf{r}, t) dV. \quad (1.29)$$

This principle will be instrumental in determining the energy structure and the relative rates of different electron transitions.

The expectation value of the energy for eigenstate nlm is obtained using the Hamiltonian operator (Eq. 1.18). We have

$$E = \langle \Psi_{nlm}(\mathbf{r}) | H | \Psi_{nlm}(\mathbf{r}) \rangle, \quad (1.30)$$

where $\Psi_{nlm}(\mathbf{r})$ is given by Eq. 1.19. Performing the integration, we find that the energies of the stationary states, $\psi_{nlm}(\mathbf{r})$, depend only on the principle quantum number, n , and that the energy for state n is identical to that of the Bohr model (Eq. 1.5), namely

$$E_n = -\frac{\hat{m}e^4 Z^2}{2\hbar^2 n^2} = -\frac{R_z}{n^2}. \quad (1.31)$$

As shown in Eq. 1.28, the multiplicity of states for level n is $g_n = n^2$ for the Schrödinger model, i.e., there are n^2 possible lm states that can be occupied by an electron in level n . Eq. 1.31 indicates that the energy of state n is independent of the lm state of the electron. The lm states are energy degenerate. Consequently, the non-relativistic wave mechanics approach to describing the bound stationary states of electrons in the hydrogen atom yields the same spectral features and the same energy structure as the Bohr model (§ 1.1.1).

The transition energies in the Schrödinger model are identical to those of the Bohr model, namely,

$$E_n^{n'} = E_{n'}^n = E_n - E_{n'}. \quad (1.32)$$

However, not all nlm states can transition to all $n'l'm'$ states, and visa versa.

An advantage of Schrödinger's wave-function formalism for the hydrogenic atomic model is that one can work out the relative probabilities, rates, and intensities of all the possible transitions. We discuss this formalism in Chapter 2, where we show that, under certain physical assumptions, not all transitions between two stationary states nlm and $n'l'm'$ are "allowed". The non-vanishing transition probabilities provide what are known as selection rules (see Table 2.1).

Under the so-called "electric dipole moment" assumption (§ 2.1.1), which

accounts for the dominant terms in the physics of transitions, the electron is restricted to those transitions in which $\Delta l = l' - l = \pm 1$ and $\Delta m = m' - m = 0, \pm 1$. In reality, when higher-order moments of the electric field and the magnetic field are included in the physics, the transitions “forbidden” by the dipole selection rules can (and do) occur. The transition probabilities and selection rules will be discussed in § 2.1.1.

Grotrian Diagram

Because the transition energies depend only on n and are degenerate with l and m , and because the energies are identical to those of the Bohr model, the energy structure and predicted spectrum from the Schrödinger model are identical to those of the Bohr model (see Table 1.1). However, because of the l and m energy degeneracy, more than one of several permitted transitions from upper state n to lower state n' can give rise to spectral lines with the same frequency or wavelength. For this reason, a spectroscopic notation for the initial and final states of each transition was developed.

Each n level is called a shell, and historically, observational spectroscopists named the $n = 1$ level the K shell, the $n = 2$ level the L shell, the $n = 3$ level the M shell, the $n = 4$ level the N shell, etc. For a given shell, n , each angular momentum state, l , is called a subshell. The subshells are also described by spectroscopic notation, “s” for $l = 0$, “p” for $l = 1$, “d” for $l = 2$, and “f” for $l = 3$. These respectively stand for “sharp”, “principle”, “diffuse”, and “fundamental”.

A given subshell, nl , for example $n = 1$ with $l = 0$, is denoted “1s”. For $n = 2$ and $l = 1$, the state is denoted “2p”. The magnetic quantum number is not included in the spectroscopic notation. As an example, there is only a single 1s state ($m = 0$) and there are three 2p states ($m = -1, 0, +1$). For the Schrödinger model, the permitted Ly α transition is denoted 2p–1s, whereas the Ly α transition 2s–1s is forbidden by the dipole selection rules.

In Figure 1.5, we present the Grotrian diagram for $n = 1, 2$, and 3. The downward permitted transitions for Ly α (2p–1s) and H α (3p–2s, 3s–2p, and 3d–2p) are indicated and labeled using spectroscopic notation. Note that there are three possible channels by which an H α photon can be emitted or absorbed. However, the transition probabilities of these three transition (m states not shown) depend upon the initial and final lm states. Thus, given an ensemble of hydrogen atoms, the final emission or absorption intensity of the H α spectral line will be a weighted average of the transition probabilities of the three transitions. This will be discussed further in § 2.2.

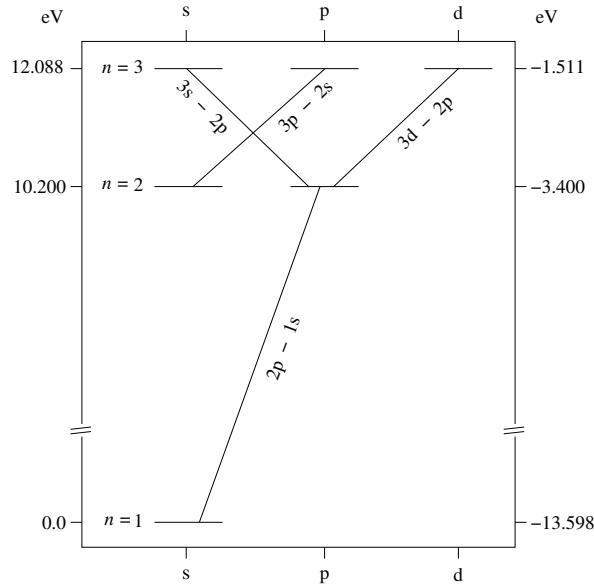


Figure 1.5 Grotrian diagram for the first three principle n levels for the Schrödinger model. The l states increase from left to right, where s denotes $l = 0$, p denotes $l = 1$, and d denotes $l = 2$. The excitation energy is given on the left and the binding energy is given on the right in units of electron volts [eV]. The dipole selection rules dictate that permitted transitions obey $\Delta l = \pm 1$ and $\Delta m = 0, \pm 1$ (m states are not shown in the Grotrian diagram). The permitted $2p \rightarrow 1s$ Ly α transition and H α multiplet ($3p \rightarrow 2s$, $3s \rightarrow 2p$, and $3d \rightarrow 2p$) are illustrated. Not illustrated is the permitted $3p \rightarrow 1s$ Ly β transition.

1.3 The Dirac Atom

As revolutionary and powerful as the Schrödinger model is, it still has substantial shortcomings: (1) the Hamiltonian accounted for non-relativistic energies only, (2) as with the Bohr model, the predicted spectrum of hydrogen does not match the observed spectrum, including mismatches with the observed transition energies and the lack of a prediction of fine splitting of the lines, (3) a property known as electron spin had been discovered and experiments demonstrated that it played a role in the energy structure of atoms; spin was not incorporated into the Schrödinger model.

1.3.1 Spin

In 1922, Otto Stern and Walther Gerlach discovered that silver atoms beamed through a magnetic field were deflected into two discrete beams (they were predicting only a single deflection in proportion to the magnetic moment of

the atom). In 1925, Samuel Goudsmit and George Uhlenbeck showed that when atoms are placed in a magnetic field, the spectral lines split in proportion to the field strength (the phenomenon is now called the Zeeman effect). These results suggested that the electron has an intrinsic magnetic moment proportional to an intrinsic angular momentum, or spin.

The experiments indicate spin has a multiplicity of states $g_s = 2s + 1$ with $s = 1/2$. Thus, $g_s = 2$; there are two spin states. The result is that the total magnetic moment of the atom is due to the combined magnetic moment of the electron orbit and the intrinsic magnetic moment of the electron. Thus, we introduce two additional quantum numbers, $s = 1/2$ and $m_s = \pm 1/2$, and the spin wave function, denoted χ_{m_s} (which obeys all orthonormal properties). When $m_s = +1/2$, the function $\chi_{+1/2}$ is referred to as a spin “up” state, and when $m_s = -1/2$, the function $\chi_{-1/2}$ is referred to as a spin “down” state. For clarification, the magnetic quantum number (z component of the orbital angular momentum) will be written m_l in the context of the Dirac model.

Similar to the orbital angular momentum vector, \mathbf{L} , which has magnitude $L = \sqrt{l(l+1)}\hbar$ with z component $L_z = m_l\hbar$, the electron spin vector, \mathbf{S} , has magnitude $S = \sqrt{s(s+1)}\hbar = \sqrt{3/4}\hbar$ with z component magnitude $S_z = m_s\hbar = \pm(1/2)\hbar$.

1.3.2 Spin-Orbit Coupling

The orbital angular momentum and the spin combine such that the z component of the orbital angular momentum, L_z , no longer retains its status as a constant of motion. The new constants of motion are the total angular momentum (orbital plus spin), denoted \mathbf{J} , and the z component of \mathbf{J} , i.e., J_z . This new quantum number is the result of spin-orbit coupling,

$$\mathbf{J} = \mathbf{L} + \mathbf{S}. \quad (1.33)$$

For hydrogenic atoms, the magnitude of \mathbf{J} is $J = \sqrt{j(j+1)}\hbar$, where the total angular momentum quantum number is given by

$$j = \begin{cases} s & l = 0 \\ l \pm s & l \geq 1, \end{cases} \quad (1.34)$$

where $s = 1/2$, and where the z component is in multiples $J_z = m_j\hbar$, with

$$m_j = m_l + m_s = -j, -j+1, \dots, j-1, j. \quad (1.35)$$

Since l takes on n possible integer values from $l = 0$ to $n - 1$, we see that j takes on n possible values, $j = 1/2, 3/2, \dots, n - 1/2$. Note that for a single electron, $j \neq l$ and m_j must always be a half integer.

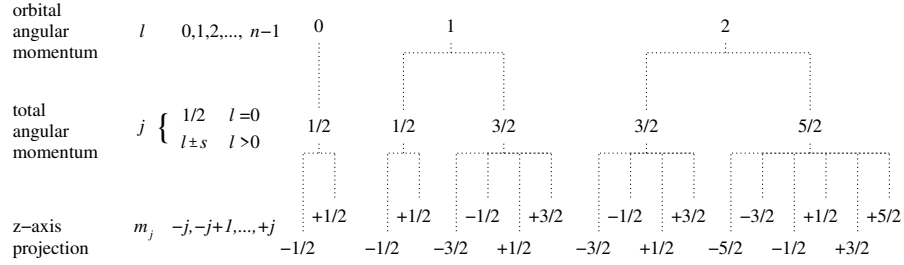


Figure 1.6 A schematic chart, analogous to Figure 1.3, of the total angular momentum quantum numbers, j , and z projection, m_j , for the $l = 0, 1$, and 2 states of the $n = 3$ level due to spin-orbit coupling. Note there is $n = 3$ unique j states, $1/2, 3/2$, and $5/2$. Relative to the Schrödinger model, the number of states for level n are doubled, so that the multiplicity of states for level n is $g_n = 2n^2$.

In Table 1.3, we summarize the quantum numbers. In Figure 1.6, the resulting quantum states for the total angular momentum are illustrated for $l = 1, 2$, and 3 . Due to the multiplicity of states of the electron, $g_s = 2s + 1$, the multiplicity of states for state nl is $g_{nl} = (2s + 1)(2l + 1)$. Effectively, the two spin states of the electron doubles the number of allowed stationary states, which doubles the multiplicity of states for level n (see Eq. 1.28) from n^2 to

$$g_n = \sum_{l=0}^{n-1} (2s + 1)(2l + 1) = 2n^2. \quad (1.36)$$

The spin-orbit coupling (vector addition) rules, given by Eqs. 1.34 and 1.35, are schematically illustrated in Figure 1.7(a) for the state $l = 1$ and spin up state resulting in $j = 3/2$ and $m_j = +3/2$. The vector \mathbf{J} precesses about the z axis in one of $2j + 1 = 4$ possible projections $m_j \hbar$, with $m_j = -3/2, -1/2, +1/2$, or $+3/2$. Only the $m_j = +3/2$ case is illustrated. The \mathbf{L} vector precesses about the \mathbf{J} vector and its z component, L_z , is no longer a constant of motion; the constant of motion is J_z .

The spin, \mathbf{S} , adds to \mathbf{L} , but with only two possible z axis projections $m_s \hbar$, where $m_s = -1/2, +1/2$. The magnitudes of the \mathbf{J} , \mathbf{L} , and \mathbf{S} vectors are illustrated Figure 1.7(b) for $l = 1$ and $j = +3/2$. The magnitude of the orbital angular momentum vector is $\sqrt{l(l+1)} \hbar = \sqrt{2} \hbar$. The magnitude of the spin vector is always $\sqrt{s(s+1)} \hbar = \sqrt{3/4} \hbar$. The magnitude of the total

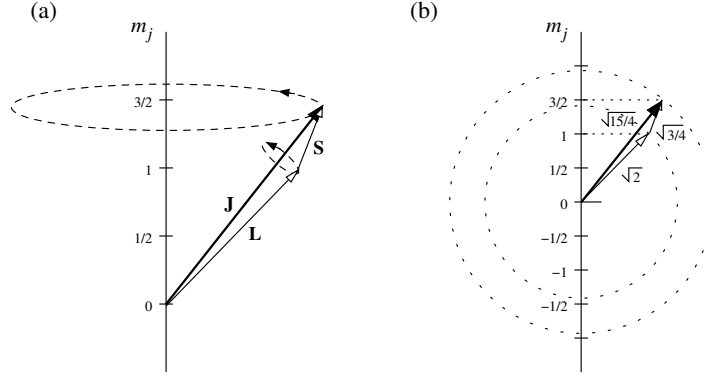


Figure 1.7 (a) The total angular momentum vector, \mathbf{J} , as the vector sum of \mathbf{L} and \mathbf{S} . For this example, $l = 1$ and the electron is in the spin up state, leading to $j = 3/2$. The z axis projection $m_j = +3/2$ was chosen for this illustration. The new constant of motion, \mathbf{J} , precesses azimuthally about the z axis with projection m_j , whereas \mathbf{L} precesses about \mathbf{J} . Note that the projection of \mathbf{L} on the z axis, L_z , is no longer a constant of motion; however \mathbf{S} and m_s remain constants of motion. (b) Schematic of the vector addition emphasizing the magnitudes of the vectors. The total angular momentum, \mathbf{J} , is represented by the thick line with the solid arrow and has magnitude $\sqrt{j(j+1)}\hbar = \sqrt{15/4}\hbar$. The orbital angular momentum and electron spin are shown as thin lines with open arrows, where the magnitude of the angular momentum is $\sqrt{l(l+1)}\hbar = \sqrt{2}\hbar$ and the magnitude of the spin is $\sqrt{3/4}\hbar$, as always. The other possible projections of \mathbf{J} on the z axis (not shown) are $m_j = -3/2, -1/2$, and $+1/2$.

Table 1.3 *Dirac model bound-state quantum numbers*

	Description	Allowed ranges	Number of states
n	principle	$1 \leq n \leq \infty$	∞
l	orbital angular momentum	$0 \leq l \leq n-1$	n
j	total angular momentum	$1/2 \leq j \leq n-1/2$	n
m_l	orbital magnetic	$-l \leq m_l \leq l$	$2l+1$
m_j	total magnetic	$-j \leq m_j \leq j$	$2j+1$
m_s	spin	$m_s = \pm 1/2$	2

All values are integers or integer multiples of $1/2$.

angular momentum vector (illustrated as thick lines with solid arrow), is $\sqrt{j(j+1)}\hbar = \sqrt{15/4}\hbar$. There are four possible z axis projections $m_j = -3/2, -1/2, +1/2, +3/2$, though only the latter is shown. The spin vector is

represented originating at the head of the angular momentum vector (but it does not precess around \mathbf{J} ; it precesses about the z axis).

1.3.3 Energy Structure and Transitions

The Dirac wave functions are represented by the quantum numbers l , j , and m_j , and m_s . These wave functions are spinors, which are beyond the scope of this text.

The Dirac energies are determined using the method presented in Eq. 1.30. Because the interaction energies due to relativistic and spin-orbit effects are on the order of $10^{-5}E_n$, the Hamiltonian can be written in terms of “low-order correction terms” to the Schrödinger Hamiltonian,

$$H = -\frac{\hbar^2}{2\hat{m}}\nabla^2 + V(\mathbf{r}, t) - \frac{p^4}{8\hat{m}^3c^2} + \frac{1}{2\hat{m}^2c^2} \frac{1}{r^2} \frac{dV}{dr} \mathbf{L}\mathbf{S} + \frac{\hbar^2}{8\hat{m}^2c^2} \nabla^2 V(r, t). \quad (1.37)$$

The first two terms are the Schrödinger Hamiltonian. The third term is a relativistic energy correction of order v^2/c^2 . The fourth term is the spin-orbit interaction energy for $l \geq 1$ electrons, which includes a relativistic correction to the electromagnetic potential. The final term is a correction called the Darwin term for $l = 0$ electrons (to account for “zitterbewegung”, a jittering precession of the electron spin due to interaction with the angular momentum of the photon field transporting the electromagnetic force between the nucleus and the electron).

Inserting Eq. 1.37 into Eq. 1.30, yields,

$$E_{nj} = \frac{\hat{m}c^2}{2} \frac{(Z\alpha)^2}{n^2} \left[1 + \frac{(Z\alpha)^2}{n^2} \left(\frac{n}{j+1/2} - \frac{3}{4} \right) \right], \quad (1.38)$$

where $\alpha = e^2/\hbar c$ is the fine structure constant. Note that the energies now depend upon both the n and the j states.

The exact solution was obtained by Dirac using full treatment of relativistic energies, electron spin, and interaction of the electron and nuclear magnetic moments as applied to spinors (two-by-two matrix wave equations). The treatment is well beyond the scope of this text. Dirac’s result is

$$E_{nj} = \hat{m}c^2 \left\{ \left[1 + \left[\frac{Z\alpha}{n - (j-1/2) + [(j+1/2)^2 - (Z\alpha)^2]^{1/2}} \right]^2 \right]^{-1/2} - 1 \right\}, \quad (1.39)$$

which is equivalent to Eq. 1.38 to order $(Z\alpha)^2$.

Eq. 1.38 clearly shows that, for level n , the Dirac energies are shifted

from the Schrödinger energies and that these shifts depend on both n and j . Furthermore, since there are n possible j states for level n , Eq. 1.38 shows that there are n unique energy states in level n . The magnitude of the shifts and splittings are on the order of $\alpha^2 E_n$ ($\simeq 10^{-5} E_n$) and decrease with increasing n . The resulting energy splittings are referred to as “fine structure”.

In Figure 1.8, the quantity $\Delta E/E_n = (E_{nj} - E_n)/E_n$ from Eq. 1.38 is schematically illustrated for $n = 1, 2$, and 3 for the hydrogen atom. We have invoked the spectroscopic notation, i.e., the $n = 2, l = 1$ state is written 2p, etc., with the added convention of including the subscript j to denote the total angular momentum state. Thus, the $n = 2, l = 0, j = 1/2$ state is written 2p_{1/2}, etc.

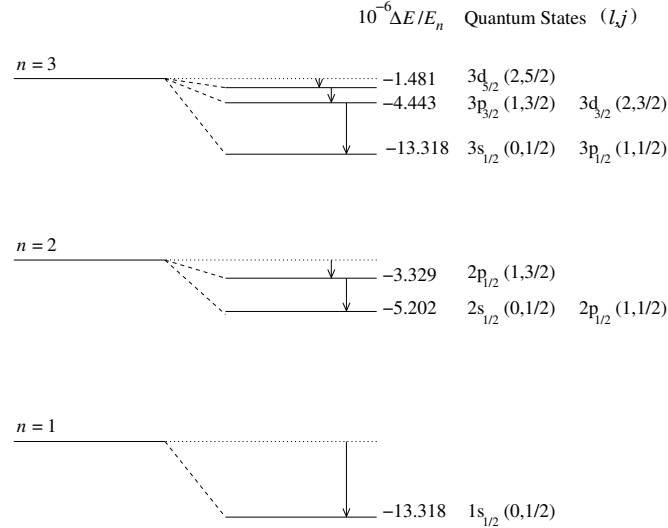


Figure 1.8 The hydrogen fine structure energy shifts and splittings, $10^{-6} \Delta E/E_n$, for the first three principle levels $n = 1, 2$, and 3. Note that there are n fine-structure states for level n . The quantum states are identified using spectroscopic notation. The energy differences between fine structure states are given relative to the Schrödinger states (left).

Transition Energies and Transitions

The result of relativistic spin-orbit coupling is that the binding energy of the electron is increased (the energy is more negative) relative to the non-relativistic Schrödinger treatment. In the Dirac model, since there are $2j + 1$ possible projections for m_j , the E_{nj} are $g_{nj} = 2j + 1$ energy degenerate.

The transition energies in the Dirac model are

$$E_{nj}^{n'j'} = E_{n'j'}^{nj} = E_{nj} - E_{n'j'}, \quad (1.40)$$

where the energies are given by Eq. 1.38 or 1.39. Given the fine structure multiplets, it is clear that the transition energies will cluster in multiplets as well. The closely spaced spectral lines that result from transitions between fine structure multiplets associated with states nl and $n'l'$ are also simply called “a multiplet”.

As with the Schrödinger model, the non-vanishing terms of the dipole moment between stationary states provide the selection rules for permitted transitions. The dipole moment vanishes for all transitions except for those with $\Delta l = l' - l = \pm 1$ and $\Delta m_l = m'_l - m_l = 0, \pm 1$; these selection rules immediately apply to the Dirac model. Since $j = l \pm 1/2$, it also immediately follows that $\Delta j = j' - j = \pm 1$ is a selection rule. An additional rule is that there must be a parity change in the stationary states; this is known as LaPorte’s rule. Parity and LaPorte’s rule will be discussed further in § 3.1.4. The consequence of LaPorte’s rule is that there is no spin flip ($\Delta m_s = m'_s - m_s = 0$) for dipole transitions. The Dirac dipole selection rules are discussed in § 2.3 and listed in Table 2.3

Grotrian Diagram: Doublets and Multiplets

In Figure 1.9, the permitted transitions for hydrogen $\text{Ly}\alpha$ and $\text{H}\alpha$ are illustrated in a Grotrian diagram of the Dirac model. There is only a single $1s$ state with $j = 1/2$ (though its energy is shifted relative to the Schrödinger energy). There are three j states for $n = 2$, but the $2s_{1/2}$ and $2p_{1/2}$ states are energy degenerate (we will see that this degeneracy is broken when we introduce radiative corrections). According to the selection rules, a $2s_{1/2} - 1s_{1/2}$ transition is forbidden, since it would require $\Delta l = 0$. Since the path for an electron in the $2s_{1/2}$ state to transition to the $1s_{1/2}$ state is forbidden, and $1s_{1/2}$ is the only $n = 1$ state, the $2s_{1/2}$ state is “metastable”.

Both the $2p_{3/2} - 1s_{1/2}$ and the $2p_{1/2} - 1s_{1/2}$ transitions are permitted with $\Delta l = \pm 1$ and with $\Delta j = \pm 1$ and 0 , respectively. Thus, the fine structure $\text{Ly}\alpha$ spectral feature is a doublet. In fact, in the Dirac model, all Lyman series spectral features are doublets due to the transitions $np_{1/2} - 1s_{1/2}$ and $np_{3/2} - 1s_{1/2}$, where $n = 2$ is $\text{Ly}\alpha$, $n = 3$ is $\text{Ly}\beta$, etc. The Lyman series doublets are

$$\text{Lyman doublets} \begin{cases} np_{1/2} - 1s_{1/2} \\ np_{3/2} - 1s_{1/2} . \end{cases}$$

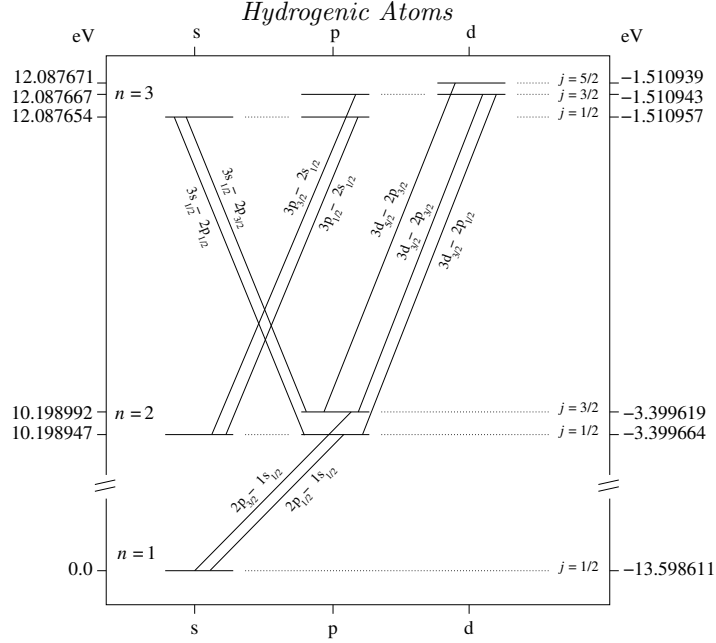


Figure 1.9 The Grotrian diagram for the Dirac model of hydrogen for $n = 1, 2$, and 3 . Excitation energy is shown on the left and binding energy is shown on the right. The energy levels are shifted relative to the Schrödinger model and exhibit fine structure splitting due to the j states. The dipole selection rules are $\Delta l = \pm 1$ and $\Delta j = 0, \pm 1$ for permitted transitions. The Ly α doublet ($2p_{3/2}-1s_{1/2}$ and $2p_{1/2}-1s_{1/2}$) and H α multiplet transitions are illustrated. Not illustrated is the permitted Ly β doublet ($3p_{3/2}-1s_{1/2}$ and $3p_{1/2}-1s_{1/2}$).

Similar discussion applies to the H α transition, which can occur as any one of seven transitions. The $3d_{5/2}-2p_{1/2}$ transition is forbidden, because it would require $\Delta j = \pm 2$, even though $\Delta l = \pm 1$ would otherwise be allowed. Note that the transitions $3p_{3/2}-2s_{1/2}$ and $3d_{3/2}-1p_{1/2}$ are energy degenerate, as are the $3p_{1/2}-2s_{1/2}$ and $3s_{1/2}-2p_{1/2}$ transitions. Thus, the Dirac theory predicts that observed spectra will show five unique features in the H α multiplet. In fact, all Balmer series spectral features are five-fold multiplets from seven transitions,

$$\text{Balmer multiplets} \left\{ \begin{array}{ll} np_{3/2}-2s_{1/2} & nd_{3/2}-2p_{1/2} \quad (\text{energy degenerate}) \\ np_{1/2}-2s_{1/2} & ns_{1/2}-2p_{1/2} \quad (\text{energy degenerate}) \\ nd_{3/2}-2p_{3/2} & \\ nd_{5/2}-2p_{3/2} & \\ ns_{1/2}-2p_{3/2}, & \end{array} \right.$$

where $n = 3$ is H α , $n = 4$ is H β , etc.

Note that fine structure energy splittings decrease with increasing n such that spectral feature multiplets from a high n to a much lower n (except to $n = 1$) are dominated by the low n fine structure. The transitions between fine structure multiplets do not have equal probability of occurring. Thus, the observed spectral multiplets have different intensities (see § 2.3.1).

1.4 Radiative Corrections

Precise measurements of the hydrogen spectrum indicate that the Dirac transition energies do not match observations within one percent. As we discuss here, the solution to improving the model of the atom required quantization of the electromagnetic field, which was not treated by Dirac. Noting that the dipole selection rules do not rule out $\Delta n = 0$ transitions, Willis Lamb and Robert Retherford demonstrated in 1947 that transition frequency between the $2s_{1/2}$ and $2p_{1/2}$ states, which are energy degenerate in the Dirac model because it predicts that two states in the same n level with the same j states have identical energies, is roughly 1000 MHz. The currently accepted transition energy is 4.3×10^{-6} eV corresponding to a photon wavelength of 28.34 cm and a frequency of 1057.77 MHz. This energy difference corresponds to an additional $\sim 1\%$ energy shift, dominated by the shift in the $2s_{1/2}$ state, and has been named the Lamb shift.

The quest to explain the Lamb shift facilitated the development of the theory of quantum electrodynamics, which takes into account the quantized nature of the electromagnetic field. The mathematical formalism of the theory is beyond the scope of this book. Conceptually, the quantum electrodynamics treatment is that the binding force is carried by virtual photons. According to the Heisenberg uncertainty principle, the quantized radiation field can fluctuate in energy on timescales $\Delta t \leq \Delta E/\hbar$. This fluctuating energy within the virtual photon field can take the form of virtual particles as long as the uncertainty principle is obeyed and energy, momentum, charge, and all quantum numbers are conserved.

While transmitting the electromagnetic force, virtual photons often convert into a virtual electron-positron pair, which immediately self-annihilate and convert back into a virtual photon. The lifetime of the electron-positron pair is $\simeq E_\gamma/\hbar$, where E_γ is the energy of the virtual photon. This first-order phenomenon is known as a polarization loop (higher order phenomenon are also occurring).

A “sea” of these loops comprise the “vacuum” energy within the atom. As the “real” electron attracts the virtual positrons and repulses the virtual electrons, this spatially polarizes the electric field of the sea of virtual

electron-positron pairs in the immediate vicinity of the electron, effectively “screening” the bare electron charge. This screening decreases the magnitude of the sensible charge of the electron with distance. Thus, vacuum polarization modifies the electromagnetic force between the electron and the positively charged atomic nucleus.

These modifications to the forces between particles are known as radiative corrections. Whereas the fine structure splittings are on the order $(Z^2/n)\alpha^2$, the radiative corrections are on the order $(Z^2/n)\alpha^3$, less than a percent of the fine structure separations.

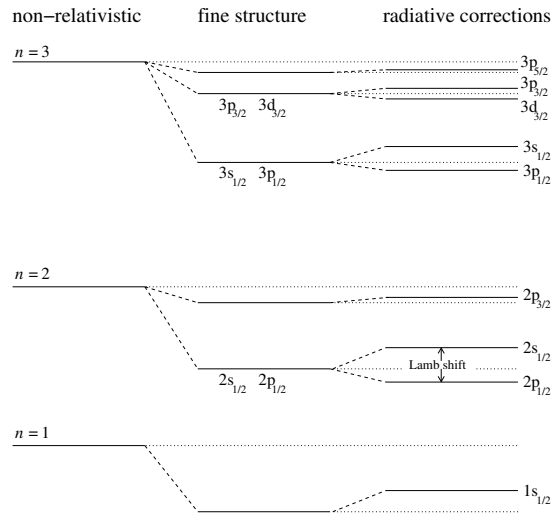


Figure 1.10 The energy shifts in the $n = 1, 2$, and 3 levels as compared to the Schrödinger energies (left) due to the Dirac fine structure (center) and radiative corrections (right). The Dirac energy shifts are on the order $(Z^2/n)\alpha^2$ and the radiative corrections are on the order $(Z^2/n)\alpha^3$; the illustration is not to scale. The dipole selection rules based upon the Dirac model remain intact.

In Figure 1.10, the energy shifts following radiative corrections are illustrated for levels $n = 1, 2$, and 3 . The left most levels are the non-relativistic energies from the Schrödinger model, whereas the central levels are the Dirac fine structure energies. The radiative corrections are shown on the right, though the scale is exaggerated relative to the fine structure splittings. The dipole selection rules for the Dirac model remain intact. Thus, the number of transitions between states are identical to those of the Dirac model. However, the energy degeneracy of identical nj states is broken. Thus, for example, all seven transitions of the $H\alpha$ transition (see Figure 1.9) have unique energies, so that $H\alpha$ is, in reality, a seven-fold spectral multiplet.

1.5 Isotope Shifts

Isotope shifts do not result in energy splittings, but in energy shifts. There are two types of isotope shifts. The first is a mass effect and is simply an energy shift due to differences in the center of electron-nucleus mass between isotopes of the same element. The second is called the mass polarization effect, which is a correction due to the fact that the nucleus is not a point charge, but has a charge distribution, the form of which depends upon the volume of the nucleus, which is isotope sensitive. Since the mass polarization effect results in tiny energy shifts on the order of $10^{-10} A^{2/3} Z^2/n$, where A is the atomic mass number, we do not discuss it further.

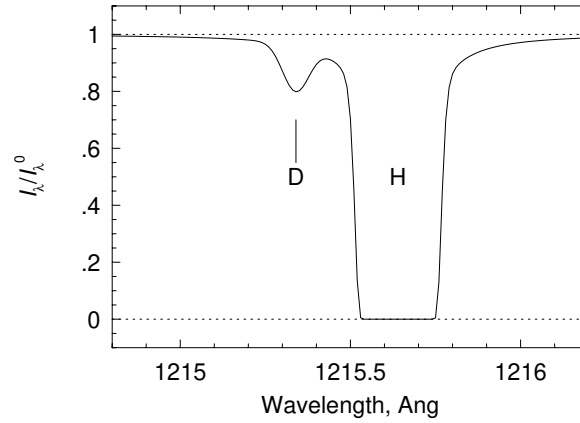


Figure 1.11 Schematic absorption spectrum illustrating the isotopic shift of deuterium relative to that of hydrogen for the Ly α transition. The cosmic deuterium to hydrogen abundance ratio is roughly 10^{-5} , which was adopted for this illustration (thus explaining the weaker absorption for deuterium).

The Bohr, Schrödinger, and Dirac energies are directly proportional to \hat{m} , the reduced mass of the electron,

$$\hat{m} = \frac{m_e}{1 + m_e/m_N}, \quad (1.41)$$

which is derived from the relations $r_N + r_e = r_n$ and $m_e r_e = m_N r_N$, where r_n , r_e and r_N are defined and illustrated in Figure 1.1(a). Thus, isotopic energy shifts for hydrogenic atoms are directly related to m_N , the sum of the proton and neutron masses in the nucleus. Since the addition or subtraction of a single neutron from a nucleus is a proportionally larger fraction of the nuclear mass in the lightest elements, the lightest elements have the greatest energy shifts between isotopes. We discuss heavier nuclei with multiple electrons in § 3.4.

Consider the isotopes hydrogen (1 proton) and deuterium (1 proton + 1 neutron). Since transition energies have proportionality $E \propto \hat{m} \propto 1/\lambda$, the observed fractional wavelength shift between hydrogen and deuterium for any given transition is

$$\frac{\Delta\lambda}{\lambda} = \frac{\lambda_D - \lambda_H}{\lambda_H} = \frac{1/\hat{m}_D - 1/\hat{m}_H}{1/\hat{m}_H} = \frac{\hat{m}_H}{\hat{m}_D} - 1. \quad (1.42)$$

Adopting $m_H = 1.00794$, $m_D = 2.04102$, and $m_e = 5.485799909 \times 10^{-4}$ in atomic mass units, we obtain $\Delta\lambda/\lambda = -0.0275\%$. As illustrated in Figure 1.11, the Ly α transition at $\lambda_H = 1215.6701 \text{ \AA}$ is shifted by -0.3343 \AA for deuterium. For the Ly β transition at $\lambda_H = 1025.7723 \text{ \AA}$, this corresponds to an observed wavelength shift of -0.2825 \AA .

1.6 Continuum States

The continuum states are characterized by positive (unbound) energies. In this case the solution to the Schrödinger equation is not a set of discrete eigenfunctions and eigenenergies; the functions are a continuous family of solutions as a function of energy, which can range from $0 \leq E \leq \infty$. In a semi-classical sense, the continuum states of the electron can be viewed as unbound hyperbolic orbits with a continuum of discrete energy levels analogous to the discrete bound states.

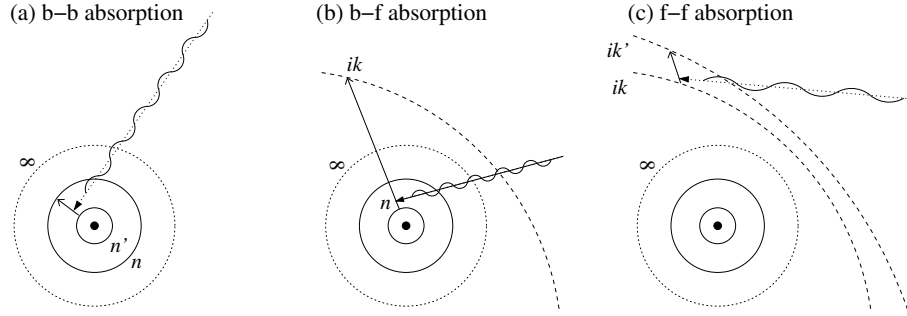


Figure 1.12 Schematic of three primary atomic absorption transitions and the principle quantum numbers employed to quantitatively describe the processes. (a) Bound-bound (b-b) absorption from bound discrete state n' to bound discrete state n , where n' and n are real integers with $n > n'$ and $n < \infty$. (b) Bound-free (b-f) absorption (ionization) from discrete bound state n to free state ik . The liberated electron enters a hyperbolic orbit of the host atom. The value of k is not necessarily an integer. (c) Free-free (f-f) absorption viewed as the transition from hyperbolic orbit ik to higher energy hyperbolic orbit ik' near a host atom, where $k' < k$.

There are two main types of interactions involving unbound states. These are called bound-free, and free-free transitions. Bound-free absorption is the ionization process. Bound-free emission is called recombination; an incident electron is captured to a bound state and emits a photon. A free-free process is a three-body interaction. A free electron in proximity to an atom may absorb an incident photon. An isolated electron cannot absorb a photon; the atom is required to conserve angular momentum.

The distinction between an atom just prior to and just following ionization becomes a formality. For example, an ionized atom can be formally treated as an ion with its electron in an unbound hyperbolic orbit. Similarly, an atom assisting in free-free absorption can be formally treated as having its electron undergoing a transition from one hyperbolic orbit to another. In Figure 1.12, we schematically illustrates the formalism of (a) bound-bound, (b) bound-free, and (c) free-free absorption.

1.6.1 Energy Structure and Transitions

The solution to the Schrödinger equation for positive energies has the same general form as for bound states in that it consists of a radial function and harmonic functions,

$$\psi_{klm}(\mathbf{r}) = R_{kl}(r)Y_{lm}(\theta, \phi), \quad (1.43)$$

where Eq. 1.43 obeys both the orthonormal properties of bound wave functions (see Eq. 1.21).

The radial function is obtained by employing imaginary quantum numbers. For the unbound state the imaginary quantum number ik , which is defined as $ik = (R_z/E)^{1/2}$, which relates the principle unbound state to the kinetic energy of the free electron (e.g., Menzel & Pekeris, 1935). We have

$$E = \frac{1}{2}m_e v^2 = \frac{R_z}{k^2}, \quad (1.44)$$

where $R_z = Z^2(\hat{m}/m_e)R$ is the Rydberg constant for a hydrogenic atom with nuclear charge Z (see Eq. 1.7). The value of k is not restricted to an integer and can range continuously over $0 \leq k \leq \infty$. Note that smaller k indicates a higher energy electron and larger k indicates a lower energy electron. For $k = \infty$, the electron is at rest in the rest-frame of the atom. Clearly, Eq. 1.44 is written for non-relativistic velocities, which would correspond to a non-vanishing k .

In Figure 1.12(a), bound-bound absorption is shown for a transition from bound orbit n to bound orbit n' , where both n and n' are real integers. The

transition energy is given by Eq. 1.10,

$$E_n^{n'} = E_{n'}^n = R_z \left[\frac{1}{n'^2} - \frac{1}{n^2} \right]. \quad (1.45)$$

Bound-Free

In Figure 1.12(b), bound-free absorption (ionization) is shown for a transition from bound orbit n to unbound orbit ik . The energy balance is such that the energy of the incident photon must overcome the binding energy of state n following which any remaining energy is converted into kinetic energy of the electron. The relationship is

$$\frac{1}{2}m_e v^2 = E_\gamma - I_n = h\nu - \frac{R_z}{n^2}. \quad (1.46)$$

From $\frac{1}{2}m_e v^2 = R_z/k^2$, we obtain the bound-free transition energy,

$$h\nu = R_z \left[\frac{1}{n^2} + \frac{1}{k^2} \right]. \quad (1.47)$$

Near the energy threshold of ionization, i.e., as $k \rightarrow \infty$, the electron has a vanishing amount of kinetic energy. Thus, we see that the photon frequency (or wavelength) at the ionization threshold ($k = \infty$) is

$$\nu = \frac{R_z}{hn^2}, \quad \text{or} \quad \lambda = \frac{hcn^2}{R_z}. \quad (1.48)$$

In the ultraviolet and optical regime of the electromagnetic spectrum, bound-free processes are observed as “ionization edges”. Ionization edges are absorption features that reduce the continuum in a spectrum in the wavelength range given by Eq. 1.48 due to the absorption of ionizing photons. As such they have a definite onset wavelength, above which the continuum is not modified by the ionization process. We further discuss the absorption behavior due to bound-free processes in §§ 2.4 and 2.6.3.

Free-Free

In Figure 1.12(c), free-free absorption is shown for a transition from unbound orbit ik to unbound orbit ik' . Note that the $k' < k$, because the energy of the electron has increased. The energy balance is such that the final kinetic energy of the electron is the sum of its initial kinetic energy and the energy of the incident photon

$$\frac{1}{2}m_e v_{k'}^2 = \frac{1}{2}m_e v_k^2 + h\nu. \quad (1.49)$$

From $\frac{1}{2}m_e v^2 = R_z/k^2$, the free-free transition energy is

$$h\nu = R_z \left[\frac{1}{k'^2} - \frac{1}{k^2} \right]. \quad (1.50)$$

Spectral features from free-free absorption is predominantly observed in the infrared, micro-wave, and radio bands of the electromagnetic spectra. The processes does not give rise to spectra features in the ultraviolet and optical bands. For these reasons, we will not discuss free-free absorption further.

2

Transitions of Hydrogenic Atoms

One of the great successes of the Schrödinger and Dirac models of hydrogenic atoms is the built-in formalism for theoretically describing transition rates and therefore the relative strength of emission lines and absorption lines. The formalism also provides for an understanding of the continuum states and therefore the strength of absorption via the ionization process.

We will show that the theoretical treatment of transitions requires spatial integrals over the time-dependent wave functions for both discrete bound states (bound-bound transitions) and continuum states (bound-free transitions). These “overlap” integrals lead to a series of transition rates for spontaneous emission, stimulated emission, and absorption, as summarized by the Einstein coefficients. Simplifying assumptions are discussed that lead to the so-called dipole selection rules for “permitted” transitions. The relative strengths of the transitions are parameterized using a unitless quantity called the oscillator strength. We will derive and discuss the oscillator strengths for both the Schrödinger states and for the fine-structure multiplets, and derive the relative strengths of doublet and triplet transitions.

For bound-bound transitions, we derive the natural broadening function, which follows from the quantum mechanics that “blurs” the bound-state energy levels and leads to a distribution in transition frequencies. The broadening for each transition is characterized by a quantity called the damping constant, which we also derive. Finally, we derive the emission and absorption power (energy per unit time) of transitions, and introduce the absorption cross sections for both bound-bound and bound-free absorption. Throughout, the expressions for hydrogenic Lyman series lines are provided.

2.1 Transition Probabilities

Within the atom, which comprises time-dependent charge densities, there naturally persists a time varying electromagnetic field. Including this interaction and assuming the weak field approximation, the Hamiltonian is written,

$$H = -\frac{\hbar^2}{2\hat{m}}\nabla^2 + V(\mathbf{r}, t) - i\hbar\frac{e}{\hat{m}}\mathbf{A}(\mathbf{r}, t) \cdot \nabla, \quad (2.1)$$

where $\mathbf{A}(\mathbf{r}, t)$ is the electromagnetic vector potential. The transition probabilities $[\text{s}^{-1}]$ for spontaneous downward transitions from upper state nlm to lower state $n'l'm'$, are the expectation values of the weak field operator (according to Eq. 1.29), and can be written

$$A_{nlm}^{n'l'm'} = \oint W_{nlm}^{n'l'm'}(\theta, \phi) d\Omega, \quad (2.2)$$

where

$$W_{nlm}^{n'l'm'}(\theta, \phi) = \frac{e^2\hbar\nu_n^{n'}}{\hat{m}^2c^3} |\langle \Psi_{n'l'm'}^*(\mathbf{r}, t) | \mathbf{A}(\mathbf{r}, t) \cdot \nabla | \Psi_{nlm}(\mathbf{r}, t) \rangle|^2, \quad (2.3)$$

where $\nu_n^{n'} = (E_n - E_{n'})/h$.

2.1.1 The Dipole Approximation

In the weak field approximation, the magnitude of the vector potential can be written $|\mathbf{A}(\mathbf{r}, t)| \simeq 1 + (i\mathbf{k} \cdot \mathbf{r}) + \frac{1}{2!}(i\mathbf{k} \cdot \mathbf{r})^2 + \dots$, where \mathbf{k} is wave number in the propagation direction of the radiation. Since $|r| \simeq a_0 \simeq 10^{-9}$ cm, in the optical and ultraviolet regime where $|k| \propto 1/\lambda \simeq 10^5$ cm, we have $(\mathbf{k} \cdot \mathbf{r}) \ll 1$, so that $|\mathbf{A}(\mathbf{r}, t)| \sim 1$. This approximation does not hold in the x-ray and gamma-ray spectral region.

Under these conditions, the integral in Eq. 2.2 can be simplified,

$$A_{nlm}^{n'l'm'} = \frac{32\pi^3 e^2 (\nu_n^{n'})^3}{3\hbar c^3} |\mathbf{r}_{nlm}^{n'l'm'}|^2. \quad (2.4)$$

where,

$$\begin{aligned} \mathbf{r}_{nlm}^{n'l'm'} &= \langle \psi_{n'l'm'}(\mathbf{r}) | \mathbf{r} | \psi_{nlm}(\mathbf{r}) \rangle \\ &= \hat{\mathbf{r}} \int_0^\infty R_{n'l'}(r) R_{nl}(r) r^3 dr \oint Y_{l'm'}^*(\phi, \theta) Y_{lm}(\phi, \theta) d\Omega \\ &= R_{nl}^{n'l'} Y_{lm}^{l'm'} \hat{\mathbf{r}}, \end{aligned} \quad (2.5)$$

where $R_{nl}^{n'l'}$ is the integral over r , called the radial overlap integral, and

$Y_{lm}^{l'm'}$ is the integral over all solid angle. An interpretation of Eq. 2.5 is that it is the spatial moment of the charge density between two stationary states from which the dipole moment of the charge densities, $\mathbf{d}_{nlm}^{n'l'm'} = e \mathbf{r}_{nlm}^{n'l'm'}$, is computed. For this reason, the above treatment is called the dipole approximation.

The quantity $A_{nlm}^{n'l'm'}$ is the so-called Einstein A coefficient for a spontaneous transition from state nml to state $n'l'm'$ (see § 2.1.2). The radial overlap integral of state nl vanishes, i.e., $R_{nl}^{nl} = 0$; thus, a single stationary state has no transition probability and does not radiate (this resolved long-standing questions regarding the stability of the atom).

2.1.2 Einstein Coefficients

Though the Schrödinger transitions are energy degenerate for lm states, their transition probabilities (e.g., Eq. 2.4) are not. As shown in Figure 2.1, in addition to spontaneous emission, atoms undergo transitions that are induced by incident photons. The first is stimulated emission, where an incident photon with energy very close to a transition energy can induce a downward emitting transition, and the second is absorption, where a photon with energy very close to a transition energy can induce an upward transition by converting its energy into excitation energy (the photon itself vanishes). In 1916, Albert Einstein elegantly summarized these transition rates into a form that remains highly useful.

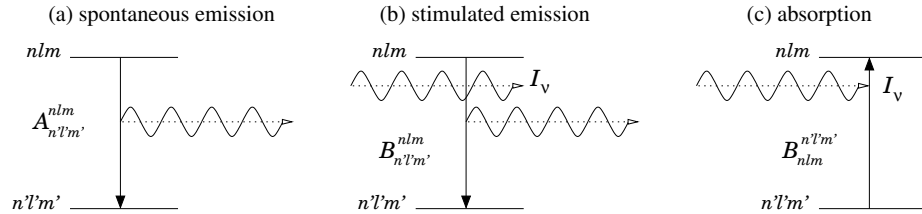


Figure 2.1 Schematics of (a) spontaneous emission, (b) stimulated emission, and (c) absorption. The rate of spontaneous emission is given by the Einstein A coefficient, $A_{n'l'm'}^{nlm}$. For stimulated emission and absorption, incident photons of specific intensity \mathcal{I}_{ν} induce the transition. The rates are given by the Einstein B coefficients, where the rate of stimulated emission is $4\pi B_{n'l'm'}^{nlm} \mathcal{I}_{\nu}$ and the rate of absorption is $4\pi B_{nlm}^{n'l'm'} \mathcal{I}_{\nu}$. These rates are expressed as inverse time [s^{-1}].

Accounting for spontaneous emission, stimulated emission, and absorption, and defining the rates in inverse seconds, Einstein formulated the def-

initions

$$\begin{aligned}
A_{nlm}^{n'l'm'} &= \text{spontaneous emission rate} & A_{nlm}^{n'l'm'} [\text{s}^{-1}] \\
4\pi B_{nlm}^{n'l'm'} \mathcal{I}_\nu &= \text{stimulated emission rate} & B_{nlm}^{n'l'm'} [\text{cm}^2 \text{ erg}^{-1} \text{ s}^{-1}] \\
4\pi B_{n'l'm'}^{nlm} \mathcal{I}_\nu &= \text{absorption rate} & B_{n'l'm'}^{nlm} [\text{cm}^2 \text{ erg}^{-1} \text{ s}^{-1}],
\end{aligned} \tag{2.6}$$

where \mathcal{I}_ν is the specific intensity¹ of the ambient radiation field. These are the famous Einstein A and B coefficients². The units of each coefficient are given to the right of Eq. 2.6. Physically, the Einstein A coefficient is the inverse of half the e -folding time, $\Delta t/2 = 1/A_{nlm}^{n'l'm'}$, for state nlm to spontaneous decay to state $n'l'm'$ and emit radiation, i.e.,

$$\Psi_{nlm}(\mathbf{r}, t) \propto \exp \left\{ -(A_{nlm}^{n'l'm'}/2) t \right\}. \tag{2.7}$$

The Einstein B coefficients are interpreted as the cross-sections per unit energy per unit time for stimulated emission and absorption, respectively. The e -folding times are the inverse rates, $\Delta t/2 = 1/4\pi B_a^b \mathcal{I}_\nu$, yielding

$$\begin{aligned}
\Psi_{nlm}(\mathbf{r}, t) &\propto \exp \left\{ -(2\pi B_{nlm}^{n'l'm'} \mathcal{I}_\nu) t \right\} \\
\Psi_{n'l'm'}(\mathbf{r}, t) &\propto \exp \left\{ -(2\pi B_{n'l'm'}^{nlm} \mathcal{I}_\nu) t \right\},
\end{aligned} \tag{2.8}$$

the time dependence for stimulated emission and absorption transitions, respectively, between states nlm and $n'l'm'$.

Once $A_{nlm}^{n'l'm'}$ is determined (see Eq. 2.4), the Einstein B coefficients can be determined using the Einstein relations³

$$\begin{aligned}
B_{nlm}^{n'l'm'} &= \frac{c^2}{8\pi h \nu^3} A_{nlm}^{n'l'm'} & \text{stimulated} \leftrightarrow \text{spontaneous emission} \\
B_{n'l'm'}^{nlm} &= \frac{g_{nl}}{g_{n'l'}} B_{nlm}^{n'l'm'} & \text{absorption} \leftrightarrow \text{stimulated emission},
\end{aligned} \tag{2.9}$$

where g is the multiplicity of states and where ν is the frequency of the incident photon that stimulates emission ($\nu \simeq E_n^{n'}/h$).

¹ The units of specific intensity are $\text{erg cm}^{-2} \text{ s}^{-1} \text{ str}^{-1} \text{ Hz}^{-1}$. See Chapter 7 for the full definition.

² As defined here, the Einstein coefficients are averaged over all solid angle and are thus isotropic with θ and ϕ . Some texts account for anisotropy in the emission and absorption by defining the coefficients per unit steradian. In such cases, the units stated in Eq. 2.6 also include str^{-1} , there is an additional factor of 4π in the rate for spontaneous emission, and the 4π does not appear in the rates for stimulated emission and absorption.

³ The quantity $8\pi h \nu^3/c^2$ has units $[\text{erg cm}^{-2}]$. Note that if the coefficients are defined per unit steradian, the 8π is replaced by 2.

2.2 Schrödinger Oscillator Strengths

When quantifying relative transition strengths, it is common practice to invoke the oscillator strength, which, for spontaneous emission, is related to the Einstein A coefficient via the relation

$$A_{n'l'm'}^{n'l'm'} = \frac{8\pi^2 e^2 (\nu_n^{n'})^2}{\hat{m} c^3} f_{n'l'm'}^{n'l'm'}, \quad (2.10)$$

where the oscillator strength for spontaneous emission is defined by,

$$f_{n'l'm'}^{n'l'm'} = \frac{4\pi\hat{m}}{3\hbar} \nu_n^{n'} \left| \mathbf{r}_{n'l'm'}^{n'l'm'} \right|^2. \quad (2.11)$$

Substitution of Eq. 2.11 into Eq. 2.10 recovers Eq. 2.4. The oscillator strength is a unitless number that quantifies transition rates, and therefore emission line intensities and absorption power.

To obtain the oscillator strength for spontaneous emission defined in Eq. 2.11; we are required to calculate $|\mathbf{r}_{n'l'm'}^{n'l'm'}|^2$, the squared magnitude of the dipole moment, which requires that we evaluate the two integrals $R_{n'l'm'}^{n'l'm'}$ and $Y_{n'l'm'}^{n'l'm'}$ given by Eq. 2.5. Evaluation of these integrals (see Bethe & Salpeter, 1957) reveals that the $Y_{n'l'm'}^{n'l'm'}$ vanish for all values of l' , l , m' , and m except for the conditions $l' = l \pm 1$, $m' = m \pm 1$, and $m' = m$. The non-vanishing terms of the dipole moment between stationary states provide the dipole selection rules for transitions. Employing $\Delta m = m' - m$, and $\Delta l = l' - l$, we have

Table 2.1 *Schrödinger Dipole Selection Rules*

Transition	Rule
Δl	± 1
Δm	$0, \pm 1$

We remind the reader that the dipole selection rules are based upon application of the weak-field dipole approximation described above. Under this approximation, only transitions obeying the rules stated in Table 2.1 have non-zero oscillator strengths and therefore non-zero transition probabilities. Such transitions are called “E1”, “allowed” and/or “permitted”.

2.2.1 Averaged Oscillator Strengths

For the Schrödinger model, the $2l + 1$ m states are energy degenerate for a given n level. Therefore, the strength of a single spectral line is the sum of the respective $nl \rightarrow n'l'$ transitions in proportion to their $2l + 1$ and $2l' + 1$ multiplicities. We thus invoke the average oscillator strength,

$$f_{nl}^{n'l'} = \frac{4\pi\hat{m}}{3\hbar} \nu_n^{n'} \left| \mathbf{r}_{nl}^{n'l'} \right|^2, \quad (2.12)$$

where

$$\left| \mathbf{r}_{nl}^{n'l'} \right|^2 = \sum_{m'=-l'}^{l'} \left| \mathbf{r}_{nlm}^{n'l'm'} \right|^2, \quad (2.13)$$

is the squared magnitude of the average dipole moment summed over all lower states for permitted transitions. The average rate for spontaneous emission is the sum

$$A_{nl}^{n'l'} = \frac{8\pi^2 e^2 (\nu_n^{n'})^2}{\hat{m} c^3} f_{nl}^{n'l'}. \quad (2.14)$$

Using the Einstein relations (Eq. 2.9), we can then obtain the average Einstein B coefficients, $B_{nl}^{n'l'}$ $B_{n'l'}^{nl}$, respectively.

Following the methods of Bethe & Salpeter (1957), evaluation of Eq. 2.13 yields

$$\left| \mathbf{r}_{nl}^{n'l'} \right|^2 = \frac{\max(l, l')}{2l+1} \left(R_{nl}^{n'l'} \right)^2, \quad (2.15)$$

where $l' = l + 1$ or $l' = l - 1$, and $g_{nl} = 2l + 1$ is the multiplicity of the initial state. Substituting Eq. 2.15 into Eq. 2.12,

$$\begin{aligned} f_{nl}^{n'l'} &= \frac{4\pi\hat{m}}{3\hbar} \nu_n^{n'} \frac{\max(l, l')}{2l+1} \left(R_{nl}^{n'l'} \right)^2 \\ &= \frac{1}{3} \frac{\max(l, l')}{2l+1} \frac{(n+n')(n-n')}{n^2 n'^2} \left(\frac{R_{nl}^{n'l'}}{\hat{a}_0} \right)^2, \end{aligned} \quad (2.16)$$

where we have invoked $\nu_n^{n'} = E_n^{n'}/h$, $E_n^{n'} = R_z(1/n^2 - 1/n'^2)$, $R_z = e^2/2\hat{a}_0$, and $\hat{a}_0 = \hbar^2/\hat{m}e^2$, which follow from Eqs. 1.10, 1.6, and 1.4, respectively. The overlap integral, $R_{nl}^{n'l'}$, will be evaluated in § 2.2.2 (see Eq. 2.21).

Note that the multiplicity of *initial* state, $g_{nl} = 2l + 1$, is the factor that effectively averages the sum for the transition $nl \rightarrow n'l'$. When the average oscillator strength is computed for the reverse transition, $n'l' \rightarrow nl$, the results is identical to Eq. 2.16, except that the denominator is $g_{n'l'} = 2l' + 1$.

Thus, we have the general relationship,

$$g_{nl} f_{nl}^{n'l'} = g_{n'l'} f_{n'l'}^{nl}, \quad (2.17)$$

where the statistical weights represent the initial states of the transitions. We see that if $l' = l + 1$, then $f_{nl}^{n'l'} > f_{n'l'}^{nl}$, and if $l' = l - 1$, then $f_{nl}^{n'l'} < f_{n'l'}^{nl}$. For example, the downward $2p \rightarrow 1s$ Ly α transition has a larger oscillator strength than the upward $1s \rightarrow 2p$ transition by the ratio 3 : 1. However, the downward $3s \rightarrow 2p$ H α transition oscillator strength is smaller than the upward $2p \rightarrow 3s$ transition by 1 : 3. For a given transition, whether the emission or absorption has the larger oscillator strength depends upon the direction in which l changes.

2.2.2 The Radial Overlap Integrals

In order to obtain the full expressions for the oscillator strengths, we are required to evaluate the overlap integrals $R_{n'l+1}^{nl}$ and $R_{n'l-1}^{nl}$ appearing in Eq. 2.16, where

$$R_{nl}^{n'l'} = \int_0^\infty R_{n'l'}(r) R_{nl}(r) r^3 dr. \quad (2.18)$$

Accounting for the dipole selection rule $\Delta l = \pm 1$, we note that the overlap integral is symmetric with respect to the change in n and l ,

$$\int_0^\infty R_{n'l-1}(r) R_{nl}(r) r^3 dr = \int_0^\infty R_{nl}(r) R_{n'l-1}(r) r^3 dr. \quad (2.19)$$

Since the choice of $n'l'$ as the upper state and nl as the lower state is an arbitrary convention, we have the symmetry

$$R_{n'l-1}^{nl} = R_{nl}^{n'l-1}. \quad (2.20)$$

Evaluation of $R_{nl}^{n'l-1}$ was worked out in closed form by Niel Gordon (1929) using generalized hypergeometric functions. For $n' \neq n$,

$$\frac{R_{nl}^{n'l-1}}{\hat{a}_0} = \frac{(-1)^{n'-l}}{4(2l-1)!} \sqrt{\frac{(n+l)!(n'+l-1)!}{(n-l-1)!(n'-l)!}} \frac{(4n'n)^{l+1} (n-n')^{n+n'-2l-2}}{(n+n')^{n+n'}} M_{nm}^{n'l-1} \quad (2.21)$$

where

$$M_{nm}^{n'l-1} = F(-n_r, -n'_r, 2l, -u) - \left[\frac{n-n'}{n+n'} \right]^2 F(-n_r-2, -n'_r, 2l, -u) \quad (2.22)$$

where $u = 4n'n/(n-n')^2$, and $n_r = n-l-1$ with $n'_r = n'-l$ (recall $l = l' - 1$). The F are the generalized hypergeometric functions,

$$F(\alpha, \beta, \gamma, x) = \frac{\Gamma(\gamma)}{\Gamma(\beta)\Gamma(\gamma-\beta)} \int_0^1 \frac{y^{\beta-1} (1-y)^{\gamma-\beta-1}}{(1-xy)^\alpha} dy, \quad (2.23)$$

where $\Gamma(x)$ is the Gamma function (see Eq. 2.38). The geometric function can be evaluated using the series (cf. Abramowitz & Stegun, 1972),

$$F(\alpha, \beta, \gamma, x) = \sum_{j=0}^{\infty} \frac{(\alpha)_j (\beta)_j}{(\gamma)_j} \frac{x^j}{j!} = 1 + \frac{\alpha\beta}{\gamma} x + \frac{\alpha(\alpha+1)\beta(\beta+1)}{\gamma(\gamma+1)} \frac{x^2}{2!} + \cdots, \quad (2.24)$$

where $(x)_j$ is the Pochhammer symbol as defined

$$(y)_j = \prod_{i=0}^{j-1} (y+i) = y(y+1)(y+2) \cdots (y+j-1). \quad (2.25)$$

Note that α and β are negative integers, which ensures that the hypergeometric series in Eq. 2.24 terminates.

2.2.3 Lyman and Balmer Series Oscillator Strengths

Consider the Lyman series transitions between the $1s$ – np states, for various n . Evaluating Eq. 2.21, we have

$$\left(\frac{R_{n1}^{10}}{\hat{a}_0} \right)^2 = \left(\frac{R_{10}^{n1}}{\hat{a}_0} \right)^2 = 2^8 n^7 \frac{(n-1)^{2n-5}}{(n+1)^{2n+5}}. \quad (2.26)$$

For the emission (downward) transition, we employ Eq. 2.16, and obtain

$$f_{n1}^{10} = \frac{1}{3^2} \frac{(n+1)(n-1)}{n^2} \left(\frac{R_{n1}^{10}}{\hat{a}_0} \right)^2 = \frac{2^8 n^5 (n-1)^{2n-4}}{3^2 (n+1)^{2n+4}}. \quad (2.27)$$

For $\text{Ly}\alpha$ ($n = 2$), we obtain $f_{21}^{10} = 2^{13}/3^{10} = 8192/59,049 = 0.13873$. To obtain the mean oscillator strength for absorption, we apply Eq. 2.17,

$$f_{10}^{n1} = \frac{g_{n1}}{g_{10}} f_{n1}^{10} = 3 f_{n1}^{10} \quad (2.28)$$

which yields $f_{10}^{21} = 0.41618$.

In Table 2.2, we list the expressions for the average absorption and emission oscillator strengths for the Lyman and Balmer series transitions of hydrogen, computed directly from Eq. 2.16 and Eq. 2.17. For each transitions, the ratio $(R_{n'l'}^{nl}/\hat{a}_0)^2 = (R_{nl}^{n'l'}/\hat{a}_0)^2$ is also given.

Table 2.2 *Lyman and Balmer series average oscillator strengths*

lower $n'l'$	upper nl	$(R_{nl}^{n'l'}/\hat{a}_0)^2$	absorption $f_{nl}^{nl'}$	g_{nl}	$g_{n'l'}$	emission $f_{n'l'}^{nl}$
1s	np	$\frac{2^8 n^7 (n-1)^{2n-5}}{(n+1)^{2n+5}}$	$\frac{2^8 n^5 (n-1)^{2n-4}}{3(n+1)^{2n+4}}$	1	3	$\frac{1}{3} f_{nl}^{nl'}$
2s	np	$\frac{2^{17} n^7 (n^2-1)(n-2)^{2n-6}}{(n+2)^{2n+6}}$	$\frac{2^{17} n^7 (n^2-1)(n-2)^{2n-5}}{3 \cdot 4(n+2)^{2n+5}}$	1	3	$\frac{1}{3} f_{nl}^{nl'}$
2p	nd	$\frac{2^{19} n^9 (n^2-1)(n-2)^{2n-7}}{3(n+2)^{2n+7}}$	$\frac{2^{20} n^9 (n^2-1)(n-2)^{2n-6}}{3^3 \cdot 4(n+2)^{2n+6}}$	3	5	$\frac{3}{5} f_{nl}^{nl'}$
2p	ns	$\frac{2^{15} n^9 (n-2)^{2n-6}}{3(n+2)^{2n+6}}$	$\frac{2^{15} n^9 (n-1)^{2n-5}}{3^3 \cdot 4(n+1)^{2n+5}}$	3	1	$\frac{3}{1} f_{nl}^{nl'}$

2.3 Fine-Structure Oscillator Strengths

An important relationship within the Dirac model is that

$$\left| \mathbf{r}_{nl}^{n'l'} \right|^2 = \sum_{j'} \left| \mathbf{r}_{nlj}^{n'l'j'} \right|^2 = \sum_{j'} \sum_{m'_j} \left| \mathbf{r}_{nljm_j}^{n'l'j'm'_j} \right|^2. \quad (2.29)$$

This sum rule shows that the jm_j averaged dipole moment for fine structure states is equivalent to the m_l averaged dipole moment of the Schrödinger states (Eq. 2.13). Again, following the methods of Bethe & Salpeter (1957), we find that the integrals for the $Y_{nljm_j}^{n'l'j'm'_j}$ vanish except for certain relationships between the quantum numbers.

We adopt the relationships $\Delta l = l' - l$, $\Delta s = s' - s$, $\Delta j = j' - j$, and $\Delta m_j = m'_j - m_j$. As with the Schrödinger model, the non-vanishing terms of the components of the dipole moment between stationary states provide the selection rules for transitions. From examination of Eq. 2.29, and the results presented in Table 2.1, we immediately see that the dipole moment vanishes for all transitions except for $\Delta l = \pm 1$ and $\Delta m_l = 0, \pm 1$. Since $j = l \pm 1/2$, it also immediately follows that $\Delta j = \pm 1$ is a selection rule. Full evaluation of the fine structure dipole moment also yields non-vanishing terms for $\Delta j = 0$, and for $\Delta m_j = 0, \pm 1$.

An additional rule of transitions is that there must be a parity change in the stationary states. We defer further discussion of parity and its role in atomic transitions until § 3.1.1. Suffice it to say for now that, for hydrogenic atoms and ions, when $\Delta l = \pm 1$ or when there is a spin flip, the parity changes. Since $\Delta l = \pm 1$ is required by the dipole selection rules, it follows that $m'_s = m_s$ (no spin flip for dipole transitions). We summarize the fine-structure dipole transition rules in Table 2.3.

Table 2.3 *Fine-Structure Dipole Selection Rules*

Transition	Rule
Δl	± 1
Δj	$0, \pm 1$
Δm_j	$0, \pm 1$
Δm_s	0

2.3.1 Oscillator Strengths for Multiplets

Here, we show that the oscillator strengths of the individual transitions comprising a fine-structure multiplet can be determined using simple sum rules. The sum rule given by Eq. 2.29 yields,

$$f_{nl}^{n'l'} = \sum_{j'} f_{nlj}^{n'l'j'}. \quad (2.30)$$

Thus, each fine-structure oscillator strength is fraction of the total oscillator strength and their sum conserves the average oscillator strength for a transition between states nl and $n'l'$ in the Schrödinger model. The fractional contribution each multiplet transition contributes to the total is derived by considering the relative statistical weights of each fine-structure transition in the multiplet.

The *s-p* Doublets

From Bethe & Salpeter (1957), the fine-structure oscillator strengths for s-p doublets are

$$\begin{aligned} n's_{1/2} - np_{3/2} \quad f_{n1\frac{3}{2}}^{n'0\frac{1}{2}} &= \frac{2}{3} f_{n1}^{n'0} \\ n's_{1/2} - np_{1/2} \quad f_{n1\frac{1}{2}}^{n'0\frac{1}{2}} &= \frac{1}{3} f_{n1}^{n'0}, \end{aligned} \quad (2.31)$$

where $f_{n1}^{n'0}$ is computed from Eq. 2.16. We immediately see the universal rule of s-p doublets when the upper state is the p state:

$$f_{n1\frac{3}{2}}^{n'0\frac{1}{2}} = 2f_{n1\frac{1}{2}}^{n'0\frac{1}{2}}, \quad (2.32)$$

which shows that the $n's_{1/2} - np_{3/2}$ member of the doublet has twice the transition strength as the $n's_{1/2} - np_{1/2}$ member. Since the transition energy from the $p_{3/2}$ state is greater than that of the $p_{1/2}$ state, we also see that

the shorter wavelength member of the doublet observed in a spectrum will have twice the strength as the longer wavelength member. This condition is independent of n and n' . Thus, for example, all Lyman series lines are s-p doublets, so Ly α , Ly β , Ly γ , etc., all exhibit 2:1 line ratios with the shorter wavelength member ($1s_{1/2} - np_{3/2}$, $\Delta l = -1$, $\Delta j = -1$) being twice that of the longer wavelength member ($1s_{1/2} - np_{1/2}$, $\Delta l = -1$, $\Delta j = 0$).

For s-p doublets in which the upper state is the s state ($l = 0$) and the lower state is the p state ($l' = l + 1 = 1$), the results are different due to the relative statistical weights; the oscillator strengths are equal for both the $\Delta j = 0$ and the $\Delta j = +1$ fine-structure transitions.

The p-d and d-f Triplets

For p-d and d-f transitions there are three components to the multiplet (called a triplet). The p-d transitions are $d_{5/2} - p_{3/2}$, $d_{3/2} - p_{3/2}$, and $d_{3/2} - p_{1/2}$, whereas the d-f transitions are $f_{7/2} - d_{5/2}$, $f_{5/2} - d_{5/2}$, and $f_{5/2} - d_{3/2}$. Note that, under the dipole approximation, the p-d transition $d_{5/2} - p_{1/2}$ and the d-f transition $f_{7/2} - d_{3/2}$ do not occur, since they would require $\Delta j = 2$, which is not allowed by the selection rules.

For p-d transitions, we obtain,

$$\begin{aligned} n'p_{3/2} - nd_{5/2} \quad f_{n2\frac{5}{2}}^{n'1\frac{3}{2}} &= \frac{9}{15} f_{n2}^{n'1} \\ n'p_{3/2} - nd_{3/2} \quad f_{n2\frac{3}{2}}^{n'1\frac{3}{2}} &= \frac{1}{15} f_{n2}^{n'1} \\ n'p_{1/2} - nd_{3/2} \quad f_{n2\frac{3}{2}}^{n'1\frac{1}{2}} &= \frac{5}{15} f_{n2}^{n'1}, \end{aligned} \quad (2.33)$$

where $f_{n2}^{n'1}$ is computed from Eq. 2.16 using $l = 2$ and $l' = 1$. Note that the transition strengths for the s-p triplet have a 9:1:5 proportionality for its member transitions.

For d-f triplets, we obtain

$$\begin{aligned} n'f_{7/2} - nd_{5/2} \quad f_{n3\frac{7}{2}}^{n'2\frac{5}{2}} &= \frac{20}{35} f_{n3}^{n'2} \\ n'f_{5/2} - nd_{5/2} \quad f_{n3\frac{5}{2}}^{n'2\frac{5}{2}} &= \frac{1}{35} f_{n3}^{n'2} \\ n'f_{5/2} - nd_{3/2} \quad f_{n3\frac{5}{2}}^{n'2\frac{3}{2}} &= \frac{14}{35} f_{n3}^{n'2}, \end{aligned} \quad (2.34)$$

where $f_{n3}^{n'2}$ is computed from Eq. 2.16 using $l = 3$ and $l' = 2$. Note that the transition strengths for the d-f triplet have a 20:1:14 proportionality for its member transitions.

Remarkably, the ratios of the oscillator strengths for both the s–p and d–f triplets are independent of $\Delta n = n - n'$. However, the absolute magnitudes of the strengths do depend upon Δn .

2.3.2 The Spectrum

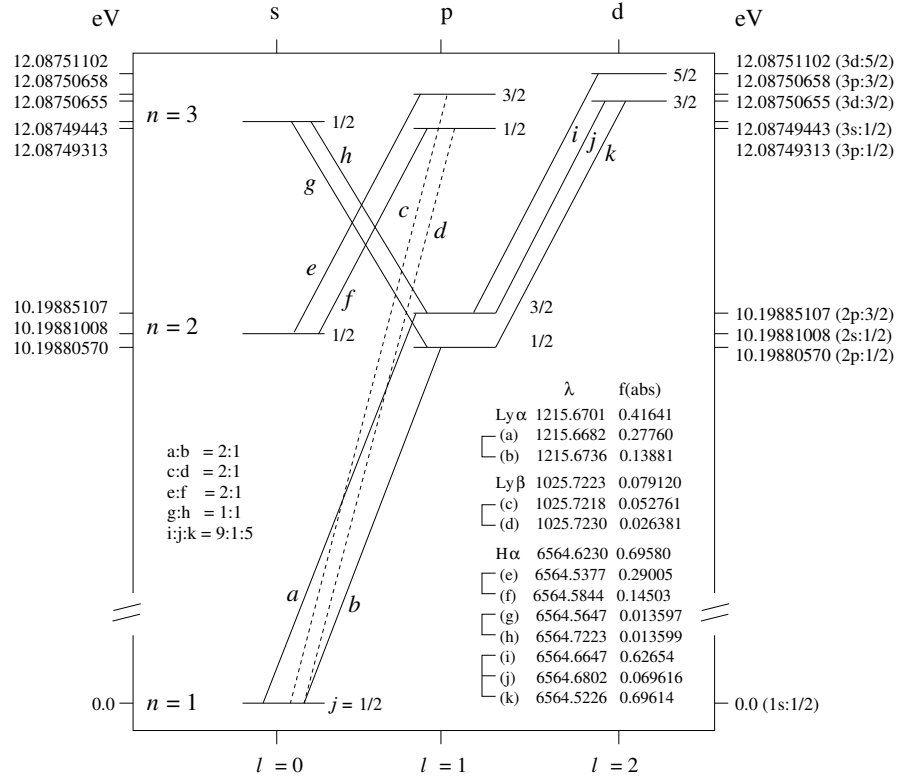


Figure 2.2 Grotrian diagram of hydrogen that incorporates fine structure and radiative corrections. A schematic of the Ly α doublet (a,b), Ly β doublet (c,d), and the seven-fold H α multiplet (e–k). Note that the energy degeneracy of the nj states is broken by the radiative corrections so that none of the transitions is energy degenerate. The inset on the right lists the transition wavelengths and absorption oscillator strengths. The inset on the left provides the relative multiplet strengths.

Taking into account fine structure and radiative corrections, we present, in Figure 2.2, the Grotrian diagram for hydrogen for transitions between levels $n = 1, 2$, and 3 . The excitation energies of the states are given in eV on the vertical axes.

Note that states of the same nj are no longer energy degenerate due to radiative corrections. Thus, each transition in a multiplet will give rise to a

unique spectral line. The transitions are indexed and identified as the Ly α doublets (a,b), the Ly β doublet (c,d), and the H α multiplet (e–k), which comprises two doublets (e,f and g,h) and a triplet (i,j,k). Also listed are the vacuum transition wavelengths and the absorption oscillator strengths. The average weighted wavelength of the transitions are also given, as are their total absorption oscillator strengths. These data are taken from Kramida (2010).

In Figure 2.3, we present the spectra for hydrogen for the Ly α , Ly β , and H α transitions. The weighted average of each multiplet is shown as a dotted line, whereas the multiplet components are solid lines. The strength of the transitions is given in terms of the absorption oscillator strengths. Since the Ly α (a,b) and Ly β (c,d) doublets are s–p transitions with $\Delta l = l - l' = -1$, the ratio of their oscillator strengths is $a:b = c:d = 2:1$. The H α doublet g,h is also s–p transitions, but with $\Delta l = -1$, so the ratio of their oscillator strengths is $g:h = 1:1$. The H α triplet (i,j,k) is a p–d transition with $\Delta l = -1$ with oscillator strengths in the proportion $i:j:k = 9:1:5$.

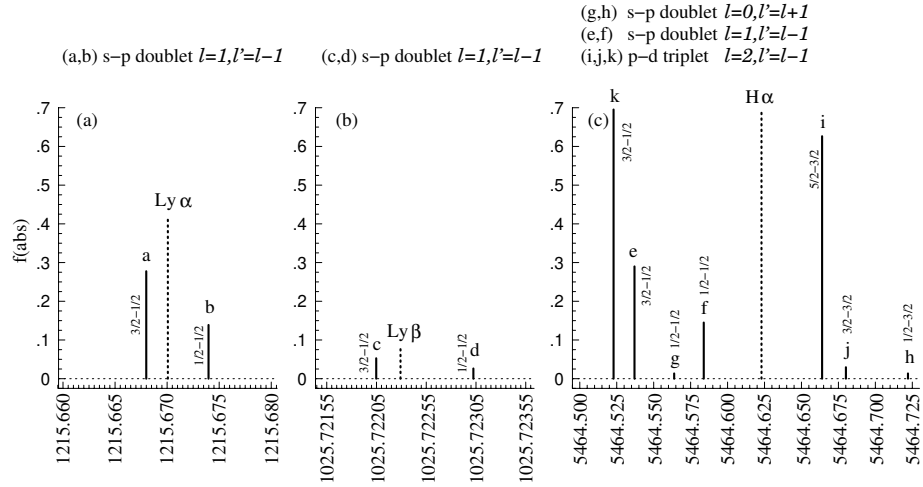


Figure 2.3 Selected transitions of the hydrogen spectrum incorporating fine structure and radiative corrections. (a) The Ly α doublet, labeled a,b according to Figure 2.2. (b) The Ly β doublet, labeled c,d. (c) The seven-fold H α multiplet, labeled e–k. The weighted average transition is shown as the dotted line. The height of each line is the absorption oscillator strength. The initial and final j states of each transition are also marked next to each line.

2.4 Bound-Free Oscillator Strengths

The bound-free oscillator strength quantifies the absorption strength for ionization, which can occur from any bound state nl as long as the incident photon has energy greater than the ionization energy for that state, i.e., $h\nu > E_n^i$, where $E_n^i = R_z/n^2$. Angular momentum between the photon, ejected electron, and the host atom must be conserved and since this is a three-body interaction, the dipole selection rules do not apply (the atom can change angular momentum during recoil). We remind the reader that the energy relation (see Eq. 1.47) is

$$h\nu = R_z \left[\frac{1}{n^2} + \frac{1}{k^2} \right], \quad (2.35)$$

for a transition from bound state n to unbound state ik , where $ik = (R_z/E)^{1/2}$ is the relationship between the kinetic energy of the liberated electron and the quantum number k .

Using the methods employed in § 2.2, the average oscillator strength for bound-free absorption can be derived. Following evaluation of the average dipole moment (Eq. 2.29), summing it over all m_l states and evaluating the overlap integral of the bound state and the continuum,

$$R_{nl}^{kl'} = \int_0^\infty R_{kl'} R_{nl} r^3 dr, \quad (2.36)$$

where the continuum wave function can be expressed (Green et al., 1958),

$$R_{kl}(\rho) = \frac{M_{kl}(z)}{2} \left[\frac{\exp\{i\pi k/2\} \Gamma(l+1-k)}{\Gamma(2l+2)} \exp\{-(i\pi/2)(l+1)\} \right], \quad (2.37)$$

where l' is the angular momentum of the unbound electron. For continuum states it is convenient to adopt atomic units, so that the radial coordinate and the energy are written $\rho = r/\hat{a}_0$ with $\epsilon = E/R_z$ where $\hat{a}_0 = (m_e/\hat{m})a_0$ is the reduced Bohr radius (Eq. 1.4). Thus, $\rho = 1$ is a Bohr radius, and $\epsilon = 1$ is $\simeq 1$ Ryd, or $\simeq 13.6$ eV. Note that $ik = 1/\epsilon^{1/2}$. The function $\Gamma(x)$ is an Euler's integral of the second kind called the Gamma function,

$$\Gamma(x) = \int_0^\infty y^{x-1} \exp\{-y\} dy, \quad (2.38)$$

which can be evaluated using one of several expansion series (cf. Abramowitz & Stegun, 1972), and the function $M_{kl}(z)$ is

$$M_{kl}(z) = z^{l+1} \exp\{-z/2\} \times {}_1F_1(l+1-k, 2l+2, z) \quad (2.39)$$

with

$$z = 2i\epsilon^{1/2}\rho = \frac{2\rho}{k}, \quad (2.40)$$

and where

$${}_1F_1(\alpha, \beta, x) = \frac{\Gamma(\beta)}{\Gamma(\beta-\alpha)} \int_0^1 \exp\{xy\} y^{\alpha-1} (1-y)^{\beta-\alpha-1} dy, \quad (2.41)$$

is the confluent hypergeometric function, which can be evaluated using the series (cf. Abramowitz & Stegun, 1972)

$${}_1F_1(\alpha, \beta, x) = \sum_{j=0}^{\infty} \frac{(\alpha)_j}{(\beta)_j} \frac{x^j}{j!} = 1 + \frac{\alpha}{\beta} x + \frac{\alpha(\alpha+1)}{\beta(\beta+1)} \frac{x^2}{2!} + \dots \quad (2.42)$$

where the Pochhammer symbol, $(y)_j$, is defined in Eq. 2.25.

Performing the overlap integral and averaging over l states, we obtain the average strength for all free transitions from bound state n (Menzel & Pekeris, 1935),

$$f_n^k = \frac{2^6}{3g_n} \frac{M_n(ik)}{k^3 n^3} \frac{\exp\{-4k \tan^{-1}(n/k)\}}{1 - \exp\{-2\pi k\}} \left[\frac{1}{n^2} + \frac{1}{k^2} \right]^{-7/2}, \quad (2.43)$$

where we have invoked $|(ik)^3| = k^3$, and where $g_n = 2n^2$ is the statistical weight of bound state n , and where

$$M_n(ik) = F^2(\alpha_k, \beta_k, \gamma, x) - F^2(\alpha_n, \beta_n, \gamma, x), \quad (2.44)$$

where F is the generalized hypergeometric function (see Eqs. 2.24 and 2.23), with $x = -4ikn/(ik-n)^2$, $\gamma = 1$, $\alpha_k = 1-ik$, $\alpha_n = 1-n$, $\beta_k = -ik$, and $\beta_n = -n$. For these parameters, only even powers of ik are present in the evaluation of $M_n(ik)$, so the function is real and can be written as a series of polynomials

$$M_n(ik) = \frac{8nk}{(n^2+k^2)^{2n-3/2}} P_n(k), \quad (2.45)$$

where the polynomials $P_n(ik)$ are tabulated by (Menzel & Pekeris, 1935). For example, for bound-free absorption from the $n = 1, 2$, and 3 levels, we have

$$M_1(ik) = \frac{8k}{(1+k^2)^{1/2}} \quad (2.46)$$

$$M_2(ik) = \frac{16k}{(k^2+4)^{5/2}} (3k^2+4)(5k^2+4) \quad (2.47)$$

$$M_3(ik) = \frac{24k}{(k^2+9)^{9/2}} (13k^4+78k^2+81)(29k^4+126k^2+81). \quad (2.48)$$

The polynomials become more complex as n increases.

Substituting Eq. 2.45 into Eq. 2.43, the resulting expression for the bound-free oscillator strength is

$$f_n^k = \frac{2^8}{3} \frac{n^3 k^5 P_n(k)}{(n^2 + k^2)^{2n+2}} \frac{\exp\{-4k \tan^{-1}(n/k)\}}{1 - \exp\{-2\pi k\}}. \quad (2.49)$$

From the transition energy (Eq. 2.35), we see that k can be written

$$k = \left[\frac{R_z}{h\nu - R_z/n^2} \right]^{-1/2}, \quad (2.50)$$

so that f_n^k can ultimately be expressed in terms of the incident photon energy.

As a final note we emphasize that f_n^k , as presented, is averaged over all l and m states for level n . Also note that the Z dependence in Eq. 2.49 is related to k via Eq. 2.50.

2.5 Natural Broadening

The lifetime that an atom remains in an excited state is finite. As mentioned in §2.1.2, the e -folding time for a spontaneous emission transition is $\Delta t_{nl}^{n'l'} = 2/A_{nl}^{n'l'}$ (see Eq. 2.1.2), where $A_{nl}^{n'l'}$ is the average Einstein A coefficient, which is on the order of 10^8 s^{-1} . According to the Heisenberg uncertainty principle, the energy of a given state is uncertain by an amount

$$\Delta E_{nl} = \frac{\hbar}{\Delta t_{nl}^{n'l'}} = \frac{A_{nl}^{n'l'} \hbar}{2}. \quad (2.51)$$

This implies that all nl states have an intrinsic energy uncertainty. The distribution in the uncertainty is such that it has a width on the order of ΔE_{nl} , as illustrated in Figure 2.4.

The fact that each energy level has an uncertainty distribution of width ΔE_{nl} means that transitions will not necessarily have the exact transition energy $E_n^{n'}$. This yields a distribution of transition frequencies having width $\Delta\nu = \Delta E_{nl}^{n'l'}/h$ that results from the convolution of the energy distributions of the two states nl and $n'l'$. For $\Delta E_{nl}^{n'l'}/h \simeq A_{nl}^{n'l'}$, we have $\Delta\nu \simeq 10^8 \text{ Hz}$. For ultraviolet and optical frequencies this yields a fractional width of emission lines of $\Delta\nu/\nu \simeq 10^{-7}$. Employing $\lambda = c/\nu$ and carrying out the same argument, we have $\Delta\lambda \simeq 10^{-4} \text{ \AA}$, and for ultraviolet and optical wavelengths, $\Delta\lambda/\lambda \simeq 10^{-7}$.

Though the width of the natural broadening is very narrow, the broadening is not inconsequential. The full-width half-maximum is narrow indeed,

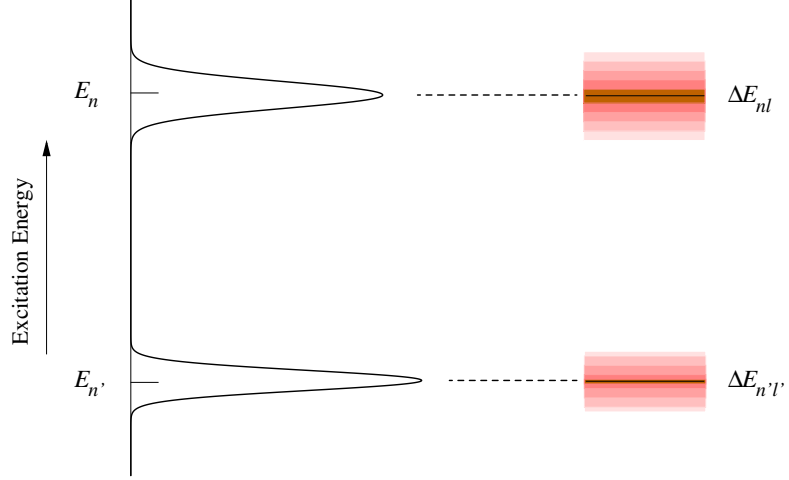


Figure 2.4 A schematic energy level diagram illustrating the relative spreads in the energies of the upper and lower levels of a possible transition $nl \leftrightarrow n'l'$. Because of the spread in the level energies the transition energy will not necessarily be exactly equal to $E_n^{n'}$. Note that since different nl states have different transition probabilities, so each state has a unique ΔE_{nl} .

but we will show that the wings of the distribution are highly extended, and this has important consequences when the number of atoms in the light path becomes large.

To determine the functional form of the line broadening, called the natural broadening function, we must consider the full treatment of the time-dependent probability of an atom in upper state nl that spontaneously decays to lower state $n'l'$. Consider the scenario in which both nl and $n'l'$ are excited states (that is, state nl can spontaneously decay to any allowed state mi where $m < n$ and $i = l \pm 1$, and state $n'l'$ could spontaneously decay to any allowed state mi , where $m < n'$ and $i = l' \pm 1$). An example of this scenario for a 5p–4s transition is illustrated in Figure 2.5, where the principle levels are indexed by m and the angular momentum quantum numbers are indexed by i . The possible decay channels of the 5p state and the 4s states are shown.

The time-dependent probability of overlap of the two states for states nl and $n'l'$ is

$$P_{nl}^{n'l'}(\mathbf{r}, t) dV = \Psi_{n'l'}^*(\mathbf{r}, t) \Psi_{nl}(\mathbf{r}, t) dV, \quad (2.52)$$

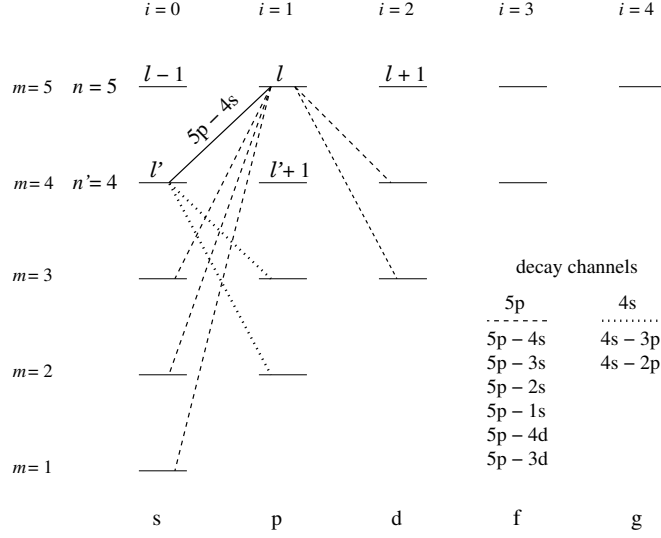


Figure 2.5 Schematic of the decay channels for the two states $nl = 5p$ and $n'l' = 4s$. The transition being considered is $5p-4s$. The decay channels are provided in the inset. For the purpose of evaluating the constants Γ_{nl} and $\Gamma_{n'l'}$, we index the principle levels by m and the angular momentum quantum states by i . For the example shown here, Eq. 2.55 is evaluated by taking the sum of the Einstein A coefficients over all decay channels from the $5p$ state and Eq. 2.55 is evaluated taking the sum over all decay channels for the $4s$ state.

where the time-dependent wave functions of states nl and $n'l'$ are

$$\begin{aligned}\Psi_{nl}(\mathbf{r}, t) &= \psi_{nl}(\mathbf{r}) \exp \left\{ - (i/\hbar) E_n t \right\} \exp \left\{ - (\Gamma_{nl}/2) t \right\} \\ \Psi_{n'l'}(\mathbf{r}, t) &= \psi_{n'l'}(\mathbf{r}) \exp \left\{ - (i/\hbar) E_{n'} t \right\} \exp \left\{ - (\Gamma_{n'l'}/2) t \right\},\end{aligned}\quad (2.53)$$

where we define $2/\Gamma_{nl}$ and $2/\Gamma_{n'l'}$ as the e -folding times for states nl and $n'l'$ to spontaneously transition via all available decay channels, respectively.

To compute Γ_{nl} , consider the decay channels from upper state nl . The time-dependent probability for decay to a given state mi is proportional to $\exp \{ - (A_{nl}^{mi}/2) t \}$. To account for all possible decay channels, we employ the joint probability,

$$\exp \{ - (\Gamma_{nl}/2) t \} = \prod_m^{n-1} \exp \left\{ - (A_{nl}^{ml-1}/2) t \right\} \prod_m^{n-1} \exp \left\{ - (A_{nl}^{ml+1}/2) t \right\}, \quad (2.54)$$

where $m = 1$ is the ground state, and where only terms with $l - 1 \geq 0$ are to be included. This product is equivalent to summing the rates over all states $m < n$ and $i = l \pm 1$, which gives the total spontaneous decay rate for upper

state nl ,

$$\Gamma_{nl} = \sum_m^{n-1} A_{nl}^{ml-1} + \sum_m^{n-1} A_{nl}^{ml+1}, \quad (2.55)$$

which yields $\Psi_{nl}(\mathbf{r}, t) \propto \exp\{-(\Gamma_{nl}/2)t\}$ and recovers Eq. 2.53. Note again that terms are retained only if $l-1 \geq 0$. Similar arguments apply for spontaneous decay of lower state $n'l'$ to states $m < n'$,

$$\Gamma_{n'l'} = \sum_m^{n'-1} A_{n'l'}^{ml'-1} + \sum_m^{n'-1} A_{n'l'}^{ml'+1}. \quad (2.56)$$

We now derive the line broadening function. Substituting Eq. 2.53 into Eq. 2.52, and invoking $E_n^{n'} = E_n - E_{n'}$, we have

$$P_{nl}^{n'l'}(\mathbf{r}, t) = \psi_{n'l'}^*(\mathbf{r})\psi_{nl}(\mathbf{r}) \exp\left\{-(i/\hbar) E_n^{n'} t\right\} \exp\left\{-(\Gamma_{nl}^{n'l'}/2)t\right\}, \quad (2.57)$$

where

$$\Gamma_{nl}^{n'l'} = \Gamma_{nl} + \Gamma_{n'l'}. \quad (2.58)$$

Eq. 2.57 indicates that the two states oscillate between one another with angular frequency $E_n^{n'}/\hbar$ and that this oscillation is damped with an e-folding time of $2/\Gamma_{nl}^{n'l'}$. We can separate Eq. 2.57 into a spatial part and a time-dependent part, where the time-dependent part is

$$f(t) = \exp\left\{-2\pi i \nu_n^{n'} t\right\} \exp\left\{-(\Gamma_{nl}^{n'l'}/2)t\right\}, \quad (2.59)$$

where we have invoked $E_n^{n'} = h\nu_n^{n'}$.

The inverse Fourier transform of $f(t)$ is a complex function, $F(\nu)$, that is the frequency power spectrum of $f(t)$. The natural line broadening function is the amplitude of the frequency power spectrum, i.e., $\phi_{nl}^{n'l'}(\nu) d\nu = F^*(\nu)F(\nu) d\nu$, which provides the relative number of transitions with transition energy $h\nu$ in the interval $\nu \rightarrow \nu + d\nu$ per unit frequency. It is thus normalized to unity,

$$\int_0^\infty \phi_{nl}^{n'l'}(\nu) d\nu = 1. \quad (2.60)$$

Note that the units of $\phi_{nl}^{n'l'}(\nu)$ is Hz^{-1} . Taking the transform,

$$F(\nu) = \frac{1}{2\pi} \int_0^\infty f(t) \exp\{2\pi i \nu t\} dt, \quad (2.61)$$

and substituting $f(t)$, we obtain

$$F(\nu) = C \int_0^\infty \exp\left\{-2\pi i (\nu_n^{n'} - \nu) t\right\} \exp\left\{-(\Gamma_{nl}^{n'l'}/2)t\right\} dt, \quad (2.62)$$

which yields

$$F(\nu) = \frac{C}{2\pi i (\nu_n^{n'} - \nu) + (\Gamma_{nl}^{n'l'}/2)}, \quad (2.63)$$

where we have introduced the constant C so that we can enforce the normalization condition given by Eq. 2.60. Applying $\phi_{nl}^{n'l'}(\nu) = F^*(\nu)F(\nu)$, we have

$$\begin{aligned} \phi_{nl}^{n'l'}(\nu) &= \frac{C^2}{4\pi^2 (\nu - \nu_n^{n'})^2 + (\Gamma_{nl}^{n'l'}/2)^2} \\ &= \frac{1}{\pi} \frac{C^2/4\pi}{(\nu - \nu_n^{n'})^2 + (\Gamma_{nl}^{n'l'}/4\pi)^2}. \end{aligned} \quad (2.64)$$

Eq. 2.64 has the functional form of the well-known Lorentzian

$$\mathcal{L}(x) = \frac{1}{\pi} \frac{y}{(x - x_0)^2 + y^2}, \quad \text{where} \quad \int_0^\infty \mathcal{L}(x) dx = 1, \quad (2.65)$$

and where x_0 is the location parameter, and $2y$ is the full-width at half maximum. The maximum is $\mathcal{L}(x_0) = 1/\pi y$.

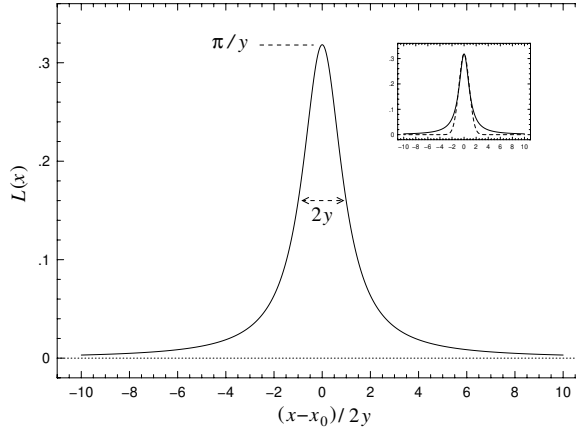


Figure 2.6 The Lorentzian distribution function given by Eq. 2.65 as a function of $\Delta x/2y$. The function has unity area, so that the amplitude at the peak is $1/\pi y$ when the full-width half-maximum is $2y$. (inset) The same Lorentzian (solid curve) with a Gaussian distribution (dashed curve) having the same full-width half-maximum and amplitude as the Lorentzian. Note that the wings of the Lorentzian are extended compared to the similar Gaussian distribution. For atomic transitions $y = \Gamma_{nl}^{n'l'}/4\pi$.

A Lorentzian centered at $\Delta x = x - x_0 = 0$ is illustrated in Figure 2.6 as a function of $\Delta x/2y$. An important characteristics of the Lorentzian distribution is that, in the regime $\Delta x \gg y$, the wings decay as Δx^{-2} , and are

therefore much more extended than those of a Gaussian distribution with the same full-width half-maximum and amplitude (see inset of Figure 2.6). Normalizing Eq. 2.64 according to Eq. 2.65, we find $C^2 = \Gamma_{nl}^{n'l'}$, yielding

$$\phi_{n'l'}^{nl}(\nu) = \frac{1}{\pi} \frac{(\Gamma_{nl}^{n'l'}/4\pi)}{(\nu - \nu_n^{n'})^2 + (\Gamma_{nl}^{n'l'}/4\pi)^2}. \quad (2.66)$$

Note that $\Gamma_{nl}^{n'l'}/2\pi$ is the full-width half-maximum of the natural broadening function. The quantity $\Gamma_{nl}^{n'l'}$ is known as the damping constant for the transition nl to $n'l'$.

The above derivation of the natural broadening function was applied for spontaneous emission. In an intense radiation field, the e -folding times of each state can be shortened due to stimulated emission and/or absorption. Using n to generically index a given nl state and m to index all principle levels with allowed Δl , the full damping constant is obtained by the joint probability of all exit channels out of a state,

$$\Gamma_n = \sum_{m=1}^{n-1} A_n^m + 4\pi \sum_{m=1}^{n-1} B_n^m \mathcal{I}_\nu + 4\pi \sum_{m=n+1}^{\infty} B_m^n \mathcal{I}_\nu, \quad (2.67)$$

where \mathcal{I}_ν is the specific intensity of incident radiation. We see that both stimulated emission and absorption can induce an electron to change states more rapidly than spontaneous decay would naturally occur. However, even in the stellar atmospheres of most stars, the stimulated emission and absorption rates are on the order of 1–0.1% of the spontaneous decay rates. In an A star ($T = 10,000$ K), where the spectral energy distribution peaks at $\simeq 10^{15}$ Hz ($\simeq 3000$ Å) with $\mathcal{I}_\nu \simeq 10^{-4}$ erg s $^{-1}$ cm $^{-2}$ Hz $^{-1}$ str $^{-1}$, the Einstein relations (Eq. 2.9) give $B_n^{n'} \simeq B_n^n \simeq 6A_n^{n'}$, so that $4\pi B_n^{n'} \mathcal{I}_\nu \simeq 4\pi B_n^n \mathcal{I}_\nu \simeq 7 \times 10^{-3} A_n^{n'}$. Thus, the expression given by Eqs. 2.55 and 2.56 suffice for computing $\Gamma_n^{n'}$ and $\Gamma_n^{n'}$ for most all applications (and this is especially true for the conditions in the interstellar, circumgalactic, and intergalactic media).

In the many astrophysical environments where stimulated emission and absorption rates are negligible compared to spontaneous emission rates, we see from Eq. 2.58 and Eqs. 2.55 and 2.56 that the damping constants for both the emission and absorption transitions are symmetric, i.e., $\Gamma_{n'l'}^{nl} = \Gamma_{nl}^{n'l'}$. And since, the emission and absorption frequencies are identical ($\nu_n^n = \nu_n^{n'}$), we see that the natural broadening function for absorption is identical to the natural broadening function for emission,

$$\phi_{nl}^{n'l'}(\nu) d\nu = \phi_{n'l}^{nl}(\nu) d\nu. \quad (2.68)$$

We re-emphasize that the units of the broadening function are Hz $^{-1}$.

2.6 Emission Power and Absorption Cross Section

Having developed the formalism of Einstein coefficients, oscillator strengths, and damping constants, we can now apply these bits and pieces to construct expressions for the emission and absorption strengths in terms of the energy rate (power) of light either released or absorbed by a single hydrogenic atom. These quantities are key to solving the radiative transfer equation (see Chapter 7), from which spectral features such as absorption lines can be physically connected back to the properties of the gas and atoms/ions that interact with the observed electromagnetic radiation. The quantities we derive in this section will be central to our analysis of absorption line data.

2.6.1 Bound-Bound Emission

In order to express the emission power for bound-bound transitions, we begin with the rates for spontaneous emission. For the Schrödinger model, we obtained the average spontaneous transition rate,

$$A_{nl}^{n'l'} = \frac{32\pi^3 e^2 (\nu_n^{n'})^3}{3\hbar c^3} |\mathbf{r}_{nl}^{n'l'}|^2, \quad (2.69)$$

from Eq. 2.4 after summing over all m states, where $\mathbf{r}_{nl}^{n'l'}$ is given by Eq. 2.15. Substitution yields

$$\begin{aligned} A_{nl}^{n'l'} &= \frac{32\pi^3 e^2}{3\hbar c^3} (\nu_n^{n'})^3 \frac{\max(l, l')}{2l+1} \left(R_{nl}^{n'l'}\right)^2 \\ &= \frac{1}{6} \frac{\hat{m} e^{10}}{\hbar^6 c^3} \frac{\max(l, l')}{2l+1} \frac{(n+n')^3 (n-n')^3}{n^6 n'^6} \left(\frac{R_{nl}^{n'l'}}{\hat{a}_0}\right)^2, \end{aligned} \quad (2.70)$$

where we have invoked $\nu_n^{n'} = E_n^{n'}/h$, $E_n^{n'} = R_z(1/n^2 - 1/n'^2)$, $R_z = e^2/2\hat{a}_0$, and $\hat{a}_0 = \hbar^2/\hat{m}e^2$, which follow from Eqs. 1.10, 1.6, and 1.4, respectively. The constants in Eq. 2.70 evaluate to $A_0 = 2.6775(\hat{m}/m_e) \times 10^9 \text{ s}^{-1}$.

For the Lyman series, we invoke Eq. 2.26, and obtain

$$A_{n1}^{10} = A_0 \frac{2^8 n(n-1)^{2n-2}}{3(n+1)^{2n+2}}. \quad (2.71)$$

For the hydrogen Ly α transition ($n = 2$), $A_{21}^{10} = (2^9/3^7)A_0 = 6.2649 \times 10^8 \text{ s}^{-1}$, which is in excellent agreement with the experimentally determined value $6.2650 \times 10^8 \text{ s}^{-1}$ (Kramida, 2010). A similar expression can be obtained for the Balmer series average transition rates from Eq. 2.70 using the overlap integrals provided in Table 2.2.

We define the line intensity for spontaneous emission as $j_{nl}^{n'l'}(\nu)$, which has

units $\text{erg s}^{-1} \text{ str}^{-1} \text{ Hz}^{-1} \text{ atom}^{-1}$; it is the energy per unit time (power) per unit frequency emitted into a solid angle $d\Omega$ per atom. The power emitted is the rate times the emitted energy, $h\nu A_{nl}^{n'l'}$. Normalizing by all solid angle and accounting for the natural broadening function, $\phi_{nl}^{n'l'}(\nu)$ [Hz^{-1}], we obtain the power per unit frequency per steradian from a single atom,

$$j_{nl}^{n'l'}(\nu) d\nu = \frac{h\nu}{4\pi} A_{nl}^{n'l'} \phi_{nl}^{n'l'}(\nu) d\nu, \quad (2.72)$$

in the frequency interval $\nu \rightarrow \nu + d\nu$. Inserting Eq. 2.66 for $\phi_{nl}^{n'l'}(\nu)$, we obtain

$$j_{nl}^{n'l'}(\nu) d\nu = \frac{h\nu}{4\pi} A_{nl}^{n'l'} \cdot \frac{1}{\pi} \frac{(\Gamma_n'/4\pi)}{(\nu - \nu_n')^2 + (\Gamma_n'/4\pi)^2} d\nu. \quad (2.73)$$

Writing Eq. 2.73 in terms of the oscillator strength (see Eq. 2.14), we have

$$j_{nl}^{n'l'}(\nu) d\nu = h\nu \frac{2\pi e^2}{\hat{m}c^3} (\nu_n')^2 f_{nl}^{n'l'} \cdot \frac{1}{\pi} \frac{(\Gamma_n'/4\pi)}{(\nu - \nu_n')^2 + (\Gamma_n'/4\pi)^2} d\nu. \quad (2.74)$$

To obtain the integrated emission power [$\text{erg s}^{-1} \text{ str}^{-1} \text{ atom}^{-1}$], Eq. 2.73 is integrated over frequency. Since the natural broadening function is extremely narrow ($\Delta\nu/\nu \sim 10^{-7}$), to an excellent approximation $\nu d\nu = \nu_n' d\nu$, so that the frequency dependence can be pulled out of the integral, yielding the integrated emission power,

$$J_{nl}^{n'l'} = \int_0^\infty j_{nl}^{n'l'}(\nu) d\nu = \frac{h\nu_n'}{4\pi} A_{nl}^{n'l'} \int_0^\infty \phi_{nl}^{n'l'}(\nu) d\nu = \frac{h\nu_n'}{4\pi} A_{nl}^{n'l'}, \quad (2.75)$$

or

$$J_{nl}^{n'l'} = \frac{2\pi h e^2}{\hat{m}c^3} (\nu_n')^3 f_{nl}^{n'l'}, \quad (2.76)$$

where we have applied Eq. 2.60.

2.6.2 Bound-Bound Absorption

The absorption power per unit frequency per unit solid angle per atom, which we denote $k_{nl}^{n'l'}(\nu)$ [$\text{erg s}^{-1} \text{ str}^{-1} \text{ Hz}^{-1} \text{ atom}^{-1}$], is similarly obtained by multiplying the absorption rate, $4\pi B_{n'l'}^{nl} \mathcal{I}_\nu$ (given by Eq. 2.6), by the energy of the absorbed photon normalized by all solid angles, $h\nu/4\pi$, which gives

$$k_{n'l'}^{nl}(\nu) d\nu = h\nu B_{n'l'}^{nl} \mathcal{I}_\nu \phi_{n'l'}^{nl}(\nu) d\nu. \quad (2.77)$$

Note that we have written the natural broadening profile for absorption, but we showed in § 2.5 that emission and absorption profile shapes are equivalent,

$\phi_{n'l'}^{nl}(\nu) = \phi_{nl}^{n'l'}(\nu)$. Applying the Einstein relations (Eq. 2.9) relating $B_{n'l'}^{nl}$ and $A_{nl}^{n'l'}$, and then eliminating $A_{nl}^{n'l'}$ via Eq. 2.14, we obtain

$$k_{n'l'}^{nl}(\nu) d\nu = \frac{\pi e^2}{\hat{m}c} \frac{g_{nl}}{g_{n'l'}} f_{nl}^{n'l'} \mathcal{I}_\nu \phi_{n'l'}^{nl}(\nu) d\nu. \quad (2.78)$$

In § 2.2.1 (see Eq. 2.17), we showed that the average oscillator strength for spontaneous emission and the average oscillator strength for absorption obey $g_{nl} f_{nl}^{n'l'} = g_{n'l'} f_{n'l'}^{nl}$, which yields the absorption power,

$$k_{n'l'}^{nl}(\nu) d\nu = \frac{\pi e^2}{\hat{m}c} f_{n'l'}^{nl} \mathcal{I}_\nu \phi_{n'l'}^{nl}(\nu) d\nu, \quad (2.79)$$

in terms of the average oscillator strength for absorption.

It is convention to quantify absorption using the frequency dependent absorption cross section, which is defined as the absorption power per unit incident radiation,

$$\sigma_{n'l'}^{nl}(\nu) d\nu = \frac{k_{n'l'}^{nl}(\nu)}{\mathcal{I}_\nu} d\nu = \frac{\pi e^2}{\hat{m}c} f_{n'l'}^{nl} \phi_{n'l'}^{nl}(\nu) d\nu, \quad (2.80)$$

which has units $\text{cm}^2 \text{ Hz}^{-1} \text{ atom}^{-1}$. Inserting Eq. 2.66 for $\phi_{n'l'}^{nl}(\nu)$, we obtain

$$\sigma_{n'l'}^{nl}(\nu) = \frac{\pi e^2}{\hat{m}c} f_{n'l'}^{nl} \cdot \frac{1}{\pi} \frac{(\Gamma_n^{n'}/4\pi)}{(\nu - \nu_n^{n'})^2 + (\Gamma_n^{n'}/4\pi)^2}. \quad (2.81)$$

Recall that $\Gamma_n^{n'} = \Gamma_n^{n'}$. The integrated cross section [$\text{cm}^2 \text{ atom}^{-1}$] is

$$\sigma_{n'l'}^{nl} = \int_0^\infty \sigma_{n'l'}^{nl}(\nu) d\nu = \frac{\pi e^2}{\hat{m}c} f_{n'l'}^{nl}, \quad (2.82)$$

where we have applied Eq. 2.60. We interpret Eq. 2.82 as the power (energy per unit time) per atom removed from a beam of specific intensity \mathcal{I}_ν across the frequency range of the total absorption profile.

In terms of wavelength, which is more commonly used for the ultraviolet and optical bands, we employ the relationship $\sigma_{n'l'}^{nl}(\nu) |d\nu| = \sigma_{n'l'}^{nl}(\lambda) |d\lambda|$, which conserves energy between a unit frequency interval and a unit wavelength interval. From $c = \lambda\nu$, we have $|d\nu| = (c/\lambda^2) |d\lambda|$, yielding the absorption cross section per unit wavelength,

$$\sigma_{n'l'}^{nl}(\lambda) = \frac{\pi e^2}{\hat{m}c^2} (\lambda_{n'l'}^{nl})^2 f_{n'l'}^{nl} \cdot \frac{1}{\pi} \frac{\Gamma_n^{n'} (\lambda_n^{n'})^2 / 4\pi c}{(\lambda - \lambda_n^{n'})^2 + [\Gamma_n^{n'} (\lambda_n^{n'})^2 / 4\pi c]^2}, \quad (2.83)$$

where the units are $\text{cm}^2 \text{ \AA}^{-1} \text{ atom}^{-1}$. Integration yields

$$\sigma_{n'l'}^{nl} = \int_0^\infty \sigma_{n'l'}^{nl}(\lambda) d\lambda = \frac{\pi e^2}{\hat{m}c^2} (\lambda_n^{n'})^2 f_{n'l'}^{nl}. \quad (2.84)$$

We interpret Eq. 2.84 as the power (energy per unit time) per atom removed from a beam of specific intensity \mathcal{I}_λ across the wavelength range of the total absorption profile. The factor of $(\lambda_n')^2/c$ in Eq. 2.83 for the wavelength dependent cross section that does not appear in the frequency dependent cross section (Eq. 2.82) conserves the total energy across the absorption profile. It follows from $\mathcal{I}_\lambda |d\lambda| = \mathcal{I}_\nu |d\nu|$, which yields $\mathcal{I}_\lambda = (c/\lambda^2)\mathcal{I}_\nu$. Thus, the factor enforces equal power is removed from the beam per unit energy for $k_{n'l'}^{nl}(\nu)$ and $k_{n'l'}^{nl}(\lambda)$.

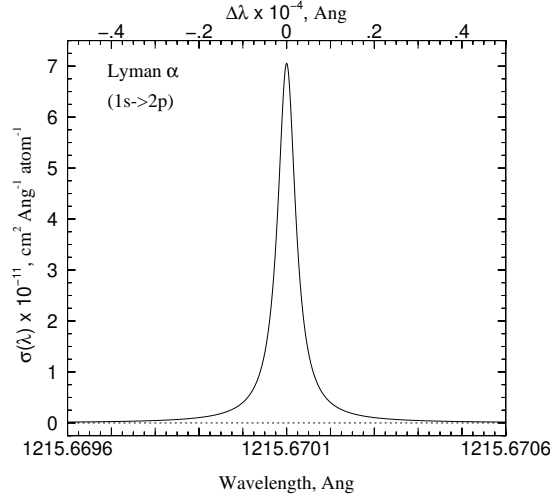


Figure 2.7 The Ly α absorption cross section, $\sigma_{21}^{10}(\lambda)$, for hydrogen. The oscillator strength is $f_{21}^{10} = 0.41641$ and the damping constant is $\Gamma_{21}^{10} = 6.265 \times 10^8 \text{ s}^{-1}$.

In Figure 2.7, we plot the Ly α cross section $\sigma_{21}^{10}(\lambda)$ for hydrogen computed using Eq. 2.83. The peak is $4\pi^2 e^2 f_{21}^{10} / \hat{m} c \Gamma_{21}^{10} = 7.055 \times 10^{-11} \text{ cm}^2 \text{ \AA}^{-1} \text{ atom}^{-1}$ and the full-width half-maximum is $\Gamma_1^2(\lambda_1^2)^2 / 2\pi c = 4.91 \times 10^{-5} \text{ \AA}$. Within the region of the full-width half-maximum, the cross section decreases as a Gaussian, $\exp\{-(\nu - \nu_n')^2\}$ or $\exp\{-(\lambda - \lambda_n')^2\}$, and in the wings the cross section decreases as the inverse square, $(\nu - \nu_n')^{-2}$ or $(\lambda - \lambda_n')^{-2}$.

2.6.3 Bound-Free Absorption

In § 1.6, we presented the wave functions and quantum numbers for continuum states and defined the quantum number of the free electron as $ik = (R_z/E)^{1/2}$, where E is the kinetic energy. For the continuum, the quantum number k is not required to be an integer because the photon frequency liberating the electron can take on any value in the interval $\nu \rightarrow \nu + d\nu$. For

an ionizing photon in frequency interval $\nu \rightarrow \nu + d\nu$ there is a finite number of continuum states for the electron between $k \rightarrow k + dk$.

Based upon the fact that there are dk continuum states into which the liberated electron can enter following ionization by a photon in the frequency interval $\nu \rightarrow \nu + d\nu$, Menzel (1930) showed that the absorption cross section per atom for absorption from bound state n to continuum state $k \rightarrow k + dk$ can be expressed

$$\sigma_n^k(\nu) = \frac{\pi e^2}{m_e c} \frac{df_n^k}{d\nu} = \frac{\pi e^2}{m_e c} f_n^k \left| \frac{dk}{d\nu} \right|, \quad (2.85)$$

where f_n^k is the average oscillator strength for absorption (as given by Eq. 2.49), and $|dk/d\nu|$ is the number of continuum states per frequency interval $d\nu$. From Eq. 2.50, we obtain,

$$\left| \frac{dk}{d\nu} \right| = \frac{hk^3}{2R_z}. \quad (2.86)$$

Substituting $R_z = Z^2(\hat{m}/m_e)R$ from Eq. 1.7 into Eq. 2.86 and then substituting Eq. 2.86 and Eq. 2.49 into Eq. 2.85, we obtain

$$\sigma_n^k(\nu) = \frac{A_0}{Z^2} \cdot \frac{n^3 k^8 P_n(k)}{(k^2 + n^2)^{2n+2}} \frac{\exp\{4 - 4k \tan^{-1}(n/k)\}}{1 - \exp\{-2\pi k\}}, \quad (2.87)$$

where

$$A_0 = \frac{2^7}{3} \frac{\pi e^2 h}{m_e c R} \frac{(m_e/\hat{m})}{\exp\{4\}} = 6.304 \times 10^{-18} \text{ cm}^2 \text{ atom}^{-1}. \quad (2.88)$$

Note that the factor of “ $\exp\{4\}$ ” has been multiplied through both the numerator and denominator of Eq. 2.87, and is absorbed into the constant A_0 . This step yields a form of the cross section consistent with that popularized by Osterbrock & Ferland (2006) as written by Hummer & Seaton (1963). The combination of physical constants are alternatively written

$$\frac{\pi e^2 h}{m_e c R} = 4\pi\alpha \cdot \pi a_0^2 = 0.28809 \pi a_0^2, \quad (2.89)$$

where α is the fine-structure constant and a_0 is the Bohr radius. Note that the latter term is the geometric “cross section” of the ground state orbit of neutral hydrogen. Thus, if desired, the bound-free absorption cross section can be evaluated in units of the hydrogen “cross section”, πa_0^2 or $\pi \hat{a}_0^2$.

A most predominant ultraviolet ionization feature in spectra is the so-called Lyman limit break (LLB), which arises from ionization at the Lyman series limit, the ground state $n_0 = 1$. We invoke Eq. 2.50 with $n = 1$, which gives $k = (\nu/\nu_0 - 1)^{-1/2}$, where ν_0 is the frequency of an incident

photon corresponding to the ionization threshold energy $E_n^i = R_z = h\nu_0$. The second term in Eq. 2.87 can now be expressed in terms of the incident photon frequency

$$\frac{k^8 |P_1(k)|}{(k^2 + 1)^4} = \frac{(k^2)^4}{(k^2 + 1)^4} = \left(\frac{\nu_0}{\nu}\right)^4, \quad (2.90)$$

where $P_1(k) = 1$ (Eq. 2.46). For hydrogen, $Z = 1$, and we obtain

$$\sigma_1^k(\nu) = A_0 \left(\frac{\nu_0}{\nu}\right)^4 \frac{\exp \left\{ 4 - \left[4/\sqrt{\nu/\nu_0 - 1} \right] \tan^{-1} \left(\sqrt{\nu/\nu_0 - 1} \right) \right\}}{1 - \exp \left\{ -2\pi/\sqrt{\nu/\nu_0 - 1} \right\}}. \quad (2.91)$$

The above expressions can be written in terms of wavelength quite simply by substituting $\nu/\nu_0 = \lambda_0/\lambda$.

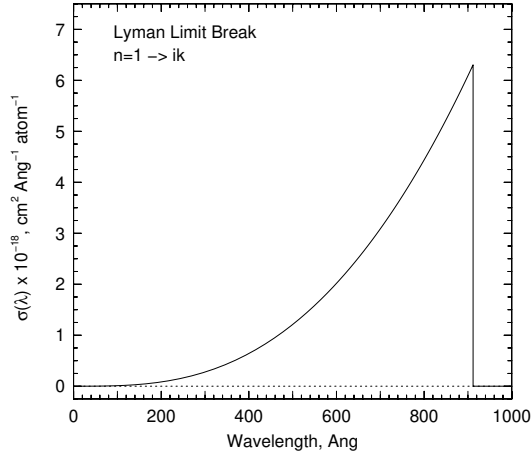


Figure 2.8 The Ly α bound-free absorption cross section, $\sigma_1^k(\lambda)$, for the Lyman limit threshold of hydrogen, $n \rightarrow ik$. The quantum number for the continuum states is defined as $ik = (R_z/E)^{1/2}$, where $E = \frac{1}{2}m_e v^2$ is the kinetic energy of the liberated electron. The energy balance with the ionizing photon yields $k = [R_z/(h\nu - R_z/n^2)]^{-1/2}$.

In general, for hydrogenic atoms with $Z \geq 1$ the bound-free cross section thresholds for level n are

$$\begin{aligned} \nu_n &= \nu_0 \frac{Z^2}{n^2} \frac{\hat{m}}{m_e} = 3.28984 \frac{Z^2}{n^2} \frac{\hat{m}}{m_e} \times 10^{15} \text{ Hz} \\ \lambda_n &= \lambda_0 \frac{n^2}{Z^2} \frac{m_e}{\hat{m}} = 911.267 \frac{n^2}{Z^2} \frac{m_e}{\hat{m}} \text{ \AA}, \end{aligned} \quad (2.92)$$

where Eq. 2.87 would need to be evaluated in full for general n to obtain the absorption cross section.

In Figure 2.8, we plot the Lyman limit cross section $\sigma_1^k(\lambda)$ for hydrogen computed using using Eq. 2.91. The peak is $A_0 = 6.304 \times 10^{-18} \text{ cm}^2 \text{ \AA}^{-1} \text{ atom}^{-1}$ at the ionization threshold wavelength 911.267 \AA . According to Eq. 2.91, near the ionization edge the cross section decreases as $(\nu/\nu_0)^{-4}$ or $(\lambda/\lambda_0)^{-4}$. Further from the edge the behavior of the cross section is well approximated as $(\nu/\nu_0)^{-3}$ or $(\lambda/\lambda_0)^{-3}$. This asymptotic behavior is often presented as the approximate expressions

$$\begin{aligned}\sigma_1^k(\nu) &\simeq 6.3 \times 10^{18} \left(\frac{\nu_0}{\nu}\right)^3 \text{ cm}^2 \\ \sigma_1^k(\lambda) &\simeq 6.3 \times 10^{18} \left(\frac{\lambda}{\lambda_0}\right)^3 \text{ cm}^2.\end{aligned}\tag{2.93}$$

3

Multi-Electron Atoms and Transitions

The spectra of multi-electron atoms and ions can be rich and complex, comprising singlets, doublets, triplets, and higher multiplets. Solving for the wave functions for atoms or ions with multiple electrons is a complicated business and involves approximations and perturbation theory. At minimum, the incorporated physics must include the kinetic energies of and Coulomb forces acting on and between all electrons, and the interaction between the angular momentum and spin of all electrons. In certain cases it is necessary to include spin-spin interactions between electrons. And of course, relativistic energies and effects, radiative corrections, and isotope shifts can be added as higher-order effects. Because the energy states depend upon the full configuration of electrons, the atom or ion takes on a total angular momentum, spin, and energy state.

We begin by discussing the “many-electron problem”, and then introduce the standard method of solution, known as the Hartree-Fock method. Applying step-by-step perturbation approximations, we illustrate how the energy structure depends on various interaction energies. In the process, we discuss the L-S coupling scheme, introduce the j - j coupling scheme, and briefly motivate the need for intermediate coupling schemes. Since bound-free spectral features are rarely observed in the cosmological spectra (with the exception of hydrogen and helium), we focus on the bound-bound transitions. We derive the emission and absorption strengths of bound-bound transitions, and describe the selection rules for transitions, including both the electric and magnetic dipole and the electric quadrupole. Finally, we discuss isotope shifts.

Throughout, we assume non-relativistic energies and an infinite-mass nucleus, with exceptions noted. Additional resources for the material covered in this chapter are Herzberg (1944), Cowan (1981), Bransden & Joachian

(2003), and Li & Blinder (2014). In particular, Herzberg provides excellent qualitative physical descriptions.

3.1 The Many-Electron Problem

Accounting for the Coulomb interactions only, the non-relativistic Hamiltonian is written,

$$H = \sum_{i=1}^{N_e} \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - \frac{Ze^2}{r_i} \right) + \sum_{i=1}^{N_e} \sum_{j=i+1}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|}, \quad (3.1)$$

where N_e is the number of bound electrons and \mathbf{r}_i is the radial vector of electron i with respect to the nucleus. The term $-Ze^2/r_i$ is the Coulomb potential of attraction, $V(r_i)$, between the nucleus of charge Ze and the i th electron. The sum of $e^2/|\mathbf{r}_i - \mathbf{r}_j|$ is the Coulomb repulsion between all individual electrons. The azimuthal and polar components to the wave function are spherical harmonics, so the problem reduces to solving for the radial wave function. The terms involving $|\mathbf{r}_i - \mathbf{r}_j|$ preclude simple separation of variables once the Hamiltonian is inserted into the Schrödinger equation.

The salient physics of multi-electron atoms can be illustrated using a step by step energy perturbation approach to capture higher-order interactions between electrons. We write the Hamiltonian as the sum of three primary interactions, i.e., $H = H_1 + H_2 + H_3$, where

$$\begin{aligned} H_1 &= \sum_{i=1}^{N_e} \left(-\frac{\hbar^2}{2m_e} \nabla_i^2 - V_c(r_i) \right), \\ H_2 &= \sum_{i=1}^{N_e} \sum_{j=i+1}^{N_e} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{N_e} \left(\frac{Ze}{r_i} - V_c(r_i) \right), \\ H_3 &= \frac{1}{2m_e^2 c^2} \sum_{i=1}^{N_e} \frac{1}{r_i} \frac{dV(r_i)}{dr_i} (\mathbf{L}_i \cdot \mathbf{S}_i). \end{aligned} \quad (3.2)$$

The first operator, H_1 , models the electrostatics as a symmetric all-angle averaged centrally concentrated potential, $V_c(r_i)$. This Hamiltonian is known as the central field approximation, which includes the angle-averaged Coulomb repulsion between electrons. The consequence of this simplification is that each electron senses a single spherically symmetric potential. The second operator, H_2 , treats spatial perturbations in $V_c(r_i)$, i.e., accounts for the residuals between $V_c(r_i)$ and the non-spherical Coulomb repulsion between electrons due to their non-isotropic spatial distributions [according to the

$Y_{lm}(\theta, \phi)]$. The third Hamiltonian, H_3 , is called the spin-orbit correction term. It accounts for interactions between orbital angular momentum and electron spin.

Invoking Eq. 3.2, the wave functions can then be solved in steps, first by solving the wave equation using $H = H_1$, then the sum $H = H_1 + H_2$, and then the sum $H = H_1 + H_2 + H_3$. This approach works well when the interaction energies (strengths) are $|E_1| > |E_2| > |E_3|$. However, as Z increases $|E_2| \sim |E_3|$, and then as Z increases further $|E_3| > |E_2|$. Thus, the order in which the perturbations are applied is reversed for large Z , while the intermediate case is more complex.

3.1.1 The Hartree-Fock Method

A highly effective approach to the problem was worked out in 1928 by Douglas Hartree. In short, the method is as follows: The full wave function for an atom is written as the product of wave functions for each electron, $\psi(\mathbf{r}_i)$. An initial guess for the wave functions for each electron, $\psi_{(0)}(\mathbf{r}_i)$, is made. From these initial guess wave functions, the electrostatic potential is self-consistently calculated from the individual charge distributions, $|\psi_{(0)}(\mathbf{r}_i)|^2$. To improve the estimate for the wave function of a given electron, say electron j , the potentials are summed over all i electrons (excluding j). This resulting potential is then inserted into the Hamiltonian and the Schrödinger equation is solved to obtain an updated guess for the wave function of electron j , i.e., $\psi_{(1)}(\mathbf{r}_j)$. This process is then repeated for all j . The initial guesses for all i electrons, $\psi_{(0)}(\mathbf{r}_i)$, will not be in agreement with the first iteration improvements, $\psi_{(1)}(\mathbf{r}_i)$. Thus, the process is repeated until the individual wave functions from the current and previous iterations are in agreement to some defined level of precision.

Note that the Hartree wave functions are not exact solutions, but numerical approximations. The Hartree wave functions have the same quantum numbers, $nlm_l m_s$, as the Dirac hydrogenic wave functions and their interpretation is effectively the same. The Hartree wave functions also account for the Pauli exclusion principle, which was stated by W. Pauli in 1925—“no electrons bound in an atom can have the same quantum numbers.”

Parity

The shortcoming of Hartree’s method was that it yielded wave functions that are all symmetric, i.e., $\psi(\mathbf{r}_i) = \psi(-\mathbf{r}_i)$, where as it is known that some atomic configurations are described by wave functions that are anti-

symmetric, i.e., $\psi(\mathbf{r}_i) = -\psi(-\mathbf{r}_i)$. Functions that are symmetric are said to have even parity and those that are antisymmetric have odd parity. In 1930, Vladimir Fock and John Slater worked out the methods that yielded properly symmetric and antisymmetric wave functions. The full treatment is now known as the Hartree-Fock method.

As we will see, the parity of the wave function is vitally important for understanding the atomic properties of multi-electron atoms and ions, and especially for understanding their spectra.

3.1.2 Coupling Schemes

The dominant electron-electron interaction physics cannot be described by a single unified model. There are a range of energetic interactions within the atom and their relative strengths depend on both the nuclear charge, Z , and the specific permutation of the electron configuration.

For example, close proximity electrons, especially those in low to moderate Z atoms, have their orbital angular momenta strongly coupled to one another and their spin strongly coupled to one another so that the atom *in toto* has a well-defined angular momentum, L , and spin state, S . The L and S weakly couple, yielding a well-defined total angular momentum state, J . This is called Russell-Saunders L-S coupling.

For electrons that are at larger average distances from one another, the orbital angular momentum of an individual electron is most strongly coupled to its spin (similar to the $j = l \pm s$ spin-orbit coupling for hydrogenic atoms). Thus, each electron has a well-defined j . The j of the electrons then weakly couple. This is called j - j coupling and it dominates for outer or excited electrons in higher Z atoms. An atom with pure j - j coupling has a well-defined J state, but not well-defined L and S states.

L-S coupling and j - j coupling bracket a range of coupling schemes. For outer electrons in intermediate Z atoms, the coupling behaves as a hybrid of L-S and j - j coupling. For example, the best models of the energy structure of some atoms invokes L-S for inner electrons and j - j for outer and excited electrons. There are other hybrid schemes such as J_1 - j or J_1 - J_2 coupling, $J_1 l$ or $J_1 L_2$ coupling, and LS_1 or LK coupling. We limit our discussions to L-S and j - j coupling.

3.1.3 The Russell-Saunders Vector Model

Russell-Saunders L–S coupling applies for the majority of atoms and ions for which astronomically interesting transitions and spectra are studied. We thus, elaborate on this model.

The eigenvalues L , S , J , M_L , M_S , and M_J describe the total quantum state of the atom or ion. The values of L , S , and J are obtained by quantized vector addition of the states of the individual electrons,

$$\mathbf{L} = \sum_{i=1}^{N_e} \boldsymbol{\ell}_i, \quad \mathbf{S} = \sum_{i=1}^{N_e} \mathbf{s}_i, \quad \mathbf{J} = |\mathbf{L} + \mathbf{S}|, \quad (3.3)$$

where $\boldsymbol{\ell}_i$ and \mathbf{s}_i are the angular momentum and spin vectors of the i th electron. The magnitude of the total orbital angular momentum of the atom is $|\mathbf{L}| = \sqrt{L(L+1)}\hbar$. The magnitude of the spin state of the atom due to the ensemble of electron spins is $|\mathbf{S}| = \sqrt{S(S+1)}\hbar$. The coupling of vectors \mathbf{L} and \mathbf{S} yields the total angular momentum $|\mathbf{J}| = \sqrt{J(J+1)}\hbar$.

The z axis projections of \mathbf{L} , \mathbf{S} and \mathbf{J} are $L_z = M_L\hbar$, $S_z = M_S\hbar$, and $J_z = M_J\hbar$. They are determined by the simple summations over all individual electrons,

$$M_L = \sum_{i=1}^{N_e} m_{l_i}, \quad M_S = \sum_{i=1}^{N_e} m_{s_i}, \quad M_J = M_L + M_S, \quad (3.4)$$

where $m_{l_i}\hbar$ and $m_{s_i}\hbar$ are the z projections of $\boldsymbol{\ell}_i$ and \mathbf{s}_i for the i th electron.

Note that the quantum description of multi-electron atoms are completely analogous to that of hydrogenic atoms. In the limit that $N_e = 1$, the hydrogenic wave functions and quantum numbers (see Chapter 1) are fully recovered.

Quantized Vector Addition

The quantized vector addition of Eq. 3.3 is constrained by the fact that the vectors \mathbf{L} , \mathbf{S} , and \mathbf{J} must be added such that the eigenvalues L , S , and J remain quantized. Since the l_i can add to or subtract from one another in unit intervals, L can take on multiple values for a given ensemble of electrons that differ by an integer and range from the difference of the individual l_i to the sum of the individual l_i . For example, consider two electrons with l_1 and l_2 , the minimum possible L state is $|l_1 - l_2|$ and the maximum is $l_1 + l_2$. In integer increments, L can take on all values between the minimum and the maximum, i.e., $|l_1 - l_2| + 1$, $|l_1 - l_2| + 2$, etc., up to $l_1 + l_2$.

For a given L , the z axis projections M_L , can take on the values,

$$M_L = -L, -L+1, \dots, 0, \dots, L-1, L. \quad (3.5)$$

Clearly, there are $2L + 1$ possible projections for M_L . Recall that M_L is the straight sum of the individual m_{l_i} (see Eq. 3.4). For any known electron configuration in which the m_{l_i} are specified, this straight sum provides the maximum L for the configuration simply by identifying the maximum $|M_L|$.

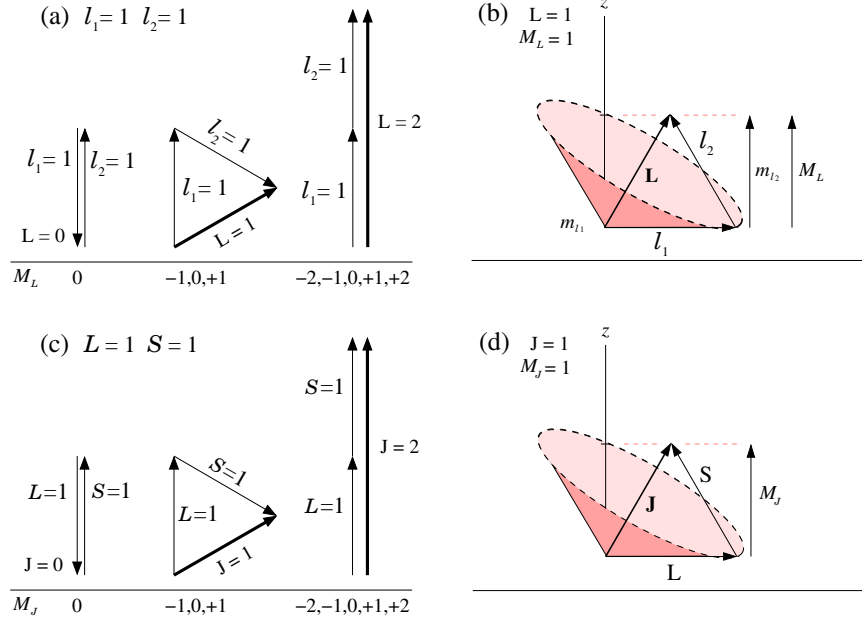


Figure 3.1 (a,b) Vector addition for $l_1 = 1$ and $l_2 = 1$ and (c,d) vector addition for $L = 1$ and $S = 1$. (a) Illustration of $\ell_1 + \ell_2$ vector addition. The possible L states are $L = |l_1 - l_2| = 0$, $L = |l_1 - l_2| + 1 = 1$ and $L = l_1 + l_2 = 2$. For $L = 0$, $M_L = m_{l_1} + m_{l_2} = 0$, or $m_{l_1} = -m_{l_2}$. For $L = 1$, M_L can be $-1, 0, +1$. For $L = 2$, M_L can be $-2, -1, 0, +1, +2$. (b) The spatial relationship for the addition yielding $L = 1$, $M_L = +1$ is illustrated for the case $m_{l_1} = 0$ and $m_{l_2} = +1$. The vectors ℓ_1 and ℓ_2 precess around \mathbf{L} . (c) Illustration of $\mathbf{L} + \mathbf{S}$ vector addition. The possible J states are $J = |L - S| = 0$, $J = |L - S| + 1 = 1$ and $J = L + S = 2$. For $J = 0$, $M_J = M_L + M_S = 0$. For $L = 1$, M_J can be $-1, 0, +1$. For $J = 2$, M_J can be $-2, -1, 0, +1, +2$. (d) The spatial relationship for the addition yielding $J = 1$, $M_J = +1$ is illustrated for the case $M_L = 0$ and $M_S = +1$. The vectors \mathbf{L} and \mathbf{S} precess around \mathbf{J} .

In Figures 3.1(a), 3.1(b), 3.1(c), we illustrate \mathbf{L} addition of two electrons. For simplicity, each electron has $l = 1$. The possible L values are 0, 1, and 2. The diagram illustrates the angles that are enforced between ℓ_1 and ℓ_2

required to achieve the various quantized values of L . The vector lengths are $|\ell_1| = \sqrt{l_1(l_1 + 1)}\hbar$, $|\ell_2| = \sqrt{l_2(l_2 + 1)}\hbar$, and $|\mathbf{L}| = \sqrt{L(L + 1)}\hbar$, respectively, though we identify the vectors by their quantum numbers. In Figure 3.1(b), we illustrate the spatial relationship between ℓ_1 , ℓ_2 , and \mathbf{L} for $L = 1$ and $M_L = +1$. The three equal length vectors form an equilateral triangle. For a z -axis projection of $M_L = +1$, the vector \mathbf{L} (of length $\sqrt{2}\hbar$) forms a 45° with the z axis. For this example, $m_{l_{i1}} = 0$ and $m_{l_{i2}} = +1$, which sum to $M_L = +1$.

In the absence of higher order physics (i.e., spin-orbit interactions), the constant of motion is \mathbf{L} . This vector precesses around the z axis with fixed M_L . The individual electron angular momentum vectors precess around \mathbf{L} and their z projections are no longer constant. The precession rate is directly proportional to the energy of the state, which is degenerate with M_L in the absence of magnetic fields.

The vector summation to obtain \mathbf{S} is also constrained by the quantization of S by the individual electron spin quantum numbers s_i . In this case, the projections of the individual electrons are either $m_s = +1/2$ or $-1/2$. As with L , S can take on multiple values for a given ensemble of electrons that differ by an integer and range from the difference of the individual s_i to the sum of the individual s_i , i.e., for two electrons the range is from $|s_1 - s_2|$ to $s_1 + s_2$. Since $s_i = 1/2$, we see that for an even number of electrons S is even and for an odd number of electrons S will be an integer multiple of $1/2$. Once S is determined, the spin vector magnitude is $|\mathbf{S}| = \sqrt{S(S + 1)}\hbar$. For a given S , the z axis projections, M_s , can take on the values,

$$M_s = -S, -S+1, \dots, S-1, S. \quad (3.6)$$

There are $2S + 1$ possible projections.

The inclusion of spin-orbit interactions yields a \mathbf{J} vector, the total angular momentum of the atom. For a given L and S , the possible J values are

$$J = |L - S|, |L - S| + 1, \dots, L + S, \quad (3.7)$$

with

$$M_J = -J, -J+1, \dots, J-1, J. \quad (3.8)$$

Since S can take on odd multiples of $1/2$ when there are an odd number of electrons, we see that this will be the case for J and M_J as well. As can be deduced from Eq. 3.7, if $L \geq S$ there are $2S + 1$ possible J states, and if $L < S$ there are $2L + 1$ possible J states. The number of J states is called the multiplicity because this gives the number of fine structure splittings in the energy states of the atom or ion for a given combination of L and S .

However, per convention, the multiplicity is always denoted $2S + 1$ even for atoms with $L < S$.

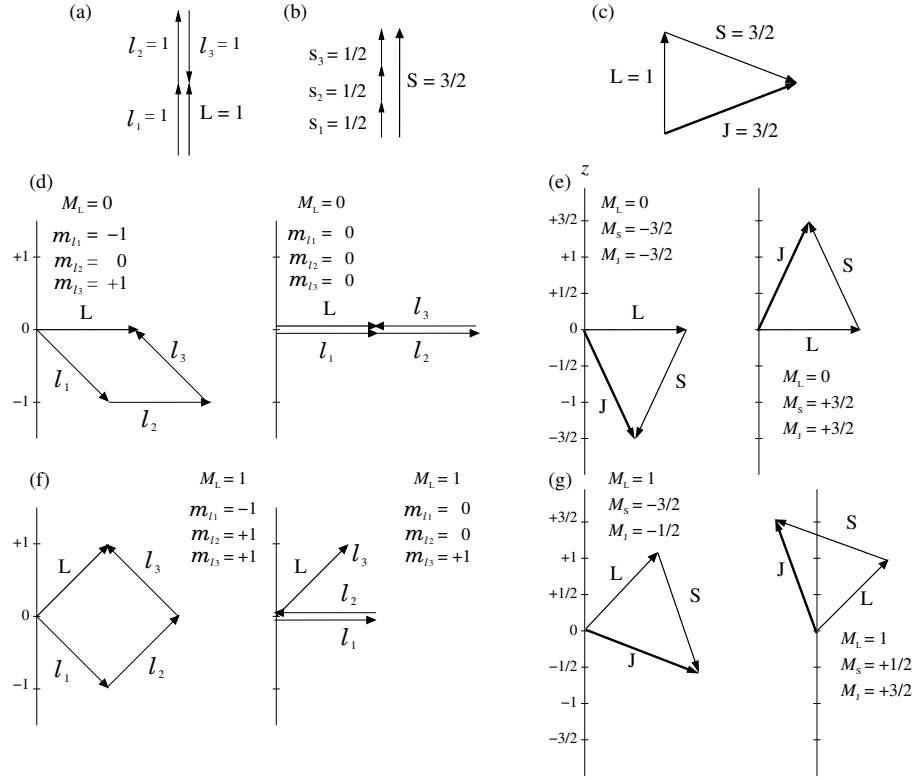


Figure 3.2 (a) Vector addition for three electrons with $l_1 = l_2 = l_3 = 1$ for the case in which $L = 1$. (b) Vector addition of the spins such that $S = 3/2$. In this case M_S can be $-3/2, -1/2, +1/2$, or $+3/2$. The possible J states are $J = 1/2, 3/2$, and $5/2$. (c) Vector addition for $J = 3/2$ with $L = 1$ and $S = 3/2$. In this case, M_J can be $-3/2, -1/2, +1/2$, or $+3/2$. (d) Two scenarios for $L = 1$ with $M_L = 0$. (e) The \mathbf{J} vector addition for $J = 3/2$ with $L = 1$ and $S = 3/2$ as constrained by (left) $M_L = 0, M_S = +3/2, M_J = -3/2$ and by (right) $M_L = 0, M_S = +3/2, M_J = +3/2$. (f) Two scenarios for $L = 1$ with $M_L = +1$. (g) The \mathbf{J} vector addition for $J = 3/2$ with $L = 1$ and $S = 3/2$ as constrained by (left) $M_L = +1, M_S = -3/2, M_J = -1/2$ and by (right) $M_L = +1, M_S = +1/2, M_J = +3/2$.

In Figure 3.1(c), we illustrate $\mathbf{L} + \mathbf{S}$ addition for $L = 1$ and $S = 1$, which yields $J = 0, 1$, or 2 depending upon the vector alignments. In Figure 3.1(d), we illustrate how J is the constant of motion and the \mathbf{J} vector precesses around the z axis with fixed M_J . The vectors \mathbf{L} and \mathbf{S} precess around \mathbf{J} so their z projections are no longer constant. The precession rate is directly

proportional to the energy of the state, which is degenerate with M_J in the absence of magnetic fields.

In Figure 3.2, we illustrate examples of \mathbf{L} , \mathbf{S} , and \mathbf{J} vector addition for an atom with three electrons. The vertical axis in each diagram is the z axis. The scenario we show is $L = 1$ for $l_1 = l_2 = l_3 = 1$ and $S = 3/2$ for $s_1 = s_2 = s_3 = 1/2$. The possible values of J are $1/2$, $3/2$, and $5/2$. We illustrate the case $J = 3/2$. This particular configuration forms an isosceles triangle with angle 75° between \mathbf{LS} and \mathbf{LJ} . Only a subset of all possible M_L , M_S , and M_J projections can yield this angular relationship between the vectors.

In Figures 3.2(d) and 3.2(e), we show two cases for $M_L = 0$, for which only the combinations $M_S, M_J = -3/2, -3/2$ and $+3/2, +3/2$ allow $J = 3/2$ for $L = 1$ and $S = 3/2$. In Figures 3.2(f) and 3.2(g), we show two cases for $M_L = +1$, for which only the combinations $M_S, M_J = -3/2, -1/2$ and $+1/2, +3/2$ allow $J = 3/2$ for $L = 1$ and $S = 3/2$. Note that for some of these scenarios, the Pauli exclusion principle would dictate that these $l = 1$ electrons would need to be distributed between different principle n levels.

3.1.4 Russell-Saunders Term and State Symbols

A useful notation for the overall state of an atom or ion was introduced by Henry Russell and Frederick Saunders in 1925. The LS state of an atom is called a term. The Russell-Saunders term symbol is written,

$$^{2S+1}\mathbf{L}^\pi, \quad (3.9)$$

where $2S + 1$ is called the multiplicity (as described in § 3.1.3). Recall that if $L < S$ then the J multiplicity is actually $2L + 1$, but it is convention to always call the multiplicity $2S + 1$. Spectroscopic notation is employed so that for $L = 0, 1, 2, 3$, etc., L is written as S, P, D, F, etc. For example, an atom with $S = 1$ and $L = 2$ is written ^3D . If $S = 0$, the term is called a singlet; if $S = 1/2$, it is a doublet; if $S = 1$, it is a triplet, etc. In Table 3.1, we list the multiplicities and J values for singlets, doublets, and triplets. In the case of singlets, $J = L$ and there is only a single state for a given term. For doublets, there are two states for a single term, and since there are an odd number of electrons, J is a multiple of $1/2$. For triplets, there are three states for a single term, and since there are an even number of electrons, J is an integer. Note that when $L < S$, the multiplicity is not $2S + 1$, but is actually $2L + 1$.

The superscript π denotes the Russell-Saunders term parity, which is either odd (antisymmetric wave functions), in which case π is written “o”,

Table 3.1 *Multiplicities*

Term	L	Singlets ($S=0$)		Doublets ($S=1/2$)			Triplets ($S=1$)					
		J		J			J					
S	0	0	1/2					1				
P	1	1	1/2	3/2			0	1	2			
D	2	2		3/2	5/2			1	2	3		
F	3	3			5/2	7/2			2	3	4	
G	4	4				7/2	9/2			3	4	5

or even (symmetric wave functions), in which case π is commonly omitted from the term symbol (though it can sometimes be written “e”). The parity is determined from

$$\pi = (-1)^\Lambda, \quad \Lambda = \sum_i l_i, \quad (3.10)$$

where the sum for Λ is over all electrons. The parity can be odd only in cases where an odd number of electrons occupy orbitals with odd l . The parity will otherwise be even.

The combination L , S , and J constitute the full state of an atom or ion. The Russell-Saunders *state* symbol is written

$$^{2S+1}L_J^\pi. \quad (3.11)$$

If $S = 1$ and $L = 1$, then there are three states, one with $J = |L - S| = 0$, one with $J = |L - S| + 1 = (L + S) - 1 = 1$, and one with $L + S = 2$. In this case the states are 3P_0 , 3P_1 , 3P_2 .

In Figure 3.3, we show $\mathbf{L} + \mathbf{S}$ vector addition that gives rise to several atomic states. For integer values of S there are an even number of electrons and for S that is a multiple of $1/2$ there are an odd number of electrons. Below each vector addition the Russell-Saunders state term is given (with the parity omitted).

3.2 Energy Structure

We now consider the energy structure of multi-electron atoms and ions by first applying the central field approximation via H_1 (Eq. 3.2) and then by adding the interaction energies from the Hamiltonians H_2 and H_3 as perturbations. For low to intermediate Z atoms and ions, where $|E_2| < |E_3|$ we first apply the Hamiltonian $H_1 + H_2$, which introduces L–S coupling

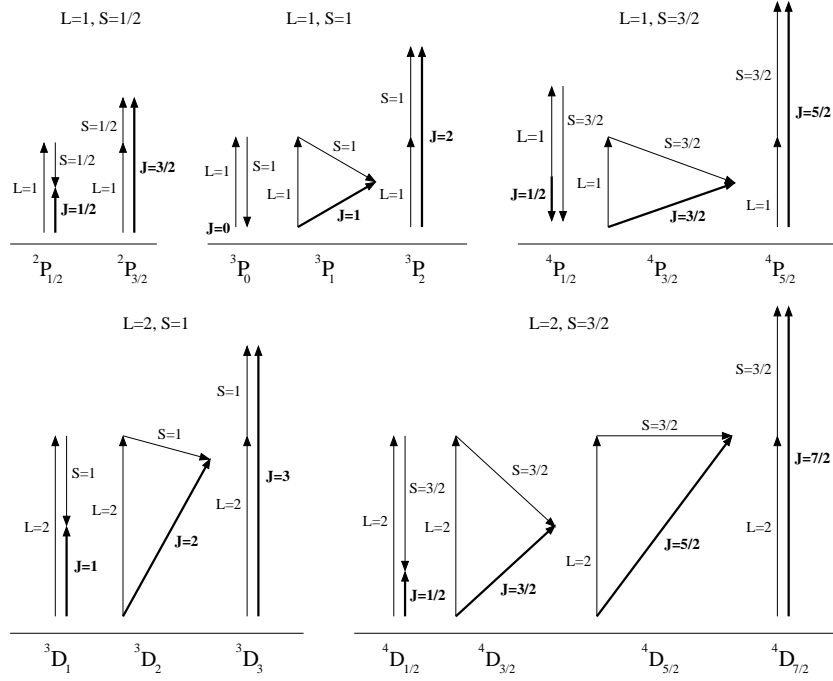


Figure 3.3 Examples of $\mathbf{L} + \mathbf{S}$ vector addition and their Russell-Saunders state terms. For $L \geq S$, the multiplicity of states is $2S + 1$, whereas for $L < S$ there are $2L + 1$ states. It is convention to always refer to the multiplicity as $2S + 1$ even in cases where $L < S$. For example, the terms with $L = 1$ and $S = 3/2$ (^4P) splits into the three states $^4\text{P}_{1/2}$, $^4\text{P}_{3/2}$, and $^4\text{P}_{5/2}$, but note that the multiplicity is still written as “4”.

via H_2 . We then apply the Hamiltonian $H_1 + H_2 + H_3$ and obtain fine structure splittings under the L–S coupling scheme. For higher Z elements where $|E_3| < |E_2|$, we first apply $H_1 + H_3$, which shows that the electrons interact most strongly via j – j coupling, and then apply $H_1 + H_3 + H_2$. The resulting general characteristics of ground-state atoms and ions are summarized in Chapter 4.

3.2.1 Central Field Approximation

For the central field approximation, we employ $H = H_1$ and average the Coulomb potential from the nucleus and the electron-electron Coulomb repulsions over all angles. This results in a spherically symmetric potential. The individual electrons occupy the available states $n l m_l m_s$ in accordance with the Pauli exclusion principle. However, whereas these quantum states

are constants of motion for hydrogenic atoms, they are not constants of motion for multi-electron atoms.

We adopt the terminology that all lm_l states corresponding to a given level n is called a shell. Shells are historically named K for $n = 1$, N for $n = 2$, L for $n = 3$, M for $n = 4$, and so on. All m_l states with a given nl are called a subshell. As with the hydrogenic atoms, subshells are denoted using the spectroscopic notation introduced in § 1.2.4, i.e., $n = 3$ and $l = 2$ is the 3d subshell, etc.

In Figure 3.4(a–d), we present selected Hartree-Fock ground-state radial distribution functions, $\sum_i r_i^2 R_{nl}^2(r_i)$, that result from the central field approximation. Shown are hydrogen and three atoms for which the shells are fully occupied: helium ($Z = 2$; K shell), neon ($Z = 10$; K and L shells), and argon ($Z = 18$; K, L, and M shells). The distributions are normalized so that the area under the curves equals the number of bound electrons. Note that the peak probability of a given shell occurs at smaller radius as Z increases. This is due to the increase in the attractive Coulomb potential for larger nuclear charge, Ze .

Whereas the binding energy of the electron in a hydrogenic atom depends only on which shell the electron occupies (i.e., principle quantum number, n) and is degenerate with the l state, this is not the case for multi-electron atoms. As obtained from H_1 , the binding energy depends on the nl state of the electron, i.e., $E = E_{nl}$. This is due to the electron-electron Coulomb repulsions. In a given shell, n , electrons in higher l states spend a greater amount of time farther from the nucleus than do the lower l state electrons. Thus, the electrons in higher l states are, on average, more effectively shielded from the nuclear potential. This renders electrons in higher l states in the same shell less bound than those with smaller l , e.g., $E_{n,l+1} > E_{nl}$ (recall that the $E_{nl} < 0$). The energies remain degenerate with m_l .

In Figure 3.4(e), we show the relative energies (increasing upward) of the nl states as a function of Z . Note that for each principle quantum number, n , higher l states have higher energy (they are less bound). However, some subshells have higher energy than the lowest l subshell in the next shell. For example, the 3d subshell has higher energy than the 4s subshell¹. However, as Z increases these subshells exchange order at $Z = 24$. There is a general trend that $l > 0$ subshells decrease in energy as Z increases, and the higher l the steeper is this decrease. The relative energies level off again around $Z \simeq 60$. Electrons will occupy the $nlm_l m_s$ states in order from the lowest energy to the highest energy, which means, in general, lowest n shells are

¹ The reason for this behavior is addressed when we present the ionization potentials in § 4.3.

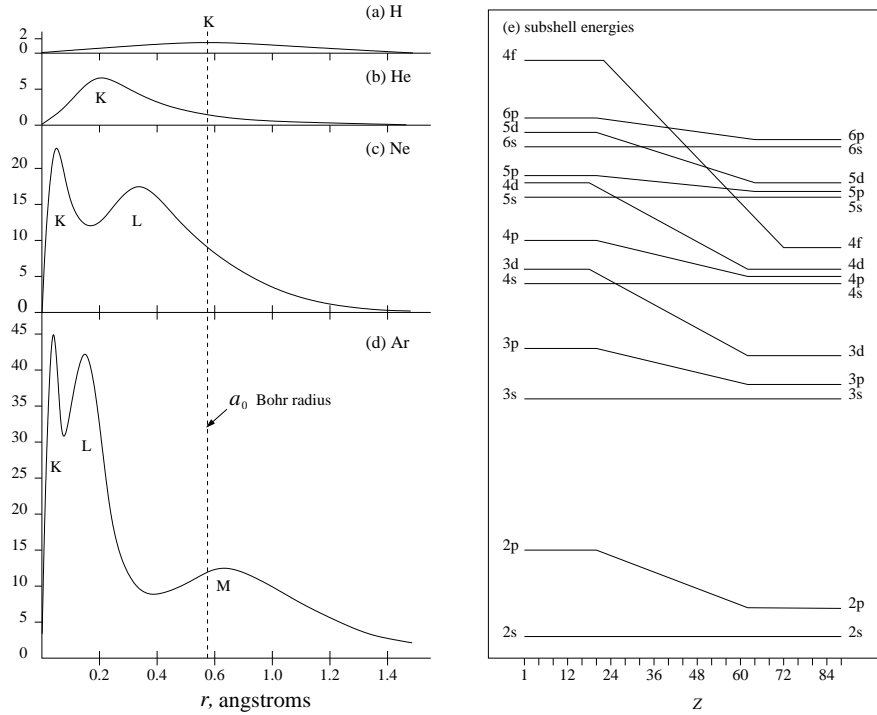


Figure 3.4 (a–d) The Hartree-Fock radial distribution functions for hydrogen ($Z = 1$), helium ($Z = 2$), neon ($Z = 10$) and argon ($Z = 18$). Note that the distributions for the individual shells overlap but are well defined. For a given shell, the distance r at which the probability of finding electrons peaks decreases with increasing Z . (e) The relative binding energies, E_{nl} , of the individual electrons as a function of nl (labeled using spectroscopic notation). Note that as Z increases, the binding energies for p ($l = 1$), d ($l = 2$), and f ($l = 3$) subshells decrease relative to the s ($l = 0$) subshells.

occupied first and each n level is occupied in order of increasing l ; this occupation is subject to the Pauli exclusion principle. However, exceptions in this ordering occur starting at $Z = 24$.

In Figure 3.5, we illustrate the overall structure of the occupation of the first three shells. First, the $n = 1$ shell is occupied. Since $l = 0$ and $m_l = 0$, there is no distinction between the shell and the subshell. The shell is populated with two electrons of opposite spin. Then the $n = 2$ shell is occupied. There are two subshells, $2s$ ($l = 0$), and $2p$ ($l = 1$). The $2s$ subshell is occupied first with two electrons of opposite spin, then the $2p$ shell is occupied with up to six electrons, such that each m_l state has two electrons of opposite spin. The process is repeated for the shell $n = 3$, the occupation order being $3s$, $3p$, and then $3d$.

There are two available states in each s subshell. With six available states

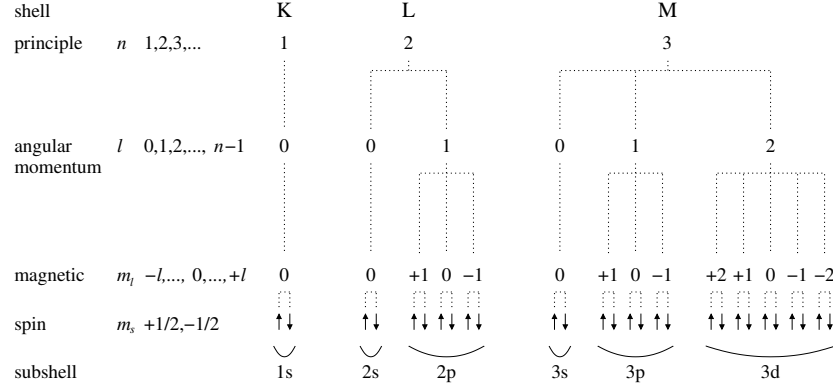


Figure 3.5 A schematic chart of the occupation of electrons into available quantum states, $nlm_l m_s$ for the central potential approximation. Because $E_{n,l+1} > E_{n,l}$, in a given shell, the low l subshells populate prior to the higher l subshells. For the Hamiltonian, H_1 , the $m_l m_s$ states in each subshell are energy degenerate. Electrons occupying the same subshell are called “equivalent” electrons, whereas electrons in different subshells are called “non-equivalent” electrons.

in the p subshell, there are a total of 10 available states that can be occupied in the $n = 2$ shell. Because there is a total of 10 available states in the 3d subshell, 6 in the 3p subshell, and 2 in the 3s subshell, there are total of 18 states in the $n = 3$ shell that can be occupied. In general, a subshell with l can be occupied with up to $g_s g_{nl} = (2s + 1)(2l + 1) = 2(2l + 1)$ electrons and a shell n can be occupied with up to $\sum_{l=0}^{n-1} g_s g_{nl} = 2n^2$ electrons.

For the Hamiltonian H_1 , the $m_l m_s$ states in a subshell are energy degenerate. Therefore, in this approximation of the multi-electron atomic structure, there are no rules to the order in which m_l and m_s states are occupied. When we introduce the Hamiltonian H_2 , we will see that the S state of the atom, coupled with the restrictions enforced by the Pauli exclusion principle, result in preferred occupation ordering of the m_l and m_s states.

In Table 3.2, we list the order of the occupation of states for neutral atoms up to $Z = 30$ (zinc). As the subshells are occupied, a superscript is used to denote the number of electrons in the subshell. For example, $2p^3$ corresponds to three electrons occupying the $l = 1$ subshell of the $n = 2$ shell. Note that the notation makes no distinction with regard to the m_l and m_s states of the electrons within the subshell. Electrons sharing a given subshell are called “equivalent electrons”. When the terms of all occupying electrons are catenated in order of their occupation, we obtain the electronic sequence. Once a shell is fully occupied, it is common to represent that portion of

Table 3.2 *Electronic sequences*

Z	Element	Electron Configuration	Z	Element	Electron Configuration
1	H.....	1s	16	S.....	[Ne]3s ² 3p ⁴
2	He.....	1s ²	17	Cl.....	[Ne]3s ² 3p ⁵
3	Li.....	[He]2s	18	Ar.....	[Ne]3s ² 3p ⁶
4	Be.....	[He]2s ²	19	K.....	[Ar]4s
5	B.....	[He]2s ² 2p	20	Ca.....	[Ar]4s ²
6	C.....	[He]2s ² 2p ²	21	Sc.....	[Ar]3d4s ²
7	N.....	[He]2s ² 2p ³	22	Ti.....	[Ar]3d ² 4s ²
8	O.....	[He]2s ² 2p ⁴	23	V.....	[Ar]3d ³ 4s ²
9	F.....	[He]2s ² 2p ⁵	24	Cr.....	[Ar]3d ⁵ 4s
10	Ne.....	[He]2s ² 2p ⁶	25	Mn.....	[Ar]3d ⁵ 4s ²
11	Na.....	[Ne]3s	26	Fe.....	[Ar]3d ⁶ 4s ²
12	Mg.....	[Ne]3s ²	27	Co.....	[Ar]3d ⁷ 4s ²
13	Al.....	[Ne]3s ² 3p	28	Ni.....	[Ar]3d ⁸ 4s ²
14	Si.....	[Ne]3s ² 3p ²	29	Cu.....	[Ar]3d ¹⁰ 4s
15	P.....	[Ne]3s ² 3p ³	30	Zn.....	[Ar]3d ¹⁰ 4s ²

the electronic sequence by the element symbol corresponding to the fully occupied shells.

Recall that for $Z < 24$, the 4s subshell has lower energy than the 3d subshell. Thus for $Z = 19$ and 20, the 4s subshell is occupied first. Since the energy of the 3d shell decreases below that of the 4s subshell for $Z \geq 24$, the ground-state sequence for chromium is [Ar]3d⁵4s rather than the next seemingly natural sequence [Ar]3d⁴4s². A similar anomaly occurs for copper ($Z = 29$).

3.2.2 *L-S Coupling*

Because electrons in the same subshell interact more strongly than electrons across subshells and/or shells, it is clear that the Coulomb interaction cannot be described accurately by a single spherically symmetric potential. The Hamiltonian H_2 includes the interaction energies due to non-spherical components of the potential. The adopted Hamiltonian is $H = H_1 + H_2$.

Recall that for the Hamiltonian H_1 , the energies depend on nl , or for a given shell n , on L alone. However, the Hamiltonian H_2 removes the degeneracy in S for a given L so that the eigenenergies are characterized by LS . This is known as Russell-Saunders L-S coupling.

It is important to realize that L-S coupling is not a direct spin-orbit “angular momentum” type of interaction; the Hamiltonian H_2 does not operate

on spin-orbit interactions, it only accounts for the residual asymmetries in the central field approximation (see Eq. 3.2). L–S coupling is a Coulomb perturbation and spherical asymmetry effect that is strongest for equivalent electrons (those in the same subshell) whose arrangement within the subshell is constrained via the Pauli exclusion principle.

In L–S coupling, a high total spin state, S , is the lowest energy (most tightly bound) and most stable configuration. Consider two electrons in the same nl subshell, i.e., equivalent electrons. Since they have the same nl , no two electrons can have the same m_l and m_s states or they violate the Pauli exclusion principle. That is, equivalent electrons that have identical spin states (same m_s) are forced to have different m_l states. For $l > 0$ subshells, the different m_l states have spatially orthogonal orientations as dictated by the spherical harmonic components of the wave function. Consider the p subshell.

Thus, equivalent electrons forced into different m_l states maintain a greater average distance from one another, whereas equivalent electrons sharing the same m_l state have a substantially smaller average distance from one another. That is, since electrons with identical nlm_s are forced to have different m_l via the Pauli exclusion principle, the average distance between electrons with identical nlm_s is a maximum for the subshell even though their average distance *from the nucleus* is the same. As a consequence, in partially occupied subshells, equivalent electrons with the same m_s (spins aligned) are less effective at perturbing the centralized potential they each sense from the nucleus; these electrons are thus more tightly bound than electrons with identical nlm_l , which must have opposite spin states and therefore a smaller S . Because the asymmetries in the central potential sensed by different m_l electrons are minimized, the energy state is also more stable.

In 1925, Friedrich Hund summarized this behavior in what are now referred to as Hund’s rules. Empirically noting that energy decreases with increasing S , he stated that the Russell-Saunders term with the largest S (greatest multiplicity) has lowest energy, and that for fixed S , the term with the largest value of L has the lowest energy.

The overall energy dependence of the atom or ion on the values of L and S , and therefore the specific l , m_l , and m_s electron configuration implies a preferred m_l and m_s ordering of the occupation of available states *within* subshells. We revisit this important point in Chapter 4 when describing the ground states of elements and ions with equivalent valence electronic sequences.

3.2.3 Spin-Orbit Coupling: Fine Structure

Incorporating spin-orbit interactions we employ $H = H_1 + H_2 + H_3$. The energy degeneracy associated with the term ^{2S+1}L is broken. This arises because the total orbital angular momentum from orbit-orbit coupling of electrons, \mathbf{L} , and the total spin resulting from spin-spin coupling of electrons, \mathbf{S} , weakly couple via the sum over all electrons of the dot products of their angular momentum and spin operators (see Eq. 3.2).

The constants of motion are the total angular momentum, $|\mathbf{J}| = \sqrt{J(J+1)} \hbar$, and its z projection $J_z = M_J \hbar$, where $\mathbf{J} = \mathbf{L} + \mathbf{S}$. The vector addition of \mathbf{L} and \mathbf{S} was reviewed in § 3.1.3. As illustrated in Figure 3.1, the vector \mathbf{J} precesses about the z axis with projection J_z , and the vectors \mathbf{L} and \mathbf{S} precess around \mathbf{J} . The term ^{2S+1}L is broken into $2S+1$ states (if $L \geq S$) or $2L+1$ states (if $L < S$). Analogous to the hydrogenic case, these splittings are known as fine structure.

		L	S	J	m_J			$^{2S+1}L_J$
					+1	0	-1	
np^2	$L=0, S=0, ^1S$	0	0	0	$\uparrow\downarrow$			1S_0
	$J=0, ^1S_0$	1	1	0	\uparrow	\uparrow		3P_0
		1	1	0	\uparrow		\uparrow	3P_0
	$L=2, S=0, ^1D$	1	1	0		\uparrow	\uparrow	3P_0
	$J=2, ^1D_2$	1	1	1	\downarrow	\uparrow		3P_1
		1	1	1	\uparrow		\downarrow	3P_1
		1	1	1		\uparrow	\downarrow	3P_1
		1	1	2	\downarrow	\downarrow		3P_2
		1	1	2	\downarrow		\downarrow	3P_2
		1	1	2		\downarrow	\downarrow	3P_2
	$H_1 + H_2$	2	0	2	$\uparrow\downarrow$			1D_2
		2	0	2	\uparrow	\downarrow		1D_2
		2	0	2	\downarrow		\uparrow	1D_2
	$L=1, S=1, ^3P$	2	0	2		\downarrow	\uparrow	1D_2
		2	0	2		$\uparrow\downarrow$		1D_2

Figure 3.6 The fine-structure splitting of LS terms for carbon ($1s^2 2s^2 2p^2$). The 15 permutations of occupation for 2 equivalent electrons in the p subshell are given in the table with their Russell-Saunders state symbols. Note that the ground-state of neutral carbon and of all np^2 atoms (and ions) is 3P_0 under the L-S coupling scheme.

In Figure 3.6, we show the fine-structure energy states of neutral carbon based on the Hamiltonian $H = H_1 + H_2 + H_3$. The energy of the levels increases upward. This structure applies for all atoms and ions with an np^2 valence electronic sequence (assuming all lower shell and subshells are fully occupied). We emphasize that any single atom or ion exists in only one of

these energy states at a given time. Thus, the energy structure refers to the states in which the atom *can* exist. For equivalent electrons, the state depends on the permutation of the electron occupation in the subshell and the realization of the vector sums of the l_i and s_i .

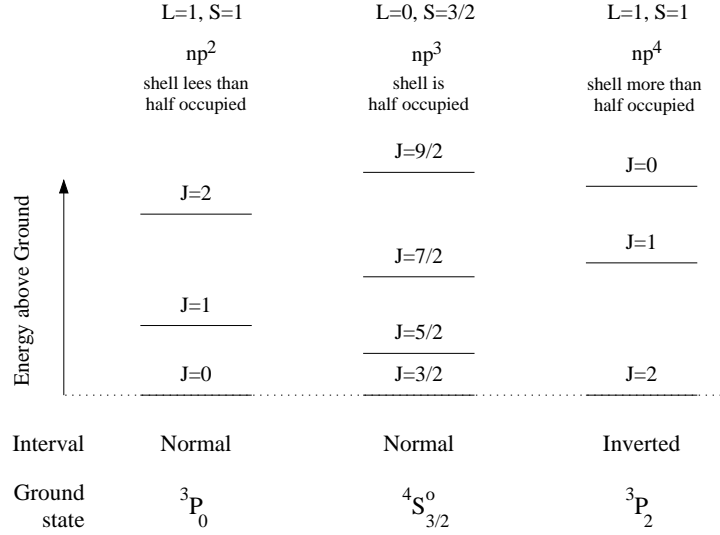


Figure 3.7 Schematics of the fine-structure energy intervals and their J dependence. The energy intervals are $J\Delta E(L, S)$, where $\Delta E(L, S)$ is known as the Landé interval. (left, center) For shells that are less than half occupied or exactly half occupied, the Landé interval yields “normal” multiplets in which higher J values have higher energies. (right) For shells that are more than half occupied or are fully occupied, the multiplet is “inverted”, in that higher J values have lower energies.

This J -dependent energy difference is known as the Landé interval. For a given atom/ion and with a given LS term, the Landé energy interval, $\Delta E(L, S)$, is a fixed value, but the energy difference between different J states is proportional to J , i.e., $E(L, S, J) - E(L, S, J-1) = J\Delta E(L, S)$. When a subshell is less than half occupied or exactly half occupied, $\Delta E(L, S) > 0$, so lower values of J have lower energies, i.e., $E(L, S, J-1) < E(L, S, J)$, and the multiplets are known as “normal” multiplets. The lowest possible value of J and thus the lowest energy multiplet is $J = |L - S|$. When a subshell is exactly half occupied, $L = 0$ and thus $J = S$. When a subshell is more than half occupied, $\Delta E(L, S) < 0$, so higher values of J have lower energies, i.e., $E(L, S, J) < E(L, S, J-1)$, and the multiplets are known as “inverted” multiplets. In this case, the highest possible value of J , and thus highest energy multiplet, is always $J = L + S$. In Figure 3.7, we have illustrated the behavior of the Landé interval for the lowest energy LS terms

for carbon ($L = 1, S = 1, np^2$), nitrogen ($L = 0, S = 3/2, np^3$), and oxygen ($L = 1, S = 1, np^4$). These respectively illustrate a less than half-occupied, half-occupied, and more than half-occupied np subshell. Note the inverted multiplets for oxygen.

Under the condition of statistical equilibrium, for a large number of these atoms or ions in a gas, one would find that each of these state would be represented. For a state LSJ , the relative proportion of the number of atoms or ions in a given state is the ratio of the number of energy degenerate M_J states to the total number of M_J states for all possible LSJ . Since there are $2J + 1$ possible M_J states for a given J , the proportion of atoms in state LSJ is

$$\frac{2J + 1}{\sum_{LS} \sum_J (2J + 1)} = \frac{2J + 1}{\sum_{LS} (2S + 1)(2L + 1)}, \quad (3.12)$$

where the sum is taken over the L and S states, i.e., the three states $(L, S) = (0, 0), (1, 1),$ and $(2, 0)$. For an ensemble of atoms with the np^2 configuration discussed above, 6.6% would be in the 1S_0 state 33.3% in the 1D_2 state 33.3% in the 3P_2 state 20% in the 3P_1 state, and 6.6% in the 3P_0 state. However, in a wide range of astrophysical settings, the atoms are not in statistical equilibrium; the atoms and ions spend the vast majority of their time in the ground state. Thus, virtually all the atoms and ions will be in the 3P_0 state.

Recapping, for the central potential approximation, H_1 , shown on the left of Figure 3.6, the energy depends only on nl , and is degenerate for all permutations of the subshell occupation. For L-S coupling, H_2 , the states split into the three terms $^1S, ^3P,$ and 1D , and these depend on the permutation of the electron configuration. The Hamiltonian operator H_3 introduces fine-structure splitting.

3.2.4 $j-j$ and Intermediate Coupling Schemes

Up to this point, we have discussed only pure L-S coupling. However, as mentioned in § 3.1.2, L-S coupling does not adequately describe the energy levels of all elements, which is known because the theoretically predicted energy levels do not match the spectra of these atoms and their ions.

As shown in Eq. 1.38 for hydrogenic atoms, the spin-orbit multiplet splittings (fine structure) scale as $E_n(Z/n)^2 \simeq Z^4/n^2$, whereas the Coulomb potential scales as Z . Thus, as Z increases, we see that the $j = l \pm s$ spin-orbit coupling becomes stronger more rapidly than does the electrostatic L-S coupling component. Examination of the Hamiltonian H_3 (Eq. 3.2) for

spin-orbit coupling indicates similar behavior is expected for multi-electron atoms and ions. Thus, as Z increases, the coupling of an electron having angular momentum ℓ_i with its own spin, \mathbf{s}_i , grows in relative strength over the electrostatic component. In the limit that spin-orbit coupling dominates, the theoretical formalism is called j - j coupling.

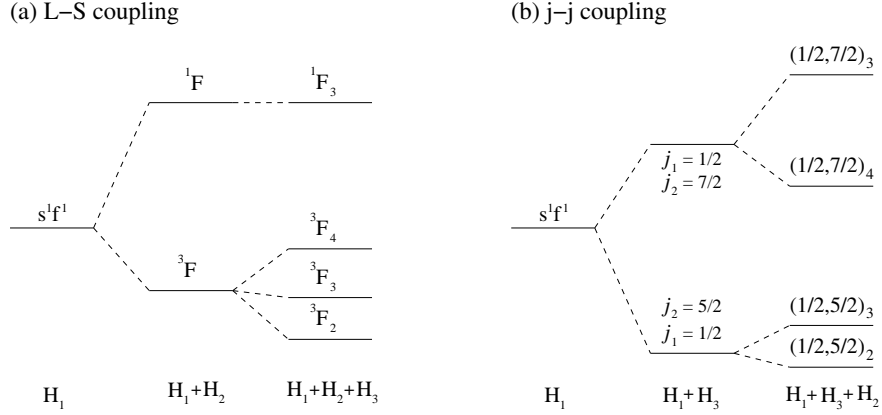


Figure 3.8 A comparison between pure L-S coupling and pure j - j coupling for two non-equivalent electrons, one in the ground state s subshell and the second excited up to the f subshell. Perturbations to the central field approximation, H_1 , are carried out in order from largest to smallest interaction energies. Larger interaction energies result in larger energy shifts and splittings. (a) Under the condition of L-S coupling, the electron-electron Coulomb repulsion, H_2 , dominates over the spin-orbit interactions, H_3 . (b) Under the condition of j - j coupling, the spin-orbit interactions dominate over the electron-electron Coulomb repulsion. The same number of J states result, however, the different strengths of the interactions result in different energy structures.

This is particularly the case for electrons that are in excited states occupying an outer shell and subshell. In pure spin-orbit coupling, ℓ_i is most strongly coupled to \mathbf{s}_i via the Hamiltonian H_3 , yielding $\mathbf{j}_i = \ell_i + \mathbf{s}_i$ for the total angular momentum of the i th electron. This interaction results in energy splits of the E_{nl} , which now become E_{nlj} . Via the Hamiltonian H_2 , the \mathbf{j}_i of the individual electrons then weakly couple according to

$$\mathbf{J} = \sum_i^{N_e} \mathbf{j}_i = \sum_i^{N_e} (\ell_i + \mathbf{s}_i) . \quad (3.13)$$

Thus, the atom has a well-defined total angular momentum, \mathbf{J} , but note that it does not have a well-defined \mathbf{L} or \mathbf{S} .

In Figure 3.8, we illustrate the L-S and j - j coupling schemes for excited

atoms with ground-state ns^2 electron sequences for which one electron occupies the s shell and the excited electron occupies the f shell. As shown in Figure 3.8(a), the operation $H_1 + H_2$ splits the energy, following which the operation $H_1 + H_2 + H_3$ introduces finer energy splits. As shown in Figure 3.8(b), the operation $H_1 + H_3$ splits the energy, following which the operation $H_1 + H_3 + H_2$ introduces finer energy splits. Note that the predicted energy structures are quite different, even though the final J states and the number of multiplets are the same for both schemes. For $j-j$ coupling, the notation is $X(j_1, j_2, \dots, j_N)_J$ for the N electrons in the valance shells and subshells, where X denotes the electronic sequence of these N electrons (for example 3s4f).

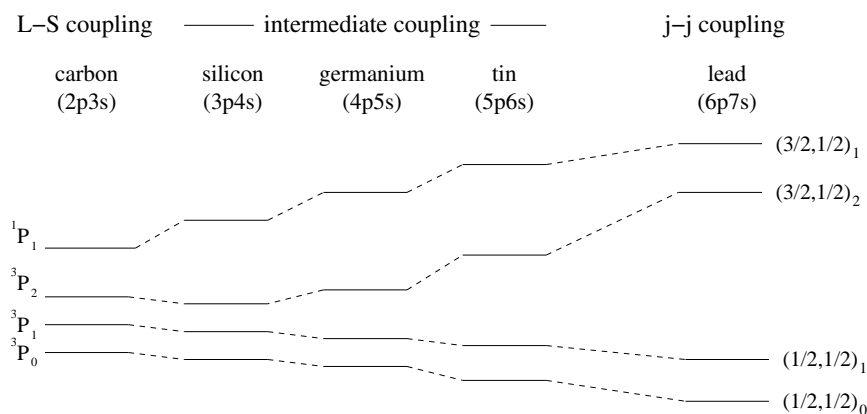


Figure 3.9 The “carbon group” of elements (those having ground-state valence electron configurations np^2) with two non-equivalent electrons $npn's$ (one electron is excited to the $n's$ subshell, where $n' > n$). Based upon its spectrum, carbon is well described by L-S coupling, whereas lead is well explained by $j-j$ coupling. The intermediate elements have spectra that indicate energy levels that are neither pure L-S nor pure $j-j$ coupling and that there is a transition from the former to the latter as Z increases. Since the number of J states is always the same for the different coupling schemes and the energies of the individual J state energies remain in the same relative order, it is convention to adopt a 1:1 correspondence to the J states as the relative strengths of the coupling physics change. It is also convention to adopt the Russell-Saunders state symbol for all atoms and ions, even if $j-j$ or other coupling schemes best reflect the energy levels based upon the observed spectra.

For some atoms and ions a mixture of L-S and $j-j$ coupling can be adopted to best theoretically understand the atomic energy structure based upon observed spectra. For example, inner core electrons can be described by L-S coupling, while outer electrons can be described by $j-j$ coupling.

The behavior of energy splittings for excited carbon-like neutral atoms

are shown in Figure 3.9. The excited states are all assumed to be $n'p$'s, where $n' = n + 1$. Carbon is well described by pure L-S coupling and lead is well described by pure $j-j$ coupling. Silicon, germanium, and tin indicate that intermediate coupling schemes are required, with silicon dominated by L-S coupling and tin dominated by $j-j$ coupling. Note that the number of splittings is the same and the J states of the atoms have a one-to-one correspondence. As such, it is convention to employ the Russell-Saunders state symbol to describe an atomic state.

3.3 Bound-Bound Transitions

The transition rates and oscillator strengths of hydrogenic atoms and ions were discussed in Chapter 2. The treatment of multi-electron atoms follows similar formalism, so we draw extensively from the results for hydrogenic atoms.

Invoking the weak-field approximation, the additional term in the Hamiltonian for inclusion of interactions with the electric field is given by,

$$H'(t) = \frac{e}{m_e} \sum_{i=1}^{N_e} \mathbf{A}(\mathbf{r}_i, t) \cdot \mathbf{p}_i = -i\hbar \frac{e}{m_e} \sum_{i=1}^{N_e} \mathbf{A}(\mathbf{r}_i, t) \cdot \nabla_i, \quad (3.14)$$

where $\mathbf{A}(\mathbf{r}_i, t)$ is the electric field vector.

Consider the transition between state k and state k' , where k' denotes the lower state and k denotes the upper state. For L-S coupling, the notation is shorthand for denoting the atomic states $k' \equiv \gamma' L' S' J' M'_J$ and $k \equiv \gamma L S J M_J$. For the $j-j$ coupling scheme, L and S are not “good” quantum numbers (i.e., not constants of motion), and the two states correspond to $k' \equiv \gamma' J' M'_J$ and $k \equiv \gamma J M_J$. The γ notation is a general place holder representing the shell level n and the form of the Hamiltonian applied to obtain the wave functions.

For first-order perturbation treatment, we assume radiative transitions involve a single electron per photon emitted or absorbed. Under this assumption, the methods employed for hydrogenic atoms in Chapter 2 apply analogously.

Following the steps to derive Eq. 2.4 for hydrogenic atoms, we obtain the Einstein A coefficient for spontaneous emission [s^{-1}],

$$A_k^{k'} = \frac{32\pi^3 e^2 (\nu_k^{k'})^3}{3\hbar c^3} |\mathbf{R}_k^{k'}|^2. \quad (3.15)$$

where the integration over all space is

$$\mathbf{R}_k^{k'} = \sum_{i=1}^{N_e} \langle \psi_k(\mathbf{r}_i) | \mathbf{r}_i | \psi_{k'}(\mathbf{r}_i) \rangle \quad \text{with} \quad |\mathbf{R}_k^{k'}| = |\mathbf{R}_k^k|, \quad (3.16)$$

which is obtained by applying the same procedural steps used to derive the hydrogenic dipole approximation. Unlike the hydrogenic case, for which the wave functions are closed expressions, the multi-electron wave functions are Hartree-Fock numerical approximations. Thus, closed-form expressions cannot be derived for the overlap integrals.

The Einstein A coefficient, $A_k^{k'}$, and the Einstein B coefficients (see Eq. 2.6) for stimulated emission, $B_k^{k'}$, and absorption, $B_{k'}^k$, have the same interpretations as for hydrogenic atoms. Furthermore, the Einstein relations given in Eq. 2.9 apply as well,

$$\begin{aligned} B_k^{k'} &= \frac{c^2}{8\pi h \nu^3} A_k^{k'} \quad \text{stimulated} \leftrightarrow \text{spontaneous emission} \\ B_{k'}^k &= \frac{g_k}{g_{k'}} B_k^{k'} \quad \text{absorption} \leftrightarrow \text{stimulated emission}, \end{aligned} \quad (3.17)$$

where g_k and $g_{k'}$ are the statistical weights, which we discuss further below.

3.3.1 Oscillator Strengths and Line Strengths

As with hydrogenic atoms (see Eqs. 2.11 and 2.10), we define the oscillator strength as

$$f_k^{k'} = \frac{4\pi m_e}{3\hbar} \nu_k^{k'} |\mathbf{R}_k^{k'}|^2, \quad (3.18)$$

such that

$$A_k^{k'} = \frac{8\pi^2 e^2 (\nu_k^{k'})^2}{m_e c^3} f_k^{k'}. \quad (3.19)$$

Fine Structure Transitions

In what follows, we assume L–S coupling, however, the treatment is identical for arbitrary coupling scheme, including j – j coupling.

In the absence of magnetic fields, the $2J+1$ M_J z -axis projections of \mathbf{J} for state γLSJ are energy degenerate. Thus, for a given fine-structure spectral line (transition energy), the oscillator strengths and transition probabilities are obtained by averaging all transitions between between fine-structure

states γLSJ and $\gamma' L'S'J'$ over the energy degenerate M_J and M'_J states, giving that average fine-structure oscillator strength

$$\bar{f}_k^{k'} = \frac{1}{\bar{g}_k} \sum_{M_J \downarrow M'_J} f_k^{k'}, \quad (3.20)$$

where the bar indicates the averaged value corresponding to fine-structure terms $k' \equiv \gamma' L'S'J'$ and $k \equiv \gamma LSJ$, and where the \downarrow indicates the sum is over all M_J states of γLSJ that transition down to states M'_J of $\gamma' L'S'J'$. There are $2J+1$ initial M_J states (and therefore elements to the sum), so that the averaging is obtained by dividing by the statistical weight $\bar{g}_k = 2J+1$.

Invoking the definitions in Eqs. 3.18, we have

$$\bar{f}_k^{k'} = \frac{4\pi m_e \nu_k^{k'}}{3\hbar(2J+1)} \sum_{M_J \downarrow M'_J} |\mathbf{R}_k^{k'}|^2, \quad (3.21)$$

where

$$\bar{A}_k^{k'} = \frac{8\pi^2 e^2 (\nu_k^{k'})^2}{m_e c^3} \bar{f}_k^{k'}. \quad (3.22)$$

Eqs. 3.21 and 3.22 are the dipole-approximation fine-structure oscillator strengths and Einstein A coefficients, respectively. The average for the reverse sense of the transition (absorption) is normalized by $\bar{g}_{k'}$, and because of the symmetry property given by Eq. 3.16, this yields

$$\bar{g}_k \bar{f}_k^{k'} = \bar{g}_{k'} \bar{f}_{k'}^k, \quad (3.23)$$

where $\bar{g}_k = 2J+1$ and $\bar{g}_{k'} = 2J'+1$.

Term Averaged Transitions

When fine-structure multiplets are blended in observational data, the recorded line is the average over all J states. In this case, the observed spectral line oscillator strengths and transition probabilities are obtained by averaging all transitions between term states γLS and $\gamma' L'S'$ over the energy dependent J and J' states, giving the term averaged oscillator strength. We sum the fine-structure oscillator strengths over all transitions from J to J' , obtaining

$$\begin{aligned} \langle \bar{f}_k^{k'} \rangle &= \frac{4\pi m_e}{3\hbar \langle \bar{g}_k \rangle} \sum_{J \downarrow J'} \nu_k^{k'} \sum_{M_J \downarrow M'_J} |\mathbf{R}_k^{k'}|^2 \\ &= \frac{4\pi m_e}{3\hbar^2 c \alpha \langle \bar{g}_k \rangle} \sum_{J \downarrow J'} \nu_k^{k'} S_k^{k'}, \end{aligned} \quad (3.24)$$

where, for notational convenience, we introduce the line strength function

$$S_k^{k'} = \hbar c \alpha \sum_{M_J \downarrow M_J'} \left| \mathbf{R}_k^{k'} \right|^2. \quad (3.25)$$

With regards to notation in Eq. 3.24, the bar plus bracket indicates the term averaged value corresponding to $k' \equiv \gamma' L' S'$ and $k \equiv \gamma L S$. A given LS term is split into $(2S+1)(2L+1)$ J states each of which have M_J energy degenerate states, so that the averaging is obtained by dividing by the statistical weight $\langle \bar{g}_k \rangle = (2S+1)(2L+1)$. The transition energies (frequencies) remain in the sum because the J state energies for a given LS term are not energy degenerate; the same applies for J' states). The mean transition energy for term-averaged transitions is obtained by the line-strength weighted average,

$$\langle \nu_k^{k'} \rangle = \frac{\sum_{J \downarrow J'} \nu_k^{k'} S_k^{k'}}{S_k^{k'}} \quad \text{where} \quad S_k^{k'} = \sum_{J \downarrow J'} S_k^{k'}, \quad (3.26)$$

Solving for the sum in the numerator of Eq. 3.26 and substituting into Eq. 3.24, we obtain,

$$\langle \bar{f}_k^{k'} \rangle = \frac{4\pi m_e \langle \nu_k^{k'} \rangle}{3\hbar^2 c \alpha} \frac{S_k^{k'}}{(2S+1)(2L+1)}. \quad (3.27)$$

The average for the reverse sense of the transition (absorption) is normalized by $\langle \bar{g}_{k'} \rangle = (2S'+1)(2L'+1)$, which yields the analogous relation to Eq. 3.23,

$$\langle \bar{g}_k \rangle \langle \bar{f}_k^{k'} \rangle = \langle \bar{g}_{k'} \rangle \langle \bar{f}_{k'}^k \rangle \quad (3.28)$$

Following the same procedures used above to obtain $\langle \bar{f}_k^{k'} \rangle$, we derive the Einstein A coefficient,

$$\langle \bar{A}_k^{k'} \rangle = \frac{32\pi^3 \langle \nu_k^{k'} \rangle^3}{3\hbar c^3} \frac{S_k^{k'}}{(2S+1)(2L+1)}. \quad (3.29)$$

In Table 3.3, we present selected examples of term averaged absorption transitions for H I Ly α , Na I $\lambda\lambda 5891, 5897$, and Mg II $\lambda\lambda 2796, 2803$ doublets and the quadruplet from Si II for which the Si II $\lambda\lambda 1190, 1193$ doublet is often observed (see §§ 4.4.1 and 4.4.2 for details). Whereas the metal-line multiplets are commonly resolved in observational spectra, the H I Ly α doublet is not. For each multiplet, the first row provides the averaged quantities for the transition terms $^S L$ and $^{S'} L'$, the mean wavelength, excitation energy of the term (from which the transition energy is $h\nu_k^{k'} = |E_k' - E_k|$), statistical weight, Einstein A coefficient, and oscillator strength. For the term ^{2S+1}L ,

Table 3.3 *Selected term-averaged line strengths*

Term/State $^{2S+1}L - ^{2S'+1}L'$	λ Å	E_k eV	$E_{k'}$ eV	g_k	$g_{k'}$	$A_k^{k'}$ 10^8 s^{-1}	$f_{k'}^k$
— H I $\lambda 1215$ —							
$^2S - ^2P^o$	1215.67004	0.000000	10.198836	2	6	6.2650	0.4164
$^2S_{1/2} - ^2P_{3/2}^o$	1215.66824	0.000000	10.198851	2	4	6.2648	0.2776
$^2S_{1/2} - ^2P_{1/2}^o$	1215.67364	0.000000	10.198806	2	2	6.2649	0.1388
— Na I $\lambda 5893$ —							
$^2S - ^2P^o$	5893.573	0.000000	2.103718	2	6	0.615	0.961
$^2S_{1/2} - ^2P_{3/2}^o$	5891.583	0.000000	2.104429	2	4	0.616	0.641
$^2S_{1/2} - ^2P_{1/2}^o$	5897.558	0.000000	2.102297	2	2	0.614	0.320
— Mg II $\lambda 2798$ —							
$^2S - ^2P^o$	2798.740	0.000000	4.430000	2	6	2.59	0.912
$^2S_{1/2} - ^2P_{3/2}^o$	2796.352	0.000000	4.433784	2	4	2.60	0.608
$^2S_{1/2} - ^2P_{1/2}^o$	2803.531	0.000000	4.422431	2	2	2.57	0.303
— Si II $\lambda 1194$ —							
$^2P^o - ^2P$	1194.100	0.023742	10.406839	6	6	246	0.876
$^2P_{3/2}^o - ^2P_{1/2}$	1197.394	0.035613	10.390117	4	2	280	0.150
$^2P_{3/2}^o - ^2P_{3/2}$	1194.500	0.035613	10.415200	4	4	138	0.737
$^2P_{1/2}^o - ^2P_{1/2}$	1193.290	0.000000	10.390117	2	2	53.8	0.575
$^2P_{1/2}^o - ^2P_{3/2}$	1190.416	0.000000	10.415200	2	4	26.1	0.277

Term average transitions have term symbols; fine-structure transitions have state symbols.
 $g_k = (2S+1)(2L+1)$ for term averaged transitions and $2J+1$ for fine-structure transitions.
 All data are obtained from Kramida et al. (2014).

the term-averaged statistical weight is $g_k = (2S+1)(2L+1)$. The subsequent rows list these quantities for the individual fine-structure transition for the states $^{2S+1}L_J$ and $^{2S'+1}L'_{J'}$. The statistical weights are $g_k = 2J+1$ and $g_{k'} = 2J'+1$, respectively.

3.3.2 Emission Power and Absorption Cross Section

Following the reasoning presented in § 2.6.1 for hydrogenic atoms and ions, we obtain the emission power per unit frequency per steradian from a single atom [$\text{erg s}^{-1} \text{ Hz}^{-1} \text{ str}^{-1} \text{ atom}^{-1}$],

$$j_k^{k'}(\nu) d\nu = \frac{h\nu}{4\pi} A_k^{k'} \phi_k^{k'}(\nu) d\nu = \frac{h\nu}{4\pi} A_k^{k'} \cdot \frac{1}{\pi} \frac{(\Gamma_k^{k'}/4\pi)}{(\nu - \nu_k^{k'})^2 + (\Gamma_k^{k'}/4\pi)^2} d\nu, \quad (3.30)$$

where $\phi_k^{k'}(\nu)$ is the natural broadening function (see Eq. 2.66), and $\Gamma_k^{k'}$ is the damping constant (see Eqs. 2.58, 2.56, and 2.55). Following the arguments

to obtain Eq. 2.75, we have the integrated emission power [$\text{erg s}^{-1} \text{ str}^{-1} \text{ atom}^{-1}$],

$$J_k^{k'} = \frac{h\nu_k^{k'}}{4\pi} A_k^{k'} \int_0^\infty \phi_k^{k'}(\nu) d\nu = \frac{h\nu_k^{k'}}{4\pi} A_k^{k'} = \frac{2\pi h e^2}{m_e c^3} (\nu_k^{k'})^3 f_k^{k'}. \quad (3.31)$$

Following the reasoning presented in § 2.6.2 for hydrogenic atoms and ions, we obtain the absorption power per unit frequency per unit solid angle per atom [$\text{erg s}^{-1} \text{ Hz}^{-1} \text{ str}^{-1} \text{ atom}^{-1}$],

$$k_{k'}^k(\nu) d\nu = h\nu B_{k'}^k \mathcal{I}_\nu \phi_{k'}^k(\nu) d\nu, \quad (3.32)$$

where \mathcal{I}_ν is the specific intensity of the incident radiation field (see Chapter 7). Carrying out the steps to derive Eq. 2.81 for hydrogenic atoms and ions, we obtain the absorption cross section [$\text{cm}^2 \text{ Hz}^{-1} \text{ atom}^{-1}$],

$$\sigma_{k'}^k(\nu) = \frac{\pi e^2}{m_e c} f_{k'}^k \cdot \frac{1}{\pi} \frac{(\Gamma_k^{k'}/4\pi)}{(\nu - \nu_k^{k'})^2 + (\Gamma_k^{k'}/4\pi)^2}. \quad (3.33)$$

As stated in § 2.6, the emission power and absorption cross sections are key quantities for solving the radiative transfer equation and analyzing absorption lines in spectra. We will be invoking these quantities in subsequent chapters.

3.3.3 Selection Rules

In § 2.1 we introduced the dipole approximation for hydrogenic atoms and ions. Having generalized the hydrogenic treatment for multi-electron atoms and ions, we obtained the moment between the probability densities of the wave functions for the upper and lower states of a transition as expressed in Eq. 3.16. This is known as the dipole approximation because $\mathbf{D}_k^{k'} = e\mathbf{R}_k^{k'}$ is the electric dipole moment between the charge distributions of states k and k' .

In Eq. 3.16, note that in the integral $\langle \psi_k(\mathbf{r}_i) | \mathbf{r}_i | \psi_{k'}(\mathbf{r}_i) \rangle$, when the operator \mathbf{r}_i is replaced with $-\mathbf{r}_i$, the result is $-\langle \psi_k(\mathbf{r}_i) | -\mathbf{r}_i | \psi_{k'}(\mathbf{r}_i) \rangle$, so that when the dipole operator acts, the parity of the operation is odd, yielding a sign change in $\mathbf{D}_k^{k'}$. This means that the dipole operator requires a parity change in the state of the atom during a dipole transition. This is known as LaPorte's rule: "the atomic state before and after a transition must have opposite parity".

The dipole transition is often referred to as an E1 transition. For the dipole approximation, the dipole moment $\mathbf{D}_k^{k'}$, and therefore the transition probability, is null between all atomic states except those for which $\Delta J =$

$J - J' = 0, \pm 1$ and $\Delta M_J = 0, \pm 1$, with the exception $\Delta J = 0$ when $J = 0$ and $J' = 0$. Finally, for E1 transitions, the parity must change. These conditions apply for all spin-orbit coupling schemes.

In the case of L-S coupling, the additional criteria for non-zero transition probabilities are $\Delta L = 0, \pm 1$ and $\Delta M_L = 0, \pm 1$, with the exception $\Delta L = 0$ when $L = 0$ and $L' = 0$. Note however, that even for $\Delta L = 0$, the parity (Eq. 3.10) can change since parity depends upon the sum of the l_i , which can change from odd to even (or vice versa) even when the quantized vector sum for L does not change. Since the dipole operator does not operate on spin, the additional rule of E1 transitions is that $\Delta S = 0$ (no change in the overall spin). This latter rule (also referred to as LaPorte's rule) will be instrumental in determining the spectra from multi-electron atoms and ions (see § 4.4).

Per the weak-field approximation, and true generally in nature, a single electron actively engages in the transition. Thus, in addition to the atomic system obeying the above rules, the dipole selection rules for single electrons will apply to the “active” electron. If this is electron i , then $\Delta m_i = 0, \pm 1$ and $\Delta l_i = \pm 1$. Note that $\Delta l_i = 0$ is not permitted because the parity of the atom must change during the transition while $\Delta s_i = 0$ is enforced by $\Delta S = 0$.

Transitions not permitted via the dipole approximation (called “forbidden” transitions) can occur for higher moments of the charge density distributions between the two transition states. These higher-order moments are obtained by the inclusion of the second term in a Taylor expansion of the electric field vector. From this treatment, one can derive two additional terms to the transition probabilities.

The first term is known as the electric quadrupole moment and is often denoted E2. Quadrupole transitions arise when the charge distributions of two states are oscillating (like electric dipoles), but out of phase such that the dipole contribution vanishes. The second term is known as the magnetic dipole moment and is denoted M1. It arises from a magnetic dipole moment between the charge distributions, which can be seen by the fact that the magnetic dipole is defined as $\boldsymbol{\mu} = e\hbar/(2m_e c)(\mathbf{r} \times \mathbf{p})$, the cross product of the position and momentum vectors.

The strength of E2 and M1 transitions are on the same order, both being a factor of $\sim k\langle r \rangle$ smaller than the dipole transitions, where $k = 2\pi/\lambda$ is the wavenumber of the transition energy and $\langle r \rangle$ is the mean radial distance of the active electron from the nucleus. Whereas E1 dipole transitions are called “permitted” or “allowed”, E2 and M1 transitions are called “forbidden”. Forbidden transitions include the additional notation of square brackets around

Table 3.4 *Transition selection rules*

	E1 dipole (permitted)	M1 dipole (forbidden)	E2 quadrupole (forbidden)	Additional rules or constraints
— general rules independent of coupling scheme —				
Parity	changes	unchanged	unchanged	E1: $J=0 \leftrightarrow J'=0$ forbidden
ΔJ	$0, \pm 1$	$0, \pm 1$	$0, \pm 1, \pm 2$	M1: $J=0 \leftrightarrow J'=0$ forbidden
ΔM_J	$0, \pm 1$	$0, \pm 1$	$0, \pm 1, \pm 2$	E2: $J+J' \geq 2$ required
— L-S coupling scheme —				
ΔL	$0, \pm 1$	0	$0, \pm 1, \pm 2$	E1: $L=0 \leftrightarrow L'=0$ forbidden
ΔM_L	$0, \pm 1$	any	$0, \pm 1, \pm 2$	E2: $L+L' \geq 2$ required
ΔS	0	0	0	
— L-S coupling scheme single active electron —				
Δl_i	± 1	0	$0, \pm 2$	E2: $l_i=0 \leftrightarrow l'_i=0$ forbidden
Δm_{l_i}	$0, \pm 1$	any	$0, \pm 1, \pm 2$	
Δs_i	0	0	0	
Δn_i	any	0	any	
— Probabilities and time scales —				
$A_k^{k'}$	$\sim 10^8$	$\sim 10^{-2}$	$\sim 10^{-2}$	[s ⁻¹]
$\tau_k^{k'}$	$\sim 10^{-8}$	$\sim 10^2$	$\sim 10^2$	[s]

the ion name, for example, in the infrared, the E2 transition [CII] 157 μm . The higher order terms in the electric field vector do not induce a parity change between the initial and final states. Thus E2 and M1 transitions have no parity change. In nature, forbidden transitions are exclusively observed in emission.

In Table 3.4, we summarize the selection rules for E1, M1, and E2 transitions. We also provide the order of magnitude of the transition rates. Once an electron is excited, a spontaneous E1 transition will occur within ~ 100 millionth of a second. In the case that the excited electron has no energy states to which it can transition via E1, the M1 or E2 spontaneous transitions may not take place for up to ~ 100 seconds.

For M1 (magnetic dipole) transitions, $\Delta J = J - J' = 0, \pm 1$ and $\Delta M_J = 0, \pm 1$, with the exception $\Delta J = 0$ when $J = 0$ and $J' = 0$. Also, the parity must remain the same. These conditions apply for all spin-orbit coupling schemes. M1 transitions take place only between fine-structure states *within* a multiplet. The transition energy is therefore very small, residing in the microwave and radio bands, and the transitions are very weak.

For E2 (electric quadrupole) transitions, $\Delta J = J - J' = 0, \pm 1, \pm 2$ and $\Delta M_J = 0, \pm 1, \pm 2$, with the constraint $J + J' \geq 2$. As with M1 transitions, the parity cannot change. A consequence of these selection rules is that

all transitions that have $\Delta J = \pm 2$, $\Delta L = \pm 2$, and/or $\Delta l_i = \pm 2$, are E2 transitions.

3.4 Isotope Shifts

Consider an atom or ion with finite nuclear mass m_N and N_e bound electrons. In an arbitrary coordinate system, the kinetic energy component of the Hamiltonian operator is

$$T = -\frac{\hbar^2}{2m_N} \nabla_{\mathbf{r}_N}^2 + \sum_{i=1}^{N_e} -\frac{\hbar^2}{2m_e} \nabla_{\mathbf{r}_i}^2 \quad (3.34)$$

In this coordinate system with arbitrary origin (not centered on the nucleus), the nucleus will “orbit” the origin at \mathbf{r}_N , and the i th electron will have radial vector \mathbf{r}_i with respect to the origin.

If we place the origin at the center of mass, both the nucleus and the electrons will “orbit” the center of mass. Following the coordinate transfer, we obtain

$$T = -\frac{\hbar^2}{2(m_N + N_e m_e)} \nabla_{\mathbf{q}_N}^2 - \frac{\hbar^2}{2\hat{m}} \sum_{i=1}^{N_e} \nabla_{\mathbf{q}_i}^2 - \frac{\hbar^2}{m_N} \sum_{i=1}^{N_e} \sum_{j=i+1}^{N_e} (\nabla_{\mathbf{q}_i} \cdot \nabla_{\mathbf{q}_j}) \quad (3.35)$$

The first term is the kinetic energy operator for the nucleus with respect to the center of mass. The second term is the sum of the kinetic energies of the electrons corrected for their reduced mass $\hat{m} = m_N m_e / (m_N + N_e m_e)$. The third term is known as the mass polarization, because it accounts for the motion of the nucleus with respect to the center of mass.

Table 3.5 *Isotope shifts for MgI and MgII*

Isotope	MgII $\lambda 2796$		MgII $\lambda 2803$		MgI $\lambda 2852$	
	E , eV	λ , Å	E , eV	λ , Å	E , eV	λ , Å
24 (79%)	4.434344	2796.3553	4.422992	2803.5324	4.346358	2852.9636
25 (10%)	4.434350	2796.3512	4.422998	2803.5282	4.346361	2852.9616
26 (11%)	4.434356	2796.3473	4.423004	2803.5244	4.346363	2852.9598

Data taken from Hallstadius (1979). All wavelengths are vacuum.

Analogous components to the mass polarization term can be derived for the Coulomb interactions and for the spin-orbit interactions, etc. These “mass polarizations” are the dominate terms that yield shifts in transition energies, and they are dominated by the mass of the nucleus. As such,

isotopes of the same atom or ion will have different transition energies in proportion to the nuclear mass.

In Table 3.5, we present selected example transition from MgI and MgII for three naturally occurring isotopes of magnesium, ^{24}Mg , ^{25}Mg , and ^{26}Mg . The isotopic abundances are 79%, 10%, and 11%, respectively. We present the MgII $\lambda\lambda 2796, 2803$ doublet ($^2\text{S}_{1/2} - ^2\text{P}_{3/2}$ and $^2\text{S}_{1/2} - ^2\text{P}_{1/2}$, respectively), and the MgI $\lambda 2852$ transition ($^1\text{S}_0 - ^1\text{P}_1^o$). Note that as the nuclear mass increases, the transition energy increases. The observed spectrum of these transition will be the isotopic abundance weighted values. Note that these energy shifts are significantly smaller than the shifts for hydrogenic atoms and ions, as discussed in § 1.5.

4

Atomic Properties and Spectra

The Russell-Saunders state terms, electronic configurations, and the origin of the patterns of the periodic table are all central to understanding the transitions and spectra of different elements and ions. We will show that the spectra of atoms and ions with “similar” electron configurations have common characteristics and can therefore be grouped. An understanding of the relationships within and between these groups is key for forging intuitive insight into the spectra of different elements on the periodic table.

We begin by illustrating how the ground-state Russell-Saunders term and state symbols of atoms and ions are determined. We show that atoms and ions with the same term and state symbols have the similar spectra but with different transition energies. We build the periodic table of elements and discuss the phenomenon of iso-electronic sequences and “iso-sequence ions”.

Finally, we present and discuss several Grotrian diagrams and the spectra of atoms and ions for which absorption transitions are commonly observed in astronomical and cosmological spectra.

4.1 Ground-State Russell-Saunders Symbols

As electrons are added to neutral atoms of ever increasing Z , the electrons occupy the lowest available energy state, which first and foremost is the least shielded subshell, which typically corresponds to the lowest nl . At a fixed Z , the electrons populate subshells in the order of increasing energy shown in Figure 3.4. We further illustrate this in the schematic provided in Figure 4.1(a). The arrows provide the order of occupation for the subshells; in each subshell panel the fraction provides the number of electrons that can occupy the subshell (numerator) to the total number of electrons bound in neutral atoms (denominator). As discussed in § 3.2.1 and shown in Table 3.2,

this simple pattern begins to break down at $Z = 24$ (chromium), when the 3d subshell is partially occupied.

In addition, the $m_l m_s$ ordering of occupation is dictated by Russell-Saunders L-S coupling. Recall that, in partially occupied subshells, electrons in different m_l states are more tightly bound and that this configuration occurs for the highest total spin state S as a result of the Pauli exclusion principle (see § 3.2.2). Recalling Hund's rule: "the lowest energy has the largest S , and for a given S has the largest L ." Stated another way, "the ground-state Russell-Saunders term symbol corresponds to the maximum S state having the maximum L ."

There are a set of simple rules for establishing the ground-state Russell-Saunders term and state symbols. They are based on the premise that we wish to maximize S and then maximize L for this maximum S .

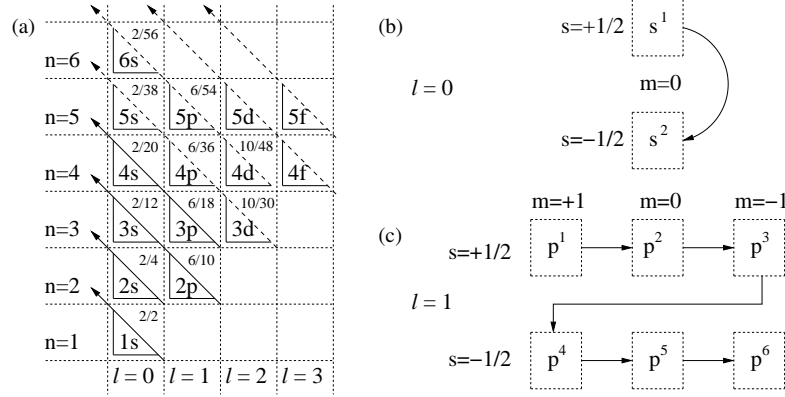


Figure 4.1 Schematics for the rules of populating shells and subshells for determining the ground-state Russell-Saunders term and state symbols. (a) Following the solid arrows, the order for populating shells is $1s \rightarrow 2s \rightarrow 2p \rightarrow 3s \rightarrow 3p \rightarrow 4s \rightarrow 3d$. Exceptions occur as the 3d shell is partially occupied, as indicated by the dash-lined arrows. (b) The s shells, which have only $m_l = 0$ projections of L_z , are populated with a spin-up electron followed by a spin-down electron. (c) The p shells, which have $m_l = +1, 0$, and -1 projections of L_z are populated with spin-up electrons in the order $m = +1, 0, -1$ and then with spin-down electrons in the order $m_l = +1, 0, -1$.

First, consider the simplest case of fully occupied shells and/or subshells, for which the vector sums are always $\mathbf{L} = 0$ and $\mathbf{S} = 0$. Therefore, $L = S = 0$ and the Russell-Saunders term symbol is 1S . Since $J = 0$, the Russell-Saunders state symbol is 1S_0 . For partially occupied subshells, we first wish to maximize S . Thus, we populate the electrons first with spin-up states and then spin-down states. We next wish to ensure that L is a maximum

for a given S , which we do by populating the electrons from the highest m_l to the lowest m_l . This technique works because L always must be equal to the maximum $|M_L|$ for a given permutation of the valence shell/subshell occupation.

n	l	m_l, m_s	L	S	J	$^{2S+1}L_J$	Element
n=4	3d $l=2$						
	4s $l=0$	(0, -1/2) $1s^2 2s^2 2p^6 4s^2$	0	0	0	1S_0	Ca
		(0, +1/2) $1s^2 2s^2 2p^6 4s$	0	1/2	1/2	$^2S_{1/2}$	K
		(-1, -1/2) $1s^2 2s^2 2p^6 3s^2 3p^6$	0	0	0	1S_0	Ar
n=3	3p $l=1$	(0, -1/2) $1s^2 2s^2 2p^6 3s^2 3p^5$	+1	1/2	3/2	$^2P_{3/2}$	Cl
		(+1, -1/2) $1s^2 2s^2 2p^6 3s^2 3p^4$	+1	1	2	3P_2	S
		(-1, +1/2) $1s^2 2s^2 2p^6 3s^2 3p^3$	0	3/2	3/2	$^4S_{3/2}$	P
		(0, +1/2) $1s^2 2s^2 2p^6 3s^2 3p^2$	+1	1	0	3P_0	Si
	3s $l=0$	(+1, +1/2) $1s^2 2s^2 2p^6 3s^2 3p$	+1	1/2	1/2	$^2P_{1/2}$	Al
		(0, -1/2) $1s^2 2s^2 2p^6 3s^2$	0	0	0	1S_0	Mg
		(0, +1/2) $1s^2 2s^2 2p^6 3s$	0	1/2	1/2	$^2S_{1/2}$	Na
		(-1, -1/2) $1s^2 2s^2 2p^6$	0	0	0	1S_0	Ne
n=2	2p $l=1$	(0, -1/2) $1s^2 2s^2 2p^5$	+1	1/2	3/2	$^2P_{3/2}$	F
		(+1, -1/2) $1s^2 2s^2 2p^4$	+1	1	2	3P_2	O
		(-1, +1/2) $1s^2 2s^2 2p^3$	0	3/2	3/2	$^4S_{3/2}$	N
		(0, +1/2) $1s^2 2s^2 2p^2$	+1	1	0	3P_0	C
	2s $l=0$	(+1, +1/2) $1s^2 2s^2 2p$	+1	1/2	1/2	$^2P_{1/2}$	B
		(0, -1/2) $1s^2 2s^2$	0	0	0	1S_0	Be
		(0, +1/2) $1s^2 2s$	0	1/2	1/2	$^2S_{1/2}$	Li
		(-1, -1/2) $1s^2$	0	0	0	1S_0	He
n=1	1s $l=0$	(0, +1/2) $1s$	0	1/2	1/2	$^2S_{1/2}$	H

Figure 4.2 The ground-state electronic configurations for hydrogen through calcium, $Z = 20$, showing the order of populating shells and subshells according to the rules illustrated in Figure 4.1. The total S , L , and J are tabulated for each configuration. They are each obtained by keeping the running sums of Eqs. 4.1 and the rules of Eq. 4.2 as each electron is populated.

A schematic of these rules is illustrated in Figure 4.1(b,c) for $l = 0$ and $l = 1$ subshells. The $l = 0$ subshell has only $m_l = 0$, so $L = M_L = 0$. This state is first populated by an $m_s = +1/2$ electron followed by an $m_s = -1/2$ electron. The $l = 1$ subshell has the m_l states -1 , 0 , and $+1$. The subshell is first populated with an electron in the $m_l = +1$ state with $m_s = +1/2$, yielding $L = M_L = 1$ and $S = 1/2$. Next, $m_l = 0$ with $m_s = +1/2$ is populated, yielding $L = M_L = 1$ and $S = 1$. Next, $m_l = -1$

with $m_s = +1/2$ is populated, yielding $L = M_L = 0$ and $S = 3/2$. The process is then repeated with spin-down electrons populating the subshell in the order $m_l = +1$ to $m_l = 0$ to $m_l = -1$. Note that this ordering always yields the largest L for a given S .

To obtain J for the ground state, we recall the Landé intervals. Subshells that are less than half occupied, the multiplets are “normal”, such that the minimum J corresponds to the lowest energy. Thus, we employ $J = |L - S|$. For half-occupied subshells $J = S$. For subshells that are more than half occupied, the multiplets are inverted so that the lowest energy corresponds to the largest J ; we employ the maximum J , which is $L + S$.

Following the order of populating the subshell as described above, the Russell-Saunders term and state symbols can be determined by the simple running sums that are incremented with each additional electron,

$$L = M_L = \sum_{i=1}^N m_{l_i}, \quad S = \sum_{i=1}^N m_{s_i}, \quad (4.1)$$

where N is the number of occupying electrons, and by the behavior of the Landé interval,

$$\begin{aligned} J &= |L - S| && \text{subshell less than half occupied} \\ J &= S && \text{subshell is half occupied } (L = 0) \\ J &= L + S && \text{subshell more than half occupied.} \end{aligned} \quad (4.2)$$

In Figure 4.2, we present the $nlm_l m_s$ states, and the iso-electron configurations for the first 20 elements (up to the 4s subshell). Also presented are the L , S , and J values and the Russell-Saunders state symbols, as computed using the running sums given by Eq. 4.1 and the Landé criterion given by Eq. 4.2. Note that helium, beryllium, neon, magnesium, argon, and calcium all have fully occupied subshells and thus have state symbols 1S_0 . Elements with half-occupied s subshells (e.g., hydrogen, lithium, sodium, and potassium) all have state symbols $^2S_{1/2}$. Elements with half-occupied p subshells (e.g., nitrogen and phosphorous) all have state symbols $^4S_{3/2}$.

4.2 The Periodic Table

An examination of Figure 4.2 reveals a repeating pattern that results from a correspondence between electronic sequence and multiplicity of energy states. If the elements are organized in tabular form with each row corresponding to a principle level n and columns aligned by electronic sequence, we find that the elements naturally organize into columns with equivalent

ground-state Russell-Saunders state symbols. Since the state symbols reflect the energy structure of the atom, this means that the energy structures, and therefore spectra of atoms, in a given column are similar. The overall arrangement of the elements is known as the periodic table. In Figure 4.3, we show a partial periodic table with the state terms provided across the top of each column. The atomic number (number of protons) and electronic sequences are also given for each element.

Since the overall LS term and fine-structure multiplicity is the same for elements of a given column, elements in a given column are considered to be similar members of a “group”. Below the periodic table in Figure 4.3, we list the groups by their designations under the Chemical Abstract Service (CAS) system. For example, in the first column of the periodic table are elements in the 1S_0 state with ns valence electronic sequences. These comprise Group IA and are also commonly called the “Alkali metals”. Since Group IA all have a single outer s subshell electron, and this outer electron will be the “optically active” electron, all ground-state Group IA metals will have spectra that share characteristics with those of hydrogenic atoms. For example, the ground-state $ns_{1/2}$ electron can undergo fine-structure doublet transitions analogous to the Lyman series.

It is convention to adopt spectroscopic notation for the ionization stages of each element when discussing absorption and emission features in spectra. Neutral atoms are denoted with the roman numeral “I”, so neutral hydrogen (H^0) is written HI. Triply ionized carbon (C^{+3}) is written CIV, etc. The hydrogenic ions are HI, HeII, LiIII, BeIV, BV, CVI, NVII, OVIII, etc. We will invoke this notation as appropriate throughout the remainder of this chapter.

4.2.1 Iso-Sequence Ions

The inter-related behavior of ions can also be directly read from the periodic table. As atoms are consecutively stripped of electrons they will have the state symbols and electronic sequences of consecutively lower Z neutral atoms. When a neutral atom is ionized, its electronic sequence now corresponds to the neutral atom lying one position to the left on the periodic table. With each successive ionization, the position of the resulting sequence lies successively leftward on the periodic table. Because the Russell-Saunders symbols, $^{2S+1}L_J$, of the ground state and excited states that an atom or ion can take on are dictated by the electron configuration of the valence electrons, atoms and ions with the same electronic sequence have

IA $^2S_{1/2}$												VIIIA 1S_0
1s H hydrogen 1												1s ² He helium 2
2s ² 2s Li lithium 3	2s ² 2s ² Be beryllium 4											
[Ne]3s Na sodium 11	[Ne]3s ² Mg magnesium 12											
[Ar]4s K potassium 19	[Ar]4s ² Ca calcium 20											
		IIIA $^2P_{1/2}^o$	IVA 3P_0	VA $^4S_{3/2}^o$	VIA 3P_2	VIIA $^2P_{3/2}^o$						
		2s ² 2s ² 2p B boron 5	2s ² 2s ² 2p ² C carbon 6	2s ² 2s ² 2p ³ N nitrogen 7	2s ² 2s ² 2p ⁴ O oxygen 8	2s ² 2s ² 2p ⁵ F fluorine 9	2s ² 2s ² 2p ⁶ Ne neon 10					
		[Ne]3s ² 3p Al aluminum 13	[Ne]3s ² 3p ² Si silicon 14	[Ne]3s ² 3p ³ P phosphorus 15	[Ne]3s ² 3p ⁴ S sulfur 16	[Ne]3s ² 3p ⁵ Cl chlorine 17	[Ne]3s ² 3p ⁶ Ar argon 18					
		[Ar]3d ¹⁰ 4s ² 4p Ga gallium 31	[Ar]3d ¹⁰ 4s ² 4p ² Ge germanium 32	[Ar]3d ¹⁰ 4s ² 4p ³ As arsenic 33	[Ar]3d ¹⁰ 4s ² 4p ⁴ Se selenium 34	[Ar]3d ¹⁰ 4s ² 4p ⁵ Br bromine 35	[Ar]3d ¹⁰ 4s ² 4p ⁶ Kr krypton 36					
		$^3D_{3/2}$	3F_2	$^4F_{3/2}$	7S_3	$^6S_{5/2}$	5D_4	$^4F_{9/2}$	3F_4	$^2S_{1/2}$	1S_0	
[Ar]3d 4s ² Sc scandium 21	[Ar]3d ² 4s ² Ti titanium 22	[Ar]3d ³ 4s ² V vanadium 23	[Ar]3d ⁴ 4s Cr chromium 24	[Ar]3d ⁵ 4s Mn manganese 25	[Ar]3d ⁶ 4s ² Fe iron 26	[Ar]3d ⁷ 4s ² Co cobalt 27	[Ar]3d ⁸ 4s ² Ni nickel 28	[Ar]3d ⁹ 4s Cu copper 29	[Ar]3d ¹⁰ 4s Zn zinc 30			
IIIB	IVB	VB	VIB	VII B	VIII B			IB	IIB			

Group	Name	Configuration	State	Elements
IA	Alkali metals	ns^1	$^1S_{1/2}$	Li, Na, K, Rb, ...
IIA	Alkali-earth metals	ns^2	1S_0	Be, Mg, Ca, Sr, ...
IIIA	Boron group	ns^2np^1	$^2P_{1/2}^o$	B, Al, Ga, In, ...
IVA	Carbon group	ns^2np^2	3P_0	C, Si, Ge, Sn, ...
VA	Nitrogen group	ns^2np^3	$^4S_{1/2}^o$	N, P, As, Sb, ...
VIA	Oxygen group	ns^2np^4	3P_2	O, S, Se, Te, ...
VIIA	Fluorine group	ns^2np^5	$^2P_{3/2}^o$	F, Cl, Br, I, ...
VIIIA	Nobel gases	ns^2np^6	1S_0	He, Ne, Ar, Kr, Xe, ...
IIIB	Scandium group	$nd^1n's^2$	$^2D_{3/2}$	Sc, Y, ...
IVB	Titanium group	$nd^2n's^2$	3F_2	Ti, Zr, ...
VB	Vanadium group	$nd^3n's^2$	$^4F_{3/2}$	V, Nb, ...
VIB	Chromium group	$nd^5n's^1$	7F_3	Cr, Mo, ...
VII B	Manganese group	$nd^5n's^2$	$^6F_{5/2}$	Mn, Tc, ...
VIII B	Iron group	$nd^6n's^2$	5D_4	Fe, Ru, ...
VIII B	Cobalt group	$nd^7n's^2$	$^4F_{9/2}$	Co, Rh, ...
VIII B	Nickel group	$nd^8n's^2$	4F_4	Ni, Pd, ...
IB	Coinage metals	$nd^{10}n's^1$	$^2S_{1/2}$	Cu, Ag, ...
IIB	Volatile metals	$nd^{10}n's^2$	1S_0	Zn, Cd, Hg, ...

Figure 4.3 A partial periodic table of elements from hydrogen to krypton ($Z = 36$). Note that the elements form groups with the same valence electronic sequence and ground-state Russell-Saunders state symbol. Thus, the table immediately informs us that there are similarities in the characteristics of the energy structure and energy transitions (and thus observed spectral features) exhibited by elements in the same group.

characteristics. Ionized three times, SiIV is sodium-like and exhibits Group IA spectral characteristics.

In Figure 4.4, we present a schematic of ions with the same electronic sequence (iso-sequence ions), which are indicated by the dotted lines. Scanning to the right along the periodic table, we find that BeII, BIII, CIV, NV, OVI, FVII, and NeVIII are all lithium-like, and that MgII, AlIII, SiIV, PV, SVI, ClVII, and ArVIII are all sodium-like. These ions will exhibit spectra characteristic of Group IA Alkali metals (which we describe in more detail in §4.4). Ions for which transitions are commonly seen in absorption in astronomical and cosmological spectra are indicated in boldface font on the schematic in Figure 4.4. These ions are more frequently observed due to a combination of their high abundance in the universe, their internal energy structure, and the large number of astrophysical environments that have favorable ionization conditions. An additional consideration is also whether the observed spectral features (including redshifted features) reside in a wavelength range for which spectroscopic facilities exist and are sufficiently sensitive.

4.3 Binding Energy and Ionization Potential

The central field approximation, H_1 (see § 3.2.1), yields the basic physical and energy structure of multi-electron atoms and ions. The inclusion of L-S coupling and spin-orbit coupling modifies the energy structure of the central field approximation, but only by small shifts and splitting. Since the ionization energy is the energy required to overcome the binding energy, the ionization energy is the negative (or absolute value) of the binding energy.

In Table 4.1, we present the first eight ground-state ionization potentials for the first 30 elemental species (hydrogen through zinc). If a given atom or ion is in an excited state, the ionization potential will differ from the values listed. The energy units are electron volts [eV]. All values are provided to the accuracy published by the National Institute of Standards and Technology (NIST, Kramida et al., 2014)

A graphical representation of the ionization potentials listed in Table 4.1 is presented in Figure 4.5 in order of atomic number (Z). The shaded elements are those for which transitions from several ionization levels are very commonly observed in ultraviolet and optical spectra. The ionization potential of HI is emphasized with the dashed horizontal line at 13.598 eV. Note that several elemental species have their first ionization at energies less than that of HI. A larger ionization potential equates to a more tightly bound electron and vice versa.

The overall pattern in the ground-state ionization potentials from ion to

Table 4.1 *Ground state ionization potentials (eV)*^a

<i>Z</i>	Element	I	II	III	IV	V	VI	VII	VIII
1	H.....	13.5984340							
2	He.....	24.5873876	54.4177630						
3	Li.....	5.3917191	75.64015	122.454353					
4	Be.....	9.322699	18.211153	153.8944	217.718572				
5	B.....	8.298026	25.15483	37.390620	259.37510	340.225993			
6	C.....	11.26030	24.38331	47.88778	64.49390	392.0886	489.99312		
7	N.....	14.53413	29.601287	47.44922	77.4735	97.89013	552.0716	667.04602	
8	O.....	13.618054	35.12111	54.93550	77.41350	113.8989	138.1196	739.29293	871.40969
9	F.....	17.42282	34.97081	62.7084	87.1398	114.243	157.1631	185.1868	953.9068
10	Ne.....	21.564539	40.96296	63.42331	97.1168	126.2159	157.9311	207.2706	239.0969
11	Na.....	5.139075	47.28636	71.6200	98.91	138.40	172.183	208.50	264.25
12	Mg.....	7.646231	15.035255	80.1436	109.265	141.270	186.76	225.02	264.96
13	Al.....	5.985755	18.82857	28.44764	119.992	153.8252	190.477	241.76	284.66
14	Si.....	8.151683	16.345845	33.49300	45.14179	166.7674	205.2670	246.4806	303.5381
15	P.....	10.48669	19.76947	30.20264	51.44387	65.02511	220.422	263.57	309.60
16	S.....	10.3600	23.33788	34.7900	47.2219	72.59449	88.05292	280.948	328.74
17	Cl.....	12.96763	23.8136	39.9105	53.4652	67.8194	97.0300	114.201	348.2840
18	Ar.....	15.759606	27.62965	40.7350	59.6860	75.1344	91.00	124.3279	143.4567
19	K.....	4.34066345	31.6284	45.8060	60.9134	82.66	99.4	117.6	154.88
20	Ca.....	6.113158	11.87172	50.9131	67.27	84.51	108.8	127.2	147.24
21	Sc.....	6.56149	12.79977	24.756837	73.4894	91.69	110.7	138.0	158.1
22	Ti.....	6.8204	13.5755	27.49171	43.26717	99.299	119.533	140.8460	170.4
23	V.....	6.7398	14.6549	29.3111	46.7086	65.2816	128.1253	150.6408	173.4539
24	Cr.....	6.76651	16.4857	30.9589	49.1597	69.4559	90.6349	160.1752	185
25	Mn.....	7.434018	15.63999	33.668	51.2	72.4	95.60	119.203	194.5
26	Fe.....	7.9024	16.1877	30.6514	54.8010	75.0104	99.0634	124.9761	151.0599
27	Co.....	7.8643	17.0844	33.50	51.27	79.5	102	129	157.8319
28	Ni.....	7.6374	18.168837	35.1867	54.9250	76.06	107.87	133	162
29	Cu.....	7.726377	20.29239	36.8407	57.3799	79.8458	103.031	138.8623	...
30	Zn.....	9.394199	17.96439	39.7233	59.5732	82.5735	133.9029	160.1752	173.9498

^a Source: National Institute of Standards and Technology (NIST) (www.nist.gov/)
All data quoted to the accuracy given by NIST

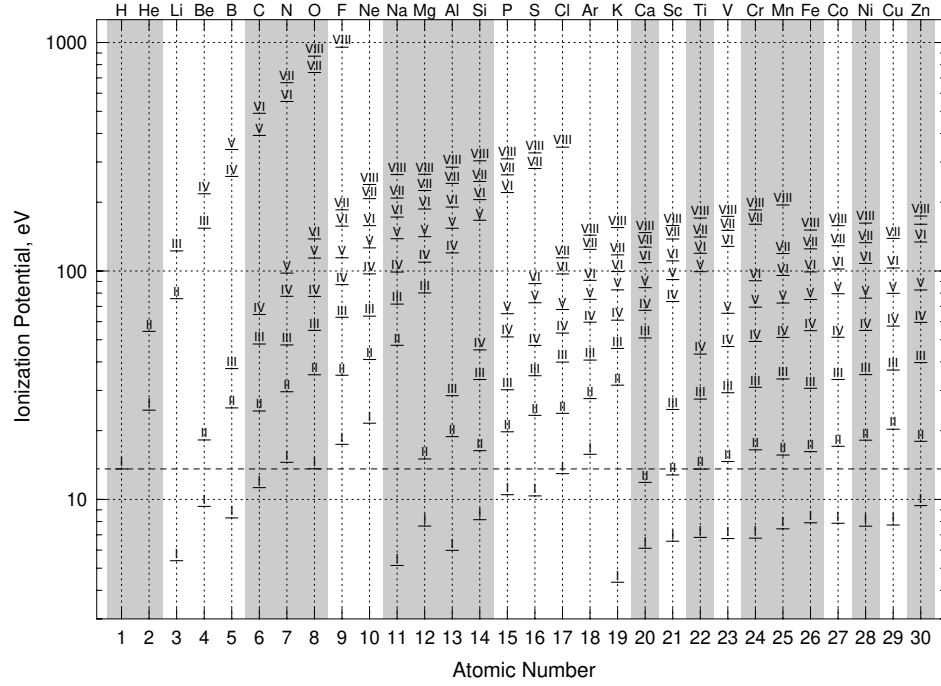


Figure 4.5 A visual representation of the ground-state ionization potentials presented in Table 4.1 for the first thirty elements on the periodic table. Elements of astrophysical interest are shaded. The dashed horizontal line highlights the ionization potential of neutral hydrogen and the dotted horizontal lines are to guide the eye at each order of magnitude.

ion can be qualitatively understood in terms of the nl dependence on the binding energies. Consider ionization from neutral atoms (I). The ionization potentials increase with increasing Z until a shell is fully occupied with electrons, and then the ionization potential decreases significantly as Z is incremented such that the next electron is a single electron in the next subshell. Note on Figure 4.5 the increase in ionization potential from H I to He I for the $1s$ and $1s^2$ electron configurations, and then the sudden decrease in ionization potential from He I to Li I, which has its outer electron in the $2p$ subshell. Also note the increase from Li I to Ne I as the $2p$ shell is incrementally occupied. This is because the nuclear charge is increasing, but the shielding of the electrons is negligibly changed, so the binding energy increases. Then, from Ne I ($1s^2 2s^2 2p^6$) to Na I ($1s^2 2s^2 2p^6 3s$) the next electron occupies the s subshell in the next shell. This electron is well shielded by the core of electrons in the $n = 1$ and $n = 2$ shells, and so the binding energy is much lower. This pattern continues from Na I to K I.

Thus, for atoms with partially occupied subshells, we see that as Z is incremented (increasing the Coulomb attraction) and an additional electron populates the partially occupied subshell, the binding energy increases. This is because electrons in the same subshell are roughly at the same distance from the nucleus and do not effectively shield each other from the nucleus. That is, as Z is incremented by $+1$, the nuclear attraction increases the binding energy, but the least-bound electron is no better shielded than the electrons that already occupy the subshell. However, in the case where Z is incremented and the added electron is forced to occupy the next higher shell, the outermost electron is at a greater radial distance and is thus more effectively shielded from the nucleus by the fully occupied subshell beneath it. Thus, its binding energy is lower than the electrons in the subshell beneath even though it is bound to a higher Z atom.

Now, note that the same pattern exists for singly ionized ions (II), but that the decrease in the ionization potential is offset by $Z + 1$ relative to the neutral (I) atoms. The pattern continues with increasing ionization stage; for triply ionized ions, the significant decrease in the ionization potential is offset by $Z + 1$ with respect to the singly ionized ions, and by $Z + 2$ from the neutral atoms.

4.4 Grotrian Diagrams and Spectra

Partial Grotrian diagrams and accompanying tables are presented for several ions for which absorption transitions are commonly seen in astronomical spectra. As such, the diagrams emphasize only a limited number of energy levels, states, and transitions. In Table 4.2, we provide a key to the diagrams. An excellent resource for diagrams covering additional transitions and elements and ions is the document “Partial Grotrian Diagrams of Astrophysical Interest” compiled by Moore & Merrill (1968).

On each diagram, the vertical axis is excitation energy increasing upward; all energies are given in eV. We break from the convention of providing a generic energy scale and instead provide the energies of the states themselves. The horizontal axis provides the Russell-Saunders terms. The states themselves are represented as thick horizontal lines aligned with their Russell-Saunders term and excitation energy. Many of the states are multiplets, but only the general nl level (subshell) is indicated. The multiplets have different J , which are not resolved on the diagrams but are indicated in the accompanying tables.

The solid, dashed and dotted lines between states represent the transitions. Solid and dashed line transitions are labeled “A”, “B”, etc and a

Atomic Properties and Spectra
Table 4.2 *Key to Grotrian diagrams*

Diagram	Group/Level	Iso-sequence	Ions
Fig. 4.6	IA $n_0 = 2$	[He]2s	Li I, C IV, N V, O VI, Ne VIII
Fig. 4.7	IA $n_0 = 3$	[Ne]3s	Na I, Mg II, Al III, Si IV
Fig. 4.8	IA $n_0 = 4$	[Ar]4s	Ca II
Fig. 4.10	IIA $n_0 = 1$	1s ²	He I
Fig. 4.11	IIA $n_0 = 2$	[He]2s ²	C III, N IV, O V
Fig. 4.12	IIA $n_0 = 3$	[Ne]3s ²	Mg I, Al II, Si III
Fig. 4.15	IIIA $n_0 = 2$	[He]2s ² 2p	C II, N III, O IV
Fig. 4.16	IIIA $n_0 = 3$	[Ne]3s ² 3p	Al I, Si II
Fig. 4.17	VIIB $n = 3, 4$	[Ar]3d ⁶ 4s	Fe II
Fig. 4.19	VIIB $n = 3, 4$	[Ar]3d ⁹	Ni II
Fig. 4.20	VIIB $n = 3, 4$	[Ar]3d4s ²	Ti II
Fig. 4.21	VB $n = 3, 4$	[Ar]3d ⁵	Cr II
Fig. 4.21	IB $n = 3, 4$	[Ar]3d ⁵ 4s	Mn II

key is provided in the accompanying tables. For each transition, the table provides the valence electron configuration, the transition states, and the vacuum wavelength of the spectral feature.

All presented diagrams apply to the condition in which a single active valence electron participates in the transition. Thus, ions with the same iso-electronic sequence will share a common set of transitional states. We therefore employ a single “representative” diagram for several ions having the same iso-electronic sequence. While this reduces the total number of diagrams we present, it has the drawback that the relative energy spacings from ion to ion are not accurately illustrated on the diagram. The energy levels presented on each diagram apply for the lowest Z atom or ion (with the exception of the IIIA $n_0 = 4$ group). For the higher Z ions, the illustrated energy spacing will differ, but the general overall structure is preserved. Exceptions are noted in the captions. The term-averaged excitation energies of represented ions are given the right of each diagram.

4.4.1 Group IA: Alkali Metals

Group IA Alkali metals include lithium (Li, $Z = 3$), sodium (Na, $Z = 11$), potassium (K, $Z = 19$), rubidium (Rb, $Z = 37$), cesium (Cs, $Z = 55$), and francium (Fr, $Z = 87$). Ions that have the iso-sequences of lithium, sodium, or potassium are called lithium-like, sodium-like, and potassium-like, respec-

tively. Commonly observed lithium-like ions include Be II, C IV, N V, O VI, and Ne VIII. Sodium-like ions include Mg II, Al III, Si IV, P V, and S VI. The most commonly observed potassium-like ion is Ca II. The energy structures and transitions are illustrated in Figures 4.6, 4.7, and 4.8, respectively.

The ground-state structure of Group IA atoms and ions is a single valence electron orbiting in the s subshell just outside closed shells. For lithium and lithium-like ions, the closed shell is $1s^2$ and the valence electron occupies the 2s subshell. For sodium and sodium-like ions, the closed shell is $1s^2 2s^2 2p^6$ and the valence electron occupies 3s. For potassium and potassium-like ions, the core is $1s^2 2s^2 2p^6 3s^2 3p^6$ and the valence electron occupies 4s. From the point of view of the single valence electron, these inner closed shells comprise a bound core of electrons for which $L = S = J = 0$. The vector sum rules (Eq. 3.3) for the states of the atom therefore greatly simplify so that $L = l$, $S = s$, and $J = j$, where l , s , and j are the quantum numbers of the valence electron. The ground state for all Alkali metals and ions is thus $L = 0$, $S = 1/2$, and $J = 1/2$, giving the state symbol $^2S_{1/2}$. Because the vector addition reduces to that of a single electron for the Alkali metals, the hydrogenic dipole selection rules can be adopted, i.e., $\Delta l = \pm 1$, $\Delta j = 0, \pm 1$, and Laporte's rule, $\Delta s = 0$.

The energy levels have a strong l dependence that results in large energy differences between the subshells of the lowest lying valence shell. As can be seen in Figures 4.6, 4.7, and 4.8, this yields an especially large energy difference between the ground state n_0s and the lowest lying subshell n_0p , where $n_0 = 2, 3$, and 4 for lithium-like, sodium-like, and potassium-like ions, respectively. For higher excited states ($n > n_0$), for which the valence electron resides at a larger average distance from the electron core, the energy separations of subshells within a shell converge toward l degeneracy, so that the s, p, d, and f subshells are very close to being energy degenerate as n increases.

The pattern in the spectra of Alkali metals and ions is a series with doublets and triplets similar (in principle) to that of hydrogenic atoms. Absorption transitions from the ground state (which has excitation energy 0 eV) are called 0-volt transitions, the strongest for which the active valence electron quantum states transition between $n_0s_{1/2} - n_0p_{1/2}$ and $n_0s_{1/2} - n_0p_{3/2}$. Note that $\Delta n = 0$. These particular doublets are known as 0-volt resonance doublets. In Figures 4.6, 4.7, and 4.8, and their accompanying tables, the 0-volt resonant doublets are labeled "A". Of these, the Mg II $\lambda\lambda 2796, 2803$, C IV $\lambda\lambda 1548, 1550$, and O VI $\lambda\lambda 1031, 1037$ doublets are among the most commonly observed transitions in astronomical and cosmological spectra.

The fine structure splittings are too small to resolve on the Grotrian dia-

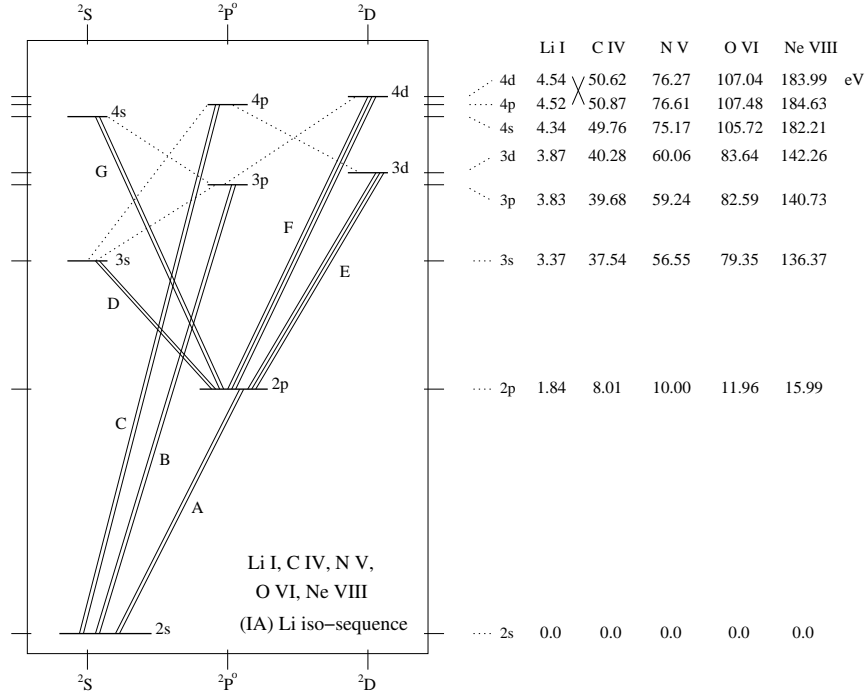


Figure 4.6 Partial Grotrian diagram for Group IA $n_0 = 2$ ions (lithium-like; $1s^2 2s$). Excitation energies for each state are given to the right. For multiplicity states, the averaged excitation energies are given. The 0-volt doublets are A, B, and C. The A doublets are 0-volt resonance transitions. The first excited state doublets are D and G, and the triplets are E and F. Note that the 4d and 4p subshell energies for the higher Z ions are inverted relative to lithium. The transitions and transition wavelengths are listed in the accompanying table. All data are obtained from Kramida et al. (2014). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	Li I λ , Å	C IV λ , Å	N V λ , Å	O VI λ , Å	Ne VIII λ , Å
A	2s–2p	$^2S_{1/2}$ – $^2P_{3/2}$	6709.61	1548.187	1238.821	1031.912	770.409
		$^2S_{1/2}$ – $^2P_{1/2}$	6709.76	1550.772	1242.804	1037.613	780.324
B	2s–3p	$^2S_{1/2}$ – $^2P_{3/2}$	3233.590	312.420	209.274	150.089	88.080
		$^2S_{1/2}$ – $^2P_{1/2}$	3233.590	312.451	209.308	150.125	88.118
C	2s–4p	$^2S_{1/2}$ – $^2P_{3/2}$	2742.014	244.903	162.556	115.821	67.380
		$^2S_{1/2}$ – $^2P_{1/2}$	2742.014	244.911	162.564	115.830	67.389
D	2p–3s	$^2P_{1/2}$ – $^2S_{1/2}$	8128.45	419.525	266.196	183.937	102.909
		$^2P_{3/2}$ – $^2S_{1/2}$	8128.68	419.714	266.379	184.117	103.084
E	2p–3d	$^2P_{1/2}$ – $^2D_{3/2}$	6105.22	384.031	247.561	172.935	98.117
		$^2P_{3/2}$ – $^2D_{5/2}$	6105.33	384.174	247.706	173.079	98.262
		$^2P_{3/2}$ – $^2D_{3/2}$	6105.35	384.190	247.719	173.095	98.276
F	2p–4d	$^2P_{1/2}$ – $^2D_{3/2}$	4604.110	289.141	186.063	129.785	73.478
		$^2P_{3/2}$ – $^2D_{5/2}$	4604.178	289.228	186.149	129.871	73.564
		$^2P_{3/2}$ – $^2D_{3/2}$	4604.182	289.231	186.152	129.875	73.567
G	2p–4s	$^2P_{1/2}$ – $^2S_{1/2}$	4973.046	296.856	190.155	132.219	74.545
		$^2P_{3/2}$ – $^2S_{1/2}$	4973.130	296.951	190.249	132.312	74.637

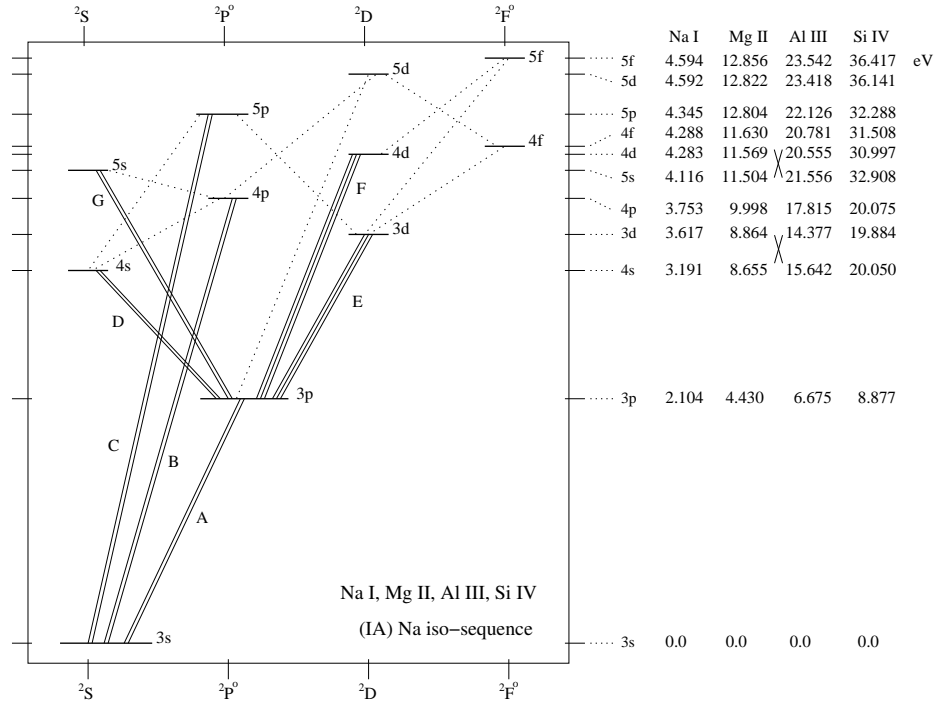


Figure 4.7 Same as for Figure 4.6 but for Group IA $n_0 = 3$ (sodium-like) ions with iso-electronic sequence $1s^2 2s^2 2p^6 3s$. The A doublets are 0-volt resonance transitions. Note that the 4s and 3d subshell energies and the 4d and 5s subshell energies are inverted as one applies the diagram for AlIII and SiIV transitions. All data are obtained from Kramida et al. (2014) except MgII transition A and B and AlIII transition A wavelengths, which are obtained from Aldenius, Johansson, & Murphy (2006) for MgII and from Griesmann & Kling (2000) for AlIII. All wavelengths are vacuum.

Key	Valence Config.	Multiplet	NaI λ , Å	MgII λ , Å	AlIII λ , Å	SiIV λ , Å
A	3s–3p	$^2S_{1/2}$ – $^2P_{3/2}$	5891.583	2796.354	1854.718	1393.755
		$^2S_{1/2}$ – $^2P_{1/2}$	5897.558	2803.532	1862.791	1402.770
B	3s–4p	$^2S_{1/2}$ – $^2P_{3/2}$	3303.319	1239.925	695.828	457.815
		$^2S_{1/2}$ – $^2P_{1/2}$	3303.929	1240.394	696.217	458.155
C	3s–5p	$^2S_{1/2}$ – $^2P_{3/2}$	2853.649	1025.968	560.317	361.560
		$^2S_{1/2}$ – $^2P_{1/2}$	2853.850	1026.113	560.433	361.659
D	3p–4s	$^2P_{1/2}$ – $^2S_{1/2}$	11384.557	2929.490	1379.670	815.053
		$^2P_{3/2}$ – $^2S_{1/2}$	11406.888	2937.369	1384.132	818.128
E	3p–3d	$^2P_{1/2}$ – $^2S_{3/2}$	8185.505	2791.600	1605.776	1122.485
		$^2P_{3/2}$ – $^2S_{5/2}$	8197.043	2798.755	1611.814	1128.325
		$^2P_{3/2}$ – $^2S_{3/2}$	8197.077	2798.823	1611.873	1128.340
F	3p–4d	$^2P_{1/2}$ – $^2S_{3/2}$	5684.210	1734.852	892.024	559.533
		$^2P_{3/2}$ – $^2S_{5/2}$	5689.771	1737.613	893.887	560.980
		$^2P_{3/2}$ – $^2S_{3/2}$	5689.783	1737.613	893.896	560.980
G	3p–5s	$^2P_{1/2}$ – $^2S_{1/2}$	6155.928	1750.663	855.034	515.118
		$^2P_{3/2}$ – $^2S_{1/2}$	6162.452	1753.474	856.745	516.344

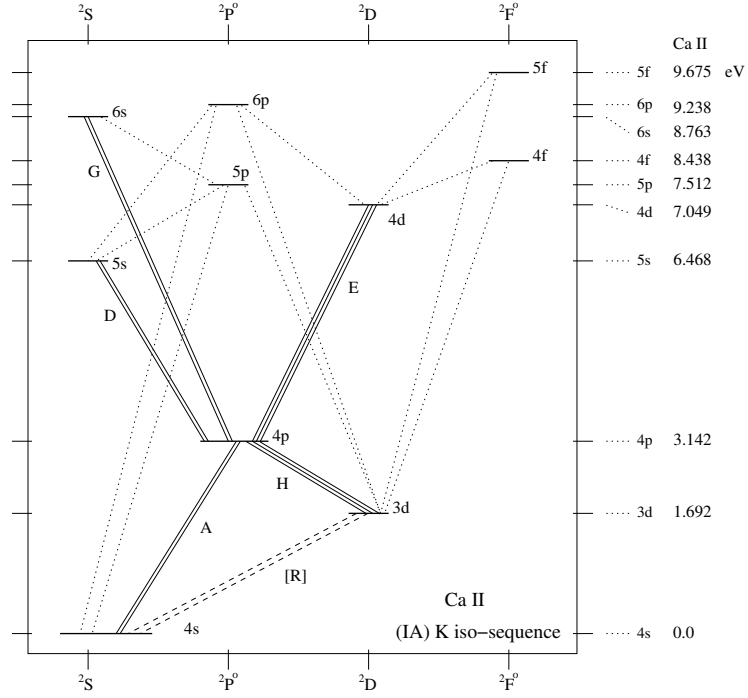


Figure 4.8 Same as for Figure 4.6 but Ca II, which has the Group IA $n_0 = 4$ (potassium-like) iso-electronic sequence $1s^2 2s^2 2p^6 3s^2 3p^6 4s$. The A doublets are 0-volt resonance transitions. Doublet R is forbidden by the dipole selection rules; it occurs via the electric quadrupole moment, E2. All data are obtained from Kramida et al. (2014). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	Ca II λ , Å
A	4s–4p	$^2S_{1/2} - ^2P_{3/2}^o$	3934.777
		$^2S_{1/2} - ^2P_{1/2}^o$	3969.592
D	4p–5s	$^2P_{1/2}^o - ^2S_{1/2}$	3707.078
		$^2P_{3/2}^o - ^2S_{1/2}$	3737.964
E	4p–4d	$^2P_{1/2}^o - ^2D_{3/2}$	3159.783
		$^2P_{3/2}^o - ^2D_{5/2}$	3180.251
		$^2P_{3/2}^o - ^2D_{3/2}$	3182.195
G	4p–6s	$^2P_{1/2}^o - ^2S_{1/2}$	2198.473
		$^2P_{3/2}^o - ^2S_{1/2}$	2209.300
H	3d–4p	$^2D_{3/2} - ^2P_{3/2}^o$	8500.35
		$^2D_{5/2} - ^2P_{3/2}^o$	8544.44
		$^2D_{3/2} - ^2P_{1/2}^o$	8564.52
[R]	4s–3d	$^2S_{1/2} - ^2D_{5/2}$	7293.48
		$^2S_{1/2} - ^2D_{3/2}$	7325.91

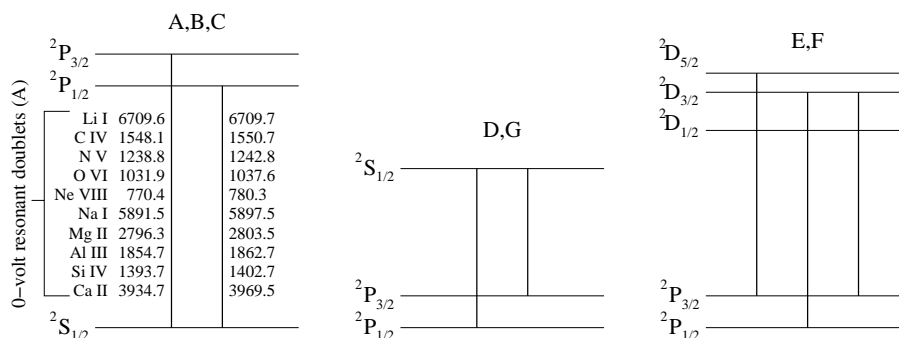


Figure 4.9 Schematics of the fine-structure splitting for selected transitions of the Alkali metals as labeled and presented in Figures 4.6, 4.7, and 4.8. Multiplet A is the 0-volt resonant doublet $(n_0s)^2S_{1/2}-(n_0p)^2P_{3/2}$ and $(n_0s)^2S_{1/2}-(n_0p)^2P_{1/2}$, for which the wavelengths of selected lithium-like, sodium-like, and potassium-like ions are given.

grams. We show a schematic of the energy splittings in Figure 4.9 for multiplets A, B, and C (left), D and G (center), and E and F (right), as labeled on the Grotrian diagrams, Figures 4.6–4.8. Vertical lines provide the individual transitions of the multiplets. For the 0-volt resonant doublets (multiplet A), we list the wavelength of each transition for commonly observed absorption lines from Alkali metals and Alkali-like ions.

4.4.2 Group IIA: Alkali-Earth Metals (and Helium)

Group IIA Alkali-earth metals include beryllium (Be, $Z = 4$), magnesium (Mg, $Z = 12$), calcium (Ca, $Z = 20$), strontium (Sr, $Z = 38$), barium (Ba, $Z = 56$, and radium (Ra, $Z = 88$). Though helium (He, $Z = 2$) is conventionally grouped with the noble gases (Group VIIIA), it arguably belongs with the Alkali-earth metals as well. A commonly observed beryllium-like ion is CIII, and less commonly NIV and OV. Commonly observed magnesium-like ions include AlII and SiIII.

The ground-state structure of these atoms is two valence electrons orbiting in an s subshell just outside inner closed shells. In the case of helium there is no inner shell; the valence shell is $1s^2$. For beryllium and beryllium-like ions, the closed shell is $1s^2$ and the valence electrons occupy the $2s$ subshell. For magnesium and magnesium-like ions, the closed shell is $1s^22s^22p^6$ and the valence electrons occupy the $3s$ subshell. For calcium and calcium-like ions, the core is $1s^22s^22p^63s^23p^6$ and the valence electrons occupy the $4s$ subshell.

There is a great number of possible states if all cases in which both elec-

trons can be excited are considered. In typical cosmological environments, however, the atoms and ions spend the majority of their time in the ground state and their transitions are such that only one of the valence electrons is active in transitions. We will first consider states in which one electron is inactive and the second undergoes transitions. For this assumption, the energy structure for neutral helium is illustrated in Figure 4.10, and partial Grotrian diagrams for beryllium-like ions and magnesium-like ions are given in Figures 4.11 and 4.12, respectively.

Helium and Helium-like Ions

The spectra of Group IIA Alkali-earth metals and similar ions can be understood by first considering helium. With two valence electrons, the total spin is either $S = 0$ or $S = 1$. Thus, the multiplicities are either singlets or triplets. Recall that the parity of the atom changes during a dipole transition, which dictates $\Delta L = \pm 1$. Thus, per Laporte's rule, $\Delta S = 0$ is enforced on transitions governed by the dipole selection rules. This means that singlet states transition only to other singlet states and triplet states transition only to other triplet states ($S + 1$ is preserved). The Grotrian diagram for helium transitions can therefore be divided into two independent diagrams, one with $S = 0$ (singlets) and one with $S = 1$ (triplets).

In Figure 4.10, we show the Grotrian diagram for helium for the condition in which one electron is in the 1s shell and the second electron is excited to some general subshell nl .

The ground state is the 0-volt singlet $1s^2$ with symbol 1S_0 . For dipole transitions, the first excited singlet state, $1s2p$ (1P_0), is roughly 21 eV above the ground state. This singlet state is at slightly higher energy than the first excited $1s2p$ triplet (3P_0 , 3P_1 , 3P_2), for which the multiplet is unresolved in Figure 4.10. As previously mentioned, due to Laporte's rule, $\Delta S = 0$, helium cannot transition from the singlet ground-state to a triplet excited state (which requires a spin flip) by absorption of a photon¹. Absorption, being most common from the ground state, is therefore almost exclusively observed as singlet transitions. In the absorption spectrum of neutral helium (and helium-like ions), there is a full series (similar to the hydrogenic series) of 0-volt $(1s)S_0-(np)P_1$ transitions. Note, however, that unlike the doublets in the hydrogenic ground-state series, the helium series lines are singlets.

As Z increases for helium-like ions, L-S coupling becomes a progressively

¹ Such an excitation would require a collisional excitation. Collisional excitation requires an astrophysical environment with high density and temperature given the large excitation energy of the 2p suborbital.

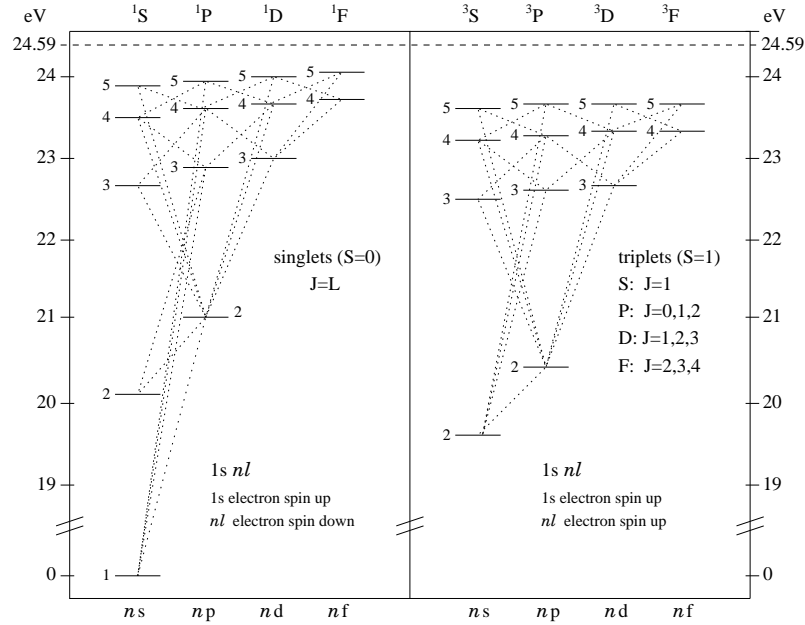


Figure 4.10 Partial helium Grotrian diagram for the electron configuration $1s nl$. Due to Laporte's rule, which dictates $\Delta S = 0$, transitions between triplet and singlet are not allowed under the electric dipole selection rules. Thus, helium can be treated as having two independent energy structures, those with $S = 0$ (singlet states) and those with $S = 1$ (triplet states). The excitation energies are given on the vertical axis. The J values for each LS term are listed in the legends. The 0-volt ground-state is a singlet. Absorption lines from helium are almost exclusively due to singlet transitions from the ground state. Note that the first excited state is highly elevated at ~ 20 eV.

less accurate description of the atomic states. This means that L and S are not as well defined for such ions and coupling schemes intermediate to L-S and $j-j$ coupling must be invoked. As such, LaPorte's rule is less strictly adhered to as Z increases for helium-like ions and for other Alkali-earth atoms and like ions. Emission lines from transitions between singlet and triplet states are extremely rare and weak for helium, and to a lesser extent, for higher Z helium-like ions and Group IIA-like ions. These spectral lines are called intercombination lines; they are not dipole transitions. They are virtually non-existent in absorption.

Beryllium-like and Magnesium-like Ions

As with helium and helium-like ions, Group IIA-like ions have either singlet or triplet multiplicities and dipole transitions take place between singlets state or between triplet states. As mentioned above, intercombination lines have been seen in emission for higher Z atoms and ions, but are virtually non-existent in absorption. Since the ground states are singlets, the absorption spectra of Group IIA-like ions are virtually all singlet transitions. We thus consider only the singlet energy configurations and spectra of these atoms and ions.

Similar to helium, a large energy difference between the ground state and the first excited state (between the s and p subshells in the ground shell, n_0) is a common characteristic of the Alkali-earth metals and their iso-sequence ions. This energy difference decreases with increasing Z . In Figures 4.11 and 4.12, we provide partial Grotrian diagrams for the beryllium-like ions CIII, NIV, and OV and for magnesium (MgI) and the magnesium-like ions AlII and SiIII.

The first excited state of these ions is $n_0s n_0p$, and the transition that occurs between this state and the ground state is $n_0s^2 - n_0s n_0p$. That is, single valence-electron transitions between $n_0s - n_0p$ while the second remains in the n_0s state. Per LaPorte's rule, there is no spin flip. The atomic state change is $^1S_0 - ^1P_1^o$. This transition is labeled "A" on the Grotrian diagrams; it comprises the most commonly observed absorption lines from CIII, NIV, OV, MgI, AlII, and SiIII. The wavelengths of these 0-volt transitions are given in the tables accompanying the partial Grotrian diagrams. The transition labeled "B" is the identical state change as transition A except the active valence electron transitions between $n_0s - (n_0 + 1)p$. Clearly, there is a series of singlet transitions $n_0s - (n_0 + m)p$, where $m = 1, 2, 3, \dots, \infty$, all having the state transition $^1S_0 - ^1P_1^o$.

In Figure 4.11, we also consider selected transitions for excited states. These are transitions C, D, and E, which have lower-state valence configurations $2s2p$ and excited states $2p^2$ and $2s3s$ respectively. In Figure 4.13, we illustrate the electron occupation for both the 0-volt transition A and the excited transition C. In the accompanying table, we tabulate the occupation for all transitions presented in Figure 4.11.

Examination of Figure 4.12 ($n_0 = 3$) reveals that singlet $3s3p - 3p^2$ transitions are not represented for Group IIA ions MgI, AlII, and SiIII. In fact, there is one singlet transition $(2p^6 3s^2) ^1S_0 - (2p^5 3s 3p^2) ^1P_2^o$ that involves two active electrons. This highlights an important difference between beryllium-like and magnesium-like ions; the more complex electron core inside the

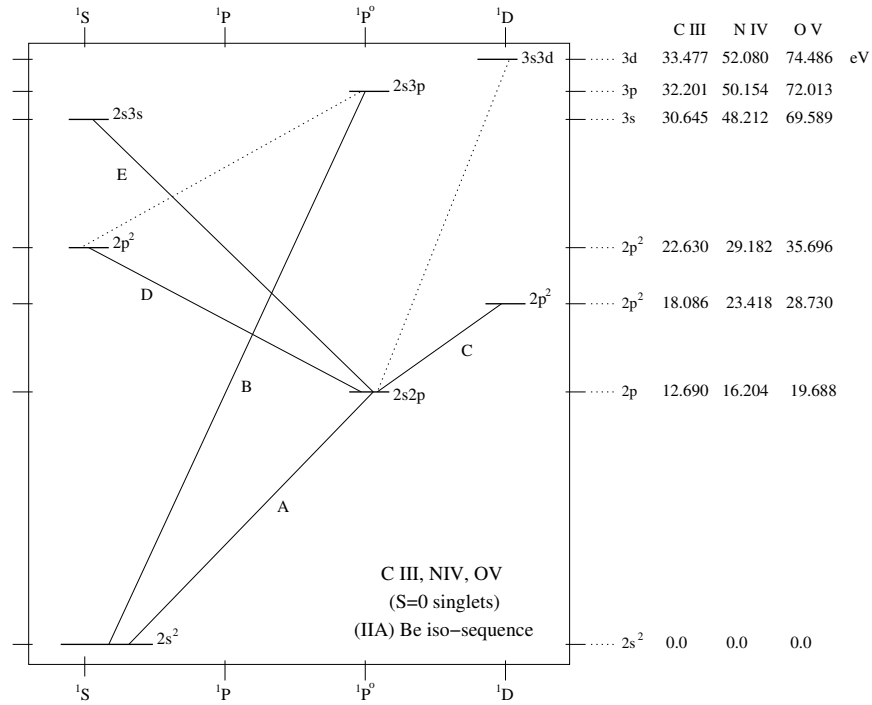


Figure 4.11 Same as for Figure 4.6 but for Group IIA $n_0 = 2$ (beryllium-like) ions with iso-electronic sequence $1s^2 2s^2$. Only the singlet states are shown. All data are obtained from Kramida et al. (2014). All wavelengths are vacuum.

Key	Valence Config.	Singlet	C III λ , Å	N IV λ , Å	O V λ , Å
A	$2s^2-2s2p$	$^1S_0-^1P_1^o$	977.020	765.147	629.732
B	$2s^2-2s3p$	$^1S_0-^1P_1^o$	386.203	247.205	172.169
C	$2s2p-2p^2$	$^1P_1^o-^1D_2$	2297.578	1718.550	1371.296
D	$2s2p-2p^2$	$^1P_1^o-^1S_0$	1247.383	955.334	774.518
E	$2s2p-2s3s$	$^1P_1^o-^1S_0$	690.521	387.356	248.460

$n_0 = 3$ valence shell includes six 2p electrons that can partake in transitions. We do not represent multi-valence electron transitions on the Grotrian diagrams.

Again, we emphasize that transition A, the 0-volt resonance transition between the ground state and the lowest lying excited state is the strongest absorption transition for Alkali-earth metals and Group IIA-like ions. As

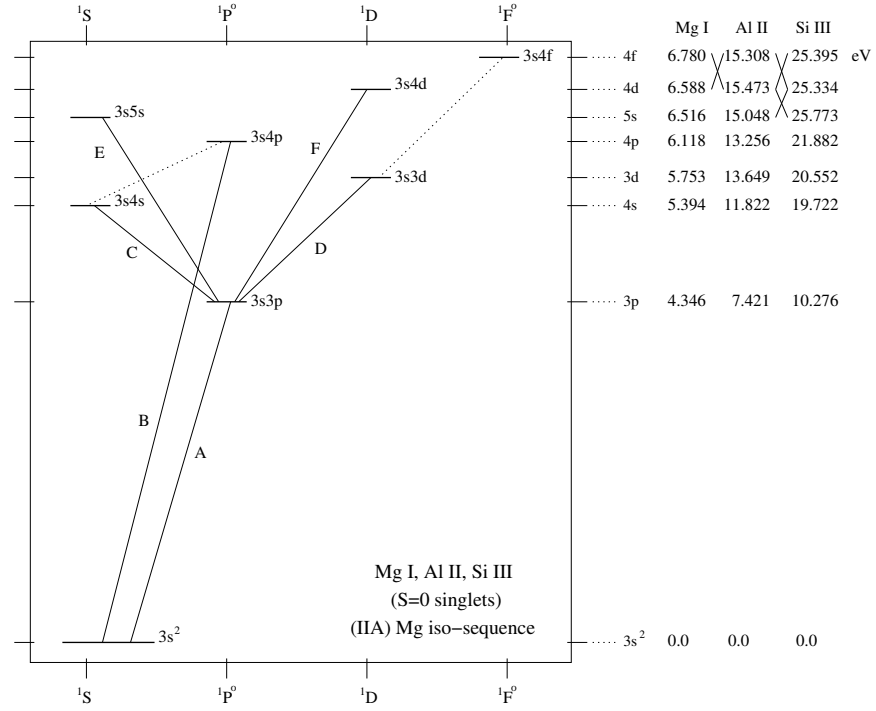


Figure 4.12 Same as for Figure 4.6 but for Group IIA $n_0 = 3$ (magnesium-like) ions with iso-electronic sequence $1s^2 2s^2 2p^6 3s^2$. Only the singlet states are shown. Note that the energies of the 4d and 5s subshells are inverted for SiII and that the 4d and 4f subshells energies are inverted for AlII with respect to the diagram. All data are obtained from Kramida et al. (2014) except the transition A and B wavelengths for MgI, which are taken from Aldenius, Johansson, & Murphy (2006), and for the A transition of AlII, which is taken from Griesmann & Kling (2000). All wavelengths are vacuum.

Key	Valence Config.	Singlet	MgI λ , Å	AlII λ , Å	SiIII λ , Å
A	$3s^2-3s3p$	$^1S_0-^1P_1^o$	2852.9628	1670.7887	1206.500
B	$3s^2-3s4p$	$^1S_0-^1P_1^o$	2026.4750	935.2738	566.613
C	$3s3p-3s4s$	$^1P_1^o-^1S_0$	11831.423	2817.014	1312.591
D	$3s3p-3s3d$	$^1P_1^o-^1D_2$	8809.175	1990.533	1206.555
E	$3s3p-3s5s$	$^1P_1^o-^1S_0$	5712.6725	1625.6276	800.066
F	$3s3p-3s4d$	$^1P_1^o-^1D_2$	5529.9403	1539.8329	823.409

such, the most common transitions from this group observed in astronomical spectra are C III $\lambda 977$, Mg I $\lambda 2852$, Al II $\lambda 1670$, and Si III $\lambda 1206$. Less frequently, N IV $\lambda 765$ and O V $\lambda 629$ are observed; this is in part to a lack of observing facilities that cover these far-ultraviolet wavelengths with the requisite sensitivity, even when the spectral lines are cosmologically redshifted.

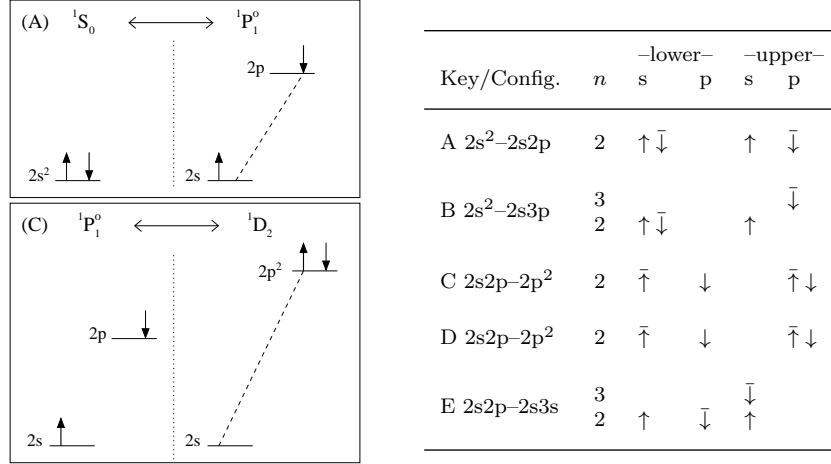


Figure 4.13 Schematics of the electron occupation for transitions A and C presented in Figure 4.11 for C III, N IV, and O V ($n_0 = 2$). Transition A is the transition from the 0-volt ground state, $(2s^2)^1S_0$ to the first excited state $(2s2p)^1P_1$. Transition C is an excited-state transition from $(2s2p)^1P_1$ to $(2p^2)^1D_2$. The transitions are further illustrated in the accompanying table, along with the 0-volt transition B, and the excited transitions D and E. For each transition, the active valence electron is identified with a bar above the arrow symbol.

4.4.3 Group IIIA: Boron Group

Group IIIA metals include boron (B, $Z = 5$), aluminum (Al, $Z = 13$), gallium (Ga, $Z = 31$), indium (In, $Z = 49$), thallium (Tl, $Z = 81$), and the not-fully characterized ununtrium (Uut, $Z = 113$). A commonly observed boron-like ion is C II, and somewhat less commonly N III and O IV. Commonly observed aluminum-like ions include Al I and Si II.

The ground-state structure of these atoms is three valence electrons, two orbiting in the n_0s subshell and one orbiting in the n_0p subshell just outside inner closed shells. The ground-state valence electronic configuration is $n_0s^2 n_0p$, where $n_0 = 2$ for boron and boron-like ions (with closed inner shell $1s^2 2s^2$) and $n_0 = 3$ for aluminum and aluminum-like ions (with closed inner

shell $1s^2 2s^2 2p^6$). Absorption from gallium, indium, and the higher- Z Group IIIA metals and their iso-sequence ions is exceedingly rare in astronomical spectra.

As with the Group IIA Alkali-earth metals and their iso-sequence ions, there is a large number of possible states if all cases in which the three valence electrons can be excited are considered. We will consider transitions in which one valence electron is active and the other two are inactive. For this assumption, the partial Grotrian diagrams for selected boron-like ions and aluminum-like ions are given in Figures 4.15 and 4.16, respectively.

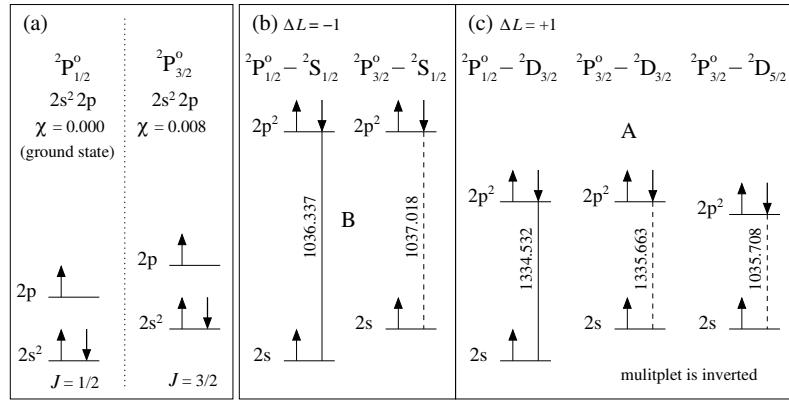


Figure 4.14 Schematics of the fine-structure splitting for C II. (a) The $J = 1/2$ ground-state electron occupation and the $J = 3/2$ fine-structure state, which has excitation energy $\chi = 0.007863$ eV. (b) The doublet B as shown in Figure 4.15. The spin-down $2s$ valence electron transitions to a spin-down state in the $2p$ subshell with $\Delta L = -1$. The transition from the $2s^2 2p \ J = 1/2$ (ground) state is shown as the solid line and the transition from the $2s^2 2p \ J = 3/2$ state is shown as a dashed line (following the convention in Figure 4.15). (c) The triplet A as shown in Figure 4.15. The spin-down $2s$ valence electron transition to a spin-down state in the $2p$ subshell with $\Delta L = +1$. The transition from the ground-state is the solid line. In astronomical spectra, the ground-state transitions are by far the strongest because the vast majority of ions in astrophysical environments are in their ground state.

Because the $n_0 s^2$ subshell for Boron-group atoms and iso-sequence ions is closed, the ground-state term symbol is given by $L = l$ and $S = s$ of the $n_0 p$ electron. With $L = l = 1$ and $S = 1/2$, the ground state Russell-Saunders term symbol is $^2P^o$; thus, the ground-state multiplicity is a doublet with $J = 1/2$ and $3/2$. As schematically illustrated in Figure 4.14(a) for the boron-like ion C II, the doublet is normal, so that the $^2P_{1/2}^o$ state is the 0-volt ground state with the $^2P_{3/2}^o$ state having slightly elevated energy. In combi-

nation with the multiplicity of the excited states, this yields doublet, triplet, and quadruplet multiplets in the spectra of boron-like ions, aluminum, and aluminum-like ions.

In Figure 4.14(b), we show the $\Delta L = -1$ doublet labeled B, the 0-volt transition $(2s^22p)^2P_{1/2}^o - (2s2p^2)^2S_{1/2}$ (solid line) and the corresponding $(2s^22p)^2P_{3/2}^o - (2s2p^2)^2S_{1/2}$ transition (dashed line). For each, the spin-down 2s valence electron transitions spin down to the 2p subshell. Figure 4.14(c), we show the $\Delta L = +1$ triplet labeled A. The 0-volt transition is indicated with the solid line. Note that the 2D multiplet is inverted. Not shown on Figure 4.14 is the $\Delta L = 0$ quadruplet (labeled C), which is a normal multiplet. Note that, unlike the case with Group IIA $2p^2$ transitions, $\Delta L = 0$ is permitted for Group IIIA $2p^2$ transitions because $\Delta l = +1$.

The multiplets of the Group IIIA metals and iso-sequence ions are finely split in energy, especially for CII. However, for absorption lines, the strongest transitions of these multiplets are those from the 0-volt ground state, $^2P_{1/2}^o$. In the tables accompanying Figures 4.15 and 4.16, we have written the transition wavelengths for the 0-volt transitions in bold face font. For CII, the most commonly observed absorption lines in order from strongest to weakest are CII $\lambda 1036$ and CII $\lambda 1334$. For NIII and OIV, the most common lines in absorption are NIII $\lambda 989$ and OIV $\lambda 787$. The most commonly observed SiII absorption lines in order of decreasing strength are SiII $\lambda 1260$, SiII $\lambda 1526$, and the SiII $\lambda \lambda 1190, 1193$ doublet (which is inverted). Finally, AlI has several lines of similar strength, including AlI $\lambda 2367$, AlI $\lambda 3083$, and AlI $\lambda 2367$.

Whereas the lowest lying excited states for the Alkali metals and Alkali-earth metals are quite high in energy, on the order of 2 to 15 eV, the fine-structure $^2P_{1/2}, ^2P_{3/2}$ doublet of the ground $n_0s^2n_0p$ valence configuration for Group IIIA atoms and ions provides a very low lying excited energy state on the order of 0.01 to 0.05 eV. It is for this reason that, under favorable conditions², absorption lines from the $n_0s^2n_0p(^2P_{3/2})$ lower state can sometimes be observed. This is especially the case for CII, and SiII, which have their $n_0s^2n_0p(^2P_{3/2})$ states at $\simeq 0.008$ and 0.036 eV, respectively. Examples of these transitions are CII* $\lambda 1335.6$, $\lambda 1335.7$, and $\lambda 1337.0$, and SiII* $\lambda 1533.4$, $\lambda 1265.0$, and $\lambda 1264.7$, where the * denotes that the initial state was excited.

² Exciting the active ground-state valence electron from $^2P_{1/2}$ to $^2P_{3/2}$ requires a spin flip with no parity change. Thus, the CII* and SiII* transitions are forbidden by the dipole selection rules and do not occur due to radiative absorption; they occur following collisional excitation, which requires relatively high density gas.

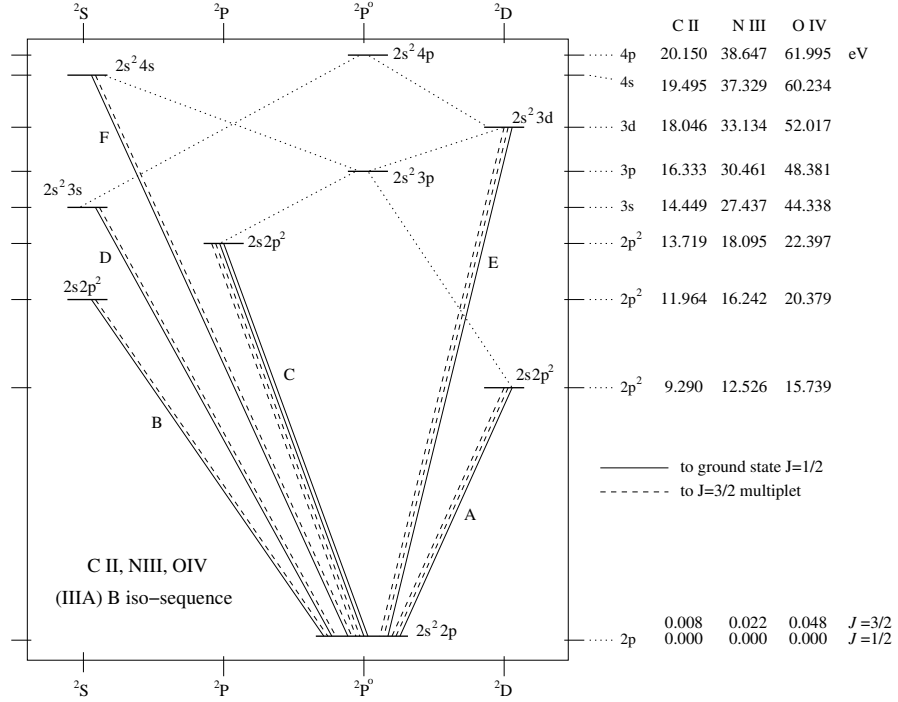
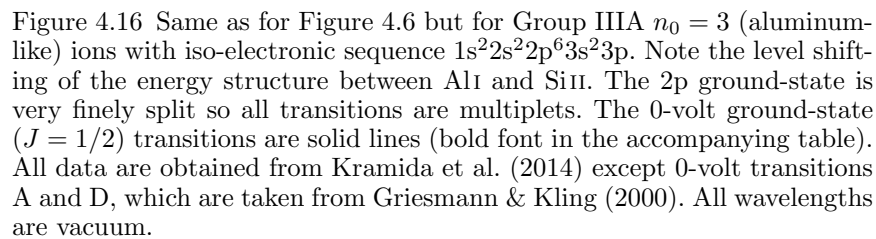


Figure 4.15 Same as for Figure 4.6 but for Group IIIA $n_0 = 2$ (boron-like) ions with iso-electronic sequence $1s^2 2s^2 2p$. The $2p$ ground-state is very finely split by spin-orbit coupling, so all transitions are multiplets. The 0-volt ground-state ($J = 1/2$) transitions are solid lines and are written in bold font in the accompanying table. All data are obtained from Kramida et al. (2014). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	C II λ , Å	N III λ , Å	O IV λ , Å
A	$2s^2 2p - 2s^2 2p^2$	$2P_{3/2}^o - 2D_{3/2}$	1335.663	991.511	790.122
		$2P_{3/2}^o - 2D_{5/2}$	1335.708	991.577	790.119
		$2P_{1/2}^o - 2D_{3/2}$	1334.532	989.799	787.710
B	$2s^2 2p - 2s^2 2p^2$	$2P_{3/2}^o - 2S_{1/2}$	1037.018	764.351	609.829
		$2P_{1/2}^o - 2S_{1/2}$	1036.337	763.334	608.397
C	$2s^2 2p - 2s^2 2p^2$	$2P_{3/2}^o - 2P_{1/2}$	904.480	686.336	555.263
		$2P_{3/2}^o - 2P_{3/2}$	904.142	685.517	555.513
		$2P_{1/2}^o - 2P_{1/2}$	903.962	685.515	554.076
		$2P_{1/2}^o - 2P_{3/2}$	903.623	684.998	553.329
D	$2s^2 2p - 2s^2 3s$	$2P_{3/2}^o - 2S_{1/2}$	858.599	452.227	279.933
		$2P_{1/2}^o - 2S_{1/2}$	858.092	451.871	279.631
E	$2s^2 2p - 2s^2 3d$	$2P_{3/2}^o - 2D_{3/2}$	687.352	374.434	238.579
		$2P_{3/2}^o - 2D_{5/2}$	687.245	374.442	238.570
		$2P_{1/2}^o - 2D_{3/2}$	687.053	374.198	238.360
F	$2s^2 2p - 2s^2 4s$	$2P_{3/2}^o - 2S_{1/2}$	636.251	332.327	...
		$2P_{1/2}^o - 2S_{1/2}$	635.994	332.135	...



Key	Valence Config.	Multiplet	Si II λ , Å	Al I λ Å
A	$3s^23p-3s3p^2$	$2P_{3/2}^o-2D_{3/2}$	1817.451	...
		$2P_{3/2}^o-2D_{5/2}$	1816.928	...
		$2P_{1/2}^o-2D_{3/2}$	1808.013	...
B	$3s^22p-3s3p^2$	$2P_{3/2}^o-2S_{1/2}$	1309.276	1936.459
		$2P_{1/2}^o-2S_{1/2}$	1304.370	1932.266
C	$3s^23p-3s3p^2$	$2P_{3/2}^o-2P_{1/2}$	1197.349	1769.1327
		$2P_{3/2}^o-2P_{3/2}$	1194.500	1766.3813
		$2P_{1/2}^o-2P_{1/2}$	1193.290	1765.6323
		$2P_{1/2}^o-2P_{3/2}$	1190.416	1762.8918
D	$3s^23p-3s^24s$	$2P_{3/2}^o-2S_{1/2}$	1533.431	3962.6410
		$2P_{1/2}^o-2S_{1/2}$	1526.707	3945.1224
E	$3s^23p-3s^24d$	$2P_{3/2}^o-2D_{3/2}$	992.696	2374.0743
		$2P_{3/2}^o-2D_{5/2}$	992.683	2373.8466
		$2P_{1/2}^o-2D_{3/2}$	989.873	2367.7750
F	$3s^23p-3s^23d$	$2P_{3/2}^o-2D_{3/2}$	1265.002	3093.7347
		$2P_{3/2}^o-2D_{5/2}$	1264.738	3039.6062
		$2P_{1/2}^o-2D_{3/2}$	1260.422	3038.0462

4.4.4 Iron Family Ions

In addition to the commonly observed transitions from the ground-states of Group IA Alkali metals, Group IIA Alkali-earth metals, and Group IIIA metals, transitions from the “iron-family” are common. The iron-family comprises the ten elements ranging from scandium (Sc, $Z = 21$) through zinc (Zn, $Z = 30$). These elements are characterized by the electronic sequences from $[\text{Ar}]3d4s^2$ to $[\text{Ar}]3d^{10}4s^2$.

Rather than there being similarities between the spectra of these elements and their ions, each has its own unique characteristics. The added variety arises because the valence 3d subshell has binding energies that bracket the closed outer 4s subshell. For example, in scandium, the single 3d electron is less bound than the 4s electrons, but once the 3d shell is about half occupied, i.e., $3d^4$, as it is for chromium and manganese, the 4s electrons are less bound than the 3d electrons. The reasons for this behavior were discussed in § 4.3. As such, there is a wide range of transition channels, some being dominated by an active 3d valence electron and some being dominated by a 4s electron. Furthermore, during a transition, the term (the L and/or S states) of the atom or ion can change. We discuss this further below.

The most commonly observed ultraviolet absorption lines are from the ions TiII , CrII , MnII , FeII , NiII , and ZnII . We limit our discussion to these ions, beginning with iron so that we can develop additional notation to describe the energy levels of atoms and ions with partially occupied 3d subshells.

Fe II

Neutral iron has the electron configuration $[\text{Ar}]3d^64s^2$ and ground state Russell-Saunders symbol 5D_4 . Note that the fully occupied 4s subshell makes no contribution to the quantized vector addition. When ionized, one of the 4s electrons is removed, so that the electron configuration of FeII is $[\text{Ar}]3d^64s$. The number of occupation permutations of the six equivalent 3d electrons is $210!$ Working through these occupations, the terms for the 3d subshell are 5D , 3P , 3D , 3F , 3G , 3H , 1S , 1D , 1F , 1G , and 1F . According to Hund’s rule, the lowest energy term has the highest multiplicity, and this term is 5D ($L = 2$ and $S = 2$). Since the subshell is more than half occupied, the multiplets are inverted, and the lowest energy state has $J = L + S = 4$, giving 5D_4 . Adding the 4s electron ($l = 0$, $m_s = +1/2$) to this lowest energy term, we obtain $L = 2$, $S = 5/2$, and $J = 9/2$, which yields the FeII ground state Russell-Saunders symbol $^6D_{9/2}$.

Because the 3d subshell can have multiple terms, it is convention to in-

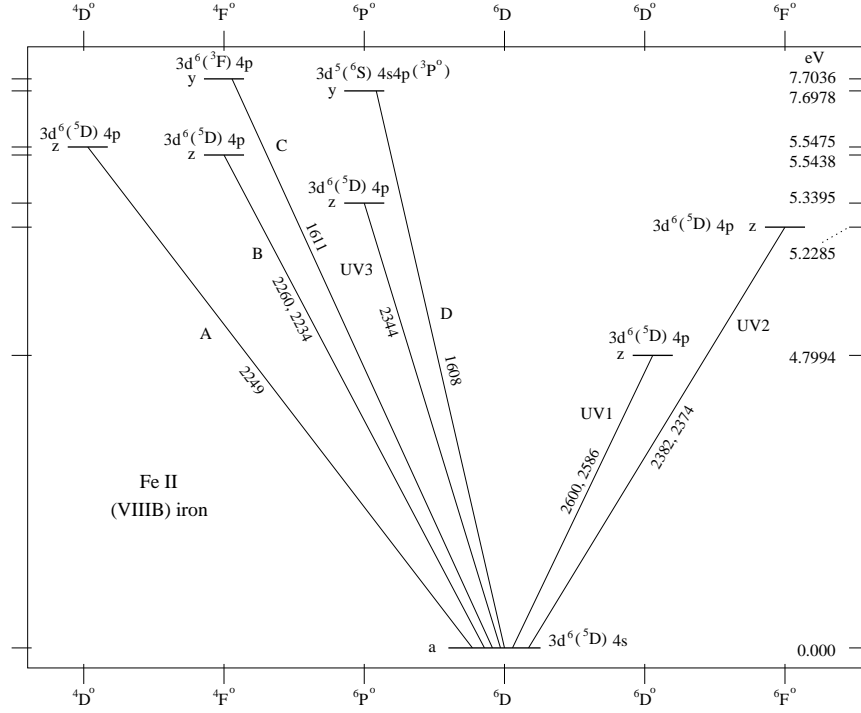


Figure 4.17 Same as for Figure 4.6 but for the FeII ion with ground-state electron sequence $1s^2 2s^2 2p^6 3s^2 3p^6 3d^6 4s$. Each of the transitions shown are fine-structure multiplets, which are unresolved on the diagram. The ground-state is ${}^6D_{9/2}$. The wavelengths of the 0-volt transitions are given for their respective terms. The J states of the UV1, UV2, and UV3 multiplets are illustrated in Figure 4.18. Wavelength data are obtained from Aldenius, Johansson, & Murphy (2006); all other data are from Kramida et al. (2014). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	FeII λ , Å
UV1	$3d^6({}^5D)4s-3d^6({}^5D)4p$	$a^6D_{9/2}-z^6D_{9/2}^o$	2600.1722
		$a^6D_{9/2}-z^6D_{7/2}^o$	2586.6494
UV2	$3d^6({}^5D)4s-3d^6({}^5D)4p$	$a^6D_{9/2}-z^6F_{11/2}^o$	2382.7641
		$a^6D_{9/2}-z^6F_{9/2}^o$	2374.4601
UV3	$3d^6({}^5D)4s-3d^6({}^5D)4p$	$a^6D_{9/2}-z^6P_{7/2}^o$	2344.2128
A	$3d^6({}^5D)4s-3d^6({}^5D)4p$	$a^6D_{9/2}-z^4D_{7/2}^o$	2249.8754
		$a^6D_{9/2}-z^4F_{9/2}^o$	2260.7793
B	$3d^6({}^5D)4s-3d^6({}^5D)4p$	$a^6D_{9/2}-z^4F_{7/2}^o$	2234.4462
		$a^6D_{9/2}-y^4F_{7/2}^o$	1611.2003
C	$3d^6({}^5D)4s-3d^6({}^3F)4p$	$a^6D_{9/2}-y^4F_{7/2}^o$	1611.2003
D	$3d^6({}^5D)4s-3d^5({}^6S)4s4p({}^3P^o)$	$a^6D_{9/2}-y^6P_{7/2}^o$	1608.4508

clude the term symbol in the electron configuration, such that the ground state configuration of FeII is written $[\text{Ar}]3d^6(^5D)4s$. Most often, the active valence electron during a transition is the 4s electron, however, one of the 3d electrons can also be the active electron, in which case the 4s electron is not active. Furthermore, during one of these transitions, the vector addition of the 3d electrons can change, say from 5D to 3F . Due to the number of transition channels available between the 3d or 4s active electron, multiple transitions between identical terms can occur.

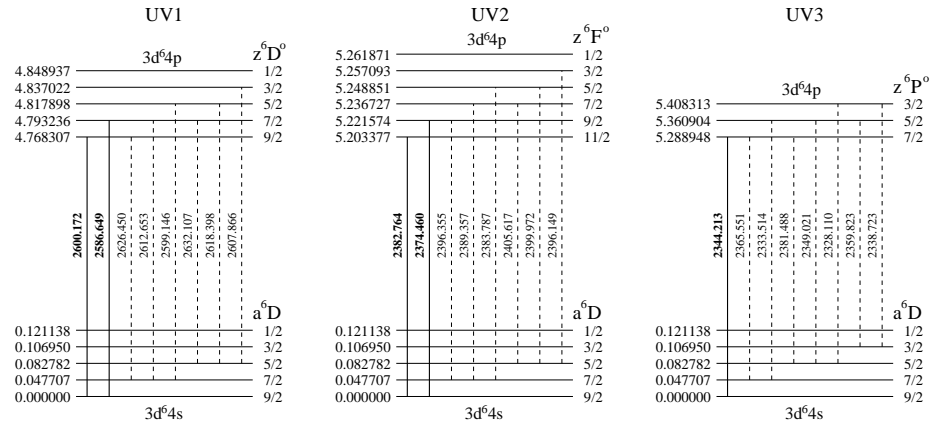


Figure 4.18 The lowest energy term for FeII is 6D . The lowest energy state (the 0 -volt ground state) has the maximum J , which is $9/2$. For the UV1 transition, two of the multiplets transition between the 0 -volt state and the $^6D^o$ term. For the UV2 transition, two of the multiplets transition between the 0 -volt state and the $^6F^o$ term. For the UV3 transition, one of the multiplets transition between the 0 -volt state and the $^6P^o$ term. The 0 -volt transitions are indicated as thick solid lines and the other members of the multiplet are shown as dashed lines. Note, not all members of the multiplets are shown. All data are from Kramida et al. (2014). All wavelengths are vacuum.

For example, from the ground state, the 4s electron can transition to the the 4p suborbital. One of the resulting terms for this excited state is $^4F^o$. This term can result from either the valence configuration $3d^6(^5D)4p$ or $3d^6(^3F)4p$, the latter excited state having also involved a change in L and S within the 3d subshell during the transition. These two $^4F^o$ terms have different energies. To distinguish them with a short-hand notation, the convention is to label the lowest energy term for the upper states with “z” and the next lowest energy term with “y”. Thus, for the $3d^6(^5D)4p$ upper state the term is z^4F^o and for the $3d^6(^3F)4p$ upper state the term is y^4F^o . Successively higher energy states for the given term are denoted “x”, “w”,

“v”, etc. A similar scheme is applied for lower states of transitions having the same term. In this case, the lowest energy state is “a”, then next lowest “b”, followed by “c”, “d”, etc. The ground state is thus always labeled with “a”.

In Figure 4.17, we illustrate a partial Grotrian diagram of selected transitions for FeII. On the diagram, each of the shown energy states is a multiplet, so the transitions are multiplets. Those that are 0-volt transitions are the most commonly observed in absorption in astronomical spectra; the wavelength of these transitions are indicated on the diagram. The accompanying table provides a key to the 0-volt transitions. In Figure 4.18, we show the multiplet structure for the transitions labeled UV1, UV2, and UV3 in Figure 4.17. This figure highlights the suite of five lines, FeII λ 2344, λ 2374, λ 2382, λ 2586, and λ 2600 that are very commonly seen in cosmological spectra.

Ni II

Neutral nickel has the ground state electron configuration [Ar]3d⁸4s² in the state ³F₄. When ionized, the configuration is [Ar]3d⁹ and the ground state is ²D_{5/2}. Thus, 0-volt transitions involve a 3d electron. There are literally hundreds of NiII transitions with spectral features in the ultraviolet and optical (see Shenstone, 1970, for a thorough tabulation).

Most 0-volt transitions involve a single active 3d valence electron transitioning to the 4p state. In Figure 4.19, we present a partial Grotrian diagram, showing only the most commonly observed NiII absorption line transitions. The accompanying table provides the key to the transitions and the transition wavelengths. As with FeII, all states are multiplets. For NiII, however, we partially resolve the multiplets for the UV4 and UV5 transitions. For both these multiplets, two members of the multiplets transition between the ground state. The first, UV4, gives rise to the pair NiII λ 1709 and λ 1751. The second, UV5 gives rise to the commonly observed absorption line NiII λ 1741. The NiII λ 1703 0-volt member of the latter multiplet has an oscillator strength that is about 10% of the λ 1741 line and is not as readily observed.

Ti II

Neutral titanium has the electron configuration [Ar]3d²4s². For TiII, the ionized electron is removed from the 4s state. Having only two 3d electrons, the energy structure of TiII is simpler than that of FeII. The ground-state occupation of the 3d² subshell is the state ³F₂ (which is the ground-state

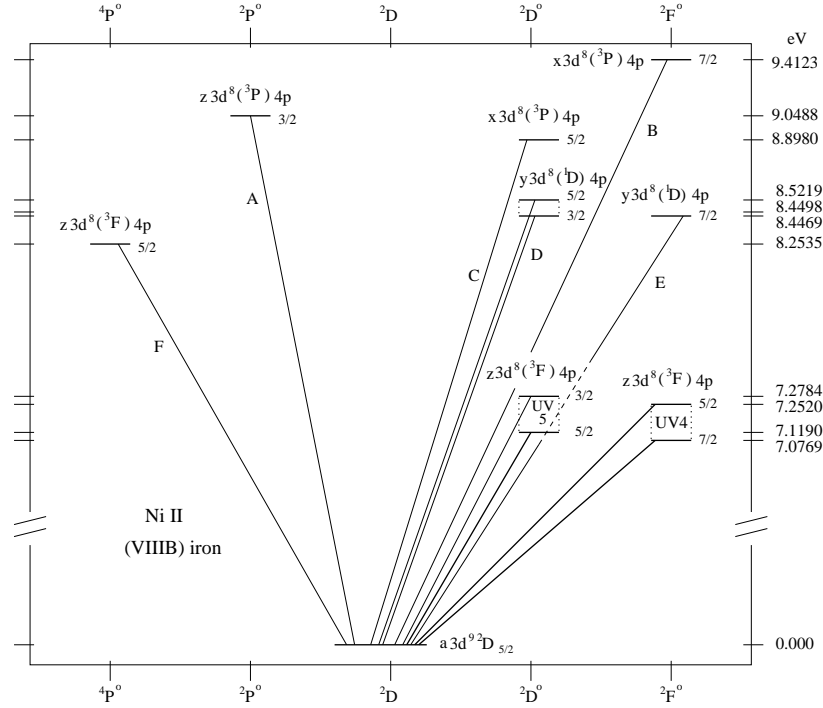


Figure 4.19 Same as for Figure 4.6 but for the NiII ion with ground-state electron sequence $1s^22s^22p^63s^23p^63d^9$. Each of the transitions shown are fine-structure multiplets, most of which are unresolved on the diagram. The ground-state is $^2D_{5/2}$. The most commonly observed transitions are shown as bold line and are indicated in the below table in bold face font. Energy, term, and state data are obtained from Shenstone (1970). Wavelength data are obtained by Shenstone (1970) except for UV4 and UV5 transitions, which are obtained from Pickering et al. (2000). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	NiII λ , Å
UV4	$3d^9(^2D)-3d^8(^3F)4p$	$a^2D_{5/2}-z^3F_{5/2}^\circ$	1709.6042
		$a^2D_{5/2}-z^3F_{7/2}^\circ$	1751.9157
UV5	$3d^9(^2D)-3d^8(^3F)4p$	$a^2D_{5/2}-z^2D_{5/2}^\circ$	1741.5531
		$a^2D_{5/2}-z^2D_{3/2}^\circ$	1703.4119
A	$3d^9(^2D)-3d^8(^3P)4p$	$a^2D_{5/2}-z^2P_{3/2}^\circ$	1370.136
B	$3d^9(^2D)-3d^8(^1G)4p$	$a^2D_{5/2}-x^2F_{7/2}^\circ$	1317.220
C	$3d^9(^2D)-3d^8(^3P)4p$	$a^2D_{5/2}-x^2D_{5/2}^\circ$	1393.330
D	$3d^9(^2D)-3d^8(^1D)4p$	$a^2D_{5/2}-y^2D_{5/2}^\circ$	1454.852
		$a^2D_{5/2}-y^2D_{3/2}^\circ$	1467.256
E	$3d^9(^2D)-3d^8(^1D)4p$	$a^2D_{5/2}-y^2F_{7/2}^\circ$	1467.762
F	$3d^9(^2D)-3d^8(^3P)4p$	$a^2D_{5/2}-z^4P_{5/2}^\circ$	1502.150

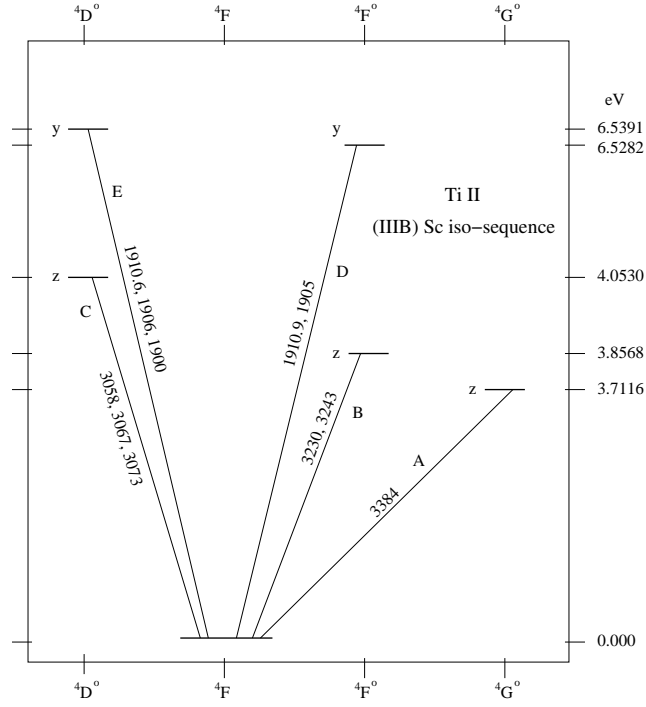


Figure 4.20 Same as for Figure 4.6 but for the TiII ion with ground-state electron sequence $1s^2 2s^2 2p^6 3s^2 3p^6 3d^2 4s$. Each of the transitions shown are fine-structure multiplets, which are unresolved on the diagram. The ground-state is $4F_{3/2}$. The wavelengths of the 0-volt transitions are given for their respective terms. All data are obtained from Kramida et al. (2014) except wavelengths for transitions A, B, and C, which are obtained from Aldenius, Johansson, & Murphy (2006). All wavelengths are vacuum.

Key	Valence Config.	Multiplet	TiII λ , Å
A	$3d^2(^3F)4s-3d^2(^3F)4p$	$a^4F_{3/2}-z^4G_{5/2}^o$	3384.7300
B	$3d^2(^3F)4s-3d^2(^3F)4p$	$a^4F_{3/2}-z^4F_{3/2}^o$	3242.9180
		$a^4F_{3/2}-z^4F_{5/2}^o$	3230.1217
C	$3d^2(^3F)4s-3d^2(^3F)4p$	$a^4F_{3/2}-z^4D_{5/2}^o$	3058.2819
		$a^4F_{3/2}-z^4D_{3/2}^o$	3067.2375
		$a^4F_{3/2}-z^4D_{1/2}^o$	3073.8629
D	$3d^2(^3F)4s-3d(^2D)4s4p(^3P^o)$	$a^4F_{3/2}-y^4F_{5/2}^o$	1905.7819
		$a^4F_{3/2}-y^4F_{3/2}^o$	1910.9538
E	$3d^2(^3F)4s-3d(^2D)4s4p(^3P^o)$	$a^4F_{3/2}-y^4D_{1/2}^o$	1910.6123
		$a^4F_{3/2}-y^4D_{3/2}^o$	1906.2362
		$a^4F_{3/2}-y^4D_{5/2}^o$	1900.0252

term for TiI because the $4s^2$ electrons makes no contribution to the quantized vector addition). Accounting for the single $4s$ electron in TiII, the ion's ground state is $^4F_{3/2}$.

In Figure 4.20, we present the partial Grotrian diagram for the multiplets transitioning from the 4F term, which has four-fold multiplicity with $J = 3/2, 5/2, 7/2$, and $9/2$. The multiplet obeys the Landé interval and is normal, i.e., $J = 3/2$ is the lowest energy state. The multiplets are not resolved on the diagram. Furthermore, only the commonly observed 0-volt transitions are labeled. The accompanying table provides the key to the transitions. Note that the upper excited states are also multiplets, so there are many non-0-volt transitions between the shown states.

CrII, MnII, and ZnII

In Figure 4.21 we present partial Grotrian diagrams for the strongest 0-volt transitions arising from CrII and MnII. The transitions are selected based upon their oscillator (or line) strengths. The CrII transitions all have oscillator strengths greater than 0.01 and those for MnII have values greater than 0.1. Both ions exhibit triplets in their spectra.

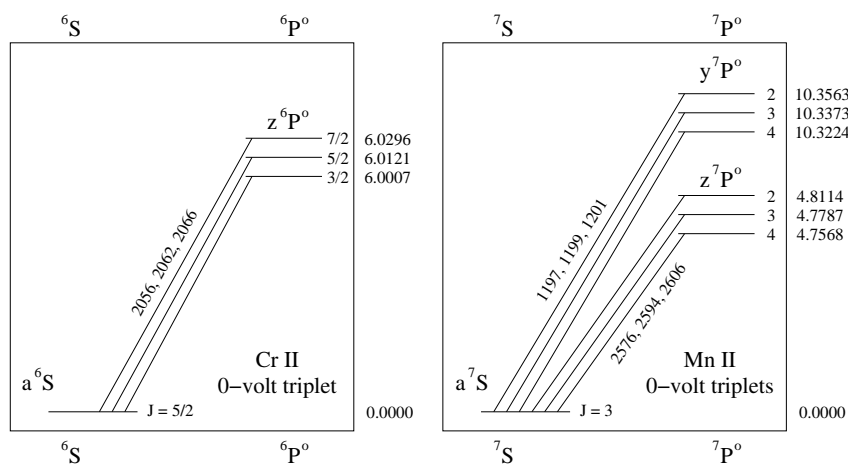


Figure 4.21 Partial Grotrian diagram for the strongest 0-volt CrII and MnII triplets. The shown CrII transitions are selected to have oscillator strengths greater than 0.01. The MnII transitions are selected to have oscillator strengths greater than 0.1. All data are obtained from Kramida et al. (2014).

For CrII, note that the ground state, $[Ar]3d^5$ with term 6S , has multiplicity $2S + 1 = 6$. However, for this term, $L < S$, so the actual multiplicity is

$2L + 1 = 1$. Thus, the ground state is a singlet state, ${}^6S_{5/2}$. The lowest lying excited state, with valence configuration $3d^4({}^5D)4p$ and term ${}^6P^\circ$ also has $L < S$; the multiplicity is therefore $2L + 1 = 3$ and the state is a triplet. For this triplet, the active valence electron is a 3d electron that transitions to a 4p electron. The transition multiplets are $a^6S_{5/2}-z^6P_{3/2}^\circ$, $a^6S_{5/2}-z^6P_{5/2}^\circ$, and $a^6S_{5/2}-z^6P_{7/2}^\circ$, yielding the spectral features 2066.1639, 2062.2359, and 2056.2568 Å, respectively.

The ground-state electron configuration of Mn II is that of neutral chromium, $[Ar]3d^54s$ with term 7S . Though the multiplicity is $2S + 1 = 7$, $L < S$, so the multiplicity is $2L + 1 = 1$. The ground state is 7S_3 . For Mn II, there are two strong triples. The first is the triplet $3d^5({}^6S)4s-3d^4({}^6S)4p$, for which the 4s active valence electron transitions to the 4p state while the 3d electrons are inactive. The transition multiplets are $a^7S_3-z^7P_4^\circ$, $a^7S_3-z^7P_3^\circ$, and $a^7S_3-z^7P_2^\circ$, yielding the spectral features 2606.4588, 2594.4967, and 2576.8753 Å, respectively. The second triplet is $3d^5({}^6S)4s-3d^4({}^6S)4s4p({}^3P^\circ)$, for which a 3d active valence electron transitions to the 4p state while the 4s electron is inactive. The transition multiplets are $a^7S_3-y^7P_4^\circ$, $a^7S_3-y^7P_3^\circ$, and $a^7S_3-y^7P_2^\circ$, yielding the spectral features 1201.1148, 1199.3888, 1197.1815 Å, respectively.

The ground-state electron configuration of Zn II is that of neutral copper, $[Ar]3d^{10}4s$ with atomic state ${}^2S_{1/2}$. Since all subshells are fully occupied except for the 4s valence subshell, Zn II has an energy structure mirroring Group IA Alkali ions. Thus, the 0-volt resonant doublet is $4s({}^2S_{1/2})-4p({}^2P_{3/2})$ and $4s({}^2S_{1/2})-4p({}^2P_{1/2})$. The vacuum wavelengths are 2026.1369 and 2062.6603 Å, respectively.

4.5 Common Absorption Lines

In Table 4.5, we present a list of the absorption lines typically observed in cosmological spectra that probe the interstellar, circumgalactic, and intergalactic media. The lines are grouped first by ion in atomic number order and then by ionization stage from lowest to highest. For a given ion, the lines are ordered from lowest to highest transition energy. The upper state is denoted k and the lower state is denoted k' . Only E1 (“allowed” electric dipole) transitions are listed.

Also tabulated are the valence electron configurations, the Russell-Saunders state symbols, the transition vacuum wavelengths ($\lambda_k^{k'}$, in angströms), the absorption oscillator strengths, $f_k^{k'}$, and the damping constants, $\Gamma_k^{k'}$. These latter three quantities constitute the three atomic constants required to char-

acterize an absorption line for purposes of spectral analysis. Also listed are the statistical weights of the lower and upper states. In the final column, the figure number for the Grotrian diagram illustrating the transition is provided.

For H I, the transitions are the term averages (see § 3.3.1), so the Russell-Saunders term symbol is tabulated. In this case, the statistical weight is $g_k = (2S + 1)(2L + 1)$. For the metal lines, the transitions are averaged over all M_J states, which are energy degenerate in the absence of a magnetic field. In this case, the statistical weights are $g_k = 2J + 1$.

The data for the vacuum wavelengths are taken from several sources, primarily from the NIST database (Kramida et al., 2014). Those of Mg I, Mg II, Ti II, Cr II, Mn II, Fe II, and Zn II are taken from Aldenius, Johansson, & Murphy (2006), and those of Ni II are taken from Pickering et al. (2000). Those of Al I and Al II are obtained from Griesmann & Kling (2000). Oscillator strengths and damping constants are taken from Moore (1970) except the H I, which are taken from (Kramida et al., 2014). We also refer the reader to the publication “Atomic Data Revisions for Transitions Relevant to Observations of Interstellar, Circumgalactic, and Intergalactic Matter” by Cashman et al. (2017).

Table 4.3 *Common absorption lines*

Ion/Tran	Transition		State		$\lambda_k^{k'}$	$f_k^{k'}$	$\Gamma_k^{k'}$	g_k	$g_{k'}$	Diagram
	k	k'	k	k'	Å		10^8 s^{-1}			
H I $\lambda 1215$ (Ly α)	1s	2p	$2S^o$	$2P^o$	1215.6701	0.4164	6.265	2	6	2.2
D I $\lambda 1215$ (D α)	1s	2p	$2S^o$	$2P^o$	1215.3394	0.4165	6.270	2	6	
H I $\lambda 1025$ (Ly β)	1s	3p	$2S^o$	$2P^o$	1025.7223	0.07912	1.897	2	6	
H I $\lambda 972$ (Ly γ)	1s	4p	$2S^o$	$2P^o$	972.5368	0.02901	0.8126	2	6	
H I $\lambda 949$ (Ly δ)	1s	5p	$2S^o$	$2P^o$	949.7431	0.01395	0.7640	2	6	
H I $\lambda 937$ (Ly ϵ)	1s	6p	$2S^o$	$2P^o$	937.8035	0.007804	0.4423	2	6	
C II $\lambda 1334$	$2s^2 2p$	$2s 2p^2$	$2P_{1/2}^o$	$2D_{3/2}$	1334.5323	0.1278	2.870	2	4	4.15
C II $\lambda 1036$	$2s^2 2p$	$2s 2p^2$	$2P_{1/2}^o$	$2S_{1/2}$	1036.3367	0.1231	2.290	2	2	4.15
C III $\lambda 977$	$2s^2$	$2s 2p$	$1S_0$	$1P_1^o$	977.020	0.7620	17.75	1	3	4.11
C IV $\lambda 1550$	2s	2p	$2S_{1/2}$	$2P_{1/2}^o$	1550.772	0.09522	2.640	1	2	4.6
C IV $\lambda 1548$	2s	2p	$2S_{1/2}$	$2P_{3/2}^o$	1548.187	0.1908	2.654	1	4	4.6
N II $\lambda 1083$	$2s^2 2p^2$	$2s 2p^3$	$3P_0$	$3D_1^o$	1083.990	0.1031	3.508	1	3	
N II $\lambda 916$	$2s^2 2p^2$	$2s 2p^3$	$3P_0$	$3P_1^o$	916.612	0.1449	11.51	1	3	
N III $\lambda 989$	$2s^2 2p$	$2s 2p^2$	$2P_{1/2}^o$	$2D_{3/2}$	989.799	0.1066	4.352	2	4	4.15
N V $\lambda 1242$	2s	2p	$2S_{1/2}$	$2P_{1/2}^o$	1242.804	0.07823	3.411	1	2	4.6
N V $\lambda 1238$	2s	2p	$2S_{1/2}$	$2P_{3/2}^o$	1238.821	0.1570	3.378	1	4	4.6
O IV $\lambda 787$	$2s^2 2p$	$2s 2p^2$	$2P_{1/2}^o$	$2D_{3/2}$	787.711	0.111	5.95	2	4	4.15
O VI $\lambda 1037$	2s	2p	$2S_{1/2}$	$2P_{1/2}^o$	1037.613	0.06609	4.095	2	2	4.6
O VI $\lambda 1031$	2s	2p	$2S_{1/2}$	$2P_{3/2}^o$	1031.912	0.1329	4.163	2	4	4.6
Ne VIII $\lambda 780$	2s	2p	$2S_{1/2}$	$2P_{1/2}^o$	780.324	0.0502	5.500	1	2	4.6
Ne VIII $\lambda 770$	2s	2p	$2S_{1/2}$	$2P_{3/2}^o$	770.409	0.1020	5.720	1	4	4.6

Common absorption lines, cont.

Ion/Tran	Transition		State		$\lambda_k^{k'}$	$f_k^{k'}$	$\Gamma_k^{k'}$	g_k	$g_{k'}$	Diagram
	k	k'	k	k'	\AA		10^8 s^{-1}			
NaI $\lambda 5897$	3s	3p	$2S_{1/2}$	$2P_{1/2}^o$	5897.558	0.3270	6.280	1	2	4.7
NaI $\lambda 5891$	3s	3p	$2S_{1/2}$	$2P_{3/2}^o$	5891.583	0.6550	6.300	1	4	4.7
NaI $\lambda 3303.9$	3s	4p	$2S_{1/2}$	$2P_{1/2}^o$	3303.929	0.0067	2.000	1	2	4.7
NaI $\lambda 3303.3$	3s	4p	$2S_{1/2}$	$2P_{3/2}^o$	3303.319	0.0134	2.000	1	4	4.7
MgI $\lambda 2852$	$3s^2$	3s3p	$1S_0$	$1P_1^o$	2852.9628	1.810	4.950	1	3	4.12
MgI $\lambda 2026$	$3s^2$	3s4p	$1S_0$	$1P_1^o$	2026.4750	0.112	1.200	1	3	4.12
MgII $\lambda 2803$	3s	3p	$2S_{1/2}$	$2P_{1/2}^o$	2803.5311	0.3054	2.592	1	2	4.7
MgII $\lambda 2796$	3s	3p	$2S_{1/2}$	$2P_{3/2}^o$	2796.5340	0.6123	2.612	1	4	4.7
MgII $\lambda 1240$	3s	4p	$2S_{1/2}$	$2P_{1/2}^o$	1240.3947	0.0001337	0.005796	1	2	4.7
MgII $\lambda 1239$	3s	4p	$2S_{1/2}$	$2P_{3/2}^o$	1239.9253	0.0002675	0.005802	1	4	4.7
AlI $\lambda 3945$	$3s^2 3p$	$3s^2 4s$	$2P_{1/2}^o$	$2S_{1/2}$	3945.1224	0.116934	2.000	2	2	4.16
AlI $\lambda 3038$	$3s^2 3p$	$3s^2 3d$	$2P_{1/2}^o$	$2D_{3/2}$	3038.0462	0.177427	2.000	2	4	4.16
AlI $\lambda 2367$	$3s^2 3p$	$3s^2 4d$	$2P_{1/2}^o$	$2D_{3/2}$	2367.7750	0.1211	0.7204	2	4	4.16
AlII $\lambda 1670$	$3s^2$	3s3p	$1S_0$	$1P_1^o$	1670.7887	1.880	14.60	1	4	4.12
AlIII $\lambda 1862$	3s	3p	$2S_{1/2}$	$2P_{1/2}^o$	1862.7912	0.2680	2.000	1	2	4.7
AlIII $\lambda 1854$	3s	3p	$2S_{1/2}$	$2P_{3/2}^o$	1854.7184	0.5390	2.000	1	4	4.7
SiII $\lambda 1808$	$3s^2 3p$	$3s2p^2$	$2P_{1/2}^o$	$2D_{3/2}$	1808.0130	0.00208	0.06749	2	4	4.16
SiII $\lambda 1526$	$3s^2 3p$	$3s^2 4s$	$2P_{1/2}^o$	$2S_{1/2}$	1526.7071	0.2303	19.60	2	2	4.16
SiII $\lambda 1304$	$3s^2 3p$	$3s2p^2$	$2P_{1/2}^o$	$2S_{1/2}$	1304.370	0.1437	17.20	2	2	4.16
SiII $\lambda 1260$	$3s^2 3p$	$3s^2 3d$	$2P_{1/2}^o$	$2D_{3/2}$	1260.422	1.0070	25.33	2	4	4.16
SiII $\lambda 989$	$3s^2 3p$	$3s^2 4d$	$2P_{1/2}^o$	$2D_{3/2}$	989.873	0.1330	6.673	2	4	4.16
SiII $\lambda 1193$	$3s^2 3p$	$3s2p^2$	$2P_{1/2}^o$	$2P_{1/2}$	1193.290	0.4991	34.95	2	2	4.16
SiII $\lambda 1190$	$3s^2 3p$	$3s2p^2$	$2P_{1/2}^o$	$2P_{3/2}$	1190.416	0.2502	35.03	2	4	4.16
SiIII $\lambda 1206$	$3s^2$	3s3p	$1S_0$	$1P_1^o$	1206.500	1.6600	25.90	2	4	4.12
SiIV $\lambda 1393$	3s	3p	$2S_{1/2}$	$2P_{3/2}^o$	1393.755	0.5280	9.200	1	4	4.7
SiIV $\lambda 1402$	3s	3p	$2S_{1/2}$	$2P_{1/2}^o$	1402.770	0.2620	9.030	1	2	4.7

Common absorption lines, cont.

Ion/Tran	Transition		State		$\lambda_k^{k'}$	$f_k^{k'}$	$\Gamma_k^{k'}$	g_k	$g_{k'}$	Diagram
	k	k'	k	k'	Å		10^8 s^{-1}			
Ca II $\lambda 3969$	3s	3p	$2S_{1/2}$	$2P_{1/2}^o$	3969.592	0.3145	1.414	2	2	4.8
Ca II $\lambda 3934$	3s	3p	$2S_{1/2}$	$2P_{3/2}^o$	3934.777	0.6346	1.456	2	4	4.8
Ti II $\lambda 3384$	$3d^2(^3F)4s$	$3d^2(^3F)4p$	$a^4F_{3/2}$	$z^6G_{5/2}^o$	3384.7300	0.3401	1.730	4	6	4.20
Ti II $\lambda 3073$	$3d^2(^3F)4s$	$3d^2(^3F)4p$	$a^4F_{3/2}$	$z^6D_{1/2}^o$	3073.8629	0.1091	2.500	4	2	4.20
Ti II $\lambda 3067$	$3d^2(^3F)4s$	$3d^2(^3F)4p$	$a^4F_{3/2}$	$z^6D_{3/2}^o$	3067.2375	0.09057	2.500	4	4	4.20
Ti II $\lambda 3242$	$3d^2(^3F)4s$	$3d^2(^3F)4p$	$a^4F_{3/2}$	$z^6F_{3/2}^o$	3242.9180	0.1832	2.350	4	4	4.20
Ti II $\lambda 3230$	$3d^2(^3F)4s$	$3d^2(^3F)4p$	$a^4F_{3/2}$	$z^6F_{5/2}^o$	3230.1217	0.05861	2.350	4	6	4.20
Cr II $\lambda 2066$	$3d^5$	$3d^4(^5D)4p$	$a^6S_{5/2}$	$z^6P_{3/2}^o$	2066.1639	0.06982	1.636	6	4	4.21
Cr II $\lambda 2062$	$3d^5$	$3d^4(^5D)4p$	$a^6S_{5/2}$	$z^6P_{5/2}^o$	2062.2375	0.1049	1.646	6	6	4.21
Cr II $\lambda 2056$	$3d^5$	$3d^4(^5D)4p$	$a^6S_{5/2}$	$z^6P_{7/2}^o$	2056.2568	0.1403	1.660	6	8	4.21
Mn II $\lambda 2606$	$3d^5(^6S)4s$	$3d^5(^6S)4p$	a^7S_3	$z^7P_2^o$	2606.4588	0.1927	2.648	7	5	4.21
Mn II $\lambda 2594$	$3d^5(^6S)4s$	$3d^5(^6S)4p$	a^7S_3	$z^7P_3^o$	2594.4967	0.2710	2.685	7	7	4.21
Mn II $\lambda 2576$	$3d^5(^6S)4s$	$3d^5(^6S)4p$	a^7S_3	$z^7P_5^o$	2576.8753	0.3508	2.741	7	9	4.21
Mn II $\lambda 1201$	$3d^5(^6S)4s$	$3d^4(^5D)4s4p(^3P^o)$	a^7S_3	$y^7P_2^o$	1201.118	0.08809	5.702	7	5	4.21
Mn II $\lambda 1199$	$3d^5(^6S)4s$	$3d^4(^5D)4s4p(^3P^o)$	a^7S_3	$y^7P_3^o$	1199.391	0.1059	4.911	7	7	4.21
Mn II $\lambda 1197$	$3d^5(^6S)4s$	$3d^4(^5D)4s4p(^3P^o)$	a^7S_3	$y^7P_4^o$	1197.184	0.1566	5.670	7	9	4.21
Fe II $\lambda 2600$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6D_{9/2}^o$	2600.1722	0.2239	2.700	10	10	4.17
Fe II $\lambda 2586$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6D_{7/2}^o$	2586.6494	0.06457	2.720	10	8	4.17
Fe II $\lambda 2382$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6F_{11/2}^o$	2382.7641	0.3006	3.100	10	12	4.17
Fe II $\lambda 2374$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6F_{9/2}^o$	2374.4601	0.02818	0.02818	10	10	4.17
Fe II $\lambda 2344$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6P_{7/2}^o$	2344.2128	0.1097	0.2.680	10	8	4.17
Fe II $\lambda 2249$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6D_{7/2}^o$	2249.8754	0.00182	0.04138	10	8	4.17
Fe II $\lambda 2260$	$3d^6(^5D)4s$	$3d^6(^5D)4p$	$a^6D_{9/2}$	$z^6F_{9/2}^o$	2260.7793	0.00244	0.04849	10	10	4.17
Fe II $\lambda 1611$	$3d^6(^5D)4s$	$3d^6(^3F)4p$	$a^6D_{9/2}$	$y^6F_{7/2}^o$	1611.2003	0.001345	0.007141	10	8	4.17
Fe II $\lambda 1608$	$3d^6(^5D)4s$	$3d^5(^6S)4s4p(^3P^o)$	$a^6D_{9/2}$	$y^6P_{7/2}^o$	1608.4508	0.05545	1.996	10	8	4.17

Common absorption lines, cont.

Ion/Tran	Transition		State		$\lambda_k^{k'}$ Å	$f_k^{k'}$	$\Gamma_k^{k'}$ 10^8 s^{-1}	g_k	$g_{k'}$	Diagram
	k	k'	k	k'						
NiII $\lambda 1751$	3d ⁹ (² D)4s	3d ⁸ (³ F)4p	a ² D _{5/2}	z ³ F _{7/2} ^o	1751.9157	0.0620	1.0400	6	8	4.19
NiII $\lambda 1741$	3d ⁹ (² D)4s	3d ⁸ (³ F)4p	a ² D _{5/2}	z ³ D _{5/2} ^o	1741.5531	0.1035	2.2760	6	8	4.19
NiII $\lambda 1709$	3d ⁹ (² D)4s	3d ⁸ (³ F)4p	a ² D _{5/2}	z ³ F _{5/2} ^o	1709.6042	0.06884	1.5710	6	6	4.19
NiII $\lambda 1703$	3d ⁹ (² D)4s	3d ⁸ (³ F)4p	a ² D _{5/2}	z ³ D _{3/2} ^o	1703.4119	0.01224	0.4221	6	8	4.19
ZnII $\lambda 2062$	4s	4p	² S _{1/2}	² P _{1/2} ^o	2062.6603	0.2529	3.965	1	2	
ZnII $\lambda 2026$	4s	4p	² S _{1/2}	² P _{3/2} ^o	2026.1369	0.5150	4.1840	1	4	

Common absorption lines from excited states

Ion/Tran	Transition		State		$\lambda_k^{k'}$ Å	$f_k^{k'}$	$\Gamma_k^{k'}$ 10^8 s^{-1}	g_k	$g_{k'}$	Diagram
	k	k'	k	k'						
CII* $\lambda 1335.6$	2s ² 2p	2s2p ²	² P _{3/2} ^o	² D _{3/2}	1335.6627	0.01277	2.870	4	4	4.15
CII* $\lambda 1335.7$	2s ² 2p	2s2p ²	² P _{3/2} ^o	² D _{5/2}	1335.7077	0.1149	2.846	4	6	4.15
CII* $\lambda 1037$	2s ² 2p	2s2p ²	² P _{3/2} ^o	² S _{1/2}	1037.0182	0.1230	22.900	6	4	4.15
SiII* $\lambda 1197$	3s ² 3p	3s3p ²	² P _{3/2} ^o	² P _{1/2}	1197.349	0.1243	34.95	4	2	4.16
SiII* $\lambda 1194$	3s ² 3p	3s3p ²	² P _{3/2} ^o	² P _{3/2}	1194.500	0.6233	35.03	4	4	4.16
SiI* $\lambda 1265$	3s ² 3p	3s ² 3d	² P _{3/2} ^o	² D _{3/2}	1265.002	0.1004	25.33	4	4	4.16
SiII* $\lambda 1264$	3s ² 3p	3s ² 3d	² P _{3/2} ^o	² D _{5/2}	1264.738	0.9034	25.12	4	6	4.16

Cosmological Paradigm

Our best working model of cosmology, the “Big bang”, has been highly successful in that no observations have clearly ruled it out as a viable scenario for the evolution of the universe. It is experimentally established (Hubble’s Law) that the universe is expanding, and the existence of the the cosmic microwave background (CMB) leaves little doubt that the universe was denser and hotter in the past. It is fascinating that the theoretical formalism describing a relativistic dynamic spacetime geometry, i.e., Einstein’s theory of General Relativity, was introduced prior to any experimental evidence that the universe was dynamic. Together, the Big bang model as governed by the theory of General Relativity comprise our current cosmological paradigm.

Most all models have dependence on fixed constants, or parameters; often these parameters are experimentally constrained boundary conditions, either temporal or spatial. Cosmology is no different in that it is parametrized by several constants based upon measurements performed at the present epoch. These constants are known as the present-epoch cosmological parameters, and they quantify the current dynamical behavior of the universe and the current energy density of its constituents. Key to connecting the cosmological model to real world observations, is the metric. The mathematical form of the metric reflects the curvature of the spacetime geometry and provides a means of calculating temporal and spatial intervals between events in a dynamical relativistic spacetime geometry.

In this chapter, we begin by introducing the metric for a generalized homogeneous and isotropic dynamic spacetime. The Robertson-Walker metric is the standard used for the current paradigm, and we outline its derivation. The spacetime dynamics are described by Einstein’s General Relativity field equations, which are presented and then rewritten in terms of observational cosmological parameters. We then rewrite the metric as a function of the cosmological parameters. The present epoch values of the cosmological pa-

rameters and their evolution are briefly discussed. This material serves as the foundation for Chapter 6, in which we develop practical expressions for quantitative application of the metric to observations. Recommended resources that extend beyond the scope of this chapter are the books by Peebles (1993), *Principles of Physical Cosmology*, and Peacock (1999), *Cosmological Physics*.

5.1 Geometry and the Metric

The metric supplies the formalism for computing temporal and spatial coordinate positions and separations in a given generalized geometry, of which Euclidean geometry is a special case. Even if the geometry is not static, the metric can be applied with little generalization if the dynamics of the geometry are isotropic and homogeneous.

Consider a static geometric space with arbitrary curvature, where the radius of curvature is

$$R^2 = x^2 + y^2 + z^2 + w^2 = r^2 + w^2, \quad (5.1)$$

where r is the spatial distance and w represents the departure from Euclidean distances. The distance r is mapped employing spherical coordinates with the observer at the origin. One can view r and w as orthogonal “axes”. The formalism is as follows. If space is flat, then $w = 0$, the space is simply the flat Euclidean geometry, and the radius of curvature reduces to $R^2 = r^2 = x^2 + y^2 + z^2$. If space is curved, as illustrated in Figure 5.1, then r and the radius of curvature are related through the w axis, which is orthogonal to r .

In terms of spherical coordinates,

$$\begin{aligned} x &= r \sin \theta \cos \phi & y &= r \sin \theta \sin \phi \\ z &= r \cos \theta & w &= \sqrt{R^2 - r^2}, \end{aligned} \quad (5.2)$$

where θ and ϕ are the polar and azimuthal angles in the spherical coordinate system, with $0 \leq \theta \leq \pi$ and $0 \leq \phi \leq 2\pi$. The static distance element is

$$dD^2 = dx^2 + dy^2 + dz^2 + dw^2. \quad (5.3)$$

In an isotropic and homogeneous dynamic spacetime geometry, the full metric, which is the “distance” element between events in spacetime, is written

$$ds^2 = c^2 dt^2 - a^2(t) dD^2, \quad (5.4)$$

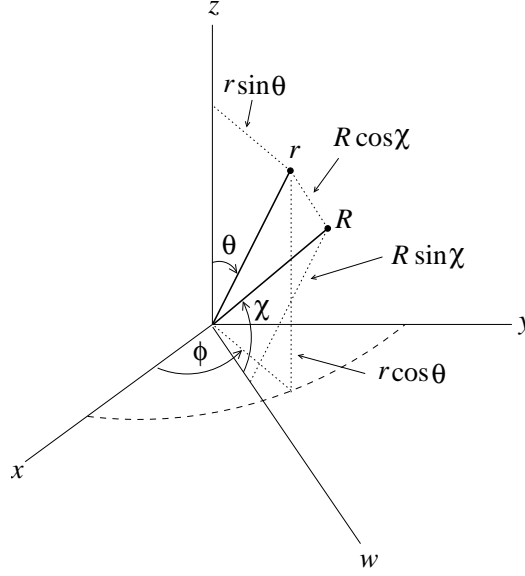


Figure 5.1 A schematic of a curved space with axes, x , y , z , and w . The angles θ and ϕ are the standard polar and azimuthal angles of the three dimensional spherical coordinate system. For a non-Euclidean (curved) space, the geometric radius of curvature is expressed $R^2 = r^2 + w^2$. The w axis is orthogonal to the r “axis” (for a flat space, the w axes has no projection). The “angle” χ parameterizes the projection of R on the r and w axes (a positive curvature is shown; the geometries are further illustrated in Figure 5.2).

where dt is the time element, c is the speed of light, giving the temporal element $d\tau$ between events, and $a(t)dD$ is the time dependent spatial distance element.

The function of proportionality, $a(t)$, is the time dependent scale factor that accounts for isotropic and homogeneous dynamics of the spatial element (i.e., expansion or contraction). The evolution of the scale factor is governed by the dynamics of the spacetime geometry, whereas the metric provides the means to compute the relative locations of and separations between spacetime events.

5.1.1 Time and Distance

In a generalized dynamic geometry, space and time are coupled through the “distance” element ds . Thus, time and distance elements are not absolute and separable quantities. Time and distance are measured quantities, and being such can only be measured in the reference frame of the observer.

The value of t appearing in $a(t)$ applies strictly in the observer's frame. One cannot talk about an absolute time element or an absolute time throughout the geometry, only the time as measured at the observer within the geometry. If we define the present time for the observer as t_0 , we can define the present value of the scale factor as $a_0 = a(t_0)$. Then, the present-time proper distance element in the observer's frame is $dl_0 = a_0 dD$. In general, the proper distance element as measured by an observer at time t is $dl = a(t) dD$, which is to say that the proper distance between events in the observer's frame change with time.

The present-time proper distance defines the physical size of the static spatial distance element, dD , for the geometry. It is convention to define $a_0 = 1$, giving $dD = dl_0$. This non-changing quantity is called the co-moving distance element, and it has the magnitude of the proper distance element at the present time of the observer.

If we consider a photon propagating through the dynamical spacetime, we can determine the relationship between the time element associated with a distant event that occurred at an earlier time as measured from an observer and the time element for the observer at the present time. For photons on radial trajectories (line of sight to the observer), we set $ds = 0$ in the metric (photons travel on so-called geodesics for which the "distance" element is null).

Applied to the metric as expressed in Eq. 5.4, the time element at the distance-past event is defined via

$$c dt = a(t) dD, \quad (5.5)$$

whereas the time element at the current epoch for the observer is

$$c dt_0 = a_0 dD. \quad (5.6)$$

Recalling that $dl_0 = dD$, so that dD is a constant of the generalized geometry, we have

$$\frac{dt_0}{dt} = \frac{a_0}{a(t)}. \quad (5.7)$$

If $a(t)$ increases with time, then $a_0/a(t) > 1$ and we see that the time element in the observer's inertial reference frame is longer than the time element in the inertial reference frame of the distant event that happened in the observer's past. This is a time dilation effect that is a general result for dynamically expanding spacetime geometries.

For example, if a distance source is pulsing photons at an interval Δt , then a present epoch observer will measure the pulse interval as $[a_0/a(t)]\Delta t$,

where t is the time at which the pulses were emitted in the observer's frame. Since time elements are relative, frequency intervals will also be relative.

5.1.2 Hubble's Law

A consequence of a dynamical spacetime is that the space local to a distant-past event will be in motion relative to the space local to the observer. Thus, objects and events will have measured velocities of recession in an expanding generalize geometry, and this velocity will increase in proportion to the spacetime interval between the event and the observation. Edwin Hubble was the first to observe this phenomenon for galaxies for which the recessional velocities were not relativistic, i.e., a few percent of the speed of light. Hubble expressed his now famous law in the form $v = H_0 l$, where v is the measured recessional velocity, l is the proper distance to the galaxy, and H_0 is the Hubble constant.

Hubble's law can be derived by expanding the scale factor near its present epoch value,

$$a(t) = a_0 + \dot{a}_0(t - t_0) + \frac{1}{2}\ddot{a}_0(t - t_0)^2 + \cdots, \quad (5.8)$$

where t_0 is the present time and $t_0 > t$. Writing Hubble's constant as $H_0 = \dot{a}_0/a_0$, and writing $\Delta t = t_0 - t$, the series can be rewritten

$$\frac{a(t)}{a_0} = 1 - H_0\Delta t - \frac{1}{2}q_0H_0^2\Delta t^2 + \cdots, \quad (5.9)$$

which provides the origin of the famous deceleration parameter,

$$q_0 = -\frac{1}{H_0^2} \frac{\ddot{a}_0}{a_0} = -\frac{\ddot{a}_0 a_0}{\dot{a}_0^2}. \quad (5.10)$$

Noting that $\Delta t = l/c$, where l is the proper distance, and dropping higher order terms, we have

$$\frac{a(t)}{a_0} = 1 - \frac{H_0}{c} l. \quad (5.11)$$

Assuming that $a(t)$ increases with time, for non-relativistic cases, the fractional recession velocity observed at t_0 from an object observed at t is given by

$$\frac{v}{c} = \frac{a_0 - a(t)}{a(t)}. \quad (5.12)$$

Rearranging we have

$$\frac{a(t)}{a_0} = \left(1 + \frac{v}{c}\right)^{-1} = 1 - \frac{v}{c}, \quad (5.13)$$

where the last step assumes $v/c \ll 1$. Equating Eqs. 5.11 and 5.13, we obtain the desired relation

$$v = H_0 l. \quad (5.14)$$

Note that the Hubble law is a direct consequence of the expanding space-time geometry. The linearity between recessional velocity and proper distance holds only for regions that are local enough to the observer that the velocities are nonrelativistic. The proportionality constant, H_0 , depends upon the time dependence of the scale factor, which will be discussed in § 5.3.

5.2 The Robertson-Walker Metric

The Robertson-Walker metric has become the standard for describing the spacetime geometry of the expanding universe. Here, we outline its derivation. The spatial coordinate elements are

$$\begin{aligned} dx &= dr \sin \theta \cos \phi + r(\cos \theta \cos \phi d\theta - \sin \theta \sin \phi d\phi) \\ dy &= dr \sin \theta \sin \phi - r(\cos \theta \sin \phi d\theta + \sin \theta \cos \phi d\phi) \\ dz &= dr \cos \theta - r \sin \theta d\theta \\ dw &= \frac{r dr}{w}. \end{aligned} \quad (5.15)$$

After some manipulation, the generalized spatial distance element is written

$$dD^2 = \frac{R^2 dr^2}{w^2} + r^2 d\psi^2, \quad (5.16)$$

where $d\psi^2 = d\theta^2 + \sin^2 \theta d\phi^2$ is the angular element between two points on a surface of constant r as viewed by an observer at the origin. For a flat geometry, $w = 0$, so that the algebraic manipulation to obtain Eq. 5.16 is not required; the distance element simply is the Euclidean element, $dD^2 = dr^2 + r^2 d\psi^2$.

It is customary to introduce the “development angle”, χ , in the r - w plane. The r - w plane and the development angle are illustrated in Figure 5.2 for both a positively curved and a negatively curved space. In a positively curved space, χ is the circular trigonometric angle that provides the projection of the geometric curvature, R , on the r and w axes. The projections of R are $r = R \sin \chi$ and $w = R \cos \chi$, and the radius of curvature is the radius of the circle,

$$r^2 + w^2 = R^2 \sin^2 \chi + R^2 \cos^2 \chi = R^2 (\sin^2 \chi + \cos^2 \chi) = R^2. \quad (5.17)$$

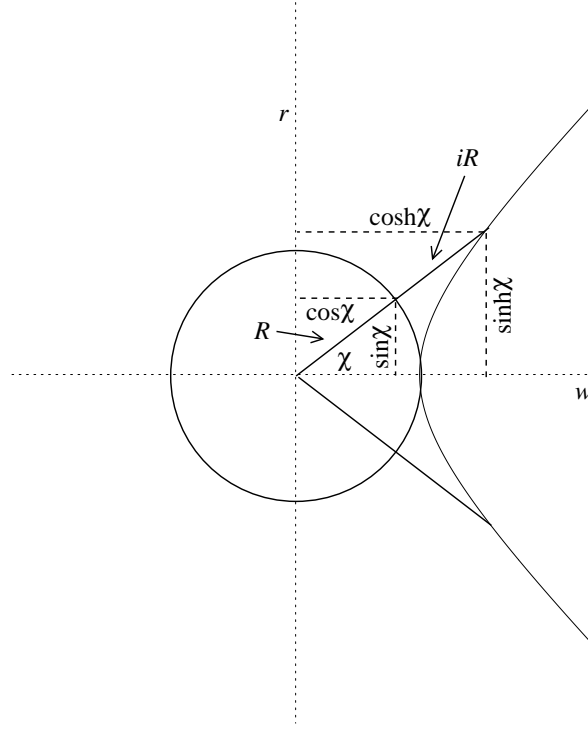


Figure 5.2 The geometric representation of the development angle χ in the r - w plane of a curved space (the context of this plane is illustrated in Figure 5.1). For a positive curvature, R and χ are real, χ is the standard trigonometric circular angle, $r = R \cos \chi$, and $w = R \sin \chi$. For a negative curvature, R and χ are imaginary, χ is the hyperbolic angle, $r = -R \sinh \chi$, and $w = iR \cosh \chi$.

A negatively curved space is described by invoking an imaginary development angle, $\chi \rightarrow i\chi$, and radius of curvature, $R \rightarrow iR$, so that $R^2 < 0$. In the imaginary r - w plane, the projections of iR are $r = iR \sin i\chi$ and $w = iR \cos i\chi$. Invoking the identities between hyperbolic and trigonometric functions, $\sinh \chi = -i \sin i\chi$ and $\cosh \chi = \cos i\chi$, we have $r = -R \sinh \chi$ and $w = iR \cosh \chi$. We see that w is the imaginary axis. The radius of curvature is defined by the hyperbola in the complex plane

$$r^2 + w^2 = R^2 \sinh^2 \chi - R^2 \cosh^2 \chi = R^2 (\sinh^2 \chi - \cosh^2 \chi) = -R^2, \quad (5.18)$$

where the real unit hyperbola is $\cosh^2 \chi - \sinh^2 \chi = 1$. Since $R^2 < 0$, we have $r^2 + w^2 > 0$. The development angle, χ , is the hyperbolic angle. The hyperbolic angle represents twice the area bounded by the hyperbola $r^2 + w^2 = -R^2$ between the w axis and the line iR (the radius of curvature) extending from the origin to the point of interception with the hyperbola.

For Euclidean geometry, $w = 0$, so that R and r are coincident. In this case, the development angle serves as a proportionality between R and r , i.e., $r = \chi R$.

The generalized parameterization is summarized as follows:

$$r = \begin{cases} R \sin \chi & \text{positive} \\ R\chi & \text{flat} \\ -R \sinh \chi & \text{negative,} \end{cases} \quad (5.19)$$

and

$$w = \begin{cases} R \cos \chi & \text{positive} \\ 0 & \text{flat} \\ iR \cosh \chi & \text{negative.} \end{cases} \quad (5.20)$$

Though χ is not a direct observable, it can be computed in terms of observables (as will be elaborated upon below). The radius of curvature is commonly expressed

$$\kappa = \frac{1}{R^2}, \quad (5.21)$$

from which the unitless curvature constant, k , is defined:

$$k = \text{sign}(\kappa) = \begin{cases} +1 & : \quad \kappa = (1/R)^2 > 0 & \text{positive} \\ 0 & : \quad \kappa = (1/R)^2 = 0 & \text{flat} \\ -1 & : \quad \kappa = (1/iR)^2 < 0 & \text{negative.} \end{cases} \quad (5.22)$$

Consider the positive curvature case; substituting Eqs. 5.19 and 5.20 for R in Eq. 5.16,

$$\begin{aligned} dD^2 &= R^2 \frac{dr^2}{w^2} + r^2 d\psi^2 \\ &= R^2 \frac{R^2 \cos^2 \chi d\chi^2}{R^2 \cos^2 \chi} + R^2 \sin^2 \chi d\psi^2 \\ &= R^2 (d\chi^2 + \sin^2 \chi d\psi^2), \end{aligned} \quad (5.23)$$

where the similar result, holds for the negative curvature case

$$\begin{aligned} dD^2 &= (iR)^2 \frac{(iR)^2 \cosh^2 \chi d\chi^2}{(iR)^2 \cosh^2 \chi} + (iR)^2 \sinh^2 \chi d\psi^2 \\ &= -R^2 (d\chi^2 + \sinh^2 \chi d\psi^2). \end{aligned} \quad (5.24)$$

Recall that $R^2 < 0$, so that $dD > 0$ holds true. For the flat case, we have

$$dD^2 = R^2(d\chi^2 + \chi^2 d\psi^2). \quad (5.25)$$

The full Robertson-Walker metric can then be written in the compact form for positive, flat, and negative curvature,

$$ds^2 = c^2 dt^2 - a^2(t) |R^2| \left[d\chi^2 + f_k^2(\chi) d\psi^2 \right], \quad (5.26)$$

where the radius of curvature can be written as an absolute value to eliminate modifying the sign in front of the spatial element, and where

$$f_k(\chi) = \begin{cases} \sin \chi & \text{positive } k = +1 \\ \chi & \text{flat } k = 0 \\ \sinh \chi & \text{negative } k = -1. \end{cases} \quad (5.27)$$

The Robertson-Walker metric will be applied throughout Chapter 6 to derive time and distance coordinates, and velocities associated with objects carried by cosmological expansion. In order to do so, we need to determine the time evolution of the scale factor, $a(t)$, and to express the radius of curvature, R , and the co-moving coordinate, χ , in terms of observables. This will be the focus of the following sections.

5.3 Relativistic Spacetime Dynamics

The dynamics of the scale factor is governed by Einstein's general relativity field equations, in which the dynamics of spacetime are coupled to the energy density and pressure of the material universe. The temporal component of the spacetime dynamics is given by

$$3 \frac{\dot{a}^2(t)}{a^2(t)} + 3 \frac{\kappa c^2}{a^2(t)} - \Lambda c^2 = 8\pi G \rho(t), \quad (5.28)$$

and the spatial component is given by

$$-2 \frac{\ddot{a}(t)}{a(t)} - \frac{\dot{a}^2(t)}{a^2(t)} - \frac{\kappa c^2}{a^2(t)} + \Lambda c^2 = 8\pi G p(t), \quad (5.29)$$

where $\ddot{a}(t)$ is the acceleration of the expansion at time t , $\dot{a}(t)$ is the expansion velocity, $a(t)$ is the scale factor, $\rho(t)$ is the energy-matter density, $p(t)$ is the pressure, and Λ , known as the cosmological constant (Carroll, Press, & Turner, 1992), is a negative density/pressure term (originally introduced to force a static solution!). Note that the units of Eqs. 5.28 and 5.29 are the inverse square of time [sec^{-2}].

Equations 5.28 and 5.29 are known as the Friedmann equations. These form a set of equations in physical cosmology that govern the expansion of space in homogeneous and isotropic models of the universe within the context of general relativity.

Rearranging Eq. 5.28, and substituting into Eq. 5.29, we have the familiar forms of the field equations,

$$\begin{aligned}\frac{\ddot{a}(t)}{a(t)} &= -\frac{4\pi G}{3} \left[\rho(t) + 3p(t) \right] + \frac{\Lambda c^2}{3} \\ \frac{\dot{a}^2(t)}{a^2(t)} &= \frac{8\pi G \rho(t)}{3} - \frac{\kappa c^2}{a^2(t)} + \frac{\Lambda c^2}{3}.\end{aligned}\tag{5.30}$$

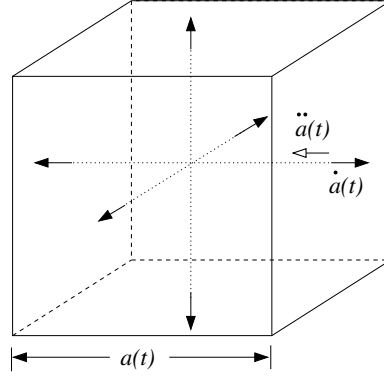


Figure 5.3 A representation of the cosmological expansion. The scale factor, $a(t)$, parameterizes the evolving metric of space between two arbitrary points as a function time. The expansion is quantified by its rate, $\dot{a}(t)$ and by its acceleration, $\ddot{a}(t)$.

A schematic of the field equations (Eq. 5.30) is presented in Figure 5.3. At time t , a representative cube of space with scale $a(t)$ is expanding at the rate $\dot{a}(t)$. The time domain is $0 \leq t \leq t_0$, where t_0 is the present time (age) of the universe. Note that $\dot{a}(t < t_0) > 0$; $a(t)$ has been increasing with time and was always smaller at earlier times. The rate of change in $\dot{a}(t)$ is $\ddot{a}(t)$, which can in principle be positive or negative (deceleration, as shown). It is the physical response of the energy density and pressure to an expanding geometry that will govern the specific time evolution of the scale factor.

The total energy density and pressure are the sums of various components,

$$\rho(t) = \sum_i \rho_i(t) \quad p(t) = \sum_i p_i(t), \tag{5.31}$$

where the indices denote

$$\begin{aligned} i = 1 & \quad \rho_m(t) \quad \text{matter} \\ i = 2 & \quad \rho_r(t) \quad \text{radiation} \\ i = 3 & \quad \rho_\Lambda(t) \quad \text{cosmological constant} . \end{aligned} \quad (5.32)$$

(Though the cosmological constant appears explicitly in the field equations, Eq. 5.30, we will see that it behaves as a negative energy density and pressure and can be formally treated in that fashion.)

In order to complete the description of the relativistic dynamics, we must incorporate the continuity equation, which describes the conservation of matter-energy in a dynamic volume element, and the equation of state relating the density and pressure of the matter-energy “fluid”. The continuity equation can be derived via differentiation and manipulation of Eqs. 5.28 and 5.29,

$$\dot{\rho}_i(t) - 3 \frac{\dot{a}(t)}{a(t)} \left[\rho_i(t) + p_i(t) \right] = 0 . \quad (5.33)$$

The equation of state for a relativistic fluid is

$$p_i(t) = \omega_i \rho_i(t) , \quad (5.34)$$

where $\omega_i = \sqrt{dp_i/d\rho_i}$ is the sound speed in the fluid. The values of ω_i are

$$\omega_i = \begin{cases} 0 & \text{matter} \\ 1/3 & \text{radiation} \\ -1 & \text{cosmological constant} . \end{cases} \quad (5.35)$$

Substituting the equation of state (Eq. 5.34) into the continuity equation (Eq. 5.33) yields

$$\dot{\rho}_i(t) = 3\rho_i(t) \frac{\dot{a}(t)}{a(t)} (1 + \omega_i) . \quad (5.36)$$

for each component. Integrating gives

$$\rho_i(t) = \rho_i \left(\frac{a(t)}{a_0} \right)^{-3(1+\omega_i)} , \quad (5.37)$$

where $\rho_i = \rho_i(t_0)$ is the energy density of the i th form at the present epoch, $t = t_0$, and a_0 is the scale factor at the present epoch. Eq. 5.37 describes the time evolution of the matter-energy density in a geometry undergoing isotropic, homogeneous contraction or expansion.

Writing out the field equations (Eq. 5.30) in terms of the sum of the

matter-density components, gives

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{4\pi G}{3} \sum_i (1 + 3\omega_i) \rho_i(t) \quad (5.38)$$

$$\frac{\dot{a}^2(t)}{a^2(t)} = \frac{8\pi G}{3} \sum_i \rho_i(t) - \frac{\kappa c^2}{a^2(t)}. \quad (5.39)$$

Carrying out the summations and applying Eq. 5.35, we have

$$\frac{\ddot{a}(t)}{a(t)} = -\frac{8\pi G}{3} \left[\frac{\rho_m}{2} \left(\frac{a(t)}{a_0} \right)^{-3} + \rho_r \left(\frac{a(t)}{a_0} \right)^{-4} - \rho_\Lambda \right], \quad (5.40)$$

$$\frac{\dot{a}^2(t)}{a^2(t)} = \frac{8\pi G}{3} \left[\rho_m \left(\frac{a(t)}{a_0} \right)^{-3} + \rho_r \left(\frac{a(t)}{a_0} \right)^{-4} + \rho_\Lambda \right] - \frac{\kappa c^2}{a^2(t)}, \quad (5.41)$$

where $\rho_m = \rho_m(t_0)$ is the matter density at the present epoch, $\rho_r = \rho_r(t_0)$ is the radiation energy at the present epoch, and $\rho_\Lambda = \rho_\Lambda(t_0)$ is the energy density of the cosmological constant at the present epoch. From inspection of Eq. 5.30, we see that

$$\Lambda c^2 = 8\pi G \rho_\Lambda. \quad (5.42)$$

5.4 Parameterized Cosmology

Since $a(t)$ appears in the Robertson-Walker metric, it is clear that the time evolution of distance intervals are linked to the field equations through the relativistic dynamics of the expansion of the universe. However, there are quantities appearing in the field equations whose values must be constrained by observations. These include ρ_m , ρ_r , ρ_Λ , and the radius of curvature $\kappa = 1/R^2$.

5.4.1 Dynamics

Evaluation of Eq. 5.41 at the present epoch, i.e., $t = t_0$, provides convenient parameterization of the present epoch energy densities and has the added benefit of yielding an expression for the radius of curvature in terms of present epoch observable quantities. We have,

$$\frac{\dot{a}_0^2}{a_0^2} \equiv H_0^2 = \frac{8\pi G}{3} [\rho_m + \rho_r + \rho_\Lambda] - \frac{\kappa c^2}{a_0^2}, \quad (5.43)$$

where the present epoch velocity of expansion per scale length is defined as the Hubble constant, $H_0 = \dot{a}_0/a_0$. Dividing Eq. 5.43 by H_0^2 yields

$$1 = \frac{8\pi G}{3H_0^2} [\rho_m + \rho_r + \rho_\Lambda] - \frac{\kappa c^2}{a_0^2 H_0^2}. \quad (5.44)$$

The form of Eq. 5.44 is the origin of the well-known present epoch dimensionless cosmological density parameters

$$\begin{aligned} \Omega_m &= \frac{8\pi G \rho_m}{3H_0^2} & \Omega_r &= \frac{8\pi G \rho_r}{3H_0^2} \\ \Omega_\Lambda &= \frac{\Lambda c^2}{3H_0^2} & \Omega_k &= \frac{-\kappa c^2}{a_0^2 H_0^2}, \end{aligned} \quad (5.45)$$

where Ω_Λ follows from Eq. 5.42. Employing this notation, Eq. 5.44 is rewritten

$$\Omega_m + \Omega_r + \Omega_\Lambda + \Omega_k = 1. \quad (5.46)$$

Given that Ω_k depends upon the radius of curvature, we can rearrange Eq. 5.46 to obtain R in terms of Ω_m , Ω_r , and Ω_Λ ,

$$\Omega_k = \frac{-\kappa c^2}{a_0^2 H_0^2} = 1 - (\Omega_m + \Omega_r + \Omega_\Lambda). \quad (5.47)$$

where

$$\begin{aligned} \Omega_k &< 0 & (k = +1, \text{positive, closed}) \\ \Omega_k &= 0 & (k = 0, \text{flat}) \\ \Omega_k &> 0 & (k = -1, \text{negative, open}). \end{aligned} \quad (5.48)$$

Invoking Eq. 5.21, we have

$$R^2 = \left(\frac{c}{a_0 H_0} \right)^2 \frac{1}{\Omega_k} = \frac{1}{a_0^2} \frac{D_H^2}{\Omega_k} = \frac{D_H^2}{1 - \Omega_m - \Omega_r - \Omega_\Lambda}, \quad (5.49)$$

where we have adopted the convention $a_0 = 1$ and defined

$$D_H = \frac{c}{H_0}, \quad (5.50)$$

which is called the Hubble distance. The cosmological density parameters given in Eq. 5.45-5.49 apply to the present epoch. Note that R is a constant of the cosmology.

The full expression of the field equations, Eqs. 5.40 and 5.41, can now be highly simplified in form. Multiplying Eq. 5.41 by H_0^2/H_0^2 and then taking the square root, the expansion velocity per unit scale length is written

$$\frac{\dot{a}(t)}{a(t)} \equiv H(t) = H_0 E(t), \quad (5.51)$$

where $H(t)$ is the time evolution of the Hubble constant, and where

$$E(t) = \left[\Omega_m \left(\frac{a_0}{a(t)} \right)^3 + \Omega_r \left(\frac{a_0}{a(t)} \right)^4 + \Omega_k \left(\frac{a_0}{a(t)} \right)^2 + \Omega_\Lambda \right]^{1/2}, \quad (5.52)$$

which has the limiting value $E(t_0) = 1$, which recovers Eq. 5.46. The range of $E(t)$ is $E(t_0) = 1$ and $E(t < t_0) > 1$. Multiplying Eq. 5.40 by H_0^2/H_0^2 , the acceleration per unit scale length is written

$$\frac{\ddot{a}(t)}{a(t)} \equiv \dot{H}(t) = H_0^2 E'(t), \quad (5.53)$$

where $\dot{H}(t)$ is the derivative of the time evolution of the Hubble constant, and where

$$E'(t) = \Omega_\Lambda - \frac{1}{2} \Omega_m \left(\frac{a_0}{a(t)} \right)^3 - \Omega_r \left(\frac{a_0}{a(t)} \right)^4. \quad (5.54)$$

The allowed range of $E'(t)$ is, in principle, $-\infty \leq E'(t) \leq \infty$. Note that the time evolution of Eqs. 5.51 and 5.53 is expressed in terms of the *present epoch* values of Ω_m , Ω_r , Ω_k , and Ω_Λ scaled by the time evolution of their respective energy densities via the continuity equation, Eq. 5.37.

The values of each Ω are addressed in § 5.5. As we will discuss in § 5.5.1, $\Omega_r \ll \Omega_m < \Omega_\Lambda$. As such, it is common practice to omit Ω_r in the application of Eqs 5.52 and 5.54.

5.4.2 Friedmann-Einstein Cosmologies

When Einstein set the cosmological constant to zero in the Friedmann equations (Eqs. 5.28 and 5.29), the resulting cosmological a model became known as the Friedmann-Einstein universe. Note that there is no constraint on the curvature in this model.

On account that the Friedmann-Einstein ($\Lambda = 0$) cosmology has been applied often in the scientific literature prior to the year 2000, it is useful to review the expansion evolution and commonly used cosmological parameters for this model. Eq. 5.52 reduces to

$$E(t) = \left[\Omega_m \left(\frac{a_0}{a(t)} \right)^3 + \Omega_k \left(\frac{a_0}{a(t)} \right)^2 \right]^{1/2}, \quad (5.55)$$

and Eq. 5.54 reduces to

$$E'(t) = -\frac{1}{2} \Omega_m \left(\frac{a_0}{a(t)} \right)^3, \quad (5.56)$$

where we have omitted Ω_r . Applying the historical use of $\Omega_0 = 1 - \Omega_k$ and the relation $\Omega_m + \Omega_k = 1$, yields

$$\Omega_0 = \Omega_m. \quad (5.57)$$

That is, in a Friedmann-Einstein cosmology, the curvature of the universe was dependent strictly upon the matter density parameter. It has been a long time convention to write $\Omega_0 = \rho_m/\rho_c$, where

$$\rho_c = \frac{3H_0^2}{8\pi G} = 1.88 h^2 \times 10^{-29} \text{ g cm}^{-3}, \quad (5.58)$$

($h = H_0/100$) defines the critical matter density for which $\Omega_0 = 1$ (zero curvature universe).

In this special case, the terms $E(t)$ and $E'(t)$ that describe the expansion velocity (Eq. 5.55) and acceleration (Eq. 5.56) depend only upon Ω_0 , which is often expressed as the deceleration parameter (see Eq. 5.10)

$$q_0 \equiv -\frac{\ddot{a}_0 a_0}{\dot{a}_0^2} = \frac{1}{2}\Omega_0, \quad (5.59)$$

where \ddot{a}_0 , \dot{a}_0 , and a_0 are the acceleration, velocity, and scale factor of cosmological expansion at the present epoch.

Invoking $\Omega_0 = 1 - \Omega_k$ and substituting Eqs. 5.57 and 5.59 into Eqs. 5.55 and 5.56, yields

$$E(t) = \frac{a_0}{a(t)} \left[1 + 2q_0 \left(\frac{a_0}{a(t)} - 1 \right) \right]^{1/2}, \quad (5.60)$$

where $E(t_0) = 1$, and

$$E'(t) = -q_0 \left(\frac{a_0}{a(t)} \right)^3, \quad (5.61)$$

which are substituted into Eq. 5.51 to obtain the time evolution of the expansion velocity per unit scale length and Eq. 5.53 to obtain the time evolution of the expansion acceleration per unit scale length.

Following the steps to derive the curvature radius, R , in terms of the cosmological constants (Eq. 5.49), we have

$$R^2 = \left(\frac{c}{a_0 H_0} \right)^2 \frac{1}{\Omega_k} = \frac{1}{a_0^2} \frac{D_H^2}{\Omega_k} = \frac{D_H^2}{1 - \Omega_m}, \quad (5.62)$$

which, again, is a constant of the cosmology.

Prior to the more precise measurements of the cosmological parameters, the value of Ω_0 was not well constrained. Observations at that time indicated that the matter density was closer to zero than to unity. Thus, many

researchers favored a low-density or open cosmology, described by the two parameters

$$\begin{aligned}\Omega_0 &= 0.1 && \text{Low Density (open)} \\ q_0 &= 0.05,\end{aligned}\tag{5.63}$$

for which $\Omega_k = 1 - \Omega_0 = 0.9$. On the other hand, inflationary theory mandated $\Omega_0 = 1$, and many researchers favored the so called Einstein-de Sitter cosmology, which is a zero curvature, $\Omega_\Lambda = 0$, matter dominated cosmology defined by

$$\begin{aligned}\Omega_0 &= 1.0 && \text{Einstein-de Sitter (flat)} \\ q_0 &= 0.5.\end{aligned}\tag{5.64}$$

5.4.3 The Robertson-Walker Metric

The compact form of the Robertson-Walker metric (Eq. 5.26) is

$$ds^2 = c^2 dt^2 - a^2(t) |R^2| \left[d\chi^2 + f_k^2(\chi) d\psi^2 \right],\tag{5.65}$$

where $f_k^2(\chi)$ is given by Eq. 5.27. Substituting Eq. 5.49 for R^2 , we have

$$ds^2 = c^2 dt^2 - \frac{a^2(t)}{a_0^2} \frac{D_{\text{H}}^2}{\Omega_k} \left[d\chi^2 + f_k^2(\chi) d\psi^2 \right].\tag{5.66}$$

We now examine each of the individual terms in Eq. 5.66. Rewriting the metric, we define three main components,

$$ds^2 = d\tau^2 - (dl^2 + l_{\text{T}}^2 d\psi^2),\tag{5.67}$$

where

$$\begin{aligned}d\tau &= c dt && \text{temporal} \\ dl &= \frac{a(t)}{a_0} \frac{D_{\text{H}}}{\sqrt{|\Omega_k|}} d\chi && \text{spatial radial (proper)} \\ l_{\text{T}} d\psi &= \frac{a(t)}{a_0} \frac{D_{\text{H}}}{\sqrt{|\Omega_k|}} f_k(\chi) d\psi && \text{spatial transverse (proper)}.\end{aligned}\tag{5.68}$$

Each of these terms have specific applications. The temporal component provides the the time difference between events from the point of view of the observer. The radial term is employed to determine line of sight distances to events and objects, whereas the transverse term is employed to determine the distances between events and objects that are aligned perpendicular to the line of sight (it is implied that they reside in the same inertial frame). In

the transverse case, the events or objects will appear separated by an angle $\Theta = \int d\psi$. Note that the radial proper distance element is proportional to the development angle element, $d\chi$, and the transverse proper distance element is proportional to the curvature dependent function of the development angle, $f_k(\chi)$.

The expression for the temporal component, $d\tau$, can be obtained directly from the definition $da/dt = \dot{a}(t)$. We have

$$d\tau = c dt = c \frac{da}{\dot{a}(t)} = c \frac{a(t)}{\dot{a}(t)} \frac{da}{a(t)} = \frac{c}{H(t)} \frac{da}{a(t)} = D_H \frac{da}{a(t)E(t)}, \quad (5.69)$$

where we invoked Eqs. 5.51 and 5.50.

Since the operational difference between proper and co-moving coordinates is simply the proportionality of the scale factor, we can refine the distinction between proper and co-moving elements by writing

$$dl = \frac{a(t)}{a_0} dD_c, \quad l_\tau d\psi = \frac{a(t)}{a_0} D_\tau d\psi, \quad (5.70)$$

where

$$\begin{aligned} dD_c &= \frac{D_H}{\sqrt{|\Omega_k|}} d\chi && \text{radial (co-moving)} \\ D_\tau &= \frac{D_H}{\sqrt{|\Omega_k|}} f_k(\chi) && \text{transverse (co-moving)}. \end{aligned} \quad (5.71)$$

As with the proper distances, the definitions of proper and transverse co-moving distances are applied to determine radial and non-radial spatial intervals, respectively. Note that the co-moving distance elements are time independent and are equivalent to the proper distance elements at the present epoch.

5.4.4 The Development Angle

The functional form of the development angle, χ , is required to complete the formalism of the metric as described in terms of the proper and co-moving radial and transverse distances. The development angle quantifies the degree of departure of the geometric curvature from Euclidean space (see §§ 5.1 and 5.2); therefore the development angle describes the spacetime geodesic, or equivalently, the path of a photon.

Consider the metric for a photon traveling on a radial trajectory ($d\psi = 0$) to the observer. Since photons propagate along geodesics, we have $ds = 0$ in the metric. Setting $ds = 0$ and $d\psi = 0$ in Eq. 5.66, we see that the coupling

of spacetime for a geodesic is therefore

$$d\tau = \frac{a(t)}{a_0} \frac{D_{\text{H}}}{\sqrt{|\Omega_k|}} d\chi, \quad (5.72)$$

which, after substituting Eq. 5.69 for $d\tau$, yields the element of the development angle in terms of the cosmological parameters and the scale factor,

$$d\chi = \sqrt{|\Omega_k|} \frac{a_0 da}{a^2(t)E(t)}. \quad (5.73)$$

Integrating $d\chi$ yields

$$\chi = \sqrt{|\Omega_k|} \int_{a_0}^{a(t)} \frac{a_0 da}{a^2(t)E(t)}, \quad (5.74)$$

where t is the time (as measured by the observer) at which the photon was emitted or absorbed (the time of this event will be in the observer's past).

In Chapter 6, we will write the above metric elements and integrated quantities in terms of an observable quantity, the redshift; and, as we shall see, the radial co-moving distance is directly expressible as a function of redshift¹.

5.5 Λ CDM, H_0 , and the Ω s

The history of the increasingly more accurate and precise measurements of Hubble's constant and the various density parameters, as well as the the history of our improved understanding of dark matter, is very rich and laced with highly technical details far beyond the scope of this treatment. Here, only a very thin surface coverage of these results are presented (see Huterer & Shafer, 2018, for a thorough summary of the current status of observational cosmology) .

Of the cosmological parameters appearing in the field equations, the *present epoch* expansion velocity per unit scale length, Hubble's constant, H_0 (with fundamental unit sec^{-1}), is of central importance. The factor $1/H_0^2$ is a factor of proportionality for each of the density parameters (Eq. 5.45), and is thus a very critical quantity. It is customary to parameterize the Hubble constant by the dimensionless Hubble parameter,

$$h = \frac{H_0}{(100 \text{ km s}^{-1} \text{ Mpc}^{-1})}, \quad (5.75)$$

¹ For a flat geometry, the curvature is $R = \infty$ and $\Omega_k = 0$. It would seem that the metric, as written in Eqs. 5.65 and 5.66, explodes for a zero curvature cosmology. However, as we shall see in Chapter 6, the Ω_k cancel for a flat geometry.

where H_0 is expressed in units $\text{km s}^{-1} \text{Mpc}^{-1}$. For the majority of the latter half of the twentieth century, h was known to range between 0.5 and 1.0. The problem was that the distribution of measured values was bimodal, with one mode near 0.5 and the second mode near 1.0. With the advent of the *Hubble Space Telescope* in the 1990s, highly detailed work suggested that $H_0 \simeq 0.72 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Freedman et al., 2001), but has more recently yielded $H_0 = 74.3 \pm 2.1 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Freedman et al., 2012). A slightly smaller value of $H_0 = 73.48 \pm 1.66 \text{ km s}^{-1} \text{Mpc}^{-1}$ has been reported (Riess et al., 2016, 2018). The nine-year *WMAP* combined data results is $H_0 = 69.33 \pm 0.88 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Hinshaw et al., 2013), a downward change from the five-year result of $H_0 = 70.5 \pm 1.3 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Komatsu et al., 2009). Most recently, the *Planck* combined data result is $H_0 = 67.74 \pm 0.0046 \text{ km s}^{-1} \text{Mpc}^{-1}$ (Planck Collaboration et al., 2016b). On occasion, we will adopt the latter value².

Observational constraints on baryonic matter and observational evidence for dark matter suggests that the total matter density is the sum of baryonic and non-baryonic (dark) components,

$$\Omega_m = \Omega_b + \Omega_c, \quad (5.76)$$

where Ω_b and Ω_c are the baryonic and dark matter density parameters, respectively. The identification of the dark matter remains an unsolved mystery.

Via its gravitational influence, dark matter is key to the formation and evolution of structure in the universe. The velocity dispersion of this mystery material strongly affects the mass scales on which gravitational instabilities form and the order in which they grow. Higher velocity dispersion dark matter, which suppresses the early formation of smaller mass scale structures, was dubbed “hot dark matter”, or HDM, and lower velocity dispersion dark matter, which promotes the development of smaller mass scale structures at earlier times, was dubbed “cold dark matter”, or CDM³.

After several years of research employing computerized cosmological simulations of structure growth, which are constrained by observations of the $\text{Ly}\alpha$ forest and the large scale distribution of luminous matter, cold dark

² Note, however, that the *Hubble* values (local universe measurements) are systematically larger than the *WMAP* and *Planck* values (high-redshift measurements); at the time of this writing, this is a tension with a statistical significance as high as 3.7σ (Riess et al., 2018).

³ Hot and Cold actually refer to whether the velocities are relativistic ($v \simeq c$) at the epoch of radiation-matter equality (see below). For CDM, $v \ll c$ at this epoch. In addition, dark matter is non-baryonic, dissipationless (cannot cool via radiation), and collisionless (does not interact with itself, baryonic matter, or other matter). Dark matter is only detectable via its gravitational influence.

matter as become strongly favored, but is not without its challenges. These models became known as CDM cosmologies.

Evidence for a non-zero cosmological constant emerged in the late 1990s, when it was discovered that the magnitudes of supernovae at cosmological distances did not match the expectations for $\Lambda = 0$ (e.g., Riess et al., 1998; Schmidt et al., 1998; Perlmutter et al., 1999a); supernovae reside at cosmological distances larger than predicted by a non-zero Λ . In Einstein's field equations, Λ acts as an acceleration term that counters gravitation deceleration; presently the cosmological expansion is accelerating. The energy density of the cosmological constant has been dubbed the “dark energy”. As the supernova results began to be accepted, the reigning Friedmann-Einstein cosmologies were replaced by the so-called Λ CDM cosmology (see Frieman et al., 2008, for a comprehensive review).

The combined constraints from Big Bang Nucleosynthesis (BBN), measurements of the matter distribution of galaxies and galaxy clusters, observations of Type Ia supernovae at cosmological distances, and analysis of the angular anisotropy and baryon acoustic oscillations of the cosmic background radiation (CMB), have provided the coveted precision in the cosmological parameters.

Deuterium abundances in quasar spectra have been applied to constrain Ω_b in the context of the standard BBN model and the distributions of galaxies on large scales and in clusters have been applied to constrain both Ω_b and Ω_c . CMB anisotropies, from experiments such as the *Cosmic Background Explorer* (COBE, Boggess et al., 1992), the balloon-borne *Millimeter Anisotropy eXperiment IMaging Array* (MAXIMA, Hanany et al., 2000), the two *Boomerang* flights (Netterfield et al., 2002), the *Degree Angular Scale Interferometer* (DASI, Halverson et al., 2002), the Cosmic Background Imager (CBI, Sievers et al., 2003), the *Wilkinson Microwave Anisotropy Probe* (WMAP, Spergel et al., 2003), and *Planck* (Planck Collaboration et al., 2016a) have ushered in the age of precision cosmology by measuring h , Ω_b , Ω_c , and Ω_k (with uncertainties on the order of 1% or better), and placing stringent constraints on Ω_r and Ω_Λ .

Many of the experiments yield a correlated dependency with the Hubble parameter. For example, BBN is sensitive to the product of baryon density and the square of the Hubble parameter, $\Omega_b h^2$, via the cosmic deuterium abundance. Both the baryon density, $\Omega_b h^2$, and the baryon to total matter ratio, Ω_b/Ω_m , can be individually measured in large volume galaxy surveys from the inhomogeneity of baryons. The anisotropy of the CMB has similar, yet unique, dependencies upon $\Omega_b h^2$ and $\Omega_m h^2$.

In Table 5.1, we list the current values of the cosmological parameters

with their 68% confidence limits for the standard Λ CDM model. Unless noted, the values are taken from Planck CMB power spectra (Planck Collaboration et al., 2016b) in combination with lensing, baryon acoustic oscillation, supernovae, and H_0 data.

Table 5.1 Λ CDM Cosmological Parameters ^a

Parameter	Symbol	Value	CL
Hubble Parameter	h	0.6774 ± 0.0046	68%
Baryonic Matter	Ω_b	0.0486 ± 0.0010	68%
Dark Matter	Ω_c	0.2589 ± 0.0057	68%
Total Matter	Ω_m	0.3089 ± 0.0062	68%
Cosmological Constant	Ω_Λ	0.6911 ± 0.0062	68%
Critical Density, g cm^{-3}	ρ_c	$(8.62 \pm 0.12) \times 10^{-30}$	68%
Curvature ^b	Ω_k	$0.0027^{+0.0039}_{-0.00389}$	95%
Age, Gyr	t_0	13.799 ± 0.021	68%

^a Planck Collaboration et al. (2016b), unless noted

^b Hinshaw et al. (2013)

The values presented in Table 5.1 are the present-epoch boundary conditions describing the matter and energy density of the universe, which govern both the dynamics of the scale factor (via Eqs. 5.52 and 5.54) and provide the curvature of the spacetime geometry (via Eq. 5.49). Since they appear in the metric, the cosmological parameters also govern the calculation of time and distance intervals. Note that $\Omega_k = 0$ to high precision, indicating that the geometry is flat ($R = \infty$). Also note that Ω_Λ dominates the energy density of the universe at the present epoch. As we describe in the next section, these parameters have evolved.

5.5.1 Evolution of the Ω s

The cosmological evolution of each Ω is not only coupled to the equation of state of their respective densities (Eq. 5.37) but also to the evolution of the Hubble constant, $H(t)$. Consider the definition of the present epoch Ω_m and the time dependent equation of state of the matter density,

$$\Omega_m = \frac{8\pi G}{3H_0^2} \rho_m, \quad \rho_m(t) = \rho_m \left(\frac{a_0}{a(t)} \right)^3, \quad (5.77)$$

from which we obtain

$$\rho_m(t) = \Omega_m \frac{3H_0^2}{8\pi G} \left(\frac{a_0}{a(t)} \right)^3. \quad (5.78)$$

Explicitly writing the time dependence for Ω_m ,

$$\Omega_m(t) = \frac{8\pi G}{3H^2(t)} \rho_m(t), \quad (5.79)$$

and substituting $\rho_m(t)$ in terms of Ω_m , we have

$$\Omega_m(t) = \left(\frac{a_0}{a(t)} \right)^3 \frac{H_0^2}{H^2(t)} \Omega_m. \quad (5.80)$$

From Eq. 5.51, the ratio $H_0^2/H^2(t) = 1/E^2(t)$. Applying similar steps to each Ω yields

$$\begin{aligned} \Omega_m(t) &= \left(\frac{a_0}{a(t)} \right)^3 \frac{\Omega_m}{E^2(t)} & \Omega_r(t) &= \left(\frac{a_0}{a(t)} \right)^4 \frac{\Omega_r}{E^2(t)} \\ \Omega_\Lambda(t) &= \frac{\Omega_\Lambda}{E^2(t)} & \Omega_k(t) &= \left(\frac{a_0}{a(t)} \right)^2 \frac{\Omega_k}{E^2(t)}. \end{aligned} \quad (5.81)$$

The present epoch value of Ω_r , as listed in Table 5.1, deserves some additional discussion. The radiation density must account for the contribution made by relativistic neutrinos. The temperature dependence of the neutrino density is $\rho_\nu = (7\pi^2/120)N_\nu T_\nu^4$, where N_ν is the effective number of neutrino species, and the neutrino temperature follows $T_\nu = (4/11)^{1/3}T_\gamma$, where T_γ is the photon temperature. We have

$$\Omega_r = (1 + 0.2271N_\nu) \Omega_\gamma, \quad (5.82)$$

where $\Omega_\gamma = 4.968 \times 10^{-5}$ is the present epoch energy density of photons for $T_{\text{cmb}} = 2.725$ K. The standard value for N_ν is 3.04, yielding $\Omega_r = 1.69 \Omega_\gamma$, the value presented in Table 5.1. However, if we adopt, for example, the best value from the five-year⁴ *WMAP* results is $N_\nu = 4.4 \pm 1.5$, yielding $\Omega_r = (2.00 \pm 0.34) \Omega_\gamma$.

A fundamental physical property of the universe is the epoch of radiation and matter equality. The scale factor at this epoch, $a(t_{\text{eq}})$, occurs at

$$\frac{\Omega_r(t_{\text{eq}})}{\Omega_m(t_{\text{eq}})} = \frac{\Omega_r}{\Omega_m} \frac{a_0}{a(t_{\text{eq}})} = 1, \quad (5.83)$$

⁴ Note that the nine-year result is $N_\nu = 3.84 \pm 0.40$ (Hinshaw et al., 2013).

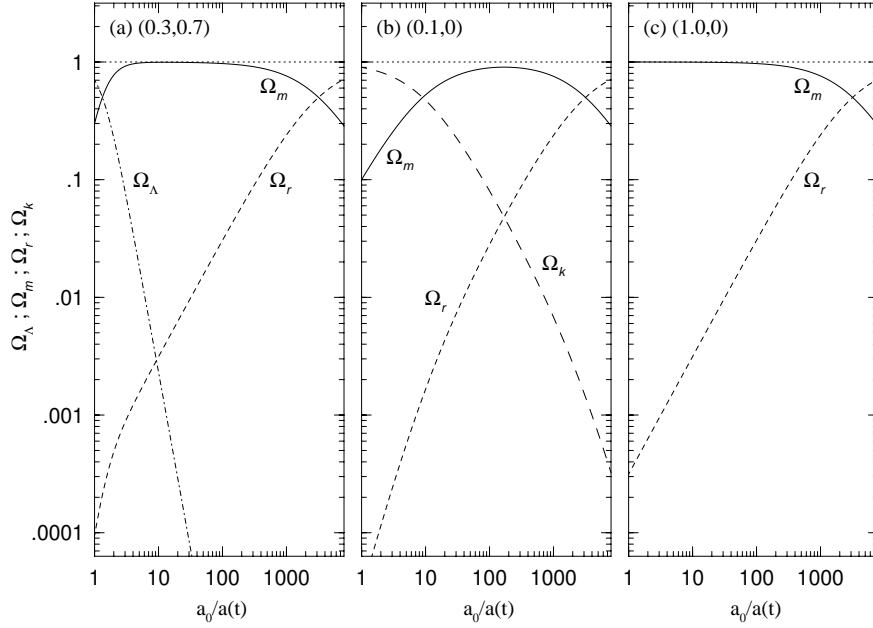


Figure 5.4 The evolution of Ω_m (solid curves), Ω_r (short dashed curves), Ω_k (long dashed curve), and Ω_Λ (dashed-dotted curve) from $a_0/a(t) = 8000$ to the present epoch, $a_0/a(t_0) = 1$. The relationship between Ω_r and Ω_m is constrained by $a_0/a(t_{\text{eq}}) = 3180$ and $N_\nu = 3.04$ (see Eq. 5.84). Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$. (a) the Λ cosmology (0.3, 0.7). (b) the low-density cosmology (0.1, 0). (c) Einstein-de Sitter cosmology (1.0, 0).

which yields (Komatsu et al., 2009)

$$\frac{a_0}{a(t_{\text{eq}})} = \frac{\Omega_m}{\Omega_r} = \frac{\Omega_m}{\Omega_\gamma} \frac{1}{1 + 0.2271 N_\nu} = 3177^{+151}_{-150}. \quad (5.84)$$

The ratio of radiation energy density to matter energy density, $\Omega_r(t)/\Omega_m(t)$ was unity when $a_0/a(t_{\text{eq}}) = 3180$. For the present epoch value of Ω_m the ratio is roughly $\Omega_r/\Omega_m \simeq 3 \times 10^{-4}$ and it is decreasing. Since Ω_r equals only 0.03% of Ω_m , and $\Omega_k = 0$ to high precision, it is common practice to omit Ω_r and Ω_k from Eqs. 5.52 and 5.54. Observational spectroscopy is firmly limited to the domain $a_0/a(t) < 7$. Thus, Ω_r can be neglected for such work.

The time evolution of $\Omega_m(t)$, $\Omega_r(t)$, and $\Omega_\Lambda(t)$ is illustrated in Figure 5.4(a) over the range $a_0/a(t) = 8000$ to the present epoch, $a_0/a(t_0) = 1$. For presentation purposes (refer to § 5.6), we adopt $\Omega_m = 0.3$, $\Omega_\Lambda = 0.7$, $\Omega_k = 0$, and $a_0/a(t_{\text{eq}}) = 3180$. Note that because the equation of state of the cosmological constant density is independent of the scale factor, the evolution of $\Omega_\Lambda(t)$ is governed strictly by the evolution of $E(t)$, given by Eq. 5.52.

In Figures 5.4(b) and 5.4(c), the time evolution of $\Omega_m(t) = \Omega_0(t)$, $\Omega_r(t)$,

and $\Omega_k(t)$ is plotted for the low density ($\Omega_m = 0.1, \Omega_\Lambda = 0$) and Einstein-de Sitter (1.0, 0) cosmologies, respectively. As with the Λ cosmology, we adopt a scale factor for matter radiation equality of $a_0/a(t_{\text{eq}}) = 3180$. For the Einstein-de Sitter cosmology, $\Omega_k(t) = 0$ and the sum the matter and radiation evolve according to $\Omega_m(t) + \Omega_r(t) = 1$. For the low density universe, the geometry is open, so $\Omega_k(t) > 0$ and the curvature density evolves to dominate at the present epoch, $\Omega_k = 0.9$

In each of the three cosmologies illustrated in Figure 5.4, the dominant Ω changes with redshift. For the highest redshifts, all three cosmologies are characterized as radiation dominated universes, following which they experience an epoch of matter dominance. The Einstein-de Sitter cosmology (Figure 5.4(c)) remains matter dominated to the present epoch. In the low-density cosmology (Figure 5.4(b)), the epoch of matter dominance ended near $z \sim 10$, following which that universe is curvature dominated to the present epoch. In the Λ cosmology (Figure 5.4(a)), the epoch of matter dominance ended near $z \sim 0.7$ and is followed by Λ dominance.

5.6 The “737 Cosmology”

In the latter half of the twentieth century, the adoption of a fairly broad range of cosmological parameters was further exasperated by poor constraints on the Hubble parameter, which ranged between $h = 0.5\text{--}1.0$; thus, various research groups applied different cosmological models to their results, which propagated to the implications of their findings (time lines, volumes, densities, distances, luminosities, masses, and velocities, etc.).

Commonly, but not exclusively, those who applied the low-density $\Lambda = 0$ cosmology, adopted $h \simeq 0.5$, where as those who applied the Einstein-de Sitter cosmology adopted $h = 1$. That is, there was the $H_0 = 50 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0.5 \text{ camp}$ and the $H_0 = 100 \text{ km s}^{-1} \text{ Mpc}^{-1}$, $q_0 = 0.05 \text{ camp}$.

This alone was a less than satisfying state of affairs. However, with the adoption of a non-zero Λ cosmology, there is now the added burden of translating previously published quantities from these two commonly adopted $\Lambda = 0$ cosmologies.

As the adopted values of the cosmological parameters presented in Table 5.1 have been continuously refined (and no doubt will have been further refined from the time of this writing and the time of your reading!). To avoid continually having to translate results in the literature, many researchers adopted cosmological parameters rounded to the nearest tenth, i.e.,

$$h = 0.7 \quad \Omega_m = 0.3 \quad \Omega_\Lambda = 0.7 \quad \Omega_k = 0.0 \quad \Omega_r = 0.0. \quad (5.85)$$

This set of parameters has been coined the “737” cosmology⁵. Throughout this work, we will adopt the “737” cosmological parameter set for schematics and illustrations of the Λ cosmology, though we occasionally adopt the value published by Planck Collaboration et al. (2016b) when quoting precise numerical values.

⁵ To the best of the author’s knowledge, D. Turnshek and S. Rao should be credited.

6

Applied Cosmology

The context for interpreting cosmological absorption line data is found within the cosmological paradigm of the universe. The primary postulates of the cosmological paradigm are that space and time comprise a four-dimensional spacetime, and that space is dynamic in reflex to the matter-energy density content of the universe. It is also assumed the dynamics are isotropic and homogeneous.

The spatial dynamics are parameterized by the time-dependent scale factor, $a(t)$, the evolution of which obeys Einstein's field equations of General Relativity. In compact form, the time evolution of the velocity and acceleration of the scale factor is (Eqs. 5.51 and 5.53)

$$\frac{\dot{a}(t)}{a(t)} = H(t) \quad \text{and} \quad \frac{\ddot{a}(t)}{a(t)} = \dot{H}(t), \quad (6.1)$$

respectively, where $H(t)$ is proportional to Hubble's constant, $H_0 = \dot{a}_0/a_0$, the present day value, and $\dot{H}(t)$ is the time derivative of $H(t)$. Observations reveal that $a(t)$ has steadily increased with time, i.e., there is no observational evidence that it has ever experienced an epoch of contraction (however, the expansion rate has not been constant with time).

The scale factor is a time dependent multiplicative factor to the spatial component of the metric, which relates temporal and spatial intervals in a dynamic spacetime geometry (with arbitrary curvature). Armed with the time evolution of $H(t)$ and $\dot{H}(t)$, and thus of $a(t)$, time and distance intervals between events and objects can be determined as a function of "cosmic time" (a misnomer we clarify in § 6.2).

In this chapter, useful relations applied in the study of cosmological objects will be presented. An additional recommended resource is the short paper by D. Hogg (1999), "Distance measures in cosmology".

6.1 Redshift

Time is not a direct observable. However, the redshifting of photons due to cosmological expansion is a directly observable quantity. The cosmological redshift, z , as measured in spectra is

$$\frac{\lambda_o}{\lambda_r} = 1 + z, \quad (6.2)$$

where λ_o is the observed wavelength, and λ_r is the emitted rest-frame wavelength.

It is incorrect to interpret the observed wavelength shift as a classical velocity Doppler shift, $(\lambda_o - \lambda_r)/\lambda_r = v/c$, or even a relativistic Doppler shift. The cosmological redshift is induced by the cosmological expansion of space itself. As a photon propagates over a very long period of “cosmic time”, the expansion of space also expands the photon wavelength (visualize a “stretching” of the light).

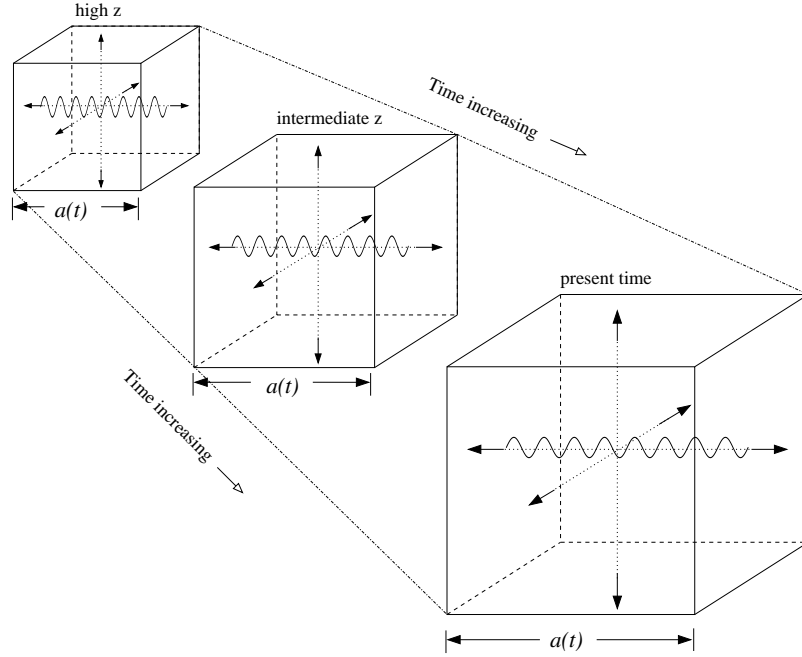


Figure 6.1 A schematic of cosmological expansion illustrating the increase in the wavelength of a photon as it propagates. Photons are locked into the cosmological expansion. As the universe expands, the wavelength increases in lock step with the scale factor, i.e. if the photon has wavelength λ_r at the epoch with scale factor $a(t)$, it will be observed to have wavelength λ_o at the present epoch with $a(t) = a_0$, such that $\lambda_o/\lambda_r = a_0/a(t)$.

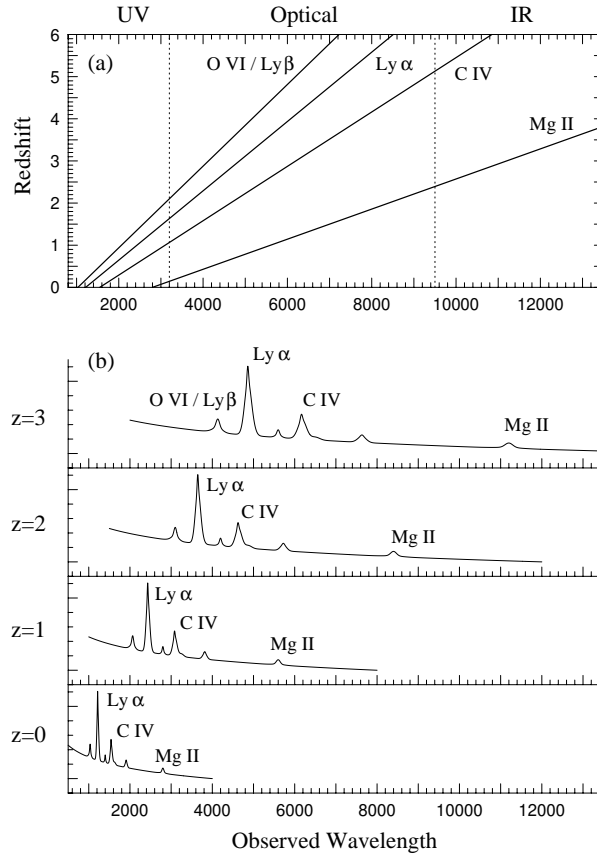


Figure 6.2 (a) The observed wavelength of redshifted absorption or emission lines having rest frame wavelengths λ_{1030} (O VI / Ly β), λ_{1215} (Ly α), λ_{1550} (C IV), and λ_{2800} (Mg II) as a function of redshift from $z = 0$ to $z = 6$. The ultraviolet, optical, and infrared approximate band passes are shown. (b) An example of a redshifted quasar spectrum from $z = 0$ to $z = 3$ illustrating the shifting of the emission lines. Note that the spectrum is significantly stretched.

As illustrated in Figure 6.1, consider a photon of wavelength λ_r emitted at time t when the scale factor of the universe was $a(t)$. As time passes, $a(t)$ increases; at the present epoch the scale factor is a_0 . Photons are locked into the cosmological expansion. The photon wavelength will increase with the expansion of space such that $\lambda_o/\lambda_r = a_0/a(t)$. Substituting for λ_o/λ_r in Eq. 6.2, we see that the cosmological redshift is simply the ratio of the present epoch scale factor to the scale factor at the time the photon was

emitted (or absorbed),

$$\frac{a_0}{a(t)} = 1 + z. \quad (6.3)$$

The redshift element is related to the scale factor element by

$$dz = -\frac{a_0}{a^2(t)} da. \quad (6.4)$$

The observed wavelength of four rest-frame wavelengths commonly observed as emission lines in quasar spectra, $\lambda_r = \lambda 1030, \lambda 1215, \lambda 1550, \lambda 2800$ (labeled OVI/Ly β , Ly α , CIV, and MgII, respectively) are illustrated Figure 6.2(a) as a function of redshift from $z = 0$ to $z = 6$ over the observed wavelength range 500–13,500 Å. For reference, the three observational band passes are indicated. In Figure 6.2(b), the redshifting of a “typical” quasar spectrum is illustrated for the redshifts $z = 0, 1, 2$ and 3 illustrating the shifting of the emission lines. Note that the spectrum is significantly stretched.

Because of this relation between z and the direct observable quantity, λ_o/λ_r , redshift is a most convenient parameter for describing cosmological relationships. However, the factor $1 + z$ should always be viewed as the ratio $a_0/a(t)$, and be thought of as the factor by which the universe has expanded since the photon was emitted (or absorbed) at the source.

6.2 Observer Time

The expression “cosmic time” would imply there is a standard “absolute” time that applies to all observers in the universe. This is not the case. As stated above, time is not a direct observable, and as discussed in § 5.4.3, time in a relativistic spacetime geometry has meaning only in the inertial reference frame of the observer. In other words, the term “cosmic time” is a misnomer. When we speak of the age of the universe, or the time that has passed since a given event, the true meaning is strictly limited to our reference frame as observers at the present epoch.

Having made an attempt to clarify the meaning of time, we now consider the relationship between time in the observer’s frame and the actual observable quantity, redshift. Because the evolution of the expansion parameter is governed by the relativistic dynamics of Einstein’s field equations, which in turn depend on the values of the cosmological parameters, we find that time can only be mathematically deduced through the formalism of the cosmological model. Technically, time, t , is simply a mathematical parameter employed to describe the evolution of the cosmological model. Different val-

ues of the cosmological parameters yield different values and derivatives of the expansion factor as a function of t .

Redshift is a measure of the expansion parameter at a given “time” with respect to the observer at t_0 , for which $a(t_0) = a_0$ is a boundary condition of the evolution. Thus, redshift and “observer time” can be directly related through the metric and the relativistic dynamics. Substituting $a(t)/a_0 = 1 + z$ into Eq. 5.69, we have the time element for an observer at $z = 0$,

$$dt = \frac{1}{H_0} \frac{dz}{(1+z)E(z)} = t_H \frac{dz}{(1+z)E(z)}, \quad (6.5)$$

where $t_H = 1/H_0$ is known as the Hubble time,

$$t_H = 9.7776 h^{-1} \text{ Gyr}, \quad (6.6)$$

which for $h = 0.6774$, yields $t_H = 14.434$ Gyr (Planck Collaboration et al., 2016b).

For this observer at $z = 0$, the time elapsed in the universe at an arbitrary redshift is simply the integral of the time element, dt , from redshift z to “infinite” redshift

$$t(z) = t_H \int_z^\infty \frac{dz}{(1+z)E(z)}. \quad (6.7)$$

Substituting the lower limit of the integral with $z = 0$ gives, t_0 , the deduced present age of the universe,

$$t_0 = t_H \int_0^\infty \frac{dz}{(1+z)E(z)}. \quad (6.8)$$

The current best estimate for the age of the universe is (Planck Collaboration et al., 2016b)

$$t_0 = 13.799 \pm 0.021 \text{ Gyr}. \quad (6.9)$$

To determine the time elapsed since an event at redshift z , we compute the so-called “look-back time”, which is defined as $t_{\text{LB}}(z) = t_0 - t(z)$. This quantity is simply given by the integral of the time element (Eq. 6.5) over the redshift the photon traveled,

$$t_{\text{LB}}(z) = t_H \int_0^z \frac{dz}{(1+z)E(z)}. \quad (6.10)$$

For a $\Omega_\Lambda = 0$ cosmology, Eq. 6.5 reduces to

$$dt = t_H \frac{dz}{(1+z)^2 \sqrt{1+2q_0 z}}, \quad (6.11)$$

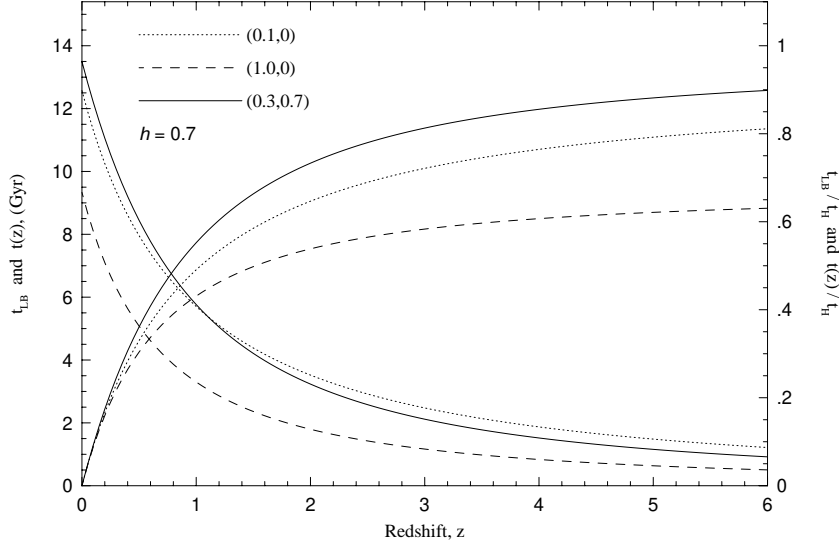


Figure 6.3 The time elapsed at redshift z , $t(z)$, and the look-back time, $t_{\text{LB}}(z)$, versus redshift. The left axis gives the times in Gigayears for $h = 0.7$. The right axis gives the times in units of the Hubble time, t_{H} . Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. [Adapted from Fig. 6 of Hogg (1999)]

and Eqs. 6.8 and 6.10 have the analytic solutions

$$\begin{aligned}
 t_0 &= t_{\text{H}} & t_{\text{LB}}(z) &= t_{\text{H}} [1 - (1+z)^{-1}] & (q_0 = 0) \\
 t_0 &= \frac{2}{3}t_{\text{H}} & t_{\text{LB}}(z) &= \frac{2}{3}t_{\text{H}} [1 - (1+z)^{-3/2}] & (q_0 = 0.5).
 \end{aligned} \tag{6.12}$$

In Figure 6.3, the quantities $t(z)$ and t_{LB} are shown for the low-density, Einstein-de Sitter, and Λ cosmologies for $h = 0.7$. At $z = 0$, the value of $t(z)$ is the age of the universe. The redshift at which the universe is half its present age occurs where the curves intersect for each cosmology.

6.2.1 Time Dilation

The proper time interval between events at the cosmologically redshifted emitter or absorber, Δt , is increased to a longer observed time interval, Δt_o , for an observer at the present time. The relationship is (see discussion in § 5.1.1)

$$\Delta t_o = (1+z) \Delta t. \tag{6.13}$$

Time dilation is a real and observable consequence of the relativistic dynamics of the expansion of space and must be taken into account when deducing the time interval between events in the inertial frame of a cosmologically redshifted object.

6.3 Expansion Dynamics

Having established the relationship between observer time and redshift, such that $t = t(z)$, we can express the time rate of change in the scale length per unit scale length (Eq. 5.53) directly as a function of the observable redshift

$$\frac{\dot{a}(t)}{a(t)} = \frac{\dot{a}[t(z)]}{a[t(z)]} = H(z) = H_0 E(z), \quad (6.14)$$

where

$$E(z) = \left[\Omega_r(1+z)^4 + \Omega_m(1+z)^3 + \Omega_k(1+z)^2 + \Omega_\Lambda \right]^{1/2}. \quad (6.15)$$

Eq. 6.14 defines the Hubble constant, $H(z) = H_0 E(z)$, as measured at redshift z . The acceleration per unit scale length (Eq. 5.53) is expressed

$$\frac{\ddot{a}(t)}{a(t)} = \frac{\ddot{a}[t(z)]}{a[t(z)]} = \dot{H}(z) = H_0^2 E'(z), \quad (6.16)$$

where

$$E'(z) = \Omega_\Lambda - \Omega_r(1+z)^4 - \frac{\Omega_m}{2}(1+z)^3. \quad (6.17)$$

Note that $\Omega_k = 0$ to high precision. At the present epoch, Ω_r equals only 0.03% of Ω_m and is negligible for the redshifts over which quasars and galaxies are observed.

For a $\Omega_\Lambda = 0$ cosmology, we have

$$E(z) = (1+z)\sqrt{1+2q_0z}, \quad (6.18)$$

and

$$E'(z) = -q_0(1+z)^3. \quad (6.19)$$

Eq. 6.18 had been commonly employed in the literature prior to acceptance of the Λ cosmology.

The cosmological expansion rate per unit scale factor, $\dot{a}(t)/a(t)$, is illustrated in Figure 6.4(a) for three cosmological models. Denoting the models by $(\Omega_m, \Omega_\Lambda)$, the three cosmologies are the low-density (0.1,0) Einstein-de

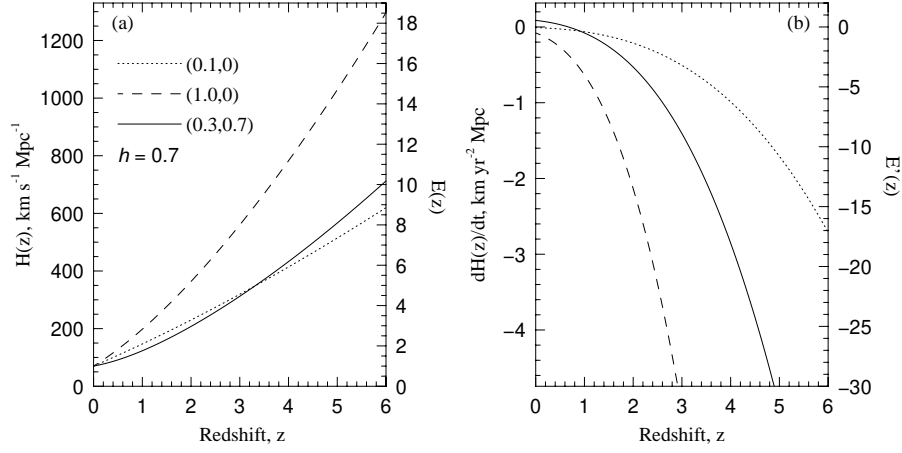


Figure 6.4 (a) The cosmological expansion rate, $\dot{a}(t)/a(t)$, given by Eq. 6.14, versus redshift. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. The left hand scale provides the expansion rate in $\text{km s}^{-1} \text{Mpc}^{-1}$ for $h = 0.7$. The right hand scale provides the expansion rate in units of H_0 , which is the function $E(z)$. (b) The cosmological acceleration rate, $\ddot{a}(t)/a(t)$, given by Eq. 6.16, versus redshift for the same cosmologies. The left hand scale provides the acceleration rate in $\text{km yr}^{-2} \text{Mpc}^{-1}$ for $h = 0.7$. The right hand scale provides the expansion rate in units of H_0^2 , which is the function $E'(z)$. Note the positive acceleration at low redshift in the Λ cosmology.

Sitter (1.0,0), and the Λ cosmology (0.3,0.7). The left axis gives the expansion rate as $H(z)$ for $h = 0.7$. The right axis gives the expansion rate normalized by H_0 , which is the function $E(z)$, given by Eq. 6.15.

The cosmological acceleration rate per unit scale factor, $\ddot{a}(t)/a(t)$, is illustrated in Figure 6.4(b) for the three same cosmological models. The left axis gives the acceleration in units $\text{km yr}^{-2} \text{Mpc}^{-1}$. Note $\ddot{a}(t)/a(t) > 0$ in the Λ cosmology at small z . The transition from deceleration to acceleration is obtained by nulling Eq. 6.16,

$$z = \left(\frac{2\Omega_\Lambda}{\Omega_m} \right)^{1/3} - 1, \quad (6.20)$$

yields $z = 0.7456$ for the cosmological parameters listed in Table. 5.1. At this redshift, the dark energy became the dominate energy density in the universe. The present epoch acceleration rate per scale factor is a mild $\ddot{a}_0/a_0 = H_0^2 = 0.17 \text{ km yr}^{-2} \text{Mpc}^{-1}$ for $h = 0.6774$.

6.4 Distances, Separations, and Angles in Cosmology

The characteristic distance of the universe is called the Hubble distance,

$$D_{\text{H}} = \frac{c}{H_0} = 2997.9 h^{-1} \text{ Mpc}, \quad (6.21)$$

which evaluates to $D_{\text{H}} = 4425.6 \text{ Mpc}$ for $h = 0.6774$. The Hubble distance is a constant of proportionality for all cosmological distance measures.

There are multiple definitions for distances in the cosmological setting and they are primarily operational in nature, i.e., they are based upon the geometrical relationship between the observer and the specific event or between multiple events. These operational definitions of distance are derived from the spatial components of the Robertson-Walker metric (see § 5.4.3),

$$ds^2 = d\tau^2 - (dl^2 + l_{\text{T}}^2 d\psi^2)$$

$$dl = \frac{a(t)}{a_0} dD_c = \frac{a(t)}{a_0} \frac{D_{\text{H}}}{\sqrt{|\Omega_k|}} d\chi \quad (\text{radial}) \quad (6.22)$$

$$l_{\text{T}} d\psi = \frac{a(t)}{a_0} D_{\text{T}} d\psi = \frac{a(t)}{a_0} \frac{D_{\text{H}}}{\sqrt{|\Omega_k|}} f_k(\chi) d\psi \quad (\text{transverse}),$$

where dl is the proper radial element, dD_c is the radial co-moving element, and the transverse co-moving, $l_{\text{T}} d\psi$ is the proper transverse element, and $D_{\text{T}} d\psi$ is the transverse co-moving element.

The radial components apply for observations requiring the distance to be determined to an event along a single line of sight. The transverse components apply for observations requiring the transverse separations to be determined between two equally distant events separated by an angle $\Theta = \int d\psi$. We provide a brief synopsis of the operational definitions of the various distance measures:

- **Proper distance:** The radial proper distance, l , is the line of sight separation between an observer and an object or between two objects that share the same time coordinate. For radial distances, $d\psi = 0$ and for proper distances $d\tau = 0$, yielding $l = \int dl = \int [a(t)/a_0] dD_c$. Note that radial proper distance is smaller at higher redshifts (in the observer's past) by the factor $a(t)/a_0 = (1+z)^{-1}$.
- **Co-moving distance:** The radial co-moving distance, $D_c = \int dD_c$, is also the line of sight separation between objects, but it is a fixed quantity (a ruler that “stretches” with the expanding geometry). It is equivalent to the radial proper distance that two objects would have at the present

epoch (i.e., $t = t_0$, $z = 0$), when $a(t_0)/a_0 = 1$. We further discuss the interpretation of co-moving distance below.

- **Transverse co-moving distance:** The transverse co-moving distance, D_T , is a misnomer; it is a curvature corrected radial distance to an object. It is employed to compute the co-moving transverse separation, $S_c = D_T \Theta$, between two objects or events observed to be separated by an angle Θ (i.e., observed along lines of sight that converge at the observer with angle $\Theta = \int d\psi$).
- **Proper separation:** The proper separation, $S = [a(t)/a_0]S_c = l_T \Theta$, is the proper transverse distance between two objects separated by an observed angle, $\Theta = \int d\psi$. Note that proper separation is smaller at higher redshifts (in the observer's past) by the factor $a(t)/a_0 = (1+z)^{-1}$.
- **Angular diameter distance:** The angular diameter “distance” is not a physical distance. Denoted D_A , it is defined as the ratio S/Θ , the proper separation divided by the observed angular separation, Θ . As such, D_A supplies the “rate” per unit angular separation at which the proper separation changes as a function of redshift. It is a utility definition that facilitates computing proper separations, the transverse physical size of resolved objects, line of sight impact parameters, and line of sight proper transverse separations in gravitational lenses as a function of redshift.
- **Luminosity distances:** The luminosity “distance”, D_L , is also not a physical distance. It is a “distance” that corrects for cosmological expansion effects on the observed flux of an object with a given luminosity. Because measured flux is a quantity integrated over the projected physical extent of the source, luminosity distance is proportional to the angular diameter distance. The proportionality derives from both space expansion and time dilation.
- **Absorption distance:** The absorption “distance”, $X(z)$, is also not a physical distance. It is a modified redshift path length, Δz , over which the probability of intersecting objects along a line of sight is made constant at different cosmic epochs (for non-evolving objects, i.e., those with unchanging physical cross sections and co-moving cosmic density). The absorption distance accounts for both the expansion of space and the rate of the expansion.

6.4.1 Co-Moving and Proper Radial Distances

To obtain the co-moving and proper radial distances from an observer at the present epoch to a distant-past object or event, we set $d\psi = 0$ in the spatial component of the Robertson-Walker metric and integrate the co-moving

radial distance element, dD_c , as demonstrated in § 5.4.3. The co-moving distance element is proportional to the element of the development angle, $d\chi$, as given by Eqs. 5.71 and 5.73,

$$dD_c = \frac{D_H}{\sqrt{|\Omega_k|}} d\chi = D_H \frac{a_0 da}{a^2(t)E(t)}. \quad (6.23)$$

Integrating,

$$D_c = \frac{D_H}{\sqrt{|\Omega_k|}} \chi = D_H \int_{a_0}^{a(t)} \frac{a_0 da}{a^2(t)E(t)}, \quad (6.24)$$

and applying Eq. 6.4 and invoking $t = t(z)$, we have

$$\chi(z) = \frac{D_H}{\sqrt{|\Omega_k|}} \int_0^z \frac{dz}{E(z)}, \quad (6.25)$$

which provides the co-moving radial distance out to redshift z

$$D_c(z) = D_H \int_0^z \frac{dz}{E(z)}. \quad (6.26)$$

From Eq. 6.26, can write the development angle in the compact notation

$$\chi(z) = \sqrt{|\Omega_k|} \frac{D_c(z)}{D_H}. \quad (6.27)$$

For a $\Omega_\Lambda = 0$ cosmology, Eq. 6.26 can be integrated to yield the analytical expression

$$D_c(z) = D_H \left[\frac{q_0 z + (q_0 - 1)(\sqrt{1 + 2q_0 z} - 1)}{q_0^2(1 + z)} \right] \quad (\Omega_\Lambda = 0). \quad (6.28)$$

From the definition of the proper distance elements given in Eq. 6.22, and invoking Eq. 6.4 and $t = t(z)$, we have the proper distance from the observer to a distant-past event or object

$$l(z) = \frac{D_H}{1 + z} \int_0^z \frac{dz}{E(z)}. \quad (6.29)$$

Note that the ratio $a(t)/a_0 = (1 + z)^{-1}$ has remained outside the integral over redshift. The expansion factor appearing in the metric is evaluated only at the upper limit of the integration (see the following discussion contrasting proper and co-moving distances in cosmology). In compact form we have

$$l(z) = \frac{D_c(z)}{1 + z}. \quad (6.30)$$

A schematic comparing the radial proper and co-moving distances is provided in Figure 6.5. Five redshifts are shown, with time and scale factor

increasing downward. Co-moving distance intervals, ΔD_c (ruled), are locked into the expansion; the number of co-moving distance intervals separating two objects (shaded circles) is unchanged as space expands. The proper interval lengths expand in exact step with $a(t)/a_0$ as time increases. Thus, the co-moving distance between the objects is constant (an unchanging, fixed number), but the proper distance increases with time.

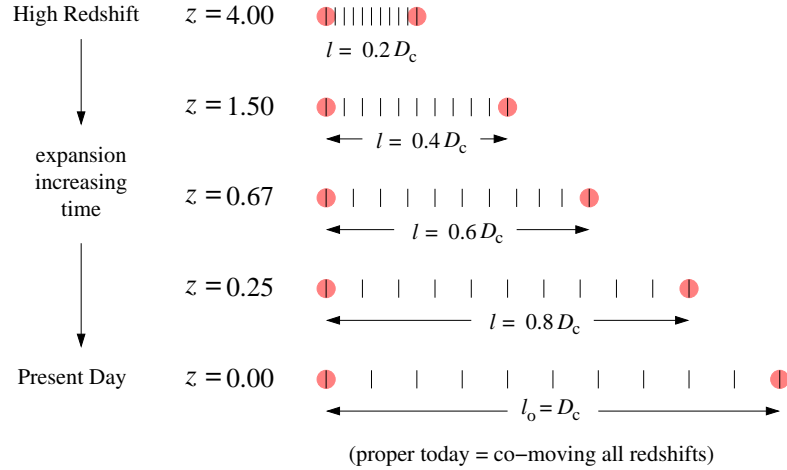


Figure 6.5 A schematic illustrating the relationship between proper distance, l , and co-moving distance, D_c . Shown are five snapshots in intervals of $a(t)/a_0 = 0.2$ of the increasing proper distance between two objects (shaded circles) at the corresponding redshifts, z . Co-moving distance intervals are marked with vertical ticks; there are 10 co-moving intervals between the two objects and this number of intervals remains constant because co-moving intervals move with the expansion. The proper distance at the present cosmic epoch, l_o , equals the co-moving distance; therefore, a proper distance measured between two $z = 0$ objects is their co-moving distance at all redshifts. See text for further description of this figure.

To make the example illustrated in Figure 6.5 concrete, there are 10 co-moving distance intervals separating the objects at all redshifts. At the present epoch, the number of proper intervals equals the number of co-moving intervals. From $l = (a(t)/a_0)D_c = D_c/(1+z)$, we see that at $z = 0.25$ a proper interval was equal to 0.8 of a co-moving interval and at $z = 1.5$ was equal to 0.4 of a co-moving interval.

From an operational point of view, the co-moving distance normalizes out the cosmological expansion in order to place distance intervals on a common scale for all cosmic times. This allows direct comparison of distance intervals at different epochs by standardizing them to the present epoch. In co-moving coordinates, space is static, time does not dilate, objects locked into the

cosmological expansion i.e., carried by the “Hubble flow”¹, are at rest with respect to one another.

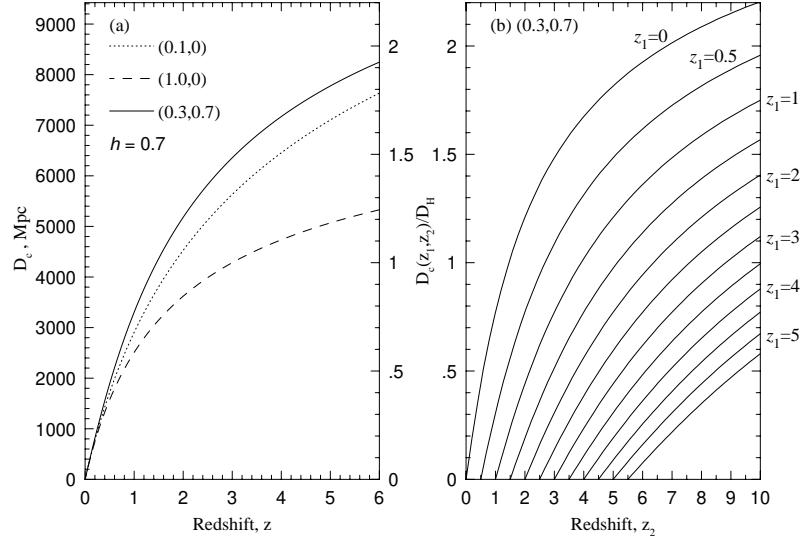


Figure 6.6 (a) The total radial co-moving distance, given by Eq. 6.26. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density $(0.1, 0)$, dotted; (ii) Einstein-de Sitter $(1.0, 0)$, dashed; and the Λ cosmology $(0.3, 0.7)$, solid. The left hand scale provides the distance in Megaparsecs for $h = 0.7$. The right hand scale provides the distance in units of D_H (Figure adapted from Hogg, 1999). (b) The co-moving radial separation, given by Eq. 6.33 in units of D_H as a function of z_2 for the Λ cosmology. Each curve is for a different z_1 .

The total radial co-moving distance, $D_c(z)$, which is illustrated in Figure 6.6(a) for the low-density, Einstein-de Sitter, and Λ cosmologies, is a fundamental distance measure. The left axis of Figure 6.6(a) gives $D_c(z)$ in Megaparsecs for $h = 0.7$. The right axis gives $D_c(z)/D_H$.

6.4.2 Radial Separations

Consider an object at $z = z_1$ and a second higher redshift object at $z = z_2$. The co-moving distance between the objects must properly account for the scale factor expansion from the epoch at z_2 to the epoch at z_1 . Let the expansion factor at z_2 be denoted $a_2 = a(t_2)$ and the expansion factor at z_1 be denoted $a_1 = a(t_1)$. Then the expansion rate $E(t)$ (Eq. 5.52), which

¹ The term “Hubble flow” is often used to describe the motion of objects locked into the cosmological expansion; for example, “... objects carried by the Hubble flow.” We will use this term throughout this work.

appears in the definition of the radial co-moving distance, is now

$$E(t_1, t_2) = \left[\Omega_m \left(\frac{a_1}{a_2} \right)^3 + \Omega_k \left(\frac{a_1}{a_2} \right)^2 + \Omega_\Lambda \right]^{1/2}, \quad (6.31)$$

where

$$\frac{a_1}{a_2} = \frac{(a_0/a_2)}{(a_0/a_1)} = \frac{1+z_2}{1+z_1} = 1+z_{12}, \quad (6.32)$$

is the reduced redshift. Thus, the radial co-moving separation between two high redshift objects is mathematically equivalent to obtaining the total radial co-moving distance from $z = 0$ to $z = z_{12}$, giving

$$D_c(z_1, z_2) = D_H \int_0^{z_{12}} \frac{dz}{E(z)}, \quad (6.33)$$

where $E(z)$ is given by Eq. 6.15 as if the observer were at $z = 0$. Eq. 6.33 does *not* yield a co-moving distance equivalent to subtracting the total radial co-moving distance of the object from $z = 0$ to $z = z_1$ from the total radial co-moving distance of the object from $z = 0$ to $z = z_2$. The radial proper separation is then $l(z_1, z_2) = D_c(z_1, z_2)/(1+z_{12})$.

The radial co-moving separation, in units of D_H , are plotted for various z_1 in Figure 6.6(b) as a function of z_2 for the Λ cosmology. For example, the co-moving distance between an object at $z_1 = 2.0$ and an object at $z_2 = 3.0$ is $D_c(2, 3) \simeq 0.3D_H$ and the proper distance is $l(2, 3) \simeq 0.3/(1+0.33) = 0.23D_H$.

6.4.3 Transverse Separations

A transverse separation is the distance between two observed events or objects at the same redshift but that are toward different lines of sight with angular separation Θ . The observed angle is dependent upon the geometry and curvature of the universe, since these properties govern the angle at which the two sightlines converge at the observer.

Employing the transverse component of the metric, the co-moving transverse separation element is

$$ds_c = D_T d\psi = \frac{D_H}{\sqrt{\Omega_k}} f_k(\chi) d\psi, \quad (6.34)$$

where (Eq. 5.27),

$$f_k(\chi) = \begin{cases} \sin \chi & \Omega_k < 0 \text{ (closed)} \\ \chi & \Omega_k = 0 \text{ (flat)} \\ \sinh \chi & \Omega_k > 0 \text{ (open)}. \end{cases} \quad (6.35)$$

where the development angle, χ , is (Eq. 6.27),

$$\chi(z) = \sqrt{\Omega_k} \frac{D_c(z)}{D_H}, \quad (6.36)$$

and where $D_c(z)$ is the radial co-moving distance given by Eq. 6.26. From examination of Eq. 6.34 and inserting Eq. 6.35, we have

$$D_T(z) = \begin{cases} \frac{D_H}{\sqrt{|\Omega_k|}} \sin\left(\sqrt{|\Omega_k|} \frac{D_c(z)}{D_H}\right) & \Omega_k < 0 \text{ (closed)} \\ D_c(z) & \Omega_k = 0 \text{ (flat)} \\ \frac{D_H}{\sqrt{\Omega_k}} \sinh\left(\sqrt{\Omega_k} \frac{D_c(z)}{D_H}\right) & \Omega_k > 0 \text{ (open)}. \end{cases} \quad (6.37)$$

Since $\Omega_k = 0$ to good precision, $D_T(z) = D_c(z)$ applies.

Integrating Eq. 6.34, we obtain the transverse co-moving separation

$$S_c(z) = D_T(z) \iint_{\phi, \theta} d\psi = D_T(z) \Theta. \quad (6.38)$$

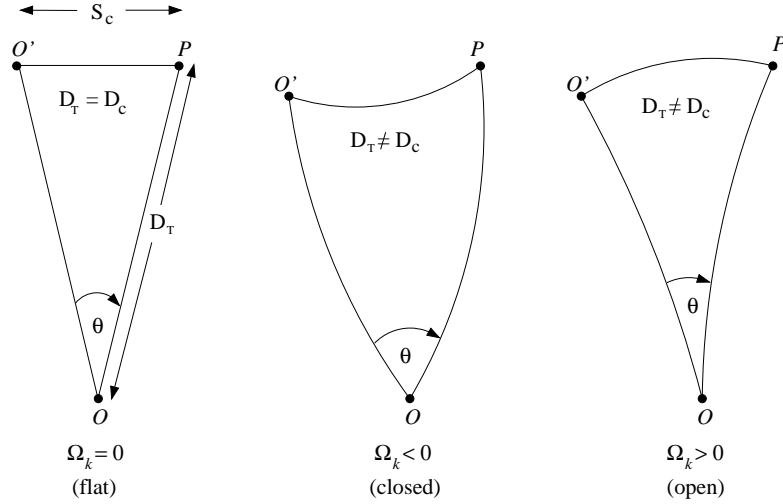


Figure 6.7 The transverse co-moving distance, D_T , is used to compute the co-moving separation, S_c , between two objects, O' and P , at the same redshifts as observed at angular separation, Θ , by observer O . The observed angle Θ will depend upon the curvature of the universe. The functional form of D_T accounts for the curvature. [Adapted from Fig. 12.1 of Peebles (1993)]

The geometric configuration relating the co-moving separation, $S_c(z)$, the

transverse co-moving distance, $D_T(z)$, and the observed angle separating the two lines of sight is illustrated in Figure 6.7. The observed angular separation Θ for two events or objects with separation $S_c(z)$ is dependent upon the curvature of the universe, which dictates the angle at which the light paths from the objects converge at the observer. As illustrated schematically in Figure 6.7, the functional form of $D_T(z)$ accounts for this geometric behavior.

Consider an object P at redshift z as observed by observer O at $z = 0$. Consider a second object O' at the same redshift as object P separated by angle Θ as observed by O . The co-moving distance to P as observed by observer O' , which is what we desire to measure, is the co-moving separation, S_c , as observed by O . This is the geometric configuration that gives the transverse co-moving distance its meaning; in a non-zero curvature geometry, $D_T(z)$ is not the radial co-moving distance—it is “curvature corrected” path length of the photon traveling from an object that has angle of incidence Θ with respect to a photon traveling from another object.

We define the transverse proper separation by the relation

$$S(z) = \frac{a(t)}{a_0} S_c(z) = \frac{D_T(z)}{1+z} \Theta, \quad (6.39)$$

which is interpreted as the proper distance between the two objects at redshift z , though they are measured to have angular separation Θ by an observer at the present epoch. The transverse proper distance can alternatively be interpreted as the radial proper distance between the objects from the point of view of an observer located on one of the objects.

6.4.4 Angular Diameter “Distance”

Defining a quantity known as the angular diameter distance,

$$D_A(z) \equiv \frac{S(z)}{\Theta} = \frac{D_T(z)}{1+z}, \quad (6.40)$$

we can rewrite the transverse proper separation

$$S(z) = D_A(z) \Theta. \quad (6.41)$$

As defined in Eq. 6.40, $D_A(z)$ is the ratio of the transverse proper separation per unit angular separation. This is the common operational definition of the angular diameter distance; it is the factor or proportionality between the observed angle Θ and $S(z)$. However, it is also equivalent to $D_T(z)/(1+z)$, which can be viewed as a transverse proper distance, l_T , as originally discusses in § 5.4.3, and appearing in the transverse component in

Eq. 6.22. Though l_T as defined herein has no historical convention, we see that $l_T(z) = D_A(z)$.

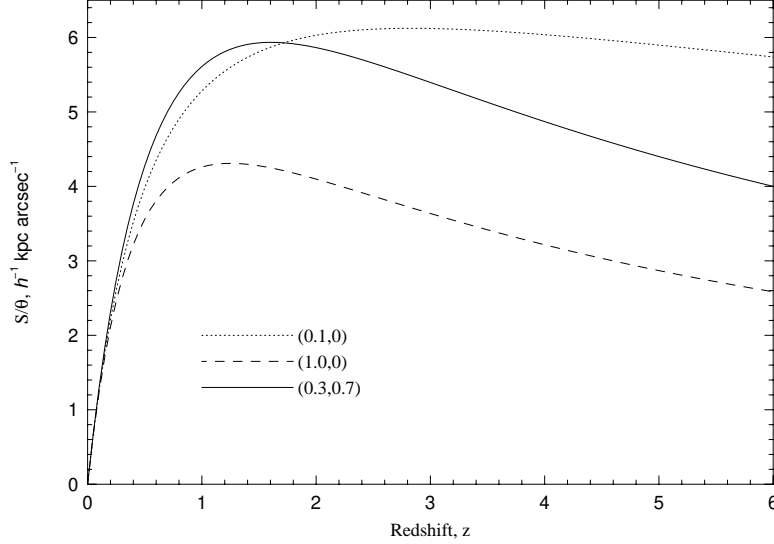


Figure 6.8 The ratio S/Θ between two objects at the same redshift given in the units $h^{-1} \text{ kpc arcsec}^{-1}$, as given by Eq. 6.42. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. [Adapted from Fig. 2 of Hogg (1999)]

Expressing the transverse proper separation in terms of angular separations measured in arc seconds yields,

$$S(z) = 14.54 \left(\frac{\Theta''}{1''} \right) \left(\frac{D_A(z)}{D_H} \right) h^{-1} \text{ kpc}. \quad (6.42)$$

For $\Omega_k = 0$, we have $D_A(z)/D_H = (1+z)^{-1}(D_c(z)/D_H)$, where $D_c(z)/D_H$ is given by Eq. 6.26. The ratio S/Θ in units $h^{-1} \text{ kpc arcsec}^{-1}$, is shown in Figure 6.8 for the low-density, Einstein-de Sitter, and Λ cosmologies. A rule of thumb is that two objects at $z \simeq 1$ separated by $1''$ have a physical separation of $\simeq 5 h^{-1} \text{ kpc}$. Note that the proper separation maximizes for $z \sim 1$, and then decreases for higher redshifts.

Examples of the application of Eq. 6.41 or Eq. 6.42 include the field of quasar absorption lines in which it is desired to compute the impact parameter between a luminous object, such as a galaxy, and an absorption system at the same redshift. An additional example is to compute the line of sight separation as a function of redshift for galaxy or quasar pairs (not gravitational lenses) having close angular separation Θ on the sky. It is not

required that the sources have similar redshifts (in this scenario, the proper separation of the two sight lines can be determined for all redshifts up to that of the lower redshift source, at which point the proper separation of the sight line of the higher redshift source is the impact parameter to the lower redshift source). Similarly, the physical transverse size of an object with measured angular extent Θ is can be determined from Eq. 6.42.

6.4.5 Angular Separations

The transverse co-moving distance, denote D_T in the metric, and the angular diameter “distance”, D_A , deserve a few extra words of discussion. For two distant-past events that share an inertial frame (or, essentially reside at the same redshift) but are separated by an observed angle $\Theta = \int d\psi$ from the point of view of the observer, their proper separation in the fame of the events is

$$S(z) = D_A(z) \iint_{\phi, \theta} d\psi = D_A(z) \int_{\theta_1}^{\theta_2} \int_{\phi_1}^{\phi_2} (d\theta^2 + \sin^2 \theta d\phi^2)^{1/2}. \quad (6.43)$$

This relation is equivalent to $S(z) = l_T(z) \Theta$ or $S = [D_T(z)/(1+z)]\Theta$, expressions that are analogous to those utilized to compute the arc length, s , between two points on the circumference of a circle of radius r that subtend an angle θ , i.e., $s = r\theta$. Thus, the so-called transverse distances of the metric are more intuitively viewed as being analogous to the radial distances to the two events, but they account for the spatial curvature, which, as illustrated in Figure 6.7, governs the angle Θ at which the two lines of sight converge; the accounting of curvature is manifest in the function $f_k(\chi)$ (Eq. 6.35) appearing in D_T (Eq. 6.37).

In the ecliptic coordinate systems, we have right ascension $\alpha = \phi$ and declination $\delta = \pi/2 - \theta$. Thus, $d\phi = d\alpha$, $d\theta = -d\delta$, and $\sin \theta = \cos \delta$, which yields $d\psi = (d\delta^2 + \cos^2 \delta d\alpha^2)^{1/2}$. Under the approximation that the separation in $\Delta\alpha$ and $\Delta\delta$ between the two events or objects are small, integration of $d\psi$ over the limits $\alpha_1 \rightarrow \alpha_2$ and $\delta_1 \rightarrow \delta_2$ yields

$$\Theta \simeq [(\Delta\delta)^2 + (\Delta\alpha)^2 \cos^2 \bar{\delta}]^{1/2} \quad (6.44)$$

where $\bar{\delta} = (\delta_1 + \delta_2)/2$ and $\Delta\alpha = \alpha_2 - \alpha_1$. Proper evaluation of the integral yields

$$\cos \Theta = \sin \delta_1 \sin \delta_2 + \cos \delta_1 \cos \delta_2 \cos \Delta\alpha. \quad (6.45)$$

6.4.6 Proper Separations in Lenses

In a gravitational lens, multiple images of the light source will be seen. Thus, there will be multiple lines of sight that can be observed simultaneously. For the majority of the time that the light travels, the paths follow the curvature of universe. However, in the proximity of the lens, the paths are deflected by the local curvature induced by the lensing mass. In cosmological applications, the path length of the deflection is negligible. Thus, to a good approximation, the deflection can be treated as instantaneous; that is, the light paths can be treated as following the curvature of the universe from source to lens and then, post deflection, from lens to observer.

A schematic of the geometry of a lens system is shown in Eq. 6.9, where the observer O is at $z = 0$, the lens l is at $z = z_l$, and the source Q is at $z = z_s$. The observer measures the two images, A and B , to have angular separation Θ . Consider some arbitrary redshift, z , such that $0 \leq z \leq z_s$, i.e., between the source and the observer. We wish to compute the proper separation, $S(z)$, of the lines of sight for all redshifts $0 \leq z \leq z_s$.

The proper separation of the light paths as a function of redshift on the observer side of the lens simply is

$$S(z) = D_{oz}\Theta \quad (0 \leq z \leq z_l), \quad (6.46)$$

where D_{oz} is the angular diameter distance from the observer to the redshift under consideration (this result is identical to Eq. 6.41). In § 6.4.4, it was shown that the angular diameter distance from the observer to redshift z is computed from Eq. 6.40, i.e.,

$$D_{oz} = D_A(z) = \frac{D_T(z)}{1+z}, \quad (6.47)$$

where $D_T(z)$ is the transverse co-moving distance to redshift z (Eq. 6.37), depending upon the sign of Ω_k .

To obtain the proper separation of the light paths on the far side of the lens, $z_l \leq z \leq z_s$, consider the light paths from the view point of the source, Q . The proper separation of the light paths is

$$S(z) = D_{sz}\Theta', \quad (6.48)$$

where Θ' is the separation angle of the light paths from Q , and D_{sz} is the angular diameter distance *from* the source *to* arbitrary redshift z (yes, oriented from high redshift to low redshift). Since we cannot measure Θ' , we need to obtain the geometric relationship between Θ' and Θ . From Q , the angular diameter distance to the lens is D_{sl} (notice the direction of the

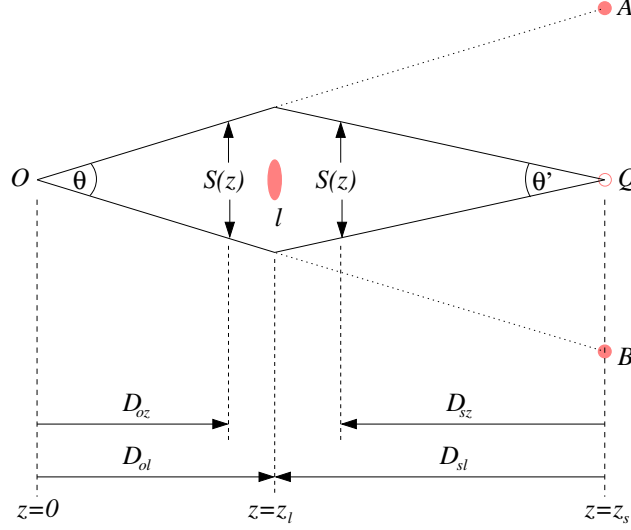


Figure 6.9 A schematic of the gravitational lens geometry and the quantities involved in the computation of $S(z)$, the proper separation of the two lines of sight, A and B , at arbitrary redshift z . From the point of view of the observer, O , the lines of sight have angular separation Θ and the source S , is seen as the images A and B . From the point of view of the source, Q , the lines of sight have angular separation Θ' . The various distances that are employed to compute the proper separation of the light paths at redshift z are: (1) for $z \leq z_l$, the observer to redshift angular diameter distance, D_{oz} , and the observer to lens angular diameter distance, D_{ol} , (2) for $z \geq z_l$, the source to redshift angular diameter distance, D_{sz} , and the source to lens angular diameter distance, D_{sl} . Note that the direction of the arrows for these respective distances are important (see text). [Adapted from Fig. 3 of Smette et al. (1992)]

arrow denoting D_{sl} in Figure 6.9) and the proper separation of the light paths at the lens is

$$S(z_l) = \Theta' D_{sl}. \quad (6.49)$$

From O , the proper separation of the light paths at the lens is

$$S(z_l) = D_{ol} \Theta, \quad (6.50)$$

where D_{ol} is the angular diameter distance to the lens from O . Equating Eqs. 6.50 and 6.49, yields the geometric relationship,

$$\Theta' = \frac{D_{ol}}{D_{sl}} \Theta. \quad (6.51)$$

Substituting Eq. 6.51 into Eq. 6.48, we have the proper separation of the

light paths on the far side of the lens from the observer,

$$S(z) = \frac{D_{sz}}{D_{sl}} D_{ol} \Theta \quad (z_l \leq z \leq z_s), \quad (6.52)$$

which is the known as the lens equation. Together, Eqs. 6.46 and 6.52 describe the proper separation of the lines of sight from observer to source. For $z \simeq z_l$, there is some uncertainty due to the treatment of instantaneous deflection.

In the lens equation, Eq. 6.52, angular diameter distances appear that are measured from z_s to z_l and from z_s to z (i.e., from higher to lower redshift). In general, the angular diameter distance to redshift z_2 as observed from redshift z_1 , where $z_2 > z_1$, is

$$D_{12} = \frac{D_2 Q_1 - D_1 Q_2}{1 + z_2}, \quad (6.53)$$

where

$$Q_1 = \left(1 + \Omega_k \frac{D_1^2}{D_H^2}\right)^{1/2} \quad Q_2 = \left(1 + \Omega_k \frac{D_2^2}{D_H^2}\right)^{1/2}, \quad (6.54)$$

are the curvature terms, and where $D_1 = D_T(z_1)$ and $D_2 = D_T(z_2)$ are the transverse co-moving distances from the observer to z_1 and z_2 , respectively. In the lens equation, two of the terms, D_{sl} and D_{sz} , are the angular diameter distance as observed *from* the source, Q , which is the higher redshift perspective. The inverse relationship to Eq. 6.53, in which an “observer” at a higher redshift, z_2 , requires the angular diameter distance to a lower redshift, z_1 , is

$$D_{21} = \left(\frac{1 + z_2}{1 + z_1}\right) D_{12}. \quad (6.55)$$

It is important to apply the inverse relationship for the terms D_{sl} and D_{sz} in the lens equation, or the separation will be incorrect.

For a $\Omega_\Lambda = 0$ cosmology, D_{12} simplifies to

$$D_{12} = 2D_H \left[\frac{(1 - 2q_0)(Q_1 - Q_2) + (Q_1 Q_2^2 - Q_1^2 Q_2)}{(2q_0)^2 (1 + z_1)(1 + z_2)^2} \right], \quad (6.56)$$

where the curvature terms are $Q_1 = \sqrt{1 + 2q_0 z_1}$ and $Q_2 = \sqrt{1 + 2q_0 z_2}$, and where Eq. 6.55 applies when writing D_{21} .

Applying Eqs. 6.53, 6.54, and 6.55, the ingredients for the lens equation

(Eq. 6.52) are

$$\left. \begin{aligned} D_{ol} &= \frac{D_T(z_l)}{1+z_l} \\ D_{sl} &= \frac{Q(z_l)D_T(z_s) - Q(z_s)D_T(z_l)}{1+z_l} \\ D_{sz} &= \frac{Q(z)D_T(z_s) - Q(z_s)D_T(z)}{1+z} \end{aligned} \right\} (z_l \leq z \leq z_s), \quad (6.57)$$

where the inverse relationship to Eq. 6.53 applies for D_{sl} and D_{sz} , and where

$$Q(z') = \left(1 + \Omega_k \frac{D_T^2(z')}{D_H^2} \right)^{1/2} \quad (z' = z, z_l, z_s). \quad (6.58)$$

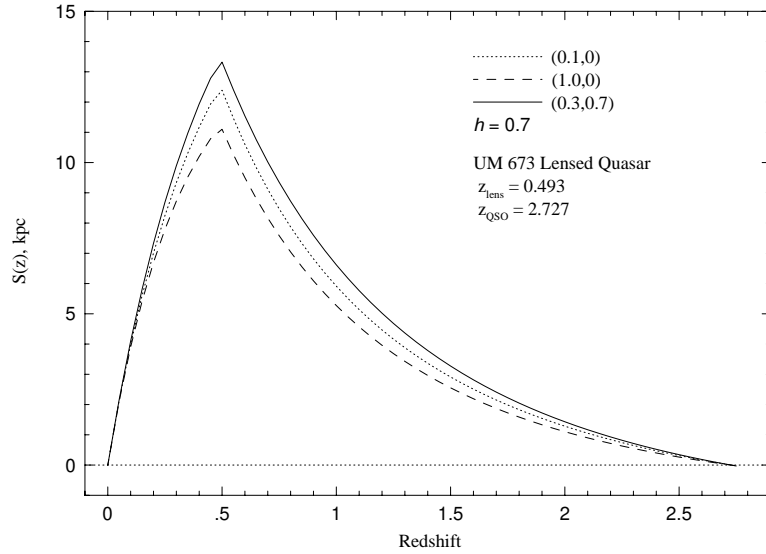


Figure 6.10 The co-moving line of sight separation, $S(z)$, as a function of redshift for the UM 673 lensed quasar for $h = 0.7$. The quasar is at $z_s = 2.727$, and the lens is at $z_l = 0.493$. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. [Adapted from Fig. 4 of Smette et al. (1992)]

The full expressions for the proper separation of the lensed sight lines are

then

$$\begin{aligned}
 S(z) &= \Theta \frac{D_T(z)}{1+z} & (0 \leq z \leq z_l) \\
 S(z) &= \Theta \frac{D_T(z_l)}{1+z} \left[\frac{Q(z)D_T(z_s) - Q(z_s)D_T(z)}{Q(z_l)D_T(z_s) - Q(z_s)D_T(z_l)} \right] & (z_l \leq z \leq z_s).
 \end{aligned} \tag{6.59}$$

Since $\Omega_k = 0$ to good precision, the angular diameter distances (Eq. 6.57) can be highly simplified. The curvature terms, Eq. 6.58, simplify to $Q(z) = 1$. Also, the transverse co-moving distance, $D_T(z)$, reduces to the radial co-moving distance, $D_c(z)$, as given by Eq. 6.37.

In Figure 6.10, the proper separation, $S(z)$, of the light paths for images *A* and *B* of the lensed quasar UM 673 are plotted for the low-density, Einstein-de Sitter, and Λ cosmologies for $h = 0.7$. The lens parameters are $z_s = 2.727$, $z_l = 0.493$, and $\Theta = 2.22''$ (Smette et al., 1992). For both the Einstein-de Sitter and Λ cosmologies, $\Omega_k = 0$, so that $D_T = D_c$ and the curvature terms are unity. However, for the low-density case, $\Omega_k = 1 - \Omega_m = 0.9$, so that D_T must be computed from Eq. 6.37 with $\Omega_k > 0$ and the curvature terms must be included in the computation of $S(z)$.

6.4.7 Source Beam Cross Section

A consequence of Eq. 6.41 is that the proper cross section of a light beam, σ_b , emitted from a source of cross section σ_s at redshift z_s , is not constant with redshift. The functional form of $\sigma_b(z)$ can be derived directly from Eq. 6.41 by realizing that the observed angle subtended by the diameter of the source is the angle subtended by the diameter of the beam for all $z \leq z_s$. This is equivalent to stating that the solid angle subtended by the cross section of a source having radius R_s ,

$$\Omega_s = \frac{\pi R_s^2}{D_A^2(z_s)} = \frac{\sigma_s}{D_A^2(z_s)}, \tag{6.60}$$

is equal to the solid angle subtended by the beam for all $z \leq z_s$,

$$\Omega_b(z) = \Omega_s, \tag{6.61}$$

where at redshift z

$$\Omega_b(z) = \frac{\sigma_b(z)}{D_A^2(z)}. \tag{6.62}$$

Equating the solid angles and solving for $\sigma_b(z)$, we have

$$\sigma_b(z) = \sigma_s \left[\frac{D_A(z)}{D_A(z_s)} \right]^2. \quad (6.63)$$

The ratio $\sigma_b(z)/\sigma_s$ is plotted in Figures 6.11(a), 6.11(b), and 6.11(c) for the low-density, Einstein-de Sitter, and Λ cosmology, respectively, for a source at $z_s = 1, 2, 3, 4, 5$, and 6. Solid curves represent even numbered z_s and dotted curves represent odd numbered z_s .

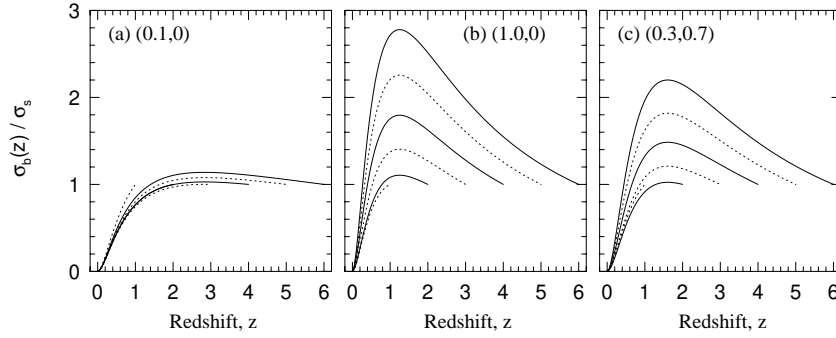


Figure 6.11 The redshift dependence of the cross section of a beam, $\sigma_b(z)$ for sources with cross section σ_s at redshift z_s , given by Eq. 6.63. Six source redshifts are illustrated, $z_s = 1, 2, 3, 4, 5$, and 6. Solid curves represent even z_s and dotted curves represent odd numbered z_s . (a) the low-density cosmology (0.1,0). (b) the Einstein-de Sitter (1.0,0) cosmology. (c) the Λ cosmology (0.3,0.7).

For quasars, the majority of the spectrum is emitted from the central compact region. However, quasars also have significantly extended regions that give rise to emission lines over limited wavelength ranges. Therefore, the cross section of the light beam from a quasar is wavelength dependent. The functional form of $\sigma_b(z)$ has several applications, one of which includes partial covering of the source. If the physical size of the source is known, then the beam size can be determined at the redshift of an absorbing structure. If, for example, an absorption feature shows characteristics of not covering the source entirely, then constraints can be placed on the covering fraction (see § 7.4.5), geometry, and or size of the absorbing material.

6.4.8 Luminosity Distance

Having measured the flux of a cosmological source, computing the luminosity requires an application of the inverse square law of light in the context

of a relativistic dynamical geometry. This requires a unique measure of “distance” called the luminosity distance, D_L .

Luminosity is energy emitted per unit time, i.e., power. Consider the bolometric luminosity of a cosmological source, L_s , which is the emitted energy per unit time integrated over all wavelengths over all solid angle (having units erg sec^{-1}). Both energy and time are affected by the expansion of space, so that the observed flux measured by an observer at $z = 0$ does not correctly correspond to the source luminosity without correction for these effects.

If the energy of a given photon at the source is $E_s(\lambda_r) = hc/\lambda_r$, where λ_r is the rest-frame wavelength at the source, then the energy of the photon received by an observer at $z = 0$ is reduced such that

$$E_o(\lambda_o) = \frac{hc}{\lambda_o} = \frac{1}{1+z} \frac{hc}{\lambda_r} = \frac{E_s(\lambda_r)}{1+z}, \quad (6.64)$$

where z is the redshift of the source.

In addition, the time interval between the arrival of photons at the observer is dilated relative to the time interval between the emission of the photons. From Eq. 6.13,

$$\Delta t_o = (1+z) \Delta t_s, \quad (6.65)$$

where Δt_o and Δt_s are the time intervals at the observer and source, respectively. Thus the photon emission rate at the source is higher than the reception rate at the observer. Since the bolometric luminosity is the integral over all wavelengths of the energy per unit time, then the observed bolometric luminosity from an object at redshift z is

$$L_o(z) = \int \frac{E_o(\lambda)}{\Delta t_o} d\lambda = \frac{1}{(1+z)^2} \int \frac{E_s(\lambda)}{\Delta t_s} d\lambda = \frac{L_s}{(1+z)^2}. \quad (6.66)$$

The observed bolometric flux of the redshifted source, $F(z)$, is the measured quantity after the luminosity from a source is spread over a sphere of co-moving area $4\pi D_T^2(z)$, where $D_T(z)$ is the transverse co-moving distance given by Eq. 6.37. Thus, the observed flux for an object at redshift z is

$$F(z) = \frac{L_o(z)}{4\pi D_T^2(z)} = \frac{L_s}{4\pi(1+z)^2 D_T^2(z)} = \frac{L_s}{4\pi D_L^2(z)}, \quad (6.67)$$

which provides the definition of the luminosity distance

$$D_L(z) = (1+z) D_T(z) = (1+z)^2 D_A(z). \quad (6.68)$$

where $D_A(z)$ is the angular diameter distance given by Eq. 6.40.

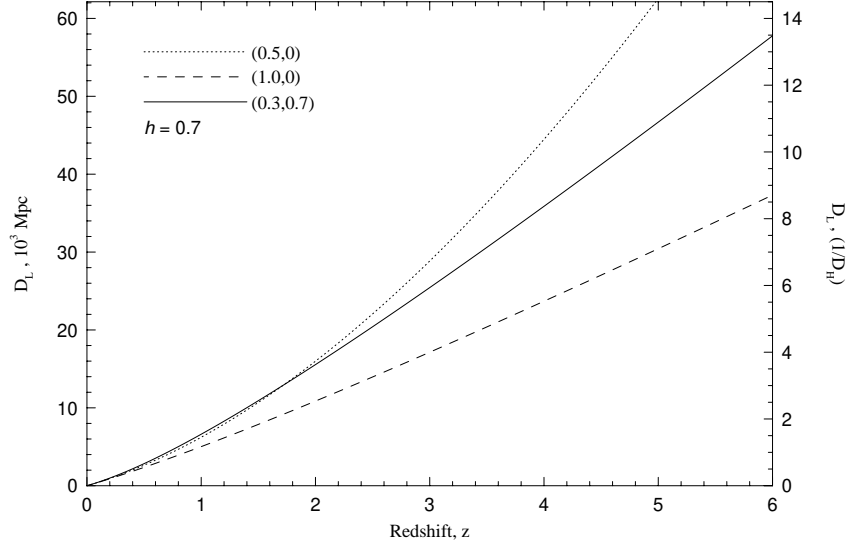


Figure 6.12 The luminosity distance, given by Eq. 6.68. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. The left hand scale provides the distance in Gigaparsecs for $h = 0.7$. The right hand scale provides the distance in units of D_H . [Adapted from Fig. 3 of Hogg (1999)]

The luminosity distance is shown in Figure 6.12 for the low-density, Einstein-de Sitter, and Λ cosmologies. The left axis gives the distance in Gigaparsecs for $h = 0.7$. The right axis gives the distance in units of the Hubble distance, which is ~ 4.3 Gpc for $h = 0.70$.

6.4.9 Surface Brightness

The intrinsic surface brightness of an extended source, such as a galaxy, is defined as its luminosity per unit area, A_s , per 4π steradians of solid angle,

$$\mu_0 = \frac{1}{4\pi} \frac{L_s}{A_s}. \quad (6.69)$$

The observed surface brightness of a high redshift extended object is defined as the collected energy per unit time per unit area of the collector. Thus, surface brightness is the observed flux, $F(z)$, per unit solid angle of the object, $\Delta\Omega$, or

$$\mu(z) = \frac{F(z)}{\Delta\Omega} = \frac{1}{\Delta\Omega} \frac{L_s}{4\pi D_L^2(z)}. \quad (6.70)$$

where Eq. 6.67 has been substituted for $F(z)$. The proper area, A_s , on the object over a unit solid angle $\Delta\Omega$ is

$$A_s = D_A^2(z) \Delta\Omega. \quad (6.71)$$

The expression for A_s follows from the same principle as the linear proper separations given by Eq. 6.41. Substituting for $\Delta\Omega$ in Eq. 6.70 yields

$$\mu(z) = \frac{D_A^2(z)}{D_L^2(z)} \left(\frac{1}{4\pi} \frac{L_s}{A_s} \right) = \frac{D_A^2(z)}{D_L^2(z)} \mu_0. \quad (6.72)$$

From Eq. 6.68, $D_L(z)/D_A(z) = (1+z)^2$, which yields

$$\mu(z) = \frac{\mu_0}{(1+z)^4}. \quad (6.73)$$

Note that the observed surface brightness is fully independent of the cosmological parameters; it depends only upon the evolution in the scale factor to the fourth power.

6.4.10 Absorption “Distance”

The absorption distance, $X(z)$, is an operational definition; it is a measure of “distance” such that the probability of a line of sight intersecting a population of object with non-evolving number density and cross section is constant. It is an essential quantity for testing redshift evolution.

Consider a population of a single type of gas structure with present epoch number density n_0 and cross section σ_0 . Since n_0 is the present epoch density, it is equivalent to the co-moving density at all redshifts. Because of universal expansion, the proper number density at earlier times was higher at earlier cosmic times, having redshift evolution $n(z) = n_0(1+z)^3$.

The scenario of line of sight interception in an expanding volume of space is illustrated in Figure 6.13. The cross section of the source beam, $\sigma_b(z)$, is not constant with redshift. However, if we neglect geometric and orientation effects from the gas structures the beam cross section cancels from the probability of intersection.

In a cosmic volume segment along the proper distance increment, dl , centered on redshift z , the probable number of “target” structures that can potentially interact with the beam is

$$dN_T(z) = n(z)\sigma_b(z) dl = n_0\sigma_b(z)(1+z)^3 dl, \quad (6.74)$$

where

$$dl = c dt = \frac{c}{H_0} \frac{dz}{(1+z)E(z)} = D_H \frac{dz}{(1+z)E(z)}, \quad (6.75)$$

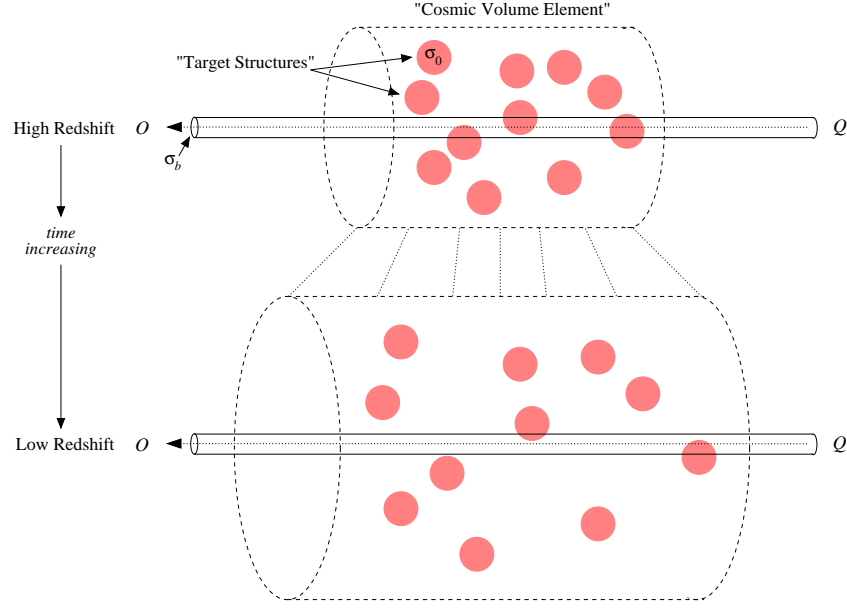


Figure 6.13 An schematic of a source beam of cross section σ_b through a representative cosmic volume element of the universe populated by target structures, such as absorbers, with number density n_0 and cross sections σ_0 . The source is at location Q , and the observer is at location O . The upper panel represents earlier cosmic times at higher redshifts and the lower panel represents later cosmic times at lower redshifts. For illustration, the dotted vertical lines bridging the two cosmic times provides the co-moving behavior of the expansion.

follows from Eq. 6.5. From probability considerations, the number of potential target structures with which the beam will actually interact within this cosmic volume segment, is $dN = dN_T(z) [\sigma_0/\sigma_b(z)]$, where $\sigma_0/\sigma_b(z)$ is the fractional interaction area per structure. Substitution yields

$$\frac{dN}{dz} = D_H n_0 \sigma_0 \frac{(1+z)^2}{E(z)}, \quad (6.76)$$

a quantity known as the redshift path density.

The redshift path density is a centrally important quantity for constraining the redshift evolution of the product $n_0 \sigma_0$ of a carefully selected population of absorbers. The redshift dependence of dN/dz can be written as the dimensionless quantity

$$\frac{dX}{dz} = \frac{(1+z)^2}{E(z)}, \quad (6.77)$$

Integrating, we obtain the absorption distance

$$X(z) = \int_0^z \frac{dX}{dz} dz. \quad (6.78)$$

The quantities dX/dz and $X(z)$ are illustrated in Figures 6.14(a) and 6.14(b), respectively, for the low-density, Einstein-de Sitter, and Λ cosmologies.

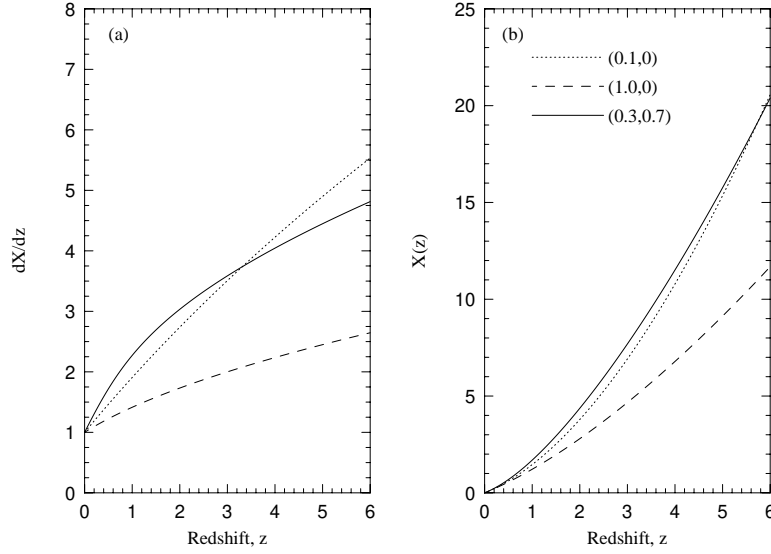


Figure 6.14 (a) The functional form of dX/dz . Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and Λ cosmology (0.3,0.7), solid. [Adapted from Fig. 7 of Hogg (1999)] (b) The absorption distance, $X(z)$, given by Eq. 6.78.

A non-evolving population of absorbers with constant co-moving density and cross section will have a non-changing probability of interception per unit absorption distance,

$$\frac{dN}{dX} = \frac{dN}{dz} \frac{dz}{dX} = D_H n_0 \sigma_0. \quad (6.79)$$

Thus, any evolution in the product $n_0 \sigma_0$ will result in a non-constant dN/dX .

In a $\Omega_\Lambda = 0$ cosmology, the expression for $X(z)$ can be greatly simplified,

$$\frac{dX}{dz} = \frac{(1+z)}{\sqrt{1+2q_0z}}, \quad (6.80)$$

where $E(z)$ is given by Eq. 6.18. Integrating (see Eq. 6.78) yields

$$X(z) = \begin{cases} \frac{1}{2} [(1+z)^2 - 1] & q_0 = 0 \\ \frac{2}{3} [(1+z)^{3/2} - 1] & q_0 = 1/2. \end{cases} \quad (6.81)$$

6.5 Cosmological Volume

The proper volume element is the product of the proper area element, dA , observed over solid angle $d\Omega$ and the proper distance element, dl ,

$$dV = dA dl. \quad (6.82)$$

The proper area is

$$dA = D_A^2(z) d\Omega, \quad (6.83)$$

where D_A is the angular diameter distance (Eq. 6.40). The expression for dA follows from the same principle as the linear proper separations given by Eq. 6.41. The proper distance element is given by Eq. 6.5,

$$dl = c dt = D_H \frac{dz}{(1+z)E(z)}. \quad (6.84)$$

Written out, the proper volume element is

$$dV = D_H \frac{D_A^2(z)}{(1+z)E(z)} d\Omega dz. \quad (6.85)$$

More commonly applied in cosmological studies is the co-moving volume, V_c . In a co-moving volume, the number density of non-evolving objects remains constant as the universe expands. The co-moving volume element, dV_c , is the product of the co-moving area element, dA_c , observed over solid angle $d\Omega$ and the radial co-moving distance element, dD_c ,

$$dV_c = dA_c dD_c. \quad (6.86)$$

The co-moving area element is

$$dA_c = D_T^2(z) d\Omega = (1+z)^2 D_A^2(z) d\Omega, \quad (6.87)$$

where $D_T(z)$ is the transverse co-moving distance (Eq. 6.37) and the right hand term follows from the definition of the angular diameter distance (Eq. 6.40). Written out, the co-moving volume element is

$$dV_c = D_H \frac{(1+z)^2 D_A^2(z)}{E(z)} d\Omega dz. \quad (6.88)$$

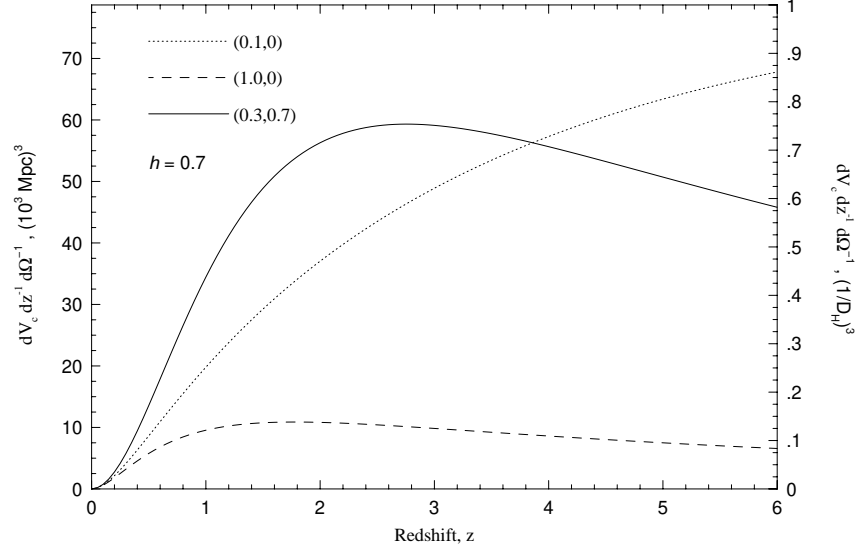


Figure 6.15 The co-moving volume element per unit redshift per unit solid angle, given by Eq. 6.88. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1,0), dotted; (ii) Einstein-de Sitter (1.0,0), dashed; and the Λ cosmology (0.3,0.7), solid. The left hand scale provides the volume in cubic Megaparsecs ($\times 10^9$) for $h = 0.7$. The right hand scale provides the distance in units of D_H (Figure adapted from Hogg, 1999). Take note that $\Omega_k = 0$ for both the Einstein-de Sitter and Λ cosmologies, so that $D_A = D_c/(1+z)$. However, for the low-density case, $\Omega_k = 1 - \Omega_m = 0.9$, so that $D_A = D_T/(1+z)$, which must be computed from Eq. 6.37 with $\Omega_k > 0$. [Adapted from Fig. 5 of Hogg (1999)]

Note that $dV = dV_c/(1+z)^3$, as might be intuitively expected.

The co-moving volume elements, Eq. 6.88 is shown in Figure 6.15 for the low-density, Einstein-de Sitter, and Λ cosmologies. The right hand scale is units of the D_H^3 , a quantity known as the Hubble volume². The left hand axis provides the scale in 10^9 Mpc^3 for $h = 0.7$. Note that the co-volume element is very sensitive to the cosmological parameters.

The co-moving volume to redshift z is

$$V_c(z) = D_H \Theta^2 \int_0^\infty \frac{(1+z)^2 D_A^2(z)}{E(z)} dz. \quad (6.89)$$

where

$$\Theta^2 = \iint d\Omega = \int_{\Delta\alpha} d\alpha \int_{\Delta\delta} \cos \delta d\delta, \quad (6.90)$$

² The Hubble volume has no geometric factors (it is technically not a volume), so it is a geometry independent concept. Per the discussion at the end of § 6.6.2, the Hubble volume is related to the volume of the Hubble sphere at the present epoch.

assuming ecliptic coordinates, where α is the right ascension and δ is the declination, such that $\Delta\alpha$ and $\Delta\delta$ are the angular region on the sky. For the “all-sky” volume, $\Theta^2 = 4\pi$. Performing the integration of Eq. 6.89 over redshift (see Carroll, Press, & Turner, 1992) gives

$$V_c(z) = \begin{cases} \Theta^2 \frac{D_c^3(z)}{3} & (\Omega_k = 0) \\ \Theta^2 \frac{D_H^3}{2\Omega_k} \left\{ \frac{D_T(z)}{D_H} [Q(z) - 1] \right\} & (\Omega_k < 0, \Omega_k > 0), \end{cases} \quad (6.91)$$

where

$$Q(z) = \left(1 + \Omega_k \frac{D_T(z)}{D_H} \right)^{1/2}, \quad (6.92)$$

and where $D_T(z)$ is the transverse co-moving distance (Eq. 6.37) appropriate for the value of Ω_k .

From the form of Eq. 6.91, it can be seen that the units are length cubed and that the dependence on the Hubble parameter is h^3 . Normally the volume is expressed in h^3 Mpc or h^3 Gpc, depending upon the size of the volume contained in a survey or study. Since $\Omega_k = 0$ to high precision, volume is directly proportional to $D_c^3(z)$.

6.6 Peculiar and Recessional Velocities

Having defined the connections between redshift and time and also various distances, it is of interest to consider the concept of velocity, which is derived from the ratio of proper distance over travel time. Velocity is the time derivative of the proper distance, $l = a(t)/a_0 D_c$,

$$v = \frac{dl}{dt} = \frac{\dot{a}(t)}{a_0} D_c + \frac{a(t)}{a_0} \dot{D}_c. \quad (6.93)$$

Writing $v = v_{\text{rec}} + v_{\text{pec}}$, we have the definitions of the velocity of recession (due to Hubble flow) and the so-called peculiar velocity,

$$v_{\text{rec}} = \frac{\dot{a}(t)}{a_0} D_c \quad v_{\text{pec}} = \frac{a(t)}{a_0} \dot{D}_c. \quad (6.94)$$

The deduced (not observed) velocity of recession is dependent upon the expansion velocity at the epoch of the observation, $\dot{a}(t)$ (to be discussed in § 6.6.2). The peculiar velocity, on the other hand, is a measurement an object’s velocity difference with respect to the Hubble flow, or with respect to other objects at the same cosmological redshift.

6.6.1 Peculiar Velocities

A peculiar velocity quantifies the line of sight velocity of an object relative to its local inertial frame. Essentially, it is a measure of the line of sight Doppler shift *relative* to the local cosmological frame. This Doppler shift is measured via redshift differences in the emission or absorption features of spectra. Since peculiar velocities are defined in the inertial frame of the object, they obey the rules of special relativity. No object can have a peculiar velocity exceeding light speed.

The peculiar velocity of a photon is always the speed of light, c . Photon velocities can be derived from the definition of v_{pec} and the fact that photons travel on geodesics, $ds = 0$. Rewriting $\dot{D}_c = dD_c/dt$, and applying the chain rule, photons have

$$v_{\text{pec}} = \frac{a(t)}{a_0} \frac{dD_c}{dt} = \frac{a(t)}{a_0} \frac{dD_c}{d\chi} \frac{d\chi}{da} \frac{da}{dt} = c, \quad (6.95)$$

which is obtained by substitution of Eqs. 5.71, 5.73, and 5.69 (rearranged), respectively,

$$\frac{dD_c}{d\chi} = \frac{D_{\text{H}}}{\sqrt{\Omega_k}}, \quad \frac{d\chi}{da} = \frac{a_0}{a^2(t)} \frac{\sqrt{\Omega_k}}{E(t)}, \quad \frac{da}{dt} = c \frac{a(t)E(t)}{D_{\text{H}}}, \quad (6.96)$$

into Eq.6.95. Application of the geodesic, i.e., $ds = 0$, is applied through the co-moving coordinate, $d\chi/da$, since this relationship in particular is derived for photons.

Consider an inertial reference frame local to cosmological redshift z , in which a source object has a redshift of z_s as measured by an observer at $z = 0$. The line of sight component of the peculiar velocity of the object induces a small observed redshift offset, $\Delta z = z_s - z$, relative to the local cosmological redshift (which can be determined only via some averaging technique using many objects at redshift z).

From the measured Δz , and assuming $v_{\text{pec}} \ll c$, the co-moving (observer frame) and proper peculiar velocities are

$$\dot{D}_c = c \Delta z, \quad \text{and} \quad v_{\text{pec}} = \frac{\dot{D}_c}{1+z} = \frac{c \Delta z}{1+z}, \quad (6.97)$$

respectively, where $v_{\text{pec}} = [a(t)/a_0]\dot{D}_c = \dot{D}_c/(1+z)$ follows from Eq. 6.94.

The non-relativistic expression for v_{pec} can also be obtained directly from the non-relativistic Doppler formula, $\Delta v/c = \Delta\lambda/\lambda$. Let $\lambda_o = \lambda_r(1+z_s)$ be the observed wavelength associated with a spectral feature of rest wavelength λ_r ; the deduced redshift of the object is z_s . Assuming the object is local to

a cosmological frame at redshift z , the peculiar velocity is

$$\frac{v_{\text{pec}}}{c} = \frac{\Delta\lambda}{\lambda} = \frac{\lambda_r(1+z_s) - \lambda_r(1+z)}{\lambda_r(1+z)} = \frac{(1+z_s) - (1+z)}{(1+z)} = \frac{\Delta z}{1+z}, \quad (6.98)$$

where $\Delta z = z_s - z$. The convention is that $v > 0$ means the object is moving away from the observer relative to its local frame, and $v < 0$ means it is moving toward the observer relative to its local frame.

Relaxing $v_{\text{pec}} \ll c$, we are required to employ the well known relativistic Doppler formula to obtain the observed and rest-frame peculiar velocity,

$$\dot{D}_c = c \frac{[1 + \Delta z]^2 - 1}{[1 + \Delta z]^2 + 1}, \quad \text{and} \quad v_{\text{pec}} = c \frac{\left[1 + \frac{\Delta z}{1+z}\right]^2 - 1}{\left[1 + \frac{\Delta z}{1+z}\right]^2 + 1}, \quad (6.99)$$

which reduces to Eq. 6.97 for $\Delta z \ll 1$.

The interpretation of \dot{D}_c is that it is the peculiar velocity of the object with respect to the Hubble flow as measured in the observer's frame ($z = 0$). It is a co-moving quantity that is directly proportional to the measured Δz . In contrast, v_{pec} is the peculiar velocity of the object with respect to the Hubble flow as would be measured in the object's frame of reference. It is a rest-frame quantity that is obtained by scaling the measured Δz by the scale factor at the redshift of the object, i.e., $(1+z)^{-1}$.

The peculiar velocity can also be applied to obtain the line of sight velocity difference between two cosmological objects (however, it must be assumed that the two objects reside in the same cosmological inertial frame).

Consider two cosmological objects along the same line of sight such that object 1 has observed redshift z_1 and object 2 has observed redshift z_2 , such that $z_1 \neq z_2$. The non-relativistic co-moving and proper velocity separations between the objects along the line of sight are

$$\Delta \dot{D}_c = c \Delta z_{12}, \quad \text{and} \quad \Delta v_{12} = \frac{c \Delta z_{12}}{1 + \bar{z}}, \quad (6.100)$$

respectively, where $\Delta z_{12} = z_1 - z_2$, and $\bar{z} = (z_1 + z_2)/2$. If $z_1 < z_2$, then Δv_{12} is negative, which is to be interpreted such that the object 1 is moving toward the observer relative to object 2. A good rule of thumb is $\Delta z_{12} = 0.0001$ yields $\Delta \dot{D}_c = 30 \text{ km s}^{-1}$ for the line of sight velocity separation in the observer frame. The proper velocity separation is computed by dividing by $1 + \bar{z}$,

$$\Delta v_{12} = \frac{\Delta z_{12}}{0.0001} \cdot \frac{30}{1 + \bar{z}} \text{ km s}^{-1}. \quad (6.101)$$

Here, the speed of light has been rounded to $3.0 \times 10^5 \text{ km s}^{-1}$. For example, for $\Delta z_{12} = 0.002$ at $\bar{z} = 1$, the proper velocity separation is approximately $\Delta v_{12} = 20 \cdot (30/2) = 300 \text{ km s}^{-1}$.

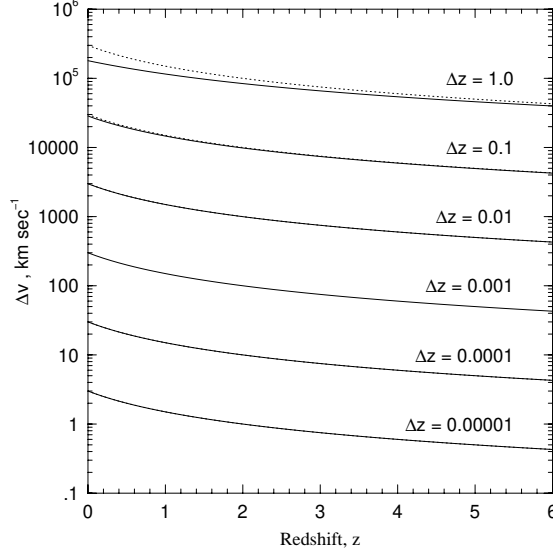


Figure 6.16 The rest-frame velocity differences, Δv_{12} , between two cosmological objects at redshift z with redshift offsets of $\Delta z = 0.00001, 0.0001, 0.001, 0.01, 0.1$, and 1.0 . The solid curves are given by Eq. 6.102, the full relativistic formula (dotted curves) are given by Eq. 6.100, for non-relativistic velocities. Note that the non-relativistic treatment is a good approximation for $\Delta z_{12} \leq 0.1$.

For a redshift offset of $\Delta z_{12} = 0.1$, Eqs. 6.100 and 6.101 over predict Δv_{12} by 5%. The relativistic expression for the proper velocity separation is

$$\Delta v_{12} = c \frac{(1 + z_1)^2 - (1 + z_2)^2}{(1 + z_1)^2 + (1 + z_2)^2}, \quad (6.102)$$

as derived directly from Eq. 6.99 for $\Delta z_{12} = z_1 - z_2$ and $z = z_2$. Both Eq. 6.100 and Eq. 6.102 are symmetric in that $\Delta v_{21} = -\Delta v_{12}$. Eqs. 6.102 (solid curves) and 6.100 (dotted curves) are illustrated in Figure 6.16 for various Δz_{12} offsets as a function of redshift. Note that the curves obey Eq. 6.101 for a large range of Δz_{12} ; it is not until $\Delta z_{12} \geq 0.1$ that the relativistic treatment becomes necessary. The interpretation of $\Delta \dot{D}_c$ is that it is the line of sight velocity differences between the two objects (assuming they are in the same cosmological reference frame), but as measured in the observer's inertial frame at $z = 0$. In contrast, Δv_{12} is the line of sight velocity differences between the two objects in the local inertial frame of the objects, i.e., in the objects' rest frame.

6.6.2 Recessional Velocities

Applying the relativistic formula (Eq. 6.102) for large redshift objects, as observers at $z_2 = 0$, we would naively assign an apparent recession velocity of

$$v_{\text{rec}} = c \frac{(1 + z_s)^2 - 1}{(1 + z_s)^2 + 1} \quad (\text{incorrect}), \quad (6.103)$$

to a cosmological object at redshift $z_s = z_1$. Wrong; don't ever do this. It should remain appreciated that the cosmological velocity of recession is not an inertial frame measurement and thus does not obey the rules of special relativity. To apply Eq. 6.103 as a means of determining a velocity of recession to a cosmological object is flat out incorrect.

Cosmological recession velocities are not interpreted as a Doppler shift. Yet, a consequence of the expanding universe paradigm is that redshift is directly related to recessional velocity (Harrison, 1993; Bunn & Hogg, 2009). There has been great confusion over the interpretation of this velocity because of the notion that no objects can recede at faster than light speed. Actually, it is a common misconception that the velocities of objects carried by the Hubble flow cannot exceed the speed of light (see Davis & Lineweaver, 2004). In the expanding universe formalism, there is no global reference frame. It is general relativity and not special relativity that is employed for computing cosmological dynamics.

In general relativity, motions outside the observer's inertial reference frame can be properly treated and can be fully consistent with faster than light speed motion. From Eq. 6.94, we have

$$v_{\text{rec}} = \frac{\dot{a}(t)}{a_0} D_c. \quad (6.104)$$

Consider a general observer at time $t = t_o$, which is not necessarily the current epoch, $t = t_0$. Assuming an arbitrary epoch of observation, we rearrange $\dot{a}(t_o)/a_0$ to obtain

$$\frac{\dot{a}(t_o)}{a_0} = \frac{\dot{a}(t_o)}{a(t_o)} \frac{a(t_o)}{a_0} = \frac{H(z_o)}{1 + z_o} = \frac{H_0 E(z_o)}{1 + z_o}, \quad (6.105)$$

where z_o is the redshift of the observer, and where $E(z_o)$ is given by Eq. 6.15. We can interpret Eq. 6.105 as a “proper expansion rate” of the universe for an observer at redshift z_o (as will be further discussed below). For an observer at z_o , the co-moving distance appearing in Eq. 6.104 is the co-moving radial separation, $D_c(z_o, z_s)$, between the observer and the source (Eq. 6.33). From Eq. 6.33, and noting that $c = H_0 D_H$, the velocity of recession of a

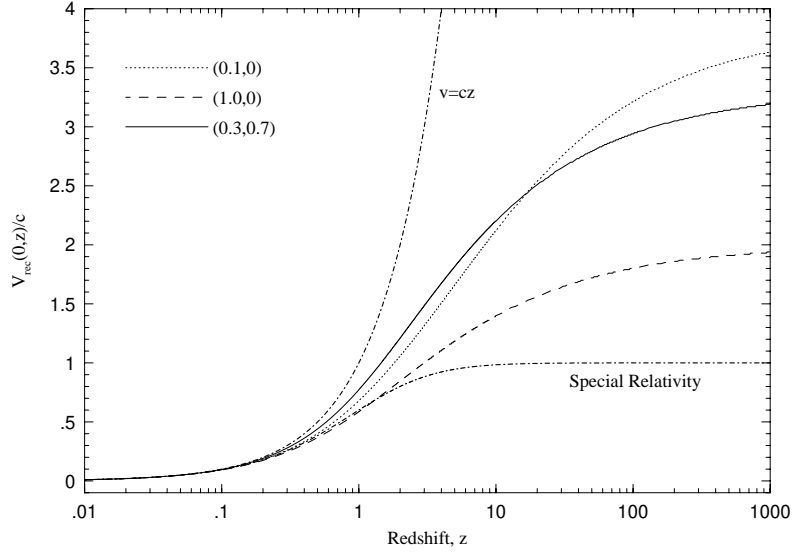


Figure 6.17 The velocity of recession, $v_{\text{rec}}(0, z)$, of objects carried by the Hubble flow as a function of redshift for a present epoch observer. Three cosmologies are shown, denoted by $(\Omega_m, \Omega_\Lambda)$; (i) the low-density (0.1, 0), dotted; (ii) Einstein-de Sitter (1.0, 0), dashed; and the Λ cosmology (0.3, 0.7), solid. The special relativity case (Eq. 6.103) and the linear relation, $v = cz$, are shown as thick dash-dot curves. For the Λ cosmology, the recession velocity of objects exceeds the speed of light at $z \simeq 1.46$ [Adapted from Fig. 2 of Davis & Lineweaver (2004)]

source or object at $z_s > z_o$ for an observer at z_o is

$$v_{\text{rec}}(z_o, z_s) = \frac{c E(z_o)}{1 + z_o} \int_0^{z_{os}} \frac{dz}{E(z)} = \frac{c E(z_o)}{1 + z_o} \frac{D_c(z_o, z_s)}{D_H}, \quad (6.106)$$

where z_{os} is the reduced redshift (Eq. 6.32),

$$1 + z_{os} = \frac{1 + z_s}{1 + z_o}, \quad (6.107)$$

for an observed source redshift, z_s .

For present epoch observers, $z_o = 0$, we have $E(z_o)/(1 + z_o) = 1$. The co-moving radial separation reduces to the total co-moving distance. Thus, Eq. 6.106 simplifies to

$$v_{\text{rec}}(0, z) = c \int_0^z \frac{dz}{E(z)} = c \frac{D_c(z)}{D_H}. \quad (6.108)$$

The recessional velocity of cosmological objects for a present epoch observer is simply a multiple of the speed of light, where that multiple is the number of Hubble distances to the object. For $D_c(z)/D_H < 1$, $v_{\text{rec}}(0, z)$ is some

fraction of light speed. For $D_c(z)/D_H > 1$, $v_{\text{rec}}(0, z)$ is greater than the speed of light. Objects at a co-moving distance equal to the Hubble distance have $v_{\text{rec}}(0, z) = c$.

In Figure 6.17, the recession velocity of cosmological objects is shown as a function of redshift for present epoch observers, i.e., $z_o = 0$. The low-density (dotted curve), Einstein-de Sitter (dashed), and Λ cosmologies (solid) are illustrated, as are the special relativity velocity law (Eq. 6.103) and the linear velocity law, $v = cz$. For $z < 0.1$, the various recession velocities are fairly consistent. For the “737” cosmology, the recession velocity exceeds the speed of light at $z \simeq 1.46$. It is important to recognize that recessional velocity is not an observed quantity; redshift is the observed quantity. Recessional velocity is a quantity interpreted only through the cosmological paradigm.

Radiative Transfer

In this chapter, we introduce and discuss the general formalism of radiative transfer, which results from the energy exchange between the particle field of a gaseous medium and the radiation field. We begin by defining the radiation field, including the specific intensity and the mean intensity. We then equate these quantities with the photon field itself and derive the energy density of the radiation field. The observed (measured) radiation from a distance emitting source is quantified as the flux. We derive the flux integral and then apply it to obtain expressions for the observed flux.

The equation of radiative transfer is derived in terms of macroscopic quantities called the extinction and emission coefficients, which are parameters that represent all the microscopic physics occurring between the radiation field and particle field that result in emission, absorption, and scattering. We then solve the transfer equation for a one-dimension geometry that applies for virtually all cosmological applications. Finally, we describe the astronomical absorption line spectrum in detail and derive the resulting spectrum for the case of partial covering of the emitting source by an intervening “cloud”.

Since radiative transfer is central to astronomical spectroscopy, and it has been covered in many excellent texts, we refer the student to a select few additional resources (Emerson, 1996; Gray, 1992; Mihalas, 1978; Novotny, 1973; Rybicki & Lightman, 2004).

7.1 The Radiation Field

A full description of a radiation field requires both the energy (i.e., magnitude) and propagation direction of photons. Thus, the radiation field is a vector field. Let the propagation direction of a photon be denoted by the unit vector $\hat{\mathbf{s}}$. In most all observer-centric applications, the $\hat{\mathbf{s}}$ direction will be the line of sight to the observer. Propagating photons will pass through

some point in position and time, the coordinate of which we will denote with the ordered pair (\mathbf{r}, t) , where \mathbf{r} represents radial vector from the coordinate system origin.

7.1.1 The Geometry

The two most convenient coordinate systems for describing the radiation field are the Cartesian, $\mathbf{r} = (x, y, z)$, and spherical, $\mathbf{r} = (r, \phi, \theta)$, systems. These two systems and the relationship between their coordinates are illustrated in Figure 7.1. The standard relationship between Cartesian and spherical coordinates is

$$\begin{aligned} x &= r \sin \theta \cos \phi & r &= (x^2 + y^2 + z^2)^{1/2} \\ y &= r \sin \theta \sin \phi & \theta &= \cos^{-1}(z/r) \\ z &= r \cos \theta & \phi &= \tan^{-1}(y/x). \end{aligned} \quad (7.1)$$

As illustrated in Figure 7.1(a), the unit base vector directions in the Cartesian system are defined parallel to the principle axes of the system. They are fixed in space and independent of the location \mathbf{r} . On the contrary, as shown in Figure 7.1(b), in the spherical coordinates system the directions of the unit base vectors depend on the ϕ and θ location, but not on r .

For a photon propagating in the $\hat{\mathbf{s}}$ direction, the unit vector is most conveniently written in terms of the orthogonal unit base vectors of the Cartesian coordinate system

$$\hat{\mathbf{s}} = \cos \alpha \hat{\mathbf{i}} + \cos \beta \hat{\mathbf{j}} + \cos \gamma \hat{\mathbf{k}}, \quad (7.2)$$

where $\cos \alpha$, $\cos \beta$, and $\cos \gamma$ are the direction cosines, which obey the relation $\cos^2 \alpha + \cos^2 \beta + \cos^2 \gamma = 1$. Further inspection of Figure 7.1(a) shows that the direction cosines are given by the inner products

$$\cos \alpha = \hat{\mathbf{s}} \cdot \hat{\mathbf{i}} \quad \cos \beta = \hat{\mathbf{s}} \cdot \hat{\mathbf{j}} \quad \cos \gamma = \hat{\mathbf{s}} \cdot \hat{\mathbf{k}}. \quad (7.3)$$

We now develop the formalism for the geometry of emitting regions from luminous sources. Consider a spherical source of radius R and total surface area A that is emitting radiation from the surface. The radiation will be emitted from infinitesimal surface area elements, as illustrated by the shaded region on the sphere in Figure 7.2.

Each area element located at R, ϕ, θ can be placed in the context of an annulus of angular width $d\theta$ about the azimuth angle ϕ (in rotation about the $\hat{\mathbf{k}}$ vector). The arc length of the area element over angle $d\theta$ is $R d\theta$. The arc length in the azimuthal direction is $R \sin \theta d\phi$. This latter arc length does not lie on a great circle, but on a circle of radius $R \sin \theta$, as illustrated

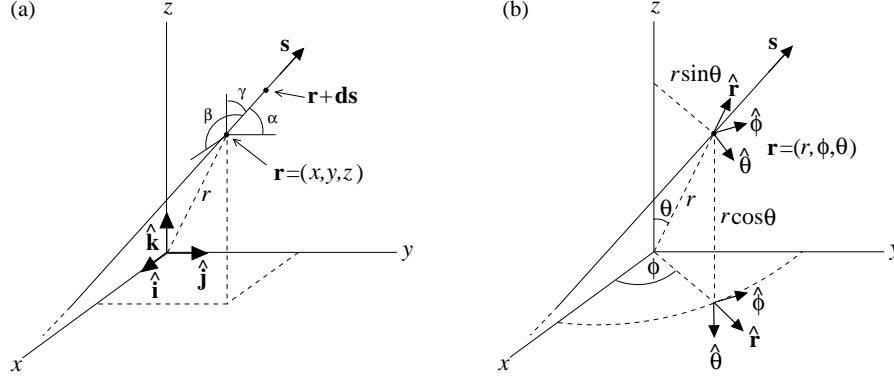


Figure 7.1 The unit vector $\hat{\mathbf{s}}$ represents propagating radiation that passes through point \mathbf{r} in some arbitrary direction. (a) In the Cartesian coordinate system the point \mathbf{r} is represented by the ordered pair (x, y, z) . Unit vector $\hat{\mathbf{s}}$ is written in terms of the orthogonal unit base vectors $\hat{\mathbf{i}}$, $\hat{\mathbf{j}}$, and $\hat{\mathbf{k}}$. (b) In the spherical coordinate system the point \mathbf{r} is represented by the ordered pair (r, ϕ, θ) , where r is the radial distance from the origin, ϕ is the azimuthal angle that sweeps counterclockwise from $0 \leq \phi \leq 2\pi$ rotated around the z axis ($\hat{\mathbf{k}}$ vector) with $\phi = 0$ in the $\hat{\mathbf{i}}$ direction, and θ is the polar angle that sweeps from the $\hat{\mathbf{k}}$ direction through the range $0 \leq \theta \leq \pi$ (to the $-\hat{\mathbf{k}}$ direction). The orthogonal unit base vectors are $\hat{\mathbf{r}}$, $\hat{\boldsymbol{\phi}}$, and $\hat{\boldsymbol{\theta}}$.

on the two-dimensional collapsed view from the $\hat{\mathbf{k}}$ point of view. The area element is the product of the two arc lengths,

$$dA = R^2 \sin \theta d\theta d\phi, \quad (7.4)$$

and has unit vector direction $\hat{\mathbf{n}} = \hat{\mathbf{r}}$ normal to the surface, i.e., $d\mathbf{A} = dA\hat{\mathbf{r}}$.

Viewed from the center of the sphere, the solid angle subtended by the surface area element is

$$d\Omega = \frac{dA}{R^2} = \sin \theta d\theta d\phi. \quad (7.5)$$

The same area element viewed in projection from the point of view of an observer located along the $+z$ axis (the $\hat{\mathbf{k}}$ direction) will be $(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) dA$. This projected area is illustrated in Figure 7.2 as the shaded region on the plane of projection to the right of the source.

For an observer at distance D along the $+z$ axis, the projected area element will subtend solid angle $d\Omega_{\text{obs}} = (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) dA / D^2$. By selecting the observer to be in the $+z$ direction, we obtain $(\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) = \cos \theta$, so that,

$$d\Omega_{\text{obs}} = \frac{R^2}{D^2} (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) \sin \theta d\theta d\phi = \frac{R^2}{D^2} \cos \theta \sin \theta d\theta d\phi. \quad (7.6)$$

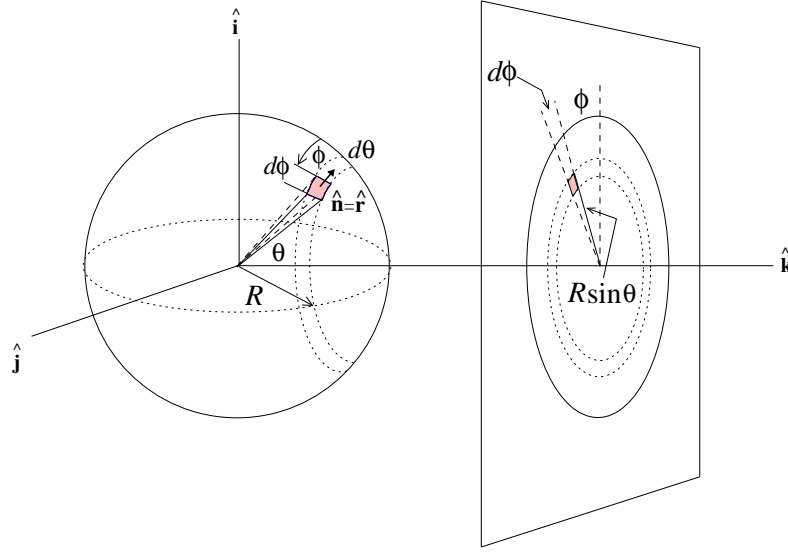


Figure 7.2 A schematic of the area element (shaded region) on a surface of a sphere with radius R . The area element is a vector given by $d\mathbf{A} = dA \hat{\mathbf{n}} = R^2 \sin \theta d\theta d\phi \hat{\mathbf{n}}$, where $\hat{\mathbf{n}} = \hat{\mathbf{r}}$ is the unit vector normal to the area surface. Viewed from the $+z$ axis, the projected area (shaded region on the plane of projection) is $\hat{\mathbf{k}} \cdot d\mathbf{A} = (\hat{\mathbf{k}} \cdot \hat{\mathbf{n}}) dA = (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) dA = dA \cos \theta$, yielding $dA_{\text{obs}} = R^2 \cos \theta \sin \theta d\theta d\phi$.

Note that, the observer and the observed projected area element and solid angle are all defined in terms of their geometric location on the source relative to the origin of the coordinate system, which is the center of the source.

7.1.2 The Specific Intensity

Within any given volume at any instant, countless photons are propagating in as many countless directions. Each photon carries and transports energy. To quantify how electromagnetic energy is transported through a given location in space and time, we consider the energy passing through a unit area centered at location $\mathbf{r} = (x, y, z) = (r, \phi, \theta)$ within the radiation field. Through this area photons are transporting energy in many directions simultaneously. We thus need a description of the radiation field that accounts for the energy, direction, location in time and space, and direction relative to the area element.

The fundamental quantity describing the propagation of radiation is called the specific intensity, which is defined as the incremental amount of energy, $d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, transported in direction $\hat{\mathbf{s}}$ by radiation per unit frequency range

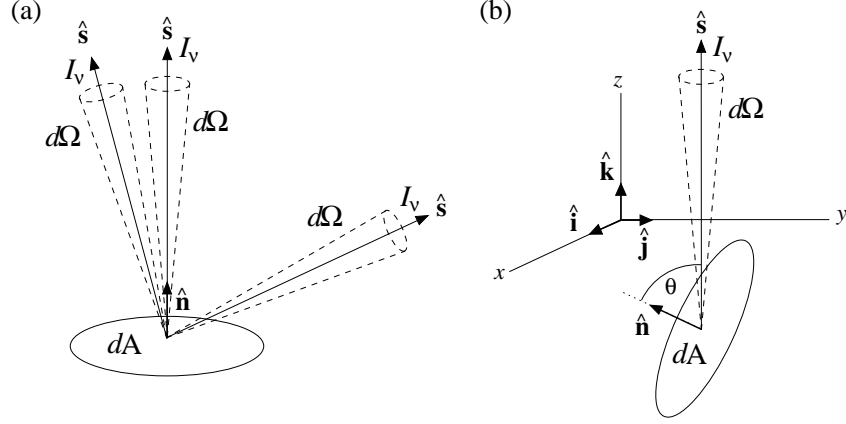


Figure 7.3 A schematic of specific intensity $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$. (a) Propagation of the radiation field in multiple $\hat{\mathbf{s}}$ directions through area element dA into solid angles $d\Omega$. (b) In cases where the area element normal, $\hat{\mathbf{n}}$, is not parallel to $\hat{\mathbf{s}}$, the increment of radiative energy $d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ transported in the $\hat{\mathbf{s}}$ direction is through the projected area $\hat{\mathbf{s}} \cdot d\mathbf{A} = (\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) dA = dA \cos \theta$.

$\nu \rightarrow \nu + d\nu$ passing through a *projected* area centered on point \mathbf{r} into solid angle $d\Omega$ in time interval $t \rightarrow t + dt$,

$$\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \frac{d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{(\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt}. \quad (7.7)$$

The units of specific intensity are $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1} \text{str}^{-1}$. In § 7.1.4, we will directly connect the specific intensity to the number distribution function of photons.

The notation $(\mathbf{r}, t; \hat{\mathbf{s}})$ accounts for the fact that at a given point in space and time (\mathbf{r}, t) the magnitude of the specific intensity can differ with propagation direction, $\hat{\mathbf{s}}$. If the specific intensity is spatially isotropic, then the magnitude is independent of direction, and we write $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \mathcal{I}_\nu(\mathbf{r}, t)$. If $\mathcal{I}_\nu(\mathbf{r}, t)$ is also constant with time, then we write $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \mathcal{I}_\nu(\mathbf{r})$. For a steady-state isotropic radiation field, the specific intensity passing through point \mathbf{r} is unchanging and of equal magnitude into all solid angles.

A schematic of the geometric relationships between $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, $\hat{\mathbf{s}}$, $d\mathbf{A}$, and $d\Omega$ is presented in Figure 7.3(a). Note that when $\hat{\mathbf{s}}$ is not parallel to $d\mathbf{A} = dA\hat{\mathbf{n}}$, the projected area element through which the energy is transported is $(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) dA = dA \cos \theta$, as illustrated in Figure 7.3(b). Thus, for a given radiation field, $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, the energy transported through the area element is modulated by the geometric factor $\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}$.

Alternatively, the specific intensity can be written in terms of the wave-

length interval $\lambda \rightarrow \lambda + d\lambda$,

$$\mathcal{I}_\lambda(\mathbf{r}, t; \hat{\mathbf{s}}) = \frac{d\epsilon_\lambda(\mathbf{r}, t; \hat{\mathbf{s}})}{(\hat{\mathbf{s}} \cdot \mathbf{dA}) d\Omega d\lambda dt}, \quad (7.8)$$

which has units $\text{erg s}^{-1} \text{cm}^{-2} \text{str}^{-1} \text{\AA}^{-1}$. Because the energy increment over a wavelength interval corresponding to the frequency interval must be conserved ($d\epsilon_\nu = d\epsilon_\lambda$), we have $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) d\nu = \mathcal{I}_\lambda(\mathbf{r}, t; \hat{\mathbf{s}}) d\lambda$, which (from $c = \nu\lambda$) yields

$$\mathcal{I}_\lambda(\mathbf{r}, t; \hat{\mathbf{s}}) = \frac{\lambda^2}{c} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \quad (7.9)$$

7.1.3 The mean Intensity

The mean intensity is the average of the specific intensity propagating into all solid angles at location \mathbf{r} at time t . It is the zeroth moment of the radiation field. We have

$$J_\nu(\mathbf{r}, t) = \frac{\int_0^{2\pi} \int_0^\pi \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) d\Omega}{\int_0^{2\pi} \int_0^\pi d\Omega} = \frac{\oint \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) d\Omega}{\oint d\Omega} = \frac{1}{4\pi} \oint \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) d\Omega. \quad (7.10)$$

When evaluating this integral for a non-isotropic specific intensity, $\hat{\mathbf{s}}$ is the direction of propagation into solid angle element $d\Omega$. At each θ and ϕ over the integration, the direction in space of this solid angle, and thus the direction of propagation of the specific intensity is (see Eq. 7.3)

$$\hat{\mathbf{s}} = \hat{\mathbf{s}}(\alpha, \beta, \gamma) = \cos \alpha \hat{\mathbf{i}} + \cos \beta \hat{\mathbf{j}} + \cos \gamma \hat{\mathbf{k}}. \quad (7.11)$$

The units of mean intensity are $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$. If the specific intensity is isotropic, then it can be factored out of the integral and we obtain $J_\nu(\mathbf{r}, t) = \mathcal{I}_\nu(\mathbf{r}, t)$.

7.1.4 The Photon Field

The radiation field is a photon field. The photon field is quantified by $n_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, the distribution of the number of photons in the frequency range $\nu \rightarrow \nu + d\nu$ at location \mathbf{r} propagating in the $\hat{\mathbf{s}}$ direction into solid angle $d\Omega$ in time interval $t \rightarrow t + dt$. Propagating at the velocity c , the number of such photons crossing through a projected area element $\hat{\mathbf{s}} \cdot \mathbf{dA}$, is

$c n_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \mathbf{dA}) d\Omega d\nu dt$. Each photon has energy $h\nu$, so the energy transported is

$$d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = (ch\nu) n_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \mathbf{dA}) d\Omega d\nu dt. \quad (7.12)$$

Equating to the definition of the specific intensity given in Eq. 7.7, we have

$$\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = (ch\nu) n_\nu(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.13)$$

7.2 The Flux Vector

The flux, $\mathbf{F}_\nu(\mathbf{r}, t)$, is a vector quantifying the rate of the net energy carried by radiation across a projected surface area element per unit frequency, i.e., $\mathbf{F}_\nu(\mathbf{r}, t) \cdot \mathbf{dA}$. The expression for the increment of energy per unit frequency per unit time that is transported by radiation propagating in the $\hat{\mathbf{s}}$ direction through an area element \mathbf{dA} is $d\epsilon_\nu/d\nu dt$. From Eq. 7.8, we have

$$\frac{d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{d\nu dt} = \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \mathbf{dA}) d\Omega. \quad (7.14)$$

From the definition of the flux and employing Eq. 7.14, we write

$$\mathbf{F}_\nu(\mathbf{r}, t) \cdot \mathbf{dA} = \iint_{\phi, \theta} \frac{d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{d\nu dt} = \iint_{\phi, \theta} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \mathbf{dA}) d\Omega, \quad (7.15)$$

where the integrals are taken over the coordinates ϕ and θ . Equating the vectors in Eq. 7.15, we see that the flux vector is

$$\mathbf{F}_\nu(\mathbf{r}, t) = \iint_{\phi, \theta} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \hat{\mathbf{s}} d\Omega, \quad (7.16)$$

with units $\text{erg s}^{-1} \text{ cm}^{-2} \text{ Hz}^{-1}$. Expressed per unit frequency, this monochromatic flux is often called the flux density. Per unit wavelength, the flux density is written $\mathbf{F}_\lambda(\mathbf{r}, t)$ with units $\text{erg s}^{-1} \text{ cm}^{-2} \text{ \AA}^{-1}$.

The flux vector can be written in terms of orthogonal flux components, for example, in Cartesian coordinates,

$$\mathbf{F}_\nu(\mathbf{r}, t) = F_\nu^x(\mathbf{r}, t) \hat{\mathbf{i}} + F_\nu^y(\mathbf{r}, t) \hat{\mathbf{j}} + F_\nu^z(\mathbf{r}, t) \hat{\mathbf{k}}, \quad (7.17)$$

where

$$\begin{aligned}
 F_{\nu}^x(\mathbf{r}, t) &= \iint_{\phi, \theta} \mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{i}}) d\Omega \\
 F_{\nu}^y(\mathbf{r}, t) &= \iint_{\phi, \theta} \mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{j}}) d\Omega \\
 F_{\nu}^z(\mathbf{r}, t) &= \iint_{\phi, \theta} \mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{k}}) d\Omega,
 \end{aligned} \tag{7.18}$$

where the first unit vector in the dot product is the propagation direction of the specific intensity and the second is the vector direction of the flux component.

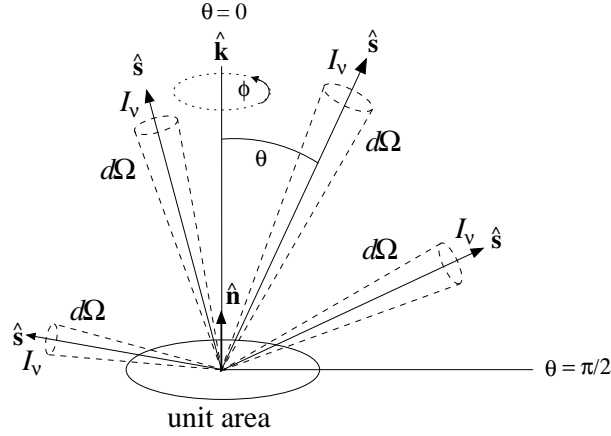


Figure 7.4 To compute the upward flux, $F_{\nu}^z(\mathbf{r}, t)$, through a unit area A , the contributions from all $\mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}})$ passing through the area surface that have a $+\hat{\mathbf{k}}$ component are included in the flux integral. The angle θ accounts for the projected area. The integration is taken over $0 \leq \theta \leq \pi/2$. The angle ϕ integrates the solid angle over the hemisphere.

Observers at great distances from astronomical sources measure only one component of the flux. Commonly, we adopt the component $F_{\nu}^z(\mathbf{r}, t)$. The choice of which component to adopt, of course, is arbitrary. However, for the z component, the polar angle θ physically represents the angle between the emitted specific intensity and this simplifies the flux integral.

For example, consider the flux passing upward through the surface of the unit area illustrated in Figure 7.4. Define the upward direction, $\hat{\mathbf{k}}$, to be parallel to $\hat{\mathbf{n}}$ and place the center of the area at the point \mathbf{r} . The flux

integral is

$$F_\nu(\mathbf{r}, t) = \iint_{\phi, \theta} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{k}}) d\Omega = \iint_{\phi, \theta} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \cos \theta d\Omega. \quad (7.19)$$

Since we wish to obtain the flux only in the upward direction we properly account for the $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ contributions to the flux by including only those having angles of passage through the area element in the range $0 \leq \theta \leq \pi/2$, which after accounting for the azimuthal integration over $0 \leq \phi \leq 2\pi$, covers the single hemisphere above the area surface. We have

$$F_\nu(\mathbf{r}, t) = \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \cos \theta \sin \theta d\theta d\phi. \quad (7.20)$$

Assuming an isotropic radiation field, the specific intensity factors out of the integral and we obtain

$$F_\nu(\mathbf{r}, t) = \pi \mathcal{I}_\nu(\mathbf{r}, t). \quad (7.21)$$

7.2.1 Defining the Observed Flux

The observed flux density is commonly written f_ν or f_λ . The observed flux of a source is the flux component in the direction of the observer; it is a measure of energy per unit time per unit area per unit frequency at the observer's location a distance D from the source. The units of f_ν are $\text{erg s}^{-1} \text{cm}^{-2} \text{Hz}^{-1}$. If measured per wavelength interval, then f_λ is in units of $\text{erg s}^{-1} \text{cm}^{-2} \text{\AA}^{-1}$.

Here, we limit our discussion to the flux measured from an unresolved emitting source. In order to simplify the problem, we will assume a spherical source of radius R_s and that the observer is at a distance $D \gg R_s$ positioned along the $+z$ axis (the $\hat{\mathbf{k}}$ direction).

It is a common misconception to envision the emission from a distant source as propagating radial outward (i.e., in the $\hat{\mathbf{r}}$ direction from every point on the surface) and then proceed to compute the flux based on the distance to and collecting area of the telescope facility by determining the solid angle of the telescope from the point of view of the distant source, i.e., $\Omega \neq A_T/D$. One reason this can seem to be an erroneous approach is that the observed flux would be proportional to the collecting area of the telescope. This would render the observed flux a meaningless quantity from observer to observer!

In reality, as illustrated in Figure 7.4, the radiation emitted through each area element is propagating into all solid angles over the full hemisphere above the local horizon of the emitting area element, and, because the source

is at such great distance, only those photons emitted in a very small solid angle directly toward the observer will be captured by their telescope.

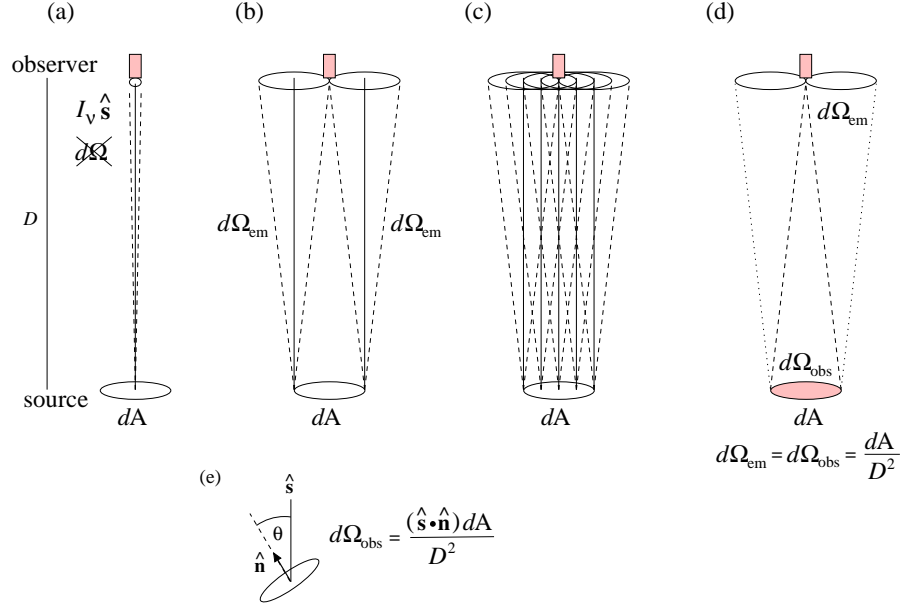


Figure 7.5 Schematic to demonstrate the solid angle, $d\Omega_{\text{em}}$, over which emitted radiation propagates to an observer telescope of area A_{T} a distance D from all points across an area element on the surface of the source. (a) Example of a naïvely and erroneous view of how radiation I_{ν} is collected from the source area element propagating toward an observer telescope in the \hat{s} direction. The solid angle shown, $d\Omega = A_{\text{T}}/D^2$, is incorrect as it does not properly account for all radiating points across the emitting area element. (b) Radiation from the edges of the area element propagating toward the observer in the \hat{s} direction. The minimum solid angles of emission, $d\Omega_{\text{em}}$, for which radiation from the edges of the area element are intercepted by the observer telescope are shown. This minimum solid angle ensures that all emitting points on the surface of the area element are captured by the telescope. (c) Same as (b) except showing emission into solid angle $d\Omega_{\text{em}}$ from multiple locations across the area element. Note that all emitting points are intercepted by the telescope. (d) Same as (b), but showing that the geometric symmetry yields $d\Omega_{\text{em}} = d\Omega_{\text{obs}} = dA/D^2$, the solid angle of the area element from the perspective of the observer. (e) If the area element normal is at an angle θ with respect to the direction to the observer, then the solid angle is $d\Omega_{\text{obs}} = (\hat{s} \cdot \hat{n}) dA/D^2 = (dA/D^2) \cos \theta$.

To determine this solid angle, consider the schematic shown in Figure 7.5. We place the observer in the \hat{s} direction at a distance D from the emitting area element, dA . The minimum solid angle, defined such that all emitting points on the surface of the area element will be captured by the observer

telescope, is denoted $d\Omega_{\text{em}}$. It is determined by the triangle of side D with vertices at the telescope and the edge of the area element (see Figure 7.5(b)). The reason that the minimum solid angle for emission is applied is that the flux is defined as the energy per unit time *per unit area* per unit frequency. If the emission solid angle were larger, then emission from adjacent area elements would overlap, and the measure flux would no longer be per unit area. As show in Figure 7.5(d), this solid angle is symmetric with the solid angle that the area element subtends from the observer point of view. Thus, $d\Omega_{\text{em}} = d\Omega_{\text{obs}}$, where $d\Omega_{\text{obs}} = dA/D^2$. If the area element normal, $\hat{\mathbf{n}}$, is at an angle to the direction to the observer, $\hat{\mathbf{s}}$, then the solid angle is $(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) dA/D^2$.

Since the solid angle over which the emission will be captured by the observer is geometrically equivalent to the observed solid angle of the emitting area element, when computing the observed flux, the solid angle to be applied in the flux integral is $d\Omega_{\text{obs}}$. We thus have,

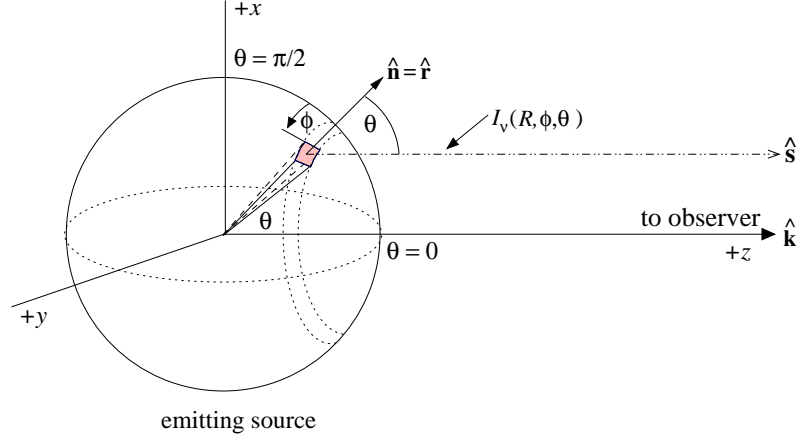
$$f_{\nu}(t) = \iint_{\phi, \theta} \mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) d\Omega_{\text{obs}}, \quad (7.22)$$

where integration is taken from the coordinate system origin over ϕ and θ , and where the dot product $\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}$ accounts for the projection of the area element and allows for possible non-spherical dependence of the area element normal vector, i.e., $\hat{\mathbf{n}} = \hat{\mathbf{n}}(\phi, \theta) \neq \hat{\mathbf{r}}$. Recalling that we locate the observer along the $+z$ axis (in the $\hat{\mathbf{k}}$ direction), so that $\hat{\mathbf{s}} = \hat{\mathbf{k}}$, and that we assume a spherical emitting source, we have $\hat{\mathbf{s}} \cdot \hat{\mathbf{n}} = \hat{\mathbf{k}} \cdot \hat{\mathbf{r}} = \cos \theta$.

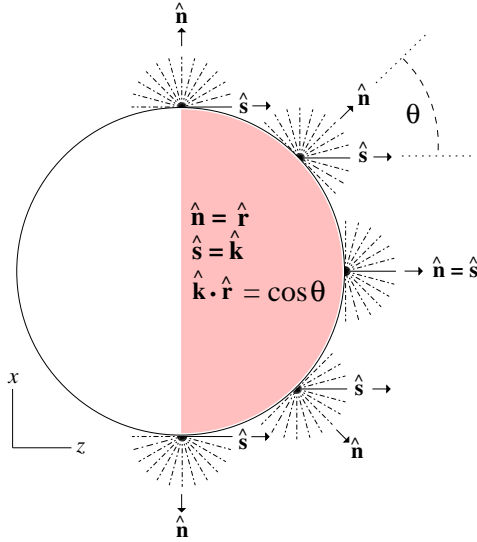
In Figure 7.6(a), we show an area element located at $\mathbf{r} = (R_s, \phi, \theta)$ emitting specific intensity $\mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) = \mathcal{I}_{\nu}(R_s, \phi, \theta, t; \hat{\mathbf{s}})$. For an observer along the $+z$ axis, only the emission across the area element that propagates within an infinitesimally small solid angle converging on the telescope along the direction $\hat{\mathbf{s}} = \hat{\mathbf{k}}$ will be observed. This is further illustrated in Figure 7.6(b) for a side-on view of the source. Each area element centered on location $\mathbf{r} = (R_s, \phi, \theta)$ emits into a hemisphere above its local horizon on the surface. The photons propagating with $\hat{\mathbf{s}} = \hat{\mathbf{k}}$ (those that are observed) are indicated with solid lines. From this schematic, it is clear that the observed flux is a single vector component of the total surface flux from the hemisphere of the source facing the observer (shaded region). The geometry of the calculation is illustrated in Figure 7.6(c). The coordinate system is centered on the source and the integration of the flux integral is taken over the hemisphere facing the observer, which is $0 \leq \theta \leq \pi/2$ and $0 \leq \phi \leq 2\pi$. We have

$$f_{\nu}(t) = \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_{\nu}(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) d\Omega_{\text{obs}}. \quad (7.23)$$

(a)



(b)



(c)

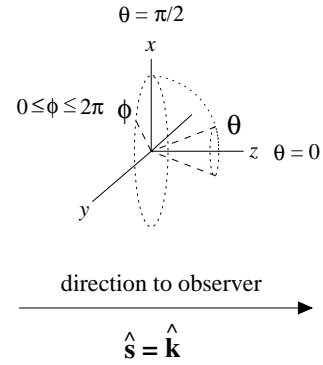


Figure 7.6 (a) A schematic of emission from the surface area element, dA (shaded region), on an emitting source of radius R . The observer is at distance $D \gg R$ along the $+z$ axis (the $\hat{\mathbf{k}}$ direction). The specific intensity is emitted over all outward directed solid angles from each area element at $\mathbf{r} = (R, \phi, \theta)$. However, because the observer is at great distance, only those photons propagating in the $\hat{\mathbf{s}} = \hat{\mathbf{k}}$ direction will be intercepted by the observer. The normal vector $\hat{\mathbf{n}} = \hat{\mathbf{r}}$ makes an angle θ with respect to the $\hat{\mathbf{s}}$ direction of the photons. Thus, the intensity emitted from the area element in the direction of the observer passes through projected area element $(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) dA = (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) dA = dA \cos \theta$. (b) Side view of the source illustrating the multi-directional emission into the local hemispheres of five example locations on the surface of the source. Though only locations on the limb are shown for this perspective, all points on the surface emit multi-directionally as shown. The shaded region is the hemisphere of the source viewable from the observer, which limits the polar angle integration in the flux integral to $0 \leq \theta \leq \pi/2$. (c) Schematic of the geometric relationships for the integration.

7.2.2 Measuring the Observed Flux

To measure the observed flux of a source, the emitted radiation is collected by a telescope of effective area A_T over a small span of time, which we call the integration time, $\Delta\tau$. For a spectrum, the resulting measurement yields the energy density in units $\text{erg s}^{-1} \text{Hz}^{-1}$, which is given by the expression

$$E_\nu = A_T \int_0^{\Delta\tau} f_\nu(t) dt. \quad (7.24)$$

Clearly, the larger the collecting area and the longer the integration time, the larger the measured energy density. If the source emission is steady state over the integration time, then $f_\nu(t) = f_\nu$ and factors out of the integral, yielding $E_\nu = (A_T \Delta\tau) f_\nu$. Correcting¹ for the telescope area and integration time, we obtained the observed flux,

$$f_\nu = \frac{E_\nu}{A_T \Delta\tau}. \quad (7.25)$$

The observed flux necessarily depends upon the solid angle of the collected beam. All instruments have a resolution limit, a minimum solid angle, Ω_r , over which they can collect focused light. If an object subtends a solid angle Ω_s such that $\Omega_s < \Omega_r$, the object is unresolved; if $\Omega_s > \Omega_r$, the object resolved. We will show below that the observed flux of an unresolved object depends upon both its distance and radius.

As illustrated in Figure 7.7, consider a telescope/instrument facility with effective focal length f_T and a limiting angular resolution that projects to an area element on the detector A_r . Let the source be spherical with radius R_s and be at a distance D from the focal point of the telescope. From the focal point, the source subtends a solid angle $\Omega_s = \pi R_s^2/D^2$, where πR_s^2 is the projected area of the source, and the detector resolution element subtends a solid angle $\Omega_r = A_r/f_T^2$. Under the condition that $\Omega_s < \Omega_r$, the source is unresolved, so all radiation captured from the source is distributed across a single resolution element on the detector.

As shown in Figure 7.5, this is given by $d\Omega_{\text{obs}} = dA/D^2$. The area element on the surface at location $\mathbf{r} = (R_s, \phi, \theta)$ is $dA = R_s^2 \sin \theta d\theta d\phi$, yielding $d\Omega_{\text{obs}} = (R_s^2/D^2) \sin \theta d\theta d\phi$. Again, we assume the observer resides on the $+z$ axis. Inserting $d\Omega_{\text{obs}}$ into Eq. 7.23 and integrating over the hemisphere of the source facing the observer, we have

$$f_\nu(t) = \frac{R_s^2}{D^2} \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) \sin \theta d\theta d\phi. \quad (7.26)$$

¹ The actual corrections must also include the sensitivity function of the telescope and instrument. See § 8.10.4.

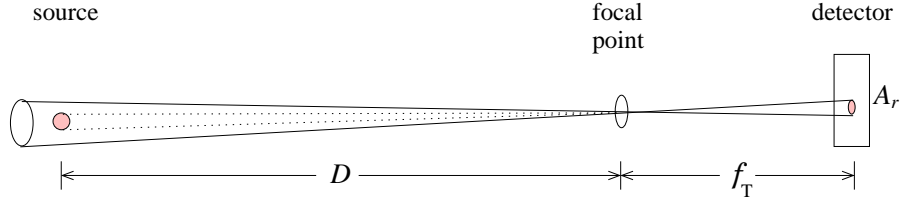


Figure 7.7 A source of radius R_s at distance D from the telescope focal point and a detector at a distance f_T (the effective focal length). This schematic illustrates the angular (solid angle) relationship between the area of the resolution element, A_r , on the detector and that of the source. For an unresolved source, relative to the focal point, the solid angle of the resolution element, $\Omega_r = A_r/f_T^2$, is greater than the solid angle of the source, $\Omega_s = \pi R_s^2/D^2$. The solid angle Ω_r is illustrated with the solid cone (which is also projected over the source) and Ω_s is illustrated with the dotted-line cone.

With the observer in the $\hat{\mathbf{k}}$ direction, we have $(\hat{\mathbf{s}} \cdot \hat{\mathbf{n}}) = (\hat{\mathbf{k}} \cdot \hat{\mathbf{r}}) = \cos \theta$, yielding,

$$f_\nu(t) = \frac{R_s^2}{D^2} \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{k}}) \cos \theta \sin \theta d\theta d\phi = \frac{R_s^2}{D^2} \mathcal{F}_\nu(t). \quad (7.27)$$

With no assumptions about the nature of the radiation field, we obtain that the observed flux is R_s^2/D^2 times the component of the object's surface flux intercepted by the observer, which is called the “astrophysical flux”, $\mathcal{F}_\nu(t)$. More conveniently, we can express the geometrically elegant result $f_\nu(t) = \Omega_s(\mathcal{F}_\nu/\pi)$.

If we assume the specific intensity is steady state and is emitted isotropically at all ϕ, θ locations across the surface of the source, then $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{k}}) = \mathcal{I}_\nu(R_s)$. In this case, the specific intensity can be factored out of the integral, yielding,

$$f_\nu = \frac{R_s^2}{D^2} \pi \mathcal{I}_\nu(R_s) = \Omega_s \mathcal{I}_\nu(R_s). \quad (7.28)$$

7.3 Radiative Transfer

Over a time interval from $t \rightarrow t + dt$, photons will propagate a distance $d\mathbf{s} = c dt \hat{\mathbf{s}}$ from location \mathbf{r} to $\mathbf{r} + d\mathbf{s}$. In a perfect vacuum or in the case that the radiation field and particle field are in steady-state thermal equilibrium, the incremental energy transported by the radiation field is unchanged along

the pathlength, i.e.,

$$\frac{\partial}{\partial s} [d\epsilon_\nu(\mathbf{r} + \mathbf{ds}, t + dt; \hat{\mathbf{s}}) - d\epsilon_\nu(\mathbf{r}, t; \hat{\mathbf{s}})] = 0. \quad (7.29)$$

Invoking Eq. 7.7, this condition can also be written

$$\frac{\partial}{\partial s} \left[\mathcal{I}_\nu(\mathbf{r} + \mathbf{ds}, t + dt; \hat{\mathbf{s}}) - \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt = 0. \quad (7.30)$$

Since at time t the specific intensity at $(\mathbf{r} + \mathbf{ds}, t + dt)$, was at location (\mathbf{r}, t) , we can write

$$\frac{\partial \mathcal{I}_\nu(\mathbf{r} + \mathbf{ds}, t + dt; \hat{\mathbf{s}})}{\partial s} = -\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t}. \quad (7.31)$$

Substituting into Eq. 7.34, we obtain

$$\left[\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial s} \right] (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt = 0. \quad (7.32)$$

The general expression for the spatial portion of the partial derivative is

$$\frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial s} = \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \quad (7.33)$$

yielding

$$\left[\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt = 0. \quad (7.34)$$

xs

In Cartesian coordinates, we have

$$\hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = (\hat{\mathbf{s}} \cdot \hat{\mathbf{i}}) \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial x} + (\hat{\mathbf{s}} \cdot \hat{\mathbf{j}}) \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial y} + (\hat{\mathbf{s}} \cdot \hat{\mathbf{k}}) \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial z}, \quad (7.35)$$

where $\hat{\mathbf{s}} \cdot \hat{\mathbf{i}} = \cos \alpha$, $\hat{\mathbf{s}} \cdot \hat{\mathbf{j}} = \cos \beta$, and $\hat{\mathbf{s}} \cdot \hat{\mathbf{k}} = \cos \gamma$ are the direction cosines (Eq. 7.3) of the propagation direction $\hat{\mathbf{s}}$.

The left hand side of Eq. 7.34 describes the transport of radiative energy in the $\hat{\mathbf{s}}$ direction through a projected volume element $(\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) ds$ per unit frequency per unit time per unit solid angle. It is called the equation of radiative transfer, or just the transfer equation. A schematic showing the geometric relationships between the terms in the transfer equation is illustrated in Figure 7.8.

As written in Eq. 7.34, the nulled right hand side of the transfer equation applies in the case that there is no temporal or spatial modification to the radiation field, i.e., the medium in which the radiation propagates is a vacuum or is in steady-state thermal equilibrium. Clearly, in a vacuum there is no particle field with which the radiation will interact and radiative energy

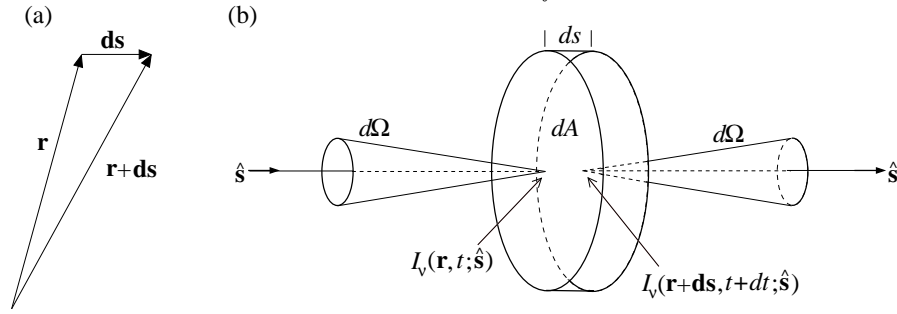


Figure 7.8 A schematic of the geometric configuration for the transfer equation. (a) In a time interval dt , photons travel a distance $ds = c dt$ from location \mathbf{r} to $\mathbf{r} + d\mathbf{s}$. (b) The specific intensity $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ propagating in the $\hat{\mathbf{s}}$ direction within solid angle $d\Omega$ travels a distance ds through a volume element of cross sectional area dA . In a medium in which absorption, scattering, and/or emission processes occur due to interactions with the particle field between the point \mathbf{r} , located at the volume element surface of entry, and the point $\mathbf{r} + d\mathbf{s}$, located at the surface of exit, the resulting modified specific intensity propagating into the same solid angle $d\Omega$ emerges as $\mathcal{I}_\nu(\mathbf{r} + d\mathbf{s}, t + dt; \hat{\mathbf{s}})$.

is conserved as the photons propagate. In a medium in thermal equilibrium, the radiation field and the particle field balance the exchange of energy between them; at all locations along the interval ds , there is no change in this balance, so the radiation field is unaltered as it propagates. If the medium is not in thermal equilibrium, the right hand side of the transfer equation is nonzero; there are temporal and spatial energy exchange between the radiation field and the particle field and this results in both a changing specific intensity in the radiation field and a changing energy within the particle field.

The microscopic physics of the interactions between the particle field and the radiation field can be extremely complex and involve detailed calculations. In the final analysis, along a given propagation direction $\hat{\mathbf{s}}$, these complex physical processes either reduce the energy transported by the radiation or they increase the energy transported. Reduction of the energy transport is either due to absorption of the radiation by the medium, or by scattering of the radiation out of the $\hat{\mathbf{s}}$ direction. An increase in the transported energy is due to emission of radiation from the medium into the $\hat{\mathbf{s}}$ direction.

In order to develop the complete formalism for the microscopic physics of absorption, scattering, and emission by the particle field, we first begin by defining macroscopic quantities to describe these processes. The macroscopic

quantities describe the sum total of the microscopic physics and they can be fully described and quantified in terms of the microscopic physics. Below we will introduce these quantities, which are called the extinction coefficient and the emission coefficient. For what follows, we will adopt the simplifying notation

$$\frac{d\mathcal{I}_\nu}{ds} \equiv \frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}), \quad (7.36)$$

for the differential terms appearing in Eq. 7.34, which allows the transfer equation for a vacuum and/or for steady-state thermal equilibrium to be written

$$\frac{d\mathcal{I}_\nu}{ds} (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt = 0. \quad (7.37)$$

7.3.1 The Extinction Coefficient and Optical Depth

The extinction coefficient, $\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, also called the opacity, is the macroscopic quantity describing the rate (slope of the gradient) at which radiative energy is absorbed and/or scattering out of the light path in the $\hat{\mathbf{s}}$ direction over the pathlength element ds ,

$$\frac{d\mathcal{I}_\nu}{ds} (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt = -\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}) d\Omega d\nu dt, \quad (7.38)$$

Thus, the extinction coefficient is defined such that a volume element of material of projected cross section $\hat{\mathbf{s}} \cdot \mathbf{d}\hat{\mathbf{A}}$ and length ds , through which radiation of specific intensity $\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ is propagating in the $\hat{\mathbf{s}}$ direction, will remove energy from the radiation field in the solid angle $d\Omega$ in the frequency range $\nu \rightarrow \nu + d\nu$ in time interval dt . The units of $\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ are cm^{-1} . Eq. 7.38 can be written in simplified form as

$$\frac{d\mathcal{I}_\nu}{ds} = -\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.39)$$

In order to make a distinction between absorption and scattering, the extinction coefficient is often written as the sum of an absorption coefficient, $K_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, and a scattering coefficient, $\Sigma_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$,

$$\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = K_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) + \Sigma_\nu(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.40)$$

The extinction coefficient can be interpreted as the inverse of the mean free path $\ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ of photons at location \mathbf{r} at time t propagating in the $\hat{\mathbf{s}}$ direction, i.e.,

$$\ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \chi_\nu^{-1}(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.41)$$

Given that a single absorption or scattering event is probabilistic in nature,

the mean free path is interpreted as a characteristic distance that photons of frequency ν propagate before being absorbed or scattered from the $\hat{\mathbf{s}}$ direction.

We now define the absorptivity along the pathlength as $d\tau_\nu = \chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) ds$. The integral of absorptivity over the pathlength is called the optical depth. Integrating from location \mathbf{r}_1 to location \mathbf{r}_2 along the $\hat{\mathbf{s}}$ direction, we have

$$\tau_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \int_{\mathbf{r}_1}^{\mathbf{r}_2} \chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) ds. \quad (7.42)$$

The geometry of Eq. 7.42 is illustrated in Figure 7.9(a). Note that the

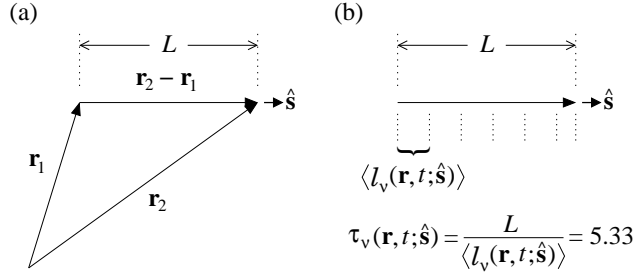


Figure 7.9 (a) A schematic for the integration to obtain the optical depth via Eq. 7.42, where L is the total pathlength of the integration. (b) The optical depth is the number of average mean free paths, $\langle \ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \rangle$, of the pathlength, L . In this example, $L = 5.33 \langle \ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \rangle$, indicating that $\tau_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = 5.33$.

integrand of Eq. 7.42 is equivalent to the ratio of the pathlength element to the mean free path, $ds/\ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$. The interpretation of the optical depth follows from the *average* mean free path over the pathlength $L = |\mathbf{r}_1 - \mathbf{r}_2|$, which is

$$\frac{1}{\langle \ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \rangle} = \frac{\int_{\mathbf{r}_1}^{\mathbf{r}_2} ds/\ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\int_{\mathbf{r}_1}^{\mathbf{r}_2} ds} = \frac{\tau_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{L}. \quad (7.43)$$

From Eq. 7.43, we have $\tau_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = L/\langle \ell_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \rangle$. Thus, as illustrated in Figure 7.9(b), the optical depth is a statistical quantity interpreted as the number of average mean free paths of the pathlength L over which photons of frequency $\nu \rightarrow \nu + d\nu$ have traveled. Note that the optical depth is unitless and is frequency dependent.

7.3.2 The Emission Coefficient

Radiative energy can be redirected into the $\hat{\mathbf{s}}$ direction via scattering by the particle field. It can also be introduced into the $\hat{\mathbf{s}}$ direction via emission from the particle field. The rate at which radiative energy is incremented over pathlength ds in the $\hat{\mathbf{s}}$ direction across area element $\hat{\mathbf{s}} \cdot d\mathbf{A}$ into the solid angle $d\Omega$ in the frequency range $\nu \rightarrow \nu + d\nu$ in time interval dt is quantified by the emission coefficient, $\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$, which is defined by the relation

$$\frac{d\mathcal{I}_\nu}{ds}(\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt = \xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) (\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt, \quad (7.44)$$

Eq. 7.44 can be written in simplified form as

$$\frac{d\mathcal{I}_\nu}{ds} = \xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.45)$$

The units of $\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})$ are $\text{erg s}^{-1} \text{ cm}^{-3} \text{ Hz}^{-1} \text{ str}^{-1}$.

7.3.3 The Transfer Equation and Source Function

In order to generalize Eq. 7.34 to account for extinction and emission processes that are not necessarily in steady-state thermal equilibrium, we equate the left hand side of Eq. 7.37 to the sum of Eqs. 7.38. and 7.44. We have

$$\frac{d\mathcal{I}_\nu}{ds}(\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt = \left[\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) - \chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] (\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt, \quad (7.46)$$

which written in full by invoking Eq. 7.36, yields transfer equation is

$$\left[\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] (\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt = \left[\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) - \chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] (\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt. \quad (7.47)$$

In a static medium, the geometry, frequency interval, and time interval are identical on both sides of the equation, the factors $(\hat{\mathbf{s}} \cdot d\mathbf{A}) d\Omega d\nu dt$ cancel, and we write,

$$\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) - \chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}). \quad (7.48)$$

It is common practice to divide Eq. 7.48, by the extinction coefficient, and write the transfer equation in the form

$$\chi_\nu^{-1}(\mathbf{r}, t; \hat{\mathbf{s}}) \left[\frac{1}{c} \frac{\partial \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\partial t} + \hat{\mathbf{s}} \cdot \nabla \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) \right] = S_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) - \mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}}), \quad (7.49)$$

where the ratio of the emission coefficient to the extinction coefficient is generally defined as the source function

$$S_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) = \frac{\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{\chi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})} = \frac{\xi_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}{K_\nu(\mathbf{r}, t; \hat{\mathbf{s}}) + \Sigma_\nu(\mathbf{r}, t; \hat{\mathbf{s}})}. \quad (7.50)$$

Under the very special condition of steady-state thermal equilibrium, in which the net energy exchange between absorption and emission processes and the gas thermal state is zero, the left hand side of Eq. 7.48 vanishes. If we further assume that the absorption processes dominate over scattering, i.e., $\Sigma_\nu(\mathbf{r}; \hat{\mathbf{s}}) \ll K_\nu(\mathbf{r}; \hat{\mathbf{s}})$, we obtain the relationship

$$\xi_\nu^\varepsilon(\mathbf{r}; \hat{\mathbf{s}}) = K_\nu^\varepsilon(\mathbf{r}; \hat{\mathbf{s}}) \mathcal{I}_\nu^\varepsilon(\mathbf{r}; \hat{\mathbf{s}}), \quad (7.51)$$

where we have employed the superscript ε to indicate the thermal equilibrium condition. Thus, in steady-state thermal equilibrium, the ratio of the emission coefficient to the absorption coefficient is the specific intensity. As shown in § 7.1.3, if the radiation field is isotropic then $J_\nu^\varepsilon(\mathbf{r}) = \mathcal{I}_\nu^\varepsilon(\mathbf{r})$, and from Eq. 7.51, we find

$$J_\nu^\varepsilon(\mathbf{r}) = \mathcal{I}_\nu^\varepsilon(\mathbf{r}) = \frac{\xi_\nu^\varepsilon(\mathbf{r})}{K_\nu^\varepsilon(\mathbf{r})}. \quad (7.52)$$

Finally, from Eq. 7.50, we have $S_\nu^\varepsilon(\mathbf{r}) = J_\nu^\varepsilon(\mathbf{r}) = \mathcal{I}_\nu^\varepsilon(\mathbf{r})$.

Under the above highly specific conditions, the source function is the Planck function [$\text{erg cm}^{-2} \text{ sec}^{-1} \text{ Hz}^{-1} \text{ rad}^{-2}$],

$$S_\nu^\varepsilon(\mathbf{r}, t) = B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{\exp(h\nu/kT) - 1}, \quad (7.53)$$

where k is the Boltzmann constant. Note then, that the spectral energy density of the radiation field in thermal equilibrium depends only on the equilibrium temperature. Integrating over all frequencies we obtain the flux per unit solid angle [$\text{erg cm}^{-2} \text{ sec}^{-1} \text{ rad}^{-2}$],

$$B(T) = \int_0^\infty B_\nu(T) d\nu = \frac{\sigma}{\pi} T^4 = \frac{ac}{4\pi} T^4 \quad (7.54)$$

where $\sigma = 2\pi^5 k / (15h^3 c^2)$, and where $a = 4\sigma/c$.

The Planck function is an ideal source function in that the physical conditions described above are idealized. However, the condition is well approximated in the interiors of stars where the optical depth is larger, i.e., the mean free paths of the photons are very short compared to and spatial variations in the thermal conditions of the gas. In the low density intergalactic, circumgalactic, and interstellar medium, the radiation field is rarely in thermal equilibrium with the gas because the radiation field originates outside

the gaseous medium itself. That is not to imply that the gas itself is rarely in an equilibrium state; it is just that the radiation field and the gas are not in thermal equilibrium. In Chapter 12, we describe the physics of low density gas.

7.3.4 Solution in One Dimension

For spectroscopy of cosmological sources, normally the only known geometric information is the line of sight direction from the source to the observer, \hat{s} . The geometry of the source is usually assumed to be spherical. In the case that the light path passes through an intervening gaseous medium, there usually is no *a priori* knowledge of the geometry of this medium. In the case of an intergalactic cloud, or a galactic halo or interstellar cloud, it is common practice to solve the transfer equation assuming the cloud has plane parallel geometry. Thus, Eq. 7.49 reduces to a single dimensional.

We will assume the line of sight from the background emitting source to the observer is in the \hat{k} direction relative to the center of the emitting source (basically, the same geometry we adopted to compute the observed flux in § 7.2.1). As illustrated in Figure 7.10, place an intervening cloud medium at arbitrary distance from the observer, and assume it has a plane parallel geometry and thickness L . Let the depth into the cloud be measured by the coordinate s in the Cartesian system, where $s_1 = 0$ is the cloud face and $s_2 = L$ is the cloud backside. Per this definition, the cloud face is the plane on which the background source radiation is incident on the cloud, and the backside is the plane closest to the observer.

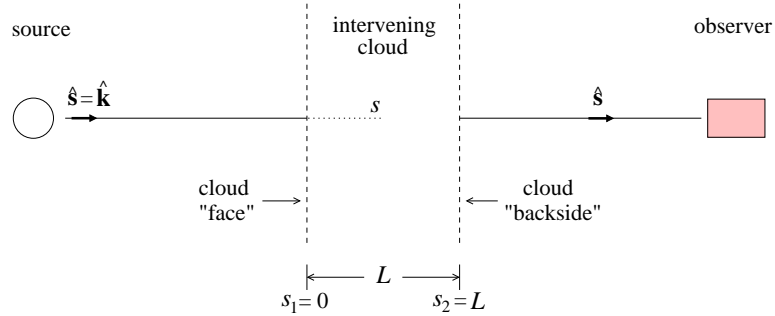


Figure 7.10 A schematic of the one-dimensional geometry employed for studies of the radiative transfer of material intervening between the observer and the background emitting source.

For purposes of illustration, assume a steady-state condition in the cloud

medium so that $c^{-1}\partial\mathcal{I}_\nu(\mathbf{r}, t; \hat{\mathbf{s}})/\partial t = 0$ in the transfer equation. Because the photon propagation direction (line of sight being measured) is in the $\hat{\mathbf{s}}$ direction, such that $\hat{\mathbf{s}} = \hat{\mathbf{k}}$, their inner product is unity. Thus, the left hand side of Eq. 7.49 simplifies to $\chi_\nu^{-1}(s)(d\mathcal{I}_\nu(s)/ds)$, and the transfer equation is now written

$$\frac{1}{\chi_\nu(s)} \frac{d\mathcal{I}_\nu(s)}{ds} = S_\nu(s) - \mathcal{I}_\nu(s). \quad (7.55)$$

This is actually an infinite number of equations, since there is a single equation for each frequency interval $\nu \rightarrow \nu + d\nu$. All of the equations must be solved simultaneously. As such, the solution for $\mathcal{I}_\nu(s)$ at all frequencies is not possible unless all of the microphysics of emission and absorption/scattering in the cloud are known as a function of location s so that $\chi_\nu(s)$ and $S_\nu(s)$ can be specified. However, both these quantities will also depend on $\mathcal{I}_\nu(s)$, so an iterative method of solution is required!

We can, however, solve for the general form of the solution of Eq. 7.55 in terms of the optical depth, $\tau_\nu(s)$. Invoking the definition of the absorptivity, we have $d\tau_\nu(s) = \chi_\nu(s) ds$. From Eq. 7.42, the optical depth at physical depth s is then the integral from $s_1 = 0$ to s (with $s_1 \leq s \leq s_2$),

$$\tau_\nu(s) = \int_0^s \chi_\nu(s') ds'. \quad (7.56)$$

Applying $d\tau_\nu(s) = \chi_\nu(s) ds$, we rewrite the transfer equation as a function of the optical depth as the independent variable,

$$\frac{d\mathcal{I}_\nu(\tau_\nu)}{d\tau_\nu} = S_\nu(\tau_\nu) - \mathcal{I}_\nu(\tau_\nu). \quad (7.57)$$

where the s dependence of the optical depth is implied.

The form of the transfer equation written as Eq. 7.57 is now such that we can apply the standard integrating factor technique and assume a solution of the form,

$$\mathcal{I}_\nu(\tau_\nu) = H_\nu(\tau_\nu) \cdot \exp\{h\tau_\nu\}. \quad (7.58)$$

where $H_\nu(\tau_\nu)$ and h are a yet to be determined function and constant, respectively. Differentiating, we have

$$\frac{d\mathcal{I}_\nu(\tau_\nu)}{d\tau_\nu} = \frac{dH_\nu(\tau_\nu)}{d\tau_\nu} \exp\{h\tau_\nu\} + h H_\nu(\tau_\nu) \cdot \exp\{h\tau_\nu\}. \quad (7.59)$$

Equating the terms in Eqs. 7.57 and 7.59, we have the relations

$$S_\nu(\tau_\nu) = \frac{dH_\nu(\tau_\nu)}{d\tau_\nu} \exp \{h \tau_\nu\} \quad (7.60)$$

$$-\mathcal{I}_\nu(\tau_\nu) = h H_\nu(\tau_\nu) \cdot \exp \{h \tau_\nu\} .$$

From Eq. 7.58, the latter relation yields $h = -1$, which after substitution into the former relation for $S_\nu(\tau_\nu)$, we have $dH_\nu(\tau_\nu)/d\tau_\nu = S_\nu(\tau_\nu) \exp \{\tau_\nu\}$. Integrating in order to obtain the function $H_\nu(\tau_\nu)$ yields

$$H_\nu(\tau_\nu) = C + \int_0^{\tau_\nu} S_\nu(t_\nu) \cdot \exp \{t_\nu\} dt_\nu . \quad (7.61)$$

Substituting Eq. 7.61 into Eq. 7.58 gives

$$\mathcal{I}_\nu(\tau_\nu) = C \exp \{-\tau_\nu\} + \exp \{-\tau_\nu\} \int_0^{\tau_\nu} S_\nu(t_\nu) \exp \{t_\nu\} dt_\nu . \quad (7.62)$$

To obtain the constant C , we evaluate Eq. 7.62 at the cloud face where $\tau_\nu(0) = 0$. Assuming that the specific intensity incident upon the cloud face from the background emitting source is $\mathcal{I}_\nu(0)$, we obtain $C = \mathcal{I}_\nu(0)$. Following substitution, and rearranging the second term, the solution to the one-dimensional transfer equation is

$$\mathcal{I}_\nu(\tau_\nu) = \mathcal{I}_\nu(0) \exp \{-\tau_\nu\} + \int_0^{\tau_\nu} S_\nu(t_\nu) \exp \{-(\tau_\nu - t_\nu)\} dt_\nu . \quad (7.63)$$

The term $\mathcal{I}_\nu(0) \exp \{-\tau_\nu\}$ represents the absorption of the specific intensity from the background emitting source that is incident on the cloud face. Recall that $\tau_\nu(s)$ is a function of the physical depth into the cloud as given by Eq. 7.56. Knowing the relationship between optical depth and physical depth requires knowledge of the extinction coefficient as a function of physical depth, $\chi_\nu(s)$. This requires details of the physical state of the particle field in the cloud medium.

The integral term in Eq. 7.63 represents the sum total of the increment in $\mathcal{I}_\nu(\tau_\nu)$ due to emission at physical depth s corresponding to $\tau_\nu(s)$. The exponential factor in the integral accounts for absorption of the emitted radiation [the source function $S_\nu(\tau_\nu)$] introduced into the propagation direction from all points in the cloud between s and the cloud face. Evaluation of the integral term in Eq. 7.63 requires knowledge of the source function as a function of physical depth, which requires estimates of both the emission and extinction coefficients, i.e. $S_\nu(\tau_\nu) = \xi_\nu(\tau_\nu)/\chi_\nu(\tau_\nu)$. Again, this requires details of the physical state of the particle field in the cloud medium.

7.4 Absorption Lines and Optical Depth

For clouds in the intergalactic medium and in the circumgalactic medium, the hydrogen density typically ranges from $10^{-6} \leq n_{\text{H}} \leq 10^{-1} \text{ cm}^{-3}$, which is too low for any significant emission into the line of sight. Thus, these clouds classify as pure absorption clouds (we will adopt the term “absorbing clouds”). In this section, we will briefly describe the observed consequences of radiative transfer through low density absorbing clouds.

7.4.1 Transforming Energy Densities

To this stage in our discussion, which has been primarily theoretical in nature, we have described the energy density of the radiation field in units per frequency interval, i.e., \mathcal{I}_ν [$\text{erg cm}^{-2} \text{ sec}^{-1} \text{ Hz}^{-1} \text{ rad}^{-2}$]. In the world of the observer, spectra are almost exclusively presented and analyzed in units per wavelength interval, i.e., \mathcal{I}_λ [$\text{erg cm}^{-2} \text{ sec}^{-1} \text{ \AA}^{-1} \text{ rad}^{-2}$].

Conservation of energy density requires $\mathcal{I}_\lambda d\lambda = \mathcal{I}_\nu d\nu$; the energy density per unit wavelength interval $\lambda \rightarrow \lambda + d\lambda$ equals the energy density per frequency interval $\nu \rightarrow \nu + d\nu$. We thus have $\mathcal{I}_\lambda = \mathcal{I}_\nu |d\nu/d\lambda|$. From the relationship $c = \nu\lambda$, we obtain

$$\mathcal{I}_\lambda = \frac{c}{\lambda^2} \mathcal{I}_\nu \quad (7.64)$$

Note that Eq. 7.64 applies to all energy densities, including, for example, J_ν , S_ν , and f_ν . Within the transfer equation itself, the emission coefficient per unit wavelength is also related to the per frequency coefficient via $\xi_\lambda = (c/\lambda^2)\xi_\nu$. Note that, since the absorption coefficient, χ_ν is defined in units cm^{-1} , we have $\chi_\lambda = \chi_\nu$. Also note that the absorptivity, and therefore optical depth, is invariant, i.e., $d\tau_\lambda = \chi_\lambda ds = \chi_\nu ds = d\tau_\nu$.

7.4.2 Pure Absorption

In the case of pure absorption, $S_\lambda(\tau_\lambda) = 0$ everywhere. and Eq. 7.63 simplifies to the well-known result

$$\mathcal{I}_\lambda(\tau_\lambda) = \mathcal{I}_\lambda(0) \exp \{-\tau_\lambda\} . \quad (7.65)$$

From the observer point of view, the light from source has traveled through the entire cloud along the path length $s_1 = 0$ to $s_2 = L$, where L is the cloud depth along the line of sight. So, in practice, how does one apply and/or interpret Eq. 7.65? To simplify the following illustration, we will invoke the

path length averaged extinction coefficient, $\bar{\chi}_\lambda$, which theoretically represents

$$\bar{\chi}_\lambda = \frac{\int_0^L \chi_\lambda(s) ds}{\int_0^L ds} = \frac{1}{L} \int_0^L \chi_\lambda(s) ds. \quad (7.66)$$

Note that the frequency dependence of the extinction coefficient is preserved. Adopting $\tau_\lambda(s) = \bar{\chi}_\lambda s$, we write Eq. 7.65 as a function of physical depth,

$$\mathcal{I}_\lambda(s) = \mathcal{I}_\lambda(0) \exp \{-\bar{\chi}_\lambda s\}. \quad (7.67)$$

Eq. 7.67 shows that when $s = \bar{\chi}_\lambda^{-1} = \langle \ell_\lambda \rangle$, photons in the frequency range $\lambda \rightarrow \lambda + d\lambda$ have penetrated one complete mean free path into the cloud and $\mathcal{I}_\lambda(0)$ has suffered extinction by the factor $e^{-1} = 0.367$. For every multiple additional mean free path the remaining unabsorbed photons penetrate into the cloud, the extinction is an additional factor of 0.367. For the full path length through the absorbing cloud, $s = L$, the total “mean” optical depth is $\bar{\tau}_\lambda = \bar{\chi}_\lambda L$. This is the optical depth measured by the observer, for which the extinction in each frequency interval is

$$\mathcal{I}_\lambda(L) = \mathcal{I}_\lambda(0) \exp \{-\bar{\tau}_\lambda\}. \quad (7.68)$$

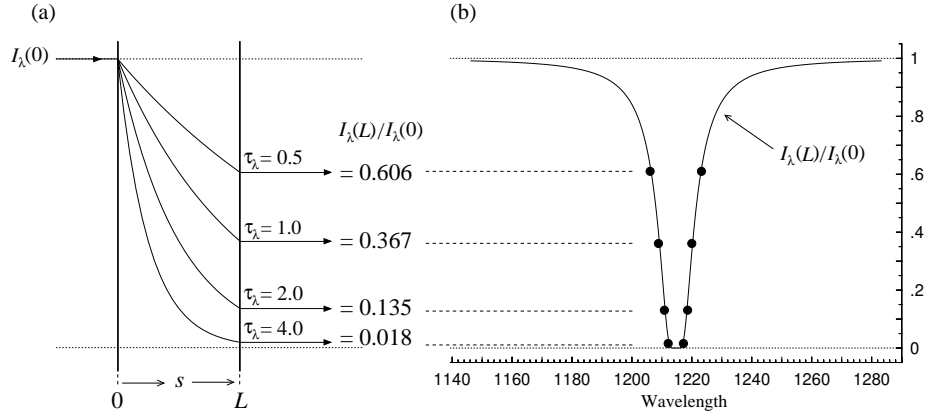


Figure 7.11 (a) A schematic of radiative transfer through a pure absorbing cloud, i.e., $S_\lambda(\tau_\lambda) = 0$, of total pathlength L , for which the total optical depth is $\bar{\tau}_\lambda = 0.5, 1.0, 2.0$, and 4.0 , depending upon frequency or wavelength. The incident specific intensity $\mathcal{I}_\lambda(0)$ (propagating from left to right) suffers extinction through the cloud according to Eq. 7.67. (b) The four example $\bar{\tau}_\lambda$ correspond to opacities at different frequencies. On the resulting absorption profile in the spectrum, $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0)$, the points superimposed on the profile correspond to the four optical depths.

In Figure 7.11(a), the behavior of Eq. 7.67 is illustrated as a function of path length for four different $\bar{\chi}_\lambda$ in an absorbing cloud through which the total pathlength is L . The different extinction coefficients (at different frequencies) result in different extinction rates and therefore in different total optical depths, $\bar{\tau}_\lambda$, at each frequency. The four cases illustrated are $\bar{\tau}_\lambda = 0.5, 1.0, 2.0$, and 4.0 . The normalized specific intensity of the emerging radiation, $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0)$, are given for the four example total optical depths.

In Figure 7.11(b), an idealized normalized absorption profile $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0)$, is shown as a function of wavelength for the Ly α $\lambda 1215.67$ transition of neutral hydrogen. For this transition, the extinction coefficient is narrow in frequency/wavelength space, symmetric, and centered on the wavelength 1215.6701 \AA . This results in an absorption profile for which the optical depth varies with wavelength. In the absence of instrumental effects, the observed total optical depth at each wavelength can be directly determined by inverting Eq. 7.68,

$$\bar{\tau}_\lambda = -\ln \left(\frac{\mathcal{I}_\lambda(L)}{\mathcal{I}_\lambda(0)} \right), \quad (7.69)$$

where $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0)$ is taken from the observed profile.

The measured normalized intensity and wavelength values along the profile that correspond to the four example total optical depths $\bar{\tau}_\lambda = 0.5, 1.0, 2.0$, and 4.0 , are represented by points superimposed on the profile. Photons with wavelengths near the profile center suffer a higher optical depth and are more readily absorbed. For photons having wavelengths where $\bar{\tau}_\lambda = 1$, the mean free path through the absorbing cloud is equal to the cloud depth L . In general, across the profile at each wavelength interval $\lambda \rightarrow \lambda + d\lambda$, the average mean free path is $\langle \ell_\lambda \rangle = \bar{\chi}_\lambda^{-1} = L/\bar{\tau}_\lambda$.

Clearly L must be known to measure the mean free path for each wavelength. This requires modeling of the physical state of the absorbing gas cloud (see Chapter 12). Nonetheless, from the example illustrated in Figure 7.11, it becomes clear that a great deal about the frequency and/or wavelength dependence of the behavior of the radiative transfer (and thus interaction between the radiation field and particle field in the context of the absorbing cloud line-of-sight physical depth) can be gleaned directly from the observed absorption profile.

Often, the two terms “optically thick” and “optically thin” are employed to describe the absorbing cloud based upon the observed absorption profile.

The term optically thick applies when the mean free path of the photons are smaller than the cloud line of sight physical depth. i.e. $\langle \ell_\lambda \rangle < L$. This corresponds to the wavelength regions of the absorption profile where $\bar{\tau}_\lambda > 1$

and $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0) < 0.367$. The term of optically thin thus applies across the wavelength regions of the absorption profile where we observe $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0) > 0.367$, corresponding to $\bar{\tau}_\lambda < 1$ and $\langle \ell_\lambda \rangle > L$.

For the example Ly α profile presented in in Figure 7.11(b), the profile is optically thin in the approximated wavelength region 1208–1222 Å. In practice, if any wavelength region of the observed absorption profile has $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0) < 0.367$ the absorbing cloud is referred to as being optically thick. Conversely, if $\mathcal{I}_\lambda(L)/\mathcal{I}_\lambda(0) > 0.367$ at all wavelengths across the observed absorption profile, the absorption is referred to as optically thin.

Though this example cloud is optically thick to absorption by neutral hydrogen, it may not be optically thick to absorption by other elements, or even other transitions from neutral hydrogen. Thus, when labeling an absorbing cloud as optically thick, one must keep in mind that this label applies for a given transition, for example, “this cloud is optically thick to neutral hydrogen Ly α absorption, but is optically thin to neutral sodium λ 5891 absorption .”

7.4.3 The Astronomical Absorption Spectrum

In § 7.4, we showed that the outcome of the radiative transfer through an intervening pure absorbing cloud reduces the observed specific intensity as a function of wavelength according to

$$\mathcal{I}_\lambda = \mathcal{I}_\lambda(0) \exp \{ -\bar{\tau}_\lambda \} , \quad (7.70)$$

where $\mathcal{I}_\lambda(0)$ is the specific intensity emitted from the background source, and where $\bar{\tau}_\lambda$ is the line-of-sight averaged total optical depth

$$\bar{\tau}_\lambda = \int_0^L \bar{\chi}_\lambda ds = \bar{\chi}_\lambda L , \quad (7.71)$$

where L is the cloud line-of-sight physical extent.

In the absence of absorption, we determined (see § 7.2.2) that the observed flux from an unresolved source of radius R_s that is at a distance D and has specific intensity distribution $\mathcal{I}_\lambda(R_s)$, i.e., assuming isotropic emission from the source surface, is

$$f_\lambda = \frac{R_s^2}{D^2} \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_\lambda(R_s) \cos \theta \sin \theta d\theta d\phi . \quad (7.72)$$

Substituting Eq. 7.70, with $\mathcal{I}_\lambda(0) = \mathcal{I}_\lambda(R_s)$, we obtain

$$f_\lambda = \frac{R_s^2}{D^2} \int_0^{2\pi} \int_0^{\pi/2} \mathcal{I}_\lambda(R_s) \exp \{ -\bar{\tau}_\lambda \} \cos \theta \sin \theta d\theta d\phi , \quad (7.73)$$

when the flux from the background source is occulted by an intervening absorption cloud with average total optical depth $\bar{\tau}_\lambda$. In terms of the solid angle and astrophysical flux of the observed source, we have

$$f_\lambda = \Omega_s \frac{\mathcal{F}_\lambda}{\pi} \exp \{-\bar{\tau}_\lambda\} . \quad (7.74)$$

Eq. 7.74 represents the commonly adopted expression for the measured observed flux as a function of wavelength when absorption is present.

Important assumptions were invoked in the derivation of Eq. 7.74. These include (1) a spherical and isotropically steady-state emitting source, (2) a line-of-sight averaged optical depth through the absorbing cloud, and (3) that the background source is completely occulted by the intervening absorbing cloud. In the most general case, none of these assumptions may apply. Below, we will briefly illustrate the scenario in which the absorbing cloud optical depth varies across radiating area elements on the surface of the emitter. We then derive an expression for the observed flux under the condition that the absorbing cloud does not occult all radiating area elements on the surface of the emitter (called “partial covering”).

7.4.4 Cosmological Intervening Absorption

We define intervening absorption as the geometric relationship in which the emitter is at galactic scale distances (10s of kpc to a few Mpc) or cosmological distances (10s of Mpc and up) from the absorbing cloud, which itself is at galactic or cosmological distances from the observer. We refer to the emitter as the background source.

The schematic (not to scale!) of the geometric configuration of an absorbing cloud intervening to an emitting background source is presented in Figure 7.12. The overall geometry follows that employed in § 7.2.2; the observer is located a distance D from the source on the $+z$ axis and the background source is unresolved from the point of view of the observer. We will assume that the background source is a steady-state spherical emitter, but not necessarily an isotropic emitter. As illustrated, an area element located at arbitrary azimuthal and polar angular positions with respect to the geometric center of the source emits a “beam” of specific intensity $\mathcal{I}_\lambda(R_s, \phi, \theta)$ that, from the perspective of the observer, probes the intervening absorbing cloud at coordinate $(R_s \sin \theta, \phi)$ on the sky-projection of the absorbing cloud.

The absorbing cloud illustrated in Figure 7.12 has uniform thickness. However, in reality, absorbing clouds have unknown and amorphous geom-

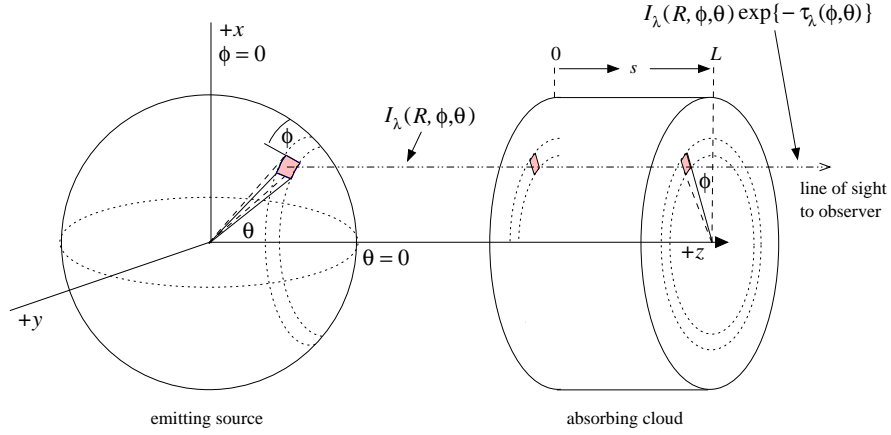


Figure 7.12 Non-isotropic specific intensity $I_\lambda(R_S, \phi, \theta; \hat{\mathbf{s}})$ is emitted from an area element on the spherical background source. The beam passes through an intervening absorbing cloud at a projected location $(R \sin \theta, \phi)$, where the coordinates correspond to the location of the emitting area element on the source.

etry. As such the line of sight path length for the example beam would be $L(R \sin \theta, \phi)$. Furthermore, the cloud internal physical properties (density, temperature, ionization, etc.) could vary spatially, and therefore the absorption coefficient would depend upon the projected location and would vary along the line of sight, $\chi_\lambda(R \sin \theta, \phi, s)$. As such, the optical depth is unique to each beam; invoking Eq. 7.42, the total optical depth for the beam emitted from the source location (R, ϕ, θ) is

$$\tau_\lambda(R \sin \theta, \phi) = \int_0^{L(R \sin \theta, \phi)} \chi_\lambda(R \sin \theta, \phi, s) ds. \quad (7.75)$$

The average optical depth employed in Eq. 7.74 is then

$$\bar{\tau}_\lambda = \int_0^{2\pi} \int_0^{\pi/2} \tau_\lambda(R \sin \theta, \phi) \sin \theta d\theta d\phi. \quad (7.76)$$

In practice, for an unresolved source, there is no observational data that can provide a spatial mapping of the source geometry nor of the absorbing cloud opacity, and therefore optical depth. When both or either the background source or intervening absorbing cloud are unresolved, a full evaluation of Eq. 7.75 is an intractable problem and one is able to measure only the averaged optical depth.

The adopted assumptions may appear obvious and even somewhat trivial; however, it is worth explicitly formulating the problem of calculating the

absorption by an intervening cloud for an unresolved astronomical source. The point here is to emphasize that virtually all published extragalactic work necessarily rests upon these simplifying assumptions. From this point forward, we will drop the notation $\bar{\tau}_\lambda$ for the line-of-sight averaged optical depth and simply write τ_λ .

7.4.5 Associated Absorption and Partial Covering

The above treatment was based upon the assumption that the background source is fully occulted by the intervening absorbing cloud. Whereas this condition likely holds in virtually every case when the background source is at cosmological distance from the intervening absorbing cloud, this condition may not always hold true when the distance between the source and the absorbing cloud is on galactic scales. An example of absorbing clouds at galactic distances would be gas comprising the surrounding medium of a luminous quasar or AGN (e.g., Hamann & Ferland, 1999). Such gas would be physically associated with the emitting source in that it would either be interacting with the quasar or AGN, or it would reside within the influenced of the quasar or AGN environment (i.e., dark matter halo). For these reasons, we call the absorption lines from this type of gas environments “associated” absorption. Historically, associated absorption is also referred to as “intrinsic” absorption.

For associated absorbing gas clouds, even if the source is unresolved from the perspective of the observer, it is possible that not all emitting area elements on the source are occulted by the absorbing cloud. This is known as partial covering. When partial covering is present, some beams of specific intensity \mathcal{I}_λ pass through the absorbing cloud and are modulated via the transfer equation, whereas the remainder of the beams propagate directly to the observer without modification.

A simplistic example of partial covering is schematically illustrated in Figures 7.13(a, face-on view) and 7.13(b, side view). In this scenario, the observed solid angle of the intervening cloud is smaller than the observed solid angle of the source (for more realistic geometric configurations of partial covering, we refer the reader to Hamann & Sabra (2004)). When partial covering is present, the behavior of the associated absorption lines can differ dramatically from intervening absorption lines or associated absorption lines without partial covering. Here, we derive the observed flux in the case of partial covering.

For heuristic purposes, we assume (1) from the observer perspective, the intervening cloud is directly aligned with the center of a spherical source

(a) observer view

(b) side view

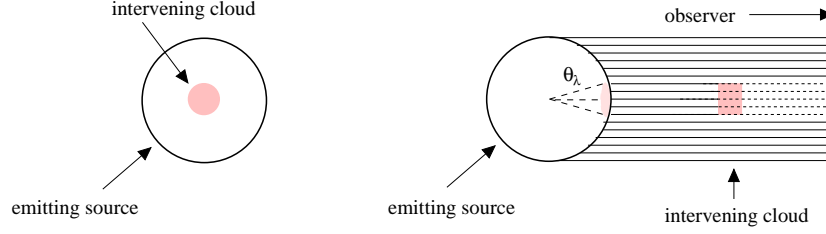


Figure 7.13 A schematic of partial covering scenarios. Solid lines from the isotropic source are specific intensity beams that are unobstructed by the intervening absorber, whereas dotted lines are beams for which some level of absorption has occurred. A highly idealized case in which the observed solid angle of the absorbing cloud (shaded) is smaller than the observed solid angle of the source. In this scenario, the intervening cloud occults a area on the source (lightly shaded) defined by a cone with angle θ_λ . (a) Face-on view. (b) Side view.

and that the projected cross section of the intervening cloud is circular, (2) the emitted specific intensity is isotropic, so that $\mathcal{I}_\lambda(R_s, \phi, \theta) = \mathcal{I}_\lambda(R_s)$, and (3) the average optical depth through the absorbing cloud provides a good approximation of the sightline to sightline optical depth variations across the cloud. As shown in Figure 7.13(b), the solid angle of the region of the source that is occulted subtends an opening angle θ_λ , where the subscript λ generalizes the possibility that the projected cross section can vary with wavelength.

From the observer point of view, the $\hat{\mathbf{s}} = \hat{\mathbf{k}}$ directed beams originating from the surface of the source within the range of polar angles $0 \leq \theta \leq \theta_\lambda$ will be occulted by the cloud, and those in the annulus $\theta_\lambda \leq \theta \leq \pi/2$ will be unocculted. From the fundamental theorem of calculus, Eq. 7.73 can be expressed as the sum

$$f_\lambda = f_\lambda^u + f_\lambda^o, \quad (7.77)$$

where f_λ^u is the observed flux from the unocculted sight lines and f_λ^o is the observed flux from the occulted sight lines. This latter term is often referred to as the transmitted flux,

$$\begin{aligned} f_\lambda^o &= \frac{R_s^2}{D^2} \int_0^{2\pi} \int_0^{\theta_\lambda} \mathcal{I}_\lambda(R_s) \exp\{-\tau_\lambda\} \cos \theta \sin \theta \, d\theta \, d\phi \\ &= \sin^2 \theta_\lambda \frac{R_s^2}{D^2} \mathcal{F}_\lambda \exp\{-\tau_\lambda\}, \end{aligned} \quad (7.78)$$

where \mathcal{F}_λ is the astrophysical flux of the source. The unocculted flux is given by

$$\begin{aligned} f_\lambda^u &= \frac{R_s^2}{D^2} \int_0^{2\pi} \int_{\theta_\lambda}^{\pi/2} \mathcal{I}_\lambda(R_s) \cos \theta \sin \theta \, d\theta \, d\phi \\ &= [1 - \sin^2 \theta_\lambda] \frac{R_s^2}{D^2} \mathcal{F}_\lambda. \end{aligned} \quad (7.79)$$

Substituting f_λ^u and f_λ^o into Eq. 7.77 yields

$$f_\lambda = [1 - \sin^2 \theta_\lambda] \frac{R_s^2}{D^2} \mathcal{F}_\lambda + \sin^2 \theta_\lambda \frac{R_s^2}{D^2} \mathcal{F}_\lambda \exp \{-\tau_\lambda\}. \quad (7.80)$$

Note that if the geometric factor is $\sin^2 \theta_\lambda = 0$, then the observed flux is unattenuated, and if $\sin^2 \theta_\lambda = 1$, then the observed flux is for a fully occulted source. Common practice is to define this geometric factor as a wavelength dependent covering fraction, $C_f(\lambda)$. Clearly, the functional form of the covering fraction depends upon the assumed geometric configuration; for the above idealized scenario $C_f(\lambda) = \sin^2 \theta_\lambda$. Defining $f_\lambda^o = (R_s^2/D^2) \mathcal{F}_\lambda$ as the observed flux of an unocculted source yields the functional form commonly quoted in the scientific literature,

$$R_\lambda = \frac{f_\lambda}{f_\lambda^o} = [1 - C_f(\lambda)] + C_f(\lambda) \exp \{-\tau_\lambda\}, \quad (7.81)$$

where the first term on the right hand side represents flux from unocculted sight lines and the second term represents the flux transmitted through the cloud.

We define the effective optical depth, $\tau_{\text{eff}}(\lambda)$, as that which is measured directly from the flux ratio in a spectrum, $R_\lambda = \exp\{-\tau_{\text{eff}}(\lambda)\}$. The relationship between a selected range of the measured $\tau_{\text{eff}}(\lambda) = \ln(1/R_\lambda)$ and the true optical depth τ_λ at the absorption line center is illustrated in Figure 7.14(a) for a range of $C_f(\lambda)$. In Figure 7.14(b), an example absorption profile from a cloud with $\tau_\lambda \gg 10$ in the absorption profile center is illustrated for $C_f(\lambda) = 0.50, 0.75, 0.95$, and 1.00 (full occultation). Note that even if τ_λ of a partially occulting cloud is very large, a non-negligible amount of flux can be transmitted to the observer for non-zero $C_f(\lambda)$. Indeed, if the line center is well resolved in the spectrum (as it is in Figure 7.14(b)), and one has reason to believe $\tau_\lambda > 10$ in the line center, then the term $C_f(\lambda) \exp \{-\tau_\lambda\} \simeq 0$ in Eq. 7.81, and we can estimate the covering fraction directly from $R_\lambda = 1 - C_f(\lambda)$. We discuss partial coverign in more detail in § 9.2.8.

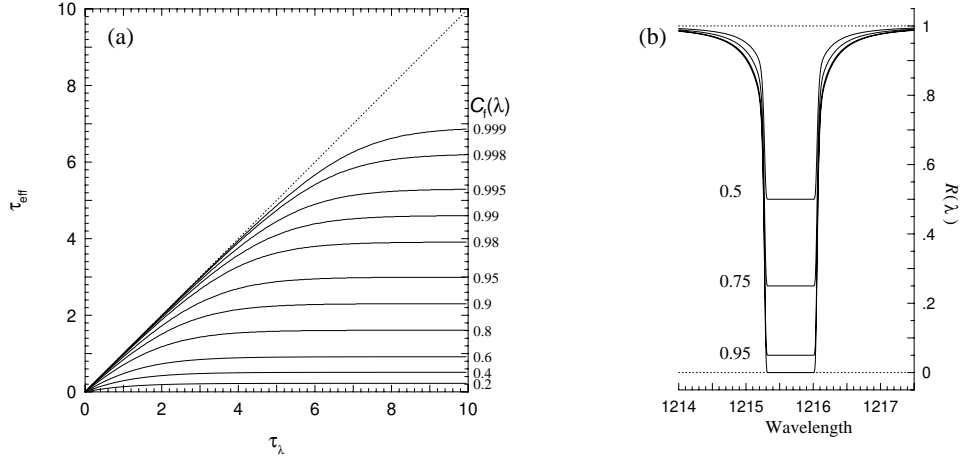


Figure 7.14 (a) The effective, or measured, optical depth τ_{eff} , obtained by inverting the flux ratio R_λ versus the mean optical depth in the cloud, τ_λ for various covering fractions, $C_f(\lambda)$. (b) An example “black-bottomed” absorption profile with $\tau_\lambda \gg 10$ in the profile center for $C_f(\lambda) = 0.50, 0.75, 0.95$, and 1.00 .

7.4.6 Column Density

As one of the goals of cosmological spectroscopy is to measure the physical properties of the gas (i.e., the kinematics, ionization conditions, chemical composition, temperature, and mass density), we must consider the link between what is measurable directly from the spectra, and what we aim to scientifically learn from the spectra.

What we can measure is the observed flux, f_λ . And, from radiative transfer we know that,

$$f_\lambda = f_\lambda^0 \exp \{-\tau_\lambda\} , \quad (7.82)$$

where the optical depth is

$$\tau_\lambda = \int_0^L \chi_\lambda(s) ds . \quad (7.83)$$

where $\chi_\lambda(s)$ is the extinction coefficient (Eq. 7.40). In Figure 7.15, we illustrate an absorbing cloud of gas being probed by the line of sight of length L to a background luminous source. What we cannot know is the path length L through the cloud. Neither can we know the the variation of $\chi_\lambda(s)$ along the path length.

For a given absorbing gas cloud, spectral absorption features arise from various transitions over a range of different ions stages from dozens of atomic

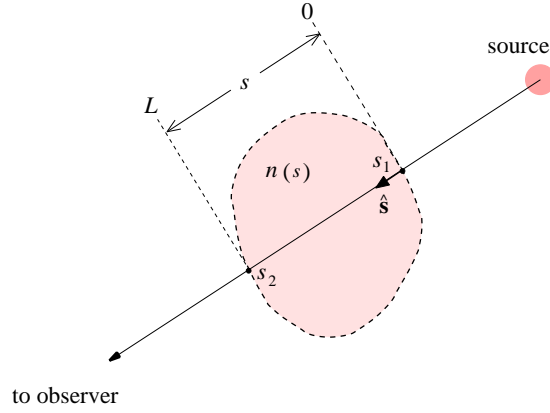


Figure 7.15 A schematic of the geometry of the line of sight intersection of an absorbing cloud. The beam passes through the cloud with unknown spatial density distribution and unknown path length through the cloud, L . The unknown density profile through the cloud along the line of sight is denoted $n(s)$.

species. However, in a finite spectral wavelength range over which any individual absorption feature is measured, the feature likely arises from a single transition from a given ion/atom. For example, a bound-bound H I Ly α transition with rest-frame transition wavelength 1215.67 Å occurring within a cloud at $z = 3$ would be observed at 4862.68 Å; if the absorption line is not contaminated, or blended, with other absorption, then $\chi_\lambda(s)$ can be quantified in terms of the absorption opacity of the Ly α transition. As such, we can simplify the extinction coefficient in terms of the absorption coefficient, $\chi_\lambda(s) = K_\lambda(s)$, by assuming continuum scattering along the path length through the cloud is negligible in the narrow spectral region across the absorption line.

In § 2.6.2 and § 3.3.2, we derived the absorption cross section for a bound-bound transition, $\sigma(\lambda)$, which is a function of the transition wavelength, oscillator strength, and damping constant (for Ly α , see Eq. 2.83 and Figure 2.7). The units of $\sigma(\lambda)$ are $\text{cm}^2 \text{ Hz}^{-1} \text{ atom}^{-1}$. To obtain the absorption coefficient, we multiply the cross section of the bound-bound transition to the number density of the ion/atom giving rise to the absorption line,

$$K_\lambda(s) = n(s)\sigma(\lambda). \quad (7.84)$$

where we emphasize that the number density, cross section, and optical depth are specific to the ion and transition corresponding to the absorption

feature and do not reflect any additional quantities in the absorbing cloud, as we have isolated the absorption line in the spectrum to have arisen from a single transition from a specific ion. Though the cross section is known from atomic physics, the number density as a function of location along the line of sight is impossible to determine.

Substituting the line absorption coefficient given by Eq. 7.84 into Eq. 7.83, we obtain

$$\tau_\lambda = \int_0^L n(s)\sigma(\lambda) ds = \sigma(\lambda) \int_0^L n(s) ds = N\sigma(\lambda), \quad (7.85)$$

where the atomic transition cross section is independent of location along the pathlength, and where we define the column density,

$$N = \int_0^L n(s) ds. \quad (7.86)$$

The column density is defined as the number of absorbing ions/atoms per unit area and has units atoms cm^{-2} . It can be interpreted as the number of ion/atoms viewed within the cross sectional area of the beam as collapsed along the path length through the absorbing cloud.

Thus, for an isolated absorption feature in a spectrum that corresponds to a single transition arising from a given ion/atom we can write the solution of the transfer equation in terms of the column density of the ion/atoms and absorption cross section of the transition,

$$f_\lambda = f_\lambda^0 \exp \{-N\sigma(\lambda)\}. \quad (7.87)$$

In such cases, one of the fundamental quantities that can be extracted directly from a spectrum is the column density of the absorbing ions/atom (say, neutral hydrogen in the case of the $\text{Ly}\alpha$ transition). As we will discuss in the following chapter, the column density is a key input for constraining ionization models from which the absorbing cloud physical properties can be determined.

8

Spectrographs and Spectra

In this chapter, the basic principles of spectrograph design are reviewed. Both low order (low-resolution) and high order (high-resolution echelle configurations) are presented. Derivations of the optical principles are not given, but only results that establish the connection between the spectrograph design and the spectral dispersion and spectral purity. Though dispersion of the light can be obtained via prisms or gratings, only the properties of gratings are treated.

Also presented are the interference function, the blaze function, the resolving power, the instrumental spread function, pixel sample rate, and the so-called “resolution element”. Further details on optics, spectrographs, diffraction gratings, and the principles of diffraction gratings can be found in Hecht & Zajac (1974), Schroeder (1987), and Gray (1992). Finally, we present the relationships between the observed flux of an astronomical object and the recorded spectrum in terms of the telescope, spectrograph, detector design, and other observational parameters.

8.1 The Spectrograph

The basic components of a spectrograph are the entrance slit, a collimator, a grating or prism, a camera, and the detector. It is common for the dispersive element in modern spectrographs to be a grating. As shown in Figure 8.1, the light first enters the slit. The slit (1) is located at the focal point of the telescope and the light beam is diverging as the light path exits the slit. The light is then collimated by the collimator (2), which is located at a distance f_{CL} from the slit. The collimated light then strikes the disperser, which is a ruled grating (3). The grating is tilted with respect to the incoming light path. The reflected and spectrally dispersed light is then focused by a camera (4) with focal length f_{CM} onto the detector (5).

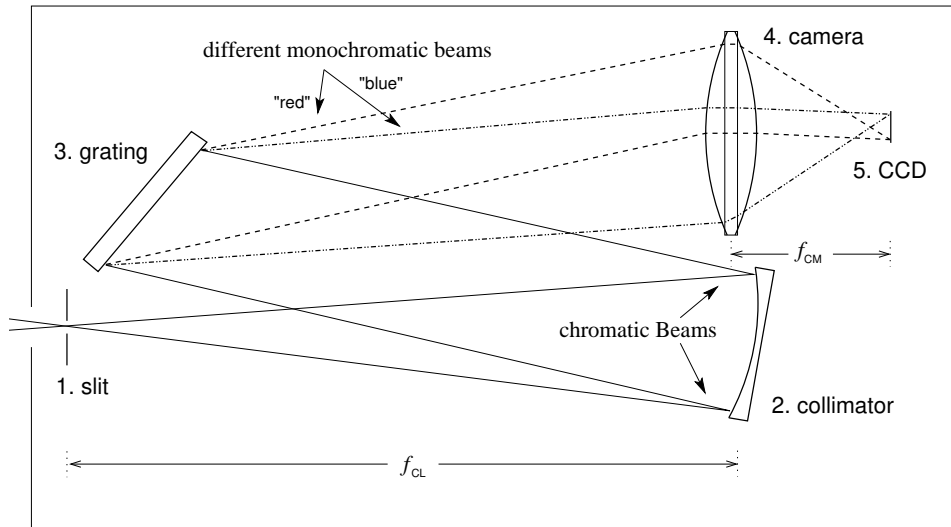


Figure 8.1 A schematic of a grating spectrograph showing the optical layout and light path from slit to detector. The distance from the collimator is the collimator focal length, f_{CL} , and the distance from the camera mirror to the detector is f_{CM} . The light is focused by the telescope at the plane of the slit. The chromatic beam (solid lines) diverges until it is parallelized by the collimator. After the light reflects off the grating, the light is organized into monochromatic beams with different wavelengths (dash-dash and dash-dot lines) due to diffraction; each monochromatic beam propagates in a unique direction. These beams are then focused by the camera at different spatial locations along the detector, forming the spectrum.

Modern detectors are usually charged coupled devices, or CCDs, which are an array of finite pixels that discretized the dispersed light. Due to the one-dimensional optical axis of the grating, the spectrum has a dispersion direction and a spatial direction. The spatial direction will be governed by the collimator and camera focal lengths, and the dispersion will be governed by both the grating properties and the collimator and camera focal lengths. There are secondary effects governing the spatial and dispersion characteristics of the spectrum, including, but not limited to, the stability of the image of the source on the slit (guiding) and the atmospheric turbulence (seeing).

8.2 The Seeing Disk

For a point source, the physical radius of the image in the telescope focal plane (at the slit) is governed by the combined effects of guiding errors and atmospheric “seeing” conditions (the atmospheric turbulence along the sight

line to the object and in the environment of the observatory and telescope dome). Guiding errors and seeing cause a point source to appear as if it has an angular extension on the sky in that the image of the object is “blurred”.

To a good approximation, the resulting surface brightness pattern of the flux of the image in the focal plane of the telescope is an axial symmetric Gaussian,

$$\mu_\lambda(r) = \frac{1}{2\pi\sigma_\lambda^2} \exp\left\{-\frac{r^2}{2\sigma_\lambda^2}\right\}, \quad (8.1)$$

in units of $\text{erg cm}^{-2} \text{ s}^{-1} \text{ \AA}^{-1} \text{ mm}^{-2}$, where r is the physical coordinate in the telescope focal plane (measured in millimeters, mm) with peak, $\mu(0) = 1/2\pi\sigma_\lambda^2$ and dispersion σ_λ . The magnitude of σ_λ can be wavelength dependent, as atmospheric dispersion and refraction decreases with increasing wavelength. Note that $\mu_\lambda(r)$ is defined such that $\int_{-\infty}^{\infty} \mu_\lambda(r) dr = 1$.

The conversion between the apparent angle subtended by the image on the sky and the physical size of the image in the telescope focal plane is given by the telescope plate scale, $\theta_T = 1/f_T$, in units of radians mm^{-1} where f_T is the effective focal length of the telescope optics. The effective focal length is often quoted in millimeters and is on the order of 10^3 – 10^4 mm.

Under the conditions of excellent guiding, the σ_λ of the imaged point source object will be dominated by atmospheric seeing; thus, $\mu_\lambda(r)$ is often referred to as the “seeing disk”. The full-width half-maximum of a point source seeing disk quantifies the seeing conditions. If the measured radius of the seeing disk is σ_λ in millimeters, then the seeing is

$$\Theta_\lambda = 2\sqrt{2\ln 2} \omega_c \sigma_\lambda \theta_T = 2\sqrt{2\ln 2} \omega_c (\sigma_\lambda/f_T), \quad (8.2)$$

in units of seconds of arc (denoted $''$) where the factor $\omega_c \simeq 206,265$ is the conversion factor from radians to arc seconds. Most astronomical facilities strive to deliver seeing on the order of $1''$. Seeing naturally improves toward redder wavelengths and adaptive optics can significantly reduce Θ_λ .

8.3 Slits

The slit is an aperture that lies in the focal plane of the telescope and receives the chromatic seeing disk of the focused image. As illustrated in Figure 8.2, slits are characterized by their physical width, w , and height, h , often in units of millimeters. For a point source, the width of the slit is chosen such that $w \simeq \sigma_\lambda$ and is positioned roughly at the center point along the slit height. Since the seeing disk is wavelength dependent, it is wise to perform the centering at a wavelength near the spectral region of scientific interest.

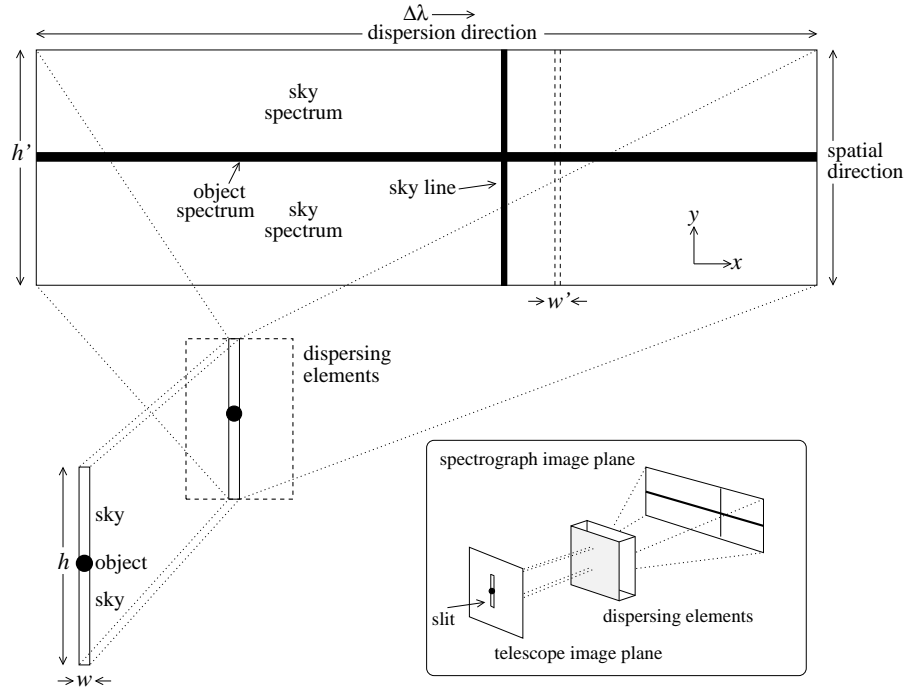


Figure 8.2 A schematic of the spatial and dispersion mapping of the slit for a point source object (the collimator and camera are omitted for simplicity of the illustration). The slit has width w and height h and lies in the telescope image plane. Sky passes through the slit along the slit height. The dispersing elements of the spectrograph disperse the light along the x coordinate (dispersion direction) in the spectrograph image plane. In the slit height direction, the object and sky are imaged in the y coordinate (spatial direction). In the spectrograph image plane, the slit width has physical projection w' in the x coordinate and the slit height has physical projection h' in y coordinate. Sky emission lines are imaged along the y coordinate (mapped along the slit height), but they are dispersed to various x positions depending upon their wavelength.

The slit height is chosen such that a desired amount of sky light is allowed to pass through the spectrograph adjacent to the image of the object. The sky not only contributes a smooth background (unwanted atmospheric scattered light), but also narrow absorption and emission lines and broad molecular absorption bands. It is desirable to obtain a good sampling of these quantities independent of the object in order that they can eventually be removed from the object spectrum. The configuration $h \gg w$ is known as long slit mode. In certain configurations, such as with echelle spectrographs (see § 8.6), a shortened slit is employed.

The angular projection of the slit width on the sky is w times the telescope plate scale $\theta_w = \omega_c(w/f_\tau) = \omega_c(w/f_\tau)$, in units of arc seconds, where w and f_τ are in units of millimeters. Similarly, the angular projection of the slit height on the sky is $\theta_h = \omega_c(h/f_\tau)$. In practice, observatories may provide the option to set the slit width and/or height in units of arc seconds rather than in millimeters. Clearly, for a point source, the chosen slit angular width should be on the order of the seeing, $w \simeq \Theta_\lambda$, in order to minimize loss of light (slit too narrow) or compromise the purity of the object light (slit too wide).

Also illustrated in Figure 8.2 is a simplified schematic of the mapping of light onto the image plane of the spectrograph. The collimator and camera have been omitted from the light path. The focused chromatic light entering the slit is wavelength dispersed along the x coordinate of the spectrograph image plane. The x coordinate is known as the dispersion direction. In the y coordinate, the object and sky are not dispersed; they are imaged. Thus, the y coordinate is known as the spatial direction; a constant y at a given x (vertical cut along the spectrograph image plane) is a monochromatic image of the slit.

In the spectrograph image plane, the slit width has physical size w' in x coordinates and the slit height has physical size h' in y coordinates. As we will show in following sections of this chapter, in the x direction, the plate scales, θ_x , the dispersion relation $d\lambda/dx$, and the magnification, w'/w , dependent on both the optical and dispersive properties of the spectrograph, whereas in the y direction, the plate scale, θ_y , and the magnification, h'/h , depend only upon the optical properties.

8.4 Diffraction Gratings

The diffraction grating is the optical element in the spectrograph that disperses the light beam and resolves the light into a spectrum. The light beam striking the grating must be collimated so that all points within the beam cross-sectional area have the same angle of incidence on the grating surface. Note then, that it is desirable for the projected area of the collimator and the projected area of the grating to both be equal to the cross-sectional area of the light beam. If the beam under fills the grating, then the efficiency of the spectrograph is not maximized; if the light beam over fills the grating, then un-dispersed stray light might be scattered back into the light beam and degrade recorded spectrum.

A reflective grating is characterized by its size and by the spacing of machined grooves on its surface. These grooves, called facets, run perpendicular

to the long axis of the grating. Consider a grating of length L , as illustrated in Figure 8.3, with facet widths s and with center-to-center spacings d , with the condition $d > s$. The ratio L/d gives the total number of facets, N_f .

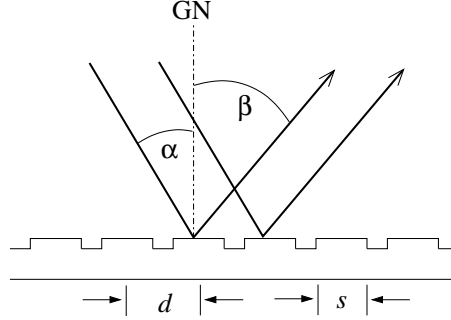


Figure 8.3 A schematic of a reflection grating showing the individual grooves, or facets. The facets are characterized by their edge-to-edge length, s , and center-to-center separation d . The incoming collimated light beam has an angle of incidence α , measured with respect to the grating normal, GN. Interference of the reflected wave fronts yields a diffraction angle, β , which is a function of wavelength phase differences (Eq. 8.3).

For a specific spectrograph design, the grating is tilted so that the angle of incidence of the incoming collimated chromatic light is α with respect to the grating normal. After interaction with the grating, the light is diffracted at wavelength dependent angles, β ; there is a unique value of β for each wavelength that corresponding to a given wavelength phase differences. As discussed in the following sections, there are two interference patterns that govern the angular dependence of the intensity function, which describes the fraction of transmitted intensity at each wavelength. The first is due to phase differences between the centers of adjacent facets, φ_{cc} . The second is due to phase differences between the center and edge of each facet, φ_{ce} .

8.4.1 The Grating Equation

Light with wavelength λ is “reflected” at an angle of diffraction that corresponds to constructive interference, or facet center to facet center phase differences of $\varphi_{cc} = n\pi$. The relationship between the diffraction angle and the wavelength is given by the grating equation,

$$\frac{n\lambda}{d} = \sin \alpha + \sin \beta, \quad (8.3)$$

where the integer n is called the diffraction order. For a fixed α , the angular dispersion with wavelength is obtained by differentiating Eq. 8.3,

$$\frac{d\beta}{d\lambda} = \frac{n}{d \cos \beta} = \frac{\sin \alpha + \sin \beta}{\lambda \cos \beta}, \quad (8.4)$$

from which it is clear that the angular dispersion increases with larger n and smaller d .

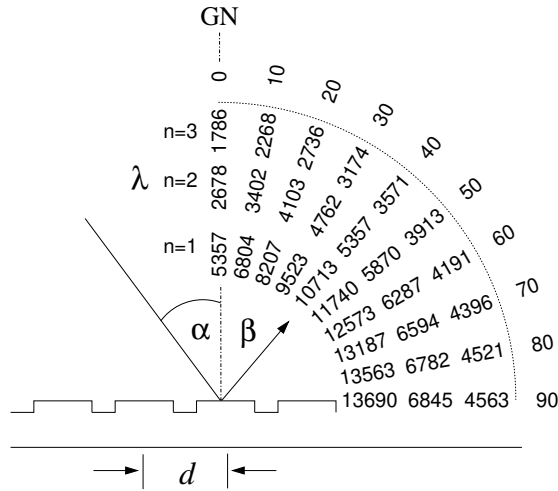


Figure 8.4 A schematic of the behavior of a grating with $1/d = 1200$ facets mm^{-1} for an incident angle $\alpha = 40^\circ$. At each diffraction angle, β , shown in intervals of 10° , the wavelengths for orders $n = 1, 2$, and 3 are tabulated, as governed by equation (Eq. 8.3). Phase differences in the light reflection off adjacent facets interfere constructively for $\varphi_{cc} = n\pi$ for each λ , which yields monochromatic beams each diffracted at a unique β . Note that the angular dispersion increases with n according to Eq. 8.4.

The behavior of the grating equation (Eq. 8.3) is illustrated in Figure 8.4, for orders $n = 1, 2$, and 3 for a grating with $1/d = 1200$ facets mm^{-1} for an incident angle $\alpha = 40^\circ$. For this grating, when $n = 0$ there is no dispersion and all wavelengths of the chromatic incident beam are diffracted at angle $\beta = \alpha$. As n increases, the angular dispersion increases. For the example grating, if one wished to observe a wavelengths in the regime of 6000 \AA , the camera would need to be placed at $\beta = 5^\circ$ for $n = 1$ or $\beta = 55^\circ$ for $n = 2$.

8.4.2 Free Spectral Range

As can be seen in Figure 8.4, the camera can be placed to capture the spectrum for any desired order. However, there are multiple $n\lambda$ that satisfy

Eq. 8.3 for a given β . Thus, there are different wavelengths from different orders that are diffracted into the same angle β . This results in spatial overlap of wavelengths satisfying $(n-1)\lambda_{n-1} = n\lambda_n$, which is obtained by the ratio of the grating equation for adjacent orders.

The free spectral range for order n is defined as the wavelength range,

$$\begin{aligned}\Delta\lambda_n^+ &= \lambda_{n-1} - \lambda_n, \\ \Delta\lambda_n^- &= \lambda_n - \lambda_{n+1},\end{aligned}\tag{8.5}$$

for a given β . After substitution of the recursive relationship, $(n-1)\lambda_{n-1} = n\lambda_n$, we have

$$\begin{aligned}\Delta\lambda_n^+ &= \lambda_n/(n-1), \\ \Delta\lambda_n^- &= \lambda_n/(n+1).\end{aligned}\tag{8.6}$$

Consider the free spectral range for the example grating illustrated in Figure 8.4 for $\beta = 40^\circ$ at $n = 2$, which diffracts wavelength $\lambda_2 = 5357 \text{ \AA}$. We have $\Delta\lambda_2^+ = 5357/(2-1) = 5357 \text{ \AA}$, which, added to λ_2 , yields 10713 \AA for $\beta = 40^\circ$ at $n = 1$, and $\Delta\lambda_2^- = 5357/(2+1) = 1786 \text{ \AA}$, which, subtracted from λ_2 yields 3571 \AA for $\beta = 40^\circ$ at $n = 3$. Examination of Figure 8.4 illustrates this case. It can be shown that, for fixed α and β , the wavelength range of the free spectra range is

$$\Delta\lambda = \Delta\lambda_n^+ + \Delta\lambda_n^- = \frac{2n}{n^2-1}\lambda_n.\tag{8.7}$$

The spatial overlap of different wavelengths can cause confusion in the recorded spectrum. There are two solutions to mitigate this confusion, depending upon whether the spectrograph is designed for low-order or high-order diffraction. In the case of small n , a blocking filter is placed in the light path (before the dispersion grating will work). For large n , a cross-dispersing grating is placed in the light path, after the dispersion grating. This latter approach is used for echelle spectrographs.

8.4.3 The Interference Function

In reality, there is a continuum of center-center interference phases that are periodic over continuous multiples of π . The generalized relationship between φ_{cc} , β , and each monochromatic diffracted λ is governed by

$$\varphi_{cc} = \frac{\pi d}{\lambda} (\sin \alpha + \sin \beta) .\tag{8.8}$$

Substituting $\varphi_{cc} = n\pi$ into Eq. 8.8, yields the grating equation (Eq. 8.3). The pattern of the normalized intensity for λ is called the interference function and is given by

$$I_i(\varphi_{cc}) = \frac{\sin^2(L/d)\varphi_{cc}}{\sin^2 \varphi_{cc}}, \quad (8.9)$$

where $L/d = N_f$, is an integer. For a given λ , there is a range of interference phases, which yield maximum intensity for $\varphi_{cc} = n\pi$ (constructive interference) and zero intensity for $\varphi_{cc} = (n/2)\pi$ (destructive interference).

In Figure 8.5(a), the intensity function (Eq. 8.9) is plotted as a function of φ_{cc} for fixed λ . It is clear that the intensity pattern recorded for a given λ is not a δ -function, but is “smeared” over a small range of phase differences $\Delta\varphi_{cc} = \lambda/N_f d = \lambda/L$. The order separation of the peaks occurs when φ_{cc} is an integer multiple of λ/d .

Consider the example grating illustrated in Figure 8.4. For $\lambda = 5357 \text{ \AA}$ in first order ($n = 1$), the diffraction angle is $\beta = 0^\circ$, and for $n = 2$ is $\beta = 40^\circ$. Naively, it would appear that one could place the camera at either location depending upon the desired dispersion (higher n location for higher dispersion). However, as we will see below, the resulting intensity of the diffracted light is not equal for all orders for this type of grating due to additional phase differences introduced by the facet sizes, s .

8.4.4 The Blaze Function

As described by Eq. 8.9 and as illustrated in Figure 8.5(a), the intensity pattern due to center-center diffraction is periodic and symmetric for all n . However, there is additional phase modulations that occur due to interference phases between the center and the edge of each facet. For center-edge interference, the relationship between β , λ , and the center-edge phase φ_{ce} is

$$\varphi_{ce} = \frac{\pi s}{\lambda} (\sin \alpha + \sin \beta). \quad (8.10)$$

The pattern of the normalized intensity for this interference, called the blaze function, is given by

$$I_b(\varphi_{ce}) = \frac{\sin^2 \varphi_{ce}}{\varphi_{ce}^2}. \quad (8.11)$$

The blaze function is a maximum when $\varphi_{ce} = 0$, which occurs when the incident and diffraction angle are symmetric about the facet normal, i.e., $\alpha = \beta$. For plane-parallel (un-tilted) facets, as shown in Figure 8.4, this condition is met for $n = 0$ only. The minima occur at $\varphi_{ce} = n\pi$, so that

the phase difference spread of the blaze function is $\Delta\varphi_{ce} = \lambda/s$. Note that this is a significantly broader width than that of the interference function, $\Delta\varphi_{cc} = \lambda/L$.

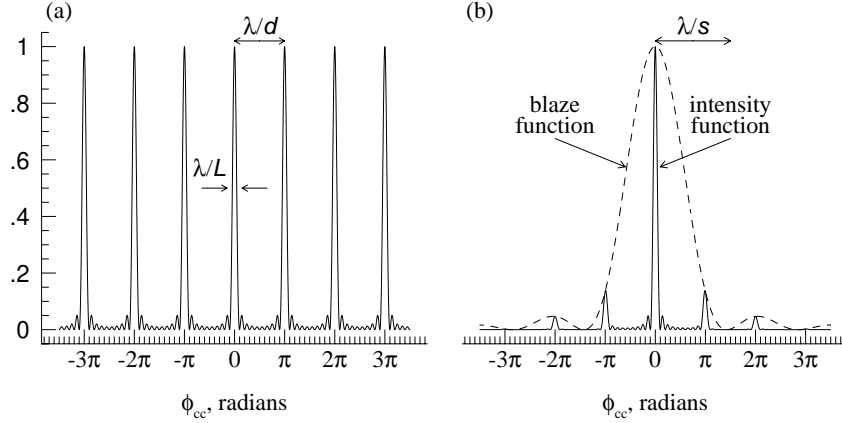


Figure 8.5 (a) The periodic interference function, Eq. 8.9, which results from the facet center-to-center interference, plotted as a function of the phase difference (in units of π) for an arbitrary wavelength λ . Note the peaks are at $n\pi$, where n is an integer. For this example, $N_f = L/d = 10$, which is an unrealistic configuration used here only for illustration purposes. A typical grating might have $L = 40$ cm with 600 facets mm^{-1} , which gives $N_f = L/d = 24,000$; thus, real gratings have much narrower $\Delta\varphi_{cc}$ peaks. Note that each peak satisfies the grating equation (Eq. 8.3) and is thus diffracted at different β . (b) The blaze function, Eq. 8.11, which results from facet center-to-edge interference also plotted as a function of the phase difference of the center-center diffraction. The width of the blaze function is governed by the facet sizes; for this example the ratio $s/d = 0.7$ was used. The blaze function modulates the interference function, giving the overall intensity pattern, Eq. 8.13.

Naively, it might appear that φ_{ce} is a periodic function and that Eq. 8.11 would have multiple maxima, but this is not the case. The behavior of the blaze function for a given wavelength at various orders is more clearly seen by solving the grating equation (Eq. 8.3) for λ and substituting into Eq. 8.10, which yields

$$I_b(n) = \frac{\sin^2 n\pi(s/d)}{[n\pi(s/d)]^2}, \quad (8.12)$$

the relative blaze intensity for order n . It is because $d \neq s$ that φ_{ce} , and thus $I_b(n)$, is not periodic. Note that $I_b(n)$ peaks at $n = 0$ and decreases as n^2 . The interpretation is that for a given λ appearing in multiple orders, the relative intensity is governed by Eq. 8.12.

It is not desirable to have the blaze function peak at $n = 0$, where there is no dispersion. The remedy is to “blaze” the grating by tilting the facets, which shifts the blaze peak. As we will show, it also yields blaze peaks for all values of n . This is discussed in § 8.4.6.

8.4.5 The Intensity Function

The resulting intensity function for light of wavelength λ , is the product of the interference and blaze functions,

$$I(\lambda) = I_i(\varphi_{cc}) \cdot I_b(\varphi_{ce}) = \frac{\sin^2(L/d)\varphi_{cc}}{\sin^2 \varphi_{cc}} \cdot \frac{\sin^2 \varphi_{ce}}{\varphi_{ce}^2}. \quad (8.13)$$

As shown in Figure 8.5(b), the resulting intensity pattern is simply the interference function modulated by the blaze function. Note that, for a non-blazed grating, the peak of the intensity occurs at $n = 0$, when $\alpha = \beta$ (see Figure 8.4) and the interference due to center-edge phase differences significantly reduces the throughput of light at higher orders. Again, this is more clearly seen by solving the grating equation (Eq. 8.3) for λ and substituting into Eq. 8.13, which yields the relative intensity for order n ,

$$I(n) = I_i(n) \cdot I_b(n) = \frac{\sin^2 n\pi(s/d)}{[n\pi(s/d)]^2}, \quad (8.14)$$

where $I_i(n) = 1$ for all n . Thus, the relative intensity of a given wavelength at order n is modulated strictly by the blaze function.

Again, consider the example grating shown in Figure 8.4 for the ratio $s/d = 0.7$. To observe wavelength 5357 Å with the lowest dispersion ($n = 1$ and $\beta = 0^\circ$) would be less than ideal because the intensity is modulated to roughly 10% its $n = 0$ throughput. Observing this wavelength at higher dispersion ($n = 2$ and $\beta = 40^\circ$), would pose an even greater challenge because the intensity is modulated down even further. Note that 5357 Å does not appear in order $n = 3$ and above. By decreasing the ratio s/d , the blaze function can be broadened so that the throughput is increased for $n > 0$. However, this is not the most effective solution.

8.4.6 Blazing a Grating

As shown above, the maximum efficiency in the intensity function occurs when the incident, α , and diffraction, β , angles are symmetric about the facet normal, which is the same as the grating normal, GN, when the grating facets are not tilted (as shown in the example grating in Figures 8.3 and 8.4).

From Eq. 8.4, we see that the dispersion, $d\beta/d\lambda$, scales as d/n , so in order to obtain a given resolving power for a spectrograph, it is desirable to maximize the intensity function (Eq. 8.13) for the required d and n combination. However, unless the blaze function is modified in some manner, the intensity pattern is maximized for $n = 0$ (no dispersion!) and is significantly reduced for large n .

In order to maximize the throughput at higher orders, the peak of the blaze function must be “shifted” along the φ_{cc} axis in Figure 8.5(b) such that it is centered at the desired n . By tilting the grating facets by an angle ϕ , as shown in Figure 8.6, the center-edge diffraction pattern can be phase shifted. In fact, blazing the grating has the added effect that the blaze function has a peak for all n when the incident and diffraction angles are symmetric about the facet normal (which occurs for different λ for each n).

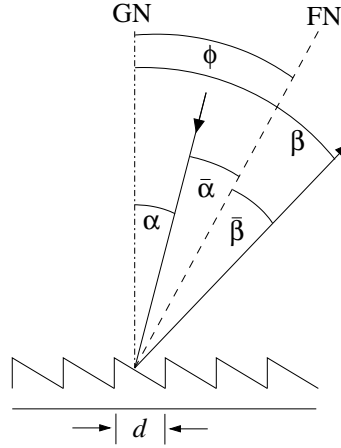


Figure 8.6 To maximize the diffraction illumination pattern for $n > 0$, a phase shift is introduced to the center-edge diffraction interference by grooving the facets at an angle ϕ , which also defines the facet normal, FN. The blaze function then peaks when $\bar{\beta} = \bar{\alpha}$.

The configuration for a blazed grating with facet tilts, ϕ , is illustrated in Figure 8.6. The facet lengths are now $s = d \cos^2 \phi$, and the center-edge diffraction is based upon the angles with respect to the facet normal (FN, as is the case for the special case $\phi = 0^\circ$). Defining the incident angle with respect to FN as $\bar{\alpha} = \alpha - \phi$ and the diffraction angle $\bar{\beta} = \beta - \phi$, Eq. 8.10 is rewritten,

$$\varphi_{ce} = \frac{\pi s}{\lambda} [\sin \bar{\alpha} + \sin \bar{\beta}], \quad (8.15)$$

$$= \frac{\pi d \cos^2 \phi}{\lambda} [\sin(\alpha - \phi) + \sin(\beta - \phi)]. \quad (8.16)$$

Recall that the blaze function (Eq. 8.11) is a maximum when $\varphi_{ce} = 0$. Thus, the optimal relationship between the facet tilts and the angles of incident and diffraction with respect to the grating normal, GN, is obtained when $\alpha - \phi = -(\beta - \phi)$, or $\phi = (\alpha + \beta)/2$. We will show that the blaze function for tilted facets peaks for all n at the λ dispersed into the $\bar{\beta} = \bar{\alpha}$ direction, which occurs when the incident and diffraction angles are symmetric about the facet normal.

The wavelength at which the blaze function peaks at each order can be computed from the grating equation (Eq. 8.3). We have

$$\frac{n\lambda}{d} = \sin \alpha + \sin \beta, \quad (8.17)$$

$$= \sin(\bar{\alpha} + \phi) + \sin(\bar{\beta} + \phi). \quad (8.18)$$

Applying $\bar{\beta} = \bar{\alpha}$, where the blaze function maximizes, yields,

$$\frac{n\lambda_b}{d} = 2 \sin \phi \cos \bar{\alpha} = 2 \sin \phi \cos(\alpha - \phi), \quad (8.19)$$

where λ_b is called the blaze wavelength for order n .

The general behavior of a blazed grating is more clearly seen by solving for d/λ_b in Eq. 8.19 and substituting into Eq. 8.15, which yields

$$\varphi_{ce} = \frac{n\pi \cos \phi [\sin \bar{\alpha} + \sin \bar{\beta}]}{2 \tan \phi \cos \bar{\alpha}}, \quad (8.20)$$

where $s = d \cos^2 \phi$ as been invoked. Clearly, the center-edge interference phase is $\varphi_{ce} = 0$ for $\bar{\beta} = \bar{\alpha}$ diffraction for all n . As before, the relative intensity for order n is the interference function modulated by the blaze function at each $\bar{\beta}$,

$$I(n, \bar{\beta}) = I_i(n) \cdot I_b(n, \bar{\beta}) = \frac{\sin^2 \left(\frac{n\pi \cos \phi [\sin \bar{\alpha} + \sin \bar{\beta}]}{2 \tan \phi \cos \bar{\alpha}} \right)}{\left(\frac{n\pi \cos \phi [\sin \bar{\alpha} + \sin \bar{\beta}]}{2 \tan \phi \cos \bar{\alpha}} \right)^2}, \quad (8.21)$$

where

$$\lambda(n, \bar{\beta}) = \frac{d}{n} (\sin \bar{\alpha} + \sin \bar{\beta}), \quad (8.22)$$

is the wavelength at each $\bar{\beta}$ for order n .

Equations 8.21 and 8.22 provide the basic relationships for spectrograph design. A central choice of the design is, of course, the spectral dispersion (resolving power), which increases with increasing n/d . As such, the choice of n and d are critical. In the most general terms, spectrographs either work

with lower dispersion in low orders, or in higher dispersions in higher orders. We discuss two example designs below.

8.5 Low-Order Spectrographs

Low-order gratings are generally blazed at $15^\circ \leq \phi \leq 20^\circ$ with $300 \leq 1/d \leq 1200$ facets mm^{-1} for orders $1 \leq n \leq 3$. It is desirable to have $\bar{\alpha} \sim \phi$. The blaze functions (Eq. 8.21) and solutions to the grating equation (Eq. 8.22) for a spectrograph with $\phi = 18^\circ$, $1/d = 600$ facets mm^{-1} , and $\bar{\alpha} = 22^\circ$ are illustrated in Figure 8.7 for orders $n = 1, 2$, and 3 .

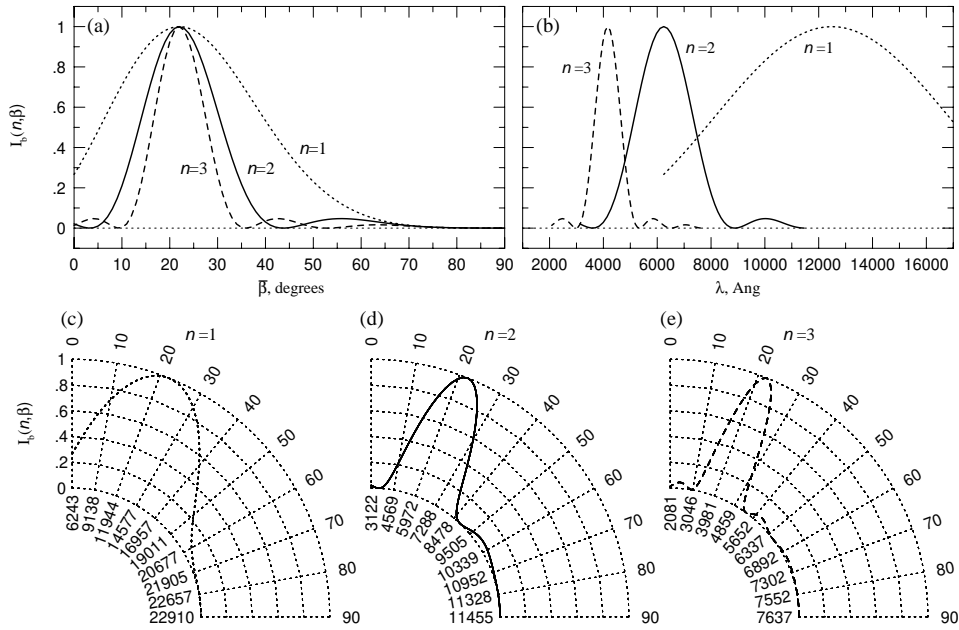


Figure 8.7 Solutions to Eqs. 8.21 and 8.22 for a spectrograph with at $\phi = 18^\circ$, $1/d = 600$ facets mm^{-1} , and $\bar{\alpha} = 22^\circ$, for orders $n = 1, 2$, and 3 . (a) The dependence of the blaze functions on the diffraction angle, $\bar{\beta}$. (b) The blaze functions with wavelength, λ . (c, d, e) The blaze functions shown in polar coordinate representation, which provides the spatial relationship between λ , $I_b(n, \bar{\beta})$, and $\bar{\beta}$.

As can be seen in Figure 8.7(a), the blaze functions peak at $\bar{\beta} = \bar{\alpha} = 22^\circ$ for all n . Note that the angular width of the blaze function scale as $1/n^2$. Only the range $0^\circ \leq \bar{\beta} \leq 90^\circ$ is presented. As seen in Figure 8.7(b), the wavelengths of the blaze peaks scale with $1/n$.

This example spectrograph is configured for an optical wavelength range of 4000–8500 Å in order $n = 2$ for a camera centered at $\bar{\beta} = 22^\circ$ that

subtends roughly 20° ($10^\circ \leq \bar{\beta} \leq 30^\circ$). Although many of these wavelengths appear in multiple orders, examination of Figures 8.7(c)–(e) reveals that they are diffracted toward different $\bar{\beta}$ for each n , which mitigates confusion. Consider the $n = 2$ blaze center, $\lambda \simeq 6500 \text{ \AA}$ at $\bar{\beta} \simeq 22^\circ$. This wavelength is diffracted toward $\bar{\beta} \simeq 5^\circ$ for $n = 1$ and toward $\bar{\beta} = 55^\circ$ for $n = 3$. Both these diffraction angles reside well beyond the angular placement of the camera for $n = 2$.

Potential confusion can occur for $\lambda \leq 4860 \text{ \AA}$ from $n = 3$ spatially overlapping with $\lambda \geq 4860 \text{ \AA}$ from $n = 2$. In this case, a “high-pass” blocking filter that transmits $\lambda \geq 4860 \text{ \AA}$ would eliminate the confusion. As such, this example spectrograph design is not optimal for blue wavelengths.

8.6 Echelle Spectrographs

Echelle gratings can have a large range of ϕ , $\bar{\alpha}$, and d . However, some useful applications fall in the range of gratings blazed at $50^\circ \leq \phi \leq 70^\circ$ with $60 \leq 1/d \leq 80 \text{ facets mm}^{-1}$ for orders $20 \leq n \leq 100$. As with lower dispersion spectrographs, it is desirable to have the incident angle $\bar{\alpha} \sim \phi$. The blaze functions (Eq. 8.21) and solutions to the grating equation (Eq. 8.22) for an echelle spectrograph with $\phi = 52^\circ$, $1/d = 70 \text{ facets mm}^{-1}$, and $\bar{\alpha} = 70^\circ$ are illustrated in Figure 8.8 for orders $n = 32$ – 70 .

In Figure 8.8(a), the blaze functions are shown as a function of wavelength. Orders 35, 45, 55, and 65 are labeled. In Figures 8.8(b)–(e), the angular relationships for these four orders are illustrated. As with the low resolution design presented above, the widths of the blaze function become narrower as $1/n^2$. However, what is most striking is that each blaze function covers a narrow range of wavelengths (high dispersion) and is spatially aligned at $\bar{\beta} = \bar{\alpha} = 70^\circ$. For a camera and detector that are sensitive to the full optical and near-infrared range (4000 – 9000 \AA), there will be significant confusion due to the spatial overlap of the wavelengths present in each order. For example, all wavelengths at the blaze peaks shown in Figure 8.8(a) will overlap at $\bar{\beta} = 70^\circ$.

To spatially separate the echelle orders, a second grating, called a cross disperser, is placed in the light path following the echelle grating. The cross disperser operates at low order, usually $m_x = 1$, and thus has grating parameters that are similar to the low order spectrograph example in § 8.5. The cross disperser is oriented such that the facets are at a right angle with respect to the echelle grating facets (on its side). Note that, if the cross disperser is a flat grating, the diverging light beam from the echelle grating must be re-collimated, which can be accomplished with a refracting lens

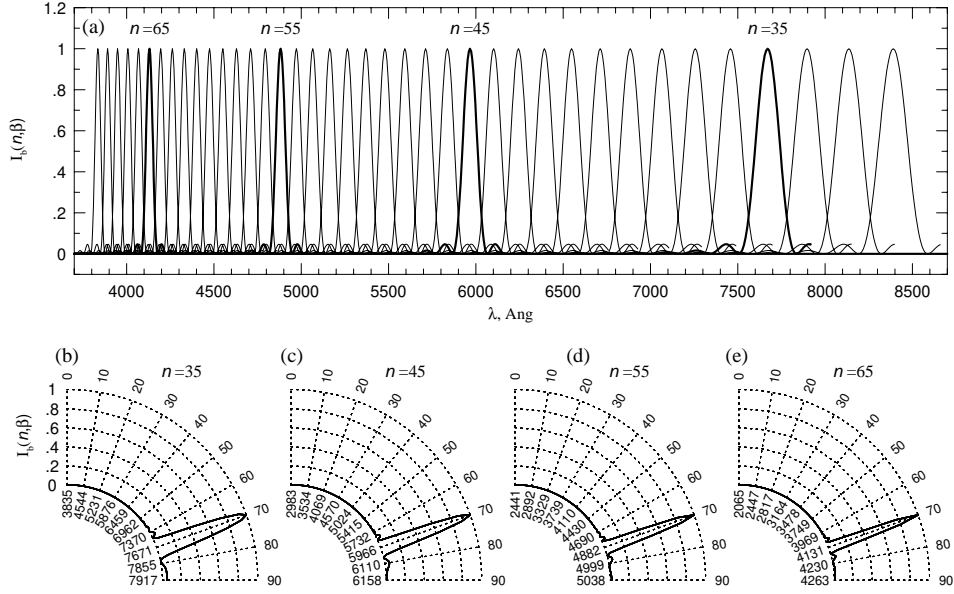


Figure 8.8 Solutions to Eqs. 8.21 and 8.22 for an echelle spectrograph with at $\phi = 52^\circ$, $1/d = 80$ facets mm^{-1} , and $\bar{\alpha} = 70^\circ$, for orders $n = 32\text{--}70$. (a) The blaze functions with wavelength, λ . (b, c, d, e) The blaze functions shown in polar coordinate representation (for $n = 35, 45, 55,$ and 65), which provide the angular relationship between λ , $I_b(n, \beta)$, and β .

(though it is possible to use a concave grating for the cross disperser) to ensure an effective constant incident angle, α_x .

Consider a cross disperser grating with facet separations d_x blazed at ϕ_x operating in order m_x and oriented so that the post-echelle grating light is incident at angle α_x . The cross disperser will refract each wavelength into the direction β_x . Following the above convention, let $\bar{\alpha}_x = \alpha_x - \phi_x$ and $\bar{\beta}_x = \beta_x - \phi_x$ be the incident and refraction angles with respect to the cross disperser facet normals. The cross disperser introduces a two dimensional mapping of wavelength in the $\bar{\beta}_x\text{--}\bar{\beta}$ plane (using the convention that the echelle grating has no subscript).

The diffraction angle off the cross disperser is

$$\sin \bar{\beta}_x = \frac{m_x}{d_x} \lambda(n, \bar{\beta}) - \sin \bar{\alpha}_x, \quad (8.23)$$

where $\lambda(n, \bar{\beta})$ is taken from Eq. 8.22 for the echelle grating. From Eq. 8.4,

the mapping will have a tilt in the $\bar{\beta}_x$ - $\bar{\beta}$ plane given by

$$\tan \Psi(\bar{\beta}_x, \bar{\beta}) = \frac{d\bar{\beta}_x/d\lambda}{d\bar{\beta}/d\lambda} = \frac{m_x}{d_x} \frac{d \cos \bar{\beta}}{n \cos \bar{\beta}_x}. \quad (8.24)$$

The cross dispersion is relatively constant (though not strictly) for a given echelle order (small $\bar{\beta}_x$ range) while the echelle dispersion is varying; as such the angle $\Psi(\bar{\beta}_x, \bar{\beta})$ will vary and the orders will have a slight curvature in the $\bar{\beta}_x$ - $\bar{\beta}$ plane. At the blaze peaks, the order separation will be

$$\sin \bar{\beta}_x = 2 \frac{m_x}{m} \frac{d}{d_x} \sin \bar{\alpha} - \sin \bar{\alpha}_x, \quad (8.25)$$

which was obtained by substituting $\bar{\beta} = \bar{\alpha}$ into Eq. 8.22.

The overall echelle blaze function is then the blaze function of the echelle grating modulated by the blaze function of the cross disperser,

$$I(m_x, \bar{\beta}_x, n, \bar{\beta}) = I(n, \bar{\beta}) \cdot I(m_x, \bar{\beta}_x), \quad (8.26)$$

where $I(n, \bar{\beta})$ and $I(m_x, \bar{\beta}_x)$ are the blaze functions for the echelle and cross disperser gratings obtained from Eq. 8.21.

The blaze functions (Eq. 8.21) and solutions to the grating equations (Eqs. 8.22 and 8.23) for a spectrograph with the above echelle grating parameters and a cross disperser with $\phi_x = 14^\circ$, $1/d_x = 1200$ facets mm^{-1} , and $\bar{\alpha} = 22^\circ$ are illustrated in Figure 8.9. Figure 8.9(a) illustrates the modulation of the echelle blaze functions by the cross disperser blaze function (dotted curve). As in Figure 8.8, orders $n = 35, 45, 55$, and 65 are shown as thick curves. Figure 8.9(b) illustrates the $\bar{\beta}_x$ dependence of the cross disperser blaze function and its wavelength range in polar coordinates.

The mapping of wavelength in the $\bar{\beta}_x$ - $\bar{\beta}$ plane is illustrated in Figure 8.9(c). The resulting mapping aligns the echelle orders in the $\bar{\beta}_x$ direction with the order dispersion mapped in the $\bar{\beta}$ direction. Longer wavelength orders are diffracted to larger $\bar{\beta}_x$, so that wavelength increases from the lower left to the upper right in the $\bar{\beta}_x$ - $\bar{\beta}$ plane. Order $n = 35$, with blaze wavelength $\lambda_b = 7671$ Å ($\bar{\beta} = 70^\circ$) is centered at $\bar{\beta}_x = 31^\circ$, where as order $n = 55$, with blaze wavelength $\lambda_b = 4882$ Å is centered at $\bar{\beta}_x = 12^\circ$. The free spectral range for each order is shown as the solid curves. Outside the free spectral range, a given wavelength is mapped in adjacent orders.

The main benefits of the echelle format is a large resolving power (high dispersion) with a large wavelength coverage while minimizing confusion; though it may be necessary to use a blocking filter for the shorter wavelengths due to second order dispersion from the cross disperser. Clearly, a shortened slit is required so that the sky does not overlap in the spatial

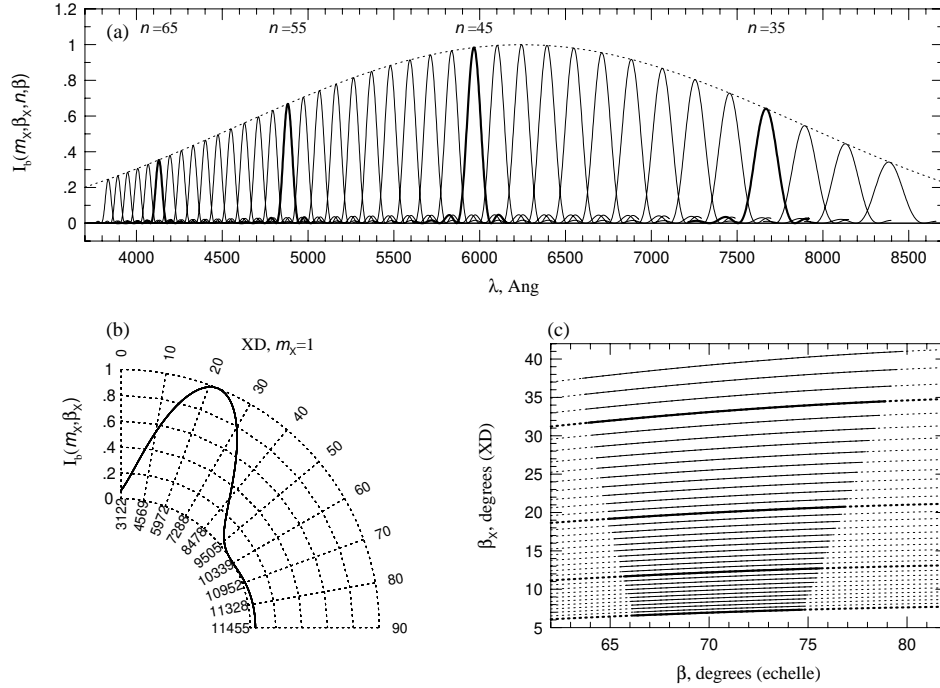


Figure 8.9 Solutions to Eqs. 8.21, 8.22, and 8.23 for the echelle grating illustrated in Figure 8.8 after being cross dispersed in order $m_x = 1$ with a grating blazed at $\phi = 14^\circ$, $1/d = 1200$ facets mm^{-1} , and $\bar{\alpha} = 22^\circ$. (a) The dependence of the echelle blaze functions on wavelength, λ . The dotted curve is the cross disperser blaze function. (b) The cross disperser blaze function shown in polar coordinate representation, which provide the angular relationship between λ , $I_b(m_x, \beta_x)$, and β_x . (c) The mapping of the echelle orders, n , in the β_x - β plane. Wavelength increases from the lower left to the upper right; orders $n = 35$ with $\lambda_b = 7671 \text{ \AA}$, $n = 45$ with $\lambda_b = 5966 \text{ \AA}$, $n = 55$ with $\lambda_b = 4882 \text{ \AA}$, and $n = 65$ with $\lambda_b = 4131 \text{ \AA}$ are presented as thick curves. The solid curves provide the free spectral range of each order.

(cross disperser) direction. If the orders are not well separated, a drawback of the echelle format is that background light from the sky may be difficult to sample. There can also be complications from scattered light from the many required optical elements.

8.7 Resolving Power

The diffraction-limited resolution of a spectrograph quantifies the wavelength difference at which the central peaks of adjacent contiguous monochromatic beams are measurably separable in the focal plane of the detector.

This difference is essentially the full-width half-maximum of their interference peaks. However, as we will discuss below, because even a point-source astronomical light source has a discernable seeing disk that fills the slit, the resolving power of a spectrograph is limited by the wavelength difference at which monochromatic images of the slit are measurably separable. This is known as the spectral purity and is given by the full-width half-maximum of the spectral images of the slit in the focal plane of the detector. We will show that the spectral purity depends not only on the spectrograph design, but also on the telescope to which it is mated.

The general definition of resolving power, denoted by the symbol R , is the ratio of the wavelength centroid divided by the full-width half-maximum of the dispersion at that wavelength,

$$R = \frac{\lambda}{\Delta\lambda}. \quad (8.27)$$

8.7.1 Diffraction Limited Resolution

As the dispersion of the light occurs at the grating, the monochromatic resolving power of the grating provides the diffraction-limited spectral resolution of the spectrograph. We begin with the grating equation (Eq. 8.3),

$$\varphi_{cc}(n) = \frac{n\lambda}{d} = \sin \alpha + \sin \beta, \quad (8.28)$$

where $\varphi_{cc}(n)$ is defined as the center-to-center phase difference from reflection off of adjacent facets for φ_{cc} in integer multiples of π . As we showed in § 8.4.5, from the intensity function (Eq. 8.13) illustrated in Figure 8.5(a), the full-width half-maxima of the interference peaks are given by

$$\Delta\varphi_{cc}(n) = \frac{\lambda}{L}, \quad (8.29)$$

which also follows from the standard theorem of Fourier transforms that data sampled with a window Δ have frequency widths of $1/\Delta$; in this case the length of the grating, L . Differencing the grating equation yields

$$\Delta\varphi_{cc}(n) = \frac{n}{d}\Delta\lambda_G, \quad (8.30)$$

where the subscript “G” refers to the grating. Equating Eqs. 8.29 and 8.30, and invoking Eq. 8.27, the grating monochromatic diffraction-limited resolving power is

$$R_G = \frac{\lambda}{\Delta\lambda_G} = n \frac{L}{d}. \quad (8.31)$$

Eq. 8.31 is the theoretical limit of the resolving power of a spectrograph, and it indicates that the larger the grating and the closer the facets are grooved, the higher the resolving power for a given order n .

8.7.2 Seeing Limited Resolution

The true resolving power of a spectrograph is lower than the diffraction-limited resolution of the grating and is dominated by wavelength dispersion of the monochromatic image of the slit at the detector (camera focal plane).

Consider a monochromatic light source with an image that spans the slit width as seen from a collimator with focal length f_{CL} . Relative to the normal of the collimator axis, light from opposite edges of the slit width arrives at the collimator over the range of incident angles $\alpha \rightarrow \alpha + d\alpha$. If the slit width is w , then the angular size at the collimator, and therefore the range of incident angles is

$$d\alpha = \frac{w}{f_{\text{CL}}} . \quad (8.32)$$

The angular spread of the collimated beam incident on the grating will also be $\alpha \rightarrow \alpha + d\alpha$. A spectrograph in which the collimated beam exactly fills the grating (i.e, the diameter of the collimator equals the length of the grating) will have $f_{\text{CL}}/L = f_{\text{T}}/D$, where D is the diameter of the telescope primary mirror and f_{T} is the effective focal length of the telescope. This relation follows from the fact that the angle of the divergent beam passing through the slit at the telescope focal plane is identical to the angle at which the focused beam converges from the telescope and assumes the collimator diameter exactly fills the beam. A telescope-spectrograph configuration as described here is the most efficient, and yields,

$$d\alpha = \frac{w}{f_{\text{T}}} \frac{D}{L} . \quad (8.33)$$

From the grating, a range of reflection angles, $d\beta$, will occur due to the range of incident angles, $d\alpha$. The relationship between $d\alpha$ and $d\beta$ follows from differentiating the grating equation, Eq. 8.3,

$$\frac{n}{d} d\lambda = \cos \alpha d\alpha + \cos \beta d\beta , \quad (8.34)$$

which, for a monochromatic source, i.e., $d\lambda = 0$, gives

$$d\beta = -\frac{\cos \alpha}{\cos \beta} d\alpha . \quad (8.35)$$

In physical coordinates, x , the size of the image of the slit, w' , at the detector

focal plane a distance f_{CM} from the camera, follows from

$$dx = f_{\text{CM}} d\beta, \quad (8.36)$$

giving,

$$w' = f_{\text{CM}} |d\beta| = f_{\text{CM}} \frac{\cos \alpha}{\cos \beta} d\alpha = f_{\text{CM}} \frac{\cos \alpha}{\cos \beta} \frac{w}{f_{\text{T}}} \frac{D}{L}, \quad (8.37)$$

where Eqs. 8.35 and 8.33 have been substituted for $d\beta$ and $d\alpha$, respectively.

The conversion of the physical coordinates, x , of the slit image on the focal plane (detector) into dispersion coordinates, λ , yields

$$\Delta\lambda_s = w' \frac{d\lambda}{dx} = w' \frac{d\lambda}{d\beta} \frac{d\beta}{dx} = \frac{w'}{f_{\text{CM}}} \frac{d\lambda}{d\beta}, \quad (8.38)$$

where Eq. 8.36 has been substituted for $d\beta/dx$. For fixed α , we have the angular dispersion from Eq. 8.34,

$$\frac{d\lambda}{d\beta} = \frac{d}{n} \cos \beta, \quad (8.39)$$

which, after substitution into Eq. 8.38, and substituting Eq. 8.37 for w' , gives the spectral purity

$$\Delta\lambda_s = \frac{w}{f_{\text{T}}} \frac{D}{L} \frac{d}{n} \cos \alpha = \frac{\theta_w}{\omega_c} \frac{D}{L} \frac{d}{n} \cos \alpha = \lambda \frac{\theta_w}{\omega_c} \frac{D}{L} \left(\frac{\cos \alpha}{\sin \alpha + \sin \beta} \right), \quad (8.40)$$

where θ_w is the angular projection of the slit on the sky in units of arc seconds, $\omega_c \simeq 206,265$ is the conversion factor for radians to arc seconds, and we have invoked the grating equation (Eq. 8.3) to show the wavelength dependence of $\Delta\lambda_s$. Note that our treatment does not provide insight into the functional shape, or distribution function, of the spectral purity, only the full-width half-maximum, $\Delta\lambda_s$.

Applying Eq. 8.27, the resolution power of the spectrograph is the ratio of the grating equation (Eq. 8.3) to the spectral purity (Eq. 8.40), giving the general expression,

$$R = \frac{\lambda}{\Delta\lambda_s} = \frac{\omega_c}{\theta_w} \frac{L}{D} \left(\frac{\sin \alpha + \sin \beta}{\cos \alpha} \right). \quad (8.41)$$

Clearly, resolving power increases with increasing grating size, L , and narrower angular projection of the slit width on the sky, θ_w . As such, resolving power improves for longer f_{T} for a fixed slit with w . Note that larger diameter telescopes actually reduced the resolving power; however, f_{T} usually scales in proportion to D .

The resolving power is not strictly constant with wavelength. As we have seen, for blazed gratings, we obtain $\beta = \alpha$ for the blazed wavelength, λ_n for

all orders n . However, from the grating equation, for which we consider α and d to be fixed quantities for a given spectrograph, we see that wavelengths dispersed $\pm\delta\lambda$ about the blaze wavelength are associated with changes in β according to

$$\delta\lambda = \frac{d}{n} \cos \beta \delta\beta = \frac{\lambda_n \cos \beta \delta\beta}{\sin \alpha + \sin \beta}. \quad (8.42)$$

Thus, from Eqs. 8.41 and 8.42, we have

$$\frac{\delta R}{R} = \frac{\cos \beta \delta\beta}{\sin \alpha + \sin \beta} = \frac{\delta\lambda}{\lambda_n}. \quad (8.43)$$

Over the free spectral range of order n , given by Eq. 8.7, we obtain

$$\frac{\delta R}{R} = \frac{\delta\lambda}{\lambda_n} = \frac{2n}{n^2 - 1}. \quad (8.44)$$

For long-slit, lower-order spectrographs, $\delta R/R$ can vary by as much as factor of 1.3 ($n = 2$). However, for high orders, such as those seen in echelle spectrographs, we see that $\delta R/R$ is roughly 0.01–0.03 across the free spectral range of the order. We can thus approximate a constant R across a given echelle order. Further, since $\beta = \alpha$ for the blaze wavelengths for all orders, we can write

$$R \simeq \frac{\omega_c}{\theta_w} \frac{L}{D} \tan \alpha, \quad (8.45)$$

approximating R as a constant (within a few percent) for all wavelengths dispersed by echelle spectrographs. Note that if the seeing disk of the source does not fill the slit width, then θ_w should be replaced with the atmospheric seeing disk, Θ_λ . Typically, echelle spectrographs on 10-meter class telescopes have $L \simeq 1$ meter and echelle angles of $65^\circ \leq \alpha \leq 70^\circ$, yielding $40,000 \leq R \leq 60,000$ for a $1''$ slit.

8.8 Instrumental Spread Function

Whereas the spectral purity (Eq. 8.40) quantifies the full-width half-maximum of the image of the slit at each monochromatic wavelength, the intensity distribution of each monochromatic image depends on small imperfections in the light path, such as off-axis optical elements, errors in facets cuts along the grating, gravity warping of optical elements, etc., that result in optical aberrations (Schroeder, 1987). The intensity distribution can also change due to variable seeing, poor centering of a point source on the slit (and therefore variable tracking), and wavelength dependent atmospheric dispersion of the seeing disk (e.g., Murphy et al., 2001).

In practice, the intensity distribution of this wavelength redistribution function is called the instrumental spread function (ISF) or the line spread function (LSF). We will adopt ISF and denote the function as $\Phi(\Delta\lambda; \Delta\lambda_s)$, which is evaluated at the wavelength difference $\Delta\lambda = \lambda' - \lambda$, where λ' is the wavelength of the redistributed light originally incident at wavelength λ . The normalization is

$$\int_{-\infty}^{\infty} \Phi(\Delta\lambda; \Delta\lambda_s) d(\Delta\lambda) = 1. \quad (8.46)$$

The ISF is often measured and published by the spectrograph commissioning team. For an ideal spectrograph, the ISF can be written as a univariate normal (Gaussian) distribution centered on wavelength λ ,

$$\Phi(\Delta\lambda; \Delta\lambda_s) = \frac{1}{\sqrt{2\pi}\Delta\sigma_s} \exp \left\{ -\frac{(\lambda' - \lambda)^2}{2(\Delta\sigma_s)^2} \right\}, \quad (8.47)$$

where $\Delta\lambda_s = 2\sqrt{2 \ln 2} \Delta\sigma_s$. In general, a spectrograph will not have an ISF given by Eq. 8.47. Examples include the Cosmic Origins Spectrograph (COS, Green et al., 2012) and the Space Telescope Imaging Spectrograph (STIS Woodgate et al., 1998) from the *Hubble Space Telescope* era, for which the ISF has asymmetric wings that are more extended than those of a normal distribution (Ghavamian et al., 2009; Kriss, 2011). However, the ISF for the High-Resolution Echelle Spectrometer on Keck-I (HIRES, Vogt et al., 1994) and for the ESO UV-visual Echelle Spectrograph on VLT-2 (UVES, Dekker et al., 2000) are both well approximated by a univariate normal distribution.

The spectral purity, $\Delta\lambda_s$, as given by Eq. 8.40, is often called the resolution element. Absorption (or emission) lines with widths much narrower than $\Delta\lambda_s$ will, after having passed through the spectrograph, be recorded with the shape closely approximating $\Phi(\Delta\lambda; \Delta\lambda_s)$ with full-width half-maximum $\Delta\lambda_s$. Such features are called unresolved lines. As such, if one measures the full-width half-maximum of an absorption line and it is consistent with resolution element, $\Delta\lambda_s$, then the line is considered to be “unresolved”. If the measured absorption line full-width half-maximum is substantially broader than the resolution element, then the line is considered to be “resolved”. The distinction is not so clear when the measured full-width half-maximum of the line is only slightly broader than the resolution element, though the term “partially resolved” is applicable.

Note that the full-width half-maximum of an unresolved line is proportional to the observed wavelength in the spectrum, $\Delta\lambda_s \propto \lambda$ (see Eq. 8.40).

8.8.1 Convolution of the ISF

When wavelength variations in the incoming flux, f_λ , are much broader than $\Delta\lambda_s$, then discernable post-optics modifications will not be apparent in the recorded flux. For example,

if the spectral energy distribution is smooth or flat across a wavelength range much broader than $\Delta\lambda_s$, then the local and global *shape* of the spectrum is virtually unaltered by the ISF (though the monochromatic wavelengths are blended on scales of $\Delta\lambda_s$).

However, an absorption (or emission) feature with a full-width half maximum on the order of or narrower than $\Delta\lambda_s$ will undergo a non-negligible flux redistribution across the line profile. This is especially true in the wings of the profiles. As such, this is why the ISF is also referred to as the line spread function, LSF, because an absorption line profile of the form,

$$f_\lambda = f_\lambda^0 \exp\{-\tau_\lambda\}, \quad (8.48)$$

will be measurably modified by the optics of the spectrograph.

Consider an absorption profile, as given by Eq 8.48, prior to its entering or passing through the optics of the spectrograph. We will denote f_λ as the “pre-optics flux”. Following passage through the spectrograph optics, the absorption profile will be recorded as the “post-optics flux”, which we will denote as \tilde{f}_λ . The relationship between the pre-optics flux and the post-optics flux is given by the convolution integral,

$$\tilde{f}_\lambda = f_\lambda * \Phi(\Delta\lambda; \Delta\lambda_s) = \int_0^\infty f_\lambda(\lambda') \Phi(\lambda' - \lambda; \Delta\lambda_s) d\lambda', \quad (8.49)$$

where $\Delta\lambda_s$ is given by Eq. 8.40, which, as we described above, is the full-width half-maximum of the redistribution of wavelengths due to the slightly different angles of the light incident on the collimator from opposite edges of the slit.

In Figure 8.10(a), we show an example of the pre-optics flux, $f_\lambda/f_\lambda^0 = \exp\{-\tau_\lambda\}$, of two closely spaced Ly α absorption lines. In Figure 8.10(b), we illustrate selected univariate normal distributions representing the ISF of various spectrographs, plotted versus $\lambda' - \lambda$. From upper to lower, they have spectral resolutions, $R = 20,000, 9000, 4500, 2250$, and 450 . The resolution elements are $\Delta\lambda_s = \lambda/R$, where $\lambda = 1215.67 \text{ \AA}$. The post-optics flux, $\tilde{f}_\lambda/f_\lambda^0$ is presented in Figure 8.10(c) for each spectrograph resolution. Note that the $R = 450$ panels cover a more extended wavelength scale. Vertical dot-dot lines provide a guide for the eye to the location of the wavelength extent of the pre-optics absorption profile. At lower resolution, the absorption lines are both blended and unresolved; the post-optics profile closely follows the

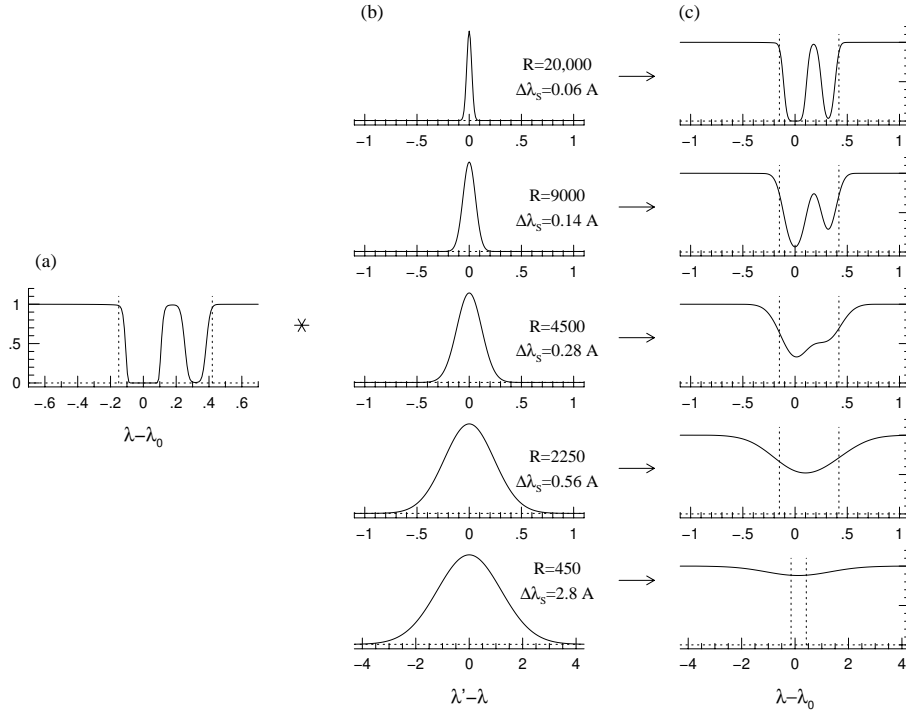


Figure 8.10 Illustrations of the degradation in spectral purity via convolution of the ISF for five somewhat typical spectroscopic resolutions, $R = 20,000, 9000, 4500, 2250$, and 450 (decreasing from the upper to lower panels). (a) The pre-optics flux illustrating two $\text{Ly}\alpha$ absorption lines. The profile is plotted versus $\lambda - \lambda_0$, where $\lambda_0 = 1215.67$ Å. (b) ISF profiles for the five resolutions, with R and $\Delta\lambda_s$ labeled for each. The ISFs are plotted versus $\lambda' - \lambda$, where λ' is the wavelength to which wavelength λ is redistributed. The ISF profiles are unit area normalized, though they are illustrated with identical peak amplitudes. (c) The post-optics flux for each resolution versus $\lambda - \lambda_0$. Vertical lines at $\Delta\lambda = -0.15$ and 0.42 provide wavelength range of the pre-optics profile.

shape of the ISF and has a full-width half-maximum equal to that of the ISF. As resolution increases, the profile shape of the post-optics flux approaches the profile shape of the pre-optics flux. Further, the full-width half-maximum of the higher-resolution profiles approach those of the pre-optics flux.

For this example, at $R = 20,000$, the post-optics profile is well resolved, where as some blending of the lines sets in for $R = 9000$. At $R = 4500$, we see the tall-tale sign of component blending in the asymmetry of the post-optics profile. By $R = 2250$, the blending is total, though the line center of the post-optics profile is shifted slightly relative to the stronger component in the pre-optics profile. For $R = 450$, which has $\Delta\lambda_s$ roughly a factor of

ten greater than the width of the pre-optics multi-component absorption profile, the absorption is redistributed to such a high degree that the profile is virtually wiped out.

As also apparent from Figure 8.10, even for a saturated profile, where the pre-optics flux is zero in the core of the profile, the post-optics flux may not reach zero, depending upon the resolution. For $R \leq 9000$, note that the profile cores are non-zero. Furthermore the level of recovery of the continuum between the two components is highly dependent upon resolution. Not until $R \geq 20,000$ is there any real hope of measuring or confidently constraining the intrinsic properties of these two absorption lines.

8.9 Pixelization

The charged couple device (CCD) is the commonly favored detector for astronomical spectrographs. A CCD is effectively a two-dimensional array of independent semi-conductors, known as pixels. We define the pixel sizes as B_x and B_y in the dispersion and spatial directions x and y along the detector (see Figure 8.2). Typically, $B_x = B_y$ on the order of 15–20 microns (μm).

8.9.1 Pixel Plate Scale

The pixel plate scale is the angular projection on the sky imaged by each pixel on the CCD. It is usually express in arc seconds pixel^{-1} . The pixel plate scale is the product of the physical pixel width, the plate scale at the telescope focal plane, and the magnification factor of the telescope image at the slit due to the spectrograph optics.

As introduced in § 8.3, $\theta_T = 1/f_T$ is the plate scale in the telescope image plane (at the slit), where f_T is the effective focal length of the telescope. The plate scale at the detector, in the focal plane of the camera, is given by the magnification, w'/w , in the dispersion direction, and h'/h in the spatial direction, giving,

$$\theta_x = \frac{w'}{w} \theta_T = \frac{\omega_c}{f_T} \frac{f_{CM}}{f_{CL}} \frac{\cos \beta}{\cos \alpha}, \quad \theta_y = \frac{h'}{h} \theta_T = \frac{\omega_c}{f_T} \frac{f_{CM}}{f_{CL}}, \quad (8.50)$$

where $\omega_c = 206,265$ is the conversion from radians to arc seconds. To convert the above plate scales from arc seconds mm^{-1} to arc seconds pixel^{-1} , we multiply by B_x for the dispersion direction and B_y for the spatial direction

and normalize length units. The pixel plate scales are then,

$$\begin{aligned}\theta_{x,\text{pix}} &= 10^{-3} B_x \theta_x = 10^{-3} B_x \frac{\omega_c}{f_T} \frac{f_{\text{CM}}}{f_{\text{CL}}} \frac{\cos \beta}{\cos \alpha}, \\ \theta_{y,\text{pix}} &= 10^{-3} B_y \theta_y = 10^{-3} B_y \frac{\omega_c}{f_T} \frac{f_{\text{CM}}}{f_{\text{CL}}},\end{aligned}\tag{8.51}$$

where the factor 10^{-3} cancels the length units between focal lengths in millimeters and pixel sizes in microns.

A well designed telescope/spectrograph facility will have a pixel plate scale on the order of $1/3$ or less of the best seeing disk of a point source. The extent of a typical seeing disk depends upon the accuracy of the guiding and the typical atmospheric seeing conditions. If the best conditions, for example, provide a seeing disk of $0.6''$, then a pixel scale of $\simeq 0.2'' \text{ pixel}^{-1}$ would be a maximum acceptable value. A pixel plate scale somewhat smaller than $\simeq 0.1'' \text{ pixel}^{-1}$ is typical for ground-based facilities. CCD pixels can be binned during or after readout. When binning, the plate scale must be corrected by the number of pixels per binned pixel.

8.9.2 Dispersion per Pixel

The process of pixelization is mathematically equivalent to multiplying the post-optics flux with a series of box functions of width B_x centered on each pixel,

$$B_x \text{III}(x) = B_x \sum_{k=0} \delta(x - kB_x/2), \tag{8.52}$$

where $\delta(0) = 1$ and is otherwise null, k is an integer, and x is the physical coordinate on the CCD in the dispersion direction. $\text{III}(x)$ is known as the Shah function. Each box function has spatial width B_x and is centered on each pixel.

The spectrograph camera focuses the post-optics spectrum such that each pixel independently collects flux at wavelength λ over a range $\Delta\lambda_{\text{pix}}$. Effectively, the flux is being binned into discrete wavelength intervals. We will refer to the resulting spectrum the post-optics “recorded flux”.

How the physical pixel widths, B_x , map to wavelength intervals depends upon the spectrograph configuration, according to Eqs. 8.36 and 8.39, giving

$$\Delta\lambda_{\text{pix}} = B_\lambda = B_x \frac{d\lambda}{dx} = B_x \frac{d \cos \beta}{n \frac{d\beta}{dx}}. \tag{8.53}$$

Since β is λ dependent via the grating equation (Eq. 8.3), the λ coordinate

will not map linearly with x nor will $\Delta\lambda_{\text{pix}}$. The mapping of the wavelength dispersion to the physical location of the pixels can be represented as $B_x III(x) \mapsto B_\lambda III(\lambda)$ yielding \AA pixel^{-1} .

The sampling rate of the spectrum is described by the number of pixels per resolution element,

$$p = \frac{\Delta\lambda_s}{\Delta\lambda_{\text{pix}}} = \frac{\theta_w}{\omega_c} \frac{D}{L} \frac{f_{\text{CM}}}{B_x} \frac{\cos \alpha}{\cos \beta}, \quad (8.54)$$

where we have invoked Eq. 8.40 for $\Delta\lambda_s$. Note that p is not necessarily an integer. Since p is the number of pixels per resolution element, p pixels sample the full-width half-maximum of the ISF. For similar reasons that the resolving power is effectively constant with wavelength for echelle spectrographs, the pixels per resolution element are effectively constant, even though $\Delta\lambda_{\text{pix}} \propto \lambda$.

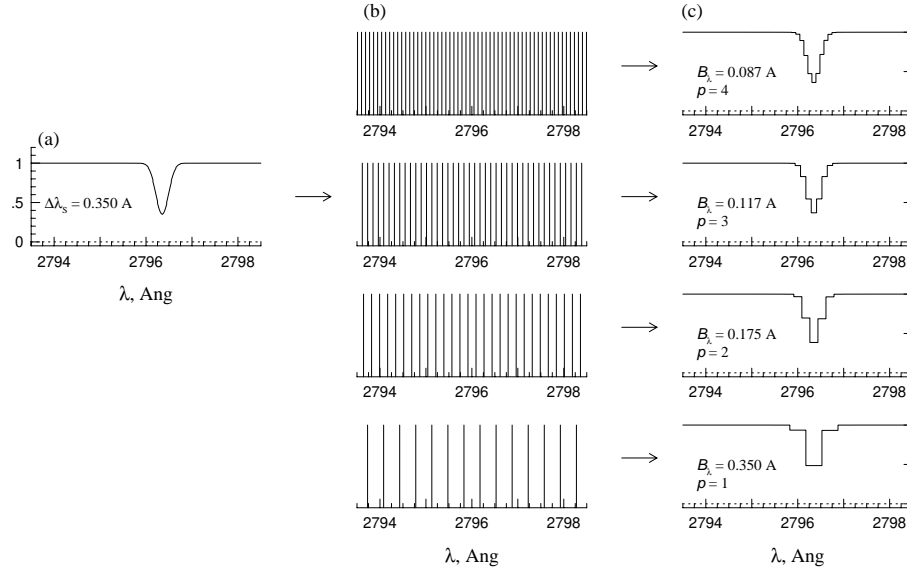


Figure 8.11 (a) A post-optics MgII $\lambda 2796$ absorption line profile from a spectrograph with $R = 8000$. The line is unresolved. (b) Four example Shah functions, Eq. 8.52, following the mapping to $B_\lambda III(\lambda)$. From lower to upper panel, these correspond to the pixel sampling rates $p = 1, 2, 3$, and 4 , where p is the number of pixels per resolution element. (c) The recorded flux for each p .

In Figure 8.11(a), we present a post-optics unresolved MgII $\lambda 2796$ absorption line from a spectrograph with $R = 8000$. In Figure 8.11(b), we present Shah functions for various B_x following the mapping to $B_\lambda III(\lambda)$. The recorded flux is presented in Figure 8.11(c) for each Shah function. From lower to upper panels, the sampling rate is $p = 1, 2, 3$, and 4 . For the

illustrated unresolved absorption line, p pixels are sampling the full-width half-maximum of the post-optics absorption profile.

Designing a spectrograph to provide a sampling rate of $p = 3$ is ideal, as this provides just enough sampling to recover the basic shape of the ISF for unresolved lines and to robustly measure the depth of line core. Raising the sampling to $p > 4$ is considered “over sampling”, as this results in the light being dispersed to such a degree that Poisson noise, scattered light, sky, and electronic noise in the detector system result in a reduced signal-to-noise ratio. A pixel sampling rate of $p = 2$ is known as “critically sampled”. In other words, it represents the minimum sampling rate in which any integrity knowledge of the instrumental spread function is preserved in the recorded spectrum. However, for $p = 2$, the shape of the pixelized absorption line is highly sensitive to where the line centroid falls with respect to the $B_{\lambda III}(\lambda)$ centroids. For presentation purposes, the Shah functions in Figure 8.11 have been centered on the post-optics absorption centroid.

8.10 Spectroscopic Data

In Chapter 7, we derived the relationship between the emission properties of a source and its observed flux density, f_{λ} . Previously in this chapter, when discussing modifications to degradation in the spectral purity of the flux as formalized in the instrumental spread function, ISF, we made the distinction between the pre-optics flux and the post-optics flux, the latter denoted \hat{f}_{λ} . To this stage, we have not incorporated modifications to the observed flux that result from atmospheric attenuation, the various reflective and transmission efficiencies of the telescope and spectrograph optical elements, and the CCD detector. In this section, we cover those topics and also discuss the relationship between the observed flux and the counts in the recorded spectrum, the uncertainty in the recorded spectrum, and flux calibration (converting from the recorded spectrum back to the observed flux). For additional introductory, yet comprehensive, resources on astronomical spectroscopy using CCDs, the reader is referred to the works of Gilliland (1992) Massey & Jacoby (1992), Wagner (1992), Churchill (1995), Martinez & Klotz (1998), and Howell (2000).

Though we do not discuss wavelength calibration of spectra, we do emphasize that, for cosmological absorption lines studies, the mapping from pixels to wavelengths is always based on vacuum wavelengths corrected to the heliocentric velocity reference system. This is simply accomplished by using vacuum line lists for the calibration lamps and then, following the application of the wavelength mapping, applying the heliocentric velocity

correction (e.g., Stumpff, 1977, 1980). These steps ensure uniformity of the cosmological redshifts across all studies.

Consider the light path beginning just above the atmosphere, where one would observe the true observed flux (including cosmological and Galactic absorption). The observed flux first passes through the atmosphere, where it suffers wavelength dependent attenuation in proportion to the path length through the atmospheric layer. Once the light enters the telescope, the attenuated flux passes through or reflects off (1) the telescope optics, (2) the slit, (3) the collimator, (4) the grating, and (5) the camera. At this point this post-optics flux is pixelized and recorded by the CCD, which counts the total number of photons focused into each pixel. Thus, the data product at the telescope is a two-dimensional “CCD frame” with photon counts at pixel dispersion positions x and spatial (or cross-dispersion) positions y .

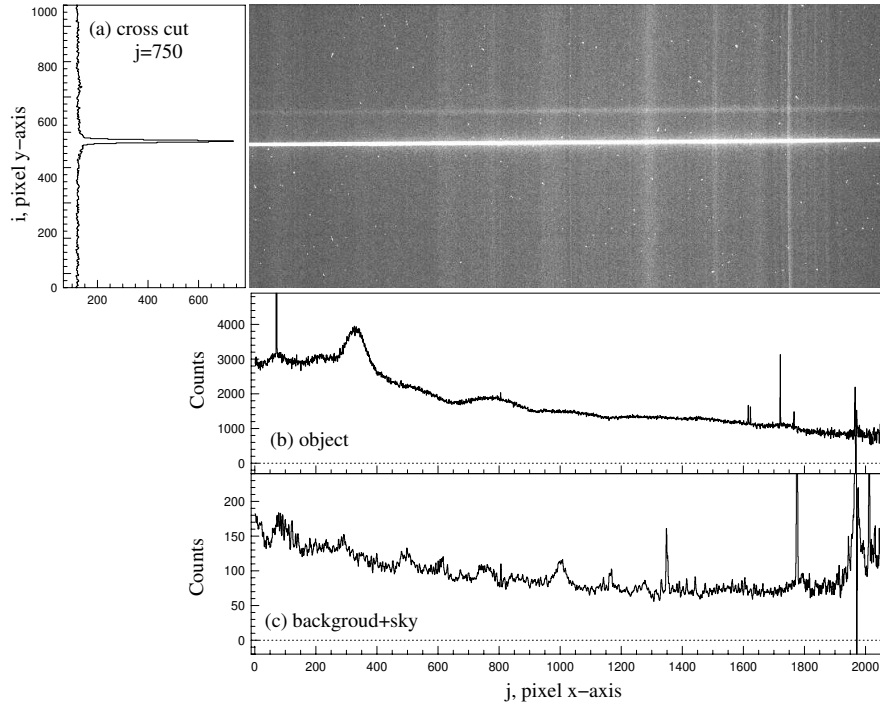


Figure 8.12 A two-dimensional CCD frame of a long-slit spectrum of an extragalactic point source showing photon counts at each pixel location. Comparing directly to Figure 8.2, the dispersion direction, x_j , is horizontal, and the spatial direction, y_i , is vertical. (a) A spatial cross-cut of the CCD frame at pixel $j = 750$. (b) The one-dimensional extracted object spectrum in units of recorded counts. (c) The spectrum of the background light including sky emission lines in units of recorded counts.

In Figure 8.12 (top right), we show the two-dimensional CCD frame of

a long-slit (low-resolution) spectrum of a point source. The desired one-dimensional spectrum must be extracted from the two-dimensional distribution of counts, which can be done using optimized algorithms (e.g., Horne, 1986; Marsh, 1989). Such algorithms optimally weight the counts in each pixel, minimizing statistical noise and maximizing photometric accuracy. Proper extractions yield three spectra: the object spectrum (panel b), the sky or background spectrum (panel c), and the uncertainty spectrum. The uncertainty spectrum (not shown in Figure 8.12, but discussed further in § 8.10.3) plays a critical role in the analysis of the spectra (see Chapter 10).

8.10.1 Atmospheric Attenuation and Throughput

Due to atmospheric attenuation and the wavelength dependent throughput efficiency of the telescope and spectrograph, the spectral energy distribution of the post-optics flux will be significantly modified from the spectral energy distribution of the observed flux (recall, we defined the observed flux as the flux of the object just before it enters the atmosphere). Consider the modifications to the observed flux, f_λ , as we follow the light path. Starting just above the atmosphere, we first have modification via attenuation/transmission of the atmosphere as quantified by the transmission ratio $\epsilon_\lambda^A(\zeta)$, where ζ is the angle between the telescope zenith and the source. The path length through the atmosphere relative to the zenith is measured as air-mass, $\sec(\zeta)$, yielding $\epsilon_\lambda^A(\zeta) = \epsilon_\lambda^A(0) \sec(\zeta)$ for $\zeta \leq 60^\circ$ where a plane-parallel atmosphere remains a good approximation. Then, along the light path, the observed flux is further modified by the reflective efficiency and/or transmission of the telescope optics, ϵ_λ^T , light loss at the slit (see Filippenko, 1982), ϵ_λ^S , the collimator, ϵ_λ^C , the grating and its intensity and blaze function, ϵ_λ^G , and the camera, ϵ_λ^O , yielding the post-optics flux

$$\tilde{f}_\lambda = \epsilon_\lambda^O \epsilon_\lambda^G \epsilon_\lambda^C \epsilon_\lambda^S \epsilon_\lambda^T \epsilon_\lambda^A(\zeta) f_\lambda = \epsilon(\lambda) \epsilon_\lambda^A(\zeta) f_\lambda, \quad (8.55)$$

where the mathematical operations occur from left to right along the light path, and where $\epsilon(\lambda) = \epsilon_\lambda^O \epsilon_\lambda^G \epsilon_\lambda^C \epsilon_\lambda^S \epsilon_\lambda^T = \tilde{f}_\lambda / \epsilon_\lambda^A(\zeta) f_\lambda$ is the throughput of the telescope plus spectrograph instrument, and is defined as the ratio of the post-optics flux to the attenuated flux. In fact, each wavelength dependent efficiency is defined as the ratio of the flux following interaction with the optical element to the flux before interaction and is a function ranging between 0 and 1.

A special note is required for the grating reflective efficiency, ϵ_λ^G , as it is dominated by the intensity function, which is given by Eq. 8.21 for low-

order spectrographs and by Eq. 8.26 for echelle spectrographs, as a function of wavelength according to Eq. 8.22.

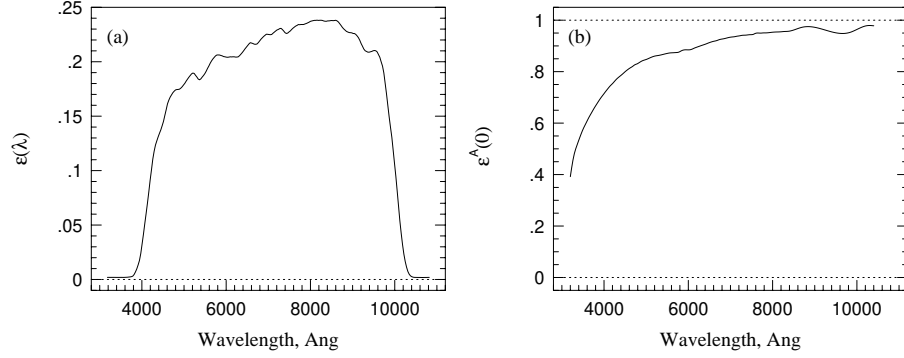


Figure 8.13 (a) A typical telescope throughput curve, $\epsilon(\lambda)$, for a hypothetical ground-based astronomical telescope and low-order optical spectrograph. Note that, at best, the post-optics flux, \tilde{f}_λ , reaching the detector is less than 25% of the attenuated flux, $\epsilon_\lambda^A(\zeta)f_\lambda$. (b) The atmospheric transparency, $\epsilon_\lambda^A(0)$, for unit airmass (toward the zenith).

A typical telescope throughput curve, $\epsilon(\lambda)$, is illustrated in Figure 8.13(a). A representative atmosphere attenuation curve for unity airmass is shown in Figure 8.13(b). The plotted curve is the atmospheric transmission, $\epsilon_\lambda^A(0)$, which is the ratio of the flux entering the telescope to the observed flux incident on the atmosphere (see Hamuy et al., 1992). Note the drop in transmission toward bluer wavelengths.

The functional form of $\epsilon(\lambda)$ is specific to the telescope and spectrograph. The functional form of $\epsilon_\lambda^A(0)$ is dependent upon the elevation of the observatory and can be affected by both atmospheric turbulence and the concentration of water vapor along the sightline to the object (quantities which are night to night variable). The observatory usually publishes average curves, where the atmospheric attenuation is for unity airmass averaged over many “photometric” nights (meaning negligible water vapor and minimal atmospheric turbulence).

On small wavelength scales, the flux also suffers degradation of its spectral purity following refraction off the grating and focusing by the camera optics. Thus, in the regions of absorption lines, i.e., $f_\lambda = f_\lambda^0 \exp\{-\tau_\lambda\}$, we must account for the ISF,

$$\tilde{f}_\lambda = \Phi(\Delta\lambda; \Delta\lambda_s) * \epsilon(\lambda) \epsilon_\lambda^A(\zeta) f_\lambda^0 \exp\{-\tau_\lambda\}. \quad (8.56)$$

As we will discuss in Chapter 9, the convolution of the ISF most strongly affects the profiles of unresolved and partially resolved absorption lines; in the

case of resolved absorption features, ISF convolution most strongly modifies the wings of the lines.

8.10.2 The Recorded Spectrum: Flux to Counts

The recording of the post-optics flux focused into a given pixel on the CCD consists of counting the *number* of incoming photons striking the pixel over some fixed time interval. Photons are counted by “summing” the number of electrons liberated in each pixel of the semi-conducting material. We must account for the fact that photons of various wavelengths liberate electrons with varying efficiency. We thus introduce the wavelength dependent quantum efficiency,

$$\epsilon_{\lambda}^{\text{QE}} = \frac{\text{number of electrons liberated}}{\text{number of incident photons}}, \quad (8.57)$$

in units counts photon⁻¹. The quantum efficiency is calibrated and published as a function of wavelength by the spectrograph design team or the manufacturer of the CCD and is usually made available by the observatory. Since each pixel is an independent semi-conducting detector, there are small pixel to pixel variations in the quantum efficiency. The published quantum efficiency curve is the average of all pixels comprising the CCD array.

To obtain the counts in each pixel, we must know the rate at which photons are incident on the pixel, which we then multiply by the time over which the photons are collected, i.e., the integration time, Δt (seconds). The rate of photons incident on a given pixel of width B_{λ} at wavelength λ is the product of the pixel width, the post-optics photon number flux, $d\tilde{n}_{\lambda}/dt$, and the collecting area of the telescope A_{T} . The photon number flux is the post-optics flux divided by the photon energy, $d\tilde{n}_{\lambda}/dt = \tilde{f}_{\lambda}/(hc/\lambda) = \epsilon(\lambda)\epsilon^{\text{A}}(\zeta)f_{\lambda}/(hc/\lambda)$. The rate of photons incident on a pixel is then

$$\frac{dN_{\lambda}}{dt} = B_{\lambda} \frac{d\tilde{n}_{\lambda}}{dt} A_{\text{T}} = B_{\lambda} \epsilon(\lambda) \epsilon^{\text{A}}(\zeta) f_{\lambda} \frac{\lambda}{hc} A_{\text{T}}, \quad (8.58)$$

which has units photons s⁻¹ pixel⁻¹. Multiplying this rate by the quantum efficiency of the CCD, we obtain the count rate per pixel,

$$\dot{I}_{\lambda} = \epsilon_{\lambda}^{\text{QE}} \frac{dN_{\lambda}}{dt} = \epsilon_{\lambda}^{\text{QE}} B_{\lambda} \epsilon(\lambda) \epsilon^{\text{A}}(\zeta) f_{\lambda} \frac{\lambda}{hc} A_{\text{T}}, \quad (8.59)$$

in units electrons $\text{s}^{-1} \text{ pixel}^{-1}$. The counts per pixel¹ are then

$$I_\lambda = \dot{I}_\lambda \Delta t. \quad (8.60)$$

The telescope and spectrograph efficiencies, and the observation specific terms appearing in Eq. 8.59 and Eq. 8.60, are typically combined into the so-called sensitivity function,

$$\mathcal{S}_\lambda = \epsilon_\lambda^{\text{QE}} B_\lambda \epsilon(\lambda) \epsilon_\lambda^{\text{A}}(\zeta) \frac{\lambda}{hc} A_\tau \Delta t, \quad (8.61)$$

which has units counts $\text{\AA cm}^2 \text{ sec erg}^{-1}$, and where the convolution with the ISF is implied (since it has negligible effect except in unresolved and partially resolved absorption and emission lines). Note that the azimuth angle, ζ , has been explicitly included in the atmospheric transmission efficiency to emphasize that the sensitivity function is airmass dependent (and thus specific to the observation). In terms of the sensitivity function, the counts are simply expressed, $I_\lambda = \mathcal{S}_\lambda f_\lambda$.

In the regions of absorption lines, we have $f_\lambda = f_\lambda^0 \exp\{-\tau_\lambda\}$, and we must explicitly account for the ISF redistribution of flux in the line,

$$I_\lambda = \Phi(\Delta\lambda; \Delta\lambda_s) * \mathcal{S}_\lambda f_\lambda^0 \exp\{-\tau_\lambda\}. \quad (8.62)$$

8.10.3 Uncertainties in Recorded Counts

It is vital that the uncertainty in the recorded counts is accurately quantified in order that uncertainties in quantities measured from analysis of the spectra can be tracked. These uncertainties are key to determining the confidence levels to which science can be gleaned from the spectra. The uncertainty in the recorded counts must not only account for the Poisson distribution of photon counts due to the object flux, but also due to the sky and other “unwanted” sources of light that creep into the system. In addition, the very act of collecting and counting the photons through digital devices controlled by analog electronics introduces counting errors.

The counting of the electrons in each pixel requires transferring the electrons across the CCD into a “register”, measuring the electric charge in the register, and then quantizing this charge into an integer number of electrons. This imperfect process results in several electronically induced errors, one of which is known as the read noise, RN. For most all astronomical detectors

¹ Note that the counts derived here are in units of electrons pixel^{-1} . Most facilities deliver data frames with the counts in digital number, DN. The relationship between electron counts and DN is $I = g \cdot \text{DN}$, where g is the gain of the CCD controller and thus applies to all pixels. The gain is adjustable and is typically between a few to several electrons per DN.

and their controlling electronics, the read noise is on the order of a few to ten electrons.

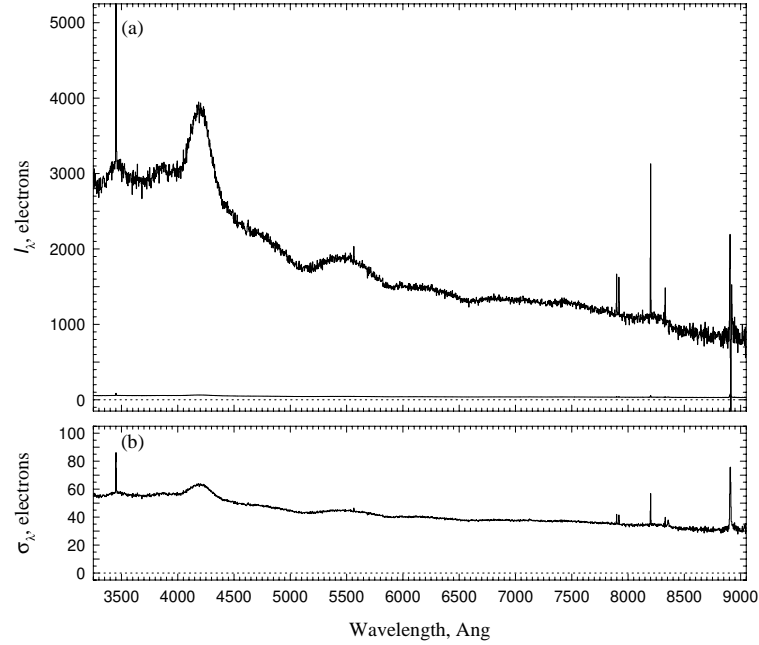


Figure 8.14 (a) The wavelength calibrated object spectrum and the uncertainty spectrum for the data frame shown in Figure 8.12. Both are shown in units of electron counts. (b) An expanded view of the uncertainty spectrum based upon Eq. 8.63.

A second important source of uncertainty in the recorded counts of the target object spectrum is background counts. There are three main backgrounds: the sky counts, dark counts, and scattered light counts. An example of the background counts is shown in Figure 8.12(c). Sky counts are due to both diffuse scattered light in the atmosphere and sky emission lines. The sky count rate per pixel at wavelength λ will be denoted \dot{S}_λ , and it is computed by knowing the flux in the sky lines and the surface brightness (and therefore flux) from the diffuse sky light and applying Eq. 8.59 with S_λ replacing f_λ . The dark current is due to thermally generated electrons in the CCD pixels. Each pixel has an individual and measurable rate at which dark current is generated, and we denote this rate corresponding to counts recorded at wavelength λ as \dot{D}_λ . Scattered light can often be negligible, but when it is not, scattered light can be directly corrected in the data (see Gray, 1974) or a smooth scattered light model can be obtained. In the case

of echelle spectra, the model may need to be fairly detailed and specific to the spectrograph (e.g., Gehren & Ponz, 1986; Churchill & Allen, 1995; Howk & Sembach, 2000).

For object and background counts, the noise model is Poissonian. Assuming that the counts in each exceed $\simeq 20$, the error distribution in the counts is a normal distribution and can be estimated as the square root of the counts. The uncertainty in the object spectrum counts is the quadrature sum of the additive and multiplicative mathematical operations required to extract the spectrum. For the extracted object counts, the uncertainty is then,

$$\sigma_{I_\lambda}^2 = \dot{I}_\lambda \Delta t + \dot{S}_\lambda \Delta t + \dot{D}_\lambda \Delta t + \text{RN}^2, \quad (8.63)$$

where we have omitted scattered light. The dark count rate is easily obtained using a series of “dark frames”, in which the CCD is integrated for various exposure times with the camera shutter closed. The slope of the dark counts versus integration time provides \dot{D}_λ . The remaining background, $S_\lambda = \dot{S}_\lambda \Delta t$, an example of which is illustrated in Figure 8.12(c), is usually dominated by the sky counts and is obtained during the extraction of the spectrum from the two-dimensional data frame².

In Figure 8.14, illustrate the uncertainty in the counts for the object shown Figure 8.12(b) and the background and sky counts shown in Figure 8.12(c). For simplicity, we have assumed $\dot{D}_\lambda = 0$ and a typical read noise of $\text{RN} = 5$ electrons. The signal-to-noise ratio, SNR, in each pixel is given by $\text{SNR} = I_\lambda / \sigma_{I_\lambda}$.

8.10.4 Flux Calibration

Flux calibration is the process of determining the observed flux from the measured counts via the relation

$$I_\lambda = \mathcal{S}_\lambda f_\lambda, \quad (8.64)$$

where \mathcal{S}_λ is the sensitivity function given by Eq. 8.61. As can be seen from the terms appearing in the sensitivity function, the telescope throughput and collecting area, the pixelization of the spectra, the airmass at the time of the observation, and the exposure time on the object all play important roles.

In Figure 8.15, the relationship between the observed flux, sensitivity function, and recorded counts is illustrated for a hypothetical telescope and

² We remind the reader that CCD data are stored as DN. For the Poisson noise model to be correct and for the units to be consistent with the read noise, the DN must be converted to electron counts by multiplying by the system gain.

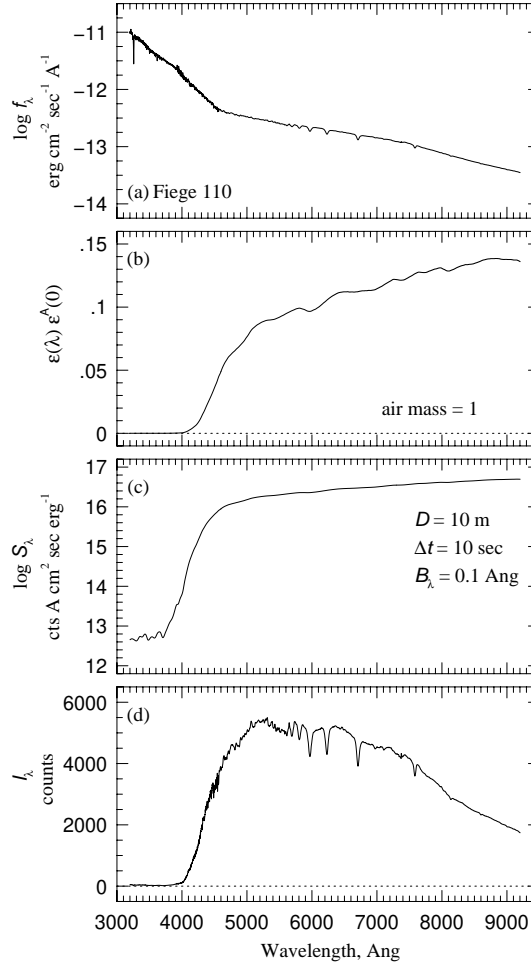


Figure 8.15 (a) The observed flux, f_λ , of the star Feige 110. (b) The total throughput, $\epsilon(\lambda) \epsilon_\lambda^A(0)$, based upon the hypothetical telescope and optical spectrograph throughput shown in Figure 8.13(a) and the atmospheric throughput curve presented in Figure 8.13(b) for unit airmass. (c) The sensitivity curve, S_λ , for a $D = 10$ meter diameter telescope for a $\Delta t = 10$ second exposure for a spectrograph with $B_\lambda = 0.1 \text{ \AA pixels}^{-1}$ at the detector. (d). The recorded counts per pixel, I_λ .

spectrograph facility. For this illustration, we have assumed a telescope diameter of $D = 10 \text{ m}$, an exposure time of $\Delta t = 10 \text{ s}$, and a spectrograph with pixel sampling rate $B_\lambda = 0.1 \text{ \AA pixel}^{-1}$. The observed flux of the star Feige 110 (Oke, 1990)³ is presented in Figure 8.15(a). The total throughput,

³ The electronic data were obtained from the online archive of optical and UV spectrophotometric flux standard stars made available by the European Southern Observatory: www.eso.org/sci/observing/tools/standards/spectra/

plotted in Figure 8.15(b), is the product of the telescope throughput and atmospheric attenuation (for unit airmass) shown in Figures 8.13(a) and 8.13(b), respectively. The sensitivity function, \mathcal{S}_λ , for the assumed observational particulars, is given in Figure 8.15(c). The total recorded counts per pixel, given by Eq. 8.64, is shown in Figure 8.15(d). Note how the drastic reduction in throughput below 4000 Å for this example results in negligible counts even though in this wavelength regime the observed flux is 1.5–2 orders of magnitude higher than where the total throughput is at its highest.

The sensitivity function is determined by observing so-called flux standard objects, usually bright stars. Flux standard stars are observed under ideal conditions for which the telescope and spectrograph sensitivity functions are well calibrated. Public lists of these stars and their spectral flux densities are often made available by the observatory (there are many papers of published flux standard stars, but a few good resources include Massey et al., 1988; Oke, 1990; Turnshek et al., 1990; Bohlin et al., 2001, and references therein).

If the science objectives are simply to measure the properties of absorption lines, then flux calibration is not required because absorption is quantified relative to the continuum; the sensitivity function cancels, i.e.,

$$\frac{I_\lambda}{I_\lambda^0} = \frac{\tilde{f}_\lambda}{\tilde{f}_\lambda^0} = \exp\{-\tilde{\tau}_\lambda\}. \quad (8.65)$$

However, if properties of the source spectrum itself are part of the science goals, then flux calibration may be necessary.

Flux calibration with echelle spectra is more challenging than calibration of lower dispersion spectra. Due to the echelle format and the free spectral range of the different orders of the grating, there can be duplicate coverage of the bluer wavelength regions. Since these overlap regions in wavelength appear in different orders of the grating, they have separate blaze functions (Eq. 8.11). For a given duplicated wavelength on orders n and $n + 1$, the value efficiencies, ϵ_λ^c , may differ by a factor of a few so that flux calibration may yield non-matching flux values without especially careful treatment (e.g. Suzuki et al., 2003)

Absorption Lines and Ionization Breaks

Observational astronomical spectra are not only a direct recording of the observed emitted flux of a distant source as a function of wavelength, f_λ , but also of the absorption, scattering, and emission that modifies f_λ due to material along the light path to the observer.

The utility of astronomical absorption spectra is that the observed flux can be employed to infer the physical conditions at the emitting source or of the gaseous medium in the proximity of the source. More commonly, the spectra are used to study gas cosmologically isolated and intervening to the emitting source. As such, high quality spectra hold great potential for revealing the chemical, ionization, structural, kinematic, and thermal conditions of the absorbing gas over a large range of cosmological epochs. The process of determining these properties involves formulating a physical model to express τ_λ as a function of both the atomic absorbers and the gas properties and then constraining the gas properties as accurately as possible from the observed spectrum.

In practice, most absorption features are spectrally isolated in wavelength so that the absorption coefficient appearing in the optical depth is based upon a single atomic or ionic state. For example, for any of the Lyman series transitions, the single atomic state being probed is neutral hydrogen with its electron in the ground state. As such, the column density measured from these absorption lines is the column density of ground state neutral hydrogen only and the atomic cross section applied in the optical depth model for each absorption line corresponds to its respective transition.

Thermal motions of the absorbing atoms/ions and large-scale dynamic motions introduce Doppler shifts along the line of sight that can significantly alter the strength of absorption lines or broaden their profile shapes. If we are to infer the dynamical conditions of the absorbing gas from an observed spectrum, our physical model of the optical depth must incorporate these

kinematic processes via the inclusion of what we will call “wavelength redistribution functions”. These functions contain free parameters describing the gas physics that can then be constrained from the spectra.

In this chapter, we briefly describe the basics of formulating an optical depth model for single spectrally isolated absorption lines and for ionization breaks. We introduce the basic formalism for incorporating physically motivated wavelength redistribution functions into the optical depth model. This mathematical process involves convolution from which the most common functional form of the optical depth is the well-known Voigt profile. As one of the central goals of absorption line analysis will be to measure the column densities of the absorbing atoms and ions, we then discuss the most basic measurement of the absorption line strengths: the equivalent width. The behavior of the equivalent width as a function of column density is known as the curve of growth, which we discuss in some detail. We also describe the behavior of the equivalent width for cosmologically redshifted absorption lines. Finally, we discuss ionization breaks in spectra and how column densities can be measured from these spectral features.

9.1 The Challenge of Spectral Analysis

Spectroscopic data provide the quantity f_λ , the observed flux density, as a function of wavelength. It is usually not difficult to obtain a reasonably accurate measurement of the continuum flux density, f_λ^0 , so that the ratio $f_\lambda/f_\lambda^0 = \exp\{-\tau_\lambda\}$ is measured.

A goal of analyzing cosmological absorption lines is to infer as much information as possible about the physical conditions in the absorbing gas, such as the kinematics and the chemical and ionization conditions. If all physical states (densities, excitation states, ionization levels, chemical abundances, line-of-sight velocities, etc.) of all absorbing atoms/ions are known, then the extinction and emission coefficients can be calculated and the observed spectrum fully predicted. Successfully predicting the observed spectra from the physical state of the gas is the goal of spectral analysis. In practice, however, the observed spectra provide an incomplete and partial view of the absorbing gas so that we can extract only limited constraints on the gaseous physical conditions.

The challenge of spectral analysis lies in the fact that the complex formalism of radiative transfer reduces to a single datum at each wavelength, i.e., the optical depth, τ_λ (see Eq. 7.42). The optical depth is an integrated quantity in which the line-of-sight geometry, dynamical state, and chemical and ionization conditions of the absorbing gas are converted into a single

number at each wavelength,

$$\tau_\lambda = \int_0^L \left[\sum_i n_i(s) \sigma_i(\lambda) \right] ds, \quad (9.1)$$

where the sum is taken over all absorbing ions, $n_i(s)$ is the number density of absorbing ion i as a function of location s through a cloud of line-of-sight depth L , and $\sigma_i(\lambda) = \sum_j \sigma_{ij}(\lambda)$ is the total absorption cross section for *all* absorption transitions j as a function of wavelength for ion i . As we will discuss in Chapter 11, the number density of ion i depends upon the chemical abundance of its atomic species and the fraction of that species that is in the appropriate ionization stage (the ionization fractions). Thus, not only is knowledge of the atomic absorption cross sections required for computing the optical depth, but detailed knowledge of the physical state of the absorbing gas is required.

In practice, a given absorption feature is often spectrally isolated in that it is the sole feature in a given wavelength range that is absorbing the continuum flux. For example, consider the Ly α 1215.67 Å absorption line from neutral hydrogen. The atomic cross section for absorption (Eq. 2.83) is shown in Figure 9.1(a). Note that the cross section has a full-width half-maximum of $\simeq 10^{-5}$ Å. It would be improbable, but not impossible, for another ion from a different species to be populated in an ionization stage that gives rise to absorption in this extremely narrow spectral region. Acknowledging that the radiative transfer across the spectral region of a given absorption line can be isolated to a single ion/transition, we can safely remove the sums from Eq 9.1, yielding

$$\tau_\lambda = \int_0^L n(s) \sigma(\lambda) ds = N \sigma(\lambda), \quad (9.2)$$

where we have dropped the subscript to denote ion i .

There is an additional complication, however. Line of sight Doppler velocity shifting due to the kinematics of the gas particles results in wavelength shifting of the absorption features; this induces absorption line broadening. To obtain the resulting optical depth, the atomic absorption cross section must be convolved with wavelength redistribution functions to obtain the total absorption cross section. Such functions may be due to the gas being in dynamic motions, such as turbulence, rotation, and winds, or simply due to thermal motions in an otherwise globally stationary medium. These redistribution functions are parameterized by quantities, such as temperature, rotation velocity, or wind speed. By using fitting techniques, such as least-squares χ^2 minimization, the observed absorption lines can be compared to

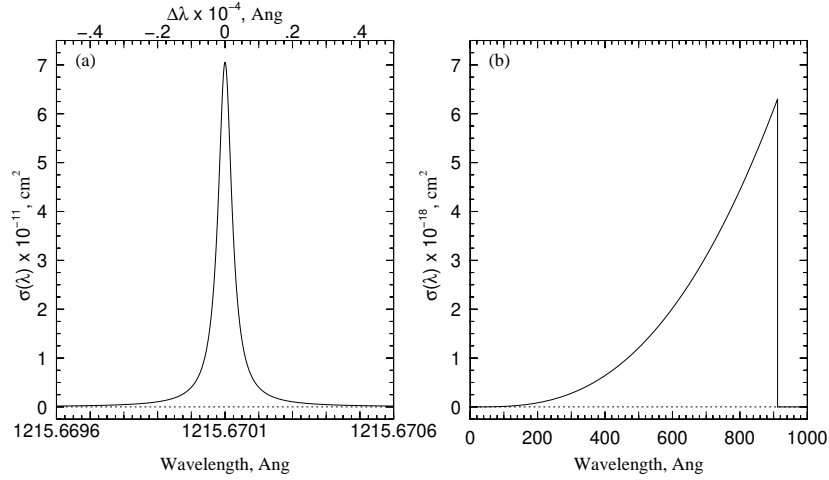


Figure 9.1 Two selected absorption cross sections for neutral hydrogen in the ground state. (a) The bound-bound absorption cross section for the Ly α λ 1215 transition, which peaks at 1215.6701 Å. (b) The bound-free absorption cross section for ionization from the ground state, which is possible only for photons having wavelengths less than $\simeq 912$ Å.

the parameterized optical depth model to provide additional insight into the physical nature of the absorbing gas.

With regard to continuum absorption due to ionization, the cross section for absorption creates a sudden and sharp decrease in f_λ/f_λ^0 , which is sometimes referred to as an “ionization break” or an “ionization edge”. The bound-free cross section for ground-state neutral hydrogen ionization (Eq. 2.91) is shown in Figure 9.1(b). Note that for wavelengths with energies lower than the ionization threshold, the cross section is null, but at the ionization energy it is a maximum. Thus, there is a discontinuity in the optical depth that yields the ionization break in the spectrum. As we shall show below, ionization breaks provide a powerful and direct means for measuring the column density of the ion.

Here, we provide a brief summary of key concepts incorporated in the analysis of astronomical spectra. We then discuss these ideas further in the remainder of this chapter and in following chapters.

- The number density of the absorbing ions, which depends on the abundance and ionization condition of the absorbing ion, and the path length through the absorbing cloud, which depends on physical size and orientation of the probed gas structure, are coupled together as the column density, N . The column density is a multiplicative factor in the optical

depth, and thus primarily governs the magnitude of τ_λ and therefore the strength of an observed absorption feature.

- Observed absorption features will never have the profile shapes of the atomic cross section, which depends only on the atomic physics of the absorbing ion. Gas physics influences and dominates the profile shape. This physics is parameterized in a “total” absorption cross section; the wavelength dependence of the total absorption cross section dictates the profile shape of the absorption feature.
- This total absorption cross section is obtained via convolution of the atomic cross section with one or more wavelength redistribution functions. These redistribution functions parameterize various gas properties. One common wavelength redistribution function is the thermal broadening function, for which the parameterized gas property is the gas temperature.
- As discussed in § 8.8, the spectrograph itself introduces a spectral “blurring” of the incoming light as described by the instrumental spread function. This instrumentally induced wavelength redistribution function must be incorporated to fully model an observed absorption profile. The instrumental spread function is *not* convolved with the atomic absorption cross section nor with the redistribution functions, as it is not part of the radiative transfer.
- To constrain the absorbing gas properties from the data, one quantitatively compares spectra to the predicted optical depths derived from ionization, chemical, and kinematic models of the gas. The key information in the spectra are the measured column densities and the absorption line profile shapes. The parameters in the wavelength redistribution functions can be adjusted to match the profile shapes, and the chemical and ionization conditions can be explored to match the column densities.

If the full suite of spectroscopic data are populated by absorption features from different transitions of several ionic stages of multiple atomic species (thus providing a sampling of a wide range of chemical and ionization conditions in the absorbing gas), then the column densities of various chemical elements in various ionization stages can be measured. As measured column densities are key for constraining the absorbing gas properties, the greater the number and variety of these measured column densities, the better the gas properties can be inferred; a process we describe in Chapter 12.

9.2 Absorption Lines

9.2.1 The Total Absorption Cross Section

In general terms, if there is redistribution of the absorbed wavelengths due to environmental physical processes that Doppler shift the absorbing atoms along the line of sight then the expression for the optical depth given by Eq. 7.85,

$$\tau_\lambda = \int_0^L n(s) \sigma(\lambda) ds = N \sigma(\lambda), \quad (9.3)$$

whose functional form with wavelength accounts only for the atomic absorption cross section of the transition, requires modification to account for the wavelength shifts.

This modification requires that the atomic cross section is convolved with a wavelength redistribution function describing the number density of absorbing atoms/ions at line of sight location s shifted by $\Delta\lambda$,

$$n(s, \Delta\lambda) ds d\lambda = n(s) f(s, \Delta\lambda; \mathbf{a}) ds d\lambda, \quad (9.4)$$

where $n(s)$ is the number density of absorbing atoms/ions at location $s \rightarrow s + ds$, $n(s, \Delta\lambda)$ is the distribution of the number densities of absorbing atoms/ions at location $s \rightarrow s + ds$ shifted in wavelength by $\Delta\lambda$ from the interval $\lambda \rightarrow \lambda + d\lambda$. Note that, the integral of the particles over all wavelength shifts in the location $s \rightarrow s + ds$,

$$n(s) ds = \int_0^\infty n(s, \Delta\lambda) d\lambda ds, \quad (9.5)$$

yields the number density in the interval $s \rightarrow s + ds$.

The functional form of the wavelength redistribution function, $f(s, \Delta\lambda; \mathbf{a})$, is unique to the physical processes giving rise to the wavelength shifting of the absorbing atoms/ions as parameterized by $\mathbf{a} = (a_1, a_2, \dots, a_N)$. For example, the parameters might include the gas temperature, T , the radiation or gas pressure, P_r or P_g , a turbulent velocity component, v_{turb} , or the luminosity of a local stellar population, L .

From Eq. 9.3,

$$\tau_\lambda = \int_0^L n(s, \Delta\lambda) * \sigma(\lambda) ds = \int_0^L n(s) [f(s, \Delta\lambda; \mathbf{a}) * \sigma(\lambda)] ds. \quad (9.6)$$

This general form can be simplified if we assume that the wavelength redistribution function is uniform throughout the gas structure, and therefore

independent of line of sight location s . We obtain

$$\tau_\lambda = (\sigma * f[\mathbf{a}])(\lambda) \int_0^L n(s) ds = N\alpha(\lambda; \mathbf{a}) \quad (9.7)$$

where $\alpha(\lambda; \mathbf{a})$ is the total absorption cross section,

$$\alpha(\lambda; \mathbf{a}) = (\sigma * f[\mathbf{a}])(\lambda) = \int_0^\infty \sigma(\lambda') f(\lambda' - \lambda; \mathbf{a}) d\lambda', \quad (9.8)$$

at wavelength λ . Computing $\alpha(\lambda; \mathbf{a})$ requires evaluating a series of integrals over λ' , each centered on λ . Visualize each integration of Eq. 9.8 as the process of centering $f(\Delta\lambda; \mathbf{a})$ at λ and integrating over λ' . The observed flux across the profile of a single feature in an astronomical spectrum can now be written in terms of the total absorption cross section,

$$f_\lambda = f_\lambda^0 \exp \{-N\alpha(\lambda; \mathbf{a})\}. \quad (9.9)$$

From Eq. 9.9, both the column density of the absorbing atomic species and the physical parameter vector \mathbf{a} can be constrained from the observed flux.

We note that, in principle, the redistribution function can be the convolution of an arbitrary number of redistribution functions,

$$f(s, \Delta\lambda; \mathbf{a}) = g(s, \Delta\lambda; \mathbf{b}) * h(s, \Delta\lambda; \mathbf{c}) * q(s, \Delta\lambda; \mathbf{d}) \cdots \quad (9.10)$$

where the parameters $\mathbf{a} = \mathbf{a}(\mathbf{b}, \mathbf{c}, \mathbf{d}, \cdots)$. For example, $g(s, \Delta\lambda; \mathbf{b})$ could be turbulent broadening with $\mathbf{b} = (v_{\text{turb}})$ the magnitude of the turbulent velocity, or $h(s, \Delta\lambda; \mathbf{c})$ could be rotational broadening with $\mathbf{c} = (v_{\text{rot}}, \sin i)$ the rotational velocity and the inclination of the rotation axis.

9.2.2 The Thermal Broadening Function

Broadening of absorption profiles due to the thermal motions of the absorbing atoms/ions is always present. If bulk kinematic motions dominate over thermal motions, then the thermal wavelength redistribution function will be insignificant compared to the redistribution due to the bulk kinematics. However, if thermal motions are comparable to or exceed any bulk kinematic motions in the gas, then the thermal redistribution function is very important; for typical temperatures of the interstellar, circumgalactic, and intergalactic medium, the observed absorption profiles will be broadened by a few orders of magnitude greater than the widths of the atomic absorption cross sections.

Consider a single absorbing ion in motion within an absorbing cloud. The line of sight component of the motion of this ion with respect to the source

results in a wavelength shifting of the atomic absorption cross section, $\sigma(\lambda)$, in the observer reference frame. Let the velocity vector of a given absorbing ion be \mathbf{v} . Assume that both the absorbing cloud and the observer are at rest with respect to the luminous continuum source. The observed line of sight samples the absorbing cloud in a single spatial direction; thus, the observed line-of-sight velocity shift is the dot product of the line of sight propagation direction of the beam, $\hat{\mathbf{s}}$, and the velocity vector of the atom, $\Delta v = -\mathbf{v} \cdot \hat{\mathbf{s}}$, where the negative sign provides the convention $\Delta v > 0$ for shifting toward the source and away from the observer and $\Delta v < 0$ for components away from the source and toward the observer.

In an isothermal parcel of gas of temperature T , if $n(v)$ is the number density of absorbing particles in the line-of-sight velocity interval $v \rightarrow v + dv$, and n is the total number density of absorbing particles integrated over all velocity intervals, then the fraction of particles with velocity difference Δv is given by $f_{\text{T}}(\Delta v; \Delta v_{\text{D}})dv = n(\Delta v)dv/n$, where

$$f_{\text{T}}(\Delta v; \Delta v_{\text{D}}) dv = \frac{1}{\sqrt{\pi}\Delta v_{\text{D}}} \exp \left\{ -(\Delta v/\Delta v_{\text{D}})^2 \right\} dv, \quad (9.11)$$

where $\Delta v_{\text{D}} = (2kT/m)^{1/2}$ is the Doppler velocity width for absorbing particles of mass m . As will be shown in § 11.2.6, the product kT is proportional to the mean non-relativistic kinetic energy of particles in a gas with temperature T . The distribution of particle velocities in any single direction (and subsequently, along any line of sight through the gas) is a simple Gaussian distribution.

The Doppler velocity width is commonly referred to as the Doppler b parameter. In Figure 9.2, we present the Doppler b parameter as a function of temperature for hydrogen, carbon, oxygen, magnesium, and iron. For a fixed temperature, the ratio of the Doppler b parameter of an ion with mass m_i to that of hydrogen is $b_i/b_{\text{H}} = (m_{\text{H}}/m_i)^{1/2}$. In Figure 9.3(a), we show the line-of-sight thermal velocity distribution functions of hydrogen for three temperatures typical of interstellar, circumgalactic, and intergalactic gas. The Gaussian standard deviation of these distributions is $\sigma(T) = b(T)/\sqrt{2}$. In order to illustrate the mass dependence, we show the same functions for hydrogen, carbon, magnesium, and iron for $T = 50,000$ K in Figure 9.3(b).

Via the Doppler shift formula, $\Delta\lambda/\lambda = \Delta v/c$, the line-of-sight temperature dependent dispersion of the redistribution of the wavelengths of the absorption cross section will be $\Delta\lambda_{\text{D}} = (\Delta v_{\text{D}}/c)\lambda_0$, where λ_0 is the rest-frame

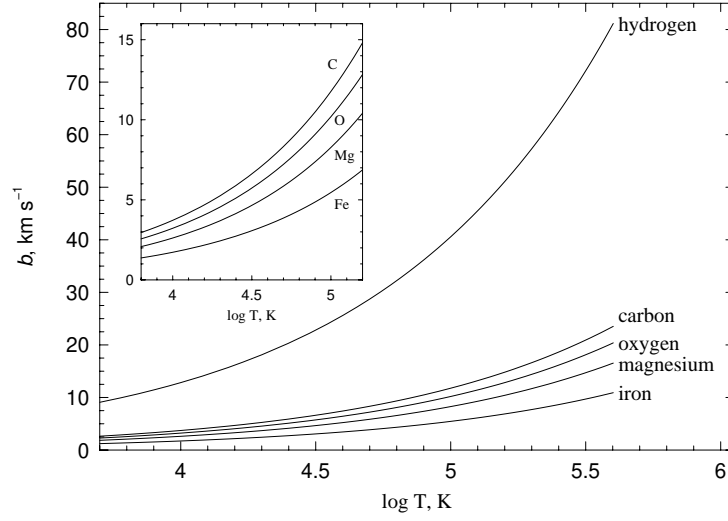


Figure 9.2 The observed line-of-sight thermal Doppler b parameter, computed from $b = \Delta v_D = (2kT/m)^{1/2}$ in km s^{-1} for hydrogen, carbon, oxygen, magnesium, and iron as a function of temperature.

wavelength of the transition. With $b = \Delta v_D$, we have

$$\Delta\lambda_D = \frac{b}{c}\lambda_0 = \frac{\lambda_0}{c} \left(\frac{2kT}{m} \right)^{1/2}, \quad (9.12)$$

which is known simply as the Doppler width. Transforming the line-of-sight thermal velocity distribution function, we obtain the line-of-sight wavelength redistribution function,

$$f_T(\Delta\lambda; \Delta\lambda_D) d\lambda = \frac{1}{\sqrt{\pi}\Delta\lambda_D} \exp \left\{ -(\Delta\lambda/\Delta\lambda_D)^2 \right\} d\lambda. \quad (9.13)$$

Selected examples of observed line-of-sight wavelength redistribution functions are illustrated in Figure 9.3(c) and 9.3(d). Because $\Delta\lambda_D$ is proportional to λ_0 , each distribution function is specific to a given absorption transition. In Figure 9.3(c), the observed distribution of $\Delta\lambda$ is shown for the Ly α $\lambda 1215$ neutral hydrogen transition at $T = 10,000, 25,000$, and $50,000$ K. In Figure 9.3(d), the observed distribution is shown for four transitions commonly observed in absorption from the interstellar and circumgalactic medium (see figure caption).

The Doppler widths of the thermal redistribution functions are a factor of a few $\times 10^4$ broader than the widths of the atomic absorption cross sections. As can be seen from Figures 9.3(c,d), we have $\Delta\lambda_D \sim 0.1 \text{ \AA}$, whereas the corresponding width of the atomic cross section is $\Gamma\lambda_0^2/4\pi c \simeq \text{few} \times 10^{-5} \text{ \AA}$

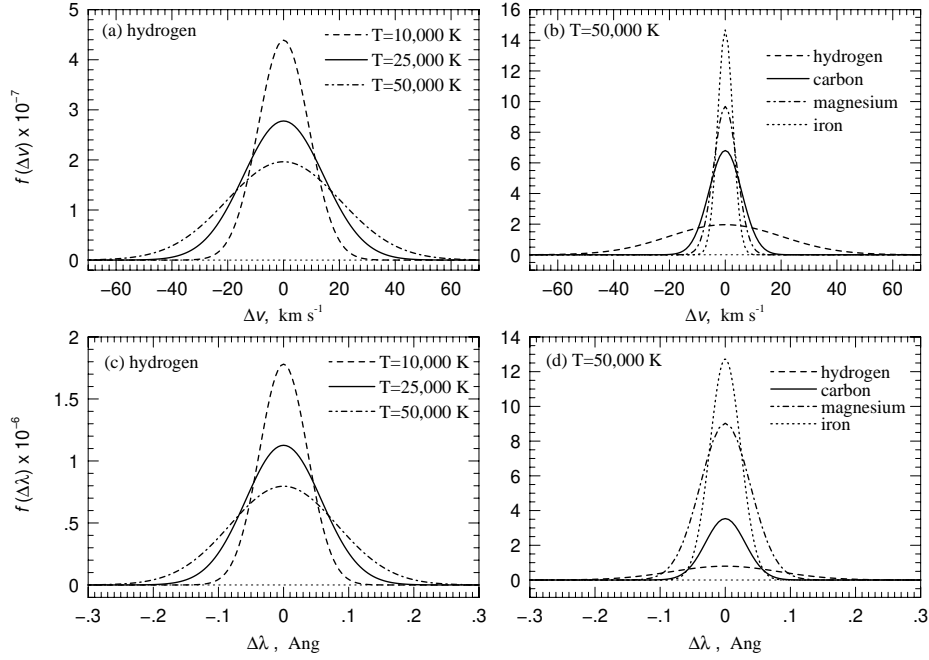


Figure 9.3 The observed line-of-sight thermal Doppler shift distribution of absorbing atoms/ions. (a) The distribution of Δv for the hydrogen at $T = 10,000, 25,000$, and $50,000$ K. (b) The distribution of Δv at $T = 50,000$ K for the selected transitions hydrogen, carbon, magnesium, and iron. (c) The wavelength redistribution function for the $\text{Ly}\alpha$ absorption line of hydrogen at $T = 10,000, 25,000$, and $50,000$ K. (d) The wavelength redistribution function at $T = 50,000$ K for the selected transitions of hydrogen ($\text{Ly}\alpha$), carbon ($\text{CIV } \lambda 1548$), magnesium ($\text{MgII } \lambda 2796$), and iron ($\text{FeII } \lambda 2600$).

(see Figure 9.1). Note that the amplitude of the line-of-sight wavelength thermal redistribution function scales with $\lambda_0^{-1}(T/m)^{-1/2}$ and the width scales as the inverse, $\lambda_0(T/m)^{1/2}$. Figures 9.3(c,d) clearly illustrate this behavior, which is influential in governing the profile shapes (depth and width) of thermally broadened absorption lines.

9.2.3 The Voigt Profile

If the absorption line profile shape is assumed to be dominated by thermal broadening, then the total absorption cross section, $\alpha(\lambda; \Delta\lambda_D)$, will be the convolution of the Lorentzian atomic absorption cross section, $\sigma(\lambda)$, and the line-of-sight Gaussian thermal redistribution of atoms, $Nf_T(\Delta\lambda; \Delta\lambda_D)$,

where N is the column density. Invoking Eq. 9.7, the optical depth is then

$$\tau_\lambda = N\alpha(\lambda; \Delta\lambda_D) = N(\sigma * f_T[\Delta\lambda_D])(\lambda). \quad (9.14)$$

The observed absorption profile will then have the shape

$$f_\lambda = f_\lambda^0 \exp\{-\tau_\lambda\}, \quad (9.15)$$

where f_λ is the observed flux density at wavelength λ and f_λ^0 is the continuum flux density (in the absence of the absorption line).

From the definition provided by Eqs. 9.7 and 9.8,

$$\tau_\lambda = N \frac{e^2 \lambda_0^2}{m_e c^2} f \int_0^\infty \frac{\Gamma \lambda_0^2 / 4\pi c}{(\lambda' - \lambda_0)^2 + (\Gamma \lambda_0^2 / 4\pi c)^2} \frac{\exp\{-(\lambda' - \lambda) / \Delta\lambda_D\}^2}{\sqrt{\pi} \Delta\lambda_D} d\lambda', \quad (9.16)$$

where λ_0 is the transition or central wavelength, f is the oscillator strength (see § 2.2.1 and § 2.3.1), and Γ is the damping constant (see § 2.5) for the transition.

The convolution of a Lorentzian and a Gaussian is a Voigt function, which has unit normalization over all wavelength. Eq. 9.16 can be written

$$\tau_\lambda = N \frac{\pi e^2 \lambda_0^2}{m_e c^2} f U(x_\lambda, y) \quad \text{where} \quad U(x_\lambda, y) = \frac{1}{\sqrt{\pi} \Delta\lambda_D} H(x_\lambda, y), \quad (9.17)$$

where $U(x_\lambda, y)$ is the Voigt function, and where the convolution is simplified to the Hjerting function,

$$H(x_\lambda, y) = \frac{y}{\pi} \int_{-\infty}^{\infty} \frac{\exp\{-t^2\}}{(x_\lambda - t)^2 + y^2} dt, \quad (9.18)$$

with normalized variables

$$x_\lambda = \frac{\lambda - \lambda_0}{\Delta\lambda_D} \quad \text{and} \quad y = \frac{\Gamma \lambda_0^2}{4\pi c} \frac{1}{\Delta\lambda_D}. \quad (9.19)$$

The Voigt function has become one of the most important tools for spectroscopic analysis and much effort has been focused on efficient evaluation of Eq. 9.18. Algorithms to compute the Hjerting function have exploited the fact that it is equivalent to the real part of the complex probability function, $H(x_\lambda, y) = \text{Re}[w(z)]$, where

$$w(z) = \exp\{-z^2\} \left(1 + \frac{2i}{\pi^{1/2}} \int_0^z \exp\{-t^2\} dt \right), \quad (9.20)$$

and where $z = x_\lambda + iy$. The power of this approach is that $w(z)$ can be expanded into a series that rapidly converges to high accuracy in as few as six terms and is easily implemented computationally (see, Humlíček, 1979; Kochanov, 2011; Boyer & Lynas-Gray, 2014).

Note that x_λ , being the difference between the wavelength along the profile and the line center in units of the Doppler width, is the independent variable. Also note that $y \simeq 10^{-5}/0.1 \simeq 10^{-4}$ is not a function of λ ; it is the half-width at half maximum of the Lorentzian in units of the Doppler width and is a constant for the absorption line. The Doppler width itself depends upon the gas temperature, the wavelength of the line center, and the mass of the absorbing atom/ion. Note that the Doppler width also appears in the normalization of the Voigt function in Eq. 9.17.

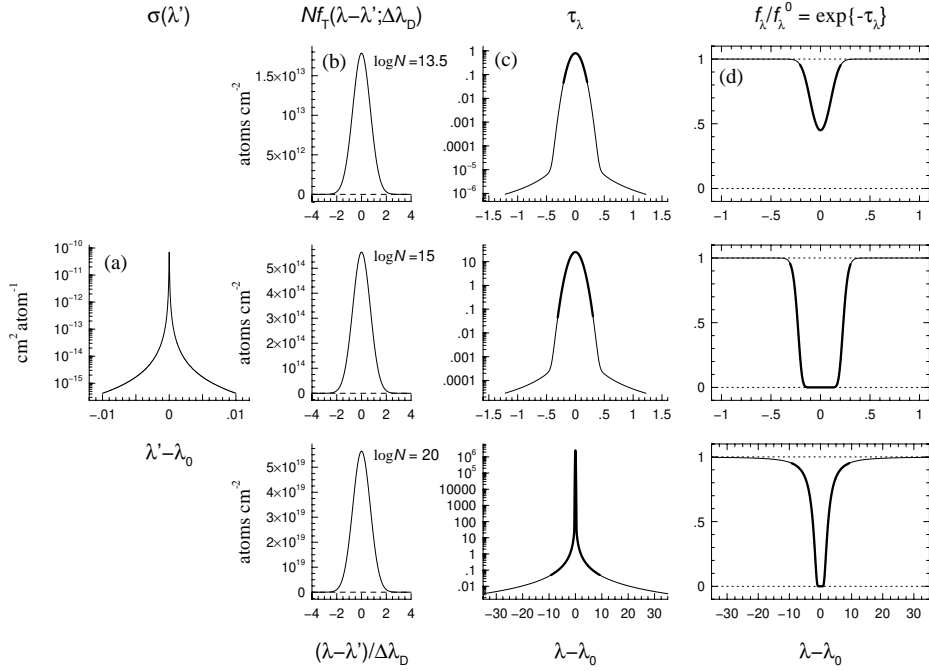


Figure 9.4 A schematic (moving from left to right) of the Voigt profile convolution process to obtain the observed absorption profile of the Ly α transition. (a) The atomic cross section $\sigma(\lambda')$. (b) The thermal redistribution function of neutral hydrogen atoms $N f_T(\lambda' - \lambda; \Delta\lambda_D)$ for $\log N = 13.5$, 15.0, and 20.0 (from top to bottom) for $b = 30 \text{ km s}^{-1}$ ($\Delta\lambda_D = 0.12 \text{ \AA}$). (c) Following convolution, the optical depth Voigt profile, τ_λ , for each respective column density. The thick portion of the curves correspond to $\tau_\lambda \geq 0.05$, the optical depth range at which $f_\lambda/f_\lambda^0 \leq 0.95$ and the absorption becomes discernible. (d) Application of Eq. 9.15 gives the absorption profile. The thick portion of the profiles correspond to the thick portions of the optical depth curves. See text for discussion.

An illustration of the convolution process for obtaining Eq. 9.15 is shown in Figure 9.4 for the Ly α transitions with $\log N = 13.5$, 15.0, and 20.0. The atomic cross section, $\sigma(\lambda')$, is shown in Figure 9.4(a). Note the ex-

treme narrowness of the profile; the width of the core is given by $\Gamma\lambda_0^2/4\pi c$. Note that the wings of the cross section scale as $(\lambda' - \lambda_0)^{-2}$. The thermal wavelength redistribution function of the neutral hydrogen column density, $Nf_{\tau}(\lambda' - \lambda; \Delta\lambda_D)$, is plotted in Figure 9.4(b) for three different total column densities, $\log N = 13.5, 15.0$, and 20.0 (upper to lower panels) assuming the same Doppler b parameter of $b = 30 \text{ km s}^{-1}$ ($\Delta\lambda_D = 0.12 \text{ \AA}$) for each (corresponding to $T \simeq 50,000 \text{ K}$). This redistribution function is Gaussian in shape and provides the line-of-sight column density distribution as a function of $\lambda' - \lambda$, where λ' is an arbitrary wavelength location along the atomic cross section, as is expressed in Eq. 9.16.

Convolution yields the optical depth, τ_{λ} , as shown in Figure 9.4(c). These curves are the Voigt profiles. For the three example column densities, the optical depth Voigt profiles all have the identical shapes because they all result from convolution of the same atomic cross section with the same wavelength redistribution function (where $b = 30 \text{ km s}^{-1}$). Though we show the profile over a much larger wavelength range for the $\log N = 20$ example, the profile shape is identical to those of the $\log N = 13.5$ and 15.0 examples; they differ only by their normalization, which is the total column density, N , of the neutral ion of hydrogen. Note that the cores of these profiles are Gaussian; however, for $|\lambda - \lambda_0| \geq 0.5 \text{ \AA}$, the Lorentzian wings of the atomic cross section dominate the Voigt profile. The exact transition from Gaussian core to Lorentzian wings depends on the Doppler b parameter; a larger (smaller) thermal broadening will increase (decrease) the width of the Gaussian core. In general the transition will occur approximately at $|\lambda - \lambda_0|/\Delta\lambda_D \simeq 4$.

Finally, the absorption profiles themselves are shown in Figure 9.4(d) for each example column density. For $\tau_{\lambda} \geq 0.05$, we obtain $f_{\lambda}/f_{\lambda}^0 \leq 0.95$, the spectral region of discernible absorption as indicated with the thicker curve. In Figure 9.4(c), we have marked thicker curves in the spectral regions of the optical depth Voigt profiles that map to the discernible absorption. We elaborate on the following important points illustrated in Figure 9.4:

- For $\log N = 13.5$, note that only a very narrow region of the core of the optical depth Voigt profile contributes to discernible absorption. The optical depths in the Gaussian wings and especially in the Lorentzian wings are too small to contribute to the absorption profile.
- For $\log N = 15.0$, note that the Gaussian core and a limited region of the Gaussian wings of the optical depth Voigt profile contributes to discernible absorption. The optical depth in the Lorentzian wings is still too small to contribute to the absorption profile. However, because discernible

absorption is coming from the wings of the Gaussian, at this larger column density, we see that the width of the absorption profile is directly dependent on the value of the thermal broadening, as a larger Doppler b parameter would yield a broader Gaussian core in the optical depth Voigt profile.

- For $\log N = 20.0$, the nature of the absorption changes such that the Lorentzian wings now have optical depth values large enough to give rise to discernible absorption. In fact, the spectral extent over which this is true is very broad relative to the Gaussian core because of the relatively flat $(\lambda - \lambda_0)^{-2}$ decline of the Lorentzian wings, even though there are six to seven orders of magnitude difference in the optical depth from the peak in the Gaussian core to the Lorentzian wings. In the absorption profile, the extended absorption is known as a damping wing in reference to the damping constant Γ appearing in the Lorentzian atomic absorption cross section.

Profiles such as those shown Figure 9.4(d) can be fit to Eq. 9.15 to objectively obtain the column density, N , and the only additional free parameter, the Doppler width, $\Delta\lambda_D$. This approach is useful only for high resolution data in which the line width induced by the instrument (see § 8.8) is significantly less than the Doppler width. Usually the fitting of the data is performed using the technique of χ^2 minimization. In high resolution spectra, the absorption lines often break up into multiple components, forming a complex and often blended profile shape. We will address this further complexity in Chapter 10.

9.2.4 Equivalent Width and the Curve of Growth

Because the line core “saturates” at $f_\lambda/f_\lambda^0 \simeq 0$, as optical depth increases, the increase in the flux density removed in the line core cannot be measured within the limitations of the noise in the data. We thus cannot quantify the “strength” of absorption directly from the ratio f_λ/f_λ^0 at the line center; we can only quantify the integrated quantity f_λ/f_λ^0 across the spectral region of detected absorption. This integrated quantity is called the equivalent width,

$$W = \int_0^\infty [1 - f_\lambda/f_\lambda^0] d\lambda = \int_0^\infty [1 - \exp\{-\tau_\lambda\}] d\lambda, \quad (9.21)$$

and the units are in wavelength, such as Å or nm. The equivalent width has a simple interpretation. It represents the width in wavelength of a hypothetical absorption feature with $f_\lambda = 0$ across the profile that yields an equivalent

W . Being an integrated quantity, W is independent of the absorption line shape and the resolution of the spectrograph.

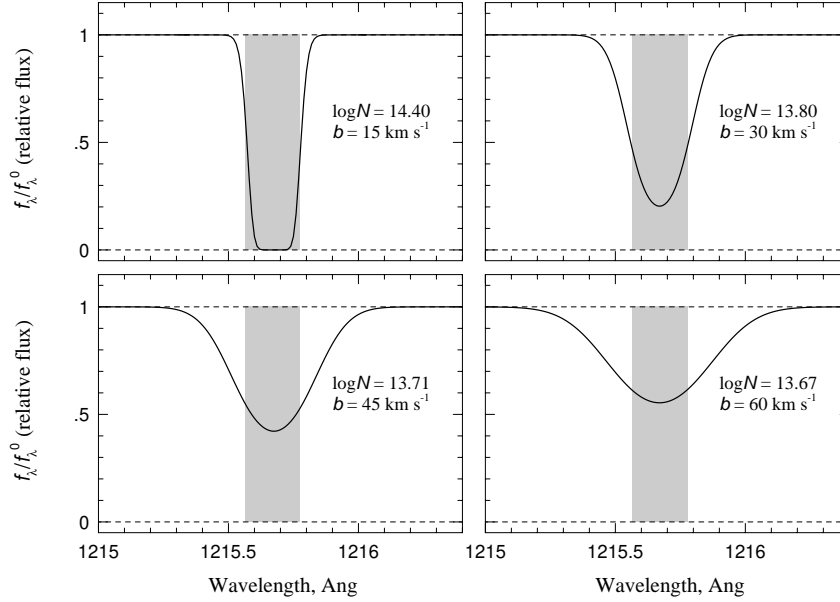


Figure 9.5 A schematic of four absorptions lines each with equivalent width $W = 0.21 \text{ \AA}$. Though the line profile shapes are quite different, the total amount of flux absorbed, as given by Eq. 9.21, are identical. The shaded grey area shows the interpretation of the equivalent width.

Assuming that the dominant physics incorporated into the total absorption cross section is the atomic natural broadening and thermal broadening, we have the Voigt profile $\alpha(\lambda; \Delta\lambda_D) = (\sigma * f_T[\Delta\lambda_D])(\lambda)$. Recalling that $b = (c/\lambda_0)\Delta\lambda_D$, we can write $\tau_\lambda = N\alpha(\lambda; b)$, which yields, $W = W(N, b)$.

A schematic illustrating the interpretation of the equivalent width is shown in Figure 9.5 for $W = 0.21 \text{ \AA}$. The shaded area represents the equivalent area removed by each absorption profile as computed using Eq. 9.21. This shaded area can be viewed as a hypothetical profile with zero flux over the wavelength interval W , which is the meaning of the term equivalent width. As can be seen in Figure 9.5, there can be a family of $N\alpha(\lambda; b)$ combinations that yield identical W , i.e., W is degenerate for various combinations of N , and $\alpha(\lambda; b)$.

The behavior of the equivalent width dependence on N and b is called the curve of growth. As the optical depth of the line increases, the equivalent width also increases, but not in linear proportion. As discussed above, for small optical depths, as the optical depth at the line core, τ_0 , increases,

the line depth increases until virtually all the photons at the line core are removed from the incoming beam. At this point, the absorption line is considered to be “saturated”. As τ_0 increases further, the detectable additional light removed from the beam occurs in the Gaussian wings of the Voigt profile, until a regime in which τ_0 is very large and the optical depth in the Lorentzian damping wings approaches unity (as illustrated in Figure 9.4(d) for $\log N = 20.0$, at these large τ_0 values, the majority of the light is removed very far from the line center).

These three regimes are called the “linear”, “flat”, and “damped” parts of the curve of growth, respectively. These are often also referred to as the “linear”, “logarithmic”, and “square root” parts due to their functional dependence upon the column density, N . Note that only in the logarithmic part will W be sensitive to the b parameter, and thus the Gaussian wings of the line broadening. In Figure 9.6, the curve of growth for the Ly α transition of H I is shown for $b = 30 \text{ km s}^{-1}$ as a function of τ_0 . The thick portions of the curves correspond to the flat, logarithmic, and square root parts of the curve of growth, based upon the slope of the curve on a log-log diagram. Insets show the Ly α absorption profiles for locations on the curve marked with filled circles.

We now examine the functional dependence of the curve of growth.

9.2.4.1 The Linear Regime

In the optically thin regime, the optical depth in the line core is small, $\tau_0 \ll 1$. This scenario corresponds to the linear part of the curve of growth as illustrated for profiles 1, 2, and 3 on Figure 9.6 and the $\log N = 13.5$ absorption line shown across the top row of Figure 9.4.

Using series expansion of the exponential, we can approximate $\exp(-\tau_\lambda) \simeq 1 - \tau_\lambda$. Substituting into Eq. 9.21, gives

$$W \simeq \int_0^\infty \tau_\lambda d\lambda = N \int_0^\infty \alpha(\lambda; \Delta\lambda_D) d\lambda. \quad (9.22)$$

Substituting Eq. 9.17 for $\alpha(\lambda; \Delta\lambda_D)$, we have

$$W \simeq \frac{\pi e^2 \lambda_0^2}{m_e c^2} N f \int_0^\infty U(x_\lambda, y) d\lambda = \tau_0 \int_0^\infty U(x_\lambda, y) d\lambda, \quad (9.23)$$

which defines τ_0 , where $U(x_\lambda, y)$ is the Voigt function given by Eq. 9.19. By definition the Voigt function has unity normalization, so the integral in Eq. 9.23 equals unity and we have the linear functional dependence,

$$W \simeq \frac{\pi e^2 \lambda_0^2}{m_e c^2} N f, \quad (9.24)$$

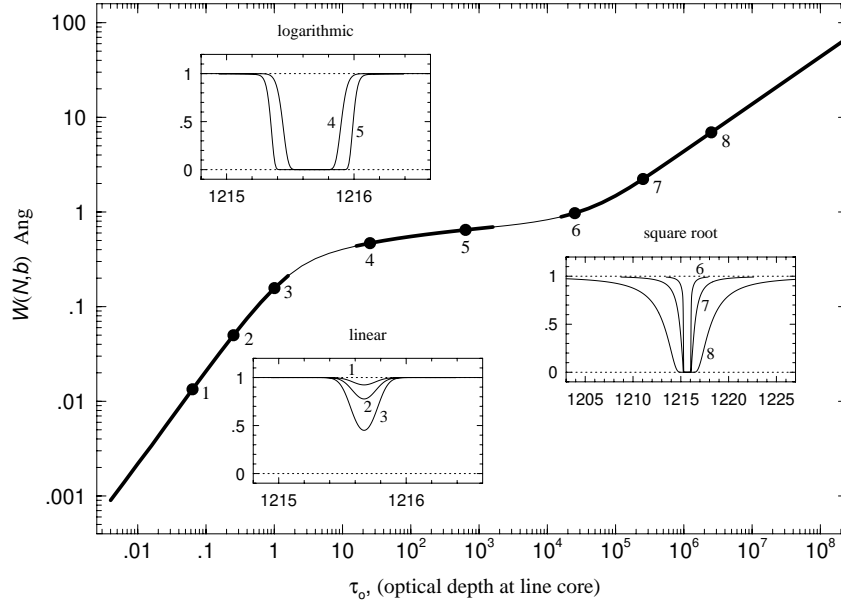


Figure 9.6 The curve of growth showing equivalent width as a function of the optical depth at line core, τ_0 for Ly α for $b = 30 \text{ km s}^{-1}$. The three regimes, “linear”, “logarithmic”, and “square root” are shown by the thick curves, respectively, as τ_0 increases. Absorption profiles are shown for each regime and their locations on the curve marked with filled points. Note the expanded wavelength scale for the profiles on the square root part of the curve of growth.

which is highly accurate for computing W .

Note that the shape of $U(x_\lambda, y)$ depends upon the Doppler width, $\Delta\lambda_D \propto b$, and thus the absorption line shape depends on the Doppler b parameter (gas temperature). However, W is an integrated quantity and, because the Voigt function has unity normalization, is independent of b in the regime of $\tau_0 \ll 1$. Physically, in the regime of small τ_0 , as more atoms are added to the absorbing gas, W grows by a deepening of the line core due to the removal of additional photons in the beam and we have $W \propto N$.

9.2.4.2 The Logarithmic Regime

In the regime where $10 \leq \tau_0 \leq 10^4$, the wings of the Gaussian thermal distribution dominate over the damping wings of the Lorentzian. This scenario corresponds to the flat part of the curve of growth as illustrated for profiles 4 and 5 on Figure 9.6 and the log $N = 15$ absorption line shown across the middle row of Figure 9.4.

We can thus approximate the Lorentzian contribution to the absorption

coefficient as a δ function and write,

$$\tau_\lambda = N\alpha(\lambda; \Delta\lambda_D) \simeq \frac{\pi e^2 \lambda_0^2}{m_e c^2} \frac{Nf}{\pi^{1/2} \Delta\lambda_D} \exp \left\{ -[(\lambda - \lambda_0) / \Delta\lambda_D]^2 \right\}. \quad (9.25)$$

Defining $\tau_0 = \pi^{1/2} e^2 \lambda_0^2 / (m_e c^2 \Delta\lambda_D) Nf$, the optical depth of the line core, and invoking $x = (\lambda - \lambda_0) / \Delta\lambda_D$, we simplify Eq. 9.25 to $\tau_\lambda = \tau_0 \exp(-x^2)$, and substituting into Eq. 9.21, approximate the equivalent width as

$$W \simeq \Delta\lambda_D \int_0^\infty \left[1 - \exp \left\{ -\tau_0 e^{-x^2} \right\} \right] dx = \Delta\lambda_D F(\tau_0). \quad (9.26)$$

In the “flat” part of the curve of growth, the line core is saturated, meaning that W does not grow due to removal of photons with small x . The function $F(\tau_0)$ provides the behavior of the line absorption strength in the regime where the width of the line is governed by a Gaussian broadening mechanism. As τ_0 is increased, the amplitude of the Gaussian increases and the wings of the line remove more flux from the beam. Physically, as more atoms are added to the gas, it is those in the tails of the Doppler velocity distribution that contribute to increasing W . Thus, as we shall see below, $W \propto b$ for a fixed N . For small τ_0 , a series expansion of Eq. 9.26 yields Eq. 9.24 for the linear part of the curve of growth. When $\tau_0 > 10$, the integral asymptotically approaches $F(\tau_0) \simeq (\ln \tau_0)^{1/2}$, which gives $W \simeq \Delta\lambda_D \sqrt{\ln \tau_0}$, yielding

$$W \simeq b \frac{\lambda_0}{c} \left[\ln \left(\frac{\pi^{1/2} e^2 \lambda_0}{m_e c} \frac{Nf}{b} \right) \right]^{1/2}. \quad (9.27)$$

where we invoked $b = (c/\lambda_0) \Delta\lambda_D$. We caution that Eq. 9.27 is not suitable for computing W as, compared to the evaluation of Eq. 9.26, there is a percent difference error of -8% for $\tau_0 = 10$, -5% for $\tau_0 = 100$, -4% for $\tau_0 = 1000$, and -3% for $\tau_0 = 10^4$. However, the general asymptotic behavior of the equivalent width in the regime of $10 \leq \tau_0 \leq 10^4$ is $W \propto b \sqrt{\ln(N/b)}$.

9.2.4.3 The Square Root Regime

In this regime, where $\tau_0 \geq 10^5$, the optical depth is large enough that the Lorentzian $(\lambda - \lambda_0)^{-2}$ broad wings dominate over the more rapidly declining wings of the thermal Gaussian line broadening. This scenario corresponds to the damped part of the curve of growth as illustrated for profiles 6, 7, and 8 on Figure 9.6 and the $\log N = 20$ absorption line shown across the bottom row of Figure 9.4.

In this case, the Gaussian contribution to the Voigt function can be treated

as if it were a δ function. Thus, we can approximate the optical depth as

$$\tau_\lambda = N\alpha(\lambda; \Delta\lambda_D) \simeq \frac{\pi e^2 \lambda_0^2}{m_e c^2} N f \frac{\Gamma \lambda_0^2 / 4\pi c}{(\lambda - \lambda_0)^2 + (\Gamma \lambda_0^2 / 4\pi c)^2}. \quad (9.28)$$

Due to the very narrow half-width at half maximum of the Lorentzian profile, i.e., $\Gamma \lambda_0^2 / 4\pi c$, the vast majority of the absorption is in the wings of the absorption line and the equivalent width is dominated by regions under the line profile that are far from the line center. Under these conditions, $|\lambda - \lambda_0| \gg \Gamma \lambda_0^2 / 4\pi c$, which leads to the approximation

$$\tau_\lambda \simeq \frac{\pi e^2 \lambda_0^2}{m_e c^2} \frac{\Gamma \lambda_0^2 / 4\pi c}{(\lambda - \lambda_0)^2} N f = \frac{\tau_0}{(\lambda - \lambda_0)^2}, \quad (9.29)$$

where we define $\tau_0 = (e^2 \Gamma \lambda_0^4 / 4m_e c^3) N f$. Substituting into Eq. 9.21,

$$W \simeq \int_0^\infty [1 - \exp \{-\tau_0 / (\lambda - \lambda_0)^2\}] d\lambda. \quad (9.30)$$

Letting $u^2 = (\lambda - \lambda_0)^2 / \tau_0$ with $du = \tau_0^{-1/2} d\lambda$, we have

$$W \simeq \tau_0^{1/2} \int_0^\infty [1 - \exp \{-1/u^2\}] du = (\pi \tau_0)^{1/2}. \quad (9.31)$$

Writing out τ_0 fully,

$$W \simeq \lambda_0^2 \left[\frac{\pi e^2 \Gamma}{4m_e c^3} N f \right]^{1/2}. \quad (9.32)$$

Note that $W \propto \sqrt{\Gamma N}$, giving this regime either the name “square root” or “damping” part of the curve of growth. Though fairly common in stellar spectra, in cosmological spectra the only observed damped absorption profiles arise from Ly α $\lambda 1215$ transitions with $\tau_0 > 10^6$. These absorption lines are common enough that they categorize as a population of absorption system in their own right. Originally thinking they arose exclusively from absorption in thick HI disks of high redshift galaxies, (Wolfe et al., 1986) defined damped Ly α absorbers (DLAs) to have $W \geq 5 \text{ \AA}$ and $b < 100 \text{ km s}^{-1}$, corresponding to $N_{\text{HI}} \geq 2 \times 10^{20} \text{ cm}^{-2}$, or $\log N_{\text{HI}} = 20.3$. We note that this column density corresponds to $W \simeq 8.8 \text{ \AA}$ independent of the b parameter.

9.2.5 Curve of Growth Methods

The above treatment for the behavior of the equivalent width as a function of the column density and Doppler b parameter is based on the assumption that the wavelength redistribution function is due to thermal broadening

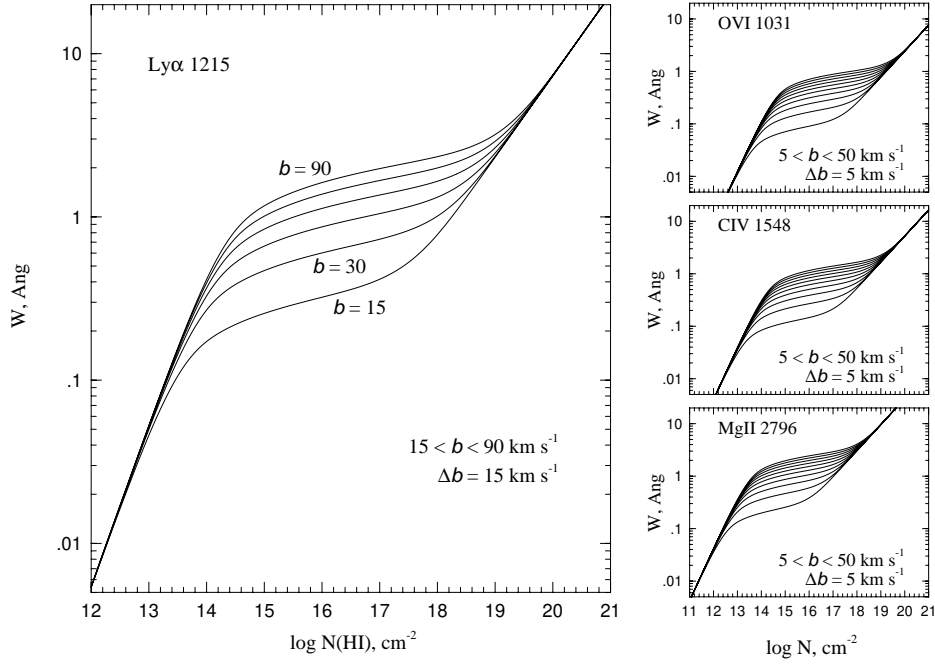


Figure 9.7 The curve of growth showing $W(N, b)$, the equivalent width as a function of the column density and Doppler b parameter for the transitions $\text{Ly}\alpha$ λ 1215 of neutral hydrogen, MgII λ 2796, CIV λ 1548, and OVI λ 1031.

only. For this limited yet essential treatment of the broadening of absorption lines, we obtained the general results,

$$\begin{aligned}
 \text{linear: } (\tau_0 \ll 1) \quad & \tau_0 = \frac{\pi e^2 \lambda_0^2}{m_e c^2} N f \quad & W \propto N \\
 \text{flat: } (10 \leq \tau_0 \leq 10^4) \quad & \tau_0 = \frac{\pi^{1/2} e^2 \lambda_0}{m_e c} \frac{N f}{b} \quad & W \propto b \sqrt{\ln(N/b)} \\
 \text{damped: } (\tau_0 \geq 10^5) \quad & \tau_0 = \frac{e^2 \Gamma \lambda_0^4}{4 m_e c^3} N f \quad & W \propto \sqrt{\Gamma N}.
 \end{aligned} \tag{9.33}$$

In Figure 9.7, we present the curve of growth as a function of column density for four commonly observed absorption line transitions. Multiple curves illustrate the dependence of W on the Doppler b parameter. Note that there is no b dependence on the linear and damped parts of the curve. On the flat part of the curve, larger b yields larger W , as the wings of the absorption profile are broader for larger b . Note, however, that as b increases, the curves converge. This behavior reflects the scaling $W \propto b \sqrt{\ln(N/b)}$. A

clear lesson from Figure 9.7, is that a measured equivalent width in the range $0.1 \leq W \leq 1.0 \text{ \AA}$ arises from absorption on the flat part of the curve of growth and that without some estimate of the Doppler b parameter, the column density is not well constrained over several orders of magnitude!

To learn about the gas physical conditions from observed absorption lines, we aim to determine the column density and the free-parameters in the wavelength redistribution function, which in the case of thermal broadening is the Doppler b parameter. If two members of a fine-structure double are observed, then the ratios of their equivalent widths can be used to estimate the column density of the absorbing ion. This is called the doublet ratio method. For doublets and especially for fine-structure multiplets or multiple transitions observed from the same ion, another technique that can be employed is the inverted curve-of-growth method. We describe these methods below.

9.2.5.1 Doublet Ratio

From Eq. 9.17, for two transitions (a) and (b) from the same ion, we see that the ratio of the optical depths at line center obeys

$$r_0 = \frac{\tau_0^{(a)}}{\tau_0^{(b)}} = \frac{\lambda_a^2 f_a}{\lambda_b^2 f_b}. \quad (9.34)$$

For fine-structure doublets, $\lambda_a \simeq \lambda_b$ so that $r_0 \simeq f_a/f_b$. Examination of Table 4.5 reveals that the ratio of the oscillator strengths of $^2S_{1/2}-^2P_{3/2}^o$ and $^2S_{1/2}-^2P_{1/2}^o$ fine-structure doublets are $f_a : f_b \simeq 2 : 1$ where (a) is the $J = 1/2 \rightarrow J = 3/2$ transition and (b) is the $J = 1/2 \rightarrow J = 1/2$ transition. Eq. 9.34 suggests that, in general, there are clear relationships between the equivalent widths of fine-structure doubles, and indeed, all transitions from the same ions. We thus define the doublet ratio, $DR = W_a/W_b$, which is a general expression for any two transitions (a) and (b) from the same ion.

Carrying through the arguments presented above to obtain Eq. 9.33, we see that the double ratio will obey the general trends

$$\begin{array}{lll} \text{linear:} & W \propto \lambda_0^2 N f & DR \propto \frac{\lambda_a^2 f_a}{\lambda_b^2 f_b} \\ \\ \text{flat:} & W \propto b \sqrt{\ln(\lambda_0 N f / b)} & DR \propto \sqrt{\frac{\ln(\lambda_a N f_a / b)}{\ln(\lambda_b N f_b / b)}} \\ \\ \text{damped:} & W \propto \lambda_0^2 \sqrt{\Gamma N f} & DR \propto \frac{\lambda_a^2}{\lambda_b^2} \sqrt{\frac{\Gamma_a f_a}{\Gamma_b f_b}}, \end{array} \quad (9.35)$$

For fine-structure doublets, $\lambda_a \simeq \lambda_b$ and $\Gamma_a \simeq \Gamma_b$. On the linear part, the doublet ratio reflects the ratio of the transition oscillator strengths and we obtain $\text{DR} \simeq 2$. On the flat part of the curve of growth, the column density does not cancel out of the doublet ratio, which will decrease to $\text{DR} \simeq 1$ as N increases; its precise functional form depends on the Doppler b parameter. As the optical depth at line center increases further, the Lorentzian damping wings dominate the equivalent width, the column density cancels out of the doublet ratio, which then asymptotically approaches $\text{DR} \simeq \sqrt{2}$.

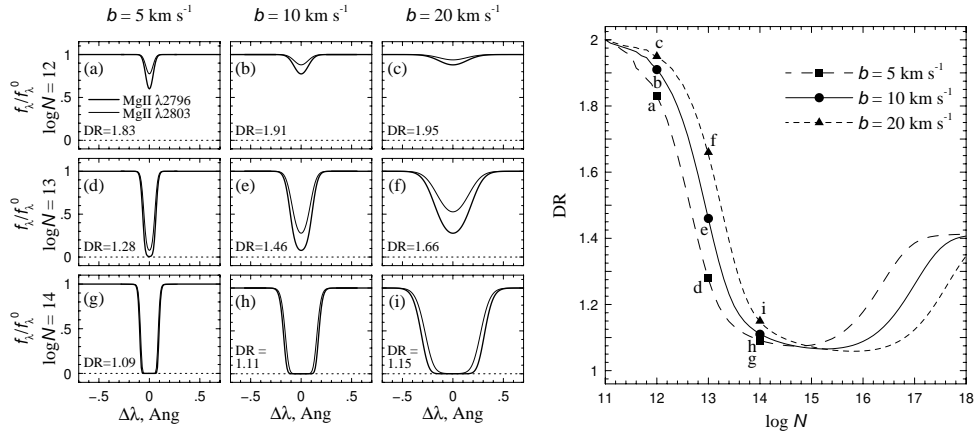


Figure 9.8 (a)–(i) Absorption profiles of the $\text{MgII } \lambda\lambda 2796, 2803$ fine-structure doublets for the column densities $\log N = 12, 13$, and 14 for the Doppler parameters $b = 5, 10$, and 20 km s^{-1} . (right) The doublet ratio of the $\text{MgII } \lambda\lambda 2796, 2803$ as a function of column density for Doppler parameters $b = 5 \text{ km s}^{-1}$ (long dash), $b = 10 \text{ km s}^{-1}$ (solid), and $b = 20 \text{ km s}^{-1}$ (short dash). The points corresponding to each of the profiles in panels (a) through (i) are labelled.

In Figure 9.8(a)–(i), we show absorption profiles for the $\text{MgII } \lambda\lambda 2796, 2803$ fine-structure doublet for three selected column densities and Doppler b parameters. In each panel, the thick curve is the $^2\text{S}_{1/2} - ^2\text{P}_{3/2}^o$ transitions at $\lambda_0 = 2796.35 \text{ \AA}$ with $f = 0.6123$ and the thin curve is the $^2\text{S}_{1/2} - ^2\text{P}_{1/2}^o$ transition at $\lambda_0 = 2803.53 \text{ \AA}$ with $f = 0.3054$. The doublet ratios for each N – b pair are written in their respective panels. In Figure 9.8, we also show the doublet ratio as a function of column density for the three selected Doppler b parameters. The labelled points on each curve correspond to their respective absorption profiles. Note the decrease from $\text{DR} \simeq 2$ on the linear part of the curve of growth to $\text{DR} \simeq 1$ on the flat part of the curve of growth. Also note the moderate dependence on the Doppler b parameter. Observationally, metal-lines such as the MgII doublet are rarely seen with $\log N \geq 15$ (e.g.,

Churchill & Vogt, 2001; Churchill et al., 2003). However, for heuristics, we present the doublet ratio well onto the damped part of the curve of growth, showing that it converges at $DR \simeq \sqrt{2}$, as predicted by Eq. 9.35, given that the damping constants of the two transitions are virtually equivalent and the ratio of the oscillator strengths is 2.005.

If the doublet ratio is measure with high precision, on the order of $\Delta DR \simeq \pm 0.05$, Then, on the flat part of the curve of growth, one can constraint the column density within about 0.5 dex and the Doppler b parameter within 5–10 km s⁻¹. On the linear part, one can only place upper limits on the column density and no constraints on the Doppler b parameter. In practice, however, the measured uncertainties in double ratios, being the quadrature sum of measurement uncertainties in the equivalent widths, are more typically on the order of $\Delta DR \simeq \pm 0.1$ –0.2, rendering the method quite imprecise for constraining N and b .

9.2.5.2 Inverted Curve of Growth

In this method, one begins with the measured W of each transition and “inverts” the curve of growth, thereby obtaining the N, b pairs that give rise to the measured W . The calculations can be performed by varying the Doppler b and solving Eq. 9.21 for the N that yields W . In Figure 9.9, we illustrate the application of the curve-of-growth method for the three transitions of neutral hydrogen, Ly α λ 1215, Ly β λ 1025, and Ly γ λ 972, which, for this example, have measured $W = 0.50, 0.31$, and 0.21 Å, respectively. The three inverted curves of growth intersect on the N – b plane at $\log N = 15.2$ and $b = 30$ km s⁻¹.

The curve-of-growth method has the advantages of being straight forward to implement and of being independent of the resolution of the spectroscopic data (because W is an integrated quantity). However, the method also has important limitations. First, the method works only if all the absorption lines are not blended with or contaminated by other possible absorption lines in the spectrum. Second, measurement uncertainty in the W propagate into a $\pm 1 \sigma$ range of N at each b for each curve, possibly yielding a large overlap region at the intersection of the curves; depending on the measurement uncertainty in each W , the constraints on N may be no better than an order of magnitude or as poor as a 50% uncertainty in b . Third, the method assumes line broadening via thermal wavelength redistribution only. It may be that additional physics is broadening the absorption lines and thus changing the ratios of the W ; this would render the method invalid. It is possible however to employ the technique if additional physics is

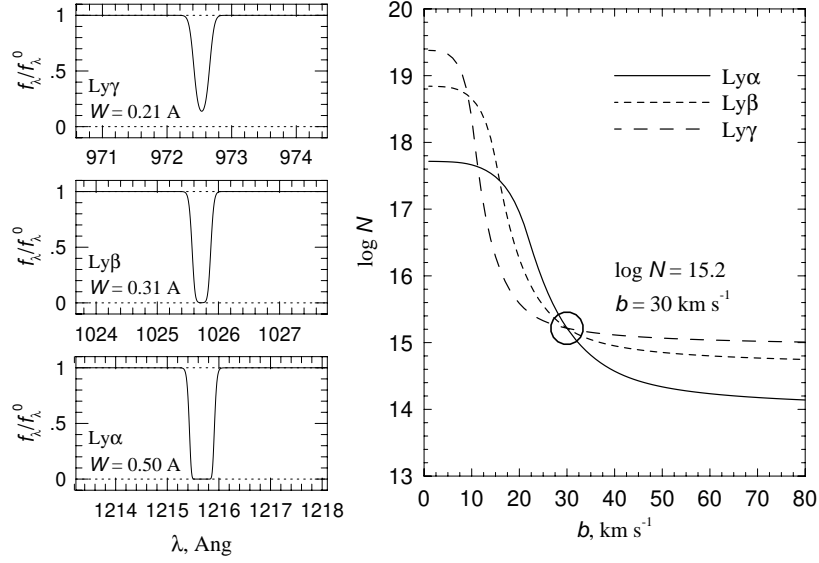


Figure 9.9 An example of the curve-of-growth method for estimating the combination of N and b for three transitions from neutral hydrogen, Ly α λ 1215, Ly β λ 1025, and Ly γ λ 972. For the measured W of each line, the inverted curve of growth is plotted. The intersection of the curves provides the N and b that gives rise to each absorption line. Note that if only two lines are used in the analysis, then there are two solutions.

included in the optical depth, i.e., other wavelength redistribution functions and free parameters are incorporated.

9.2.6 Additional Physics and the Curve of Growth

If additional physics needs to be included in our attempts to understand the absorption line data and how they reflect the absorbing gas physical properties, then multiple wavelength redistribution functions will need to be convolved with the atomic cross section (see § 9.2.1) and there will be additional free parameters.

If two or more of the redistribution functions are area-normalized Gaussian distributions of the form given by Eq. 9.13 having characteristic widths $\Delta\lambda_i$, then the resulting standard deviations of the convolved functions is the quadrature sum of the individual standard deviations,

$$(\Delta\lambda_G)^2 = \sum_i (\Delta\lambda_i)^2. \quad (9.36)$$

where in most all applications the first term in the series is thermal broad-

ening, $\Delta\lambda_1 = \Delta\lambda_D$. The total Gaussian contribution to the wavelength redistribution function is then

$$f_G(\Delta\lambda; \Delta\lambda_G) d\lambda = \frac{1}{\sqrt{\pi}\Delta\lambda_G} \exp\left\{-\left(\Delta\lambda/\Delta\lambda_G\right)^2\right\} d\lambda. \quad (9.37)$$

We then have a “total” Doppler b parameter $b_{\text{tot}} = (c/\lambda_0)\Delta\lambda_G$. In terms of b parameters, commutative properties yield $b_{\text{tot}}^2 = \sum_i b_i^2$. One example of a non-thermal Gaussian redistribution function is turbulence. If we assume a characteristic turbulent velocity of $b_{\text{turb}} = \Delta v_{\text{turb}}$, then

$$b_{\text{tot}}^2 = b_{\text{T}}^2 + b_{\text{turb}}^2 = \frac{2kT}{m} + \Delta v_{\text{turb}}^2. \quad (9.38)$$

Similarly, if two or more of the redistribution functions are area-normalized Lorentzian distributions of the form given by Eq. 2.65 having characteristic widths Y_i , then the resulting standard deviations of the convolved functions is the quadrature sum of the individual standard deviations,

$$Y_{\text{L}}^2 = \sum_i Y_i^2. \quad (9.39)$$

and the total Lorentzian contribution to the wavelength redistribution function is

$$f_{\text{L}}(\Delta\lambda; Y_{\text{L}}) d\lambda = \frac{1}{\pi} \frac{Y_{\text{L}}}{\Delta\lambda^2 + Y_{\text{L}}^2} d\lambda. \quad (9.40)$$

On example of Lorentzian wavelength redistribution is pressure broadening (e.g., Mihalas, 1978; Gray, 1992), though the gas pressure in the interstellar, circumgalactic, and intergalactic medium never approach the values required of discernible pressure broadening effects.

Convolution is commutative, so the order of convolution is arbitrary. Carrying through Eq. 9.8 to obtain the total absorption coefficient for insertion into Eq. 9.7, we first convolve the Lorentzian functions, $\alpha'(\lambda; Y) = (\sigma * f_{\text{L}}[Y_{\text{L}}])(\lambda)$, which is a Lorentzian with $Y^2 = (\Gamma\lambda_0^2/4\pi c)^2 + Y_{\text{L}}^2$ and then convolve the Gaussian functions, $\alpha(\lambda; Y, \Delta\lambda_G) = (\alpha'[Y] * f_G[\Delta\lambda_G])(\lambda)$. The result is that the optical depth is a Voigt profile precisely as defined in Eq. 9.17, with the parameters of $U(x_\lambda, y)$ given by

$$x_\lambda = \frac{\lambda - \lambda_0}{\Delta\lambda_G} \quad \text{and} \quad y = \frac{Y}{\Delta\lambda_G}. \quad (9.41)$$

As such, when the wavelength redistribution functions are limited to Gaussian and/or Lorentzian distributions, we recover the optical depth Voigt profile and that means we recover the curve of growth behavior of the equivalent width but with the curves offset relative to the thermal-only scenario. The

curve-of-growth method can then be applied but with the solution being N , $\Delta\lambda_G$, and Y , following which the individual parameters comprising $\Delta\lambda_G$ and Y might be constrained via application of additional information in the data. If desired, the solution can be obtained by implementing a least-squares minimization algorithm to locate the intersection of the inverted curves of growth while incorporating the measurement uncertainties in each W .

The curve of growth method can be very useful for low resolution spectra. As we will discuss in Chapter 10, higher resolution spectra in which the absorption profiles shapes are resolved can provide additional information for constraining multiple parameters in the redistribution functions. In these cases, determining the column density and the physical broadening parameters of the absorbing gas utilizes the technique of Voigt profile decomposition (§ 10.7.2).

9.2.7 Apparent Optical Depth Method

Examination of Eq. 9.17 suggests that the optical depth can be written

$$\tau_\lambda = \frac{\pi e^2}{m_e c^2} f \lambda_0^2 N(\lambda) \quad (9.42)$$

where f is the oscillator strength, and $N(\lambda) = NU(x_\lambda, y)$ where $U(x_\lambda, y)$ is the Voigt function. Integrating, we obtain the column density

$$N = \int_0^\infty N(\lambda) d\lambda = \frac{m_e c^2}{\pi e^2} \frac{1}{f \lambda_0^2} \int_0^\infty \tau_\lambda d\lambda, \quad (9.43)$$

where $\tau_\lambda = -\ln(f_\lambda/f_\lambda^0)$ is obtained from the measured absorption line profile.

Though this method of measuring the column density is direct and simple, it is fraught with limitations. First, in the cores of strong absorption lines τ_λ cannot be measured; only a lower limit can be estimated. Thus, in lines with saturated cores, only a lower limit on N can be obtained. Second, the method is most accurate for absorption lines that are highly resolved in high resolution spectra and is ineffective in low resolution spectra. This latter complication is due to modification of the absorption line profile shape by the instrumental spread function (see § 8.8). For this reason, the method of inverting the absorption via Eqs. 9.42 and 9.43 is known as the “apparent optical depth method” (Savage & Sembach, 1991; Jenkins, 1996).

Recall that the pre-optics flux is modified by the spectrograph according to the convolution integral given by Eq 8.49. Thus, the post-optics flux in

an absorption line will be

$$\tilde{f}_\lambda = \Phi(\Delta\lambda; \Delta\lambda_s) * f_\lambda^0 \exp\{-\tau_\lambda\}, \quad (9.44)$$

where $\Phi(\Delta\lambda, \Delta\lambda_s)$ is the instrumental spread function (ISF, see, for example, Eq. 8.47), and $\Delta\lambda_s$ is the full-width half-maximum of the ISF given by Eq. 8.40, and where we make the explicit distinction¹ that \tilde{f}_λ is the post-optics flux. Thus, inversion of the absorption profile actually measures the apparent optical depth

$$\tilde{\tau}_\lambda = -\ln\left(\frac{\tilde{f}_\lambda}{f_\lambda^0}\right) = -\ln(\Phi(\Delta\lambda; \Delta\lambda_s) * \exp\{-\tau_\lambda\}). \quad (9.45)$$

As such, Eq. 9.43, must be qualified as the apparent optical depth column density,

$$\tilde{N} = \int_0^\infty \tilde{N}(\lambda) d\lambda = \frac{m_e c^2}{\pi e^2} \frac{1}{f \lambda_0^2} \int_0^\infty \tilde{\tau}_\lambda d\lambda. \quad (9.46)$$

One may ask, why is the integrated quantity W a valid measurement independent of spectral resolution, when the integrated quantity \tilde{N} is not valid for low resolution, but is valid for high resolution (i.e., when the absorption profile is resolved, meaning is broader than the width of the instrumental spread function). The equivalent width is defined such that it accounts for saturation in the line core. However, when the line core is saturated, only a lower limit can be obtained for $\tilde{\tau}_\lambda$, and therefore \tilde{N} .

From Eq. 9.45, we note that for high-resolution spectra, where $\Phi(\Delta\lambda; \Delta\lambda_s)$ is very narrow compared to the line broadening, then $\tilde{\tau}_\lambda \simeq \tau_\lambda$. In high resolution spectra, the registered flux in a saturated line will be $\tilde{f}_\lambda/f_\lambda^0 \simeq 0$, and the measurement of the lower limit will be well defined. For weaker lines, the measured \tilde{N} will be a robust measurement. In low resolution spectra, even when the line core is saturated, the *observed* flux in the line core may still be substantially greater than zero because of the redistribution of flux from the wings of the line into the core (see Figure 8.10). Thus, one has lost the optical depth information and cannot determine an accurate lower limit on the integrated apparent column density. However, note that, even when the line is weak, nowhere along the absorption profile is $\tilde{\tau}_\lambda \simeq \tau_\lambda$, so even weak lines cannot have robustly measured \tilde{N} .

We will discuss the apparent optical depth method for high-resolution spectra in more detail in § 10.

¹ Such a distinction is not necessary for the equivalent width and curve of growth methods because W is an integrated quantity defined to be independent of spectrograph resolution.

9.2.8 Partial Covering from Doublets

Partial covering of the illumination source by the absorbing gas was discussed in § 7.4.5. We derived a simple relationship (given by Eq. 7.81) relating the residual flux, R_λ , in the absorption line and the covering fraction, $C_f(\lambda)$. Repeating that result here, we have

$$R_\lambda = \frac{f_\lambda}{f_\lambda^0} = [1 - C_f(\lambda)] + C_f(\lambda) \exp \{-\tau_\lambda\} , \quad (9.47)$$

where τ_λ is the optical depth of the occulting absorbing gas. Eq. 9.47 is based upon the source-absorber geometry illustrated in Figure 7.14. We showed that for $\tau_\lambda > 10$, one could directly estimate $C_f(\lambda) = 1 - R_\lambda$ in the core of the line. The problem is that the optical depth in the absorbing gas for the transition being measured is an unknown; thus, one cannot securely estimate the covering fraction without additional observational constraints.

If both transitions of a fine-structure doublet exhibit the absorption signature of partial covering, then the covering fraction can be estimated because each member has a unique optical depth, for which the column density factors out (e.g. Barlow & Sargent, 1997; Hamann et al., 1997; Hamann & Ferland, 1999; Crenshaw et al., 1999; Ganguly et al., 1999; Hamann & Sabra, 2004; Misawa et al., 2016). From Eq. 9.17, the ratio of the optical depths are simply $r_\lambda = \tau_\lambda^{(a)} / \tau_\lambda^{(b)} = \lambda_a^2 f_a / \lambda_b^2 f_b$, where we have adopted the notation of § 9.2.5.1 for the members of the doublet.

To estimate the covering fraction, we utilize ratio of the optical depths. We will assume that the covering fraction is the same for both members of the doublet, and that a single value of the covering fraction applies across the absorption profiles (though no generality is lost for the latter assumption). We also note that the following discussion is valid only for high-resolution spectra for the same reasons discussed with regard to the apparent optical depth method in § 9.2.7 (also see, Ganguly et al., 1999, who address the conditions of validity). It is thus implied that validity in what follows applies when $R_\lambda = \tilde{R}_\lambda = \tilde{f}_\lambda / f_\lambda^0$.

Solving Eq. 9.47 for the optical depth in the absorbing gas, we obtain

$$\tau_\lambda = -\ln \left\{ \frac{R_\lambda - 1 + C_f}{C_f} \right\} , \quad (9.48)$$

where we have dropped the wavelength dependence of $C_f(\lambda)$. We thus have

$$r_\lambda = \frac{\ln \left\{ (R_\lambda^{(a)} - 1 + C_f) / C_f \right\}}{\ln \left\{ (R_\lambda^{(b)} - 1 + C_f) / C_f \right\}} , \quad (9.49)$$

from which follows

$$\left[\frac{R_\lambda^{(b)} - 1 + C_f}{C_f} \right]^{r_\lambda} = \left[\frac{R_\lambda^{(a)} - 1 + C_f}{C_f} \right]. \quad (9.50)$$

If $r_\lambda \notin \mathbb{Z}$, then Eq. 9.50 must be solved numerically. However, as discussed in § 9.2.5.1, the fine-structure doublets that are commonly observed in cosmological spectra have $r_\lambda \simeq 2$. If we assume $r_\lambda = 2$, we obtain the closed-form solution

$$C_f = \frac{(R_\lambda^{(b)} - 1)^2}{R_\lambda^{(a)} - 2R_\lambda^{(b)} + 1}. \quad (9.51)$$

Since $0 < C_f \leq 1$, Eq. 9.51 is valid only for $R_\lambda^{(b)} \leq (R_\lambda^{(a)})^{1/2} < 1$. We again remind the reader that Eq. 9.51 is valid only in high-resolution spectra in regions of the absorption profiles where $\tilde{R}_\lambda \simeq R_\lambda$. In practice, these regions are in the cores of lines far from the wings, as the shape of the wings is modified by the instrumental spread function (see Ganguly et al., 1999).

Once a non-zero covering fraction has been estimated, the column density of the absorbing ion in the gas cloud is related to the registered residual flux, $\tilde{R}_\lambda = \tilde{f}_\lambda / f_\lambda^0$, via

$$\Phi(\Delta\lambda; \Delta\lambda_s) * \exp\{-N\alpha_\lambda\} = \frac{\tilde{R}_\lambda - 1 + C_f(\lambda)}{C_f(\lambda)}, \quad (9.52)$$

where we have explicitly included the notation that the instrument spread function must be incorporated to obtain a robust measurement of the column density. In practice, this relation is solved by using least-squares fitting of trial values of N and of the physical parameters comprising α_λ to model the observed spectrum. Within the confines of the adopted functional form of α_λ , it is alternatively possible to estimate both N and $C_f(\lambda)$ simultaneously from Eq. 9.52.

9.2.9 Observed Redshifted Equivalent Widths

When absorption lines are redshifted, the observed equivalent widths will be a factor of $1 + z$ greater than the rest-frame equivalent width, i.e.,

$$W = W_r(1 + z), \quad (9.53)$$

where the definition of observed redshift is (Eq. 6.2)

$$z = \frac{\lambda - \lambda_r}{\lambda_r}, \quad (9.54)$$

and where λ is the observed wavelength corresponding to the rest-frame wavelength λ_r . We thus have the well-known relationship $\lambda = \lambda_r(1 + z)$.

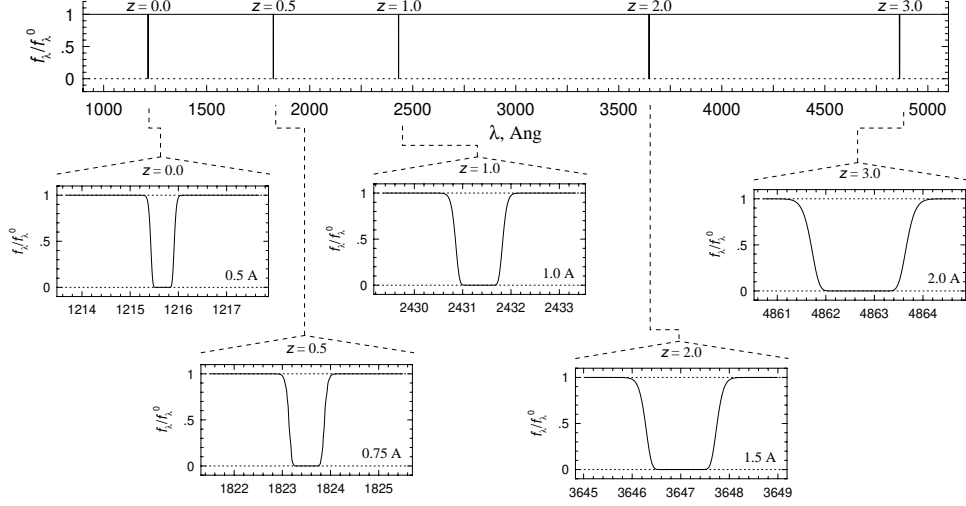


Figure 9.10 Examples of redshifted Ly α λ 1215 absorption lines with rest-frame equivalent width $W_r = 0.5 \text{ \AA}$. Absorption profiles for $z = 0.0$ (upper left), $z = 0.5$ (lower left), $z = 1.0$ (upper center), $z = 2.0$ (lower right), and $z = 3.0$ (upper right) are shown and their measured W are provided. Each panel covers the same wavelength range (4 \AA) so that the profile widths can be directly compared. The top spectrum shows the redshifting of the central wavelengths of these absorption lines.

To demonstrate the origin of Eq. 9.53, we consider an idealized absorption line in the rest frame with zero flux density over the finite wavelength range λ_- to λ_+ ,

$$f_\lambda = \begin{cases} f_\lambda^0 & \text{for } \lambda < \lambda^- \\ 0 & \text{for } \lambda^- \leq \lambda \leq \lambda^+ \\ f_\lambda^0 & \text{for } \lambda > \lambda^+ \end{cases} \quad (9.55)$$

The equivalent width in the rest-frame for our idealized absorption is

$$W_r = \int_0^\infty [1 - f_\lambda/f_\lambda^0] d\lambda = \int_{\lambda^-}^{\lambda^+} d\lambda = \lambda^+ - \lambda^-. \quad (9.56)$$

Now, let this idealized absorption line be redshifted by some arbitrary z . Every wavelength across the profile will be shifted in wavelength by the factor $1 + z$. The observed equivalent width is then

$$W = \int_0^\infty [1 - f_\lambda/f_\lambda^0] d\lambda = \int_{\lambda^-(1+z)}^{\lambda^+(1+z)} d\lambda = (\lambda^+ - \lambda^-)(1 + z). \quad (9.57)$$

Comparing Eqs. 9.56 and 9.57, we recover Eq. 9.53.

In Figure 9.10, we show examples of the redshifting of $\text{Ly}\alpha$ $\lambda 1215$ for a rest-frame equivalent $W_r = 0.5$ Å. Each panel covers 4 Å wide. For each example redshift, z , the central wavelength is shifted to $\lambda = 1215.67(1+z)$, which is shown in the central spectrum. The observed equivalent widths are increased by the factor $1+z$, such that $W = W_r(1+z)$.

Given that the equivalent width of an observed absorption line is based upon the atomic physics as measured in the laboratory, and on the wavelength redistribution function based on physics local to the absorbing gas, in order to extract information on the absorbing gas, the observed equivalent width must be converted to the rest frame of the absorption. This is accomplished by simply computing $W_r = W/(1+z)$.

9.3 Ionization Breaks

The most commonly observed ionization break observed in cosmological spectra is due to neutral hydrogen (HI) at the Lyman ionization threshold, or edge ($n = 1 \rightarrow \infty$). In the rest frame of the absorbing gas, photons with energy $hc/\lambda > 13.598$ eV can potentially ionize neutral hydrogen, the most abundant and ubiquitous element in the cosmos. Thus, for $\lambda \leq 912$ Å in the rest-frame of the gas, continuum absorption may present, and if present it will appear as a sharp discontinuity in the observed flux at the wavelength of the ionization edge. If the absorbing gas is at redshift z , the discontinuity on the spectrum will be seen at the wavelength $912(1+z)$.

This spectral feature is so commonly seen that it is given the moniker the “Lyman limit break” (LLB), the “Lyman limit” (LL), or simply the “Lyman break”. Because Lyman breaks arise in high column density gas, associated metal absorption lines are almost always present, in which case the absorbing gas is known as a “Lyman limit system”, or “LLS”.

Although the cross section for bound-free absorption, as illustrated in Figure 9.1(b), extend over a large wavelength range, we still employ Eq. 9.9 to obtain the column density of atoms undergoing ionization, i.e., $f_\lambda = f_\lambda^0 \exp\{-N\alpha(\lambda)\}$.

The atomic cross section for the bound-free transition of neutral hydrogen is (see Eq. 2.91)

$$\sigma_{\text{LL}}(\lambda) = A_0 \left(\frac{\lambda}{\lambda_0} \right)^4 \frac{\exp \left\{ 4 - \left[4/\sqrt{\lambda_0/\lambda - 1} \right] \tan^{-1} \left(\sqrt{\lambda_0/\lambda - 1} \right) \right\}}{1 - \exp \left\{ -2\pi/\sqrt{\lambda_0/\lambda - 1} \right\}}, \quad (9.58)$$

where $\lambda_0 = 912$ (see Eq. 2.92), and where $A_0 = 6.304 \times 10^{-18}$ cm² atom⁻¹

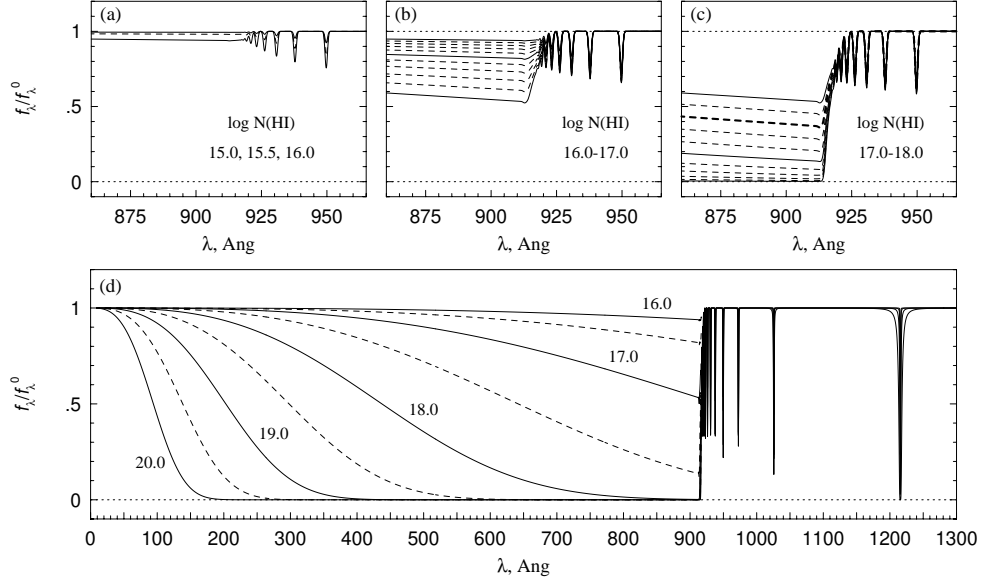


Figure 9.11 The rest-frame Lyman limit ionization break for (a) $\log N_{\text{HI}} = 15.0, 15.5$ (dashed spectrum), and 16.0. (b) $\log N_{\text{HI}} = 16.0\text{--}17.0$ in intervals of 0.1, where the solid spectra are 16.0, 16.5, and 17.0. (c) $\log N_{\text{HI}} = 17.0\text{--}18.0$ in intervals of 0.1, where the solid spectra are 17.0, 17.5, and 18.0. The thick-dashed spectrum is for $\log N_{\text{HI}} = 17.2$, corresponding to $\tau_{\text{LL}} = 1$. For the limited spectral region shown, the ionization breaks are indistinguishable for $\log N_{\text{HI}} > 18.0$. (d) Over a broad wavelength range, the ionization break “recovers” to the continuum due to the $\propto \lambda^3$ asymptotic behavior of the cross section as given in Eq. 9.63. Shown are breaks over the range $\log N_{\text{HI}} = 16.0\text{--}20.0$ in intervals of 0.5. (all panels) The Lyman series bound-bound absorption lines are also shown, but not labelled. For higher N_{HI} , the higher order lines begin to overlap near the ionization threshold resulting in a sloping flux decrement just redward of the break, which is clearly seen in panels (b) and (c).

as give by Eq. 2.88. Since the cross section varies smoothly over a very broad wavelength range, the convolution to obtain the total absorption cross section, given by Eq. 9.8, is rarely necessary. That is, physics such as thermal Doppler broadening has a negligible impact on the spectral shape of the ionization break. Doppler widths are on the order of 0.1 \AA (see Figure 9.3), whereas the bound-free cross sections vary smoothly over $100\text{--}1000 \text{ \AA}$ (see Figure 9.1). Thus, for ionization edges we can write

$$f_{\lambda} = f_{\lambda}^0 \exp\{-N_{\text{HI}}\sigma_{\text{LL}}(\lambda)\}. \quad (9.59)$$

As we shall discuss below, this greatly simplifies measuring the neutral hydrogen column density directly from the spectra.

In Figure 9.11, we show $f_\lambda = f_\lambda^0 \exp\{-N_{\text{HI}}\sigma_{\text{LL}}\}$ for several ranges of N_{HI} in incremental steps of $\Delta \log N_{\text{HI}} = 0.1$. Though the Lyman limit ionization break is often observed at redshifts $z > 0$ in cosmological spectra, we show the break in the rest frame. For $\log N_{\text{HI}} \leq 16.0$, the flux decrement at the ionization break is barely discernible. The break becomes discernible for $\log N_{\text{HI}} \geq 16.5$, but is typically indistinguishable from noise in the data. The flux decrement near the break f_λ/f_λ^0 grows rapidly in the range $17 \leq \log N_{\text{HI}} \leq 18$, asymptotically approaching $f_\lambda/f_\lambda^0 = 0$. For $\log N_{\text{HI}} \geq 18$, the Lyman break is saturated.

Because the total cross section is simply the atomic cross section, at the break wavelength we have

$$\tau_{\text{LL}} = N_{\text{HI}}\sigma_{\text{LL}}(\lambda_0) = 6.304 \times 10^{-18} N_{\text{HI}}, \quad (9.60)$$

yielding

$$N_{\text{HI}} = 1.586 \times 10^{17} \tau_{\text{LL}} \quad (9.61)$$

where $\tau_{\text{LL}} = -\ln(f_{\lambda_0}/f_{\lambda_0}^0)$ is evaluated at $\lambda_0 \simeq 912$. When $f_{\lambda_0}/f_{\lambda_0}^0 = 0.368$, we have $\tau_{\text{LL}} = 1$, yielding $\log N_{\text{HI}} = 17.2$ (shown as the thick-dashed spectrum in Figure 9.11(c)). In practice, it is convention to measure the optical depth using the ratio of the mean fluxes below and above the break,

$$\tau_{\text{LL}} = \ln \frac{\langle F(\lambda > \lambda_0) \rangle}{\langle F(\lambda < \lambda_0) \rangle}, \quad (9.62)$$

where the $\langle F \rangle$ are usually obtained by fitting a constant to a limited wavelength range, for example 875–900 below the break and 925–950 above the break. When $\langle F(\lambda < \lambda_0) \rangle \simeq 0$ (or is consistent with null flux within the uncertainties in the data), then only a lower limit on the optical depth, and there for column density can be determined.

Due to the functional form of the atomic cross section, higher energy photons have lower absorption rates than photons near the ionization edge. The asymptotic behavior of Eq. 9.58 for $\lambda \ll \lambda_0$ is

$$\sigma_{\text{LL}}(\lambda) \simeq 6.304 \times 10^{-18} \left(\frac{\lambda}{\lambda_0} \right)^3, \quad (9.63)$$

so that the flux blueward of the break will recover to the continuum flux according to $\propto \exp\{-\lambda^3\}$. We show the recovery of the break as a function of N_{HI} in Figure 9.11(d). If this recovery can be measured, then it is possible to measure the neutral hydrogen column density even for $\log N_{\text{HI}} > 18$.

We now can refine our definition of a Lyman limit system. Given that damped Ly α systems (DLAs) are defined to have $\log N_{\text{HI}} \geq 20.3$, we can

define absorption systems with $18 \leq \log N_{\text{HI}} < 20.3$ as sub-damped Ly α systems (sub-DLAs). This definition accounts for the fact that the Lyman break is saturating at $\log N_{\text{HI}} = 18$ just when damping wings in the Ly α line are becoming discernable. Then, the moniker Lyman limit system (LLS) is reserved for $16.8 \leq \log N_{\text{HI}} < 18$, the range over which the Lyman break just becomes discernable to when it becomes saturated.

Note that for $\log N_{\text{HI}} > 20$, a vast spectral region is fully absorbed by ionization. If such a high column density absorption system is observed at high redshift in a quasar spectrum, then virtually all of the quasar radiation blueward of the redshifted Lyman break wavelength is absorbed. This has the unfortunate consequence that some or all transitions from an absorbing gaseous systems at a redshift lower than the Lyman limit system cannot be probed by the quasar light beam. That is, if the Lyman limit system is at z_{LL} and the gaseous system is at redshift $z < z_{\text{LL}}$, then only those absorption lines from the system that meet the condition

$$\lambda_r(1+z) > 912(1+z_{\text{LL}}), \quad (9.64)$$

where λ_r is the rest-frame transition wavelength, can be observed. The removal of the quasar flux over a large wavelength range has the added oddity that if one were to take a narrow-band image of the quasar centered at $\lambda < 912(1+z_{\text{LL}})$, the quasar would be invisible! However, a narrow-band image centered at $\lambda > 912(1+z_{\text{LL}})$ would reveal the quasar fully. This method has been exploited to look for luminous objects (galaxies) associated with damped Ly α systems (e.g., O'Meara et al., 2006).

10

Spectral Analysis

As we have discussed in Chapter 8, for absorption lines, the relationship between observed flux and the recorded counts is

$$I_\lambda = \mathcal{S}_\lambda \Phi(\Delta\lambda; \Delta\lambda_s) * f_\lambda^0 \exp\{-\tau_\lambda\}, \quad (10.1)$$

where \mathcal{S}_λ is the sensitivity function, which accounts for the atmospheric attenuation, telescope and spectrograph throughput, pixelization sampling rate, and the quantum efficiency of the detector. As further discussed in § 8.10.4, wherever the spectral energy distribution of the source is smooth, meaning variations in the flux levels are over scales much broader than the resolution element $\Delta\lambda_s$, then to a high degree of accuracy the continuum flux need not account for the effects of the ISF, yielding,

$$I_\lambda^0 = \mathcal{S}_\lambda f_\lambda^0. \quad (10.2)$$

Across the absorption line, the continuum value must to be estimated (a topic we cover in § 10.1). We denote this estimate as \bar{I}_λ^c . Then, the measured “flux decrement” in the absorption line is written,

$$D_\lambda = 1 - I_\lambda/\bar{I}_\lambda^c = 1 - \Phi(\Delta\lambda; \Delta\lambda_s) * \exp\{-\tau_\lambda\}. \quad (10.3)$$

Since the counts at each wavelength are pixelized, we will adopt the notation that the wavelength corresponding to pixel i is λ_i . Then, in practice, integrated absorption quantities become sums. For example, the equivalent width of an absorption line is measured by summing the flux decrements over the pixels comprising the absorption,

$$W = \int_{\lambda^-}^{\lambda^+} [1 - I_\lambda/\bar{I}_\lambda^c] d\lambda = \sum_{i=i^-}^{i^+} D_i \Delta\lambda_i, \quad (10.4)$$

where $D_i = 1 - I_i/\bar{I}_i^c$ is the flux decrement in pixel i mapped to wavelength

λ_i , and where $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$ is the wavelength interval spanned by pixel i , and $\lambda^- = \lambda_{i-}$ and $\lambda^+ = \lambda_{i+}$ are the lower and upper wavelength limits corresponding to the pixels included in the sum.

The first step of spectral analysis is to obtain an estimate of the continuum counts, or in the case of flux calibrated spectra, the continuum flux. For the remainder of this chapter, we will assume counts. The second step is to locate absorption features, including their lower and upper wavelength limits. The features should be identified using objective methods that also properly measure their statistical significance. This means that the estimated uncertainties in the flux decrements must be accurate. Once absorption features are identified, we can pursue our science objectives, which normally begins by identifying absorption lines by their atomic transitions and then, applying the principles of the cosmological redshift, grouping the identified absorption lines into “absorption systems”. Finally, we wish to analyze the absorption systems, starting with measuring quantities such as the rest-frame equivalent widths, column densities, and broadening parameters. These steps are covered in this chapter. Additional analysis of the absorbing gas is discussed in Chapters 11 and 12.

10.1 Continuum Fitting

Virtually all measurements of absorption features invoke the pixelized flux decrement, D_i or the ratio I_i/\bar{I}_i^c . Thus, for the analysis of absorption lines, the true continuum level, I_λ^c , must be estimated for wavelength regions spanning absorption features.

The continuum is usually modeled using a smooth curve fit through the data employing an objective statistical treatment (such as least squares fitting). Often, the smooth curve is a polynomial or a series of splined polynomials and is known as the “continuum fit” or the “fitted continuum”.

The modeled continuum fit may have \bar{I}_i^c either slightly systematically larger or smaller than the actual true continuum. Thus, the process introduces systematic errors in the measured absorption quantities. The uncertainty in \bar{I}_i^c , denoted $\sigma_{\bar{I}_i^c}$, should be properly accounted for the uncertainty estimates of quantified absorption properties. Whereas the uncertainty in the counts, σ_{I_i} , are independent quantities for each pixelized wavelength (see Eq. 8.63), the uncertainties in the fitted continuum must account for that fact that they depend upon the counts over a *range* of wavelengths. That is, the uncertainties in the fitted continuum in each pixel depend on the values I_i and σ_{I_i} over the range of pixels used to constrain the model continuum.

Sembach & Savage (1992) showed that using normalized orthogonal (orthonormal) functions for continuum modeling provide an ideal approach to minimizing the σ_{D_i} . Here, we review their approach to continuum fitting, which is most appropriate for wavelength regions of the spectrum where the counts are not varying rapidly or for small regions where scientific analysis is to be performed.

In general, orthonormal functions are defined by the result of their inner product

$$\int_a^b P_j(x)P_k(x) dx = \begin{cases} 0 & j \neq k \\ 1 & j = k. \end{cases} \quad (10.5)$$

where a and b are the lower and upper limits of the interval over which orthogonality holds, and j and k are the orders of the functions. For purposes of continuum modeling, the choice of the function ((i.e., Hermite, Gegenbauer, Jacobi, Laguerre, Legendre, Chebyshev, etc., see Abramowitz & Stegun, 1972) is not vital, but it is best to try and select behavior in the function such that low order functions can be employed. The choice of polynomial dictates the interval $[a, b]$ over which orthogonality holds. In order to enforce orthogonality for the continuum fitting model, the wavelength interval must be temporarily transformed to the interval $[a, b]$ during the fitting process.

Following Sembach & Savage (1992), we discuss the use of Legendre polynomials, $P_k(x)$, where k is the order of x . The interval of orthogonality is $[a, b] = [-1, 1]$ and the $P_k(x)$ are symmetric about $x = 0$ for even k and are anti-symmetric about $x = 0$ for odd k . The fitted continuum, \bar{I}_λ^c , is then modeled as the sum of $M + 1$ functions,

$$\bar{I}_i^c = P_M(x_i) = \sum_{k=0}^M a_k P_k(x_i), \quad (10.6)$$

where a_k are the coefficients (fitting parameters) for each polynomial order k , and M is the highest order function included in the model. If the wavelength region of spectrum to be fitted is over the interval $[\lambda_L, \lambda_U]$, then the transformation to the variable x_i over the interval $[-1, 1]$ is given by $x_i = (2\lambda_i - \lambda_U - \lambda_L)/(\lambda_U - \lambda_L)$.

The coefficients are determined by least squares fitting. The method of χ^2 minimization is appropriate, where

$$\chi_\nu^2 = \sum_{i=1}^{N_{\text{pix}}} \frac{[I_i - P_M(x_i)]^2}{\sigma_{I_i}^2}, \quad (10.7)$$

for $\nu = N_{\text{pix}} - (M+1)$ degrees of freedom, where N_{pix} is the number of pixels included in χ^2 least squared fit. Note that not all pixels in the interval $x_i = [-1, 1]$ are required to be included in the least squares minimization (the χ^2 statistic is a simple sum; there is no information regarding wavelength or pixel location). In fact, it is important to omit pixels for which the I_i deviate significantly from the surrounding continuum. For example, the counts in the absorption lines themselves should be omitted from the fit, which can be performed using σ clipping.

An example continuum fit is illustrated in Figure 10.1(a) for a single echelle order. The upper axis provides the interval of orthogonality based upon the wavelength interval being modeled, given by the lower axis. The overall shape of the counts reflects the blaze function of the echelle grating. The absorption feature at $x = [0.33, 0.42]$ has been omitted from the fit. The uncertainty spectrum is also shown. In Figure 10.1(b), the normalized spectrum, $I_\lambda/\bar{I}_\lambda^c$, and the normalized uncertainty spectrum, $\sigma_{I_\lambda}/\bar{I}_\lambda^c$, are plotted in units of normalized (relative) counts.

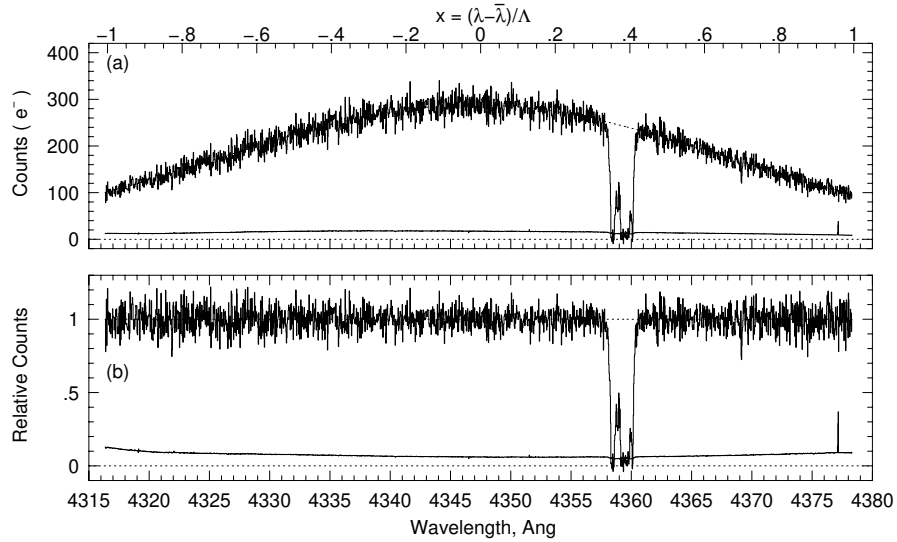


Figure 10.1 (a) An example continuum fit for a single order from an echelle spectrograph. The increasing counts toward the center of the wavelength interval reflects the blaze efficiency of the grating. The interval of orthogonality, $x = [-1, 1]$, is given by the upper axis. (b) The normalized spectrum and normalized uncertainty spectrum, σ_{I_i}/\bar{I}_i^c , based upon the recorded counts only and not including the uncertainty due to the continuum fit.

The uncertainty in the continuum fit at each pixel is computed from

$$\sigma_{\bar{I}_i^c}^2 = \sum_{j=0}^M \sum_{k=0}^M \sigma_{a_{jk}}^2 P_j(x_i) P_k(x_i), \quad (10.8)$$

where $\sigma_{a_{jk}} = C_{jk}, = \alpha_{jk}^{-1}$ are the uncertainties in the coefficients of the fitted continuum, which are the elements of the co-variance matrix, C_{jk} , which are elements of the inverse of the curvature matrix, α_{jk}^{-1} (see, e.g., Press et al., 2007). From Eq. 10.8, we see that the higher the order of the fitted polynomial, the larger the uncertainty in the continuum at each pixel. Thus, we aim to use the lowest order, M , while achieving a minimized χ^2 . As discussed by Sembach & Savage (1992), the properties of orthogonality of the polynomials allow one to apply a simple F -test that informs us when adding an additional order to the fit no longer provides a statistically significant improvement in the fit. Thus allows one to objectively determine the minimum M dictated by the noise characteristics of the data.

Along the matrix diagonal, the $\sigma_{a_{jj}}$ can be viewed as the uncertainties in successive orders of curvature in the fitted continuum. For example, $\sigma_{a_{00}}$ is the uncertainty in a constant offset shift of the fitted continuum, $\sigma_{a_{11}}$ is the uncertainty in the linear slope, and $\sigma_{a_{22}}$ is the uncertainty in the quadrature curvature. The off-diagonal elements, $\sigma_{a_{jk}}$, provide the cross-correlated uncertainties.

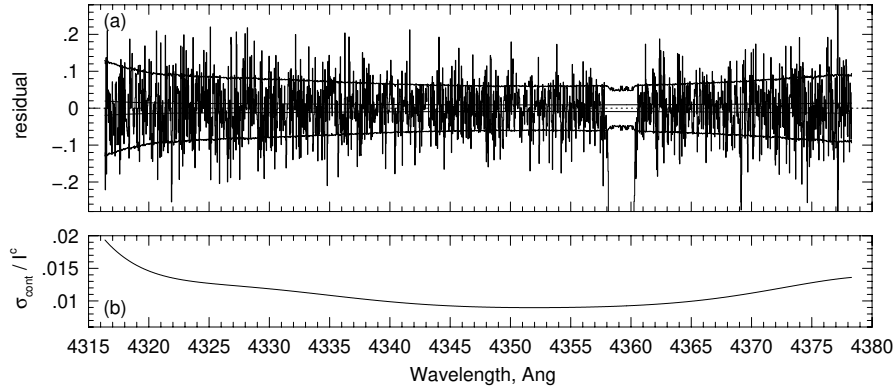


Figure 10.2 (a) The residuals of the counts about the fitted continuum, following the normalization illustrated in Figure 10.1, i.e., $(I_i - \bar{I}_i^c) / \bar{I}_i^c$. The normalized uncertainty spectrum of the counts, is shown as the two outer curves (with $\sigma_{I_i} / \bar{I}_i^c \simeq \pm 0.1$ superimposed on the residuals. The uncertainties in the fitted continuum are the two inner smooth curves superimposed on the residuals, having $\sigma_{\bar{I}_i^c} / \bar{I}_i^c \simeq \pm 0.015$. (b) Expanded view of the normalized uncertainty in the fitted continuum.

In Figure 10.2(a), we further examine the residuals of the continuum normalized spectrum presented in Figure 10.1. Superimposed are the normalized sigma spectrum, having a typical value of $\sigma_{I_i}/\bar{I}_i^c = 0.1$, and the uncertainty in the fitted continuum, having a typical value of $\sigma_{\bar{I}_i^c}/\bar{I}_i^c = 0.015$. These uncertainties are presented such they bracket the zero point of the normalized residuals. For normally distributed uncertainties and an accurate continuum fit, roughly 68% of the residuals should be bracketed by the uncertainty spectrum. In Figure 10.2(b), the uncertainty in the fitted continuum is shown with an expanded view.

Incorporation of the uncertainty due to the continuum fit is straight forward. Consider the uncertainty in flux decrement, we have,

$$\sigma_{D_i}^2 = \left[\frac{\partial D_i}{\partial I_i} \right]^2 \sigma_{I_i}^2 + \left[\frac{\partial D_i}{\partial \bar{I}_i^c} \right]^2 \sigma_{\bar{I}_i^c}^2, \quad (10.9)$$

where $\partial D_i / \partial I_i = 1 / \bar{I}_i^c$, and $\partial D_i / \partial \bar{I}_i^c = -I_i / [\bar{I}_i^c]^2$, and where $\sigma_{I_i}^2$ is obtained directly from the extracted spectrum and $\sigma_{\bar{I}_i^c}^2$ is obtained from the continuum fit model (Eq. 10.8). The resulting uncertainty can be written,

$$\sigma_{D_i}^2 = \left(\frac{I_i}{\bar{I}_i^c} \right)^2 \left[\left(\frac{\sigma_{I_i}}{I_i} \right)^2 + \left(\frac{\sigma_{\bar{I}_i^c}}{\bar{I}_i^c} \right)^2 \right], \quad (10.10)$$

We will invoke Eq. 10.10 for several applications in the remainder of this chapter.

10.2 Combining spectra

Often, in the case of long slit spectra, multiple observations have been obtained and must be combined into a single spectrum. In cases where echelle spectra have been obtained, it is desirable to combine the overlapping wavelength regions from adjacent orders. Combining the spectra requires a new binning of the wavelength pixelization. This must be done with care. First, the counts or flux must be conserved during the binning process. Second, the count level of each spectrum must in some fashion be on an equal scaling, either by flux calibration or by normalization of the continuum.

The counts in the pixels of the original, or input spectra to the combining process, are not correlated. Each pixel is an independent measurement. However, the binning process will result in counts that are correlated with adjacent pixels. Furthermore, if the pixel bin size of the output spectrum is fixed width, this linearizes the dispersion of the spectrum.

Consider a set of N_s spectra of the same object. The process begins by

accounting for the natural binning of the input spectra due to pixelization. In the j^{th} spectrum, each pixel, i , has a central wavelength λ_{ji} , with width $\Delta\lambda_{ji}$, given by

$$\Delta\lambda_{ji} = \lambda_{ji} - \lambda_{j,i-1}, \quad \lambda_{ji}^- = \lambda_{ji} - \frac{1}{2}\Delta\lambda_{ji}, \quad \lambda_{ji}^+ = \lambda_{ji} + \frac{1}{2}\Delta\lambda_{ji}, \quad (10.11)$$

where λ_{ji}^- and λ_{ji}^+ are the starting and ending wavelengths of the pixel (the pixel edges). Note that $\Delta\lambda_{ji}$ is not necessarily a constant for all i . For the combined output spectrum, it is simplest to linearize the wavelength interval of the pixels. It is advantageous to set this interval to the maximum wavelength width of the input spectra (this avoids possible aliasing),

$$\Delta\lambda = \max(\Delta\lambda_{ji}), \quad (10.12)$$

where the lack of a subscript denotes the combined output spectrum pixelization. This combined output spectrum must cover the full range of the input spectra, so that

$$\lambda^{\min} = \min(\lambda_{ji}), \quad \lambda^{\max} = \max(\lambda_{ji}), \quad (10.13)$$

where λ^{\min} and λ^{\max} are the beginning wavelength and ending wavelength of the combined output spectrum. The number of pixels in the output spectrum is then $N = \text{INT}[(\lambda^{\max} - \lambda^{\min})/\Delta\lambda]$. This information provides the wavelength pixelization of the output spectrum as follows,

$$\lambda_k = \lambda^{\min} + (k - 1)\Delta\lambda, \quad \lambda_k^- = \lambda_k - \frac{1}{2}\Delta\lambda, \quad \lambda_k^+ = \lambda_k + \frac{1}{2}\Delta\lambda, \quad (10.14)$$

where $k = 1, 2, \dots, N$, and where λ_k is the wavelength center of pixel k , and λ_k^- and λ_k^+ are the starting and ending wavelengths of pixel k . The indexing is illustrated in Figure 10.3(a). The pixelization of several input spectra, labeled j through $j + 3$ (for $N_s = 4$) are shown, with details of the pixelization binning illustrated for pixel i of the j^{th} spectrum (Eq. 10.11). The output spectrum binning is shown for pixel k (Eqs. 10.12, 10.13, and 10.14).

To obtain the output spectrum requires a two step process. The first step is to resample the data for each input spectrum, j , to have the wavelength pixelization of the output spectrum. This process requires conservation of the counts (or flux), for interpolation will be required since the pixels of each input spectrum will not necessarily align with the pixels of the output spectrum. This interpolation is based upon the fractional overlap of the input spectrum pixel and the output spectrum pixel. The flux conserving

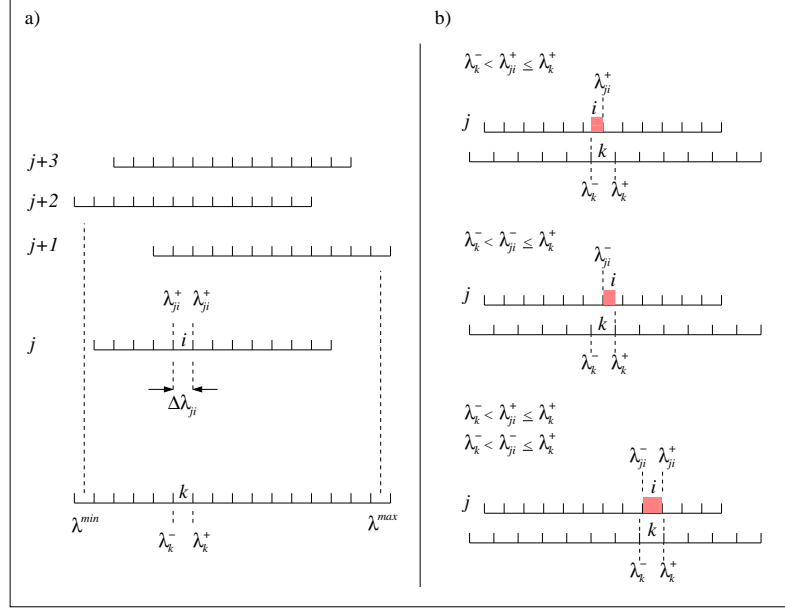


Figure 10.3 (a) The pixelization of the input spectra are shown with details of the indexing for pixel i of spectrum j . The output spectrum is shown at the bottom of the panel. The minimum and maximum wavelengths are illustrated as is the book keeping indexing for pixel k . (b) The binning of spectrum j onto the pixelization binning, k , of the output spectrum is illustrated for the three conditions of Eq. 10.15. The process accounts for fractional overlap of the pixels and conserves the counts in the binning process.

weights, w_{kji} , are

$$w_{kji} = \begin{cases} \frac{\lambda_{ji}^+ - \lambda_k^-}{\Delta\lambda_{ji}} & \lambda_k^- < \lambda_{ji}^+ \leq \lambda_k^+ \\ \frac{\lambda_k^+ - \lambda_{ji}^-}{\Delta\lambda_{ji}} & \lambda_k^- < \lambda_{ji}^- \leq \lambda_k^+ \end{cases} \quad (10.15)$$

where $w_{kji} = 1$ if both conditions are true, or $w_{kji} = 0$ if neither condition is true.

The conditions account for the fraction of the counts in pixel i of spectrum j to be resampled into pixel k based on a fractional pixel area weighting scheme. As illustrated in upper portion of Figure 10.3(b), the first condition of Eq. 10.15 provides the fraction of I_{ji} where the longer wavelength portion of pixel i of input spectrum j overlaps with pixel k . The center portion of Figure 10.3(b) illustrates the second condition of Eq. 10.15 that provides

the fraction of I_{ji} where the shorter wavelength region of pixel i of input spectrum j overlaps with pixel k . The third condition of Eq. 10.15 applies when the pixel i of spectrum j aligns fully within pixel k , as is shown in the lower portion of Figure 10.3(b). This latter condition is not common, but must be considered. The fourth condition (not illustrated) applies when there is no alignment between the input and output pixels for the given indices. A mask spectrum can also be created in which undesirable pixels can be assigned $w_{kji} = 0$.

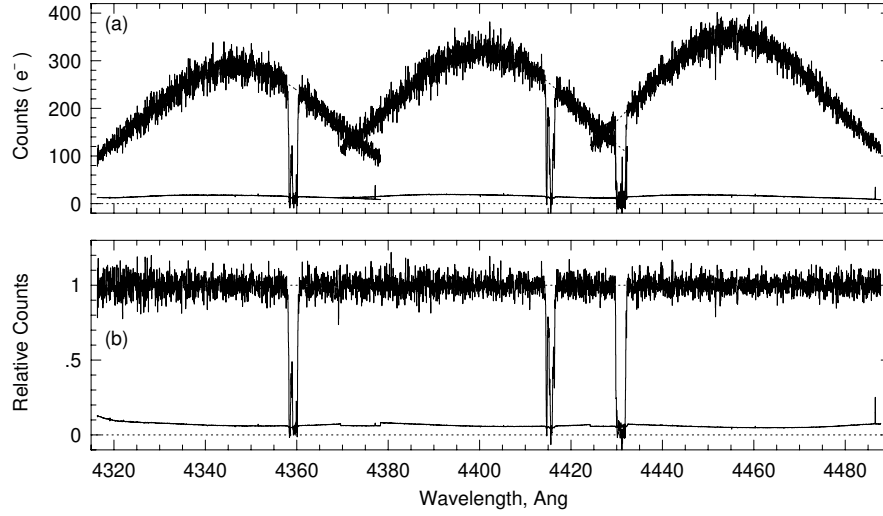


Figure 10.4 (a) Three echelle orders for which there is overlapping wavelength coverage. Each order has been continuum modeled (shown as dotted curves through the counts). (b) The resulting normalized combined spectrum. Note the slight reduction in the uncertainty spectrum in regions of overlap.

If I_{ji} is the continuum normalized counts (or calibrated flux values) and $\sigma_{I_{ji}}$ is the uncertainty in I_{ji} for pixel i in spectrum j , then the counts and the uncertainty in counts in spectrum j resampled to pixel k is given by

$$I_{kj} = \sum_{i=1}^{N_j} w_{kji} I_{ji}, \quad \sigma_{I_{kj}}^2 = \sum_{i=1}^{N_j} w_{kji} \sigma_{I_{ji}}^2, \quad (10.16)$$

where $\sum_{i=1}^{N_j} w_{kji} = 1$, and where the sums are over the N_j pixels of spectrum j . Note however, that $w_{kji} = 0$ except for the few pixels that have some wavelength overlap with pixel k of the output spectrum.

The second step, after all j spectra have been resampled to the output spectrum wavelength pixelization, is to compute the counts and uncertainty

in pixel k from the optimally weighted mean of the N_s spectra,

$$I_k = \sigma_k^2 \sum_{j=1}^{N_s} I_{kj} / \sigma_{I_{kj}}^2, \quad \sigma_k^2 = \left[\sum_{j=1}^{N_s} 1 / \sigma_{I_{kj}}^2 \right]^{-1}. \quad (10.17)$$

In Figure 10.4(a), we present three overlapping echelle orders for which individual continuum fits have been obtained. The resulting combined spectrum is shown in Figure 10.4(b) following application of Eq. 10.17. It is important to realize that the above treatment for combining individual spectra onto a common wavelength scale will have the undesirable effect of “smoothing” the flux values, which are now correlated.

10.3 Objectively Measuring Absorption Lines

Absorption (and emission) features must be statistically significant and should always be located using an objective algorithm. Ensuring statistical significance and objectivity requires proper and careful accounting of the uncertainties for the measured counts (and continuum fit) at each pixel. Thus, the uncertainty spectrum plays a critical role. Before proceeding with feature identification, it should be verified that the uncertainty spectrum is a sound representation of the fluctuations in the data.

The adopted statistical significance quantifies the detection threshold at which features are defined. Throughout, we denote the detection threshold as N_σ , for which we adopt values in the range of 3–5. An interpretation of N_σ is that it corresponds to the number of Gaussian standard deviations to which features are deemed significant, assuming that the uncertainties in the pixel counts are normally distributed. This corresponds to a confidence level of $CL = \text{erf}(N_\sigma/\sqrt{2})$, or a null-hypothesis probability of $P = 1 - CL$, where the null-hypothesis is that the feature is consistent with noise. For $N_\sigma = 3$, a 3σ significance, we obtain $P = 2.7 \times 10^{-3}$, and for $N_\sigma = 5$, we obtain $P = 5.7 \times 10^{-7}$.

There are several different quantities that can be employed for objective feature finding, including the direct measurements of the counts in each pixel. However, the fundamental quantity of each pixel is the flux decrement weighted by the pixel width, $D_i \Delta \lambda_i$. This quantity is the pixel equivalent width. We thus examine the data as an “equivalent width spectrum”. In this subsection, we describe two methods for converting the measured spectrum of pixel counts into an equivalent width spectrum and present a method useful for objectively locating and defining statistically significant absorption and/or emission features.

The first step to obtaining the equivalent width spectrum and its associated uncertainty spectrum is to compute the flux decrement and its uncertainty in each pixel,

$$D_i = 1 - \frac{I_i}{\bar{I}_i^c}. \quad (10.18)$$

In the continuum, the mean flux decrement will be $D_i = 0$ with a uncertainty,

$$\sigma_{D_i}^2 = \left(\frac{I_i}{\bar{I}_i^c} \right)^2 \left[\left(\frac{\sigma_{I_i}}{I_i} \right)^2 + \left(\frac{\sigma_{\bar{I}_i^c}}{\bar{I}_i^c} \right)^2 \right], \quad (10.19)$$

where σ_{I_i} and $\sigma_{\bar{I}_i^c}$ are the uncertainties in the counts and fitted continuum in pixel i , respectively. Thus, it is possible that $D_i < 0$. In the cores of saturated absorption lines, $D_i \simeq 1$, also with a standard deviation given by Eq. 10.19. Thus, it is possible that $D_i > 1$.

The equivalent in each pixel of the spectrum is then $e_i = D_i \Delta \lambda_i$ with uncertainty $\sigma_{e_i} = \sigma_{D_i} \Delta \lambda_i$, where wavelength interval spanned by the pixel is $\Delta \lambda_i = \lambda_{i+1} - \lambda_i$.

10.3.1 Aperture Method

The simplest method for objectively locating absorption and/or emission features in spectra is the “aperture” method (see Lanzetta, Turnshek, & Wolfe, 1987). This technique, though well suited for strong features, is not optimal for weak unresolved lines. A more sensitive technique is discussed in § 10.3.2.

Consider a spectrum for which there are p pixels per resolution element (Eq. 8.54). Since the full-width half-maximum of an unresolved feature will span p pixels, the full width will span approximately $2p$ pixels. Thus, one defines the aperture for detection to be $2p + 1$, roughly twice the number of pixels per resolution element. The equivalent width of an unresolved feature centered on pixel i is then

$$w_i = - \sum_{n=i-p}^{i+p} D_n \Delta \lambda_n, \quad (10.20)$$

Note that the negative sign in Eq. 10.20 defines w_i such that $w_i < 0$ for absorption and $w_i > 0$ for emission. The uncertainty in w_i is given by

$$\sigma_{w_i}^2 = \sum_{n=i-p}^{i+p} (\sigma_{D_n} \Delta \lambda_n)^2. \quad (10.21)$$

Since we can compute Eq. 10.20 and Eq. 10.21 for all pixels, we obtain a new spectrum that quantifies the equivalent width per resolution element. We will refer to this as the equivalent width spectrum.

Finally, we define the ratio of the equivalent width per resolution element to its uncertainty, w_i/σ_{w_i} , as the significance level. In the continuum, the condition $|w_i/\sigma_{w_i}| \leq 1$ should hold for 68.27% of the pixels, $|w_i/\sigma_{w_i}| \leq 2$ should hold for 95.45% of the pixels, and $|w_i/\sigma_{w_i}| \leq 3$ should hold for 99.73% of the pixels. This behavior suggests that as the absolute value of the significance level per resolution element increases, the probability that the pixel is consistent with the continuum (lack of an absorption or emission feature) drops.

Thus, to objectively find candidate features, one scans the equivalent width spectrum pixel by pixel starting from the bluest wavelength progressing toward redder wavelengths. The first step is to locate a candidate feature by identifying a pixel satisfying the condition

$$\frac{w_i}{\sigma_{w_i}} \leq -N_\sigma, \quad (10.22)$$

where $N_\sigma \geq 3$ for a probability of $P \leq 0.0027$ that the pixel is consistent with being drawn from the continuum. For the highest confidence level, we would adopt $N_\sigma = 5$, which yields $P = 5.7 \times 10^{-7}$. The negative sign in Eq. 10.22 applies for absorption features; the negative sign is omitted for emission features. From this point forward, we limit our discussion to absorption features.

When Eq. 10.22 is satisfied, we set $j = i$. The next step is to determine the pixels spanned by the candidate feature, i.e., to determine the “starting” pixel and “ending” pixel. The starting pixel is located by scanning the spectrum blueward, toward smaller i from pixel j , until the condition

$$\frac{w_i}{\sigma_{w_i}} \geq -1 \quad (i < j), \quad (10.23)$$

is first satisfied, at which point we set $i^- = i$, where $i^- < j$ is defined as the starting pixel of the candidate feature first identified at pixel j . The above condition simply states that the “blue edge” of the feature begins where w_{i^-} first becomes consistent with the 1σ uncertainty in the equivalent width spectrum. To find the ending pixel, the spectrum is then scanned redward, toward larger i starting from pixel j , until the condition

$$\frac{w_i}{\sigma_{w_i}} \geq -1 \quad (i > j), \quad (10.24)$$

is first satisfied, at which point we set $i^+ = i$, where $i^+ > j$ is the defined as the ending pixel of the feature first identified at pixel j .

Satisfaction of the three conditions, i.e., Eq. 10.22, followed by Eqs. 10.23 and 10.24 provide the wavelength location from pixel i^- to pixel i^+ of a candidate absorption feature. Note that pixel j , the location the feature was first identified, lies in the range $i^- < j < i^+$, but has no further significance to the search or analysis process. To determine if the candidate feature is significant we compute its significance level, defined as the ratio of the candidate feature equivalent width to its uncertainty, $SL = W/\sigma_W$, where

$$W = \sum_{i=i^-}^{i^+} D_i \Delta \lambda_i, \quad \sigma_W^2 = \sum_{i=i^-}^{i^+} (\sigma_{D_i} \Delta \lambda_i)^2. \quad (10.25)$$

respectively, where σ_{D_i} is the uncertainty in the flux decrement, D_i , as given by Eqs. 10.19 and 10.18, respectively. The candidate feature is considered statistically significant when the condition,

$$SL \geq N_\sigma, \quad (10.26)$$

is met, where, as before, N_σ is usually taken to be in the range 3–5. If the significance level of the feature is greater than or equal to the detection threshold, N_σ , the candidate feature is retained for further analysis, otherwise the candidate feature is discarded.

The scanning of the spectrum for additional features is continued starting from pixel $i^+ + 1$ (one pixel redward of the ending pixel of the previous candidate feature). A drawback of this method is that when multiple individual lines are blended into a single complex feature, the individual lines are not located separately, only the full feature is located. Breaking these into their components would require a deblending technique, such as Gaussian fitting (§ 10.7.1) or Voigt profile decomposition (§ 10.7.2)

10.3.2 An Optimized Method

The aperture method is not optimized to locate weak unresolved absorption lines. This is because the aperture method applies a simple box window that does not properly account for the function shape of the redistribution of counts (or flux) over the full-width half-maximum of the instrumental spread function (ISF, § 8.8). Schneider et al. (1993) introduced an optimized method for the purpose of objectively locating weak unresolved absorption (and emission) lines.

For computational purposes, the weighting of the redistribution of counts

across adjacent pixels is described by a discrete model, P_n , of the ISF consisting of $M = 2J_0 + 1$ elements with indices $n = 1$ to $n = M$, where J_0 is a positive integer equaling $2p$, twice the number of pixels per resolution element (see § 8.9). The indexing of the ISF model is defined such that peak value occurs at $n = J_0 + 1$. Note that M is an odd integer, providing the discrete model with an equal even number of pixels either side to the center pixel.

For simplicity of example, we will assume that the ISF is modeled as a Gaussian function,

$$P_n = A \exp \{-x_{ki}^2\}, \quad (10.27)$$

where $x_{ki} = (\lambda_i - \lambda_k)/\sigma_i^{\text{ISF}}$, and the “convolution” index,

$$k = i + (n - 1) - J_0, \quad (10.28)$$

maps how counts from pixel k are redistributed to pixel i as weighted by the model ISF at index n . The normalization constant is $1/A = \sum_{n=1}^M \exp\{-x_{ki}^2\}$, which yields,

$$\sum_{n=1}^M P_n = 1, \quad (10.29)$$

where $P_n > 0$ everywhere. From the definition of the spectrograph resolving power (Eq. 8.27), the model ISF full-width half-maximum at λ_i is $\Delta\lambda_i = \lambda_i/R$, where R is the spectrograph resolution. Thus, for a Gaussian distribution, $\sigma_i^{\text{ISF}} = \lambda_i/(2\sqrt{2\ln 2}R)$. The important point is that the full-width half-maximum of the ISF model increases linearly with wavelength and is not a constant.

The ISF weighted equivalent width per resolution element, w_i , is obtained by weighting the flux decrement of adjacent pixels by the ISF,

$$w_i = -\frac{\Delta\lambda_i}{P^2} \sum_{n=1}^M P_n D_k, \quad (10.30)$$

where $P^2 = \sum_{n=1}^M P_n^2$, and where D_k is the flux decrement, given by Eq. 10.18, and $\Delta\lambda_i = \lambda_{i+1} - \lambda_i$ is the wavelength interval spanned by pixel i . Note that, similar to the aperture method, the negative sign in Eq. 10.30 defines $w_i < 0$ for absorption and $w_i > 0$ for emission. The uncertainty in w_i is given by

$$\sigma_{w_i} = \frac{\Delta\lambda_i}{P^2} \left(\sum_{n=1}^M P_n^2 \sigma_{D_k}^2 \right)^{1/2}, \quad (10.31)$$

where $\sigma_{D_k}^2$ is given by Eq. 10.19. The σ_{w_i} serve as the 1σ equivalent width detection threshold at each pixel.

To obtain the full spectrum of the ISF weighted equivalent widths per resolution element, the ISF model is centered on each pixel and the calculations of Eqs. 10.30 and 10.31 are carried out. Note that the P_n depend upon the wavelength of the current pixel i , and therefore should be recomputed at each i . Objectively locating statistically significant candidate features using Eqs. 10.30 and 10.31 then follows the same procedure as described for the aperture method (Eq. 10.22 through Eq. 10.26). The difference is that the ISF weighting method is optimized to locate weak, unresolved features.

10.3.3 The Line List

The outcome of the aperture or optimized method is a list of objectively located absorption features, each with a starting pixel location λ_{i-} and an ending pixel location λ_{i+} . The equivalent width and its uncertainty for each feature are computed from Eq. 10.25.

The mean wavelength is computed by weighting the equivalent width in each pixel,

$$\langle\lambda\rangle = \frac{1}{W} \sum_{i=i-}^{i+} \lambda_i D_i \Delta\lambda_i = \frac{W^{(1)}}{W}, \quad (10.32)$$

where $W^{(1)}$ is the first moment of the equivalent width. The uncertainty in $\langle\lambda\rangle$ is obtained by partial differentiation of Eq. 10.32,

$$\sigma_{\langle\lambda\rangle}^2 = \sum_{i=i-}^{i+} \left[\frac{\partial\langle\lambda\rangle}{\partial D_i} \right]^2 \sigma_{D_i}^2 = \sum_{i=i-}^{i+} \left[\frac{1}{W} \frac{\partial W^{(1)}}{\partial D_i} - \frac{W^{(1)}}{W^2} \frac{\partial W}{\partial D_i} \right]^2 \sigma_{D_i}^2, \quad (10.33)$$

where

$$\frac{\partial W^{(1)}}{\partial D_i} = \lambda_i \Delta\lambda_i, \quad \frac{\partial W}{\partial D_i} = \Delta\lambda_i. \quad (10.34)$$

In final form we have,

$$\sigma_{\langle\lambda\rangle}^2 = \sum_{i=i-}^{i+} \left[\frac{\lambda_i W - W^{(1)}}{W^2} \right]^2 (\sigma_{D_i} \Delta\lambda_i)^2. \quad (10.35)$$

In Figure 10.5(a), a typical quasar spectrum and its uncertainty spectrum are shown with objectively identified absorption features at the $N_\sigma = 5$ detection threshold. The equivalent width spectrum and its uncertainty spectrum, based on the optimal method, are shown in Figure 10.5(b). The ma-

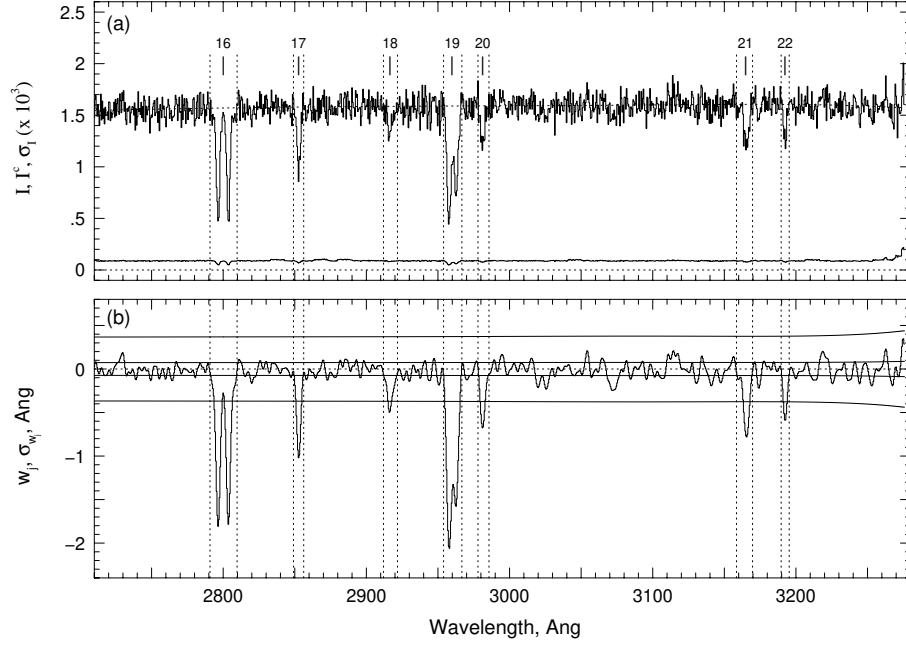


Figure 10.5 (a) An example quasar spectrum with objectively found absorption features with $N_\sigma = 5$. Ticks above the continuum identify the mean wavelength of each feature as computed from Eq. 10.32. The dotted vertical lines mark λ_{i-} and λ_{i+} , the pixel ranges over which the feature equivalent widths are measured. (b) The equivalent width spectrum. The smooth curves above and below the spectrum provide the $\pm 1 \sigma$ and $\pm 5 \sigma$ uncertainties, respectively. Regions meeting the condition given by Eq. 10.26 with $N_\sigma = 5$ are objectively identified. See Table 10.1 for the list of measured quantities.

jority of the equivalent widths scatters about $w_i = 0$, except in the regions of significant features. The inner smooth lines straddling the equivalent width spectrum are the $\pm 1 \sigma$ uncertainties and the outer smooth lines are the $\pm 5 \sigma$ uncertainties. In principle, 68% of the pixels should have w_i with values ranging between the 1σ uncertainties (once significant features are excluded). Candidate features are located at the $N_\sigma = 5$ detection threshold wherever Eq. 10.22 is satisfied. The starting and ending pixels of the features, shown as the vertical dashed lines in Figure 10.5, are defined where the equivalent widths in adjacent pixels first become consistent with the 1σ uncertainties (i.e., the first adjacent pixels satisfy the conditions of Eqs. 10.23 and 10.24). Candidate features are retained for further analysis when they satisfy the $N_\sigma = 5$ detection threshold condition of Eq. 10.26.

In Table 10.1, an example objective line list is presented for the features

numbered 16 through 22 illustrated in Figure 10.5. The measured quantities included are the mean wavelength $\langle\lambda\rangle$, mean wavelength uncertainty $\sigma_{\langle\lambda\rangle}$, equivalent width W , equivalent width uncertainty σ_W , and significance level SL. For record keeping and possible further analysis, it is useful to also include the starting wavelength λ^- , ending wavelength λ^+ . Here, we have used the notation that $\lambda^- = \lambda_{i-}$ and $\lambda^+ = \lambda_{i+}$.

Table 10.1 *Example Objective Feature List^a ($N_\sigma = 5$)*

ID	$\langle\lambda\rangle$	$\sigma_{\langle\lambda\rangle}$	W	σ_W	SL	λ^-	λ^+
16	2800.06	0.22	4.50	0.16	28.12	2790.29	2809.72
17	2852.90	0.25	1.05	0.10	10.07	2848.58	2855.74
18	2916.65	0.53	0.75	0.13	5.92	2911.48	2921.71
19	2959.84	0.12	4.43	0.12	36.35	2953.41	2966.71
20	2981.57	0.41	0.72	0.11	6.44	2977.45	2985.63
21	3165.13	0.46	1.12	0.13	8.46	3157.96	3169.72
22	3192.96	0.42	0.50	0.10	5.16	3189.15	3195.29

^a Wavelengths and equivalent widths in units of Å.

Inspection of features 18 and 22 in Table 10.1 and Figure 10.5(a) provides insight into the $N_\sigma = 5$ detection threshold, for the SL of these features are only slightly above this level. Note, for feature 22, that $w_i \sim W_i$ at the feature center; this is a characteristic of unresolved lines in that the equivalent width per resolution element at the center of an unresolved feature should be equal the feature equivalent width.

10.4 Absorption Line Systems

We define an absorption system as a gas structure giving rise to multiple absorption lines in a spectrum. Some absorption systems may give rise to a dozen or more lines representing multiple transitions from various atomic and ionic species. For example, one might detect cosmologically redshifted MgII $\lambda\lambda 2796, 2803$ doublet absorption along with associated absorption from MgI $\lambda 2852$, CII $\lambda 1036$, FeII $\lambda 2344, 2382, 2600$, SiII $\lambda 1526$, and SiIV $\lambda\lambda 1393, 1402$ all at the same redshift. Other systems may give rise to only a fine-structure doublet from one atomic species, such as the CIV $\lambda\lambda 1548, 1550$ or the OVI $\lambda\lambda 1031, 1037$ doublet. In order for an absorption system to be unambiguously established at least two to three lines must be detected. That is, either a fine-structure doublet or three singlet absorp-

tion lines must be identified such that the relationship between their rest-frame wavelengths, λ_r , and their observed wavelengths, λ , are all consistent with the same cosmological redshift according to

$$\lambda = \lambda_r(1 + z). \quad (10.36)$$

Once a candidate absorption system is identified and its redshift estimated, it is a simple task to apply Eq. 10.36 to examine whether other absorption lines in the spectrum are associated with the system by trial and error. However, one must first identify the candidate systems in a spectrum or, as the case of survey work, in multiple spectra. Numerous methods have been developed to find absorption systems in cosmological spectra, some of which are well suited for large surveys (e.g., Nestor et al., 2005; Prochter et al., 2006; Zhu & Ménard, 2013). Here, we present one approach that is well-suited for illustrating the basic concepts with some focus on fine-structure doublets.

Using fine-structure doublets from Group IA ions to locate absorption systems is a standard approach with a long and rich history. The method has several advantages. First, the sodium-like ion of MgII and the lithium-like ions of CIV and OVI are highly abundant in the interstellar and circumgalactic medium. Second, the fine-structure doublets in these ions are resonant transitions (large oscillator strengths) from the ground-states (the excitation state most highly populated in these astrophysical environments). These two facts mean that doublet absorption from MgII, CIV, and OVI is common and strong in cosmological spectra. Third, the transitions of the doublets from each ion have a unique wavelength separation, $\Delta\lambda$, and can thus be uniquely identified in absorption spectra.

To find redshifted fine-structure doublets in absorption, we will utilize the equivalent width spectrum, w_λ and its uncertainty spectrum, σ_{w_λ} , as computed from Eqs. 10.20 and 10.21 for the aperture method or Eqs. 10.30 and 10.31 for the optimized method. Recall that the equivalent width spectrum accounts for the wavelength redistribution in the recorded counts by the instrumental spread function so that in a given pixel, this spectrum provides the equivalent width for an unresolved line. Also recall that we define absorption to have $w_\lambda < 0$, and that candidate absorption features are located by $w_\lambda < -N_\sigma\sigma_{w_\lambda}$, where N_σ provides the detection confidence level in units of Gaussian standard deviations.

Consider a doublet with transition wavelengths $\lambda_r^{(a)}$ and $\lambda_r^{(b)}$. We will assume $\lambda_r^{(a)} < \lambda_r^{(b)}$ and that the oscillator strengths obey $f^{(a)} > f^{(b)}$. We hereafter will refer to transition (a) as the “primary transition”. We wish to determine the redshifts where the primary transition of the doublet is

detected at the $N_\sigma^{(a)}$ level and the secondary member is detected at $N_\sigma^{(b)}$ level. On the linear part of the curve of growth, the ratio of equivalent widths is $W^{(a)}/W^{(b)} \propto f^{(a)}/f^{(b)}$; thus, we employ $N_\sigma^{(a)}/N_\sigma^{(b)} > f^{(a)}/f^{(b)}$.

If we consider the spectrum on a pixel by pixel basis, at observed wavelength λ_i corresponding to pixel i , the cosmological redshift of the absorbing gas would be

$$z = \frac{\lambda_i}{\lambda_r^{(a)}} - 1, \quad (10.37)$$

and the observed wavelength of the second member of the doublet would be

$$\lambda_j = \lambda_r^{(b)}(1 + z) = \lambda_i \frac{\lambda_r^{(b)}}{\lambda_r^{(a)}}, \quad (10.38)$$

at pixel j , where $j > i$. At a minimum, we aim to have the conditions $w_i < -N_\sigma^{(a)}\sigma_{w_i}$ and $w_j < -N_\sigma^{(b)}\sigma_{w_j}$ met simultaneously. We thus formulate a “doublet autocorrelation spectrum” by computing

$$X_D(z) = H(|w_i| - N_\sigma^{(a)}\sigma_{w_i})H(|w_j| - N_\sigma^{(b)}\sigma_{w_j}) \quad (10.39)$$

where $H(x)$ is the Heaviside step function for which $H(x) = 1$ for $x \geq 0$ and $H(x) = 0$ for $x < 0$. We commonly adopt $N_\sigma^{(a)} = 5$ and $N_\sigma^{(b)} = 3$. The doublet autocorrelation spectrum has values $X_D(z) = 1$ where candidate double absorption is found and $X_D(z) = 0$ otherwise. Note that $X_D(z) = 1$ can occur by chance, particularly in spectral regions with many absorption features.

For additional robustness, one can multiply $X_D(z)$ by a “richness” factor, which yields an integer equal to the number of detected candidate associated lines in addition to the doublet,

$$X(z) = R(z)X_D(z), \quad R(z) = 1 + \sum_{t=3}^{N_t} H(|w_{i_t}| - N_\sigma^{(t)}\sigma_{w_{i_t}}), \quad (10.40)$$

where pixel i_t corresponds to wavelength $\lambda_{i_t} = \lambda_r^{(t)}(1 + z)$, t represents the transition number in the list of N_t transitions being searched, and where we reserve $t = 1$ and $t = 2$ to denote the primary and secondary transition of the doublet.

Generalizing, the adopted absorption system autocorrelation spectrum might be written as some permutation of Eqs. 10.39 or 10.40, for example,

$$X(z) = \left[\prod_{t=1}^{N_A} H(|w_{i_t}| - N_\sigma^{(t)}\sigma_{w_{i_t}}) \right] \left[1 + \sum_{t=N_B}^{N_t} H(|w_{i_t}| - N_\sigma^{(t)}\sigma_{w_{i_t}}) \right], \quad (10.41)$$

where the product is taken over the N_A transitions in a priority line list, and the sum from $N_B = N_A + 1$ is taken over the remaining transition in the line list, and where $t = 1$ is reserved for the primary transition for purposes of defining the candidate redshift of the system. Note that the priority line list is not necessarily a fine-structure doublet. We emphasize that Eqs. 10.39, 10.40, and 10.41 represent only one approach to locating absorption systems; ultimately, the method one adopts should be optimized for the science objectives.

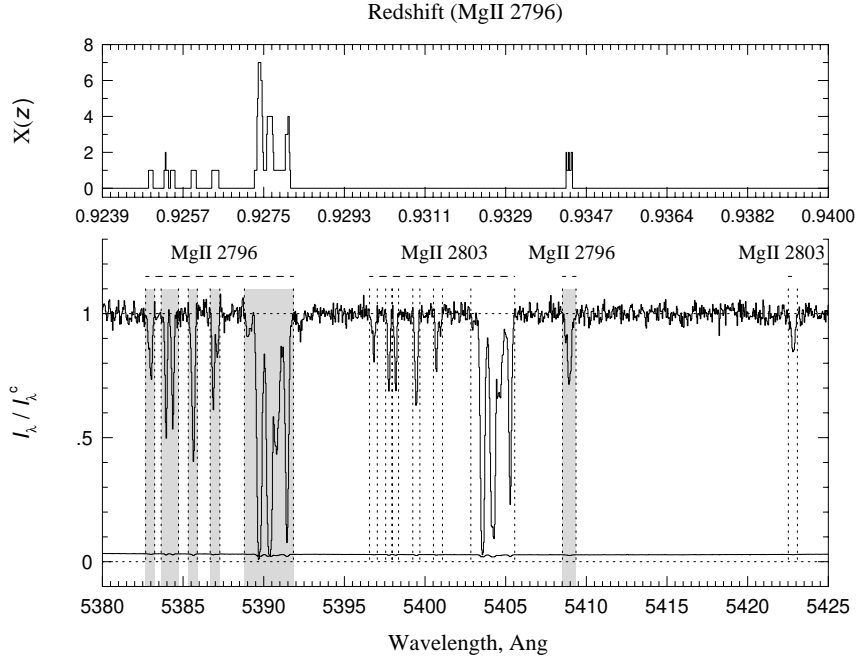


Figure 10.6 The absorption line autocorrelation spectrum (Eq. 10.40) for the quasar spectrum Q1206+456 showing identification of a rich MgII $\lambda\lambda 2796, 2803$ absorption system at $z \simeq 0.9275$ with up to six additional associated absorption lines. There are also several doublet-only system. For each detected doublet, the primary transition (MgII $\lambda 2796$) is shown in shaded gray. The vertical dashed lines provide the wavelength ranges of the objectively defined features, including the secondary MgII $\lambda 2803$ transition spanning 5396–5406 Å.

As false-positives will likely occur for the evaluation of the Heaviside function of various transitions at different redshifts (due to chance spectral alignments of an unrelated absorption features), it is useful to apply secondary filters while executing Eq. 10.39, 10.40, or 10.41. For example, consider the commonly observed FeII $\lambda\lambda 2344, 2374, 2382, 2687$, and 2600 transitions. The oscillator strengths are $f = 0.1097, 0.0282, 0.3006, 0.0646$,

and 0.2239, respectively. If a given FeII transition is detected at a given redshift, and it is not saturated, its measured $|w_{i_t}|$ should be in proportion to its oscillator strength as compared to the other FeII transitions. In other words, for example, $|w_{i_t}|$ for FeII $\lambda 2382$ should be the largest of all the FeII transitions, and it should be a factor of 1.3 ($= 0.3006/0.2239$) greater than $|w_{i_t}|$ for FeII $\lambda 2600$ and a factor of 2.7 greater than for FeII $\lambda 2344$. We thus see that the relative oscillator strengths of transitions from the same ionic species can be exploited to flag unphysical line strengths and/or eliminate false-positives while computing Eq. 10.41.

In the final analysis, the use of secondary filters has to be judiciously applied depending on the science objectives, noise characteristics of the spectra, and the density of absorption features in the spectral region being searched. For example, in spectra of high redshift quasars, the use of Eq. 10.41 will not be highly effective at wavelengths blueward of the Ly α emission of the quasar as the number of Ly α forest lines is very high. At some level visual inspection is required to verify any automated objective algorithm for locating absorption systems. A complete treatment of the process would include Monte-Carlo modeling to obtain an estimate the percentage of false positives for the adopted search methods (e.g., Nestor et al., 2005).

In Figure 10.6, we show the results of the application of Eq. 10.40, which is a doublet autocorrelation including the richness factor, $R(z)$, for the quasar spectrum Q1206+459. The primary transition is MgII $\lambda 2796$ of the MgII $\lambda\lambda 2796, 2803$ doublet. Transitions included in the richness factor are MgI $\lambda 2852$, FeII $\lambda 2344, 2374, 2382, 2687, 2600$, and MnII $\lambda 2576, 2594, 2606$. The primary transition detection threshold is $N_\sigma = 5$, whereas all other transitions have been detected at the $N_\sigma = 3$ significance level.

In the redshift range $0.9249 \leq z \leq 0.9343$, seven MgII doublets were located. The dominant absorption at $z \simeq 0.9275$ has $X(z) = 7$ in four pixels, indicating that six of the nine transitions included in the richness factor were also detected. Note that subcomponent structure in the richness is revealed. Also note several one-pixel false positives in the $z \simeq 0.9343$ system. Inspection of the richness factor for this redshift shows chance alignment of the Heaviside function for FeII $\lambda 2374$ with the FeII $\lambda 2383$ feature at $z \simeq 0.9275$. As FeII $\lambda 2374$ has the weakest oscillator strength of included FeII transitions, and stronger FeII transitions were not detected at $z \simeq 0.9343$, and the strength of the feature fits in the hierarchy of strength for the detected FeII transitions at $z \simeq 0.9275$, we can inductively rule out FeII $\lambda 2374$ absorption at $z \simeq 0.9343$.

The rich $z \simeq 0.9275$ system is in close redshift proximity to multiple systems ranging from $0.9249 \leq z \leq 0.9281$ (shown as the shaded MgII $\lambda 2796$ fea-

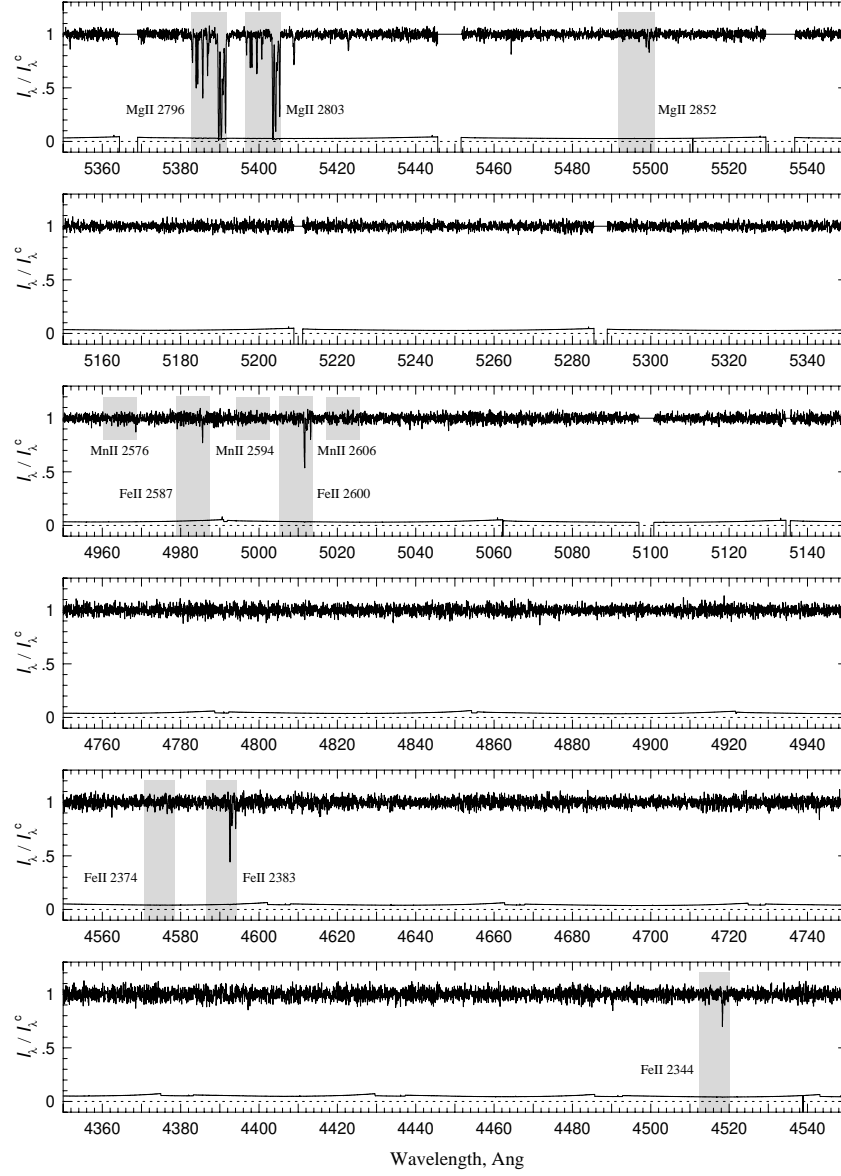


Figure 10.7 Spectral locations of the transitions included in the autocorrelation spectrum for the $z = 0.9275$ MgII-selected absorption system shown in Figure 10.6. Shading spans the spectral regions associated with the full redshift window of the primary transition, MgII $\lambda 2796$. Detected associated transitions have full shading and non-detections have truncated shading.

tures in Figure 10.6). The absorption profiles give the impression of being physically distinct gas complexes separated in line-of-sight velocity. How-

ever, how one defines a single absorption system is not always obvious in principle, but in practice a maximum velocity spread is arbitrarily adopted. For example, a system can be defined to include all absorption within an arbitrarily chosen peculiar velocity window

$$\Delta v = c \frac{z^+ - z^-}{1 + \bar{z}}, \quad (10.42)$$

where $z^- = \lambda^-/\lambda_r - 1$ and $z^+ = \lambda^+/\lambda_r - 1$ are the lowest and highest redshift extent of the absorption of the primary transition, and $\bar{z} = (z^+ + z^-)/2$. A maximum Δv spread in the range of 500–1000 km s⁻¹ is typical. Another method of defining the velocity extent of absorption systems would be to enforce a minimum velocity or redshift separation between absorption clustered in redshift. For example, one could adopt a minimum velocity separation of 1000 km s⁻¹ between absorption. The choice of these definitions should be based upon the science objectives, and, if possible, some knowledge of the astrophysical environment of the absorbing gas, such as whether one is probing an individual galaxy, cluster of galaxies, or a void or filament in the intergalactic medium.

If we group the distinct redshifts from $z^- \simeq 0.9249$ to $z^+ \simeq 0.9281$ into a single absorption system, we find $\Delta v \simeq 500$ km s⁻¹. Inclusion of the $z \simeq 0.9343$ system would yield a single absorption system with $\Delta v \simeq 1460$ km s⁻¹. For purposes of illustrating further analysis, we adopt $\Delta v \leq 800$, thus grouping the multiple MgII absorption features from $0.9249 \leq z \leq 0.9281$ into a single system. Hereafter, we refer to the full system absorption profile as a feature and the individual absorption profiles as subfeatures.

In Figure 10.7, we show the Q1206+459 spectrum over the wavelength range 4350 Å to 5550 Å and illustrate the observed locations of the transitions associated with the MgII doublet-selected absorption system at $z \simeq 0.9275$. The shaded areas guide the eye to the spectral regions corresponding to the transitions that contributed to the autocorrelation function (Eq. 10.40) at $z \simeq 0.9275$.

Presented in Figure 10.8 are the absorption profiles of the MgII doublet selected-absorber shown in Figures 10.6 and 10.7. There are $N_k = 5$ formally detected subfeatures (shaded regions). The subsystems are identified by index number, $k = 1$ –5. For each transitions presented in Figure 10.8, the equivalent width spectrum (per resolution element, histogram; Eq. 10.30) is shown for the same wavelength region. The smooth curves above and below $w_\lambda = 0$ Å are the $\pm 1 \sigma$ (inner) and $\pm 3 \sigma$ (outer) uncertainties in the w_λ (Eq. 10.31), except for the primary transition, for which $\pm 5 \sigma$ uncertainties apply. The vertical dotted lines define λ_k^- , the starting wavelength of the k^{th}

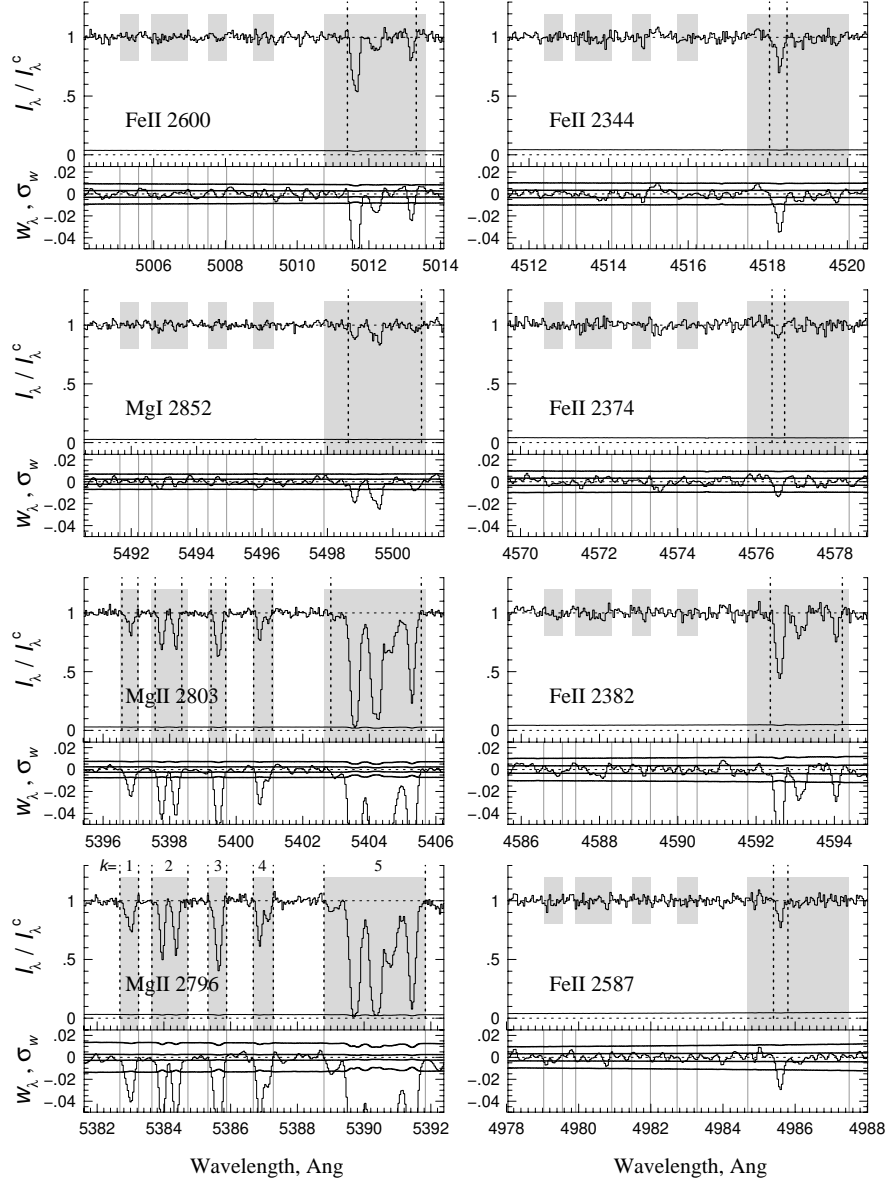


Figure 10.8 Detected transitions associated with the MgII-selected absorption system. For MgII $\lambda 2796$ (the primary transition of the primary species), five subfeatures are found (shaded gray and labeled $k = 1-5$). For each transition, the dotted vertical lines mark $\lambda_{t,k}^-$ and $\lambda_{t,k}^+$, the beginning and ending pixelized wavelengths for each subfeature k of transitions t . Full shaded regions correspond to the subfeature with detected absorption and truncated shaded regions indicate non-detections corresponding to a subfeature. Below each transition is the equivalent width spectrum (Eq. 10.30). The smooth curves straddling each spectrum provide the $\pm 1 \sigma$ and $\pm 3 \sigma$ uncertainties, except for MgII $\lambda 2796$, which shows $\pm 1 \sigma$ and $\pm 5 \sigma$ uncertainties (Eq. 10.31). Gray vertical lines demarcate the primary transition's subfeature spectra regions.

kinematic subsystem, and λ_k^+ , the ending wavelength of the k^{th} kinematic subsystem.

10.4.1 Systemic Redshift

The systemic redshift, z_{sys} , of an absorption system can be determined in a number of various ways and there is no best or incorrect approach to establishing this quantity. The main consideration that one should account for is that the rest-frame velocity zero point is established by the systemic redshift. According to the non-relativistic Doppler relationship, line-of-sight velocity shifts Δv induce observed wavelength shifts $\Delta \lambda$ via $\Delta v/c = \Delta \lambda/\lambda$. Consider a transition with rest-frame wavelength λ_r that is observed in absorption at wavelength λ and the absorbing gas is associated with an absorption system with systemic redshift z_{sys} . The relationship between the observed and rest-frame wavelengths is $\lambda = \lambda_r(1 + z)$, where $|z - z_{\text{sys}}| > 0$. We thus have $\Delta \lambda = \lambda_r(1 + z) - \lambda_r(1 + z_{\text{sys}}) = \lambda_r(z - z_{\text{sys}})$. Thus, in terms of the redshifts, the line-of-sight velocity of the absorbing gas in the rest-frame of the system is

$$v = c \frac{\Delta \lambda}{\lambda} = c \frac{z - z_{\text{sys}}}{1 + z_{\text{sys}}} . \quad (10.43)$$

If the absorbing material is measured relative to an emission line object, such as a galaxy, AGN, or quasar, then, depending on the science objectives, the systemic redshift can be set to the emission redshift. If one wishes to define the systemic redshift in terms of the absorption itself, then one method is to use the primary transition ($t = 1$) of the primary species ($s = 1$) and determine the observed wavelength that will serve as the velocity zero point.

One method is to determine the wavelength of the median of the flux decrement distribution of the primary transition. This wavelength is obtained by solving the implicit equation,

$$\int_{\lambda^-}^{\lambda_r(1+z)} D(\lambda) d\lambda = \int_{\lambda_r(1+z)}^{\lambda^+} D(\lambda) d\lambda , \quad (10.44)$$

for $z = z_{\text{abs}}$. The observed wavelength at the apparent flux decrement where λ^- and λ^+ are the extreme blueward and redward wavelengths of the full absorption profile of the primary transition. This definition places the redshift at the location where there is equal equivalent width on either side of $\lambda = \lambda_r(1 + z_{\text{abs}})$.

Eq. 10.44 can be evaluated by numerical integration using interpolation techniques. Alternatively, one could carry out Eq. 10.44 as a summation by

scanning the spectrum to locate the pixel $\langle i \rangle$ that minimizes

$$f = \sum_{i_{1,1}^-}^{\langle i \rangle} w_i D_i \Delta \lambda_i - \sum_{\langle i \rangle}^{i_{1,1}^+} w_i D_i \Delta \lambda_i \simeq 0, \quad (10.45)$$

where, if so desired, the weight w_i can be applied as a pixel mask ($w_i = 0, 1$), or optimizing weight ($w_i = 1/\sigma_{D_i}^2$).

As, the optical depth is directly related to the column density of the absorbing ion, whereas the flux decrement is a consequence of radiative transfer, one may prefer to solve Eq. 10.44 or 10.45 using optical depth, $\tau_\lambda = -\ln(I_\lambda/\bar{I}_\lambda^c)$, in place the flux decrement; the result will that z_{abs} will be established at the optical depth median of absorption profile of the primary transition. This substitution requires some care as $I_\lambda \leq 0$, which can happen due to the noise in the recorded counts, yields an unphysical optical depth. In fact, $I_\lambda < \sigma_{I_\lambda}$ is technically an upper limit on the optical depth, so that one should substitute $\sigma_\lambda/\bar{I}_\lambda^c$ when $I_\lambda \leq \sigma_{I_\lambda}$.

Hereafter, we adopt z_{sys} to designate the systemic redshift, without regard to whether it was assigned from an associated emission line object or determined from the absorption profile of the primary transition of the primary species. Applying Eq. 10.44, to the rich MgII selected-absorption system in the spectrum of PG 1206+459 (illustrated in Figures 10.6, 10.7, 10.8), we obtain $z_{\text{sys}} = 0.927668$.

10.4.2 Equivalent Widths and Mean Wavelengths

Having identified the transitions giving rise to the absorption features and their subfeatures, we can proceed to quantify the absorption properties. Within a given velocity range, each transition may have detected subfeatures, or may have a no detections. We discuss methods for quantifying individual subfeature absorption properties, which provide the ability to study variations with velocity in the absorption properties of the system.

Adopting the notation that subfeature k of transition t is measured over pixel indices $i_{t,k}^-$ to $i_{t,k}^+$, the observed equivalent width and its uncertainty are

$$W_{t,k} = \sum_{i=i_{t,k}^-}^{i_{t,k}^+} D_i \Delta \lambda_i, \quad \sigma_{W_{t,k}}^2 = \sum_{i=i_{t,k}^-}^{i_{t,k}^+} (\sigma_{D_i} \Delta \lambda_i)^2. \quad (10.46)$$

The mean wavelength subfeature k of transition t is computed from

$$\langle \lambda \rangle_{t,k} = \frac{1}{W_{t,k}} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} \lambda_i D_i \Delta \lambda_i = \frac{W_{t,k}^{(1)}}{W_{t,k}}, \quad (10.47)$$

where $W_{t,k}^{(1)}$ is the first moment of the subfeature equivalent width. The uncertainty in $\langle \lambda_{t,k} \rangle$ is obtained by partial differentiation of Eq. 10.47. Following the steps employed to derive Eq. 10.35, we have

$$\sigma_{\langle \lambda \rangle_{t,k}}^2 = \sum_{i=i_{t,k}^-}^{i_{t,k}^+} \left[\frac{\lambda_i W_{t,k} - W_{t,k}^{(1)}}{W_{t,k}^2} \right]^2 (\sigma_{D_i} \Delta \lambda_i)^2. \quad (10.48)$$

For a complete examination of the subfeatures, one obtains the limits on non detections. In the velocity window where at least one transition has a detected subfeature, the upper limits on the other transitions in this velocity window are obtained from

$$W_{t,k} < N_\sigma \sigma_{W_{t,k}}, \quad (10.49)$$

where $\sigma_{W_{t,k}}$ is computed using Eq. 10.46 where the summation is over the pixel indices that correspond to the starting and ending velocity of the detected subfeature. The factor $N_\sigma = 3$ is usually adopted to increase the confidence level to 99.73% from 68.24% ($N_\sigma = 1$). For example, in the $z_{\text{sys}} = 0.9277$ absorption system in the spectrum of PG 1206+459 shown in Figure 10.9, five subfeatures of MgII transitions are detected, but only a single feature is detected for the remaining transitions. By summing Eq. 10.46 over the velocity ranges of these other four subfeatures, the upper limits on these non-detected subfeatures can be quantified.

In Table 10.2, we present the rest-frame equivalent widths for the $z_{\text{sys}} = 0.9277$ absorption system in the spectrum of PG 1206+459 for the five subfeatures detected in MgII $\lambda 2796$ and MgII $\lambda 2803$. We present the equivalent width limits for three selected transitions of MgI and FeII. The corresponding absorption profiles are shown in Figure 10.9. As we will show in § 10.6, equivalent width limits can be converted to column density limits. The rest-frame equivalent widths are $W_r = W_{t,k}/(1+z_{\text{sys}})$, where the observed equivalent widths are computed using Eqs. 10.46 and 10.49.

Once the subfeature quantities are measured, the equivalent width and its uncertainty for transition t is then

$$W_t = \sum_{k=1}^{N_k} W_{t,k}, \quad \sigma_{W_t}^2 = \sum_{k=1}^{N_k} \sigma_{W_{t,k}}^2. \quad (10.50)$$

Table 10.2 *Equivalent Widths: PG 1206+459 $z_{\text{sys}} = 0.9277$*

k	MgII 2796 $W_r, \text{\AA}$	MgI 2853 $W_r, \text{\AA}$	FeII 2600 $W_r, \text{\AA}$	FeII 2374 $W_r, \text{\AA}$
1	0.031 ± 0.001	< 0.003	< 0.004	< 0.004
2	0.078 ± 0.001	< 0.003	< 0.004	< 0.004
3	0.066 ± 0.001	< 0.003	< 0.004	< 0.004
4	0.052 ± 0.001	< 0.003	< 0.004	< 0.004
5	0.648 ± 0.002	0.061 ± 0.003	0.089 ± 0.002	0.011 ± 0.001
Full	0.875 ± 0.003	0.061 ± 0.003	0.089 ± 0.002	0.011 ± 0.001

A benefit of the formalism of measuring the subfeatures individually and then summing them to obtain the full feature quantities is that continuum pixels between subfeatures are excluded from the sums. This approach naturally omits pixels with flux decrements consistent with the noise level in the spectrum while minimizing the uncertainty in the measured quantities because fewer pixel uncertainties are being summed in Eq. 10.50.

The mean wavelength of the feature is computed by weighting the equivalent width in each pixel,

$$\langle \lambda \rangle_t = \frac{1}{W_t} \sum_{i=1}^{N_t} W_{t,k}^{(1)} = \frac{W_t^{(1)}}{W_t}, \quad (10.51)$$

where $W_t^{(1)}$ is the first moment of the full feature equivalent width. The uncertainty in $\langle \lambda \rangle_t$ is obtained by partial differentiation of Eq. 10.51,

$$\sigma_{\langle \lambda_t \rangle}^2 = \sum_{k=1}^{N_t} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} \left[\frac{\lambda_i W_t - W_t^{(1)}}{W_t^2} \right]^2 (\sigma_{D_i} \Delta \lambda_i)^2. \quad (10.52)$$

In the last row of Table 10.2, we present the system total rest-frame equivalent widths for the $z_{\text{sys}} = 0.9277$ absorption system in the spectrum of PG 1206+459.

If a transition of interest is not detected in a region of the spectrum where it is predicted to be observed, one can determine the upper limit on its equivalent width by applying

$$W_t < N_\sigma \sigma_{W_t}, \quad (10.53)$$

where, as stated above, the value of N_σ is subjective. For example, the full wavelength coverage of the spectrum of PG 1206+459 covers the expected wavelengths where the MnII $\lambda 2576, 2595, 2606$ triplet would be observed for the $z_{\text{sys}} = 0.9277$ absorption system. Application of Eq. 10.49 applied to the five subfeature velocity ranges for, for example, the un-detected MnII $\lambda 2587$ transition yields $W_r = 0.004, 0.004, 0.003, 0.004, 0.008 \text{ \AA}$ for $k = 1$ to $k = 5$, respectively and application of Eq. 10.53 yields $W_r < 0.011 \text{ \AA}$.

10.5 Kinematics

10.5.1 Rest-Frame Velocities

Once the systemic or absorption redshift is determined, the observed wavelength scale can be converted to a rest-frame velocity scale using the Doppler relation. For a given wavelength, λ_i , calibrated at pixel i , the corresponding rest-frame velocity scale for a transition with rest-frame wavelength λ_r follows from Eq. 10.43,

$$v_i = c \frac{\lambda_i - (1 + z_{\text{sys}})\lambda_r}{(1 + z_{\text{sys}})\lambda_r}. \quad (10.54)$$

The velocity interval spanned by the corresponding pixel is

$$\Delta v_i = v_{i+1} - v_i \quad (10.55)$$

One advantage of transforming the absorption profiles for each transition to a common rest-frame velocity is that they now align and can be visually inspected on a common scale. The main advantage of a common velocity scale, however, is that it provides a direct view of the complex run of the flux decrement as a function of line-of-sight velocity. Thus, the detailed kinematics can now be quantified for each kinematic subsystem for each transition.

In Figure 10.9, transitions from the ionic species MgII, MgI, and FeII associated with the $z_{\text{sys}} = 0.9277$ MgII-selected absorption system in the spectrum of PG 1206+459 shown in Figures 10.6 and 10.8 are presented in rest-frame velocity space. The subfeatures indexed $k = 1-5$, are shown for the primary transition of the primary species, MgII $\lambda 2796$.

10.5.2 Velocity Moments

The velocity moments of each subfeature and of the total system are useful for characterizing the flux decrement weighted velocity distributions of the absorbing gas. The first four moments provide the mean velocity, $\langle v \rangle$, the

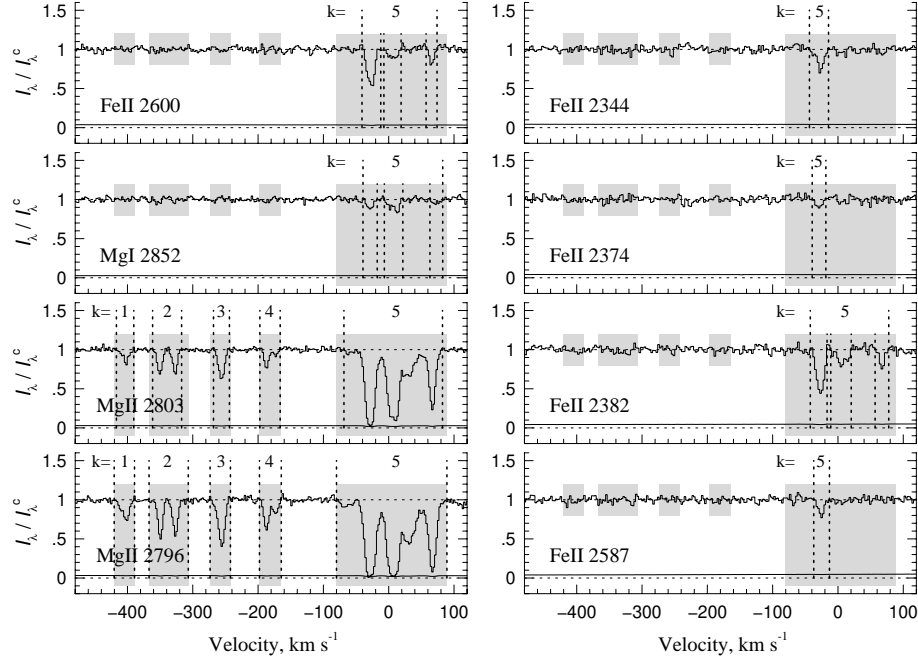


Figure 10.9 Same as Figure 10.8 except all absorption profiles are converted to rest-frame velocity for a systemic redshift of $z_{\text{sys}} = 0.9277$ based on the optical depth mean.

velocity spread, ω_v , and the velocity asymmetry, A_v . One can think of these as the mean, standard deviation, and skew of the profile kinematics. Though these quantities are computed directly from the flux decrements in the spectra, they essentially quantify the kinematic complexity of the absorption as a few statistical descriptors.

We first introduce the equivalent velocity width, $V_{t,k}^{(0)}$ for subfeature k of transition t , which can be computed by multiplying the equivalent width by the factor c/λ_0 , or by computing the sum

$$V_{t,k}^{(0)} = \sum_{i=i_{t,k}^-}^{i_{t,k}^+} D_i \Delta v_i, \quad (10.56)$$

The mean velocity of subfeature is the first velocity moment of the flux decrement,

$$\langle v \rangle_{t,k} = \frac{V_{t,k}^{(1)}}{V_{t,k}^{(0)}} = \frac{1}{V_{t,k}^{(0)}} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} v_i D_i \Delta v_i. \quad (10.57)$$

The kinematic spread, defined as the flux decrement weighted velocity dispersion, is the square root of the ratio of the second moment to the zeroth moment,

$$\omega_{t,k} = \left[\frac{V_{t,k}^{(2)}}{V_{t,k}^{(0)}} \right]^{1/2} = \left[\frac{1}{V_{t,k}^{(0)}} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} (v_i - \langle v \rangle_{t,k})^2 D_i \Delta v_i \right]^{1/2}. \quad (10.58)$$

The kinematic asymmetry, defined as the flux decrement weighted velocity skew, is the cube root of the ratio of the third to the zeroth moment,

$$A_{t,k} = \left[\frac{V_{t,k}^{(3)}}{V_{t,k}^{(0)}} \right]^{1/3} = \left[\frac{1}{V_{t,k}^{(0)}} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} (v_i - \langle v \rangle_{t,k})^3 D_i \Delta v_i \right]^{1/3}. \quad (10.59)$$

The uncertainties in the velocity moments can be expressed compactly by noting the general form of the successive moments and their uncertainties

$$U^{(n)} = \left[\frac{V^{(n)}}{V^{(0)}} \right]^{1/n}, \quad \sigma_{U^{(n)}}^2 = \sum_i \left[\frac{\partial U^{(n)}}{\partial D_i} \right]^2 \sigma_{D_i}^2, \quad (10.60)$$

where we adopt $\langle v \rangle_{t,k} = U_{t,k}^{(1)}$, $\omega_{t,k} = U_{t,k}^{(2)}$, and $A_{t,k} = U_{t,k}^{(3)}$, and the sum is over all pixels included in the calculation. Carrying out the operations of Eq. 10.60, we derive

$$\sigma_{U_{t,k}^{(n)}}^2 = \frac{1}{n^2} \frac{[V_{t,k}^{(n)}]^{(2/n)-2}}{[V_{t,k}^{(0)}]^{(2/n)+2}} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} \left[V_{t,k}^{(0)} \frac{\partial V_{t,k}^{(n)}}{\partial D_i} - V_{t,k}^{(n)} \frac{\partial V_{t,k}^{(0)}}{\partial D_i} \right]^2 \sigma_{D_i}^2, \quad (10.61)$$

where

$$\begin{aligned} \frac{\partial V_{t,k}^{(0)}}{\partial D_i} &= \Delta v_i & \frac{\partial V_{t,k}^{(2)}}{\partial D_i} &= (v_i - \langle v \rangle_{t,k})^2 \Delta v_i \\ \frac{\partial V_{t,k}^{(1)}}{\partial D_i} &= v_i \Delta v_i & \frac{\partial V_{t,k}^{(3)}}{\partial D_i} &= (v_i - \langle v \rangle_{t,k})^3 \Delta v_i. \end{aligned} \quad (10.62)$$

Similar to the equivalent widths, summing over all subsystems gives the total system velocity moments of the transition,

$$V_t^{(0)} = \sum_{k=1}^{N_k} V_{t,k}^{(0)}, \quad V_t^{(1)} = \sum_{k=1}^{N_k} V_{t,k}^{(1)}, \quad \langle v \rangle_t = \frac{V_t^{(1)}}{V_t^{(0)}}. \quad (10.63)$$

Table 10.3 *Velocity Moments: PG 1206+459* $z_{\text{sys}} = 0.9277$

Mg II 2796 Subfeatures				
k	$W_k, \text{\AA}$	$\langle v \rangle, \text{km s}^{-1}$	$\omega_k, \text{km s}^{-1}$	$A_k, \text{km s}^{-1}$
1	0.031 ± 0.001	-404.47 ± 0.47	6.09 ± 1.80	-1.79 ± 0.42
2	0.078 ± 0.001	-339.21 ± 0.25	11.88 ± 0.41	0.60 ± 0.17
3	0.066 ± 0.001	-255.80 ± 0.19	4.99 ± 0.65	0.45 ± 0.20
4	0.052 ± 0.001	-183.21 ± 0.30	7.67 ± 0.57	1.00 ± 0.18
5	0.648 ± 0.002	12.65 ± 0.13	34.82 ± 0.12	0.33 ± 0.01

Transitions: Full System				
Tran (t)	$W_t, \text{\AA}$	$\langle v \rangle_t, \text{km s}^{-1}$	$\omega_t, \text{km s}^{-1}$	$A_t, \text{km s}^{-1}$
Fe II 2344	0.040 ± 0.002	-2.55 ± 2.06	35.14 ± 1.78	6.85 ± 0.45
Fe II 2374	0.011 ± 0.001	-16.59 ± 2.55	16.70 ± 2.11	10.65 ± 2.20
Fe II 2383	0.112 ± 0.003	-11.27 ± 2.05	53.63 ± 2.88	-4.21 ± 0.35
Fe II 2587	0.030 ± 0.002	0.09 ± 3.26	37.94 ± 2.81	10.82 ± 0.99
Fe II 2600	0.089 ± 0.002	-2.98 ± 0.82	35.16 ± 0.73	3.54 ± 0.08
Mg I 2853	0.061 ± 0.003	-2.57 ± 0.98	39.43 ± 1.27	2.54 ± 0.16
Mg II 2803	0.643 ± 0.003	-52.96 ± 0.87	129.60 ± 0.98	-2.13 ± 0.01
Mg II 2796	0.875 ± 0.003	-65.23 ± 0.69	140.29 ± 0.71	-1.58 ± 0.01

The kinematic spread and uncertainty for the transition are then

$$\omega_t = \left[\frac{V_t^{(2)}}{V_t^{(0)}} \right]^{1/2} = \left[\frac{1}{V_t^{(0)}} \sum_{k=1}^{N_k} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} (v_i - \langle v \rangle_t)^2 D_i \Delta v_i \right]^{1/2}, \quad (10.64)$$

$$A_t = \left[\frac{V_t^{(3)}}{V_t^{(0)}} \right]^{1/3} = \left[\frac{1}{V_t^{(0)}} \sum_{k=1}^{N_k} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} (v_i - \langle v \rangle_t)^3 D_i \Delta v_i \right]^{1/3}, \quad (10.65)$$

respectively.

Adopting $\langle v \rangle_t = U_t^{(1)}$, $\omega_t = U_t^{(2)}$, and $A_t = U_t^{(3)}$, the expression for the uncertainties are derived from Eq. 10.60,

$$\sigma_{U_t^{(n)}}^2 = \frac{1}{n^2} \frac{[V_t^{(n)}]^{(2/n)-2}}{[V_t^{(0)}]^{(2/n)+2}} \sum_{k=1}^{N_k} \sum_{i=i_{t,k}^-}^{i_{t,k}^+} \left[V_t^{(0)} \frac{\partial V_t^{(n)}}{\partial D_i} - V_t^{(n)} \frac{\partial V_t^{(0)}}{\partial D_i} \right]^2 \sigma_{D_i}^2, \quad (10.66)$$

where

$$\begin{aligned} \frac{\partial V_t^{(0)}}{\partial D_i} &= \Delta v_i & \frac{\partial V_t^{(2)}}{\partial D_i} &= (v_i - \langle v \rangle_t)^2 \Delta v_i \\ \frac{\partial V_t^{(1)}}{\partial D_i} &= v_i \Delta v_i & \frac{\partial V_t^{(3)}}{\partial D_i} &= (v_i - \langle V \rangle_t)^3 \Delta v_i. \end{aligned} \quad (10.67)$$

Since all quantities are computed in the rest-frame velocity of the absorption system, the measured velocity moments, and therefore the computed mean velocities, velocity widths, and velocity asymmetries are rest-frame quantities. In Table 10.3, we present the velocity moments of the $z_{\text{sys}} = 0.9277$ absorption systems in the spectrum of PG 1206+459. All velocity quantities were computed using Eqs. 10.56 through 10.67. The corresponding absorption profiles are shown in Figure 10.9.

10.5.3 Fractional Velocity Width

In order to quantify the bulk of the absorption, some researchers prefer to not adopt the velocity full width of the absorption profile, as this measurement can be sensitive to the presence of weaker subfeatures at large velocities. Generalizing the concept to quantify the kinematics of the bulk of the absorption profile, we define the fractional velocity $V_{f_v} = f_v V^{(0)}$. Typical fractions quoted in the literature are $f_v = 0.5$ and $f_v = 0.9$, from which the fractional velocity widths ΔV_{50} , and ΔV_{90} , respectively, can be determined. Computing ΔV_{f_v} requires one to isolate the velocity region, v_a to v_b , contributing to V_{f_v} , from which $\Delta V_{f_v} = v_b - v_a$ is computed.

From the definition of the fractional velocity,

$$V_{f_v} = \int_{v_a}^{v_b} D(v) dv = f_v V^{(0)}. \quad (10.68)$$

For a given transition t , in terms of pixelized data,

$$\sum_{i=i_t^-}^{i_{t,a}} D_i \Delta v_i = (1 - f_v) V^{(0)}, \quad \sum_{i=i_t^-}^{i_{t,b}} D_i \Delta v_i = f_v V^{(0)}, \quad (10.69)$$

where $V^{(0)}$ is computed using Eq. 10.63. We clearly have $i_t^- < i_{t,a} < i_{t,b} < i_t^+$, where i_t^- and i_t^+ are the starting and ending pixels of the absorption profile, and where $i_{t,a}$ and $i_{t,b}$ are the pixels corresponding to velocities v_a and v_b , respectively.

For the MgII $\lambda 2796$ profile of the $z_{\text{sys}} = 0.9277$ absorption systems in the spectrum of PG 1206+459, we find $V^{(0)} = 94.6 \text{ km s}^{-1}$. Application of

Eq. 10.69 for $f_v = 0.9$ yields $v_a = -329.1 \text{ km s}^{-1}$ and $v_b = 63.5 \text{ km s}^{-1}$, giving $\Delta V_{90} = 392.6 \text{ km s}^{-1}$.

10.5.4 Two-Point Velocity Correlation Function

An examination of the maximum velocity spread of an absorption system, shows that equivalent width is a poor indicator of absorber kinematics, and though the flux decrement weighted velocity spread, ω_v , correlates with equivalent width, the scatter is significant enough that one cannot predict one quantity from the other quantity (e.g. Churchill et al., 2000; Nielsen et al., 2016).

An alternative measure of absorber gas kinematics is velocity clustering. The simplest clustering statistic is a two-point velocity difference probability distribution function, called the two-point velocity correlation function (TPCF). This function was originally applied to the cosmological line-of-sight two-point clustering of populations of various ion-select absorbing systems to discern if they were consistent with cosmological distributed objects (e.g., Sargent et al., 1980; Lanzetta, Turnshek, & Wolfe, 1987; Steidel & Sargent, 1992). With the advent of higher-resolution spectra and the development of Voigt profile decomposition methods (which we discuss in § 10.7.2), the kinematics of absorption profiles were resolved and could be modeled as multiple “clouds” (Voigt profile components). The two-velocity clustering function of the Voigt profile components was pioneered by Petitjean & Bergeron (1990, 1994).

However, pixel velocities, rather than discrete “cloud” velocities, are a dramatically improved representation of the internal kinematic dispersion of absorbers (Nielsen et al., 2016). Two-point velocity clustering of “absorbing pixels” provides significantly more data points for improved statistics, generates a virtually continuous distribution of velocity separations compared to the highly discrete velocities of Voigt profile components, bypasses the time-intensive semi-subjective analysis involved in Voigt profile decomposition, and yields a model-independent quantification of the gas kinematics.

The pixel-based TPCF of a given transition is created in two steps. The first step is to compute the velocity differences of every unique pixel pair within a given transition’s absorption profile. For a feature comprising N pixels, there will be $N(N-1)/2$ unique pixel pairings. For example, if $N = 5$, there are 10 pairs, for which the velocity differences $\Delta v_{i,j} = |v_j - v_i|$, are $\Delta v_{1,2}$, $\Delta v_{1,3}$, $\Delta v_{1,4}$, $\Delta v_{1,5}$, $\Delta v_{2,3}$, $\Delta v_{2,4}$, $\Delta v_{2,5}$, $\Delta v_{3,4}$, $\Delta v_{3,5}$, $\Delta v_{4,5}$. Having computed the list of pairs, the second step is to bin the pairs. Set up a velocity grid with bin size B_v . The center of the m^{th} bin is then $\Delta v_m =$

$(m - 1/2)B_v$ and the bin spans $\Delta v_m^- = (m - 1)B_v$ to $\Delta v_m^+ = mB_v$. The TPCF is created by counting the number of i, j pixel pairs with $\Delta v_{i,j}$ that reside in each bin m , i.e., that satisfy the condition

$$\Delta v_m^- < \Delta v_{i,j} \leq \Delta v_m^+. \quad (10.70)$$

Let $n(\Delta v_m)$ represent the number of pixel pairs with velocities differences falling in bin m and assume there are a total of M bins. The TPCF is the probability distribution function given by

$$P(\Delta v_m) \Delta v_m = \frac{n(\Delta v_m) \Delta v_m}{B_v \sum_{m=1}^M n(\Delta v_m)}, \quad (10.71)$$

where $\int_0^\infty P(\Delta v) d\Delta v = 1$.

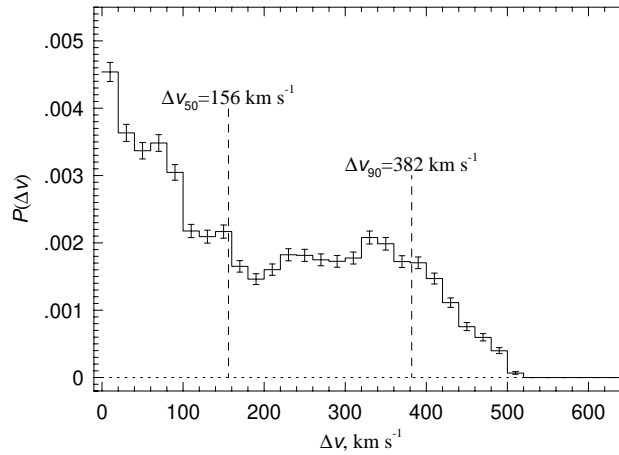


Figure 10.10 Pixel two-point correlation function for the MgII $\lambda 2796$ profile of the rich absorption line system at $z_{\text{sys}} = 0.9277$ in the spectrum of PG 1206+459. The binning is $B_v = 20 \text{ km s}^{-1}$. The uncertainties, $\sigma_{P(\Delta v)}$, assume the frequency of velocity splittings obeys a Poisson distribution. The velocity splittings are most frequent at $0 \leq \Delta v \leq 20 \text{ km s}^{-1}$, reflecting the kinematics of narrow unresolved subfeatures. However, there is power at high-velocity splittings for this kinematically complex absorbing system (see Figure 10.9), with the secondary peak being $300 \leq \Delta v \leq 320 \text{ km s}^{-1}$.

In Figure 10.10, we show the TPCF for the MgII $\lambda 2796$ profile of the rich absorption line system at $z_{\text{sys}} = 0.9277$ in the spectrum of PG 1206+459. The binning is $B_v = 20 \text{ km s}^{-1}$ and the asymmetric uncertainties are Poisson. In order to have the TPCF reflect the distribution of the velocity splittings for the absorbing gas, only pixels residing within the detected subfea-

tures are included (continuum pixels between subfeatures are omitted). The included pixels are shown in Figure 10.9 in the gray shaded regions. For this TPCF, there were 151 pixels, resulting in 11325 pixel-velocity pairs.

A further quantification of the absorption kinematics is the velocity splitting containing the fractional area f_v under the TPCF, where typical values of f_v are 0.5 and 0.9, representing the 50% and 90% percentiles of the velocity splittings. A smaller value of f_v provides a measure of the “core” of the distribution function, whereas a larger value provides insight into the “wings”. We define this velocity as Δv_{f_v} , according to the relationship

$$f_v = \int_0^{\Delta v_{f_v}} P(\Delta v) d\Delta v. \quad (10.72)$$

For the TPCF in Figure 10.10, we obtain $\Delta v_{50} = 156.1^{+14.8}_{-12.7} \text{ km s}^{-1}$ and $\Delta v_{90} = 382.9^{+36.5}_{-27.9} \text{ km s}^{-1}$. The downward and upward uncertainties are obtained evaluating Eq. 10.72 with $P(\Delta v)$ replaced by $P(\Delta v) - \sigma_{P(\Delta v)}$ and by $P(\Delta v) + \sigma_{P(\Delta v)}$, respectively.

10.6 The Apparent Optical Depth Method

In § 9.2.7, we introduced the apparent optical depth method for measuring the column densities of absorption line profiles. Here we describe the method in additional detail with focus on the practical analysis steps required to obtain the integrated apparent column densities for various ionic species. We remind the reader that the relationship between true optical depth and apparent optical depth was provided in Eq. 9.45.

10.6.1 Apparent Optical Depth Profiles

The run of optical the depth across the absorption profile from any transition t is computed in each pixel i by inverting the radiative transfer of pure absorption,

$$\tilde{\tau}_i = -\ln(R_i), \quad R_i = \frac{I_i}{\bar{I}_i^c}, \quad (10.73)$$

where \bar{I}_i^c is the fitted continuum counts, the range $0 < R_i \leq 1$ applies within uncertainties, and the notation $\tilde{\tau}_i$ emphasizes that this quantity is the apparent optical depth in pixel i calibrated to have velocity v_i . Applying error propagation to R_i yields an expression for σ_{R_i} identical to the uncertainty

in D_i ,

$$\sigma_{R_i}^2 = \sigma_{D_i}^2 = R_i^2 \left[\left(\frac{\sigma_{I_i}}{I_i} \right)^2 + \left(\frac{\sigma_{\bar{I}_i^c}}{\bar{I}_i^c} \right)^2 \right] \quad (10.74)$$

However, due to the nature of logarithms, the symmetric uncertainties in R_i do not translate to normally distributed uncertainties in optical depth (i.e., they are not symmetric). If $y = \ln x$, in cases where $x \gg \sigma_x$ we can adopt $\sigma_y = d \ln x = \sigma_x/x$, as this is in the regime of normally distributed errors. However, in the cores of absorption lines, Eq. 10.73 is evaluated in a regime where $R_i \simeq \sigma_{R_i}$. We thus, approximate the uncertainties in the apparent optical depth by assuming that the probability distribution of R_i maps to the probability distribution in $\tilde{\tau}_i$. Generally expressed, if we have the functional dependence $y = \ln x$, then we map the uncertainties using $y \pm \Delta y = \ln(x \pm \Delta x)$. Applied to Eq. 10.73, we have

$$\sigma_{\tilde{\tau}_i}^+ = \ln(R_i + \sigma_{R_i}) - \ln(R_i), \quad \sigma_{\tilde{\tau}_i}^- = \ln(R_i) - \ln(R_i - \sigma_{R_i}). \quad (10.75)$$

With high resolution spectra, resolved saturated absorption profiles can have counts in their cores consistent with $R_i = 0$. In these line cores, some pixels will, due to noise, have pixels with $R_i \leq 0$; in these cases the apparent optical depth (and its uncertainty) are ill-defined via Eq. 10.73. In such pixels, we obtain a lower limit on the apparent optical depth. This lower limit can be no better than the noise level permits in the each pixel. Thus, when the counts in a pixel meet the condition $R_i < \sigma_{R_i}$, the optical depth is treated as a lower limit and Eq. 10.73 is replaced by,

$$\tilde{\tau}_i > -\ln(\sigma_{R_i}). \quad (10.76)$$

In pixels where the counts are consistent with the continuum, i.e., $R_i > 1 - \sigma_{R_i}$ the optical depth is treated as an upper limit given by

$$\tilde{\tau}_i < -\ln(1 - \sigma_{R_i}). \quad (10.77)$$

In Figure 10.11(a), the apparent optical depth profiles are illustrated for subfeature $k = 5$ for the MgII selected absorber shown in Figure 10.9(a). Two MgII transitions, one MgI transition, and five FeII transitions are shown. Outside the detection region (at $v_i \leq -80 \text{ km s}^{-1}$ and $v_i \geq +95 \text{ km s}^{-1}$), the optical depths are nulled. Upward arrows denote lower limits (Eq. 10.76) and downward arrows denote upper limits (Eq. 10.77) on $\tilde{\tau}_i$. Only two pixels met the criterion of a lower limit ($v_i \simeq -30 \text{ km s}^{-1}$) in the MgII $\lambda 2796$ profile. As can be seen by inspection of the error bars, the uncertainties are relatively symmetric for smaller $\tilde{\tau}_i$ and tend toward greater asymmetry as $\tilde{\tau}_i$ increases.

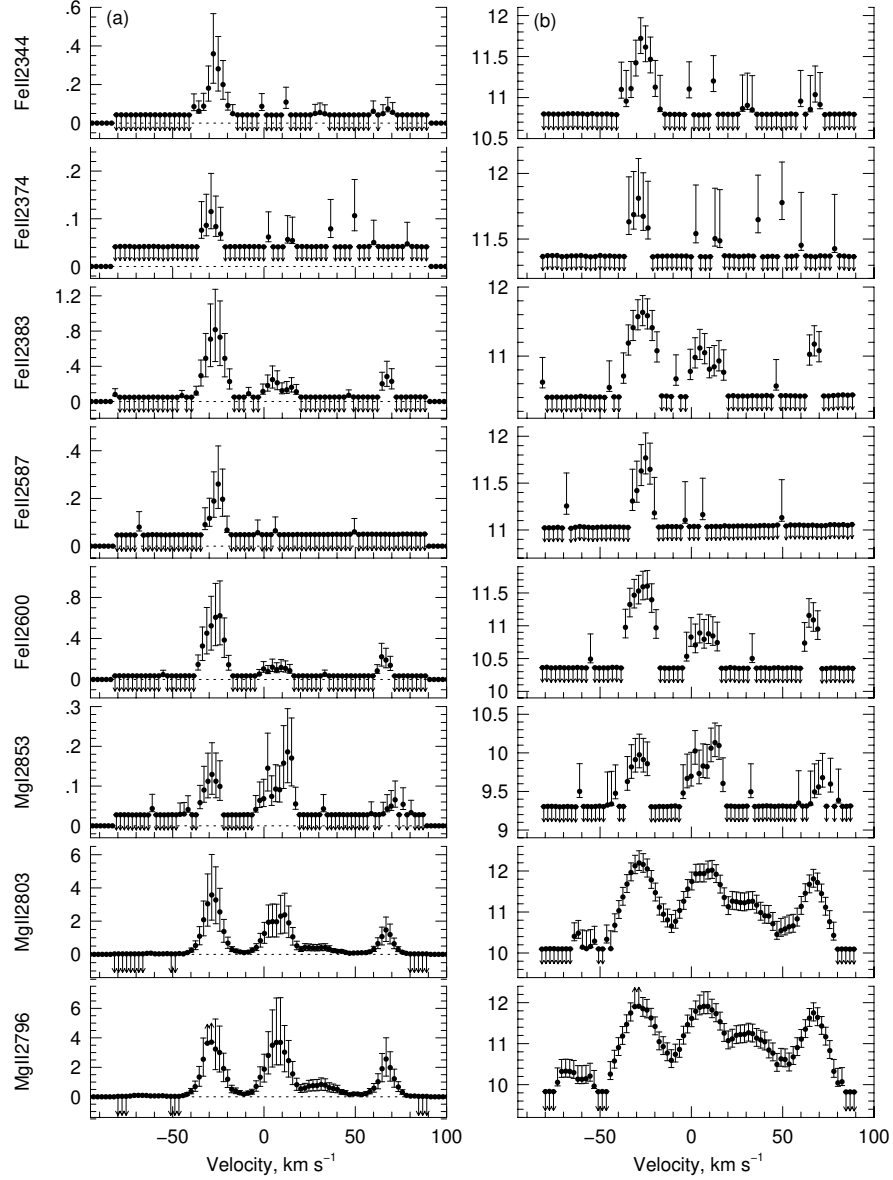


Figure 10.11 (a) The $\tilde{\tau}_i$ profiles for subfeature $k = 5$ for the absorption profiles associated with the MgII selected absorber shown in Figure 10.9(a). Downward arrows denote upper limits and upward arrows denote lower limits. The apparent optical depth are set to $\tilde{\tau}_i = 0$ outside the detection region. (b) The $\log \tilde{N}(v_i)$ profiles in units atoms cm⁻² (km s⁻¹)⁻¹ for the same transitions. As with the apparent optical depth spectra, downward arrows denote upper limits and upward arrows denote lower limits.

10.6.2 Apparent Optical Depth Column Densities

The column density per unit wavelength, $N(\lambda)$, in units $\text{atoms cm}^{-2} \text{ \AA}^{-1}$ is obtained from 9.42. Conversion to $N(v)$, in units $\text{atoms cm}^{-2} (\text{km s}^{-1})^{-1}$ is obtained via “flux” conservation per unit interval, $N(\lambda)d\lambda = N(v)dv$. We obtain $N(\lambda) = (c/\lambda)N(v)$ from $dv/c = d\lambda/\lambda$. Substituting into 9.42, we have

$$\tau(v) = \frac{\pi e^2}{m_e c} f \lambda_0 N(v), \quad (10.78)$$

which, when inverted and written in terms of the pixelization in which pixel i is calibrated to velocity v_i , we have

$$\tilde{N}(v_i) = \frac{m_e c}{\pi e^2 f \lambda_0} \tilde{\tau}_i. \quad (10.79)$$

Error propagation through Eq. 10.73, yields

$$\begin{aligned} \sigma_{\tilde{N}_i}^+ &= \frac{m_e c}{\pi e^2 f \lambda_0} [\ln(R_i + \sigma_{R_i}) - \ln(R_i)] \\ \sigma_{\tilde{N}_i}^- &= \frac{m_e c}{\pi e^2 f \lambda_0} [\ln(R_i) - \ln(R_i - \sigma_{R_i})] . \end{aligned} \quad (10.80)$$

When the apparent optical depth is an upper (lower) limit, then in $\tilde{N}(v_i)$ is an upper (lower) limit.

In Figure 10.11(b), profiles of the apparent column densities per unit velocity interval are illustrated for the apparent optical depth profiles shown in panel (a). Upward arrows denote lower limits and downward arrows denote upper limits. The value presented are $\log \tilde{N}(v_i)$. The uncertainties in the $\log \tilde{N}(v_i)$ are obtained via the relationship

$$\sigma_{\log \tilde{N}_i}^\pm = 0.4343 \frac{\sigma_{\tilde{N}_i}^\pm}{\tilde{N}(v_i)} \quad (10.81)$$

which follows from $d \log x = d \ln(x) / \ln(10) = 0.4343(dx/x)$

Each transition of a given species has a unique profile representing the column densities per unit velocity for the species. In Figure 10.11(b), the MgII $\lambda\lambda 2796, 2803$ doublet provides two representations of $\tilde{N}_{\text{MgII}}(v_i)$, and the five transitions of FeII provide multiple representations of $\tilde{N}_{\text{FeII}}(v_i)$. In order to obtain the optimal apparent column density profile for the species, an optimized mean of the measurement at each velocity pixel is computed. This calculation is complicated by several considerations.

First, following wavelength calibration, the velocity pixelization of each absorption profile is different in that the velocity centers and widths of the pixels are not exactly aligned. Thus, prior to computing the apparent optical

depth profiles, one should rebin the pixelization of the counts (or flux) in each profile in the absorption line system according to the method described in § 10.2, except that the profiles are not combined together; each profile is individually rebinned to the same fixed pixel binning using flux conservation. Once all profiles are on the same velocity binning, we then compute the apparent optical depth profiles and the apparent column density profiles. This allows the optimal mean to be computed in each velocity bin (pixel).

Second, at a given velocity, there are three possible scenarios that must be considered: the measured apparent column density at velocity v_i in each transition is (1) a clean measurements and all transitions of the species can be included in the mean; (2) such that some transitions have clean measurements and others have either an upper or lower limit; or (3) is an upper or lower limit so that only a limit can be obtained. We will describe the calculation for each scenario below.

Third, the uncertainties in the measure apparent optical depths and column densities are asymmetric. We thus must account for these asymmetric uncertainties when computing the mean. There is no one-size-fits-all treatment of asymmetric errors, as it requires modeling the probability distribution of the errors in the measurements. Here, we illustrate the second of the two simple models developed by Barlow (2003). We also point the reader to a somewhat more sophisticated model developed by D'Agostini & Raso (2000).

We first consider scenario (1) in which species s has a measured apparent column density profiles for multiple transitions, indexed by t , and all $N_{s,t}$ transitions have measured values of $\tilde{N}_t(v_i)$. We assume that the profiles are binned to a common velocity pixelization. Following Barlow (2003), we begin by defining the mean and the difference in the asymmetric uncertainties,

$$\sigma_t(v_i) = \frac{\sigma_{\tilde{N}_t(v_i)}^+ + \sigma_{\tilde{N}_t(v_i)}^-}{2}, \quad \alpha_t(v_i) = \frac{\sigma_{\tilde{N}_t(v_i)}^+ - \sigma_{\tilde{N}_t(v_i)}^-}{2}, \quad (10.82)$$

for which the model variance in each $\tilde{N}_t(v_i)$ is

$$V_t(v_i) = \sigma_t^2(v_i) + 2\alpha_t^2(v_i). \quad (10.83)$$

The optimal mean is then computed from

$$\tilde{N}_s(v_i) = \frac{\sum_{t=1}^{N_{s,t}} w_t(v_i) [\tilde{N}_t(v_i) - \alpha_t(v_i)]}{\sum_{t=1}^{N_{s,t}} w_t(v_i)}, \quad (10.84)$$

where the quantity $\tilde{N}_t(v_i) - \alpha_t(v_i)$ models a bias in the measure value for transition t at velocity v_i according to the asymmetric probability distribution of the uncertainty, and $w_t(v_i)$ is the weight for transition t at velocity v_i . The uncertainty in $\tilde{N}_s(v_i)$, the mean apparent column density for the species at velocity v_i , is

$$\sigma_{\tilde{N}_s(v_i)}^2 = \frac{\sum_{t=1}^{N_{s,t}} w_t^2(v_i) V_t(v_i)}{\left(\sum_{t=1}^{N_{s,t}} w_t(v_i) \right)^2} . \quad (10.85)$$

Barlow (2003) derives $w_t(v_i) = 1/V_t(v_i)$ as the optimal weighting.

For scenario (2), in which some of the transitions have an upper or lower limit at velocity v_i , the simplest approach is to null the weight, i.e., $w_t(v_i) = 0$, thus removing the particular velocity pixel from the optimal mean for the species in Eqs. 10.84 and 10.85. If there is only one transition with a measured apparent column density at v_i , then $\tilde{N}_s(v_i) = \tilde{N}_t(v_i) - \alpha_t(v_i)$ and $\sigma_{\tilde{N}_s(v_i)}^2 = V_t(v_i)$.

For scenario (3), in which only limits are measured at velocity v_i , we adopt the most stringent limit from the various transitions as the limit for the species. In the case of lower limits, the most stringent limit will come from the transition with the smallest value of $f\lambda_0$, as this provides the highest limit on $\tilde{N}_t(v_i)$. In the case of upper limits, the opposite is true in that the transition with the largest value of $f\lambda_0$ provides the lowest limit on $\tilde{N}_t(v_i)$.

As can be seen, the logic for the various scenarios must be applied on a velocity-pixel by velocity-pixel basis. And in some cases, one has to make higher-order decisions for compromised pixels, such as fixed pattern noise, cosmic rays, bad sky subtraction, etc., or “unresolved saturation”. The condition of unresolved saturation can occur at a given velocity within a complex profile in a region where the line broadening is much narrower than the instrument spread function.

Unresolved saturation can be identified in the cases of fine-structure doublets or multiplets. Consider a fine-structure doublet in which, at a given pixel velocity, both members of the doublet have robustly measured values of $\tilde{N}_t(v_i)$. Within uncertainties, both members of the doublet should provide the same value of $\tilde{N}_t(v_i)$. However, in the case of unresolved saturation, since

$\tilde{N}_t(v_i) \propto (f\lambda_0)^{-1}$, the ratio of the apparent column densities will be

$$\frac{\tilde{N}_t^{(b)}(v_i)}{\tilde{N}_t^{(a)}(v_i)} = \frac{f^{(a)}\lambda_0^{(a)}}{f^{(b)}\lambda_0^{(b)}} \simeq 2, \quad (10.86)$$

where we assume, typical of fine-structure doublet in lithium-like and sodium-like ions, transition (a) has oscillator strength twice that of transition (b). When this condition holds true, one should adopt the measured value in transition (b) as a lower limit on the apparent column density at velocity v_i .

The final result of Eqs. 10.84 and 10.85 is an apparent column density profile for each species. From this profile the integrated apparent column density can be computed,

$$\tilde{N}_s = \sum_{i=1}^{N_{\text{pix}}} \tilde{N}_s(v_i) \Delta v_i, \quad \sigma_{\tilde{N}_s}^2 = \sum_{i=1}^{N_{\text{pix}}} \sigma_{\tilde{N}_s(v_i)}^2, \quad (10.87)$$

where i is the pixel number corresponding to velocity v_i ranging from 1 to N_{pix} for the rebinned velocity pixelization of the apparent optical depth profiles with pixel widths Δv_i . If desired, the integrated apparent column densities of the individual subfeatures for each transition can be computed using the book keeping methods presented for the computation of the equivalent widths (§ 10.4.2) and velocity moments (§ 10.5.2).

Since it is possible that some velocity pixels will yield $\tilde{N}_s(v_i)$ that are lower limits, a criterion must be established as when to consider the integrated apparent column density a measurement or a lower limit following execution of Eq. 10.87. One option is to adopt the rule that if the saturation is present over a full resolution element (i.e., p contiguous pixels, where $p = \Delta\lambda_s/\Delta\lambda$ is the number of pixels per resolution element), then \tilde{N}_s is a lower limit.

10.7 Deblending Absorption Features

10.7.1 Gaussian Decomposition

In the case of lower-resolution spectra, in which individual absorption lines are unresolved or only partially resolved, an alternate method for measuring equivalent widths is to model the absorption with a fitted function. This function should capture the essence of the instrumental spread function (ISF). For absorption features that are resolved into complex absorption profiles, the assumption is that each apparent component comprising the profile is itself an unresolved or partially resolved absorption line. Thus, a

multi-component functional fit can be used to obtain the number of absorption lines, their relative velocities, and the equivalent widths for each.

The ISF of most spectrographs is well approximated by a Gaussian function, so unless the wings of the ISF depart significantly from a normal distribution, a Gaussian fitting function is commonly adopted. The Gaussian function is written

$$G(\lambda_i; \bar{\lambda}, A, \Sigma) = A \exp \left\{ -\frac{1}{2} \left(\frac{\lambda_i - \bar{\lambda}}{\Sigma} \right)^2 \right\}, \quad (10.88)$$

where the three free parameters describing the modeled line are the central wavelength $\bar{\lambda}$, the Gaussian width, Σ , and the amplitude A . The parameters are determined from least-squares minimization using the χ^2 statistic,

$$\chi^2 = \sum_{i=j^-}^{j^+} \left[\frac{D_i - G(\lambda_i; \bar{\lambda}, A, \Sigma)}{\sigma_{D_i}} \right]^2, \quad (10.89)$$

where σ_{D_i} is the uncertainty in the flux decrement, D_i . A proper fitting routine yields the parameter uncertainties $\sigma_{\bar{\lambda}}$, σ_A , and σ_{Σ} (for example, see Press et al., 2007). The fitted pixels should include continuum counts on either side of the absorption line. Per the objective feature finding methods discussed in § 10.3, if i^- and i^+ are the starting and ending pixels of the absorption line, then buffering two to three resolution elements should suffice, so that $j^- = i^- - 3p$ and $j^+ = i^+ + 3p$, where p is the number of pixels per resolution element (Eq. 8.54).

Physically, an absorption line cannot have a width narrower than the ISF width. However, noise in the data can give rise to a realization in which, following χ^2 minimization, we obtain $\Sigma \leq \sigma^{\text{ISF}}$. In these cases, the line should be refit with Σ fixed (no longer a fitted parameter) at the ISF width, $\Sigma = \bar{\lambda}/(2\sqrt{2\ln 2}R)$. An additional reason for this is that the amplitude A and the width Σ are anti-correlated, so an unphysical narrow width will result in an overestimate of the amplitude.

The observed equivalent width is then obtained from the integral

$$W = \int_{-\infty}^{\infty} G(\lambda; \bar{\lambda}, A, \Sigma) d\lambda = \sqrt{2\pi} A \Sigma, \quad (10.90)$$

where the constant $\sqrt{2\pi}$ arises from the standard normalization of a unit-area Gaussian function. The uncertainty in W is

$$\sigma_W^2 = 2\pi \left[(\Sigma\sigma_A)^2 + (A\sigma_{\Sigma})^2 - (A\sigma_A)(\Sigma\sigma_{\Sigma}) \right], \quad (10.91)$$

where the last term in Eq. 10.91 accounts for the anti-correlated errors in W

(for a fixed W the product $A\Sigma$ is a constant). Note that σ_W is independent of the line center. For an unresolved line, in which Σ was held constant for the minimization, we apply $\sigma_\Sigma = 0$. If the atomic transition of the absorption line has been identified at redshift z , then the rest-frame equivalent width is $W_r = W/(1+z)$.

In complex multi-component absorption profiles, the equivalent widths of individual absorption line components can be measured by decomposing the absorption profile with multiple Gaussian functions. Again, this approach is appropriate when the resolution is low enough that the line broadening is narrower or on the order of the ISF width.

The profile flux decrement is modeled as the sum of Gaussian components

$$D_i = \sum_{k=1}^N A_k \exp \left[-\frac{1}{2} \left(\frac{\lambda_i - \bar{\lambda}_k}{\Sigma_k} \right)^2 \right], \quad (10.92)$$

where N is the number of components, and the A_k , $\bar{\lambda}_k$, and Σ_k are the component depths, central wavelengths, and widths, respectively. In practice, the fitting parameters are vectorized such that $\mathbf{a} = a_1, a_2, a_3, \dots, a_{3N}$, and Eq. 10.92 is written $D_i = \mathcal{G}(\lambda_i; \mathbf{a})$, where

$$\mathcal{G}(\lambda_i; \mathbf{a}) = \sum_{k=1}^N a_j \exp \left[-\frac{1}{2} \left(\frac{\lambda_i - a_{j+1}}{a_{j+2}} \right)^2 \right], \quad (10.93)$$

where $j = 2k + (k - 2)$. The fit is performed using least-squares χ^2 minimization applying Eq. 10.89 with $\mathcal{G}(\lambda_i; \mathbf{a})$ replacing $G(\lambda; \bar{\lambda}, A, \Sigma)$. One then unpacks the vectorized parameter list according to $A_k = a_j$, $\bar{\lambda}_k = a_{j+1}$, and $\Sigma_k = a_{j+2}$, with $\sigma_{A_k} = \sigma_{a_j}$, $\sigma_{\bar{\lambda}_k} = \sigma_{a_{j+1}}$, and $\sigma_{\Sigma_k} = \sigma_{a_{j+2}}$. Again, if after the initial fit, a given component has $\Sigma_k \leq \sigma^{\text{ISF}}$, the full profile should be refitted with the fixed value $\Sigma_k = \bar{\lambda}_k / (2\sqrt{2 \ln 2} R)$.

Consistent with the philosophical principle of Occam's razor, one should always fit the data with a minimum number of free parameters. It is best if an objective method, such as an F -test, can be applied to determine when adding an additional component provides no statistical improvement in χ^2 , which is the normalized variance of the data about the fit. An example of a minimum-component Gaussian decomposition is shown in Figure 10.12(a). The solid curve is the model spectrum and the dotted curves are the individual Gaussian components with central wavelengths indicated by the ticks above the continuum. The two components at 2994.7 and 2999.8 Å are unresolved; we have fixed their widths to that of the ISF with $R = 1300$. The observed equivalent widths and uncertainties for each component are obtained using Eqs. 10.90 and 10.91, respectively. The full profile observed

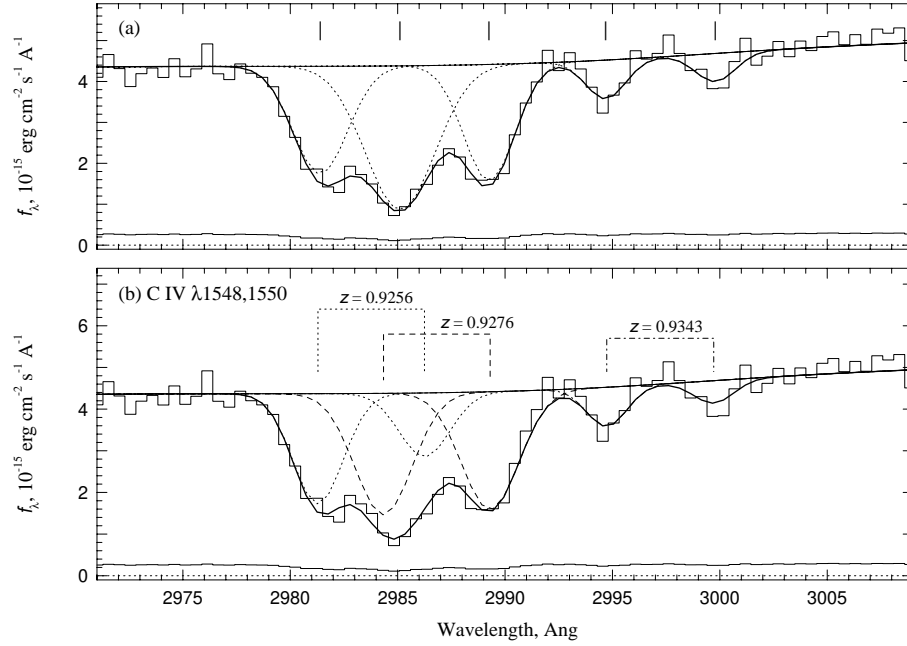


Figure 10.12 (a) Gaussian decomposition of a complex absorption feature in which the minimum number of components are employed. In this $R = 1300$ spectrum, the two components at 2994.7 and 2999.8 \AA are unresolved so that their widths are fixed at the width of the ISF. (b) Having identified the complex feature as multiple components of CIV $\lambda\lambda 1548, 1550$ absorption at $z \simeq 0.93$, the Gaussian decomposition is performed by coupling the line centers of each doublets absorption lines. The dotted, dashed, and dot-dashed fits each represent one doublet pair.

equivalent width its uncertainty is obtained from the components via the sums

$$W = \sum_{k=1}^N W_k, \quad \sigma_W^2 = \sum_{k=1}^N \sigma_{W_k}^2. \quad (10.94)$$

The absorption profile presented in Figure 10.12 comprises multiple blended CIV $\lambda\lambda 1548, 1550$ doublets at $z \sim 0.93$. The minimum-component fit presented in Figure 10.12(a) does not account for any of the atomic properties governing these doublets, and therefore the component equivalent widths cannot be utilized to estimate the CIV column densities. However, by coupling fitted components as members of CIV doublets, we introduce a physical basis to the profile decomposition. As atomic physics governs the wavelength separation of the doublet, the wavelengths are coupled. If, for example, we couple components $k = 1$ and $k = 3$ as the CIV $\lambda 1548.195$ and

$\lambda 1550.770$ transitions, respectively, then the central wavelengths couple as $\bar{\lambda}_3 = \bar{\lambda}_1 + \Delta\lambda(1+z)$, where $\Delta\lambda = \lambda^{(1550)} - \lambda^{(1548)}$, and $1+z = \bar{\lambda}_1/\lambda^{(1548)}$, which yields $\bar{\lambda}_3 = \bar{\lambda}_1(\lambda^{(1550)}/\lambda^{(1548)})$. Similarly, as the broadening mechanisms of each member of a doublet will be identical, the component widths are coupled, $\Sigma_3 = \Sigma_1$. Thus, for the $k=3$ component, the only free parameter in the fit is A_3 . Of course, for coding purposes and application of Eq. 10.93, the coupling of these parameters needs to be mapped to the vectorized fitting parameter \mathbf{a} and the least-squares solver needs to be informed which indices of \mathbf{a} are free to vary for the fit.

In Figure 10.12(b), we present a Gaussian decomposition of three CIV doublets, of which two are blended. In the blended profile, the dotted curves represent the first doublet (coupled components 1 and 3) and the dashed curves represent the second doublet (coupled components 2 and 4). Coupled components 5 and 6 are not blended, but they are unresolved and thus have widths fixed to the ISF width. Since the absorption transitions are identified as CIV doublets, we can obtain the redshift for each doublet from the central wavelengths, and upon compute the observed equivalent widths, we can obtain the rest-frame equivalent widths, $W_r = W/(1+z)$. Thus, the profile decomposition provides data from which we can apply the curve of growth and double ratio methods to estimate or constrain the CIV column densities (see §§ 9.2.5.1–9.2.5.2).

10.7.2 Voigt Profile Decomposition

For high-resolution spectra, i.e., $R \geq 20,000$, Voigt profile decomposition is an established technique for modeling complex and blended absorption profiles because it is well suited for parameterizing the absorbing gas properties into physically meaningful quantities, i.e., the number of components, their relative line-of-sight velocities, column densities, and Doppler b parameters. The most commonly used Voigt profile fitting program is VPFit¹ by R. F. Carswell and J. K. Webb.

Recall that for an absorption line recorded by an instrument having an ISF function $\Phi(\Delta\lambda; \Delta\lambda_s)$, the post-optics spectrum is given by

$$I_\lambda = \Phi(\Delta\lambda; \Delta\lambda_s) * \bar{I}_\lambda^c \exp\{-N\alpha(\lambda; \Delta\lambda_c)\}, \quad (10.95)$$

where I_λ is the recorded counts at observed wavelength λ , \bar{I}_λ^c is the counts in the fitted continuum, N is the column density of the absorbing atom or ion, and $\alpha(\lambda; \Delta\lambda_c)$ is the total absorption cross section, including the atomic

¹ VPFit is publicly available, free, modifiable software that can be accessed at www.ast.cam.ac.uk/~rfc/vpfit.html.

broadening and the Gaussian broadening component $\Delta\lambda_{\text{G}}$, and where $\Delta\lambda_{\text{s}}$ is the full-width half-maximum of the ISF.

Under the assumption the Gaussian component is thermal broadening, we have $\Delta\lambda_{\text{G}} = \Delta\lambda_{\text{D}} = (\lambda_0/c)b_{\text{T}}$, where the thermal Doppler b parameter is $b_{\text{T}} = \sqrt{2kT/m}$. As discussed in § 9.2.6, the Gaussian component to the line broadening can be modeled to include additional physics, especially a turbulent velocity component, $b_{\text{turb}} = \Delta v_{\text{turb}}$, giving $\Delta\lambda_{\text{turb}} = (\lambda_0/c)b_{\text{turb}}$, so that $\Delta\lambda_{\text{G}}^2 = \Delta\lambda_{\text{D}}^2 + \Delta\lambda_{\text{turb}}^2$. Physically, these components to the line broadening are due to the line-of-sight velocity distribution of the absorbing atomic species, whereas the magnitude of the wavelength redistribution, $\Delta\lambda_{\text{G}}$, is proportional to the transition wavelength, λ_0 . Thus, the physical parameter adopted to describe the Gaussian line broadening component is the total Doppler b parameter

$$b^2 = b_{\text{T}}^2 + b_{\text{turb}}^2 = \frac{2kT}{m} + (\Delta v_{\text{turb}})^2. \quad (10.96)$$

From Eq. 9.17, the optical depth from transition with rest-frame wavelength λ_0 due to an ion of column density N in an isothermal cloud is

$$\tau_{\lambda} = N\alpha(\lambda; b) = \frac{\sqrt{\pi}e^2}{m_e c^2} \frac{Nf\lambda_0^2}{\Delta\lambda_{\text{G}}} H(x_{\lambda}, y), \quad (10.97)$$

where $\Delta\lambda_{\text{G}} = (\lambda_0/c)b$, and where λ_0 and f are the transition wavelength and oscillator strength, and $H(x_{\lambda}, y)$ is the Hjerting function² given by Eq. 9.18.

As written, Eq. 10.97 applies an absorption line observed in the rest-frame. The modification of Eq. 10.97 for a redshifted line is embedded in the variable x_{λ} (Eq. 9.19) of the Hjerting function. For cosmologically redshifted absorption at redshift z , the line will be observed in the spectrum at wavelength $\lambda = \lambda_0(1+z)$. Since $H(x_{\lambda}, y)$ models the rest-frame properties of the absorption profile, x_{λ} is translated to the rest-frame. For an absorption line at redshift z that is associated with an absorption system at redshift z_{sys} , the line center in the rest-frame will be $\bar{\lambda} = \lambda_0(1+z)/(1+z_{\text{sys}})$. Thus, in the rest-frame,

$$x_{\lambda}(z, z_{\text{sys}}) = \frac{\lambda - \lambda_0(1+z)}{\Delta\lambda_{\text{G}}(1+z_{\text{sys}})}, \quad (10.98)$$

where $\lambda/(1+z_{\text{sys}})$ is the rest-frame wavelength corresponding to the recorded counts I_{λ} at observed wavelength λ . An alternative interpretation of Eq. 10.98 is that the Doppler width is redshifted to z_{sys} in order for x to be cosmologically co-moving quantity. The quantity y (Eq. 9.19), which describes the

² Computationally, $H(x_{\lambda}, y)$ is easily evaluated as the real part of the complex probability function as discussed in § 9.2.3.

Lorentzian atomic broadening, is unaltered for a redshifted line, i.e.,

$$y = \frac{\Gamma \lambda_0^2}{4\pi c} \frac{1}{\Delta \lambda_G}, \quad (10.99)$$

where Γ is the transition damping constant. We see that for a given Voigt profile component, there are three constants specific to the atomic transition, λ_0 , f , and Γ , and three free parameters, N , b , and z , that provide insight into the absorbing gas physical conditions. We now write the optical depth computed using Eq. 10.97 and applied in Eq. 10.95 as

$$\tau_\lambda = N\alpha(\lambda; b, z) = \tau_\lambda(N, b, z). \quad (10.100)$$

The redshift z and the relative velocity of the component in the rest-frame of the absorbing structure are related through the non-relativistic Doppler shift formula $\Delta v/c = \Delta \lambda/\lambda$, which yields

$$v = c \left[\frac{1+z}{1+z_{\text{sys}}} - 1 \right]. \quad (10.101)$$

For a complex absorption profile comprising several components, k , each having a unique column density, Doppler b parameter, and redshift, the optical depth of the full profile is the sum of the optical depths of each Voigt profile component,

$$\tau_\lambda = \sum_{k=1}^{N_{\text{VP}}} \tau_\lambda(N_k, b_k, z_k), \quad (10.102)$$

where each optical depth in the sum is evaluated using Eq. 10.97. Since there are three free parameters per component, the total number of free parameters becomes $3N_{\text{VP}}$, where N_{VP} is the number of components used to model the complex profile.

Applying the law of exponents, the Voigt profile decomposition of a complex profile modeled into N_{VP} components can be described by a single function of the form

$$\mathcal{V}(\lambda; \mathbf{a}) = \Phi(\Delta \lambda; \Delta \lambda_s) * \prod_{k=1}^{N_{\text{VP}}} \exp \{ -\tau_\lambda(N_k, b_k, z_k) \}, \quad (10.103)$$

where $\mathbf{a} = a_1, a_2, a_3, \dots, a_{3N_{\text{VP}}}$ comprises the vectorized free parameter list, which we discuss further below. As with Gaussian decomposition, Voigt profile decomposition usually involves a least-squares χ^2 minimization between

the model profile and the observed profile,

$$\chi^2 = \sum_{i=j^-}^{j^+} \left[\frac{(I_i/\bar{I}_i^c) - \mathcal{V}(\lambda_i; \mathbf{a})}{\sigma_{D_i}} \right]^2, \quad (10.104)$$

where σ_{D_i} is the uncertainty in the flux decrement, applied from $I_i/\bar{I}_i^c = 1 - D_i$ and where the starting and ending pixels j^- and j^+ are determined as discussed for the Gaussian decomposition (Eq. 10.89).

A proper fitting routine yields the parameter uncertainties for each component σ_{N_k} , σ_{b_k} , and σ_{z_k} . Compared to Gaussian decomposition, we see that iterating toward the minimum χ^2 for Voigt profile decomposition is computationally more intensive, as generating the model absorption profile from Eq. 10.103 requires convolution with the ISF. The convolution also renders computation of the uncertainties in the fitting parameters to be challenging, as the elements of the Jacobian cannot be derived symbolically; the derivatives of Eq. 10.103 must be obtained numerically.

The formalism of Eqs. 10.103 and 10.104 can be further extended, allowing one to leverage the absorption profiles of multiple transitions from the same atom or ion. This is because, for a given ion, the absorption lines from all its transitions will have the same relative line-of-sight velocity, the same column density, and the same line broadening. In terms of Voigt profile fitting parameters, for a given component k , this translates to all transitions arising from the same N_k , b_k , and z_k . The flux decrements and equivalent widths of the observed absorption profiles of the different transitions will differ, but only because of their unique λ_o , f , and Γ . In practice, this means that a greater number of pixels are being employed to constrain the same number of free parameters. If t denotes a given transition of an ion for which N_t transitions are being modeled, then

$$\chi^2 = \sum_{t=1}^{N_t} \sum_{i=j_t^-}^{j_t^+} \left[\frac{(I_i/\bar{I}_i^c)_t - \mathcal{V}_t(\lambda_i; \mathbf{a})}{(\sigma_{D_i})_t} \right]^2, \quad (10.105)$$

where j_t^- and j_t^+ are the starting and ending pixels spanning transition t , for which the Voigt profile model is \mathcal{V}_t . Note that, even though multiple transitions are being modeled, the number of components dictates the number of free parameters, so that the vector \mathbf{a} still has $3N_{\text{VP}}$ elements.

In Figure 10.13, we show the Voigt profile decomposition of kinematically complex MgII $\lambda\lambda 2796, 2803$ doublet absorption at $z_{\text{sys}} = 0.772996$ in the quasar spectrum of Q1248+401 using Eq. 10.105 (see Churchill et al., 2003). The number of components has been minimized using an F-test formalism

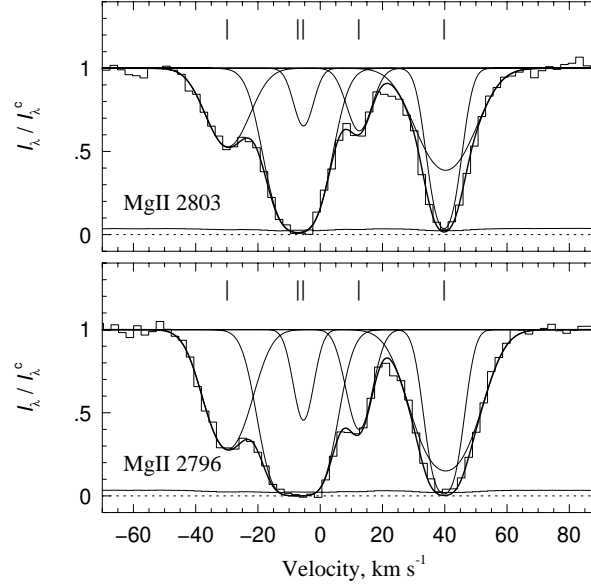


Figure 10.13 An example of Voigt profile decomposition of MgII $\lambda\lambda 2796, 2803$ doublet absorption at $z_{\text{sys}} = 0.772996$ in which both transitions are simultaneously fit to Eq 10.105. The resolution is $R = 45,000$. The thick curve is the model fit to the data. Thin curves are the individual components. Ticks above the continuum are the component centroids. Note that, even using a minimum component approach, two components can have virtually the same velocity (though different Doppler b parameters and column densities) to provide a good fit.

(e.g., Press et al., 2007). The test involves systematically removing components, starting with the component with the largest parameter uncertainties, refitting the data and testing whether the change in the minimum χ^2 value is statistically significant. If the change in χ^2 is not statistically significant for the given degrees of freedom, then the component is not considered to be significant.

In “rich” absorption systems, where absorption lines from several ions are observed in a spectrum, constraints on the number of components, column densities, and line broadening mechanisms are further improved. If we aim to model N_s atomic species, each having $N_{s,t}$ transitions, then the required minimization is

$$\chi^2 = \sum_{s=1}^{N_s} \sum_{t=1}^{N_{s,t}} \sum_{i=j_{s,t}^-}^{j_{s,t}^+} \left[\frac{(I_i/\bar{I}_i^c)_{s,t} - \mathcal{V}_{s,t}(\lambda_i; \mathbf{a})}{(\sigma_{D_i})_{s,t}} \right]^2, \quad (10.106)$$

where $j_{s,t}^-$ and $j_{s,t}^+$ are the starting and ending pixels of spanning transition t of atomic species s , where the index s denotes a given element in a given ionization stage, i.e., MgII or CIV. For each transition t for each atomic species s , the Voigt profile models, $\mathcal{V}_{s,t}$ are computed using Eq. 10.103.

At this stage, it is clear that careful book keeping is required to apply the appropriate atomic constants for each transition, and to apply the appropriate free parameters for each Voigt profile component for each species. This book keeping requires the creation of a translation table to relate the indices of the free parameters for a given absorption component k of species s to the vectorized free parameter list a_m , where $m = 1, 2, 3, \dots, N_{\text{par}}$. This translation table will be specific to the adopted assumptions of the Gaussian component to the line broadening.

For example, if we assume no turbulent contribution, $\Delta v_{\text{turb}} = 0$, the line broadening is strictly thermal. Under these conditions, the Doppler b parameters of a given component k are coupled through the species mass, $b_{k,s'} = \sqrt{m_s/m_{s'}} b_{k,s}$, where s' and s are different species. The free parameter for each component for all transitions of all species is then the absorbing gas temperature T_k , from which we then compute $b_{k,s}$. Instead of having to constrain $N_{\text{VP}} N_s$ unique broadening parameters (one for each component of each species), there are only N_{VP} temperatures to constrain (one for each component applied across all species). If, on the other hand, we assume that the turbulent contribution dominates so that $b_{k,s} = \Delta v_{\text{turb},k}$, then the Doppler parameters are coupled according to $b_{k,s'} = b_{k,s}$ and we have only N_{VP} turbulent velocities to constrain.

If the data are high enough quality and the constraints on the broadening are robust, we may adopt the approach that we wish to leverage the data to constrain both the thermal and turbulent contributions. From Eq. 9.38.

$$b_{k,s}^2 = b_{\text{T},s}^2 + b_{\text{turb},k}^2 = \frac{2kT_k}{m_s} + (\Delta v_{\text{turb},k})^2, \quad (10.107)$$

we solve for $2kT_k$ as a function of $b_{k,s}^2$ and, substituting into $b_{k,s'}^2$ and solving for the turbulent velocity, we obtain the coupling between species total Doppler b parameters for a given component,

$$(\Delta v_{\text{turb},k})^2 = \frac{b_{k,s}^2 - (m_s/m_{s'})b_{k,s'}^2}{1 - (m_s/m_{s'})}. \quad (10.108)$$

In practice, the total b parameters for each component of each species would be a free parameter and be constrained from the model fit, following which Eq. 10.108 would be applied for all s, s' pairs for a given component and the

average $\Delta v_{\text{turb},k}$ and average T_k computed, where

$$T_k = \frac{m_s}{2k} \frac{b_{k,s'}^2 - b_{k,s}^2}{(m_s/m_{s'}) - 1} \quad (10.109)$$

We have presented three scenarios for constraining the physics underlying the line broadening. For scenarios that assume “uniform broadening” across species, we have $N_{\text{par}} = N_{\text{VP}}(N_s + 2)$, accounting for $N_{\text{VP}}N_s$ unique column densities, N_{VP} unique Doppler widths, and N_{VP} unique component redshifts. If no coupling is assumed between the components of different species, then we have “independent broadening” across species, in which case we have $N_{\text{par}} = 2N_{\text{VP}}(N_s + 1)$, accounting for $N_{\text{VP}}N_s$ unique column densities, $N_{\text{VP}}N_s$ unique Doppler widths, and N_{VP} unique component redshifts.

In Table 10.4, we show two possible translation tables for the mapping of the indices of the free parameters required for each component ($k = 1, N_{\text{VP}}$) of each species ($s = 1, N_s$) to the vectorized parameter list a_m ($m = 1, N_{\text{par}}$). For independent broadening, we assume the first $N_{\text{VP}}N_s$ elements are column densities, the next $N_{\text{VP}}N_s$ elements are b parameters, and the final N_{VP} elements are component redshifts. For uniform broadening, we assume the first $N_{\text{VP}}N_s$ elements are column densities, the next N_{VP} elements are b parameters, and the final N_{VP} elements are component redshifts. The resulting translations tables are specific to these adopted structures.

Table 10.4 *Translation Table Vectorizing VP Parameters*

Independent Broadening		Uniform Broadening	
Parameter	Vector Index, m	Parameter	Vector Index, m
$N_{s,k}$	$N_{\text{VP}}(s - 1) + k$	$N_{s,k}$	$N_{\text{VP}}(s - 1) + k$
$b_{s,k}$	$N_{\text{VP}}(N_s + s - 1) + k$	b_k	$N_{\text{VP}}N_s + k$
z_k	$2N_{\text{VP}}N_s + k$	z_k	$N_{\text{VP}}(N_s + 1) + k$

In Figure 10.14 and Table 10.5, we show a Voigt profile decomposition model of the absorption line system at $z_{\text{sys}} = 0.9277$ previously presented in Figures 10.6, 10.8, 10.9, and 10.11. This system is observed to have complex MgII profiles and associated MgI and FeII transitions. We emphasize that there is no unique Voigt profile model as several subjective assumptions must be applied. As with Gaussian decomposition, one of the most fundamental assumptions is that the system profiles should be modeled with the

fewest number free parameters (number of components). We have applied this assumption for the example model presented here, having obtained 10 components.

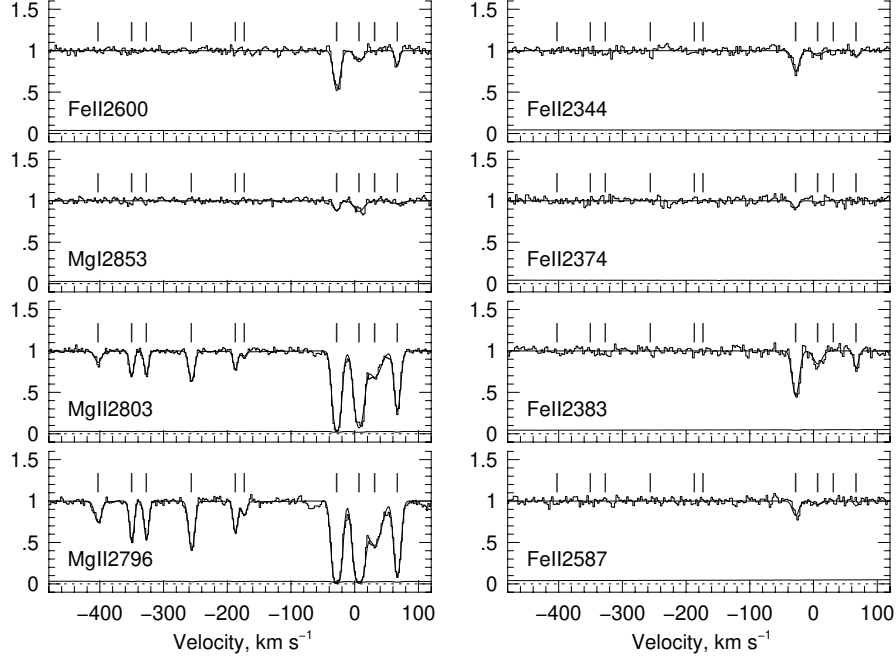


Figure 10.14 Voigt profile decomposition of a complex absorption line system at $z_{\text{sys}} = 0.9277$ selected by MgII $\lambda\lambda 2796, 2803$ absorption. The thin curves through the data are the model spectra. Ticks above the continuum provide the VP component velocities, which are coupled. The column densities and Doppler b parameters for all transitions of a given atomic species are coupled. Finally, the line broadening was modeled such that the Doppler b parameters of magnesium and iron were not coupled. The best-fit parameters are presented in Table 10.5. (This system was observed in January 1995 by the author for his dissertation research and published in Churchill et al. (2003)).

According to the model, three components have measurable absorption from the FeII and MgI ions. As some components are non-detections, the column density upper limits are determined from the curve of growth using the equivalent width limit derived from the uncertainty spectrum. For this model, the Doppler b parameters of a given component were not coupled by species; they were allowed to vary freely between species. As such, for a pure thermal broadening, we would expect $b_{\text{Mg}}/b_{\text{Fe}} = \sqrt{m_{\text{Fe}}/m_{\text{Mg}}} = 1.53$, and for pure turbulent broadening, $b_{\text{Mg}}/b_{\text{Fe}} = 1$. The three measurable component ratios range from 0.96 to 2.02, but are consistent with the physical

Table 10.5 *VP fits*

v_k km s ⁻¹	N_{MgII} cm ⁻²	b_{MgII} km s ⁻¹	N_{FeII} cm ⁻²	b_{FeII} km s ⁻¹	N_{MgI} cm ⁻²	b_{MgI} km s ⁻¹
-402.67	11.93±0.03	6.40±0.58	< 11.52	...	< 10.39	...
-349.84	12.15±0.02	3.18±0.27	< 11.52	...	< 10.40	...
-326.79	12.11±0.02	3.33±0.29	< 11.52	...	< 10.39	...
-255.87	12.34±0.01	4.92±0.21	< 11.51	...	< 10.39	...
-187.06	12.00±0.02	3.33±0.41	< 11.51	...	< 10.39	...
-173.33	11.60±0.05	3.99±1.00	< 11.50	...	< 10.40	...
-27.95	13.32±0.02	6.37±0.11	12.74±0.01	5.90±0.29	11.02±0.06	4.95±1.28
6.89	13.26±0.01	7.76±0.12	12.17±0.05	8.01±1.19	11.28±0.05	10.19±1.37
31.49	12.59±0.01	12.66±0.43	< 11.49	...	< 10.40	...
66.91	12.81±0.01	5.19±0.11	12.06±0.04	2.56±1.00	10.73±0.16	11.43±5.44

range when uncertainties in the Doppler b parameters are taken into account. One might consider if additional components should be used included at $v \simeq -28$ km s⁻¹, as the b parameter for MgI is larger than that of MgII; this would be unexpected physically as the ionization potential of MgI is lower than that of MgII, so it would be expected that MgI would have a narrower width if the line broadening is thermal. This consideration highlights the trade-off between adopting the minimum-component model and attempting to achieve a physically consistent model. One might question the model parameter values obtained for MgI in the $v = 67$ km s⁻¹ component, which may be an artifact of the noise realization in the data; perhaps the component is not above the detection threshold of the spectral region.

It is very important to know the shape and the width of the ISF profile for VP decomposition. Inferring the number of components, or absorbing “clouds”, in a kinematically complex profile, and measuring their column densities, N , and b parameters is very sensitive to the ISF profile shape and width. Adopting a profile that is too broad (narrow) can result in too few (many) components and correspondingly skewed N and b distributions.

11

Gas Physics and Processes

Given that measured absorption line data are quantified in terms of the column densities of various ions, the observed spectra do not directly yield the physical properties of the absorbing gas, such as the phases (densities and temperatures), and chemical and ionization conditions. If we seek to measure the chemical abundances or the gas phase properties, we must attempt to “unfold” several physical unknown quantities from the observables. The measured column density, $N_{k,j}$, of atomic species k in ionization stage j can be expressed in terms of the unknown physical gas quantities,

$$N_{k,j} \simeq n_{k,j}L = f_{k,j}n_kL = n_H f_{k,j} \left(\frac{n_k}{n_H} \right) L, \quad (11.1)$$

illustrating that the key underlying properties giving rise to the observed spectra are (1) the hydrogen number density, n_H , (2) the ionization fraction, $f_{k,j} = f_{k,j}(n_e, T, J_\nu) = n_{k,j}/n_k$, which itself is dependent upon the gas temperature, T , density of free electrons, n_e , and the spectral energy density of the local radiation field, J_ν , (3) the abundance of the atomic species, n_k/n_H , and (4) the path length, L , of the line of sight through the absorbing gas structure. These quantities, at a minimum, are sought to make physical sense of cosmological absorption line spectra.

Since multiple absorption lines from different species are often measured, it can be a challenge to estimate the underlying physical quantities that are consistent with (or that fit) the full set of observed column densities. For example, each species has a unique abundance and each ion has a unique ionization fraction. Furthermore, all measured column densities need to be consistent with the product $n_H L$ and with the ionizing radiation, J_ν , as the shape of the spectral energy density drives the ionization fractions of different elements and ions in differing proportions.

These considerations underscore the need for modeling the gas as con-

strained by the observed column densities. The goal is, for a best estimate of the ionizing spectrum, to determine the gas hydrogen density, equilibrium temperature, and elemental abundances. This is called ionization modeling, which we discuss further in Chapter 12. First, we must develop the formalism to quantitatively characterize the gas while accounting for the most relevant physical gas processes.

In this chapter, we review the basic properties of the radiation field and the particle field of the intergalactic and circumgalactic medium. We will assume an ideal gas that is dynamically static and will not consider the influence of a surrounding medium. Once particle and charge conservation principles are established, the gas equation of state is presented. Then, after discussing the principle of detailed balancing, we examine photo and collisional excitation, de-excitation, ionization, and recombination processes. In the circumgalactic and intergalactic medium, thermodynamic equilibrium does not hold so that the radiation field is not coupled to the thermal properties of the particle field. Finally, we discuss the heating and cooling of the gas, which is critical for establishing the gas temperature and for examining many of the assumptions adopted in this chapter (see § 12.2.1).

11.1 The Radiation Field

As the most common astrophysical environments studied in cosmological absorption line spectra are the intergalactic medium, the circumgalactic medium (and in some cases the interstellar medium), we focus our attention on the radiation fields most common to those environments. The radiation fields present in these media, which are responsible for the photoionization of the absorbing gas, almost never originate in situ. The three most common sources of the ionizing radiation is (1) the ultraviolet cosmic ultraviolet background, which is redshift dependent, (2) emission by quasars/AGN or by fairly local young stellar populations, the latter of which depend on the population total stellar mass, age, metallicity, and distance from the probed absorbing gas, and (3) ionization by both the cosmic background and young stellar populations.

11.1.1 *In the Intergalactic Medium*

The cosmic web of intergalactic gas is embedded in a diffuse background of light that covers the full electromagnetic spectrum from high-energy X-rays to the radio. Though each region of the spectrum interacts with the gas clouds and contributes to the gas physics, we will concern ourselves

primarily with the ultraviolet (UV) spectral region, as this is where photons have high enough energies to directly affect the photoionization balance of hydrogen, helium, and most metals (primarily in the range 10 to 100 eV, see Table 4.1). We call these ionizing photons, which are defined as those photons capable of ionizing neutral hydrogen. Thus they have energies $E \geq 13.598$ eV corresponding to the ultraviolet and soft X-ray portion of the electromagnetic spectrum. This region of the spectrum is often referred to as the Lyman continuum (LyC). We quantify the ionizing photons in a local volume of space by their number density,

$$n_\gamma = \frac{4\pi}{c} \int_{\nu_0}^{\infty} \frac{J_\nu}{h\nu} d\nu, \quad (11.2)$$

where J_ν is the mean intensity (see Eq. 7.10) and $h\nu_0 = 13.598$ eV is the energy at the ionization edge of neutral hydrogen. Hereafter, we generally refer to the ubiquitous background radiation as the cosmic ultraviolet background (cosmic UVB).

The luminous sources of the cosmic UVB are primarily quasars/AGN and galaxies. As these objects evolve dramatically with redshift, so does the intensity and shape of the cosmic UVB. As the cosmic UVB changes, it induces evolution in the physical state of intergalactic gas, which changes the frequency dependent opacity (and transparency) of the gas. As such, understanding the cosmic UVB as a function of redshift requires an iterative approach that accounts for the spectral evolution of quasars and galaxies, and radiative transfer through an expanding universe with an evolving partially ionized and clumpy intergalactic medium.

As posed by Haardt & Madau (1996), the mean intensity of the cosmic UVB at redshift z is obtained by integrating the cosmological radiative transfer equation (e.g., Peebles, 1993), which yields the expression

$$J_\nu(z) = \frac{1}{4\pi} \int_z^\infty \varepsilon(\nu', z') \exp \{ -\bar{\tau}_\nu(z, z') \} \left(\frac{\nu}{\nu'} \right)^3 \frac{dl}{dz'} dz', \quad (11.3)$$

where the relationship between the observed frequency ν at redshift z and the emitted frequency ν' at z' is

$$\frac{\nu}{\nu'} = \frac{1+z}{1+z'}, \quad (11.4)$$

and the general form of the proper line element is given by (see Eq. 6.29)

$$\frac{dl}{dz} = \frac{D_H}{(1+z)} \frac{1}{E(z)}, \quad (11.5)$$

where $D_H = c/H_0$ is the present epoch Hubble distance and $E(z)$ is defined in Eq. 6.15.

The first challenge of evaluating Eq. 11.3 is to estimate the emissivity from quasars and galaxies at frequency ν at all emitting redshifts

$$\varepsilon(\nu, z) = h\nu \oint n_\nu(z) d\Omega, \quad (11.6)$$

where n_ν is the number density of photon of energy ν . The second challenge is treat the radiative transfer through the intergalactic medium to estimate the effective optical depth, $\bar{\tau}_\nu(z, z')$, of photons emitted at redshift z' that are observed at redshifts z . This very delicate problem has been undertaken in a series of four papers by F. Haardt and P. Madau (Madau, 1995; Haardt & Madau, 1996; Madau et al., 1999; Haardt & Madau, 2012), from which they have provided a library of cosmic UVB spectra¹. These spectra are specified as the isotropic mean intensity, J_ν [erg s⁻¹ cm⁻² str⁻¹ Hz⁻¹] over the energy interval $-0.9 \leq \log E \leq 7.7$ [eV] and are provided for redshifts $0 \leq z \leq 15$.

In Figure 11.1, we present the modeled cosmic UVB, produced by the cosmological radiative transfer code CUBA (Haardt & Madau, 2012) for $0 \leq z \leq 6$ in steps of unit redshift. The dotted curve in each panel illustrates the present epoch ($z = 0$). Main features of note are the ionization edges of hydrogen (13.6 eV) and helium (54.4 eV), which manifest as rapid declines in J_ν at those photon energies. They are present due to the increasing effective optical depth of neutral hydrogen and singly ionized helium, due in part to a reduction in the clumpiness of the intergalactic medium for $z > 5$ as a consequence of the epoch of reionization as quasars and galaxies first begin to ionize their surroundings.

11.1.2 In the Circumgalactic Medium

The cosmic UVB is modified on local scales near quasars and galaxies. In the case of quasars, which have quasi power-law spectral energy distributions, i.e., $J_\nu \propto \nu^\alpha$, the radiation field can be significantly “harder” than the cosmic UVB, meaning that the number density of ionizing photons (Eq. 11.2) in the Lyman continuum is greater. This so-called “proximity effect” results in net higher ionization balance in the gas in the circumgalactic medium of quasars (e.g., Weymann et al., 1981; Bajtlik et al., 1988; Scott et al., 2000; Hennawi et al., 2006).

¹ The Haardt & Madau cosmic UVB spectra, created by their code CUBA, are available for download at www.icolick.org/~pmadau/CUBA/DOWNLOADS.html.

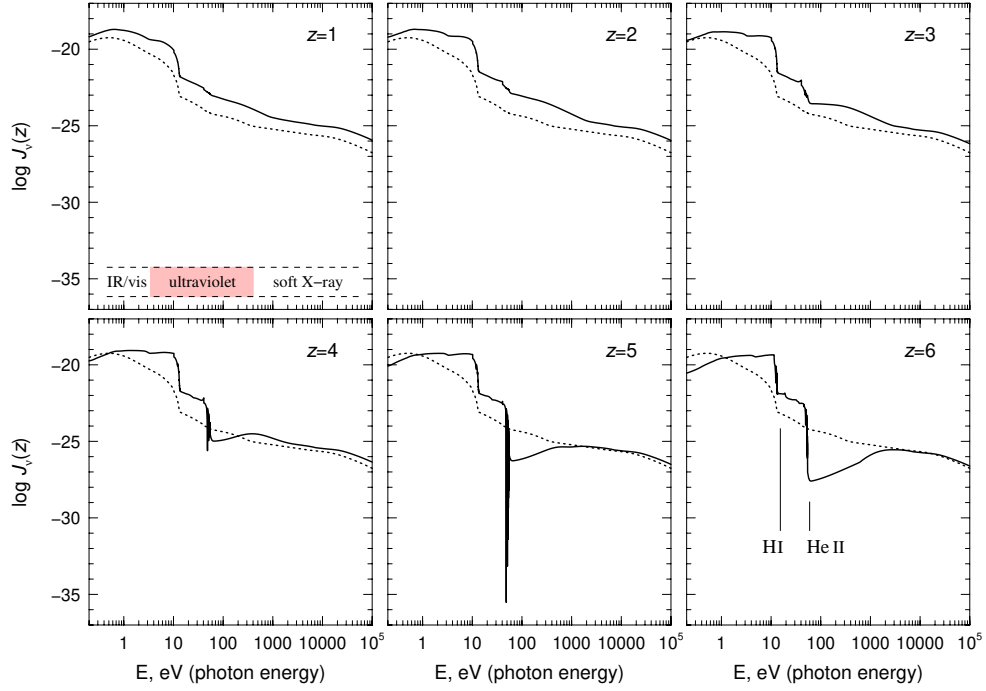


Figure 11.1 Selected redshifts of the cosmic UVB from Haardt & Madau (2012) presented as $J_\nu(z)$ versus photon energy E [eV]. The dotted curve in each panel is $J_\nu(z)$ at $z = 0$. The H I and He II ionization edges are indicated at 13.6 eV and 54.4 eV, respectively. These spectra were generated using the one-dimensional cosmological radiative transfer code called CUBA. They provide working models for the radiation field ubiquitous to the intergalactic medium.

Similarly, ionizing photons from hot O and B stars that escape the interstellar medium of galaxies can also modify the radiation field in their local surroundings. As a result, the circumgalactic medium of active star-forming galaxies can be exposed to a radiation field that is modified from the cosmic UVB. Whereas quasars emit harder spectra, and therefore significant numbers of ionizing photons, only O stars and to a lesser degree B stars emit significant numbers of ionizing photons. However, ionizing photons have a moderately low escape fraction from the interstellar medium of a galaxy. This escape fraction has been observed to be dependent on galaxy luminosity and redshift (e.g., Cassata et al., 2015; De Barros et al., 2017)

In Figure 11.2, we show the relative mean intensity, J_ν , of a composite quasar spectrum constructed by ?, which has power index $\alpha_\nu \simeq -0.75$ (shown by the dotted line). We also show “template” galaxy spectral en-

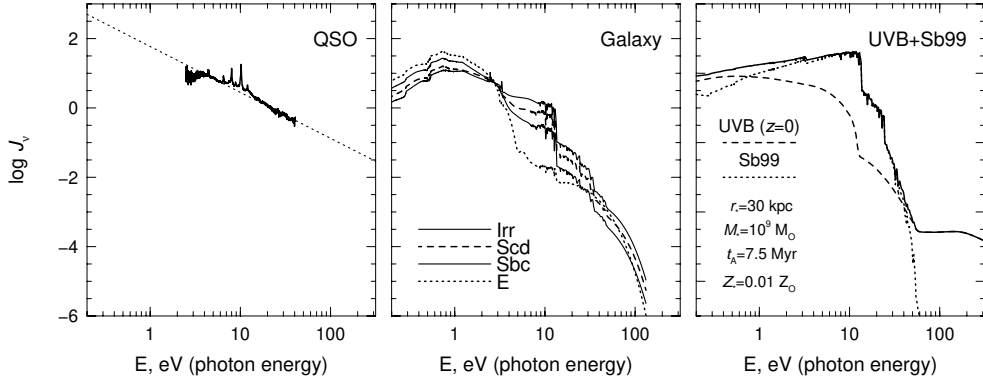


Figure 11.2 Selected spectral energy distribution, J_ν , versus photon energy E [eV] for quasars (QSO), galaxies, and the UVB with added stellar contribution (a single-age stellar population of mass $10^9 M_\odot$ with age 7.5 Myr and a metallicity of 1% solar). For this example, J_ν is calculated at 100 pc from the stellar population assuming a ubiquitous and isotropic UVB and no extinction of the stellar photons. These spectra provide examples of the radiation field throughout the circumgalactic medium for various galaxies and their environments.

ergy distributions for elliptical (E), Sbc, Scd, and Irregular (Irr) galaxies, which are taken from Coleman et al. (1980) with ultraviolet and infrared extensions added by (Bolzonella et al., 2000) using the Gissel98 synthetic spectral evolution library of Bruzual & Charlot (1993). Note that the average quasar lacks any significant self absorption at the Lyman ionization edge, which suggest that Lyman continuum ionizing photons reach into the circumgalactic medium of most quasars. On the other hand, the galaxies exhibit significant self absorption at the Lyman edge. Though Irr and Scd galaxies exhibit harder spectra than Sbc and E galaxies, there is a significant reduction in Lyman continuum ionizing photons reaching into the circumgalactic medium from the average (non-AGN/non-active) galaxy.

For active star-forming or starburst galaxies (in which the star formation rate is some 10–100 times greater than a typical spiral galaxy) the stellar populations can be rich in young O and B stars. In the circumgalactic medium of these galaxies, the cosmic UVB can be substantially modified.

In Figure 11.2, we illustrate a spectral energy distribution (UVB+Sb99) that combines contributions from both the UVB (dashed curve) and a stellar population (dotted curve). For this example, the cosmic UVB is a $z = 0$ CUBA spectrum (Haardt & Madau, 2012) and the stellar population spec-

trum is obtained from the Starburst99 package², which models spectrophotometric and related properties of star-forming galaxies (Leitherer et al., 1999; Vázquez & Leitherer, 2005; Leitherer et al., 2010, 2014). The illustrated stellar population is assumed to have total mass $M_* = 10^9 M_\odot$, age $t = 7.5$ Myr, and a metallicity 1% of solar. This mass is at the lower end typical of local star burst galaxies (e.g., Bergvall et al., 2016). At 7.5 Myr, the population still comprises O5 through O9 stars. Shown as the solid curve in Fig. 11.2 is the combined J_ν at a distance of 30 kpc from the population assuming no extinction from the interstellar or circumgalactic medium. In general, for fixed age and metallicity, the contribution by stars scales linearly with the mass of the stellar population and inversely with the square of the distance (see Churchill & Le Brun, 1998).

Star burst radiation, if luminous enough, modifies the cosmic UVB in the important spectral region between 10 and 100 eV, where hydrogen, helium, and most all metal ions have their ground-state ionization potentials (see Table 4.1). When present, starbursts can generally lead to higher ionization conditions in regions of the circumgalactic medium where photoionization dominates.

11.2 The Particle Field

The particles comprising a gas include atomic/nuclear particles and free electrons. The free electrons are donated to the gas by the ionized atoms. For a given volume of gas, we will denote the number densities in units particles cm^{-3} and mass densities in units g cm^{-3} . The atomic/nuclear densities will be governed by the relative abundances of each atomic species and the equation of state (see § 11.2.8) through the pressure and temperature of the gas. As the free electrons are donated to the gas by ions, the electron density will be dictated by the distribution of ionization stages of the atomic/nuclear particles.

11.2.1 Particle and Mass Density Conservation

Because many atomic species are present in the gas, and because they can be ionized to various levels and each of these atoms/ions can have many various excitation states, it is necessary to describe the ensemble of neutral and ionized atoms (ions) using an indexing scheme. We will denote k as the index for a given atomic species, such as $k = 1$ for hydrogen, $k = 2$ for

² Starburst99 is web based software (www.stsci.edu/science/starburst99/) that generates synthetic spectra for stellar populations of specified total mass, age, and metallicity.

helium, etc. For ionization stages, we employ the index j , which ranges from $j = 1$ for the neutral atom to $j = k + 1$ for the fully ionized atom of species k . For excitation levels, we use the index i , for which $i = 1$ is the ground state.

Using this notation, we define $n_{k,j,i}$ as the number density of atomic species k in ionization stage j and excitation state i . When referring to the number density inclusive of all excitation states of species k in ionization stage j , we invoke $n_{k,j}$. And finally, we define n_k as the number density of species k in all ionization stages and excitation states. Since the number of particles is conserved, the relationships between these densities are

$$n_{k,j} = \sum_{i=1} n_{k,j,i} \quad n_k = \sum_{j=1}^{k+1} n_{k,j} \quad n_N = \sum_k n_k, \quad (11.7)$$

where n_N is the number density of all atomic/nuclear particles in the gas. The total number density of all atoms and free electrons is

$$n = n_N + n_e, \quad (11.8)$$

where n_e is the density of free electrons.

The mass densities are related to the number densities through the mass of species, m_k . Writing $m_k = A_k m_a$, where A_k is the atomic weight of species k in atomic mass units ($m_a = 1.66054 \times 10^{-24}$ g),

$$\rho_{k,j,i} = n_{k,j,i} A_k m_a \quad \rho_{k,j} = n_{k,j} A_k m_a \quad \rho_k = n_k A_k m_a, \quad (11.9)$$

where, from mass density conservations,

$$\rho_{k,j} = \sum_{i=1} \rho_{k,j,i} \quad \rho_k = \sum_{j=1}^{k+1} \rho_{k,j} \quad \rho_N = \sum_k \rho_k. \quad (11.10)$$

11.2.2 Abundances and Mass Fractions

The relative abundance of the various species is an important constraint on the particle densities and resulting electron density. Typically, logarithmic abundances, \mathcal{A}_k , are tabulated on a scale where hydrogen is 12,

$$\mathcal{A}_k = \log \left(\frac{n_k}{n_H} \right) + 12, \quad (11.11)$$

where we explicitly write n_H for the number density of hydrogen. Solar abundance are often tabulated using this scale (e.g., Asplund et al., 2009).

However, it is most convenient to invoke the abundance fraction, defined as the ratio of density of species k to the total atomic/nuclear particle density,

$$\eta_k = \frac{n_k}{n_N}, \quad (11.12)$$

which can be computed from the logarithmic abundances by writing

$$\eta_k = \frac{(n_k/n_H)}{(n_N/n_H)} = \frac{10^{(A_k-12)}}{\sum_k 10^{(A_k-12)}}. \quad (11.13)$$

Note that $\sum_k \eta_k = 1$.

In terms of mass densities, we invoke the mass fraction, which is defined as the ratio of the mass density of species k to the total mass density of atomic/nuclear particles,

$$x_k = \frac{\rho_k}{\rho_N}, \quad (11.14)$$

which can be computed from tabulated logarithmic abundances by writing

$$x_k = \frac{(\rho_k/\rho_H)}{(\rho_N/\rho_H)} = \frac{A_k 10^{(A_k-12)}}{\sum_k A_k 10^{(A_k-12)}}, \quad (11.15)$$

where we have invoked $\rho_k = n_k A_k m_a$ from Eq. 11.9. Note that $\sum_k x_k = 1$.

Important and useful relations are that, given the total mass density of atomic/nuclear particles and the mass fractions, the number density of species k can be determined,

$$n_k A_k m_a = x_k \rho_N \quad (11.16)$$

and that given the total number density of atomic/nuclear particles and the abundance fractions, the mass density of species k can be determined,

$$\rho_k = \eta_k A_k m_a n_N \quad (11.17)$$

Note that if mass fractions are given, but abundance fractions are desired,

$$\eta_k = \frac{x_k/A_k}{\sum_k x_k/A_k}, \quad (11.18)$$

using $n_k = (x_k/A_k)(\rho_N/m_a)$. Or if abundance fractions are given and mass fractions are desired,

$$x_k = \frac{\eta_k A_k}{\sum_k \eta_k A_k}, \quad (11.19)$$

using $\rho_k = \eta_k A_k m_a n_N$.

Mass fractions for the abundance patterns observed in variety of astrophysical environments can be found in Draine (2011), for which the solar pattern is derived from Asplund et al. (2009). Often the mass fractions are written as $X = x_1 = x_H$ for hydrogen, $Y = x_2 = x_{He}$ for helium, and $Z = \sum_{k=3} x_k$ for all metals, where $X + Y + Z = 1$. From Asplund et al. (2009) we have

$$X_\odot = 0.7381 \quad Y_\odot = 0.2485 \quad Z_\odot = 0.0134. \quad (11.20)$$

Expressing abundances relative to solar abundances is a convenient and standard practice. One standard notation is $[X_k/H]$, describing the abundance of atomic species k relative to hydrogen,

$$[X_k/H] = \log \left(\frac{n_k}{n_H} \right) - \log \left(\frac{n_k}{n_H} \right)_\odot, \quad (11.21)$$

where n_k/n_H is given by Eq. 11.11. For abundances expressed via Eq. 11.21, $[X_k/H] = 0$ indicates a solar abundance ratio for atomic species k , where as $[X_k/H] = -2$ indicates a 1% solar abundance. Note that, from Eq. 11.11, we can directly write $[X_k/H] = \mathcal{A}_k - (\mathcal{A}_k)_\odot$.

Different stellar nucleosynthesis processes yield various elements in different proportions, and each of these processes pollute gas in their local astrophysical environment with their own unique proportions of a small group of elements. As such, it highly useful to quantify the relative abundances of various groups of elements on the periodic table, especially the iron group elements. Thus, an additional commonly employed abundance measure relative to solar is

$$[X_k/Fe] = [X_k/H] - [Fe/H] = \log \left(\frac{n_k}{n_{Fe}} \right) - \log \left(\frac{n_k}{n_{Fe}} \right)_\odot, \quad (11.22)$$

The usefulness of comparing $[X_k/H]$ and $[X_k/Fe]$ is that the former provides the degree to which gas is diluted with hydrogen and the latter provides the degree to which a given element of a particular stellar nucleosynthetic origin is present relative to iron group elements. For example, greater dilution by hydrogen results in smaller $[Fe/H]$, but leaves $[X_k/Fe]$ unchanged for all other elements. By contrast, if the gas is preferentially polluted with oxygen, then $[O/Fe]$ ratio will increase, whereas $[Fe/H]$ will be unchanged. The relative behavior of Eqs. 11.21 and 11.22 decode the signatures of the history of gas metallicity enrichment. Of course, based upon the astrophysical application, one can substituted any element into Eqs. 11.21 and 11.22 as the comparison element; for example, oxygen is commonly used because of its origin in stellar

nucleosynthetic origins compared to carbon and nitrogen, i.e., [C/O], [N/O], etc.

Finally, the metallicity of the gas in solar units is defined as the mass fraction of all metals relative to that of hydrogen in units of the solar ratio,

$$Z/Z_{\odot} = \frac{\sum_{k=3} (x_k/x_H)}{\sum_{k=3} (x_k/x_H)_{\odot}}, \quad (11.23)$$

where $x_H = x_1$. The denominator evaluates to $(Z/X)_{\odot} = 0.0182$.

11.2.3 Charge Conservation and Ionization Balance

To fully describe the state of the particle field in a gas, the free electron number density and the number densities for all atomic species and for all of their ionizations stages must be determined. It is because each ion contributes a number of electrons in proportion to its ionization stage that we must know the number densities, $n_{k,j}$, of all ionization stages of all atomic/nuclear species. We therefore must derive a balancing equation in which the $n_{k,j}$ and n_e can be solved for simultaneously. The resulting balancing equation is based upon the principle of charge density conservation.

In a given volume, each ion contributes $(j-1)n_{k,j}$ to the electron density. Note that the neutral stage of an atom makes no contribution. Summing over all contributions, that is over all ionization stages for all species, we have

$$n_e = \sum_k \sum_{j=1}^{k+1} (j-1)n_{k,j}. \quad (11.24)$$

We introduce the ionization fractions

$$f_{k,j} = f_{k,j}(n_e, T, J_{\nu}) = \frac{n_{k,j}}{n_k}, \quad (11.25)$$

where we explicitly write the fraction as a function of n_e , T , and J_{ν} . Computation of the ionization fractions for a gas in thermal equilibrium will be detailed in § 12.2. Adopting the notation, we obtain

$$n_e = \sum_k n_k \sum_{j=1}^{k+1} (j-1)f_{k,j}. \quad (11.26)$$

This relation provides the electron density *if* we know the ionization fractions; but, since the ionization fractions depend upon the electron density

and temperature, we need to know n_e in order to sum the $f_{k,j}$. This implies that we are required to root solve Eq. 11.26. However, we have yet to fully constrain the particle densities, which we do via the equation of state (§ 11.2.8).

If the abundance fractions, η_k , are known for all atomic species, then particle conservation is expressed

$$n_k = \eta_k n_N = \eta_k (n - n_e) , \quad (11.27)$$

where the last step follows from $n = n_N + n_e$. Substituting, we obtain the non-linear balancing equation for the electron density

$$n_e = (n - n_e) \sum_k \eta_k \sum_{j=1}^{k+1} (j-1) f_{k,j} \quad (11.28)$$

which must be root solved using iterative numerical methods.

Once the equilibrium ionization balance is determined, n_e and the $f_{k,j}$ are known, from which the individual number densities can be obtained by employing particle conservation

$$\begin{aligned} n_N &= n - n_e \\ n_k &= \eta_k n_N \\ n_{k,j} &= f_{k,j} n_k \end{aligned} \quad (11.29)$$

Note that the quantities η_k , n , and T must be specified. That is, the abundance fractions dictate the *relative* contributions of the ions for each atomic species, whereas the total particle number density constrains the *magnitude* of the individual densities. Equation 11.28 is closed by the equation of state, which provides n for a given pressure and temperature. As such, the ionization balance and free electron density is coupled to the gas hydrodynamic state.

11.2.4 Thermodynamic Equilibrium

The macro-variable that describes the thermal properties of a gas is temperature, T . As we will discuss, there are multiple simultaneous temperatures that can describe the thermal conditions of a gaseous system, including both the radiation field and the particle field. The term “thermodynamics” comprises much more than thermal properties; it also includes the mechanical, radiative, chemical, ionization, and statistical state of the gas. To say a system is in thermodynamic equilibrium (TE), means the system is in thermal

equilibrium, mechanical equilibrium, radiative equilibrium, ionization equilibrium, and statistical equilibrium. As such, hydrodynamic equilibrium is a highly strict condition.

Equilibrium means that the macro-state variables that describe the micro-state physics of the system are in a steady state. This can result only when the micro-state physics is in state of balance. Thus, in a system in a state of thermodynamic equilibrium, there are no net flows of energy or matter into or out of the system, no phase changes (ionization gradients) across the system, nor any unbalanced potentials (forces) acting across the system. The radiation field must be in equilibrium with the internal energy and the kinetic energy of the particles. Thermodynamic equilibrium is effectively an ideal state that will be naturally achieved (due to the second law of thermodynamics), but only when a system is in mechanical equilibrium and fully isolated from its surroundings.

As mentioned, several temperatures can be used to describe the thermal properties of a gas plus radiation system. Here, we describe the kinetic and radiation temperature.

A single kinetic temperature, T , can be used to describe the particle field when the particle velocity distribution reflects a steady-state equipartition of the energy density of states. The resulting velocity distribution is known as the Maxwell-Boltzmann distribution, which we describe in more detail in § 11.2.5. When the particles obey this steady-state distribution, they are said to be “thermalized”. A system that is thermalized is said to be in thermal equilibrium. Note that a system can be in thermal equilibrium while not being in thermodynamic equilibrium.

The radiation temperature, T_r , is a temperature that specifies the energy density in the radiation field, u_r , according to $u_r = aT_r^4$ [erg cm⁻³], where a is the radiation constant. Since, for a black body, $u_r = (4\pi/c)B(T_r) = aT_r^4$, where $B(T_r)$ is the Planck function integrated over all frequencies (see Eqs. 7.53 and 7.54), T_r can also be thought of as the temperature that appears in the Planck function having the same energy density as the radiation field in the system. A system is said to be in radiative equilibrium if any and all randomly selected subsystems (spatially separated volume elements) exchange equal quantities of radiative energy with each other in steady state. It is difficult for a system to be in strict radiative equilibrium if it is not in thermodynamic equilibrium.

In thermodynamic equilibrium, photons are not allowed to escape the system (and this would result in an energy loss), and we have $T = T_r$ in steady state. For this condition to be stable, photons produced in the gas (whose energy density is parameterized by T_r) must be tightly coupled to the random

kinetic motions and to the internal energy distribution of the atomic particles (whose energy density is parameterized by T). This strong coupling implies that, in order for thermodynamic equilibrium to hold, the excitation, de-excitation, and ionization of the atoms and ions must be dominated by collisions, i.e., collision occur more frequently than photo-excitations and radiative spontaneous de-excitations. If that were not the case, then T and T_{R} would diverge from equality. It would also mean that the distribution of excitations states and ionization levels of the atomic particles would not remain in energy equipartition. In § 11.4.1 we discuss the equipartitioned distribution of excitation states in thermodynamic equilibrium and in § 12.2.2 we discuss the distribution of equipartitioned ionization levels in thermodynamic equilibrium.

A system is said to be in statistical equilibrium if the excitation level populations of the atoms and ions are in steady state (the transition rate into any given level equals the rate out); it does not imply that the level populations are in energy equipartition (that would occur only in thermodynamic equilibrium). A system is said to be in ionization equilibrium if the ionization stages (or ionization fractions) of the atoms and ions are in steady state. The tracking of the micro-physical processes in statistical equilibrium is called “detailed balancing”, a topic we discuss in § 11.3. Note that the term “steady state”, and therefore the working definition of a system being in equilibrium, implies only that the timescales (for example, of excitation, de-excitation, and/or ionization) are significantly shorter than the timescale for any change in the thermal state of the system. We will address this further in § 12.2.1.

Finally, we define mechanical equilibrium to hold if no net translational or rotational external forces act on the system or internally across the system. For example, a systems undergoing gravitational collapse, exhibiting shock fronts, or undergoing acceleration due to radiative pressure or mechanical forces are examples of systems that are not in mechanical equilibrium. In our treatment of absorbing gas physics, we will be assuming mechanical equilibrium.

11.2.5 Thermalized Velocity Distributions

Assume a gas system in which the particles are “thermalized”, which is to say there is complete equipartition of the energy-density distribution of the particles and they are in steady-state, i.e., thermal equilibrium holds. In this case, the gas can be described by a single kinetic temperature, T .

Consider an isothermal gas with no bulk motion (mechanical equilibrium)

comprising particles of species k having mass m_k . We shall adopt a cartesian coordinate system for describing the particle velocities. If $n_k(v_x)$ is the number of particles of species k having mass m_k in the interval $v_x \rightarrow v_x + dv_x$, and n_k is the total number of particles of species k in the gas volume, then the fraction of particles, $f_k(v_x, T) = n_k(v_x)/n_k$, of species k in the interval $v_x \rightarrow v_x + dv_x$, is

$$f_k(v_x, T) dv_x = \left(\frac{m_k}{2\pi kT} \right)^{1/2} \exp \left\{ -\frac{m_k v_x^2}{2kT} \right\} dv_x. \quad (11.30)$$

As will be shown in § 11.2.6, the product kT is proportional to the mean non-relativistic kinetic energy of particles in a gas with temperature T . The distribution of particle velocities in any single direction (and subsequently, along any line of sight through the gas) is simply a Gaussian distribution. The normalization is

$$\int_0^\infty f_k(v_x, T) dv_x = 1. \quad (11.31)$$

The distribution of particle speeds, $f_k(v, T) = n_k(v)/n_k$ in the interval $v_x \rightarrow v_x + dv_x$, $v_y \rightarrow v_y + dv_y$, and $v_z \rightarrow v_z + dv_z$, is

$$f_k(v_x, T) f_k(v_y, T) f_k(v_z, T) dv_x dv_y dv_z = f_k(v, T) 4\pi v^2 dv \quad (11.32)$$

where $v^2 = v_x^2 + v_y^2 + v_z^2$, and $dv_x dv_y dv_z = 4\pi v^2 dv$, is obtained by replacing the cartesian velocity elements $dv_x dv_y dv_z$ by the spherical element $4\pi v^2 dv$, which can be visualized as a shell of thickness $v \rightarrow v + dv$. The distribution of particle velocities is then

$$f_k(v, T) dv = \left(\frac{m_k}{2\pi kT} \right)^{3/2} \exp \left[-\frac{m_k v^2}{2kT} \right] 4\pi v^2 dv, \quad (11.33)$$

known as a Maxwell-Boltzmann distribution, which has normalization,

$$\int_0^\infty f_k(v, T) dv = 1. \quad (11.34)$$

For a particle of mass m_k , the most probable, average, and RMS speeds are

$$v_{0k} = \left(\frac{2kT}{m_k} \right)^{1/2}, \quad \langle v_k \rangle = \left(\frac{8kT}{\pi m_k} \right)^{1/2}, \quad \text{and} \quad v_{\text{RMS}k} = \left(\frac{3kT}{m_k} \right)^{1/2}, \quad (11.35)$$

respectively. Note that more massive particles have smaller v_{0k} , $\langle v_k \rangle$, and $v_{\text{RMS}k}$ at a given T .

The Maxwell-Boltzmann speed distributions (Eq. 11.33) are illustrated in Figure 11.3(a) for hydrogen at three gas temperatures in the range of what is observed in both the intergalactic and circumgalactic medium. Note

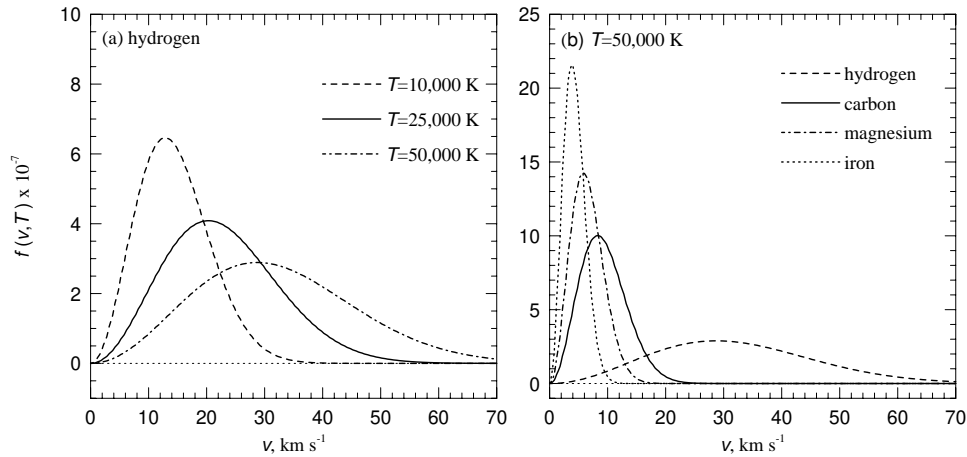


Figure 11.3 The Maxwell-Boltzmann distribution of particle velocities and speeds. (a) The distribution of the x component of velocity for hydrogen at $T = 10,000, 25,000$, and $50,000 \text{ K}$. (b) The distribution of the x component of velocity for hydrogen, carbon, magnesium, and iron at $T = 50,000 \text{ K}$. (c) The distribution of particle speeds for hydrogen at $T = 10,000, 25,000$, and $50,000 \text{ K}$. (d) The distribution of particle speeds for hydrogen, carbon, magnesium, and iron at $T = 50,000 \text{ K}$.

that, for a fixed particle mass, the velocity dispersion increases and the amplitude decreases with increasing T . In Figure 11.3(b), the dependence with species particle mass is shown for the common elements hydrogen, carbon, magnesium, and iron at $T = 50,000 \text{ K}$ (the particle mass ratios are roughly 1:12:24:56). Note that the distribution dispersion decreases and the amplitude increases with increasing particle mass.

The Maxwell-Boltzmann distribution is achieved because particle collisions redistribute and partition individual particle kinetic energies among all particles (by “collisions”, we mean Coulomb interactions). This kinetic redistribution happens on a much shorter timescale than ionization and recombination equilibrium can be achieved, so that one can regard thermalization as effectively being instantaneous. However, the thermalization process is hierarchical by mass and electric charge.

If the particle field is out of thermal equilibrium, electrons (being electrically charged and of low mass) thermalize quickly with themselves first. The electrons are also transferring energy to the electrically charged atomic ions until equipartition of energy is achieved, however this process has a longer bottom-up mass-ordered timescale because, the higher the mass difference between the ion and the electron, the less efficient the kinetic energy transfer. Finally, the neutrals (which are less influenced by Coulomb interactions)

are thermalized. Even if the atomic particle field is not quite thermalized, it is usually safe to assume that the free electron pool is thermalized.

11.2.6 Kinetic Energy

For given particle species k of mass m_k having non-relativistic speed v_k , its non-relativistic kinetic energy is $\frac{1}{2}m_k v_k^2$. The mean kinetic energy of particles of mass m_k in an isothermal gas of temperature T is

$$\langle \text{KE} \rangle = \frac{1}{2}m_k \langle v_k^2 \rangle, \quad (11.36)$$

with

$$\langle v_k^2 \rangle = \frac{\int_0^\infty v^2 f_k(v, T) dv}{\int_0^\infty f_k(v, T) dv}, \quad (11.37)$$

where $f_k(v, T)$ is the Maxwell-Boltzmann distribution for particles of species k , given by Eq. 11.33. The denominator is unity. Invoking Eq. 11.35, we have

$$\langle v_k^2 \rangle = \int_0^\infty v^2 f_k(v, T) dv = \frac{4}{\sqrt{\pi}} \frac{1}{v_{0k}^3} \int_0^\infty \exp \left\{ - (v/v_{0k})^2 \right\} v^4 dv. \quad (11.38)$$

The integral evaluates to $\frac{3}{8}\sqrt{\pi}v_{0k}^5$, which yields

$$\langle v_k^2 \rangle = \frac{3}{2}v_{0k}^2 = \frac{3kT}{m_k}, \quad (11.39)$$

from which we obtain the average kinetic energy per particle

$$\langle \text{KE} \rangle = \frac{3}{2}kT. \quad (11.40)$$

An interesting property of the mean kinetic energy is that it is independent of particle mass. Thus, the mean kinetic energy of all particle species are the same, being dependent only upon the first power of gas temperature.

The total kinetic energy density [erg cm⁻³] of the gas particles is obtained by multiplying the average kinetic energy per particle by the particle number density. The total particle kinetic energy density is

$$Q = (n_N + n_e) \langle \text{KE} \rangle = \frac{3}{2} (n_N + n_e) kT = \frac{3}{2} n kT. \quad (11.41)$$

11.2.7 Pressure

Pressure, P , is defined as force, F , per unit area, where force is the rate of momentum transfer, $F = dp/dt = m(dv/dt)$. In a gas, this can be visualized

as the rate at which particles transfer momentum across an imaginary unit area within a unit volume.

The flux of particles of species k crossing both directions through a unit area per unit time is

$$\frac{dn_k}{dt} = \frac{1}{3}n_k \langle v_k \rangle, \quad (11.42)$$

where n_k is the number density of particles of species k , and $\langle v_k \rangle$ is the mean velocity assuming a Maxwell-Boltzmann velocity distribution (Eq. 11.33). Multiplying by the momentum per particle, we obtain the pressure

$$P_k = m_k \langle v_k \rangle \frac{dn_k}{dt} = \frac{1}{3}n_k m_k \langle v_k^2 \rangle = \frac{2}{3}n_k \langle \text{KE} \rangle. \quad (11.43)$$

Note that the gas pressure is proportional to the particle mean kinetic energy. Applying Eq. 11.40, we have

$$P_k = n_k kT, \quad (11.44)$$

which is independent of m_k . The quantity P_k is the partial pressure for particles of species k . The pressure from all atomic/nuclear particles is the sum of the partial pressures,

$$P_N = \left(\sum_k n_k \right) kT = n_N kT. \quad (11.45)$$

To obtain the total gas pressure, we add the partial pressure of the free electrons,

$$P_g = P_N + P_e = n_N kT + n_e kT = (n_N + n_e) kT = nkT. \quad (11.46)$$

Alternatively, applying Eqs. 11.36 and 11.39, we have

$$P_k = \frac{1}{3}n_k m_k \langle v_k^2 \rangle = (n_k m_k) \frac{kT}{m_k} = \frac{\rho_k}{m_k} kT \quad (11.47)$$

where $\rho_k = n_k m_k$ is the mass density [g cm^{-3}] of species k . The total atomic/nuclear particle pressure can then also be written in terms of the total nuclear particle mass density, ρ_N via the summation

$$P_N = \sum_k P_k = \left(\sum_k \frac{\rho_k}{m_k} \right) kT = \frac{k}{\mu_N m_a} \rho_N T, \quad (11.48)$$

where we have defined the “mean molecular weight” for the atoms/ions

$$\frac{1}{\mu_N} = \frac{m_a}{\rho_N} \left(\sum_k \frac{\rho_k}{m_k} \right), \quad (11.49)$$

and where m_a is the atomic mass unit. From the definition of mass fraction, $x_k = \rho_k/\rho_N$, and employing $m_k = A_k m_a$, we have

$$\frac{\rho_k}{m_k} = \frac{\rho_N}{m_a} \frac{x_k}{A_k}, \quad (11.50)$$

which, after substitution into Eq. 11.49, we obtain

$$\mu_N = \left[\sum_k \left(\frac{x_k}{A_k} \right) \right]^{-1}. \quad (11.51)$$

The same principles apply to obtain the partial pressure from the free electrons in the gas in terms of the total nuclear particle mass density, ρ_N ,

$$P_e = n_e kT = \frac{k}{\mu_e m_a} \rho_N T, \quad (11.52)$$

where we define μ_e , the mean molecular weight of electrons,

$$\mu_e = \frac{\rho_N}{m_a} \frac{1}{n_e}. \quad (11.53)$$

From charge density conservation (Eq. 11.26),

$$n_e = \sum_k n_k \sum_{j=1}^{k+1} (j-1) f_{k,j}, \quad (11.54)$$

and employing $n_k = (x_k/A_k)(m_a/\rho_N)$ (see Eq. 11.16), we obtain

$$\mu_e = \left[\sum_k \left(\frac{x_k}{A_k} \right) \sum_{j=1}^{k+1} (j-1) f_{k,j} \right]^{-1}. \quad (11.55)$$

Note that μ_e depends on both the mass fractions of the atomic/nuclear particles *and* the ionization balance of the gas via Eq. 11.28.

The total gas pressure is the sum of the partial pressures of the atoms/ions and the free electrons

$$P_g = P_N + P_e = \frac{k}{m_a} \left(\frac{1}{\mu_N} + \frac{1}{\mu_e} \right) \rho_N T = \frac{k}{\mu m_a} \rho_N T \quad (11.56)$$

where we define the total mean molecular weight of the gas

$$\frac{1}{\mu} = \frac{1}{\mu_N} + \frac{1}{\mu_e}. \quad (11.57)$$

However, the “massless electron approximation” is often implicitly used to

write the gas pressure in terms of the total particle mass density, ρ . Consider the ratio

$$\frac{\rho}{\rho_N} = \frac{\rho_N + \rho_e}{\rho_N} = \left(1 + \frac{\rho_e}{\rho_N}\right). \quad (11.58)$$

The magnitude of ρ_e is a maximum when the gas is fully ionized. We now show that even in this extreme scenario, $\rho_e/\rho_N \ll 1$, which yields $\rho \simeq \rho_N$ as an excellent approximation. For a fully ionized gas, all $f_{k,j} = 0$, except for $f_{k+1,k} = 1$, i.e., there are no partially ionized atoms in the gas. Thus,

$$\rho_e = m_e n_e = m_e \left[\sum_k n_k \sum_{j=1}^{k+1} (j-1) f_{k,j} \right] = m_e \sum_k k n_k. \quad (11.59)$$

For the atoms/nuclear particles,

$$\rho_N = \sum_k m_k n_k = m_a \sum_k A_k n_k \simeq 2m_a \sum_k k n_k, \quad (11.60)$$

where we have approximated $A_k \simeq 2k$, which is fairly accurate for all species except hydrogen. We have

$$\frac{\rho}{\rho_N} = \left(1 + \frac{\rho_e}{\rho_N}\right) \simeq \left(1 + \frac{m_e}{2m_a}\right) = (1 + 0.00028) \simeq 1. \quad (11.61)$$

Therefore, even for the extreme scenario in which the gas is fully ionized, the electron mass density is no greater than 0.03% of the atomic/nuclear particle mass density. Adopting the massless electron approximation, all pressures can be written to high accuracy in terms of the total mass density, i.e., $\rho \simeq \rho_N$,

11.2.8 The Equation of State

The equation of state relates the gas density and temperature to the hydrodynamic conditions the gas is experiencing. We limit our discussion to a static ideal gas, in which case the hydrodynamics simplifies to the hydrostatic pressure. In the intergalactic medium and the majority of the circumgalactic medium, the gas is not in equilibrium with the radiation field; further, the gas is not experiencing acceleration by radiation pressure, so we neglect this condition (which can occur in galactic scale winds and in the outflows of quasars, e.g., Vilkoviskij et al., 2003; Hamann & Sabra, 2004).

Neglecting radiation pressure, the total pressure of an ideal gas is

$$P = P_g = P_N + P_e \quad (11.62)$$

where P_g is the gas pressure, P_N is the partial pressure from nuclear particles (atoms and ions), P_e is the partial pressure from free electrons.

In terms of number densities, we have $P = (n_N + n_e)kT = nkT$, yielding

$$n(P, T) = \frac{P}{kT}. \quad (11.63)$$

As such, we see that the total density of the gas is constrained by the temperature and pressure. However, the equation of state does not provide either the nuclear particle nor the free electron density. Those must be determined through particle and charge density conservation, which requires the ionization conditions to obtain n_e .

In terms of mass densities, we have $P = (k/\mu m_a)\rho T$, where we apply the massless electron approximation (Eq. 11.61). We have

$$\rho(P, T, \mu) = \frac{\mu m_a}{kT} P. \quad (11.64)$$

Since the mean molecular weight, μ , depends upon both the mass fractions (to obtain μ_N , and the ionization fractions, to obtain μ_e , we find that we require the ionization condition of the gas in order to “close” the equation of state.

Small changes in the equation of state can be determined via the relation

$$\frac{d\rho}{\rho} = \left(\frac{\partial \ln \rho}{\partial \ln P} \right) \frac{dP}{P} + \left(\frac{\partial \ln \rho}{\partial \ln T} \right) \frac{dT}{T} + \left(\frac{\partial \ln \rho}{\partial \ln \mu} \right) \frac{d\mu}{\mu}, \quad (11.65)$$

where the partial derivatives are the logarithmic slopes of the respective independent variables. As such, this equation allows us to examine a fractional change in the mass density in terms of fraction changes in the pressure, temperature, and/or mean molecular weight. The logarithmic slopes (or power-law indices) are

$$\frac{\partial \ln \rho}{\partial \ln P} = \frac{d\rho}{\rho} \frac{P}{dP} = 1 \quad \frac{\partial \ln \rho}{\partial \ln T} = \frac{d\rho}{\rho} \frac{T}{dT} = -1 \quad \frac{\partial \ln \rho}{\partial \ln \mu} = \frac{d\rho}{\rho} \frac{\mu}{d\mu} = 1, \quad (11.66)$$

which yields

$$\frac{d\rho}{\rho} = \frac{dP}{P} - \frac{dT}{T} + \frac{d\mu}{\mu}. \quad (11.67)$$

11.3 Detailed Balancing

Having described both the radiation field and the particle field in some detail, we now briefly discuss the interaction between the radiation field and the particle field (radiative processes) and the particle field with itself

(collisional processes). Due to interactions and quantized energy exchange, each atomic species in a volume element of gas will be rapidly changing ion stages and excitation states.

As illustrated in Figure 11.4(a,b,e,f), the simplest excitation, and de-excitation, processes are collisional and radiative, the latter being well-known as absorption and spontaneous emission, respectively. As illustrated in Figure 11.4(c,g), the simplest processes of ionization are photoionization and collisional ionization. The simplest recombination processes are radiative recombination, and three-body collisional recombination, as illustrated in Figure 11.4(d,h). An additional process of ionization and recombination is charge exchange (not shown in Figure 11.4).

Atomic collisional ionization, excitation, and de-excitation can be induced by free electrons, free protons, or other ions. Free electrons are the dominate collisional particle for the ions in the gas, where as other ions are the dominate particle for neutral atoms in the gas. For both ionization and recombination there are other important higher-order important processes, such as Auger ionization in which two or more electrons are simultaneously ejected during the ionization process, dielectronic recombination in which an electron recombines to an excited state and excites another electron in the process, and excitation auto-ionization, in which a collisional excitation is followed by ejection of an electron (instead of a photon) as the atom/ion spontaneously de-excites. Thus, there are various branching ratios of higher order processes that involve combinations of excitations and ionization or recombinations and excitations and/or ionization. We discuss several of the more important higher-order processes later in this chapter.

In the very special case of thermodynamic equilibrium, the particle field kinetic energy distribution and the distribution of internal energies (i.e., the occupation of excitation and ionization states of atoms and ions) is in energy equilibrium with the radiation field. In this case, the particle kinetic energy distribution is a Maxwell-Boltzmann distribution (Eq. 11.33) and the photon energy distribution is Planck function (Eq. 7.53). The known partition of energies allows the distribution of excited states and ionization levels to be expressed in simple form, i.e., the Boltzmann equation (see Eq. 11.78) and the Saha-Langmuir equation (see Eq. 12.26), respectively.

In general, the excitation, de-excitation, ionization, and recombination rates and their equilibrium balance are related through the global balance equation, a matrix of non-linear rate equations that must be solved simultaneously. Consider the complete state space, S , representing the equilibrium occupation of states for all atomic species k in ionization levels j , and excitation states i . The global balance equation is a matrix, for which each row

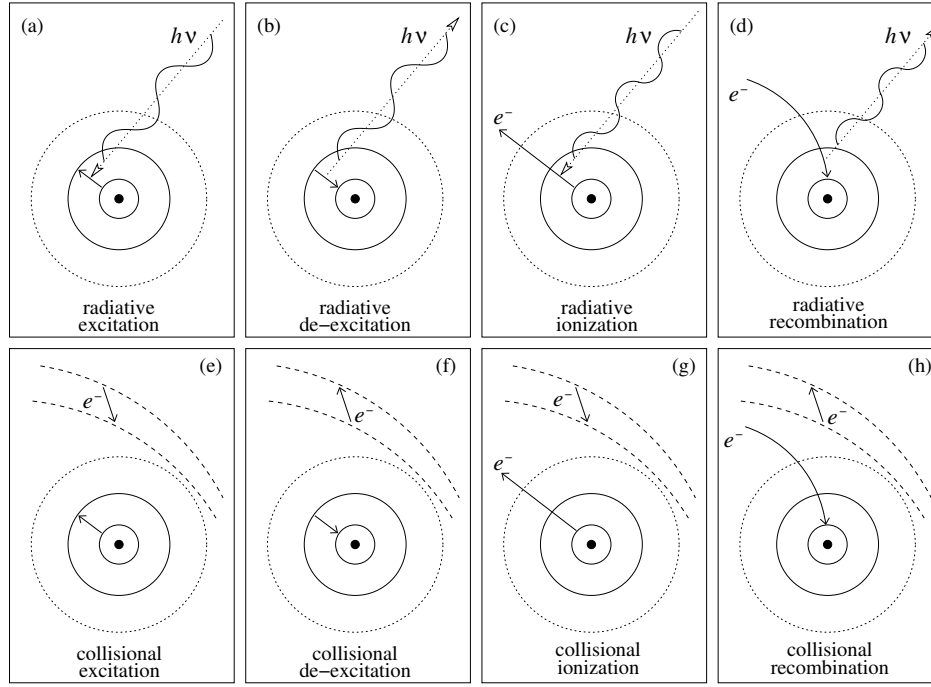


Figure 11.4 Schematics of the simplest, or primary, interactions between (top) the particle field and the radiation field and (bottom) the particle field with itself. (a) Radiative excitation (also called bound-bound absorption). (b) Radiative de-excitation (also known as spontaneous decay or emission). (c) Radiative ionization. (d) Radiative recombination. For excitation and de-excitation, internal energy in the atom/ion is exchanged with the radiation field. In ionization, radiative energy is deposited into the electron pool as kinetic energy (less the binding energy). In recombination, kinetic energy from the electron pool is converted into radiative energy. (e) Collisional excitation. (f) Collisional de-excitation. (g) Collisional ionization. (h) Collisional (three-body) recombination. For excitation and de-excitation, internal energy in the atom/ion is exchanged with the kinetic energy of the collisional electron. For collisional ionization, kinetic energy from the collisional electron is given up to ejecting a bound electron (the energy in the electron pool is reduced by the binding energy of the ejected electron). For collisional recombination, during a recombination, instead of radiative energy being emitted, a proximate electron is elevated to a higher kinetic energy.

is the equation for species k

$$\frac{dn_{k,j,i}}{dt} = \sum_{kj'i' \in S} n_{k,j',i'} R_{k,j,i}^{k,j',i} - n_{k,j,i} \sum_{kj'i' \in S} R_{k,j,i}^{k,j',i'} \equiv 0, \quad (11.68)$$

where $R_{k,j,i}^{k,j',i'}$ is the reaction rate out of state k, j, i to state k, j', i' in units of

reactions s^{-1} , and where both sums exclude the state indexed by $k, j' i' = k, j, i$. The second term on the right-hand side of Eq. 11.68 represents the rate per unit volume out of state k, j, i into states other than k, j, i , whereas the second term represents the rate per unit volume out of all states $k j' i' \neq k, j, i$ into state k, j, i . The global balance equation of the complete state space of a gas is virtually an intractable problem without simplifying assumptions.

The most common and useful simplifying assumption is detailed balancing, which is an equilibrium condition imposed on the rates of the radiative and collisional processes that govern the excitation and ionization balance. The first condition is that, for every atom/ion species k, j , the number of excitations and de-excitations into state i per unit time per unit volume equals the number of excitations and de-excitations out of state i per unit time per unit volume, expressed

$$\sum_{i'} n_{k,j,i'} R_{k,j,i'}^{k,j,i} - n_{k,j,i} \sum_{i'} R_{k,j,i}^{k,j,i'} = 0. \quad (11.69)$$

The second condition is that for every atom/ion species k, j , the number of ionization and recombinations into ionization level j per unit time per unit volume equals the number of ionizations and recombinations out of ionization level j per unit time per unit volume,

$$\sum_{j'} n_{k,j'} R_{k,j'}^{k,j} - n_{k,j} \sum_{j'} R_{k,j}^{k,j'} = 0. \quad (11.70)$$

Detailed balancing yields a state space of the gas that is steady state yet does not require thermodynamic equilibrium. In what follows, we will assume detailed balancing applies.

11.4 Excitation

It is fairly rare to observed absorption from excited states, as the conditions of the intergalactic medium and circumgalactic medium are such that virtually all ions are in their ground states. However, in cases where the line of sight to the luminous source probes the interstellar medium of an intervening galaxy or satellite galaxy, or probes the associated gaseous environment of the source itself, the density of the absorbing gas can be high enough that collisional excitations can occur.

As listed in Table 4.5, two examples of excited ions seen in absorption are CII^* and SiII^* . As we discussed in § 4.4.3 and illustrated in Figures 4.14 and 4.15, these Group IIIA metals have a $J = 3/2$ fine-structure splitting with their $J = 1/2$ ground states. Because the excitation energy to the $J = 3/2$

state is very small (for example, $\chi = 0.00786$ eV for C II*), collisions with free electrons can elevate the ground-state electron to its excited fine-structure state. The same holds for Si II* and all Group IIIA ions (Al I, C II, Si II, N III, and O IV).

As summarized by Bahcall & Wolf (1968), detailed balancing can be applied to the Group IIIA atomic fine-structure excited states to obtain the number density of free electrons in the absorbing gas. First, it is assumed that the ion is a two-level atom because the next lowest excitation level in these ions ranges from 6.9 eV to 15.7 eV and the mean kinetic energy of an electron is 0.3–1.6 eV in gas of $T = 10,000$ K to 50,000 K, respectively. So, whereas the energy is sufficient to populate the fine-structure excited state via electron collisions, it would be the rarefied high-energy tail of the Maxwell-Boltzmann distribution (Eq. 11.33) that would give rise to a negligibly small fraction of the higher excited states.

Since we can apply a two-level atom, we will employ the notation that the $J = 1/2$ ground state is level “1” and the excited $J = 3/2$ fine-structure state is level “2”. Applying Eq. 11.69 to a two-level model, we have

$$\frac{n_2}{n_1} = \frac{R_{1,2}^{\text{coll}} + R_{1,2}^{\text{rad}}}{R_{2,1}^{\text{coll}} + R_{2,1}^{\text{rad}}}, \quad (11.71)$$

where n_1 and n_2 are the number densities of the ion in level 1 and level 2, respectively, and where $R_{1,2}^{\text{coll}}$ is the collisional excitation rate and $R_{1,2}^{\text{rad}}$ is the radiative excitation rate of level 1 to level 2, and $R_{2,1}^{\text{coll}}$ is the collisional de-excitation rate and $R_{2,1}^{\text{rad}}$ is the radiative de-excitation rate of level 2 to level 1.

11.4.1 Excitation Rates and Thermodynamic Equilibrium

Under the assumption that collisional processes dominate over radiative processes i.e., $R_{1,2}^{\text{coll}} \gg R_{1,2}^{\text{rad}}$ and $R_{2,1}^{\text{coll}} \gg R_{2,1}^{\text{rad}}$, the steady-state population of excitation states will obey

$$\frac{n_2}{n_1} = \frac{R_{1,2}^{\text{coll}}}{R_{2,1}^{\text{coll}}}. \quad (11.72)$$

The rate of collisional excitations is the product of the number density of free electrons and the collisional excitation rate coefficient, $\alpha_{1,2}^{\text{ex}}$, and the rate of collisional de-excitations is the product of the electron number density and the collisional de-excitation rate coefficient, $\alpha_{2,1}^{\text{ex}}$. We have

$$R_{1,2}^{\text{coll}}(T) = n_e \alpha_{1,2}^{\text{ex}}(T), \quad R_{2,1}^{\text{coll}}(T) = n_e \alpha_{2,1}^{\text{ex}}(T), \quad (11.73)$$

where the units of the coefficients are $\text{cm}^3 \text{s}^{-1}$. The rate coefficients are the means of the product of the velocity-dependent cross sections and the incident electron velocity,

$$\alpha_{1,2}^{\text{ex}}(T) = \langle \sigma_{1,2}^{\text{ex}} \cdot v \rangle = \int_{\chi_{1,2}}^{\infty} \sigma_{1,2}^{\text{ex}}(v) f(v, T) v dv, \quad (11.74)$$

and

$$\alpha_{2,1}^{\text{ex}}(T) = \langle \sigma_{2,1}^{\text{ex}} \cdot v \rangle = \int_0^{\infty} \sigma_{2,1}^{\text{ex}}(v) f(v, T) v dv, \quad (11.75)$$

where $E = \frac{1}{2}m_e v^2$ is the kinetic energy of the incident electron, $\sigma_{1,2}^{\text{ex}}(v)$ and $\sigma_{2,1}^{\text{ex}}(v)$ are the collisional excitation and de-excitation cross sections [cm^2], respectively, $f(v, T)$ is the Maxwell-Boltzmann distribution function (Eq. 11.33) of free electrons thermalized at temperature T , and $\chi_{1,2}$ is the excitation energy between level 1 and level 2. Note that in Eq. 11.74, excitation can take place only if the kinetic energy of the electron is $\frac{1}{2}m_e v^2 \geq \chi_{1,2}$, where as de-excitation, as expressed in Eq. 11.75, can be induced even when $\frac{1}{2}m_e v^2 \simeq 0$. Applied to Eq. 11.72, we have

$$\frac{n_2}{n_1} = \frac{\alpha_{1,2}^{\text{ex}}}{\alpha_{2,1}^{\text{ex}}}. \quad (11.76)$$

In the special case of thermodynamic equilibrium in which the radiation field is a Planck function (Eq. 7.53) characterized by the same temperature as the thermalized electrons and thermalized atoms and ions in the particle field, L. Boltzmann derived the relation

$$\frac{\alpha_{1,2}^{\text{ex}}}{\alpha_{2,1}^{\text{ex}}} = \frac{g_2}{g_1} \exp \{ -\chi_{1,2}/kT \}, \quad (11.77)$$

where g_1 and g_2 are the statistical weights for the lower and upper excitation states, respectively (see Eq. 1.36 for the special case of hydrogenic ions).

Generalizing the two-level atom/ion jk in excited state i' relative to excited state i , we thus have the ratio of number densities,

$$\frac{n_{k,j,i'}}{n_{k,j,i}} = \frac{g_{k,j,i'}}{g_{k,j,i}} \exp \{ -\chi_{k,j,i'}/kT \}, \quad (11.78)$$

where $\chi_{k,j,i'}^{\text{ex}} = \chi_{k,j,i'} - \chi_{k,j,i}$ difference in the excitation energies of the two states. Since the ground state has $\chi_{k,j,1} = 0$ eV, the ratio of excited stage i to the ground state simplifies to

$$\frac{n_{k,j,i}}{n_{k,j,1}} = \frac{g_{k,j,i}}{g_{k,j,1}} \exp \{ -\chi_{k,j,i}/kT \}. \quad (11.79)$$

In thermodynamic equilibrium, the distribution of excited states will reflect the equilibrium energy partition in proportion to the statistical weights. In this case, the temperature appearing in Eqs. 11.78 and 11.79 will be equal to the kinetic and radiative temperature of the system, $T = T_{\text{R}}$. However, given that gas in the interstellar, circumgalactic, and intergalactic medium is never in thermodynamic equilibrium, one can measure the so-called excitation temperature, T_{ex} . This temperature is specific to the given excitation levels that have been measured,

$$T_{\text{ex}} = \frac{\chi_{k,j,i}^{k,j,i'}}{k} \ln \left(\frac{g_{k,j,i'} n_{k,j,i}}{g_{k,j,i} n_{k,j,i'}} \right). \quad (11.80)$$

The excitation temperature provides an additional window on the degree to which collisional and photo-excitations are in thermodynamic equilibrium.

Using particle density conservation, the Boltzmann equation can be expanded to yield the excitation fraction, $f_{k,j,i} = n_{k,j,i}/n_{k,j}$, which is the ratio of the number density of atoms/ions of species k and ionization stage j in excited state i relative to the number density of atoms/ions of species k and ionization stage j in all excitations states. Invoking particle conservation, $n_{k,j} = \sum_i n_{k,j,i}$, we write

$$f_{k,j,i} = \frac{n_{k,j,i}}{n_{k,j}} = \frac{(n_{k,j,i}/n_{k,j,1})}{\sum_{i=1} (n_{k,j,i}/n_{k,j,1})} = \frac{g_{k,j,i} \exp \{-\chi_{k,j,i}/kT\}}{\sum_{i=1} g_{k,j,i} \exp \{-\chi_{k,j,i}/kT\}}. \quad (11.81)$$

The sum is known as the partition function,

$$U_{k,j}(T) = \sum_{i=1} g_{k,j,i} \exp \{-\chi_{k,j,i}/kT\}. \quad (11.82)$$

Note that the partition function accounts for all excitation levels of the atom/ion; it can be interpreted as the total statistical weight at temperature T . Tables of partition functions are difficult to find(!), but a good, if incomplete source is Gray (1992). In final form, we write the excitation fraction as

$$f_{k,j,i} = \frac{g_{k,j,i}}{U_{k,j}(T)} \exp \{-\chi_{k,j,i}/kT\}. \quad (11.83)$$

11.4.2 Excited States and Electron Density

The circumgalactic and the interstellar medium are highly variable environments in terms of density, temperature, and radiation field. Over a range of these local conditions, there are regions where collisional excitations of the fine-structure excited state can occur whereas the radiation field is usually

insufficient to contribute significantly to the excitation rates. Conversely, for de-excitations, the spontaneous radiative decay of the excited state is on a significantly shorter timescale than the time between ion/electron collisions and/or the timescale of stimulated emission by the radiation field. When these conditions hold, then for excitations, $R_{1,2}^{\text{coll}} \gg R_{1,2}^{\text{rad}}$, but for de-excitations $R_{2,1}^{\text{coll}} \ll R_{2,1}^{\text{rad}}$. Applying this to detailed balancing for a two-level ion expressed in Eq. 11.71, we have

$$\frac{n_2}{n_1} = \frac{R_{1,2}^{\text{coll}}}{R_{2,1}^{\text{rad}}}, \quad (11.84)$$

to a high degree of accuracy.

The rate of spontaneous decay was discussed in § 2.1.2 in terms of the Einstein coefficient, and we thus set $R_{2,1}^{\text{rad}} = A_{2,1}$. As described in Eq. 11.73, the rate of collisional excitations is the product of the number density of free electrons and the collisional excitation rate coefficient,

$$R_{1,2}^{\text{coll}}(T) = n_e \alpha_{1,2}^{\text{ex}}(T). \quad (11.85)$$

As described above, the excitation rate coefficient can be related to the de-excitation rate coefficient via the Boltzmann equation (Eq. 11.77), yielding

$$R_{1,2}^{\text{coll}}(T) = n_e \alpha_{1,2}^{\text{ex}}, = n_e \alpha_{2,1}^{\text{ex}} \frac{g_2}{g_1} \exp\{-\chi_{1,2}/kT\}, \quad (11.86)$$

with

$$\alpha_{2,1}^{\text{ex}}(T) = \langle \sigma_{2,1}^{\text{ex}} \cdot v \rangle = \int_0^\infty \sigma_{2,1}^{\text{ex}}(v) f(v, T) v dv. \quad (11.87)$$

The cross section rapidly decreases with increasing electron velocity. Near the dominating low-velocity threshold for the collisional interaction, the cross section scales as v^{-2} ,

$$\sigma_{2,1}^{\text{ex}}(v) dv = \frac{\pi \hbar^2}{m_e^2} \frac{\Omega_{2,1}(v)}{g_2} \frac{dv}{v^2}, \quad (11.88)$$

where $\Omega_{1,2}(v) = \Omega_{1,2}(v)$ is the collisional strength, a quantity analogous to the line strength (Eq. 3.25) except invoking the wave function of the free electron in the overlap integral (Hebb & Menzel, 1940; Seaton, 1958). Collisional strengths vary slowly with the kinetic energy of the colliding electron, though there are resonant peaks. To a good approximation, the kinetic energy averaged collisional strength can be used

$$\bar{\Omega}_{2,1} = \frac{m_e}{kT} \int_0^\infty \Omega_{2,1}(v) \exp\{-m_e v^2/2kT\} v dv. \quad (11.89)$$

Integrating Eq. 11.87 over all electron velocities to evaluate the collisional de-excitation rate coefficient and substituting into Eq. 11.86, we obtain

$$R_{1,2}^{\text{coll}}(T) = \frac{\hbar^2}{m^{3/2}} \left(\frac{2\pi}{kT} \right)^{1/2} \frac{\bar{\Omega}_{2,1}}{g_1} \exp \{ -\chi_{1,2}/kT \} , \quad (11.90)$$

where the physical constants evaluate to 8.629×10^{-6} . Substituting $R_{1,2}^{\text{coll}}$ and $R_{2,1}^{\text{rad}} = A_{2,1}$ into Eq. 11.84, we have

$$\frac{n_2}{n_1} = n_e \frac{8.629 \times 10^{-6} \bar{\Omega}_{2,1}}{A_{2,1} T^{1/2}} \frac{1}{g_1} \exp \{ -\chi_{1,2}/kT \} . \quad (11.91)$$

Following the many applications in the literature (e.g., Spitzer & Fitzpatrick, 1993; Wood & Linsky, 1997; Oliveira et al., 2003), Eq. 11.91 can be used to estimate the electron density as a function of temperature directly from the ratio of the measured column densities, $N_2/N_1 = n_2/n_1$. For example, for the $J = 3/2$ C II* $\lambda 1037$ and the $J = 1/2$ C II $\lambda 1036$ transitions, the statistical weight of the ground state is $g_1 = 4$, the rate for spontaneous radiative decay is $A_{2,1} = 2.29 \times 10^{-6} \text{ s}^{-1}$, the excitation energy is $\chi_{1,2} = 0.00786 \text{ eV}$ (Kramida et al., 2014), and the collisional strength is $\bar{\Omega}_{1,2} = \bar{\Omega}_{2,1} = 2.90$ (Osterbrock & Ferland, 2006).

11.5 Ionization

In the case that the radiation field is not thermodynamically coupled to the particle field, a scenario common to the intergalactic, circumgalactic, and much of the interstellar medium, then a detailed treatment must be undertaken to estimate the ionization rates of the atoms/ions in the gas. Accurate estimates requires a good general knowledge of the radiation field, specifically the mean intensity, J_ν , and the temperature of the particle field, particularly the electron pool. Accurate estimates also require knowledge of the atomic cross sections for interaction with the radiation field and for collisional processes. In addition to discussing radiative and collisional ionization, we discuss three additional ionization processes that are important in the intergalactic medium and the circumgalactic medium. These are Auger ionization, excitation auto-ionization, and charge exchange ionization.

Unlike ionization, which requires knowledge of the radiation field, the physics of dominant channels of recombination require knowledge of the particle field. Again, accurate estimates require a good general knowledge of the temperature of the particle field, particularly the electron pool, and the atomic cross sections for interaction with collisional processes with the

electron pool. In addition to discussing radiative recombination, we also discuss dielectronic recombination and charge exchange recombination.

11.5.1 Photo and Auger Ionization

Photoionization and Auger ionization both begin with radiative ionization processes in which a bound electron is photo-ejected from an ion or neutral atom. The electron can be liberated from any one of the ion's populated electron shells.

To compute the rate, $R_{k,j,s}^{\text{ph}}$, at which an electron in a given shell, index s , is photo-ejected from ion k, j , we multiply the cross section for absorption, $\sigma_{k,j,s}^{\text{ph}}(E)$, for shell s at energy $E = h\nu$ by the photon number density per unit energy, $4\pi J_E/E$ [$\text{erg}^{-1} \text{ cm}^{-2} \text{ s}^{-1}$], and integrate over all energies³ greater than the binding energy of the electron,

$$R_{k,j,s}^{\text{ph}} = 4\pi \int_{I_{k,j,s}}^{\infty} J_E \sigma_{k,j,s}^{\text{ph}}(E) \frac{dE}{E}, \quad (11.92)$$

where $I_{k,j,s}$ is the ionization (binding) energy for electrons in shell s of ion k, j .

Photoionization cross sections can be computed from the fitting functions and fitting parameters tabulated by Verner & Yakovlev (1995) for inner shells and by Verner et al. (1996) for the outer shells⁴. In Figure 11.5, we present the photoionization cross sections for H^0 , and Mg^+ , and C^{+3} , and O^{+5} as a function of the photon energy as applied in Eq. 11.92. Ground-state Mg^+ has the isoelectronic sequence of neutral sodium ($1s^2 2s^2 2p^6 3s^1$) and ground-state C^{+3} , and O^{+5} have the isoelectronic sequence of neutral lithium ($1s^2 2s^1$). Note that the lowest energy thresholds correspond to the outermost populated electron shell; these are the ground-state ionization potentials listed in Table 4.1 and illustrated in Fig 4.5.

In the case of photoionization, a single electron, e_{ej}^- , is ejected and the ionization stage of the incident ion k, j is incremented by one to $k, j + 1$,

$$A_{k,j} + \gamma \rightarrow A_{k,j+1} + e_{\text{ej}}^-. \quad (11.93)$$

In cases where the incident photon has the required energy to liberate an inner shell electron some of its energy can be channeled into also liberating one or more of the less bound, higher shell electrons. This process is known

³ It is convenient to perform the integrals as a function of energy, whereas the mean intensity is often expressed as J_ν . The conversion is $J_E dE = J_\nu d\nu$, where $dE/d\nu = h$.

⁴ D. Verner provides publicly available electronic tables for photoionization, recombination, collisional ionization, auto-ionization, and some supporting computational subroutines at www.pa.uky.edu/~verner/.

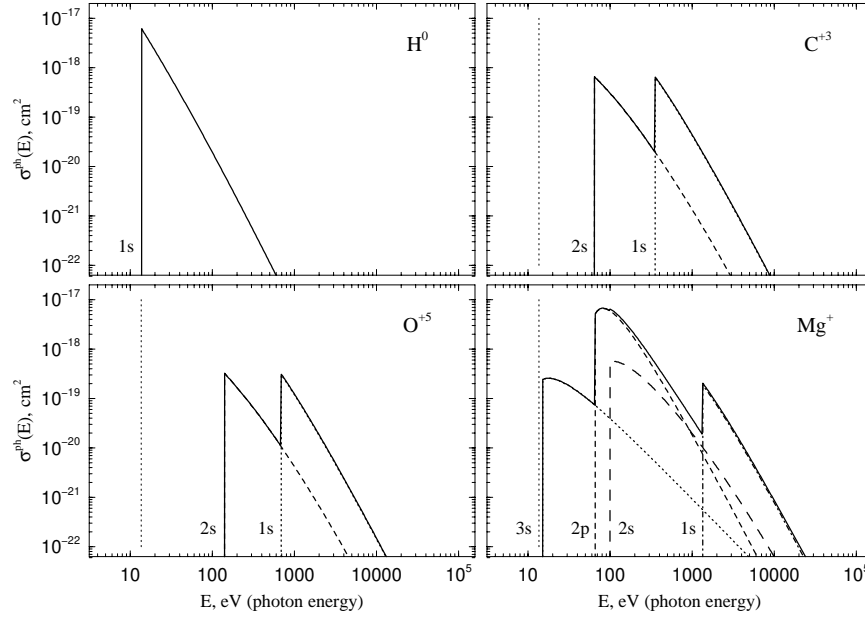


Figure 11.5 The partial and total photoionization cross sections [cm^2] for H^0 , Mg^+ , C^{+3} , and O^{+5} as a function of the photon energy [eV]. The partial cross sections for each electron shell are shown as dotted, dashed and dot-dashed curves as labeled 1s, 2s, etc. The total cross sections are given by the solid curves. For reference, the vertical dotted line at $E = 13.598$ eV is the ground-state ionization energy for H^0 .

as Auger ionization, in which the ionization stage of the incident ion k, j is incremented by two or more,

$$A_{k,j} + \gamma \rightarrow A_{k,m} + (m - j) e_{\text{ej}}^-, \quad (11.94)$$

where we use the convention that the initial ionization stage is j and the final higher ionization stage is m . Note that the number of ejected electrons is $N_e = m - j$. Because the photo-electron is included in the notation, the final stage m is always greater than or equal to $j + 2$. Photoionization is the special case in which $m = j + 1$.

In order to compute the photo and Auger ionization rates, it is necessary to know the yield probability, $P_{k,j,m-j}^s$, i.e., the probability that $N_e = m - j$ electrons in total are ejected from an ion following a photoionization of an electron originating from shell s (the photo-electron). These yield probabilities have been calculated and tabulated by Kaastra & Mewe (1993) for each electron shell.

For photoionization, the total photoionization rate, $R_{k,j}^{\text{ph}}$, for destruction

of ion k, j is given by $R_{k,j,s}^{\text{ph}}$ (see Eq. 11.92), the rate at which an electron bound in shell s of ion k, j is ionized by incident photons, weighted by the probability that only the photo-electron is ejected from the ion and summed over all electron shells,

$$R_{k,j}^{\text{ph}} = \sum_{s=1}^{N_{k,j}^s} P_{k,j,1}^s R_{k,j,s}^{\text{ph}}, \quad (11.95)$$

where $N_{k,j}^s$ is the number of shells for ion k, j .

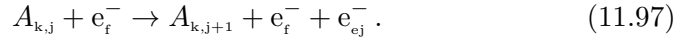
Similarly, the Auger ionization rate, $R_{k,j,m}^{\text{aug}}$, for destruction of an ion k, j that ejects $N_e = m - j$ electrons (including the photo-electron, so $N_e \geq 2$) is the sum of $R_{k,j,s}^{\text{ph}}$ over all electron shells weighted by the probability that N_e electrons in total were ejected from the ion in response to a photo-electron originating in shell s ,

$$R_{k,j,m}^{\text{aug}} = \sum_{s=1}^{N_{k,j}^s} P_{k,j,m-j}^s R_{k,j,s}^{\text{ph}}. \quad (11.96)$$

The $R_{k,j,m}^{\text{aug}}$ are unique amongst the various ionization rates, because they dictate the balance between non-adjacent ionization stages of species k . Clearly, $R_{k,k+1,m}^{\text{aug}} = R_{k,k,m}^{\text{aug}} = 0$ since fully ionized and hydrogenic ions cannot undergo Auger ionization.

11.5.2 Direct Collisional Ionization

Direct collisional ionization is the collision of an electron with an ion, which then directly ionizes from j to $j + 1$,



In the process, the colliding free electron, e_f^- , loses an energy equal to the ionization energy plus the kinetic energy of the ejected electron.

The ionization rate for destruction of ion k, j due to direct collisional ionization is obtained by multiplying the total direct collisional ionization rate coefficient, $\alpha_{k,j}^{\text{cdi}}(T)$, by the number density of free electrons,

$$R_{k,j}^{\text{cdi}}(T) = n_e \alpha_{k,j}^{\text{cdi}}(T), \quad (11.98)$$

where

$$\alpha_{k,j}^{\text{cdi}}(T) = \sum_{s=1}^{N_{k,j}^s} \alpha_{k,j,s}^{\text{cdi}}(T). \quad (11.99)$$

is the sum of the direct collisional ionization rate coefficient contributions, $\alpha_{k,j,s}^{\text{cdi}}(T)$, for ejection of an electron from shell s .

The $\alpha_{k,j,s}^{\text{cdi}}(T)$ are the expectation values $\langle \sigma_{k,j,s}^{\text{cdi}} \cdot v \rangle$, where $\sigma_{k,j,s}^{\text{cdi}}(v)$ is the

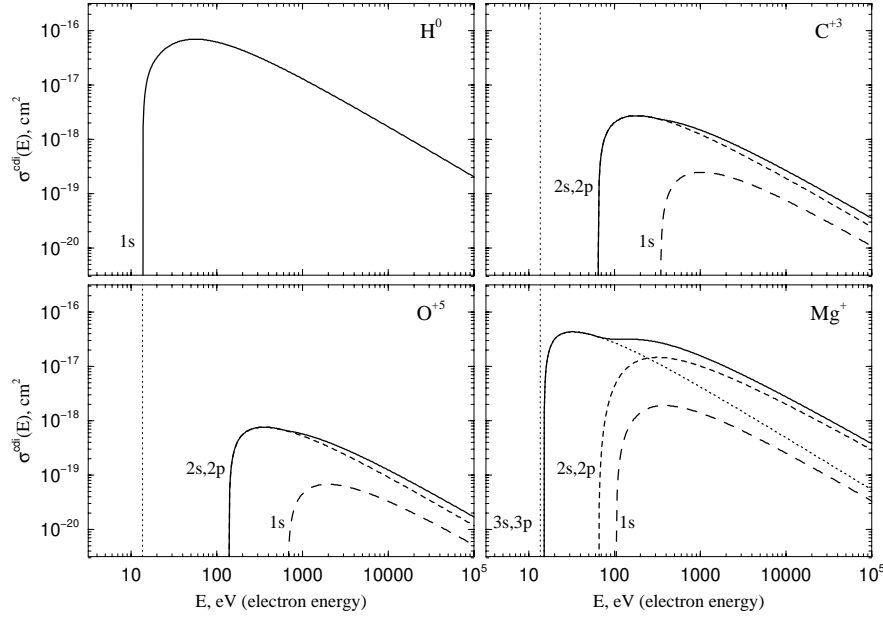


Figure 11.6 The direct collisional ionization cross sections [cm^2] for H^0 , Mg^+ , C^{+3} , and O^{+5} as a function of the electron energy [eV]. The partial cross sections for each electron shell are shown as dotted, dashed and dot-dashed curves as labeled 1s, 2s, etc. The total cross sections are given by the solid curves. For reference, the vertical dotted line at $E = 13.598$ eV is the ground-state ionization energy for H^0 .

direct collisional ionization cross section for the shell, and $v = \sqrt{2E/m_e}$ is the electron speed for kinetic energy E . The integration is over energies large enough to overcome the binding energy,

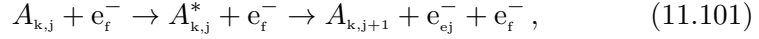
$$\alpha_{k,j,s}^{\text{cdi}}(T) = \langle \sigma_{k,j,s}^{\text{cdi}} \cdot v \rangle = \int_{v_{k,j,s}}^{\infty} \sigma_{k,j,s}^{\text{cdi}}(v) f(v, T) v dv, \quad (11.100)$$

where $f(v, T)$ is the Maxwell-Boltzmann electron velocity distribution function at equilibrium temperature T , and $v_{k,j,s} = \sqrt{2I_{k,j,s}/m_e}$ is the velocity corresponding to the ionization energy of shell s of ion k, j . The direct collisional ionization cross sections and rate coefficients can be computed from the fitting functions and parameters tabulated by Arnaud & Rothenflug (1985).

In Figure 11.6, we present the direct collisional ionization cross sections for H^0 , and Mg^+ , and C^{+3} , and O^{+5} as a function of the electron energy. The individual shell cross sections are shown as dotted and dashed curves. The total cross section is shown as the solid curve.

11.5.3 Excitation Auto-Ionization

Excitation auto-ionization (E-A) occurs in ions with many inner filled shell electrons and only a few outer shell electrons. A collision with a free electron first excites the ion. Then, during the internal de-excitation process the released energy can either channel into recombination emission lines or into liberating an outer shell electron, which is auto-ionization,



where e_f^- the free collisional electron and the e_{ej}^- is ejected auto-ionized electron.

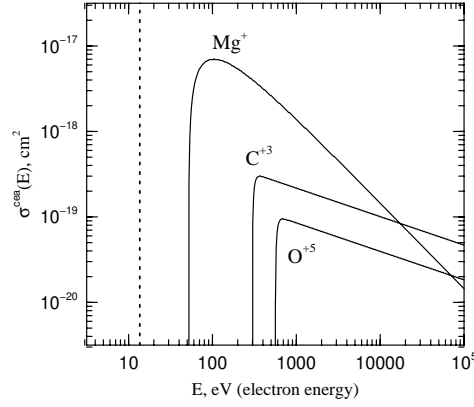


Figure 11.7 The excitation auto-ionization collisional cross sections [cm^2] for Mg^+ , C^{+3} , and O^{+5} as a function of the electron energy [eV]. For H^0 , the cross section is null. For reference, the vertical dotted line is the ground-state ionization energy for H^0 .

The ionization rate for destruction of ion k, j due to E-A collisional ionization is obtained by multiplying the E-A collisional ionization rate coefficient, $\alpha_{k,j}^{\text{cea}}(T)$, by the number density of free electrons,

$$R_{k,j}^{\text{cea}}(T) = n_e \alpha_{k,j}^{\text{cea}}(T). \quad (11.102)$$

The total E-A rate coefficient is given by the expectation value $\alpha_{k,j}^{\text{cea}}(T) = \langle \sigma_{k,j}^{\text{cea}} \cdot v \rangle$ computed from Eq. 11.100 with $I_{k,j,s}$ replaced by the E-A onset energy $\chi_{k,j}$, and $\sigma_{k,j,s}^{\text{cdi}}(E)$ replaced by the total E-A cross section $\sigma_{k,j}^{\text{cea}}(E)$. In Figure 11.7, we show the E-A cross sections for Mg^+ , C^{+3} , and O^{+5} .

The E-A cross sections and rate coefficients depend on the bound electron configuration of the ion, i.e., the isoelectronic sequence. For example, C^{+3} , N^{+4} , and O^{+5} all have the electron configuration of neutral lithium ($1s^2 2s$) and are thus lithium iso-sequence ions ($N_e = 3$). There is no E-A process for

hydrogen and helium sequence ions. For lithium sequence ions, the dominant contribution to the cross section is the 1s–2p transition. As the charge of the ion, Z , increases, the branching ratio to E-A decreases. No significant E-A contribution to the direct collisional cross section is observed for the beryllium sequence (except perhaps O^{+4}), nor for the sequences from boron to neon, which differ only in the number of 2p shell electrons. For the sodium sequence ($[\text{Ne}]3s^1$), up to 18 transitions can contribute to E-A, for which the relative importance increases with Z . The sequences from magnesium to argon differ in the number of 3p shell electrons, and the relative importance of E-A decreases as the shell fills.

Fitting functions and their parameters for the total E-A cross sections and rate coefficients for ions up to nickel have been tabulated by Arnaud & Rothenflug (1985), with the exceptions of iron ions, which can be obtained from the fitting functions and parameters presented by Arnaud & Raymond (1992), and the cross sections for the lithium sequence, which can be obtained from Hu et al. (1996). A summary and simplification of those works has been presented by Churchill et al. (2014).

11.5.4 Radiative Recombination

Radiative recombination is the capture of a free electron by ion $k, j + 1$ followed by the emission of a photon,

$$A_{k,j+1} + e_f^- \rightarrow A_{k,j} + \gamma. \quad (11.103)$$

The radiative recombination rate for creation of ion k, j due to electron recombination with ion $k, j + 1$ is obtained by multiplying the total recombination rate coefficient, $\beta_{k,j}^{\text{phr}}(T)$, by the electron number density,

$$R_{k,j+1}^{\text{phr}}(T) = n_e \beta_{k,j}^{\text{phr}}(T). \quad (11.104)$$

Note we have adopted the convention that the rate coefficients, $\beta_{k,j}$, are indexed to the final state j , whereas all rates, $R_{k,j}$, are indexed by the initial state j .

The cross section for capture of a free electron decreases with electron kinetic energy. Given the cross section, $\sigma_{k,j,s}^{\text{phr}}(E)$, for radiative recombination to shell s forming ion k, j , the radiative recombination rate coefficient for the shell, $\beta_{k,j,s}^{\text{phr}}(T)$, is obtained by integrating over all electron velocities, analogous to Eq. 11.100, i.e., with no threshold energy ($I_{k,j,s} \rightarrow 0$) and with $\sigma_{k,j,s}^{\text{cdi}}(v)$ replaced by $\sigma_{k,j,s}^{\text{phr}}(v)$. Radiation induced recombination, which, is a three-body process, is not a dominant recombination channel. The total

recombination rate coefficient is the sum over all shells,

$$\beta_{k,j}^{\text{phr}}(T) = \sum_{s=1}^{N_{k,j}^s} \beta_{k,j,s}^{\text{phr}}(T). \quad (11.105)$$

For hydrogenic ions ($N_e = 1$) the formula originally proposed by Seaton (1959), which is highly accurate (Arnaud & Rothenflug, 1985; Dopita & Sutherland, 2003),

$$\beta_{k,j}^{\text{phr}}(T) = \beta_0 Z_k \lambda^{1/2} \left[0.4288 + 0.5 \ln \lambda + 0.469 \lambda^{-1/3} \right], \quad (11.106)$$

where $\beta_0 = 5.197 \times 10^{-14}$ and where $\lambda = Z_k^2 (1.5789 \times 10^5 / T)$.

The Atomic and Molecular Diagnostic Processes in Plasmas (AMDPP) group has published fitting functions and parameters for radiative recombination rate coefficients for many non-hydrogenic ions (Badnell, 2006). Their fitting functions neglect narrow resonant spikes.

Since not all ionization stages have been tabulated by Badnell (2006), one can employ the fitting functions and parameters for the simple power-law form published by Arnaud & Rothenflug (1985), based upon work of Seaton (1959), Aldrovandi & Pequignot (1973), and Shull & Van Steenberg (1982) for cases omitted from the AMDPP tables. The expression is

$$\beta_{k,j}^{\text{phr}}(T) = a_{k,j} (T/10^4)^{-b_{k,j}}, \quad (11.107)$$

where $a_{k,j}$ and $b_{k,j}$ are the fitting parameters tabulated by Arnaud & Rothenflug (1985) for all ions of helium through nickel.

11.5.5 Dielectronic Recombination

Dielectronic recombination often dominates over radiative recombination. In this process, a high energy free electron first excites a bound deep inner shell electron prior to its capture in an elevated excited state of the ion. There are now two excited electrons and an unfilled state in an inner shell. Multiple channels of relaxation for the ion are now available (of which one is also auto-ionization).

In dielectronic recombination, the doubly excited ion works its way back to the ground state via multiple radiative cascades. At high temperatures this process usually proceeds first by the decay of one of the excited electrons to refill the empty inner shell by radiative decay followed by a downward cascade of the remaining excited electron. At low temperatures, the dominant channel occurs when the free electron is capture in a shell that is a resonant state to the emptied inner shell. The electron transitions rapidly and

is then followed by a downward cascade of the remaining excited electron. The reaction can be written

$$A_{k,j+1} + e_f^- \rightarrow A_{k,j}^{**} \rightarrow A_{k,j}^* + \gamma \rightarrow A_{k,j+1} + \sum \gamma_i, \quad (11.108)$$

where the sum indicates that several recombination photons can be emitted during the cascade process.

The dielectronic recombination rate for creation of ion k, j due to electron recombination with ion $k, j + 1$ is obtained by multiplying the total dielectronic recombination rate coefficient, $\beta_{k,j}^{\text{die}}(T)$, by the electron number density,

$$R_{k,j+1}^{\text{die}}(T) = n_e \beta_{k,j}^{\text{die}}(T). \quad (11.109)$$

Since the dominant channels for dielectronic recombination are temperature dependent, the rate coefficient is double peaked. For this reason, previous fitting functions and parameters for the rate coefficients were split into a low temperature regime (Nussbaumer & Storey, 1983, 1986, 1987) and high temperature regime (Aldrovandi & Pequignot, 1973; Shull & Van Steenberg, 1982; Arnaud & Rothenflug, 1985).

A newer fitting function and accompanying parameter list for all elements from helium to zinc and valid for temperatures ranging from $T \simeq 100$ to $T \simeq 10^7 \text{K}$ has been made available by the AMDPP group⁵. These fitting functions and parameters are described in Altun et al. (2007) and are based on a series of papers by the group (see references in Badnell et al., 2003; Altun et al., 2007).

11.5.6 Charge Exchange

Charge exchange is the transfer of an electron from one ion to another during a collision. Since hydrogen is the most abundant species, a charge exchange with a given metal ion k, j is dominated either by ionization ($k, j \rightarrow k, j + 1$) from an ionized hydrogen (in which the H^+ ion recombines with the exchanged electron), or by recombination ($k, j - 1 \rightarrow k, j$) via the ionization of neutral hydrogen,

$$A_{k,j} + \text{H}^+ \leftrightarrow A_{k,j+1} + \text{H}^0. \quad (11.110)$$

Helium is also relatively abundant and is the second most important charge exchange channel,

$$A_{k,j+1} + \text{He}^0 \leftrightarrow A_{k,j} + \text{He}^+. \quad (11.111)$$

⁵ N. Badnell, R. Bingham, G. Duxbury, and H. Summers of the Atomic and Molecular Diagnostic Processes in Plasmas group provide data for radiative and dielectronic recombination rates at <http://amdpp.phys.strath.ac.uk/tamoc/>.

The rate for destruction of ion k, j via charge exchange ionization with ionized hydrogen ($k = 1, j = 2$) is

$$R_{k,j}^{\text{xH}^+}(T) = n_{1,2} \alpha_{k,j}^{\text{xH}^+}(T), \quad (11.112)$$

where $\alpha_{k,j}^{\text{xH}^+}(T)$ is the ionization rate coefficient. The rates for creation of ion $k, j - 1$ via destruction of ion k, j via charge exchange recombination with neutral hydrogen ($k = 1, j = 1$) and with neutral helium ($k = 2, j = 1$) are given by

$$\begin{aligned} R_{k,j}^{\text{xH}}(T) &= n_{1,1} \beta_{k,j-1}^{\text{xH}}(T) \\ R_{k,j}^{\text{xHe}}(T) &= n_{2,1} \beta_{k,j-1}^{\text{xHe}}(T), \end{aligned} \quad (11.113)$$

where $\beta_{k,j-1}^{\text{xH}}(T)$ and $\beta_{k,j-1}^{\text{xHe}}(T)$ are the respective recombination rate coefficients.

The total charge exchange ionization and recombination rate coefficients can be computed using the fitting function and parameters of Kingdon & Ferland (1996)⁶. The rate coefficient for ionization of ion k, j via charge exchange to neutral hydrogen is obtained via detailed balancing,

$$\alpha_{k,j}^{\text{xH}^+}(T) = \beta_{k,j}^{\text{xH}}(T) \exp \{ -\Delta E_{k,j} / kT_4 \}, \quad (11.114)$$

where the Boltzmann factor, $\Delta E_{k,j}/k$, is also tabulated by Kingdon & Ferland (1996).

The computation of the recombination rate coefficient for charge exchange with neutral helium, $\beta_{k,j}^{\text{xHe}}(T)$, is also obtained from the fitting functions and parameters from Kingdon & Ferland (1996).

11.5.7 Astrophysical Rates and Rate Coefficients

Having reviewed the nature of the radiation field, the formalism describing the particle field, detailed balancing, and the various processes that govern the excitation and ionization balance of atoms and ions in astrophysical gaseous environments, we now proceed to discuss the typical range of ionization and recombination rates and rate coefficients.

The final ionization balance and densities of all ions and of free electrons depends on the boundary conditions of the gas cloud, i.e., the atomic abundance or mass fractions, and the hydrodynamics via the equation of state, given by Eqs. 11.63 and 11.64,

$$n = n_N + n_e = \frac{P_g}{kT}, \quad \rho = \frac{\mu m_a}{k} \frac{P_g}{T}, \quad (11.115)$$

⁶ P. Stancil, D. Schultz, J. Wang, M. Raković, J. Kingdon, and A. Dalgarno, provide data at the Oakridge National Lab UGA Charge Transfer Database for Astrophysics, www-cfadc.phy.ornl.gov/astro/ps/data/.

where μ is the mean molecular weight of the gas (Eq. 11.57). The ionization balance is quantified in terms of the ionization fractions, providing the ratio $f_{k,j} = n_{k,j}/n_k$ for all atomic species. Through charge density conservation, n_e is determined, and then all particles densities can be solved via Eq. 11.29.

The ionization fractions are governed by the photo and collisional ionization rates and the recombination rates. Whereas the photoionization rates are density independent and depend only on the spectral energy density of the ionizing radiation field, the collisional ionization and recombination rates depend on the ion and electron densities, and therefore on the hydrodynamic equation of state of the gas.

For photoionization by the cosmic UVB, we can compute the photoionization rate, $R_{k,j}^{\text{ph}}$, for species k, j at any redshift independent of the equation of state. In Figure 11.8(a), we show the redshift evolution of cosmic UVB, $J_\nu(I_{k,j}, z)$, the mean intensity at the ionization energy, $I_{k,j}$, for the ground states of neutral hydrogen, H^0 , single ionized helium, He^+ , and the commonly observed Group IA ions Mg^+ , C^{+3} , and O^{+5} . We adopt the commonly employed notation for the mean intensity at the ionization energy of neutral hydrogen as $J_{\nu_0} = J_\nu(I_{\text{H}^0})$, where $h\nu_0 = I_{\text{H}^0} = 13.598$ eV. The ionization energy for He^+ is $I_{\text{He}^+} = 54.418$ eV. Inspection of Figure 11.1 shows that these ionization energies mark the thresholds for significant absorption of photons with $h\nu > I_{\text{H}^0}$ and $h\nu > I_{\text{He}^+}$ due to the ionization of hydrogen and helium.

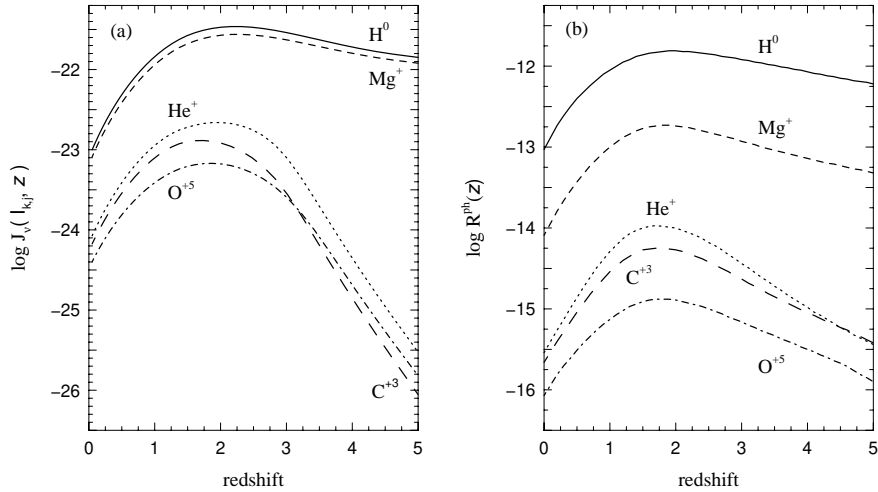


Figure 11.8 (a) The cosmic UVB mean intensity at $I_{k,j}$, the ground-state ionization energy, for H^0 , He^+ , Mg^+ , C^{+3} , and O^{+5} versus redshift. (b) The photoionization rates (Eq. 11.92) of H^0 , Mg^+ , C^{+3} , and O^{+5} versus redshift. The cosmic UVB is taken from Haardt & Madau (2012).

As we describe in more detail in § 12.3.3, these thresholds, or ionization edges, reduce the number density of ionizing photons available to ionize metal ions with ionization energies $I_{k,j} > I_{H^0}$ and $I_{k,j} > I_{He^+}$. Since $I_{Mg^+} = 15.035$, the resulting ionization fraction of Mg^+ is highly dependent on the attenuation of J_ν due to H^0 ionization. As seen in Figure 11.8(a), the ratio $J_\nu(I_{Mg^+})/J_{\nu_0}$ does not strongly evolve with redshift, indicating that as the ionization rate of H^0 evolves, the ionization of Mg^+ will evolve similarly. As illustrated in Figure 11.8(b), the photoionization rate of Mg^+ evolves similarly to that of H^0 , peaking at $z \simeq 1.9$.

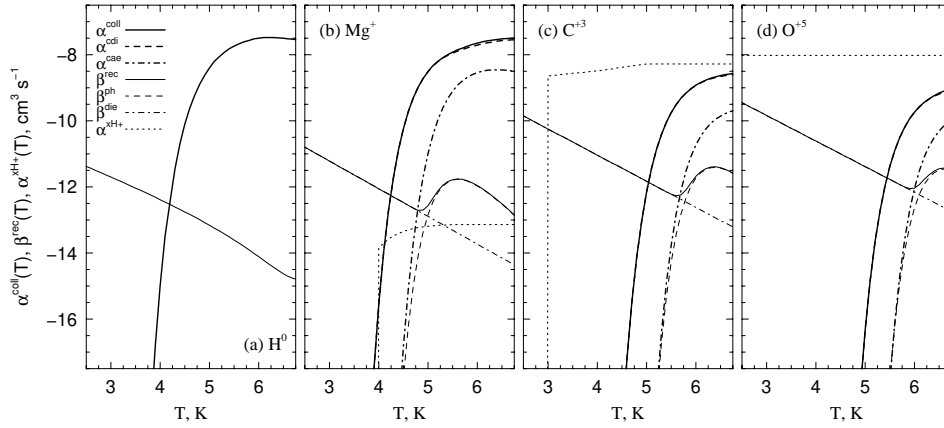


Figure 11.9 The collisional ionization, recombination, and hydrogen ionization charge exchange rate coefficients versus temperature for H^0 , Mg^+ , C^{+3} , and O^{+5} . The actual ionization and recombination rates can be obtained by multiplying the rate coefficients by the electron density, n_e (except for the charge exchange, where one multiplies by the density of H^+). The total collisional ionization rate coefficient ($\alpha_{k,j}^{coll}$, thick solid curve) is the sum of the direct collisional process ($\alpha_{k,j}^{cdi}$, thick dashed curve) and the excitation auto-ionization process ($\alpha_{k,j}^{cea}$, thick dot-dash curve). The total recombination rate coefficient ($\beta_{k,j-1}^{rec}$, thin solid curve) is the sum of the radiative recombination process ($\beta_{k,j-1}^{ph}$, thin dashed curve) and the dielectronic recombination process ($\beta_{k,j-1}^{die}$, thin dot-dash curve). The charge exchange process ($\alpha_{k,j}^{xH^+}$, dotted curve) is ionization of the metal ion by H^+ .

As can be seen in Figure 11.1, absorption due to ionization of He^+ dominates the spectral energy distribution of the cosmic UVB for energies above $I_{He^+} = 54.418$ eV. The ionization energies of both C^{+3} and O^{+5} lie in this range, with $I_{C^{+3}} = 64.494$ and $I_{O^{+5}} = 138.120$ eV. For these ions, the mean intensity at their ionization energies evolves rapidly for $z > 2$ as compared to the evolution of J_{ν_0} ; this is due to He^+ ionization in the intergalactic medium at higher redshifts. We thus expect that the ionization rates of

C^{+3} and O^{+5} will evolve in step with the ionization rates of He^+ , which is illustrated in Figure 11.8(b).

A further central insight of Figure 11.8 is that there is strong evolution in $J_\nu(I_{k,j}, z)$ for $z < 1.5$ in that the intensity of the cosmic UVB background at the ionization energies of commonly observed ions declines by an order of magnitude from $z \simeq 1.5$ to $z = 0$, and this is reflected in the photoionization rates $R_{k,j}^{ph}(z)$.

Though the collisional ionization and recombination rates depend on both gas density and temperature, the rate coefficients depend only on temperature. We can think of the collisional ionization and recombination rate coefficients as the rates normalized by the electron density, i.e., $\alpha_{k,j}^{coll} = R_{k,j}^{coll}/n_e$. Similarly, for charge exchange we can think of the rate coefficient as the rate of charge exchange per unit hydrogen atom.

In Figure 11.9, we show the total collisional ionization coefficient, $\alpha_{k,j}^{coll}$, which is the sum of the direct collisional process, $\alpha_{k,j}^{cdi}$, thick dashed curve and the excitation auto-ionization process $\alpha_{k,j}^{cea}$. We also show the total recombination rate coefficient, $\beta_{k,j-1}^{rec}$, which is the sum of the radiative recombination process, $\beta_{k,j-1}^{ph}$, and the dielectronic recombination process, $\beta_{k,j-1}^{die}$. The charge exchange process, $\alpha_{k,j}^{xH^+}$, is ionization of the metal ion by the H^+ , which recombines with the electron. We have focused on the centrally important H^0 atom, and the commonly observed Group IA ions Mg^+ , C^{+3} , and O^{+5} .

Note that as temperature increases, the rates of direct collisional (thick solid curves) ionization dominate over the rates of recombination (thin solid curves). This tends to results in higher ionization conditions. Though excitation auto-ionization (thick dot-dash curves) is not inconsequential, for these ions it makes a non-discernible contribution to the collisional ionization rate for $T < 10^{5.5}$ K, above which it make roughly a 10% contribution. Charge exchange ionization of Mg^+ by H^+ is fairly negligible, but note that it is a dominant source of ionization even to low temperatures for both C^{+3} and O^{+5} (charge exchange recombination in which the electron doner is H^0 is negligible for all three of these ions).

11.6 The Cooling Function

The cooling and heating rates of interstellar, circumgalactic, and intergalactic gas are central for understanding the thermodynamics of cosmological gas. For example, the cooling rate governs how fast a gas will evolve thermodynamically, and thus provide insight into the dynamic stability of the gas structure and whether the excitation, ionization, and recombination processes of energy exchange between the particle field and the radiation field

can settle into equilibrium. If a gas structure cools or heats more rapidly than the timescale for ionization or recombination, the gas may be “over-ionized” or “under-ionized” relative to the ionization conditions that would be present if the gas temperature were stable because these physical processes cannot “catch-up” to the current thermodynamic conditions.

Though the cooling and heating functions of cosmological gas has been addressed and/or studied by many researchers (e.g., Krolik et al., 1981; Sutherland & Dopita, 1993; Dopita & Sutherland, 2003; Ferland, 2003; Osterbrock & Ferland, 2006; Gnat & Sternberg, 2007; Wiersma et al., 2009; Gnedin & Hollon, 2012), we provide a introductory synthesis of this complicated, yet very important subject. In what follows, the term “cooling” will reflect energy lost by the particle field; often this energy is transferred directly to the radiation field (and may escape the gas structure). The term “heating” will mean the converse, that energy from the radiation field is transferred to the particle field. We do not treat cooling or heating due to dynamical work done by or on the gas, as that is related through the differential equation of state (Eq. 11.66).

We note that many past works discussing the cooling and heating of cosmological gas have applied the assumption of collisional ionization equilibrium. However, this assumption rarely applies to the interstellar medium and never applies to the intergalactic medium. That is, in the vast majority of cosmological environments, a substantive radiation field is present and it has become appreciated that this introduces non-negligible modification to the collisional equilibrium cooling and heating functions (e.g., Wiersma et al., 2009; Gnedin & Hollon, 2012). In addition to the collisional processes of excitation and de-excitation, ionization, recombination, free-free continuum scattering, the radiative processes of photoionization and Compton scattering should be included.

The cooling rate is highly specific to the physical state of the gas, meaning the intensity and spectral energy distribution of the incident radiation field, the gas density and temperature, and the chemical composition, i.e.,

$$\frac{dQ}{dt} = \frac{dQ(J_\nu, T, n_{\text{H}}, n_{\text{He}}/n_{\text{H}}, Z)}{dt}, \quad (11.116)$$

where Z is shorthand for the abundances of all metal species, $n_{\text{k}}/n_{\text{H}}$. Cooling rates are expressed as the energy exchange per unit volume per second, i.e., $\text{erg cm}^{-3} \text{ s}^{-1}$.

As mentioned above, the total cooling rate is the sum of various cooling processes. If we denote the cooling due to process i as $dQ^{(i)}/dt$, the cooling

rate can be written

$$\frac{dQ}{dt} = \sum_i \frac{dQ^{(i)}}{dt}, \quad (11.117)$$

where each process is the sum over all atoms and ions in the gas,

$$\frac{dQ^{(i)}}{dt} = \sum_{k,j} \frac{dQ_{k,j}^{(i)}}{dt}. \quad (11.118)$$

For collisional processes, we can further describe the cooling rate in terms of the so-called cooling function, $\Lambda_{k,j}^{(i)}$ [erg cm³ s⁻¹],

$$\frac{dQ_{k,j}^{(i)}}{dt} = n_{k,j} n_e \Lambda_{k,j}^{(i)} = f_{k,j} n_k n_e \Lambda_{k,j}^{(i)}, \quad (11.119)$$

where $n_e \Lambda_{k,j}^{(i)}$ [erg s⁻¹] is essentially the cooling power of process i per atom/ion of species k, j . The cooling function is directly comparable to the rate coefficient except that it quantifies the energy exchange rate. Analogous to Eqs. 11.99 and 11.100, we have $\Lambda_{k,j}^{(i)} = \sum_s \langle \Delta E \cdot \sigma_{k,j,s}^{(i)} \cdot v \rangle$, where the sum is over the electron shells of ion k, j . In general, for collisional process of type i ,

$$\Lambda_{k,j}^{(i)}(T) = \sum_{s=1}^{N_{k,j}^s} \int_{v_{k,j,s}}^{\infty} \sigma_{k,j,s}^{(i)}(v) f(v, T) \Delta E(v) v dv, \quad (11.120)$$

where $f(v, T)$ is the Maxwell-Boltzmann distribution function of electron velocities thermalized at temperature T , $\sigma_{k,j,s}^{(i)}(v)$ is the interaction cross section as a function of electron velocity, and $v_{k,j,s} = \sqrt{2I_{k,j,s}/m_e}$ is the velocity at which the energy equals the ionization energy for shell s of ion k, j . The benefit of this formalism is that one quasi-separates variables in that the cooling function is strictly temperature dependent; however, note that the ionization fraction, $f_{k,j}$, multiplying the cooling function in the cooling rate (Eq. 11.119) also depends on temperature and the radiation field spectral energy density, J_ν .

In the case of photoionization, the gas is heated due to the adding of photoionized electrons to the particle field. The heating rate for ion j, k is

$$\frac{dQ_{k,j}^{(i)}}{dt} = n_{k,j} \Upsilon_{k,j}^{\text{ph}} = f_{k,j} n_k \Upsilon_{k,j}^{\text{ph}}, \quad (11.121)$$

where $\Upsilon_{k,j}^{\text{ph}}$ is the photo-heating power in erg s⁻¹, which is derived in § 11.6.1.

Obtaining the total cooling rate of a gas requires knowing the equilibrium ionization condition, the chemical abundances, and the ΔE energy exchanged per process. In the presence of photoionization, the gas gets heated

at low temperatures and cooled at high temperatures. There is thus a value of $T = T_{\text{eq}}$ at which

$$\frac{dQ(T_{\text{eq}})}{dt} = 0. \quad (11.122)$$

If dynamical processes such as shock heating, compression, or expansion are negligible, the gas will eventually settle to this equilibrium temperature. It is then said to be in thermal equilibrium. The equilibrium temperature and therefore ionization conditions and electron density, all depend on the heating and cooling balancing.

In our treatment, the gas is assumed to be static, so dynamical cooling such as expansion and advection cooling are not treated. We also do not treat cooling due to molecules as molecular cooling makes negligible contributes above $T \simeq 10^4\text{K}$, a temperature well below the typical values found in the circumgalactic and intergalactic medium. Cooling due to conduction, which occurs across temperature gradients, is also not included, as we assume an isothermal medium.

In Chapter 12, we will discuss ionization modeling of the gas. Ionization modeling is the keystone of interpreting cosmological absorption line spectra as the absorption features are due to ions of various chemical species. Except for special cases, the astronomical context of the data simply cannot be interpreted without ionization modeling. Since the data represent ions of various ionization levels, and ionization fractions are temperature dependent, estimating the equilibrium temperature of the gas is critical to interpreting the data. As such, unless the temperature is known or assumed *a priori*, any ionization model of the absorbing gas must include the constraint of heating and cooling balance in order to be self consistent.

11.6.1 Photoionization Heating

Photoionization results in the removal of a photon from the radiation field whose energy is converted into the kinetic energy of the ionized electron. The electron enters the particle field with energy $\Delta E = h\nu - I_{\text{k,j}}$, where $h\nu$ is the photon energy and $I_{\text{k,j}}$ is the ionization energy of the atom or ion. In this case, the gas energy is increased with the introduction of the ejected electron.

To obtain the photo-heating power, $\Upsilon_{\text{k,j}}^{\text{ph}}$, one includes the kinetic energy ΔE imparted on the photoionized electron in the integral quantifying the

photoionization rate (Eq. 11.92), yielding

$$\frac{dQ_{k,j}^{\text{ph}}}{dt} = n_{k,j} \Upsilon_{k,j}^{\text{ph}} = n_{k,j} \sum_{s=1}^{N_{k,j}^s} \int_{\nu_{k,j,s}}^{\infty} \frac{4\pi J_{\nu}}{h\nu} (h\nu - I_{k,j,s}) \sigma_{k,j,s}^{\text{ph}}(\nu) d\nu, \quad (11.123)$$

where $I_{k,j,s} = h\nu_{k,j,s}$ and the photoionization cross section is $\sigma_{k,j,s}^{\text{ph}}(\nu)$. Note that $n_{k,j} = f_{k,j} n_k = n_{\text{H}} f_{k,j} (n_{k,j}/n_{\text{H}})$, which highlights the fact that the photoionization heating from each ion depends on its ionization fraction and its number density (abundance) relative to hydrogen.

As hydrogen and helium are the dominant elements by number, we can obtain a good approximation to the photoionization heating rate by assuming a pure hydrogen/helium gas. Employing Eq. 11.121, for hydrogen we have

$$\frac{dQ_{\text{H}}^{\text{ph}}}{dt} = n_{\text{H}^0} \Upsilon_{\text{H}^0}^{\text{ph}} = f_{\text{H}^0} n_{\text{H}} \Upsilon_{\text{H}^0}^{\text{ph}} \quad (11.124)$$

where the H^0 photoionization cross section is given by Eq. 2.91. Assuming ionization equilibrium and detailed balancing (Eq. 11.70), we derived f_{H^0} in § 12.2. Using that expression (Eq. 12.16), inserting Eq. 12.20, and rearranging, we have

$$\frac{dQ_{\text{H}}^{\text{ph}}}{dt} = \left(\frac{n_e \beta_{\text{H}^0}^{\text{phr}}}{n_e \beta_{\text{H}^0}^{\text{phr}} + R_{\text{H}^0}^{\text{ph}} + n_e \alpha_{\text{H}^0}^{\text{cdi}}} \right) n_{\text{H}} \Upsilon_{\text{H}^0}^{\text{ph}}, \quad (11.125)$$

where $R_{\text{H}^0}^{\text{ph}}$ is the photoionization rate, and $\beta_{\text{H}^0}^{\text{phr}}$ and $\alpha_{\text{H}^0}^{\text{cdi}}$ are the radiative recombination and direct collisional ionization rate coefficients, respectively. For helium, we have $dQ_{\text{He}}^{\text{ph}}/dt = dQ_{\text{He}^0}^{\text{ph}}/dt + dQ_{\text{He}^+}^{\text{ph}}/dt$. Following the steps for hydrogen, we have $dQ_{\text{He}^0}^{\text{ph}}/dt = f_{\text{He}^0} n_{\text{He}} \Upsilon_{\text{He}^0}^{\text{ph}}$ and $dQ_{\text{He}^+}^{\text{ph}}/dt = f_{\text{He}^+} n_{\text{He}} \Upsilon_{\text{He}^+}^{\text{ph}}$. From our derivation of the helium ionization fractions (Eqs. 12.19 and 12.21), and rearranging terms, we have

$$\begin{aligned} \frac{dQ_{\text{He}}^{\text{ph}}}{dt} &= n_{\text{H}} \left[f_{\text{He}^0} \frac{n_{\text{He}}}{n_{\text{H}}} \Upsilon_{\text{He}^0}^{\text{ph}} + f_{\text{He}^+} \frac{n_{\text{He}}}{n_{\text{H}}} \Upsilon_{\text{He}^+}^{\text{ph}} \right] \\ f_{\text{He}^0} &= \frac{n_e^2 \beta_{\text{He}^0}^{\text{phr}} \beta_{\text{He}^+}^{\text{phr}}}{n_e^2 \beta_{\text{He}^0}^{\text{phr}} \beta_{\text{He}^+}^{\text{phr}} + (R_{\text{He}^0}^{\text{ph}} + n_e \alpha_{\text{He}^0}^{\text{cdi}})(R_{\text{He}^+}^{\text{ph}} + n_e \alpha_{\text{He}^+}^{\text{cdi}} + n_e \beta_{\text{He}^+}^{\text{phr}})} \\ f_{\text{He}^+} &= \frac{n_e \beta_{\text{He}^+}^{\text{phr}} (R_{\text{He}^0}^{\text{ph}} + n_e \alpha_{\text{He}^0}^{\text{cdi}})}{n_e^2 \beta_{\text{He}^0}^{\text{phr}} \beta_{\text{He}^+}^{\text{phr}} + (R_{\text{He}^0}^{\text{ph}} + n_e \alpha_{\text{He}^0}^{\text{cdi}})(R_{\text{He}^+}^{\text{ph}} + n_e \alpha_{\text{He}^+}^{\text{cdi}} + n_e \beta_{\text{He}^+}^{\text{phr}})}. \end{aligned} \quad (11.126)$$

Note that we have written Eq. 11.126 in a form to emphasize the heating rate for helium is primarily proportional to n_{H} as scaled to the abundance

of helium $n_{\text{He}}/n_{\text{H}}$. This form allows the explicit expression of the sensitivity of the gas heating/cooling rates to abundance pattern.

Photons with energies just above the ionization threshold energy are preferentially absorbed due to the shape of the hydrogen photoionization cross sections, which scale roughly as $(\nu_0/\nu)^3$. Thus, as the radiation field penetrates further into the gas, the radiation field becomes progressively harder; deeper into the gas structure the higher mean energy of the photons yields a higher mean kinetic energy to the photoionized electrons. Thus, photoionization heats the gas to higher temperatures at larger depths into gas clouds.

11.6.2 Collisional Ionization Cooling

This term accounts for a free electron imparting part of its kinetic energy to the ionization of an atom or ion. The energy lost to the particle field is always equal to the ionization energy of the ionized electron. Consider direct collisional ionization with an incident colliding electron having kinetic energy KE_i . Following collisional ionization, the sum of the kinetic energies of the post-collision electron and the ionized electron is $\text{KE}_f = \text{KE}_i - I_{k,j,s}$, which reflects the binding energy that was overcome to liberate the ionized electron from shell s . We thus have $\Delta E = I_{k,j,s}$, which is a constant for ion k, j . Substituting into Eq. 11.120 and writing in terms of the incident velocity of the colliding electron,

$$\Lambda_{k,j}^{\text{cdi}} = \sum_{s=1}^{N_{k,j}^s} I_{k,j,s} \int_{v_{k,j,s}}^{\infty} \sigma_{k,j,s}^{\text{cdi}}(v) f(v, T) v dv = \sum_{s=1}^{N_{k,j}^s} I_{k,j,s} \alpha_{k,j,s}^{\text{cdi}}, \quad (11.127)$$

where $v_{k,j,s} = \sqrt{2I_{k,j,s}/m_e}$. For a pure hydrogen and helium gas, the expressions for $dQ_{\text{H}}^{\text{cdi}}/dt$, and for $dQ_{\text{He}}^{\text{cdi}}/dt = dQ_{\text{He}^0}^{\text{cdi}}/dt + dQ_{\text{He}^+}^{\text{cdi}}/dt$ are analogous to Eqs. 11.125 and 11.126 with $\Upsilon_{k,j}^{\text{ph}}$ replaced with $n_e \Lambda_{k,j}^{\text{cdi}}$.

For the general metal atoms and ions, we apply Eqs. 11.118 and 11.119,

$$\frac{dQ^{\text{cdi}}}{dt} = \sum_{k,j} f_{k,j} n_k n_e \Lambda_{k,j}^{\text{cdi}} = n_{\text{H}} n_e \sum_{k,j} \frac{n_k}{n_{\text{H}}} f_{k,j} \Lambda_{k,j}^{\text{cdi}}, \quad (11.128)$$

Note that, because the ionization fractions depend on the photoionization rates as well as the collisional ionization rates, the collisional cooling rates still depend on the photoionization rates. Furthermore, we see that the cooling depends on the abundance ratios of the various metals. Thus, we see that the cooling is strongly coupled to both the chemical and ionization conditions of the gas.

11.6.3 Recombination Cooling

When an electron is captured and becomes bound to atom or ion k, j , the available energy, $I_{k,j,s} + \frac{1}{2}m_e v^2$, where v is the velocity of the recombining electron, is most commonly channel into an emitted photon (radiative recombination). However, in denser gas, the energy can be transferred to a proximate free electron thereby increasing its kinetic energy (collisional recombination).

In radiative recombination, regardless of whether the electron recombines to an excited state or the ground state the energy removed from the particle field equals the kinetic energy of the recombination electron. The ion may release a single photon of energy $h\nu = I_{k,j,s} + \frac{1}{2}m_e v^2$, or a series of photons (called recombination lines) as the electron cascades via spontaneous de-excitation to the ground state. Regardless, the particle field has lost energy $\Delta E = \frac{1}{2}m_e v^2$, whereas the radiation field has gained energy $I_{k,j,s} + \frac{1}{2}m_e v^2$. We have,

$$\Lambda_{k,j}^{\text{rec}} = \sum_{s=1}^{N_{k,j}^s} \int_0^\infty [\sigma_{k,j,s}^{\text{phr}}(v) + \sigma_{k,j,s}^{\text{die}}(v)] f(v, T) \frac{1}{2}m_e v^2 v dv, \quad (11.129)$$

where we have included the dielectronic recombination channel since it also results in energy $\frac{1}{2}m_e v^2$ exchanged to the radiation field.

Globally across the gas structure, radiative recombination is a cooling process. However, if one thinks about the process in more detail, it becomes clear that local to the recombination event radiative recombination raises the average kinetic energy, $\frac{3}{2}kT$ and therefore gas temperature. This is because of the velocity dependence of the recombination cross section, which is largest for smaller velocities such that electrons on the low energy tail (those below the average kinetic energy) are preferentially removed from the particle field via recombination. In most all cosmological gas, especially environments for which optical and ultraviolet absorption lines are observed, the electron pool thermalizes to the lower average $\frac{3}{2}kT$ on short time scales.

In collisional recombination, the gas loses the kinetic energy of the recombination electron, but, for a recombination to shell s , the kinetic energy of the collisional electron is increased by the amount $\Delta E = I_{k,j,s}^{\text{rec}} + \frac{1}{2}m_e v^2$, where v is the velocity of the recombining electron prior to its capture. Since the kinetic energy returned to the particle field is equal to the recombining electron's pre-capture kinetic energy plus released binding energy, collisional recombination is always a heating process. The collisional recombination cooling function, $\Lambda_{k,j}^{\text{coll}}$, is computed identically to Eq. 11.129 with the radiative and

dielectronic cross sections replaced with the collisional recombination cross section, and the energy exchange replaced with $I_{k,j,s}^{\text{rec}} + \frac{1}{2}m_e v^2$.

Applying Eqs. 11.118 and 11.119, the total recombination cooling is

$$\frac{dQ^{\text{rec}}}{dt} = n_H n_e \sum_{k,j} \frac{n_k}{n_H} f_{k,j} (\Lambda_{k,j}^{\text{phr}} - \Lambda_{k,j}^{\text{coll}}) , \quad (11.130)$$

where the difference is taken to reflect that collisional recombination is always a heating term.

11.6.4 Collisional Excitation Cooling

As described in § 11.4.2, in both the interstellar medium and circumgalactic medium, a collisionally excited bound electron is most likely to spontaneously de-excite and emit a photon. For this scenario, kinetic energy $\Delta E = \chi_{k,j,i'} - \chi_{k,j,i}$ is lost to the electron pool and transferred to the radiation field, where $\chi_{k,j,i}$ is the excitation energy of the initial state and $\chi_{k,j,i'}$ is the excitation energy of the final state. The net effect is that the particle field loses energy with each collisional excitation.

In a higher density environment, the rate of collisional de-excitations can become competitive with the rate of spontaneous de-excitations. When a collisional de-excitation occurs, the energy lost to the original colliding electron, being transferred to a second colliding electron, is restored to the particle field. As such, collisional de-excitation is a heating term. However, since spontaneous de-excitations will also be occurring, some energy is still being lost to the radiation field. Thus, even at very high densities, the net energy exchange from collisional excitation and de-excitation results in a cooling of the particle field.

Adopting the two-level atom for reasons we described in § 11.4, and accounting for collisional de-excitations, the rate of cooling can be expressed

$$\frac{dQ_{1,2}^{\text{ex}}}{dt} = n_1 n_e \Lambda_{1,2}^{\text{ex}} - n_2 n_e \Lambda_{2,1}^{\text{ex}} = n_2 A_{2,1} \chi_{1,2} , \quad (11.131)$$

where the excitation energy $\chi_{1,2} = h\nu$ is equal to the energy of the emitted photon, and thus the energy lost to the particle field per spontaneous de-excitation. The cooling functions can be computed from

$$\Lambda_{1,2}^{\text{ex}} = \chi_{1,2} \alpha_{1,2}^{\text{ex}} = \chi_{1,2} \int_{\chi_{1,2}}^{\infty} \sigma_{1,2}^{\text{ex}}(v) f(v, T) v dv , \quad (11.132)$$

and

$$\Lambda_{2,1}^{\text{ex}} = \chi_{1,2} \alpha_{2,1}^{\text{ex}} = \chi_{1,2} \int_0^{\infty} \sigma_{2,1}^{\text{ex}}(v) f(v, T) v dv . \quad (11.133)$$

To obtain the cooling rate, we begin with the right hand side of Eq. 11.131, which informs us that the cooling is due to the population of excited ions that undergo spontaneous emission, $dQ_{1,2}^{\text{ex}}/dt = n_2 A_{2,1} \chi_{1,2}$. The number density n_2 is governed by detailed balancing,

$$R_{1,2}^{\text{ex}} = R_{2,1}^{\text{ex}} + n_2 A_{2,1}, \quad (11.134)$$

or

$$n_1 n_e \alpha_{1,2}^{\text{ex}} = n_2 n_e \alpha_{2,1}^{\text{ex}} + n_2 A_{2,1}, \quad (11.135)$$

which yields

$$n_2 = n_1 \left[\frac{n_e \alpha_{1,2}^{\text{ex}} / A_{2,1}}{1 + (n_e \alpha_{2,1}^{\text{ex}} / A_{2,1})} \right]. \quad (11.136)$$

Substituting into Eq. 11.131, we obtain

$$\frac{dQ_{1,2}^{\text{ex}}}{dt} = n_2 A_{2,1} \chi_{1,2} = \frac{n_1 n_e \alpha_{1,2}^{\text{ex}} \chi_{1,2}}{1 + (n_e \alpha_{2,1}^{\text{ex}} / A_{2,1})}. \quad (11.137)$$

In the low density limit, $n_e \rightarrow 0$, collisional de-excitations become negligible compared to spontaneous radiative de-excitations. The denominator of Eq. 11.137 approaches unity, and $dQ_{1,2}^{\text{ex}}/dt \rightarrow n_1 n_e \alpha_{1,2}^{\text{ex}} \chi_{1,2} = n_1 n_e \Lambda_{1,2}^{\text{ex}}$, where $\Lambda_{1,2}^{\text{ex}}$ is computed from Eq. 11.132. In the high density limit, $n_e \rightarrow \infty$, the ratio $n_2/n_1 \rightarrow (g_2/g_1) \exp\{-\chi_{1,2}/kT\}$, which yields

$$\frac{dQ_{1,2}^{\text{ex}}}{dt} = n_2 A_{2,1} \chi_{1,2} \rightarrow n_1 A_{2,1} \chi_{1,2} \frac{g_2}{g_1} \exp\{-\chi_{1,2}/kT\}. \quad (11.138)$$

Writing Eq. 11.137 in terms of an average cooling function to recover the formalism, $dQ_{1,2}^{\text{ex}}/dt = n_1 n_e \bar{\Lambda}_{1,2}^{\text{ex}}$, we have

$$\bar{\Lambda}_{1,2}^{\text{ex}} = \frac{\alpha_{1,2}^{\text{ex}} \chi_{1,2}}{1 + (n_e \alpha_{2,1}^{\text{ex}} / A_{2,1})} = \frac{\Lambda_{1,2}^{\text{ex}}}{1 + (n_e \Lambda_{2,1}^{\text{ex}} / A_{2,1} \chi_{1,2})}. \quad (11.139)$$

To express the cooling rate in terms of the abundance of the ion, $n_{k,j} = f_{k,j} n_k$ as opposed to the ion's ground state number density, $n_1 = n_{k,j,1}$, or excited state number density $n_2 = n_{k,j,2}$, we invoke $n_{k,j,1} + n_{k,j,2} = f_{k,j} n_k$, which yields,

$$\left. \frac{dQ_{1,2}^{\text{ex}}}{dt} \right|_{k,j} = n_1 n_e \bar{\Lambda}_{1,2}^{\text{ex}} = n_H n_e f_{k,j} \frac{(n_k/n_H)}{(1 + n_{k,j,1}/n_{k,j,2})} \bar{\Lambda}_{1,2}^{\text{ex}}. \quad (11.140)$$

where $n_{k,j,1}/n_{k,j,2} = n_1/n_2$ is given by Eq. 11.136.

For the above, we have assumed a two-level atom, which we have demonstrated is a good approximation for Group IIIA atoms and ions, such as CII, NIII, OIV, AlI, SiII, etc. However, for Group IVA, VA, and VIA ions, such as CI, NI, NII, OII, OIII, SiI, and MgIV–MgVII, multiple p subshell electrons

yield triplet fine-structure splittings with the ground state. As such, these atoms and ions cannot be treated as two-level systems. We refer the reader to the treatment by Osterbrock & Ferland (2006) for solving the equilibrium equations required to obtain the contribution to the collisional excitation cooling rate from these ions.

11.6.5 Continuum Cooling

Here, we discuss the heating and cooling due to free-free energy exchanges, which primarily involve a change in energy of a free electron in the vicinity of an atom or ion. Free-free absorption, which occurs as described in § 1.6 and illustrated in Figure 1.12, is a heating process in which a free electron absorbs a photon. The proximate nuclear particle serves to conserve angular momentum (since a free electron has no intrinsic angular momentum, and thus cannot absorb a photon without a nearby recoil event). Free-free emission, also known as *Bremsstrahlung*, or “braking radiation”, is a cooling process in which a free electron is slowed or deflected when passing through the electric field of a nuclear particle. In both cases, the energy exchanged between the radiation field and the particle field is $\Delta E = KE_f - KE_i = h\nu$, where the final kinetic energy of the free electron is $\frac{1}{2}m_e v_f^2$, the initial energy is $\frac{1}{2}m_e v_i^2$, and the energy lost to the particle field is $h\nu$. The cooling function for free-free processes is,

$$\Lambda_{k,j}^{\text{ff}} = \int_0^\infty \int_0^\infty \sigma_{k,j}^{\text{ff}}(\nu, v) \exp\{-h\nu/kT\} f(v, T) h\nu v dv d\nu, \quad (11.141)$$

where the factor $\exp\{-h\nu/kT\}$ accounts for self-absorption in optically thick gas (it can be omitted in optically thin gas where the emitted photon can escape unimpeded, but continuum cooling is negligible compared to other cooling processes in that regime). Expressed in a semi-classical form (Menzel, 1930),

$$\sigma_{k,j}^{\text{ff}}(\nu, v) dv d\nu = \frac{2^5}{3\sqrt{3}} \frac{\pi^2 e^6}{h m_e^2 c^3} Z_k^2 g_{\text{III}}(\gamma^2, u) \frac{dv}{v} \frac{d\nu}{\nu} \quad (11.142)$$

where $\gamma^2 = R_k/kT$ and $u = h\nu/kT$, and where R_k is the Rydberg constant (Eq. 1.6) for an atom or ion of species k having nuclear charge eZ_k . The function $g_{\text{III}}(\gamma^2, u)$ is the Gaunt factor (Gaunt, 1930) for free-free transitions. Essentially, the Gaunt factor is the ratio of the semi-classical form of the cross section to the full expression derived from quantum mechanics (which requires the overlap integrals of the wave functions of the initial and final

states of the free electron, see § 2.1.1). Directly substituting the Maxwell-Boltzmann electron velocity distribution function (Eq. 11.33) and invoking the change of variables $x = m_e v^2 / 2kT$ and $u = h\nu / kT$, the integrals become

$$\Lambda_{k,j}^{\text{ff}} = A_0 Z_k^2 \left(\frac{2m_e kT}{\pi} \right)^{1/2} \int_0^\infty \exp\{-u\} du \int_0^\infty g_{\text{III}}(\gamma^2, u) \exp\{-x\} dx, \quad (11.143)$$

where $A_0 = (2^5/3\sqrt{3})(\pi^2 e^6 / h m_e^2 c^3)$. Further simplification yields,

$$\Lambda_{k,j}^{\text{ff}} = A_0 Z_k^2 \left(\frac{2m_e kT}{\pi} \right)^{1/2} \int_0^\infty \bar{g}_{\text{III}}(\gamma^2, u) \exp\{-u\} du, \quad (11.144)$$

where $\bar{g}_{\text{III}}(\gamma^2, u)$ is the thermal averaged free-free Gaunt factor. Integrating over all frequencies, we obtain

$$\Lambda_{k,j}^{\text{ff}} = \frac{2^5}{3\sqrt{3}} \frac{\pi e^6}{h m_e c^3} \left(\frac{2\pi kT}{m_e} \right)^{1/2} Z_k^2 G_{\text{III}}(\gamma^2), \quad (11.145)$$

where we introduce the frequency integrated thermally averaged Gaunt factor $G_{\text{III}}(\gamma^2)$. The Gaunt factors, the thermally averaged Gaunt factors, and the frequency integrated thermally averaged Gaunt factors have been computed and tabulated in electronic form by Sutherland (1998). For most astrophysical values of γ and u , these functions remain near unity. Evaluating the constants, we have

$$\Lambda_{k,j}^{\text{ff}} = 1.426 \times 10^{-27} T^{1/2} Z_k^2 G_{\text{III}}(\gamma^2), \quad (11.146)$$

The cooling rate is then obtained by summing over the contribution from all atoms and ions,

$$\frac{dQ^{\text{ff}}}{dt} = n_{\text{H}} n_e \sum_{k,j} \frac{n_k}{n_{\text{H}}} f_{k,j} \Lambda_{k,j}^{\text{ff}}. \quad (11.147)$$

Note that $\Lambda_{k,j}^{\text{ff}} \propto T^{1/2}$, which indicates that free-free cooling increases in importance at the highest temperatures. In fact, free-free cooling dominates the cooling rate for $T > 10^7$ K, which is primarily X-ray emitting gas.

11.6.6 Compton Heating and Cooling

The Compton process (Compton, 1923) is the scattering and energy exchange between the radiation field and the free electrons. The environments where heating or cooling by Compton scattering include the hot intracluster medium, the circumgalactic medium of quasars and Seyfert galaxies, and even the regions surrounding pulsars.

Whether Compton scattering is a heating or cooling term depends on the mean energy of the radiation field relative to the mean energy of the particle field. When the thermal energy of the electrons exceeds that of the radiation field, then the Compton scattering transfers energy to the radiation field and is thus a cooling term for the particle field (this “inverse” Compton scattering is what gives rise to the Sunyaev-Zel’dovich effect where the energies of cosmic microwave background photons are elevated in the hot intracluster gas of large galaxy clusters). When the mean photon energy exceeds the mean thermal energy of the electrons, the electrons lose energy to the radiation field and the Compton scattering is a cooling effect.

The Compton heating/cooling rate is the difference of the Compton heating and cooling power [erg s⁻¹],

$$\frac{dQ^c}{dT} = n_e (\Upsilon_+^c - \Upsilon_-^c) . \quad (11.148)$$

The equation for the Compton interaction between thermalized electrons with photons was presented by Kompaneets (1957, Eq. 2). Converted to an energy equation (by multiplying by $h\nu$), and integrating over all photon phase space yields (Levich & Sunyaev, 1970; Pomraning, 1973),

$$\Upsilon_+^c = \frac{1}{m_e c} \int_0^\infty \sigma_{\text{KN}}^c(\nu) J_\nu h\nu d\nu + \frac{h^2 c^3}{8\pi m_e} \int_0^\infty \sigma_{\text{KN}}^c(\nu) (J_\nu / h\nu)^2 d\nu \quad (11.149)$$

and

$$\Upsilon_-^c = \frac{4kT}{m_e c} \int_0^\infty \sigma_{\text{KN}}^c(\nu) J_\nu d\nu , \quad (11.150)$$

where the Klein & Nashina (1929) cross section integrated over all solid angle is

$$\begin{aligned} \sigma_{\text{KN}}^c(x) = & \frac{3\sigma_{\text{T}}}{4} \frac{(1+x)}{x^2} \left[\frac{2(x+1)}{2x+1} - \frac{1}{x} \ln(2x+1) \right] \\ & + \frac{3\sigma_{\text{T}}}{4} \left[\frac{1}{2x} \ln(2x+1) - \frac{(3x+1)}{(2x+1)^2} \right] , \quad x = \frac{h\nu}{m_e c^2} , \end{aligned} \quad (11.151)$$

and where $\sigma_{\text{T}} = (8\pi/3)(e^2/m_e c^2)^2$ is the classical Thompson electron scattering cross section. For $h\nu \ll m_e c^2$, which is the condition for heating (“inverse” scattering), the cross section can be written $\sigma_{\text{KN}}^c(x) \simeq \sigma_{\text{T}}(1 - 2x + 26x^2/5)$, which shows that $\sigma_{\text{KN}}^c(x) \rightarrow \sigma_{\text{T}}$ as $x \rightarrow 0$. In this regime the energy exchange is very small. In the relativistic regime ($x \gg 1$), $\sigma_{\text{KN}}^c(x) \simeq \frac{3}{8}\sigma_{\text{T}}x^{-1}(\ln 2x + \frac{1}{2})$; though the cross section diminishes proportional to x^{-1} , the energy exchange (cooling) increases.

11.6.7 The Net Cooling Curve

The net total cooling rate can then be written

$$\frac{dQ}{dt} = n_{\text{H}} \sum_{k,j} \frac{n_k}{n_{\text{H}}} f_{k,j} \Upsilon_{k,j}^{\text{ph}} - n_{\text{H}} n_e \sum_i \sum_{k,j} \frac{n_k}{n_{\text{H}}} f_{k,j} \Lambda_{k,j}^{(i)} + n_e (\Upsilon_+^{\text{C}} - \Upsilon_-^{\text{C}}) . \quad (11.152)$$

Note that for all collisional processes (central term on right hand side), the cooling rates are expressed as the cooling function (a temperature dependent expression unique to each process) multiplied by the free electron number density and the number density of the relevant ionic species.

Plotting this quantity as a function of temperature is called the cooling curve. If the gas is confined by gravity or by ambient pressure, loss of thermal energy is usually accompanied by an increase in density that restores the pressure equilibrium with the surrounding medium. This will drive $dQ/dt \rightarrow 0$ and the gas to the equilibrium temperature.

In most presentations of the cooling curve in the literature, the photoionization heating and Compton terms are ignored; this applies only in astrophysical environments in which energy exchange with the radiation field is negligible compared to the particle-particle energy exchange. For this scenario, the cooling rate is simplified to the expression

$$dQ/dt = n_{\text{H}} n_e \Lambda_{\text{net}} , \quad (11.153)$$

where Λ_{net} would represent the double central term on the right hand side of Eq. 11.152. This is known as the “coronal approximation” (e.g., Dopita & Sutherland, 2003). An exploration of the effects of density and metallicity for these conditions was presented by Wang et al. (2014), including useful fitting functions.

We emphasize again that the coronal approximation is not valid for the circumgalactic nor the intergalactic medium, and is rarely valid for the interstellar medium. We thus caution the reader to be diligent in their attention to knowing the assumptions underlying a given cooling curve before applying it in their research. A first exploration of the effects of photoionization heating and the relative contribution of various important metal ions, including comparison to the coronal approximation cooling curve, was undertaken by Wiersma et al. (2009).

In Figure 11.10, we show cooling functions from the work of Wiersma et al. (2009), which were generated with Cloudy 07.02 (Ferland et al., 1998) for model clouds as a function of n_{H} , $n_{\text{He}}/n_{\text{H}}$, T , Z/Z_{\odot} , and redshift z using the cosmic UVB of Haardt & Madau (2001). The coronal approximation is shown in Figure 11.10(a), whereas the inclusion of photoionization is shown

in Figure 11.10(b). The model clouds are assumed to be in ionization equilibrium, dust free, optically thin, and have solar metallicity. Note that the dashed curves give the cooling function for zero-metallicity. As metallicity is reduced, the total cooling curve approaches the dashed curve.

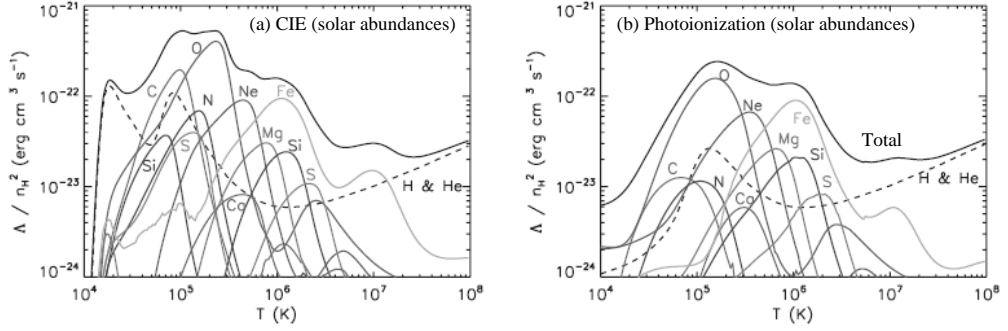


Figure 11.10 Solar abundance cooling curves normalized to n_{H}^2 for $n_{\text{H}} = 10^{-4} \text{ cm}^{-3}$ for a $z = 3$ cosmic UVB ionizing spectrum. The contribution to the total cooling is shown for each metal species (sum of all ions for that species) and for hydrogen and helium only (dashed curves), which corresponds to a zero-metallicity gas. (a, CIE) The collisional ionization equilibrium curves. Note that the normalized values of these curves are independent of n_{H} . (b, Photoionization) The cooling curve for photoionization. [Fig. 6 of Wiersma et al. (2009)]

For a given gas phase location, (n_{H}, T) , the inclusion of photoionization heating will yield a gas that is more highly ionized than collisionally ionized gas. This results in cooling rates that are diminished as compared to the rates for collisionally ionized gas (Wiersma et al., 2009; Schure et al., 2009, and references therein). Thus, the standard collisional ionization equilibrium cooling functions (e.g., Sutherland & Dopita, 1993) can substantially underestimate the cooling timescales in photoionized gas.

12

Ionization Modeling

The goal of ionization modeling gas observed in absorption is to constrain the chemical and ionization conditions of various cosmological environments and a range of cosmic epochs. It is through the tool of ionizing modeling that we are capable of charting the chemical evolution of the intergalactic medium and the circumgalactic medium, or estimating the mass of baryons in the hot gas phase as compared to the cool gas phase in the circumgalactic medium for various types of galaxies.

From the analysis of the absorption lines in the spectra of background luminous sources, i.e., quasars, gamma-ray bursts, starburst galaxies, etc., we extract estimates of the column densities for the ionization stages of several chemical species, $N_{k,j}$ from a single gas structure (which is sometimes referred to as an “absorption line system”). For example, a given system may include absorption lines from the Lyman series of H I, as well as the commonly observed Mg II $\lambda\lambda 2796, 2803$ and C IV $\lambda\lambda 1548, 1550$ doublets, or perhaps also the O VI $\lambda\lambda 1031, 1037$ doublet, all which yield¹ N_{HI} , N_{MgII} , N_{CIV} , and N_{OVI} . A “rich” system may also include C II, Si II, Si IV, and several Fe II absorption lines, yielding N_{CII} , N_{SiII} , N_{SiIV} , and N_{FeII} (see Table 4.5 for common ions seen in absorption).

Ionization modeling of the absorbing gas will predict the ionization fractions of each of these ions, allowing the densities of the atomic species to be estimated. The model cloud will have a quantifiable gas phase (n_{H}, T), physical size, L , from which insights can be gained about the mass and physical extent of the absorbing gas structure. It may be that the observed column density ratios cannot be made consistent with the assumed ionizing spectrum, so that constraints on the local radiation field can be explored.

¹ We make the transition in notation to reflect the historical convention that quantities measured in spectra adopt the roman numeral format, i.e H I for H^0 , H II for H^+ , C IV for C^{+3} , etc.

Or, the column density ratios may indicate a chemical abundance pattern associated with a known astrophysical process, thus providing insight into the enrichment history of the absorbing gas.

In this chapter, we build on the formalism developed in Chapter 11 to discuss ionization modeling and its application to observed data. The use of ionization modeling is so far reaching and varied that no single chapter can come close to covering the nuances of the steps involved in extracting insights about the absorbing gas. Partly, this is because each absorption system is unique in the limited data it yields, not only in terms of the absorption lines detected, but in terms of the quality of the observed spectra. Thus, each system requires a unique approach, and that approach is driven in great part by one's physical intuition as founded on their knowledge of gas physics, astrophysical processes, and their experience with ionization modeling.

We first cover the rate equations from the application of detailed balancing, and then present an illustrative solution for a pure hydrogen and helium gas. We then examine the timescales of ionization and recombination. Finally, we present introductory methods of ionization modeling using the publicly available code Cloudy (Ferland et al., 1998, 2013, 2017).

12.1 Rate Equations

Here, we derive the rate equations, $dn_{k,j}/dt$, for all ionization stages of all treatable species. We adopt the principle of detailed balancing for a two-level atoms, effectively the ground state and the continuum². Thus, the rate equations are taken directly from the simplified version (Eq. 11.70) of the global balance equation (Eq. 11.68). For the following, we drop the explicit temperature dependence of all rates and rate coefficients. All rates, denoted $R_{k,j}$, are referenced to the initial state j , except that we use the convention that recombination rate coefficients, $\beta_{k,j-1}$, are indexed to the final state $j-1$, having originated from the initial state j .

12.1.1 Hydrogen

Hydrogen is the simplest case because the channels for creation and destruction involve only two adjacent ionization stages. Because of this, the

² For a brief discussion of the ramification of this assumption, see Section 11.1 of *Hazy 2* (Ferland, 2002).

hydrogen rate equations for $n_{1,1}$ and $n_{1,2}$ are antisymmetric,

$$\begin{aligned}\frac{dn_{1,1}}{dt} &= n_{1,2}(R_{1,2}^{\text{rec}} + R_{1,2}^{\text{xH}^+}) - n_{1,1}(R_{1,1}^{\text{ph}} + R_{1,1}^{\text{coll}} + R_{1,1}^{\text{xH}}) \\ \frac{dn_{1,2}}{dt} &= -\frac{dn_{1,1}}{dt}.\end{aligned}\quad (12.1)$$

The creation rates of $n_{1,1}$ are due to the recombination of $n_{1,2}$ with free electrons, $R_{1,2}^{\text{rec}}$, and ionization charge exchange from metals, $R_{1,2}^{\text{xH}^+}$, where

$$R_{1,2}^{\text{rec}} = n_e \beta_{1,1}^{\text{phr}}, \quad R_{1,2}^{\text{xH}^+} = \sum_{k=2}^k \sum_{j=1} n_{k,j} \alpha_{k,j}^{\text{xH}^+}. \quad (12.2)$$

The destruction rates of $n_{1,1}$ are due to photoionization, $R_{1,1}^{\text{ph}}$, collisional ionization via free electrons, $R_{1,1}^{\text{coll}}$, and recombination charge exchange to metals, $R_{1,1}^{\text{xH}}$, which ionizes neutral hydrogen, where

$$R_{1,1}^{\text{coll}} = n_e \alpha_{1,1}^{\text{cdi}}, \quad R_{1,1}^{\text{xH}} = \sum_{k=2}^{k+1} \sum_{j=2} n_{k,j} \beta_{k,j-1}^{\text{xH}}. \quad (12.3)$$

Note that the negative of these rates are also the destruction and creation rates of $n_{1,2}$, respectively.

12.1.2 Helium

Helium has three ionization stages. There is no Auger channel directly connecting the neutral and fully ionized stages; however, non-zero dielectronic rate coefficients for the channel from singly ionized to neutral helium exist. The rate equation for neutral helium is

$$\frac{dn_{2,1}}{dt} = n_{2,2}(R_{2,2}^{\text{rec}} + R_{2,2}^{\text{xHe}^+}) - n_{2,1}(R_{2,1}^{\text{ph}} + R_{2,1}^{\text{coll}} + R_{2,1}^{\text{xHe}}). \quad (12.4)$$

The creation rates of $n_{2,1}$ are due to the recombination channels of $n_{2,2}$ with free electrons, $R_{2,2}^{\text{rec}}$, and ionization charge exchange from metals, $R_{2,2}^{\text{xHe}^+}$, where

$$R_{2,2}^{\text{rec}} = n_e (\beta_{2,1}^{\text{phr}} + \beta_{2,1}^{\text{die}}), \quad R_{2,2}^{\text{xHe}^+} = n_{1,1} \alpha_{1,1}^{\text{xHe}^+} + \sum_{k=3}^k \sum_{j=1} n_{k,j} \alpha_{k,j}^{\text{xHe}^+}. \quad (12.5)$$

The destruction rates of $n_{2,1}$ are due to photoionization, $R_{2,1}^{\text{ph}}$, collisional ionization via free electrons, $R_{2,1}^{\text{coll}}$, and recombination charge exchange to metals, $R_{2,1}^{\text{xHe}}$, which singly ionizes neutral helium, where

$$R_{2,1}^{\text{coll}} = n_e \alpha_{2,1}^{\text{cdi}}, \quad R_{2,1}^{\text{xHe}} = n_{1,2} \beta_{1,1}^{\text{xHe}} + \sum_{k=3}^{k+1} \sum_{j=2} n_{k,j} \beta_{k,j-1}^{\text{xHe}}, \quad (12.6)$$

respectively. For twice ionized helium, the rate equation is

$$\frac{dn_{2,3}}{dt} = n_{2,2}(R_{2,2}^{\text{ph}} + R_{2,2}^{\text{coll}} + R_{2,2}^{\text{xH}^+}) - n_{2,3}(R_{2,3}^{\text{rec}} + R_{2,3}^{\text{xH}}). \quad (12.7)$$

The creation rates via the destruction of $n_{2,2}$ are due to photoionization, $R_{2,2}^{\text{ph}}$, collisional ionization, $R_{2,2}^{\text{coll}}$, and ionization via charge exchange recombination to ionized hydrogen, where

$$R_{2,2}^{\text{coll}} = n_e \alpha_{2,2}^{\text{cdi}}, \quad R_{2,2}^{\text{xH}^+} = n_{1,2} \alpha_{2,2}^{\text{xH}^+}. \quad (12.8)$$

The destruction rates of $n_{2,3}$ are due to the channels of recombination with free electrons, $R_{2,3}^{\text{rec}}$, and recombination via charge exchange ionization of neutral hydrogen, where

$$R_{2,3}^{\text{rec}} = n_e \beta_{2,2}^{\text{phr}}, \quad R_{2,3}^{\text{xH}} = n_{1,1} \beta_{2,2}^{\text{xH}}. \quad (12.9)$$

For singly ionized helium, the creation and destruction rates are simply the negative of the sum of those for neutral and doubly ionized helium,

$$\frac{dn_{2,2}}{dt} = - \left(\frac{dn_{2,1}}{dt} + \frac{dn_{2,3}}{dt} \right). \quad (12.10)$$

12.1.3 Metals

Here, we write out the rate equations for all ions with $k \geq 3$. For what follows, let $R_{k,j-1}^{\text{ion}}$ denote the creation rate of $n_{k,j}$ via ionization of $n_{k,j-1}$ and let $R_{k,j+1}^{\text{rec}}$ denote the creation rate of $n_{k,j}$ via recombination from initial state $n_{k,j+1}$. Further, let $R_{k,j}^{\text{rec}}$ denote the destruction rate of $n_{k,j}$ via recombination to final state $n_{k,j-1}$ and let $R_{k,j}^{\text{ion}}$ denote the destruction of $n_{k,j}$ via ionization to final state $n_{k,j+1}$. We then write

$$\begin{aligned} \frac{dn_{k,j}}{dt} = & n_{k,j-1} R_{k,j-1}^{\text{ion}} + n_{k,j+1} R_{k,j+1}^{\text{rec}} + \sum_{i=1}^{j-2} n_{k,i} R_{k,i,j}^{\text{aug}} \\ & - n_{k,j} \left(R_{k,j}^{\text{ion}} + R_{k,j}^{\text{rec}} + \sum_{m=j+2}^{k-1} R_{k,j,m}^{\text{aug}} \right). \end{aligned} \quad (12.11)$$

The creation rate of $n_{k,j}$ via ionization destruction of adjacent ion $n_{k,j-1}$ is

$$R_{k,j-1}^{\text{ion}} = R_{k,j-1}^{\text{ph}} + n_e (\alpha_{k,j-1}^{\text{cdi}} + \alpha_{k,j-1}^{\text{cea}}) + n_{1,2} \alpha_{k,j-1}^{\text{xH}^+}. \quad (12.12)$$

The creation rate of $n_{k,j}$ via recombination destruction of adjacent ion $n_{k,j+1}$ is

$$R_{k,j+1}^{\text{rec}} = n_e (\beta_{k,j}^{\text{phr}} + \beta_{k,j}^{\text{die}}) + n_{1,1} \beta_{k,j}^{\text{xH}} + n_{2,1} \beta_{k,j}^{\text{xHe}}. \quad (12.13)$$

The recombination destruction rate of $n_{k,j}$ to adjacent stage $n_{k,j-1}$ and the ionization destruction rate of $n_{k,j}$ to adjacent stage $n_{k,j+1}$, are

$$\begin{aligned} R_{k,j}^{\text{rec}} &= n_e (\beta_{k,j-1}^{\text{phr}} + \beta_{k,j-1}^{\text{die}}) + n_{1,1} \beta_{k,j-1}^{\text{xH}} + n_{2,1} \beta_{k,j-1}^{\text{xHe}} \\ R_{k,j}^{\text{ion}} &= R_{k,j}^{\text{ph}} + n_e (\alpha_{k,j}^{\text{cdi}} + \alpha_{k,j}^{\text{cea}}) + n_{1,2} \alpha_{k,j}^{\text{xH}^+}, \end{aligned} \quad (12.14)$$

respectively. The summation terms in Eq. 12.11 account for Auger ionization processes, which skip adjacent ionization stages. All ions of species k from the neutral stage to ionization stage $i \leq j-2$ can contribute to the creation rate of $n_{k,j}$ due to their destruction via Auger ionization. Similarly, ion k, j can be destroyed by Auger ionization to high ionization final stage m , where $m \geq j+2$. In practice, Auger ionization is a viable creation and destruction channel only for $k \geq 4$.

12.2 Equilibrium Solution

The equilibrium solution is obtained when $dn_{k,j}/dt = 0$ is satisfied for the rate equation for all k and j . Depending on how many atomic species are included in the model cloud, there can be hundreds of coupled rate equations. In addition, the rate equations are non-linear because the collisional ionization and recombination rates for a given ion k, j depend on the product $n_e n_{k,j}$, and the charge exchange rates between ion k, j and ion k', j' depend on the product $n_{k,j} n_{k',j'}$. The system of equations is “closed” by enforcing charge density conservation, given by Eq. 11.26.

Discussing the steps to a full solution is beyond the scope of this work (see, e.g., Mihalas & Mihalas, 1984), however, a hydrogen and helium only gas in which no charge exchange is assumed can be used to illustrate the simplest case. Consider a static gas cloud of temperature T and total pressure P , where we will assume negligible radiation pressure, i.e., $P = P_g$. For this gas, the total particle number density is $n = P/kT$, where $n = n_N + n_e$. Further, assume the abundances fractions for hydrogen and helium are η_1 and η_2 and that the cloud is subject to a radiation field J_ν .

Writing out charge density conservation from Eq. 11.26, and rearranging to equal zero,

$$n_e - \left(\frac{P}{kT} - n_e \right) \left[\eta_1 f_{1,2} + \eta_2 (f_{2,2} + 2f_{2,3}) \right] = 0, \quad (12.15)$$

where we remind the reader that the ionization fractions are dependent on the value of n_e , i.e., $f_{k,j} = f_{k,j}(n_e, T, J_\nu)$. Thus, Eq. 12.15 must be root solved for n_e . We immediately know that the value of n_e is bracketed between the

minimum value $n_e = 0$ for a fully neutral gas, and the maximum value $n_e = (P/kT)(\eta_1 + 2\eta_2)/(1 + \eta_1 + 2\eta_2)$ for a fully ionized gas. A bisecting method, such as the Van Wijngaarden–Dekker–Brent Method (e.g., Press et al., 2007), can then be used to find the root to Eq. 12.15 by carrying out evaluation of the ionization at each test (current guess) value of n_e .

For each test value of n_e the ionization fraction must be determined. For hydrogen we have

$$f_{1,1} = \frac{n_{1,1}}{n_1} = \frac{n_{1,1}}{n_{1,1} + n_{1,2}} = \frac{1}{1 + (n_{1,2}/n_{1,1})}, \quad (12.16)$$

and

$$f_{1,2} = \frac{n_{1,2}}{n_{1,1} + n_{1,2}} = \frac{n_{1,2}/n_{1,1}}{1 + (n_{1,2}/n_{1,1})} = \frac{n_{1,2}}{n_{1,1}} f_{1,1}. \quad (12.17)$$

Similar operations for helium yield

$$f_{2,1} = \frac{1}{1 + (n_{2,2}/n_{2,1}) + (n_{2,3}/n_{2,2})(n_{2,2}/n_{2,1})}, \quad (12.18)$$

$$f_{2,2} = \frac{n_{2,2}}{n_{2,1}} f_{2,1}, \quad f_{2,3} = \frac{n_{2,3}}{n_{2,2}} f_{2,2}. \quad (12.19)$$

The required ion density ratios are obtained from the rate equations. Rearranging the hydrogen balance (Eq. 12.1), we have

$$\frac{n_{1,2}}{n_{1,1}} = \frac{R_{1,1}^{\text{ph}}(J_\nu) + R_{1,1}^{\text{coll}}(n_e, T)}{R_{1,2}^{\text{rec}}(n_e, T)} = \frac{R_{1,1}^{\text{ph}}(J_\nu) + n_e \alpha_{1,1}^{\text{cdi}}(T)}{n_e \beta_{1,1}^{\text{phr}}(T)}, \quad (12.20)$$

and for helium, we rearrange Eqs. 12.4 and 12.10, which yields

$$\frac{n_{2,2}}{n_{2,1}} = \frac{R_{2,1}^{\text{ph}}(J_\nu) + n_e \alpha_{2,1}^{\text{cdi}}(T)}{n_e \beta_{2,1}^{\text{phr}}(T)}, \quad \frac{n_{2,3}}{n_{2,2}} = \frac{R_{2,2}^{\text{ph}}(J_\nu) + n_e \alpha_{2,2}^{\text{cdi}}(T)}{n_e \beta_{2,2}^{\text{phr}}(T)}. \quad (12.21)$$

We clearly see that the ionization fractions, being functions of the ratios of adjacent ion densities, are themselves functions of J_ν , n_e , and T . Once Eq. 12.15 is satisfied, the atomic and ion number densities are computed from Eq. 11.29, where $n = P/kT$.

12.2.1 Ionization Equilibrium

Since most all ionization modeling of observational data is based upon the assumption of ionization equilibrium, it is important to understand to what degree this assumption holds. First investigations into non-equilibrium ionization have been presented by Cen & Fang (2006) and by Oppenheimer &

Schaye (2013), who both find that non-equilibrium effects are not as pronounced when photoionization is taken into account.

Whether the gas is predominantly photoionized or collisionally ionized has implications for correctly estimating the density, temperature, physical size, cooling time, and estimated mass of the absorbing gas. In observational work, one cannot directly know whether the gas is dominated by photoionization or collisional ionization. Inferences of collisional ionization are based upon the presence of large thermal velocity widths in the absorption profiles (suggestive of high temperatures), or the column density ratios of different ions in different ionization stages (e.g., Tripp et al., 2001; Simcoe et al., 2006; Tripp et al., 2011; Fox et al., 2014). However, as we will discuss here, under the assumption of ionization equilibrium, one can directly examine the degree to which the ionization balance in the absorbing gas is dominated by photoionization or collisional ionization from the ionization modeling (also see Smith et al., 2008).

The condition of ionization equilibrium is that the gas must be thermally stable over the time required for the photoionization, recombination, and collisional ionization rates to achieve a steady-state balance. If the time required to achieve ionization equilibrium is denoted τ_{eq} , the condition is often expressed $\tau_{\text{eq}} \ll \tau_{\text{cool}}$, where the cooling timescale is given by

$$\tau_{\text{cool}} = \frac{Q}{|dQ/dt|} = \frac{\frac{3}{2}(n_{\text{N}} + n_{\text{e}})kT}{|dQ/dt|}, \quad (12.22)$$

where Q is the energy density (erg cm^{-3} , see Eq. 11.41), and dQ/dt is the cooling rate ($\text{erg cm}^{-3} \text{ s}^{-1}$, see Eq. 11.52). The complexity and behavior of the cooling rate was discussed in detail in § 11.6.

The ionization and recombination timescales are different for each ion. As such, it is possible that, for a given n_{H}, T gas phase some ions are in ionization equilibrium, while other ions are not. The ionization timescales for ion k, j are obtained by rearranging the rate equations into the form $(1/n_{k,j})(dn_{k,j}/dt) = R_{k,j}$, where $R_{k,j}$ represents the net value of the creation rates and destruction rates for ion k, j on the right hand side of the rate equations. The solution is of the form $e^{-t/\tau_{k,j}}$, where $\tau_{k,j} = |R_{k,j}|^{-1}$. The absolute value reflects that the change in $n_{k,j}$ may be decreasing or increasing over the timescale interval. Thus, the timescales are the e -folding time for a change in the number density $n_{k,j}$ and are a function of the ionizing radiation, J_{ν} , ion densities, electron density, and gas temperature, all of which are assumed to be held constant so that $R_{k,j}$ can be approximated as a constant (thus, greatly simplifying the integrals yet rendering the timescales as approximations).

Accounting only for photoionization rates (omitting Auger ionization), we obtain the photoionization timescale for ion k, j ,

$$\tau_{k,j}^{\text{ph}} \simeq \left| R_{k,j}^{\text{ph}} - \frac{n_{k,j-1}}{n_{k,j}} R_{k,j-1}^{\text{ph}} \right|^{-1}, \quad (12.23)$$

where $R_{k,j}^{\text{ph}}(J_\nu)$ is the photoionization rate for ion k, j . In the case of the neutral ion ($j = 1$), the rates indexed by $j-1$ do not enter into the derivation and are omitted from Eq. 12.23.

Accounting for both photo and dielectronic recombination, we obtain the recombination timescale,

$$\tau_{k,j}^{\text{rec}} \simeq \frac{1}{n_e} \left| (\beta_{k,j-1}^{\text{phr}} + \beta_{k,j-1}^{\text{die}}) - \frac{n_{k,j+1}}{n_{k,j}} (\beta_{k,j}^{\text{phr}} + \beta_{k,j}^{\text{die}}) \right|^{-1}, \quad (12.24)$$

where $\beta_{k,j-1}^{\text{phr}}(T)$ and $\beta_{k,j-1}^{\text{die}}(T)$ are the photo and dielectronic recombination rate coefficients for ion k, j .

Accounting for both direct and excitation-auto collisional ionization processes, we obtain the collisional ionization timescale,

$$\tau_{k,j}^{\text{coll}} \simeq \frac{1}{n_e} \left| (\alpha_{k,j}^{\text{cdi}} + \alpha_{k,j}^{\text{cea}}) - \frac{n_{k,j-1}}{n_{k,j}} (\alpha_{k,j-1}^{\text{cdi}} + \alpha_{k,j-1}^{\text{cea}}) \right|^{-1}, \quad (12.25)$$

where $\alpha_{k,j}^{\text{cdi}}(T)$ and $\alpha_{k,j}^{\text{cea}}(T)$ are the direct and excitation-auto ionization rate coefficients for ion k, j . For the derivation of Eq. 12.25, we have omitted charge exchange reactions.

To estimate the timescales for a model cloud, one would need to compute the energy density, Q , and the photoionization, recombination, and collisional ionization timescales directly from the photoionization rates, equilibrium electron density, and the recombination and collisional ionization rate coefficients. In some applications, the timescale for ionization to come into balance with recombination is desired. If so, it can be computed from $\tau_{\text{eq}} = (\tau_{k,j}^{\text{rec}}/\tau_{k,j}^{\text{coll}})/[1 + (\tau_{k,j}^{\text{rec}}/\tau_{k,j}^{\text{coll}})]$ or $\tau_{\text{eq}} = (\tau_{k,j}^{\text{rec}}/\tau_{k,j}^{\text{ph}})/[1 + (\tau_{k,j}^{\text{rec}}/\tau_{k,j}^{\text{ph}})]$ using Eqs. 12.24, 12.25, and 12.23.

12.2.2 Ionization in Thermodynamic Equilibrium

Based upon the principles discussed in § 11.2.4 and § 11.4.1 regarding the dominance of collisional processes and equilibrium between the radiation field and the particle field in thermodynamic equilibrium, M. Saha derived an equation (also known as the Saha-Langmuir equation) that provides the

number density ratio of adjacent ionization stages,

$$\frac{n_{k,j+1}}{n_{k,j}} = \frac{2}{n_e} \left(\frac{2\pi m_e k}{h^2} \right)^{3/2} T^{3/2} \frac{U_{k,j+1}(T)}{U_{k,j}(T)} \exp \left\{ -\frac{\chi_{k,j,I}}{kT} \right\}, \quad (12.26)$$

where the first three terms on the right hand side follow from integrating the multiplicity of states for free electrons over the Maxwell-Boltzmann velocity distribution at the equilibrium temperature (see, for example Mihalas, 1978). The result is based upon the application of equipartition of energies and multiplicity of occupation of states for the free electrons, excitations, and ionizations.

Note that the ratio $n_{k,j+1}/n_{k,j}$ is proportional to temperature and the ratio of partition functions (Eq. 11.82) of the upper to the lower adjacent ionization stages, but inversely proportional to electron density. The ratio also scales inverse exponentially with the ratio of the ground-state ionization potential of the lower ionization stage (i.e., the negative of the binding energy of the ground state of the initial atom/ion prior to ionization) to the mean kinetic energy per particle in the gas.

Interpreting Eq. 12.26, the number density of the upper ionization stage decreases relative to the lower ionization stage with increasing electron density. This is because, in a higher density electron pool, the number of recombinations increase more rapidly for ion $k, j+1$ than for ion k, j due to the greater electron (charge) affinity of the higher ion.

As discussed in § 12.2, the electron density must be determined by enforcing charge density conservation via root solving Eq. 11.26. The iteration to solution is the same as outlined in § 12.2 to obtain the ionization fractions. However, what makes the Saha-Langmuir equation so powerful (in the rare conditions where it is applicable) is that (1) it is not required to compute the ionization and recombination rates for all ions in the gas, and (2) there is no need to solve the non-linear matrix of rate equations, making the inclusion of metals algorithmically as straight forward as solving a pure hydrogen gas. One simply computes $n_{k,j+1}/n_{k,j}$ for each guess n_e , computes the ionization fractions, and iterates until charge conservation is satisfied.

12.3 Applying Ionization Models to Data

Because the data are often limited to only one or two ionization states of different chemical species, it can be challenging to constrain both the abundance pattern *and* ionization conditions. Since the ionization balance is sensitive to the spectral energy distribution of the ionizing radiation, there can be a degeneracy between the spectrum shape and the abundance pattern.

If more than one ionization state is observed for a given species, this can help mitigate this degeneracy. For most applications, the sensitivity of the spectral energy distribution of the ionizing spectrum should be responsibly explored to help constrain the cloud properties and chemical conditions (abundance pattern).

12.3.1 *Cloudy*

For the assumption of ionization equilibrium, the industry standard ionization code is *Cloudy*, which is distributed by the *Cloudy and Associates Group* led by G. Ferland³. Among its many applications, *Cloudy* is well suited for cosmological absorption line studies. The code is highly versatile, treating all of the photo and collisional ionization and recombination processes described in Chapter 11 and additionally includes excitation and de-excitation processes in multi-level atoms/ions that are important to the heating and cooling balance of the model cloud.

The minimum required information to create a unique model cloud are:

Abundances

The metallicity in solar units, Z/Z_{\odot} , and the abundance pattern, expressed $\eta_{k,j}$, $x_{k,j}$, or n_k/n_N . A custom abundance pattern can be entered by hand or the pattern can be input from a standard list of commonly observed astrophysical patterns (solar, interstellar, supernovae yields, intergalactic, etc.) included with *Cloudy*. Additionally, the abundance of each individual atomic species can be uniquely adjusted relative to a standard input pattern.

Ionizing Spectrum

The spectral energy distribution of the ionizing spectrum, J_{ν} , incident on the model cloud. The user can manually input a custom spectral shape. Most commonly a cosmic UVB spectrum at a desired redshift is specified, from which *Cloudy* obtains $J_{\nu}(z)$ by interpolating across the included CUBA data grid (Haardt & Madau, 2012). Alternatively, a Starburst99 model (Leitherer et al., 1999) can be used, or a combined cosmic UVB and Starburst99 model.

Ionization Parameter

The intensity of the ionizing spectrum incident on the model cloud. This is quantified by the ionization parameter, U , which normalizes the intensity of the ionizing spectrum relative to the model cloud density. This quantity is the ratio of the number density of hydrogen ionizing photons, n_{γ} (see

³ *Cloudy* is publicly available at www.nublado.org/. All *Cloudy* based ionization modeling calculations presented here were performed by Elizabeth Klimek using version C07.02, last described by Ferland et al. (2017).

Eq. 11.2), to the hydrogen⁴ density, n_{H} , at the face of the model cloud, i.e., $U = n_{\gamma}/n_{\text{H}}$.

Boundary Condition

As Cloudy iterates toward a simultaneous solution of the transfer equation (Eq. 7.47) and the global balance equation (Eq. 11.68), it requires a boundary condition for the model cloud. This can be the cloud depth L , or it can be the total neutral hydrogen column density, $N_{\text{HI}} = n_{\text{H}}L$. This boundary condition is often called the “stopping condition”.

If each of these input parameters are specified, Cloudy will solve for the equilibrium ionization conditions and ionization structure of the model cloud. There are several additional input parameters that can be specified, but the minimum parameter set described above will run the default settings. These default settings are suitable for most all intergalactic and circumgalactic medium studies; the additional parameters and or physics are primarily applicable for interstellar medium studies, where gas environments are much more varied due to there proximity to a variety of stellar environments and activity. Summarizing, the simplifying assumptions include the following:

1. ionization equilibrium applies for all atoms and ions at all locations in the model cloud (includes recombination equilibrium)
2. the cloud model is dynamically static with no bulk velocity structure, which simplifies the radiative transfer through the cloud and eliminates the need to model hydrodynamics
3. the model cloud geometry is plane-parallel (the cloud thickness, L , is an output of the equilibrium model cloud)
4. throughout the model cloud the hydrogen density, n_{H} is constant, which allows the ionization parameter to be well-defined for an assumed ionizing radiation spectrum, and allows simple estimates of the column densities, $N_{\text{k,j}} = n_{\text{k,j}}L$, from the equilibrium ionization model
5. the ionizing radiation is the cosmic UVB, though variations are often explored to examine the effects of the spectral shape on the ionization balance and subsequent inferred gas properties
6. the illumination by the ionization radiation is on one face of the model cloud
7. dust physics, including heating, cooling, and radiative transfer, is omitted (though it is common to explore the effects of various dust populations on the ionization balance of the equilibrium model cloud)

⁴ We adopt the short-hand notation $n_{\text{HI}} = n_{1,2}$ and $N_{\text{HI}} = N_{1,2}$.

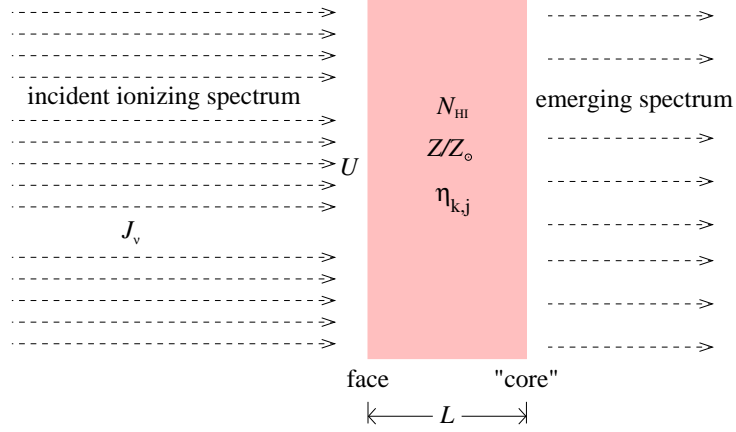


Figure 12.1 Schematic of a Cloudy model showing the geometric relationship between the cloud face and “core” and the incident ionizing radiation field, J_ν . The quantities defining the model cloud are the metallicity, Z/Z_\odot , the abundance pattern, $\eta_{k,j}$, the ionizing spectrum, J_ν , and the ionization parameter $U = n_\gamma/n_{\text{H}}$ at the cloud face. The resulting cloud thickness, L , is determined by the stopping condition, and is related to the neutral hydrogen column density via $L = N_{\text{HI}}/(f_{\text{HI}}n_{\text{H}}) = (N_{\text{HI}}/f_{\text{HI}})(U/n_\gamma)$.

In Figure 12.1, a schematic of a Cloudy model is shown. The model cloud geometry is assumed to be plane parallel with a constant density n_{H} . The ionizing spectrum, J_ν , illuminates one side (the face) of the cloud⁵. The stopping condition is the neutral hydrogen column density N_{HI} . The cloud itself is assumed to have a metallicity of Z/Z_\odot with abundance pattern $\eta_{k,j}$. The ionization parameter, U , provides a fixed ratio between the number density of hydrogen ionizing photons, n_γ , at the cloud face and the hydrogen number density, n_{H} , of the model cloud. If a cosmic UVB is used then n_γ is fixed according to its redshift via Eq. 11.2 and U and the model cloud density are fixed by the product $n_\gamma = n_{\text{H}}U$. The final cloud thickness will be determined by the equilibrium ionization conditions according to $L = N_{\text{HI}}/(f_{\text{HI}}n_{\text{H}}) = (N_{\text{HI}}/f_{\text{HI}})(U/n_\gamma)$, where f_{HI} is the ionization fraction of neutral hydrogen obtained from the model cloud.

As we will discuss below, if the hydrogen number density is assumed to be

⁵ For cosmological absorption line studies, we make the distinction between the observed radiation and the local radiation field, or ionizing spectrum. While the radiation from spectroscopically observed luminous source, such as a distant background quasar, interacts with and is absorbed by the atoms and ions in the gas, it is *not* coupled to the gas ionization conditions; it simply probes them. It is the local radiation field that governs the equilibrium ionization of the gas. If the luminous source is local to the probed gas, then such a distinction is blurred as the source radiation can then be a non-negligible fraction of the local radiation field.

constant throughout the cloud, then homology scaling relationships can be derived that allow small changes in the assumed model input parameters to be explored. This allows flexibility in building model clouds that can then be “scaled” to match observable quantities. In the following, we will assume constant density plane parallel “slab” models.

12.3.2 The Ionization Parameter

The ionizing spectrum and the ionization parameter, $U = n_\gamma/n_{\text{H}}$, can significantly alter the ionization balance in the model cloud. As the ionization parameter is increased, a given model cloud will be more highly ionized. As such, U is a key input quantity to the cloud model, as it sets the relationship between the intensity of the ionizing radiation incident on the cloud in relation to the opacity of the absorbing gas. *Ceteris paribus*, the ionization structure of the model cloud is strongly dependent upon this parameter, as the number density of ionizing photons, n_γ (Eq. 11.2), is sensitive to the intensity and shape of the spectral energy distribution, J_ν ,

$$U = \frac{n_\gamma}{n_{\text{H}}} = \frac{1}{n_{\text{H}}} \frac{4\pi}{c} \int_{\nu_0}^{\infty} \frac{J_\nu}{h\nu} d\nu, \quad (12.27)$$

where $h\nu_0 \geq 13.598$ eV is the ionization threshold energy for hydrogen, and the term $4\pi J_\nu$ is the mean flux density of photons [$\text{erg cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1}$] incident in the cloud face.

If the spectral energy distribution of the incident ionizing radiation is written $J_\nu = J_{\nu_0} f(\nu)$, where the generalized shape functions is $f(\nu)$ and J_{ν_0} is the mean intensity at the hydrogen ionization edge, we can rewrite the ionization parameter as

$$U = \frac{4\pi}{hc} \frac{J_{\nu_0}}{n_{\text{H}}} \int_{\nu_0}^{\infty} \frac{f(\nu)}{\nu} d\nu = \frac{4\pi}{hc} \frac{J_{\nu_0}}{n_{\text{H}}} F \quad (12.28)$$

In the idealized case where the ionizing spectrum is a power law of the form $f(\nu) = (\nu_0/\nu)^\alpha$, we have $F = 1/\alpha$, and obtain

$$U = \frac{4\pi}{hc} \frac{J_{\nu_0}}{n_{\text{H}}} \frac{1}{\alpha}, \quad (12.29)$$

where typical values of α are in the range $1 \leq \alpha \leq 2$ for quasars and AGN.

In general, Eq. 12.28 applies if the form of $f(\nu)$ and the normalization J_{ν_0} are known. The special case of the power-law would more readily apply if the absorbing gas were in the proximity of a quasar.

12.3.3 Ionization Structure

In practice, the only direct constraints on the model cloud properties are provided by matching the observed column densities, $N_{k,j}$. As we will discuss further below, the column density of neutral hydrogen, N_{HI} , is especially desirable for constraining cloud ionization conditions, metallicity, and even estimating the cloud physical size and mass.

Due to the nature of radiative transfer through a gaseous medium, as the incident ionizing spectrum penetrates further into the model cloud, absorption modifies the spectral energy distribution J_ν . Thus, the ionization balance changes with location in the cloud, which is to say that the ionization fractions must be determined as a function of depth. Cloudy iterates toward the equilibrium solution by discretizing the cloud into several zones, each with zone thickness Δs_i . The number of zones depends on the number of ionization fronts, or regions where the ionization fractions change rapidly with depth, as these regions need to be resolved (the optical depths are approximated as being constant in each zone). Furthermore, the zone depths are not necessarily equal thickness. For a given atomic ion k, j , absorption at frequency ν at a given depth, s , in the model cloud depends on the optical depth at that location

$$\tau_\nu(s) = \sum_{k,j} N_{k,j}(s) \sigma_{k,j}(\nu), \quad (12.30)$$

which follows from the definition given in Eq. 7.85 where $\sigma_{k,j}(\nu)$ is the absorption cross section, $N_{k,j}(s)$ is the column density at depth s , and where we have summed over all atoms and ions to obtain the total optical depth. From Eq. 7.86, we have

$$N_{k,j}(s) = \sum_i n_{k,j}^{(i)} \Delta s_i = n_k \sum_i f_{k,j}^{(i)} \Delta s_i = n_{\text{H}} \left(\frac{n_k}{n_{\text{H}}} \right) \sum_i f_{k,j}^{(i)} \Delta s_i, \quad (12.31)$$

where (n_k/n_{H}) is the model cloud abundance of atomic species k (Eq. 11.11) and the sum is over the currently solved zones to depth s .

In Figure 12.2(a), we illustrate the ionization structure for H I, He I, and He II as a function of depth into a model cloud ionized by the cosmic UVB at $z = 1$ with $\log n_{\text{H}} = -2$ and a stopping condition of $\log N_{\text{H}} = 18.5$. For these atoms/ions, the ionization front extends from $0.3 \leq s/L \leq 0.5$, so this is where the zone resolution is highest (i.e., Δs_i is smallest). The ionization cross sections are illustrated in Figure 12.2(b). The transmitted ionizing spectrum as a function of depth is presented in Figure 12.2(c), where it can be seen that the shape is strongly influenced by the cross sectional frequency

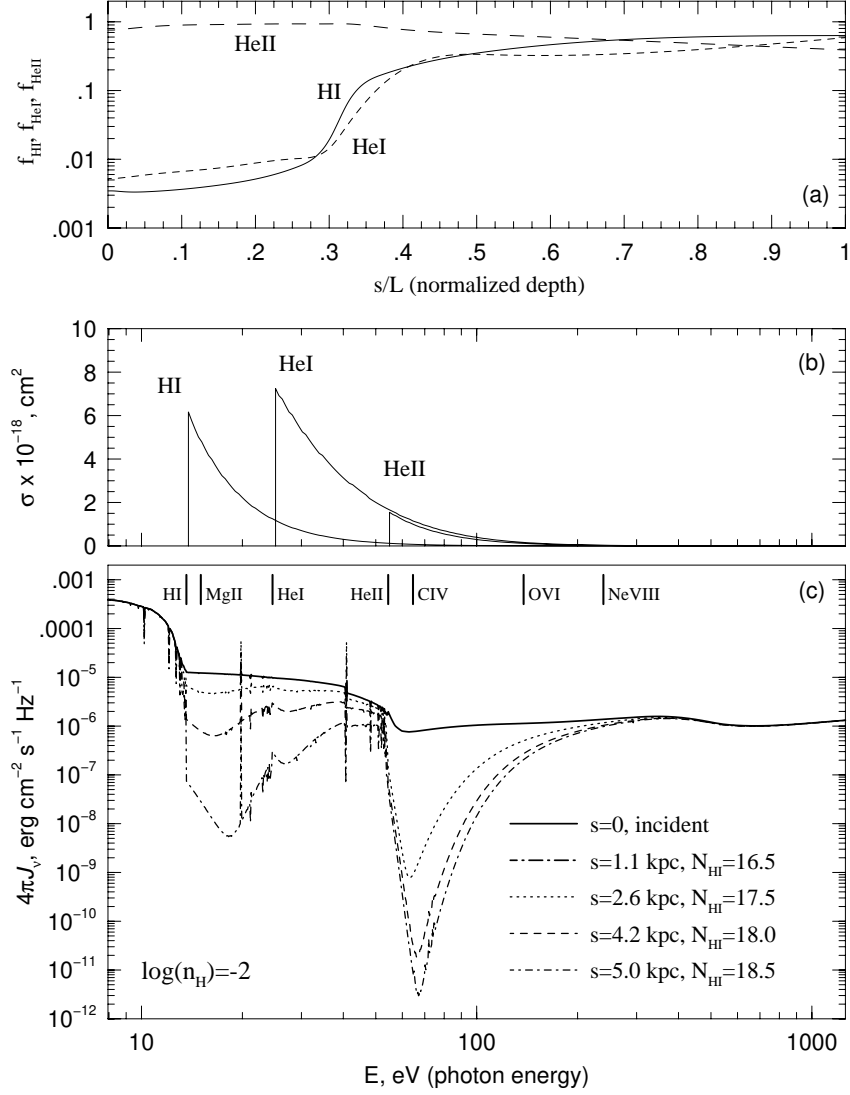


Figure 12.2 (a) The ionization fractions of the most abundant elements HI, HeI, and HeII as a function of normalized depth into the model cloud. (b) The photoionization cross sections of HI, HeI, and HeII as a function of photon energy. (c) The transmitted mean flux density, $4\pi J_\nu$, of the cosmic UVB at $z = 1.0$ as a function of depth, s , into a model cloud with $\log n_{\text{H}} = -2$. The black solid curve is the incident radiation field. As the radiation penetrates deeper into the cloud, HI ionization and HeII ionization remove flux in proportion to their optical depths (product of column densities and absorption cross section) at each depth. As the spectral energy distribution is softened with depth in the model cloud, the ionization levels decrease, which is reflected in the strengthening of the HI ionization edge and the appearance and strengthening of the HeI ionization edge for $s > 2.6$ kpc ($s/L > 0.52$). For reference, the ground-state ionization energy thresholds for HI, HeI, HeII, MgII, CIV, OVI, and NeVIII are shown above the spectrum.

dependence and the magnitude of the absorption is proportional to the column density at depth s . As depth increases into the cloud, absorption by a given atom/ion removes photons from the radiation field that would ionize these given atom/ions at larger depths into the cloud. This phenomenon is called “self-shielding”; this is the radiative transfer physics that gives rise to ionization fronts in the clouds and why the ionizing conditions become more neutral with increasing depth.

From Eqs. 2.91 and 2.88 (also see § 9.3 and Eq. 9.60), we see that $\sigma_{\text{HI}} = 6.304 \times 10^{-18} \text{ cm}^2$ at the ionization threshold of $h\nu_0 = 13.598 \text{ eV}$. Defining the optical depth $\tau_{\text{LL}} = 1$ at the hydrogen ionization edge (Lyman limit) as the value separating optically thin ($\tau_{\text{LL}} < 1$) and optically thick $\tau_{\text{LL}} > 1$ regions of the cloud, we obtain $N_{\text{HI}} = 1/\sigma_{\text{HI}} = 1.586 \times 10^{17} \text{ cm}^{-2}$ as the column density corresponding to $J_{\nu_0}(s)/J_{\nu_0}(0) = e^{-1}$. For optically thin clouds, $\log N_{\text{HI}} \leq 17.2$, there is little to no ionization structure. For $\tau_{\text{LL}} \ll 1$, the ionization fractions are effectively constant throughout the cloud and there is no ionization structure. For an optically thick model cloud with $\log N_{\text{HI}} > 17.2$, a hydrogen ionization front will arise at depth in the model cloud so that the face of the cloud has higher ionization conditions than the “core”. We write “core” in quotes because the model cloud really has a back side that is not illuminated.

Self-shielding by hydrogen and helium reduces the number density of ionizing photons, so n_γ decreasing with depth. In relative terms, a radiation field with smaller (larger) n_γ is called a softer (harder) spectral energy distribution. Due to the frequency dependence of the absorption cross sections for ionization, the highest energy photons have lower probability of participating in an ionization event, and therefore have a longer mean free path through the cloud medium and do not contribute in a significant way to the ionization balance. As illustrated in Figure 12.2(c), photons in the energy range $\sim 13\text{--}40 \text{ eV}$ and $\sim 55\text{--}200 \text{ eV}$ are preferentially absorbed by H I and He II ionization, respectively. As such, atoms/ions with ground-state ionization threshold energies slightly higher than that of H I, such as Mg II (15.03 eV), will experience shielding from H I and exhibit an ionization structure similar to that of H I. Similarly, atoms/ions with ground-state ionization threshold energies somewhat higher than that of He II, such as C IV (64.49 eV), and O VI (138.12 eV), will experience shielding from He II and exhibit an ionization structure similar to that of He II. This phenomenon is evident in the behavior of the mean intensity at the ionization edges of these ions and the behavior in the ionization rates of these ions as shown in Figure 11.8.

12.3.4 Ionization Grids

As previously mentioned, the constraints on the model clouds are the measured column densities, $N_{k,j}$, of various atoms and ions measured in an absorption spectrum. The goal is to use those observed data to infer the absorbing cloud density, metallicity, abundance pattern, and over all ionization conditions. As such, one of the important output quantities of the ionization models are the predicted column densities. From Eq. 12.31, these values are obtained by summing the individual zone column densities over all N_z model cloud zones

$$N_{k,j} = n_H \left(\frac{n_k}{n_H} \right) \sum_{i=1}^{N_z} f_{k,j}^{(i)} \Delta s_i. \quad (12.32)$$

It is also of interest to obtain the average ionization fraction that would yield $f_{k,j} = N_{k,j}/N_k$, where $N_k = \sum_j N_{k,j}$. This average can be simplified to the expression

$$f_{k,j} = \frac{1}{L} \sum_{i=1}^{N_z} f_{k,j}^{(i)} \Delta s_i. \quad (12.33)$$

One common strategy for constraining the chemical and ionization conditions of absorbing clouds by using the observed column densities is to create a grid of model clouds and output their predicted column densities and/or average ionization fractions. This method has been employed since the pioneering work of Bergeron & Stasińska (1986) as applied to quasar absorption line systems. One builds a series of cloud models over a specified range of n_H , U , Z/Z_\odot , abundance pattern, and ionizing radiation field, J_ν . One also specifies grid intervals for each input quantity, such as $\Delta \log n_H = 0.25$, for example. If the cosmic UVB is used, then the redshift is required, and so redshift can become an additional grid dimension.

In Figure 12.3, we illustrate the predicted absorption line column densities for the commonly observed MgII, FeII, CII, CIV, SiII, SiIV, NV, OVI, and NeVIII ions from a grid of 300 model clouds with $Z/Z_\odot = 0.1$ with a solar abundance pattern ionized by the cosmic UVB at $z = 1$. The grid points are $-3.75 \leq \log n_H \leq +1.00$ in steps of $\Delta \log n_H = 0.25$ and $13.5 \leq \log N_{\text{HI}} \leq 20.5$ in steps of $\Delta \log N_{\text{HI}} = 0.5$. Solid curves are lines of constant N_{HI} and dotted curves are lines of constant n_H . For the $z = 1$ cosmic UVB, we have $n_\gamma = 1.2 \times 10^{-5}$ ionizing photons cm^{-3} . From $U = n_\gamma/n_H$, the dotted curves are also lines of constant ionization parameter, where U varies from $\log U = -1.17$ (for $\log n_H = -3.75$), to $\log U = -5.9$ (for $\log n_H = +1.0$).

If one has measured the column density of only a single ion, say N_{MgII} or

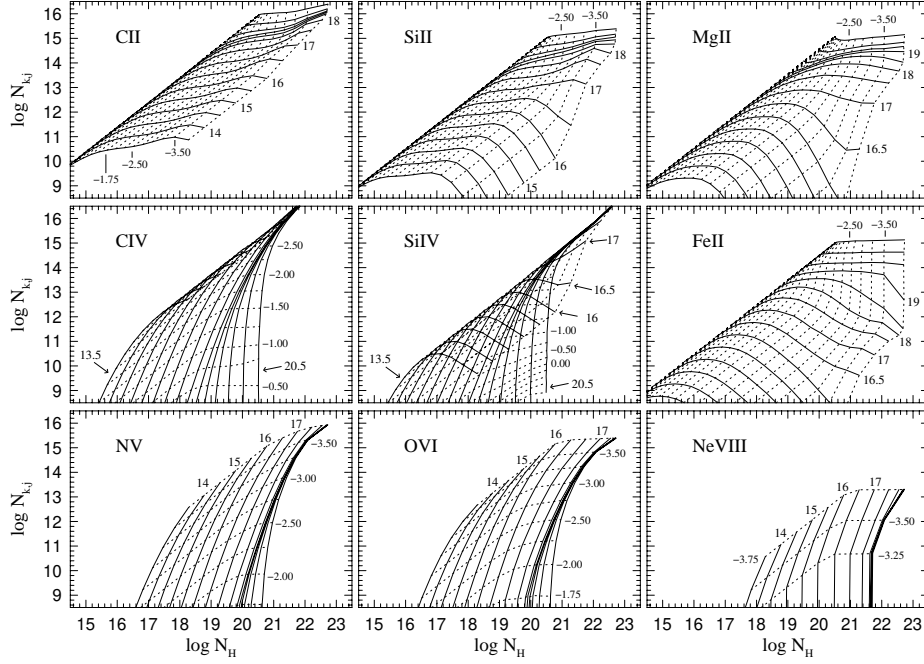


Figure 12.3 A grid of model clouds with varying n_{H} showing N_{CII} , N_{SiII} , N_{MgII} , N_{CIV} , N_{SiIV} , N_{FeII} , N_{NV} , N_{OVI} , and N_{NeVIII} versus N_{H} . For each cloud model, the inputs are a metallicity of $Z/Z_{\odot} = 0.1$ with a solar abundance pattern, a cosmic UVB ionizing spectrum at $z = 1$, the neutral hydrogen density, n_{H} , and a stopping condition at a specified N_{HI} . Solid curves are lines of constant N_{HI} ranging from $\log N_{\text{HI}} = 13.5$ to $\log N_{\text{HI}} = 20.5$ in steps of 0.5. The dotted curves are lines of constant n_{H} ranging from $\log n_{\text{H}} = -3.75$ to $\log n_{\text{H}} = +1.0$ in steps of 0.25.

N_{CIV} , and has no measurement of N_{HI} , then the usefulness of the model cloud grids are limited but still provide insight. For example, consider a measured column density of $\log N_{\text{MgII}} = 12$. Examining Figure 12.3, this would correspond to a horizontal line across the N_{MgII} grid that passes through a range of n_{H} and N_{HI} locations, implying the constraint $16.0 \leq \log N_{\text{HI}} \leq 17.0$, but not constraining n_{H} ; these are the model clouds that, for the assumed input parameters, are allowed by the measured MgII column density. The $\log N_{\text{MgII}} = 12$ line crosses the $\log N_{\text{HI}} = 16.5$ curve at $\log n_{\text{H}} \simeq -2.5$. Since n_{H} is decreasing (U increasing) from left to right, we can infer $\log N_{\text{HI}} \geq 16.5$ for $\log n_{\text{H}} \leq -2.5$ and $\log N_{\text{HI}} \leq 16.5$ for $\log n_{\text{H}} \geq -2.5$. If one also has a robust measurement of N_{HI} and it lies within the range of the allowed model clouds, then one can further constrain n_{H} . If, for instance, if we measured $\log N_{\text{HI}} = 16.75$, then we can constrain $\log n_{\text{H}} \sim -3.0$.

The above example provides meaningful constraints only to the degree that the assumed model cloud input parameters apply to the real world observed absorbing cloud. A grid built using a different metallicity or ionizing spectrum will result in different predicted column densities for a given n_{H} and N_{HI} grid location. Consider a different metallicity. In general, if the entire grid were built using a metallicity 0.5 dex lower (higher), then, to a very good approximation, the predicted column densities at all $n_{\text{H}}, N_{\text{HI}}$ points on the model cloud grid would be 0.5 dex smaller (larger). Since changing the metallicity does not change the ratio $N_{\text{HI}}/N_{\text{H}}$, this can be envisioned as a shifting of the grids downward (upward) by 0.5 along the vertical axes in Figure 12.3 (this would apply to all metal ions)⁶.

If we had measured $\log N_{\text{MgII}} = 12$ with $\log N_{\text{HI}} = 18.5$ then we find that the neutral hydrogen column density lies above the allowed range $16.0 \leq \log N_{\text{HI}} \leq 17.0$ of the model cloud grid in Figure 12.3 as constrained by $\log N_{\text{MgII}} = 12$. A grid with a lower abundance of magnesium would require larger hydrogen column densities to give rise to $\log N_{\text{MgII}} = 12$ (for a given n_{H} , a greater model cloud thickness is required to yield the same metal ion column density as a higher metallicity model cloud). As a lower metallicity grid would be equivalent to shifting the grid downward, we can visualize that a measurement of $\log N_{\text{MgII}} = 12$ with $\log N_{\text{HI}} = 18.5$ constrains the cloud model to have $\log n_{\text{H}} \sim -3.0$ with a metallicity 2 dex lower than the model grid presented Figure 12.3, or $\log Z/Z_{\odot} = -3$. Note that the hydrogen density is 1 dex lower than if the metallicity were $\log Z/Z_{\odot} = -1$, illustrating that the model cloud constrained with this different metallicity would have a different n_{H} (and therefore U) than the model shown in the Figure 12.3.

The above example highlights that obtaining a measured value of N_{HI} is critical to the exploiting the power for ionization modeling. It also highlights the limits of having only a single measured metal ion column density. Though one could in principle apply the above example for any one of the metal ions included in the model grid, the accuracy to which one can constrain the cloud model properties can vary from ion to ion. Furthermore, given that, in the real world, not only is the metallicity unknown, but the relative abundance pattern of the observed absorbing gas and the shape and intensity of the ionizing spectrum are unknown; there remains unexplored parameter space for the model cloud properties that cannot be explored with a single measured metal ion column density. Thus, it is key to measured col-

⁶ This statement holds if the net cooling rate and heating rate in the gas remains the same, as the gas temperature governs the collisional ionization and recombination rates. Only at high metallicities does the cooling function change enough with metallicity to weaken this statement.

umn densities from a range of atomic species covering a range of ionization levels in order to further constrain the cloud model physical chemical and ionization conditions that would match observations.

12.3.5 Properties of Ionization Grids

Over the decades, countless researchers have employed model cloud grids by using several ions simultaneously to constrain the allowed n_{H} , J_{ν} , U , Z/Z_{\odot} and relative abundances parameter space (for early and explanatory works, see e.g., Bergeron & Stasińska, 1986; Steidel, 1990b; Bergeron et al., 1994; Dittmann & Koeppen, 1995). The importance of constraining the ionization parameter cannot be overstated, as this single quantity more than any other correlates with the ionization fractions.

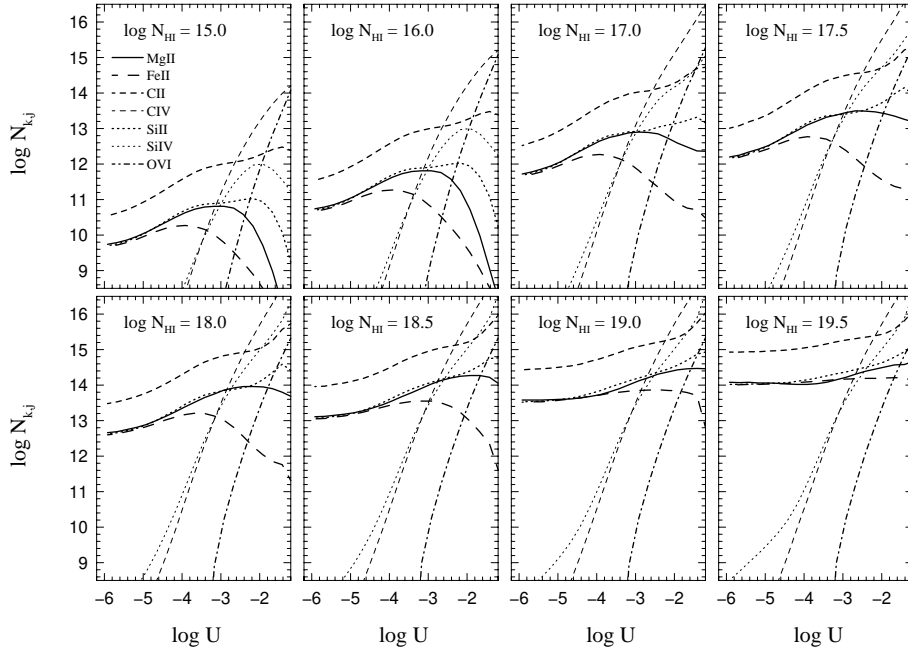


Figure 12.4 A grid of Cloudy models with varying n_{H} showing N_{MgII} , N_{FeII} , N_{CII} , N_{CIV} , N_{SiII} , N_{SiIV} , and N_{OVI} versus $U = n_{\gamma}/n_{\text{H}}$. The input parameters of the models are discussed in § 12.3.4. If one has measured column density for most of these ions, the ionization parameter and neutral hydrogen column density can be constrained, though there is some sensitivity to metallicity. Again, there is a N_{H} metallicity degeneracy, but the ion column density ratios can constrain U quite well.

In Figure 12.4, we show the predicted metal ion column densities for

commonly observed ions as a function of ionization parameter. Each panel applies to a model cloud with fixed N_{HI} . The input parameters are the same as employed for the grids illustrated in Figure 12.3. This particular presentation of the model cloud grids highlights several important points:

- For a fixed metallicity, as N_{HI} increases, the metal ion column densities increase in proportion across all U . This is primarily due to the fact that, for a given U , cloud thickness increases as N_{H} increases.
- For a given N_{HI} , the ratio of the column densities is unique at each U . This behavior is key to constraining U .
- For very optically thin clouds ($\log N_{\text{HI}} \leq 16$), the column densities of ions with low ionization thresholds (near or below the hydrogen threshold of 13.598 eV) drop rapidly with increasing U , whereas those with much higher ionization thresholds increase rapidly.
- For very optically thick clouds ($\log N_{\text{HI}} \geq 18$), the column densities of ions with low ionization thresholds remain fairly constant with increasing U , whereas those with much higher ionization thresholds increase rapidly.

The behavior of the various ions with varying U results from the development of ionization structure in the model clouds as N_{HI} , and therefore the cloud thickness, becomes larger. In optically thin clouds there is no ionization structure; the ionization fractions are constant with depth into the model cloud. Therefore, as U is increased, the column density directly reflects the ionization fraction of the ion, such that the U at which the column density maximizes is the U at which the ionization fraction maximizes. For low ionization ions, as U is increased the neutral atoms become progressively ionized, thus increasing the low ionization ion column densities. Then, as U is increased above their peak ionization fraction, the low ionization ions are ionized to higher ionization stages, thus decreasing their column densities. For high ionization ions, a larger U is required to raise the ionization stages enough to yield larger column densities.

In optically thick clouds, shielding by hydrogen and helium results in ionization fronts at various depths in the model clouds for each ion (see Figure 12.2(a)). With increasing physical depth into the cloud, the optical depths for HI and HeII increase, the ionizing radiation field becomes softer, and the ionization conditions become more neutral for the low ionization ions. As N_{HI} increases, the physical depth of this “neutral core” increases (the physical depths at which the ionization fronts exist do not change, but additional HI column increases the total physical depth of the model cloud). If we consider higher ionization parameters, such as $\log U > -2$, as N_{HI} increases, the physical depth of the neutral cloud core increases yielding

higher column densities of the low ionization ions. As such, the column densities of low ionization ions in optically thick clouds arise primarily in the self-shielded regions of the cloud models.

For the higher ionization ions, they are primarily affected by the shielding due to HeII ionization. As can be seen in Figure 12.2(c), for a fixed n_{H} (and therefore U) the CIV ion, having an ionization threshold just above that of HeII, would have an ionization structure in the cloud that changes more sensitively to HeII shielding as N_{HI} is increased from optically thin to optically thick values than would OVI. This behavior can be seen in Figure 12.4. As $\log N_{\text{HI}}$ is increased from 16.0 to 17.5, the $N_{\text{CIV}}/N_{\text{OVI}}$ ratio converges to unity as U increases, showing that the ionization fraction of CIV is decreasing relative to that of OVI. Since, as is apparent in Figure 12.2(c), the shielding from HeII changes very little as $\log N_{\text{HI}}$ is increased above 18.0, we see that the relative ionization of CIV to OVI is a constant in these highly optically thick clouds in Figure 12.4.

As mentioned in § 12.3.4, if the input metallicity of the model grid were increased (decreased), the predicted column densities would increase (decrease) in direct proportion. Graphically, this is equivalent to the column density curves for all ions in all cloud models being shifted vertically in proportion to metallicity in Figure 12.4. Similarly, if the input relative abundance pattern is changed, then the modified ions would have their curves shifted; for example, say magnesium is increased +0.5 dex over its solar abundance ratio, then the MgII column density curves would be shifted +0.5 dex upward relative to the column density curves of all other ions.

Because relative abundance variations change the ratio of column densities from different ions as a function of U , they are particularly important to consider. In the absence of knowing the abundance pattern of the observed absorbing gas cloud, we have the problem that the inferred or constrained value of U from the observed column densities may be incorrect. Again, the greater the number of observed ion column densities, the better one can break such degeneracies in the models and constrain the chemical and ionization conditions.

The above mentioned scaling of the predicted column densities with metallicity and relative abundance can be express as

$$\log N_{k,j} = \log N_{k,j}^{(m)} + \Delta Z + \Delta[X_k/H], \quad (12.34)$$

where $\Delta[X_k/H] = [X_k/H] - [X_k/H]^{(m)}$ is the scaled logarithmic abundance of ion k in solar units (see Eq. 11.21), $\Delta Z = \log(Z/Z_{\odot}) - \log(Z/Z_{\odot})^{(m)}$ is the logarithm scaled metallicity of the model cloud, where Z/Z_{\odot} is obtained via Eq. 11.23, and the superscript “(m)” represented the input value used to

create the model grid. Since there is no ionization structure in optically thin model clouds, there is an additional simple scaling relationship that safely applies for model clouds with $\log N_{\text{HI}} < 16$ (i.e., $\tau_{\text{LL}} \leq 0.06$),

$$\log N_{k,j} = \log N_{k,j}^{(m)} + \Delta N_{\text{HI}} + \Delta Z + \Delta[X_k/\text{H}], \quad (12.35)$$

where $\Delta N_{\text{HI}} = \log N_{\text{HI}} - \log N_{\text{HI}}^{(m)}$ is the logarithm scaled neutral hydrogen column density relative to the value used to create the model grid. This scaling can be immediately seen by visual comparison of the predicted ion column density curves for the $\log N_{\text{HI}} = 15$ and $\log N_{\text{HI}} = 16$ model clouds in Figure 12.4, for which we would apply $\Delta N_{\text{HI}} = +1$ or $\Delta Z = +1$ to the $\log N_{\text{HI}} = 15$ model cloud to recover all $N_{k,j}$ curves for the $\log N_{\text{HI}} = 16$ model cloud. Conversely, we could apply $\Delta N_{\text{HI}} = -1$ or $\Delta Z = -1$ to the $\log N_{\text{HI}} = 16$ model cloud to recover all $N_{k,j}$ curves for the $\log N_{\text{HI}} = 15$ model cloud. As such, *ceteris paribus*, for all cloud models there is a scaling degeneracy between $N_{k,j}$ and Z/Z_{\odot} for all ions and $[X_k/\text{H}]$ for ion k , whereas for optically thin clouds there is an additional scaling degeneracy between $N_{k,j}$ and N_{HI} for all ions.

12.3.6 Constraining Metallicity and Abundances

As we will show here, at a minimum, it is critical to have a measurement of the neutral hydrogen column density, N_{HI} and a well constrained ionization parameter, U , in order to confidently constrain the metallicity and/or abundance pattern of a given atomic species from ionization modeling. If N_{HI} is measured and multiple metal ion column densities can be leveraged to constrain the cloud model ionization parameter, then model-based ionization corrections (defined below) can be used to convert measured column densities of various ionization stages into abundances. However, even in the absence of knowing N_{HI} , there are techniques, as we describe in this section, that allow varying degrees of insight to be inferred, depending on the quality of the observed data.

To relate the model clouds to the observed data, consider the relationship between the measured N_{HI} , the model cloud hydrogen number density, n_{H} , and the thickness of the model cloud, L ,

$$N_{\text{HI}} = n_{\text{H}} L = f_{\text{HI}} n_{\text{H}} L. \quad (12.36)$$

The identical relationship follows for the column density of the measured metal ion, $N_{k,j} = n_{k,j} L = f_{k,j} n_{\text{H}} L$ for atomic species k in ionization level j .

Equating via the model cloud thickness, we have

$$\frac{N_{k,j}}{N_{\text{HI}}} = \frac{f_{k,j}}{f_{\text{HI}}} \frac{n_k}{n_{\text{H}}}, \quad (12.37)$$

where observed quantities are on the left hand side and model cloud quantities are on the right hand side. The ionization fractions are obtained by constraining the ionization parameter using the measured metal ion column densities whereas the ratio of the number densities was an input to the cloud model. It is adjustment of the input ratio n_k/n_{H} to match the observed column densities for a given U that provides insight into the metallicity and the relative abundance of atomic species k .

From Eq. 12.37, it is apparent how important the ionization fractions from the model cloud are in anchoring this adjustment. Converting to logarithms,

$$\log N_{k,j} - \log N_{\text{HI}} = \log (f_{k,j}/f_{\text{HI}}) + \log (n_{k,j}/n_{\text{H}}), \quad (12.38)$$

The first term on the right hand side of Eq. 12.38 is referred to as the ionization correction, $\text{IC}_{k,j} = \log(f_{k,j}/f_{\text{HI}})$, and it is the key quantity in abundance measurements. Note that the ionization correction applies for a fixed value of N_{HI} and U , i.e., $\text{IC}_{k,j} = \text{IC}_{k,j}(N_{\text{HI}}, U)$. Recall that if the model clouds are subjected to the cosmic UVB, then n_{γ} is known, so $U \propto 1/n_{\text{H}}$, and we can then write $\text{IC}_{k,j} = \text{IC}_{k,j}(N_{\text{HI}}, n_{\text{H}})$ for this particular scenario.

For the cloud model, we can use Eq. 12.38 and apply the logarithmic abundance $[\text{X}_k/\text{H}]$ as defined in Eq. 11.21, and obtain the relation,

$$[\text{X}_k/\text{H}] = \log N_{k,j} - \log N_{\text{HI}} - \text{IC}_{k,j} - \log (n_{k,j}/n_{\text{H}})_{\odot}. \quad (12.39)$$

Thus, armed with the ionization correction from the model cloud, the measured column density of ion k, j , and the measured column density of neutral hydrogen, we can estimate $[\text{X}_k/\text{H}]$ for the absorbing gas.

In Figure 12.5, we present a grid of ionization corrections for the commonly observed metal ions MgII, CIV, and OVI. These are based on the grids presented in § 12.3.4 for 300 model clouds. Note that the ionization correction for MgII is in fairly narrow range of $0 \leq \text{IC}_{\text{MgII}} \leq 1$ for a wide range of model cloud parameter space. This narrow range of ionization corrections is desirable, as smaller corrections applied in Eq. 12.39 provide a greater confidence that the choice of the model cloud from which the corrections are drawn is not a major source of uncertainty in estimated $[\text{X}_k/\text{H}]$. Small ionization corrections (i.e., $\text{IC}_{k,j} \leq \pm 1$) are typical of low ionization threshold metal ions such as MgII, CII, SiII, etc. However, as shown in Figure 12.5, the ionization corrections for ions such as CIV and OVI can range over ten orders of magnitude! Their value is highly sensitive to the model cloud n_{H} .

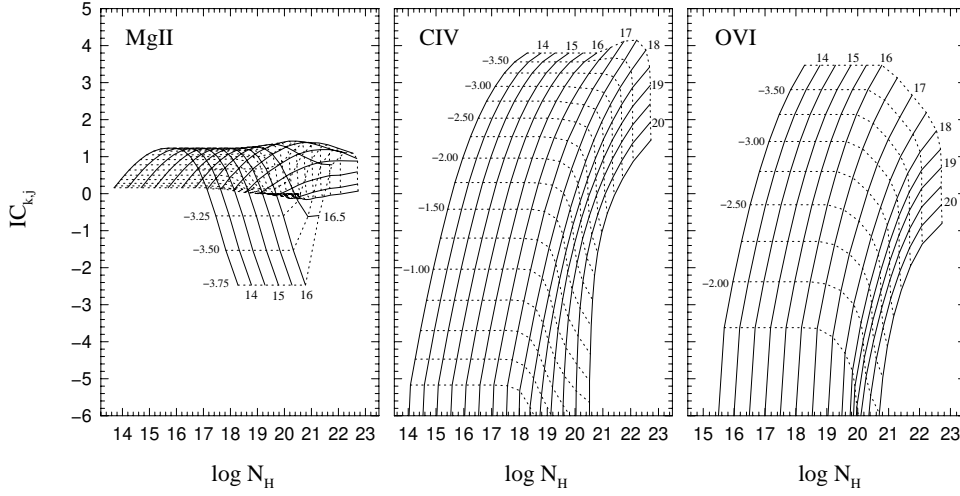


Figure 12.5 A grid of Cloudy models with varying n_{H} showing the ionization corrections, $\text{IC}_{\text{k,j}}$, for MgII, CIV, and OVI versus N_{H} . The cloud model inputs and range of n_{H} , U , and N_{HI} are the same as shown in Figure 12.3. Solid curves are lines of constant N_{HI} and dotted curves are lines of constant n_{H} (and constant U).

Further, note that for optically thin model clouds, the ionization corrections for a given n_{H} are degenerate with N_{HI} (i.e., the curves are horizontal, or flat as a function of neutral hydrogen column density). We emphasize that the ionization corrections presented in Figure 12.5 are specific to the input model cloud parameters, in particular, the ionizing radiation field, J_{ν} , which is taken at $z = 1$ for the presented model grid.

Due to the N_{HI} dependence of the ionization correction in the optically thick regime, there is an important caveat in the application of Eq. 12.39 in the case of optically thick model clouds. For measured values of $\log N_{\text{HI}} \geq 17.2$, it is imperative that the cloud model neutral hydrogen column match the observed value within uncertainties in order to obtain accurate ionization corrections. In practice, it is recommended that this threshold value be reduced to $\log N_{\text{HI}} \geq 16$, to ensure as little ionization structure as possible in the model cloud adopted for obtaining the ionization correction. On the other hand, in the case of optically thin clouds with measured values $\log N_{\text{HI}} \leq 16.5$, the ionization correction is independent of neutral hydrogen column density. Thus, in the optically thin regime, one can create a single cloud model with, for example, $\log N_{\text{HI}} = 16.0$, from which the ionization corrections can be obtained to apply Eq. 12.39.

In Figure 12.6, we illustrate a simple idealized case to show how observed data are employed to constrain the gas chemical and ionization conditions

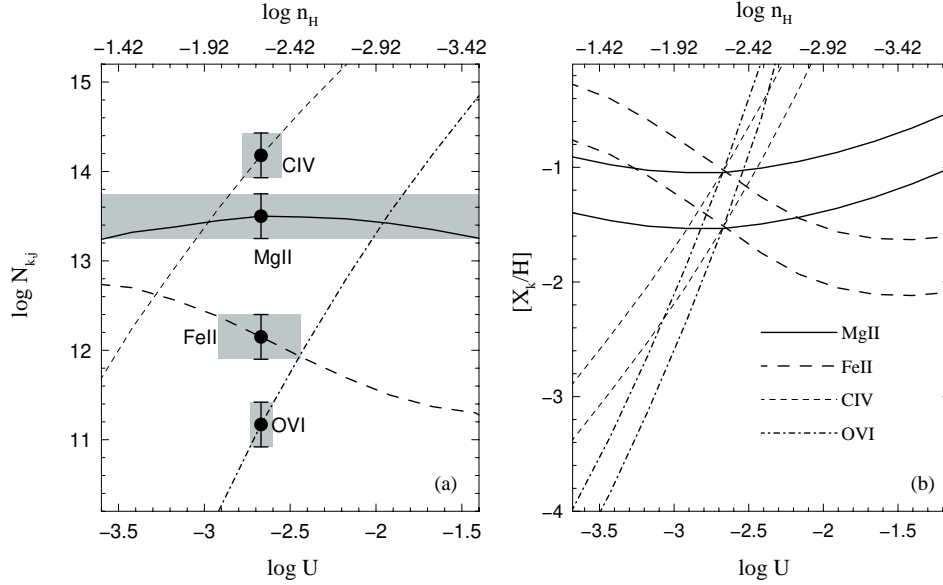


Figure 12.6 (a) Metal ion column density, $\log N_{k,j}$ versus ionization parameter, $\log U$, for a model cloud fitted to idealized measured MgII, FeII, CIV, and OVI column densities. In the absence of a measured N_{HI} , the best-fit cloud model can be determined by least square minimization using interpolation of the model grid across U and N_{HI} with a free parameter for scaling the metallicity. If N_{HI} is measured, then the minimization omits interpolation in the N_{HI} direction. The shaded regions provide the range of ionization parameter consistent within the measurement uncertainties of the observed column densities. The best-fit model has $N_{\text{HI}} = 17.7 \pm 0.3$, $\log U = -2.67 \pm 0.15$, and $\log Z/Z_{\odot} = -1.21 \pm 0.25$. (b) $[X_k/H]$ curves versus $\log U$, where the curves are computed from Eq. 12.39 for a constant $\log n_{\text{H}} = 17.7$; the upper curves account for the upper 1σ uncertainties in the measured column densities and the lower curves account for the lower 1σ uncertainty. Note that the intersection of all measured ions indicates that there is an single unambiguous best-fit model. This is an idealized case.

from model grids. In this case, we have measured column densities for four ions, MgII, FeII, CIV, and OVI. We assume that they each have uncertainties of 0.25 dex. Figure 12.6(a) shows the best-fit model cloud to the four column densities. The shaded areas indicate the range over which each ion constrains the ionization parameter based upon the 1σ uncertainties in the column densities. Note that the constraints are strongly dependent on both how large the measurement uncertainties are and how rapidly the ionization fraction of the ion changes with ionization parameter.

This best-fit cloud model, was determined using a χ^2 least-squares minimization routine to match the observed metal ion column densities assuming

N_{HI} is not known (measured). The model grid is interpolated across $\log U$ and $\log N_{\text{HI}}$ and allows metallicity, Z/Z_{\odot} , to be optimized as a free parameter based upon the scaling law given in Eq. 12.34, while assuming no relative abundance variation, i.e., $\Delta[X_k/H] = 0$. If N_{HI} is a measured quantity, then this quantity is fixed and the interpolation is performed only for $\log U$. In this case, it is more reasonable to explore variations in the relative abundances of the metal ions. But, given complicated degeneracies between ionization parameter and variations in relative abundances, it is difficult to constrain $\Delta[X_k/H]$ if N_{HI} is not measured. The uncertainties in the parameters of the best-fit model cloud are determined by the uncertainties in the measured column densities. For the example presented here, the best-fit model cloud parameters are $\log N_{\text{HI}} = 17.7$, $\log U = -2.67 \pm 0.15$, and $\log Z/Z_{\odot} = -1.21 \pm 0.25$.

An additional method for constraining the metallicity and ionization parameter, suitable when in cases where N_{HI} is measured, is shown in Figure 12.6(b). Presented are $[X_k/H]$ curves computed from Eq. 12.39 as a function of ionization parameter for a model with $\log N_{\text{HI}} = 17.7$. Determining the common locus or intersection of these curves provides an additional method for estimating the metallicity and ionization parameter when N_{HI} is a known quantity. Because our idealized observed data for this example observation are based upon a solar abundance pattern, the curves overlap with no $\Delta[X_k/H]$ adjustments. In reality, such adjustments may be required, and knowing which adjustment are appropriate should be guided by a solid understanding of interstellar and galactic nucleosynthetic abundance patterns and their evolution (e.g., Lauroesch et al., 1996; Henry & Worthey, 1999; Hamann & Ferland, 1999).

12.3.7 Estimating Cloud Sizes and Masses

As we have discussed, ionization models exhibit behaviors that allow various scaling relations, such as the scaling of column density with metallicity and abundance (Eq. 12.34) and with N_{HI} (Eq. 12.35). Such relationships have been utilized since the pioneering work of Bergeron & Stasińska (1986) and others (e.g., Steidel, 1990b; Dittmann & Koeppen, 1995).

In addition to observationally constraining absorbing cloud column chemical and ionization conditions, we often wish to make rough physical size and mass estimates. These quantities depend on the cloud hydrogen number and mass density, respectively. Since cosmological and galactic environments can vary greatly in their local sources of ionizing radiation, the normalization and shape of the ionizing spectral energy distribution is not strongly

constrained. As such, it is sound practice to explore how variations in the ionizing spectrum, and therefore ionization conditions, affect our scientific inferences based on ionization modeling.

Since our formalism is that the model clouds have constant density, n_{H} , we can derive simple functional relationships between cloud density, physical size, mass, and the unknown ionizing radiation field, J_{ν} , and ionization parameter, U . The utility of these relationships is that, in the absence of having strong constraints on J_{ν} , instead of building several model grids based on a range of ionizing spectral energy distributions, one can more simply evaluate the integral in Eq. 12.28 to obtain F and explore how the changing shape in the ionizing spectrum affects the model cloud densities, and therefore sizes and masses. There is moderate redshift evolution in J_{ν_0} , which can also be explored.

In addition, one can examine how small variations in N_{H} , hydrogen ionization conditions, f_{H} , and ionization parameter, U , modulate model cloud density, size, and mass. To express fractional changes in these quantities, one applies

$$\frac{dy}{y} = \sum_i \left(\frac{\partial \log y}{\partial \log x} \right) \frac{dx}{x}, \quad (12.40)$$

where y is the dependent variable and x_i is an independent variable. The derivatives $\partial \log y / \partial \log x = (dy/y)/(dx/x)$ are the power-law dependences of y on x , i.e., the local slope of $\log y$ vs. $\log x$.

In what follows, we draw on the methods inspired by Steidel (1990b). We also direct the reader to the works of Schaye (2001), who worked out the scaling relations and fiducial cloud densities, column densities, sizes, and masses under the conditions of hydrodynamic equilibrium in the intergalactic medium. We also mention representative works by Rees & Ostriker (1977), Mo & Miralda-Escude (1996), Maller & Bullock (2004), and Thompson et al. (2016), who worked out cloud properties and cloud survivability timescales in a two-phase circumgalactic medium.

12.3.7.1 Cloud Density

From the definition of the ionization parameter, we see that the hydrogen density of a model cloud scales as $n_{\text{H}} = n_{\gamma}/U$. In § 12.3.2, we wrote the ionizing radiation field in the form $J_{\nu} = J_{\nu_0}f(\nu)$, from which we derived the relationship between U and n_{H} given in Eq. 12.28, which rearranged gives

$$n_{\text{H}} = \frac{4\pi J_{\nu_0}}{hc U} F. \quad (12.41)$$

where F is the integral over the shape function, $f(\nu)$, and is of order unity.

For the cosmic UVB, though there is redshift evolution in the value of J_{ν_0} and F , typical values of J_{ν_0} are on the order of $10^{-22} \text{ erg cm}^{-2} \text{ s}^{-1} \text{ Hz}^{-1}$ (see Fig 11.8). Typical values of ionization parameters are roughly distributed over an order of magnitude around $U \simeq 10^{-2}$. We have

$$n_{\text{H}} \simeq 0.063 \frac{J_{-22}}{U_{-2}} F \text{ cm}^{-3}, \quad (12.42)$$

where we define $J_{\nu_0} = J_{-22} \times 10^{-22}$ and $U = U_{-2} \times 10^{-2}$. Since F is of order unity, we see that typical densities of photoionized clouds in environments subject to the cosmic UVB lie in the approximate range $10^{-3} \leq n_{\text{H}} \leq 10^{-1} \text{ cm}^{-3}$ and that density is higher for lower ionization conditions and vice versa.

From Eq. 12.40, the fractional change in the density local to the values n_{H} , J_{ν_0} , F , and U , is

$$\frac{dn_{\text{H}}}{n_{\text{H}}} = \frac{dJ_{\nu_0}}{J_{\nu_0}} + \frac{dF}{F} - \frac{dU}{U}. \quad (12.43)$$

12.3.7.2 Cloud Sizes

The size of the model cloud refers to the thickness, L , which observationally corresponds to the length of the line of sight through the cloud. This thickness is given by $L = N_{\text{H}}/n_{\text{H}}$, where N_{H} is the total hydrogen column density. Adopting the model cloud average ionization fraction, $f_{\text{HI}} = N_{\text{HI}}/N_{\text{H}}$ (see Eq. 12.33), and substituting Eq. 12.41 into L yields

$$L = \frac{hc}{4\pi} \frac{N_{\text{HI}}}{f_{\text{HI}}} \frac{U}{J_{\nu_0}} \frac{1}{F}. \quad (12.44)$$

For a cloud with an approximately unity optical depth at the hydrogen ionization threshold energy (i.e., a Lyman limit cloud), the neutral hydrogen column density is $N_{\text{HI}} \simeq 10^{17} \text{ cm}^{-2}$. Assuming an ionization fraction of 10^{-2} typical of $\tau_{\text{LL}} \simeq 1$ model clouds, we obtain

$$L \simeq 52 \frac{N_{17}}{f_{-2}} \frac{U_{-2}}{J_{-22}} \frac{1}{F} \text{ pc}, \quad (12.45)$$

where $N_{\text{HI}} = N_{17} \times 10^{17}$, and where $f_{\text{HI}} = f_{-2} \times 10^{-2}$. Thus, the typical line of sight distance through an absorbing cloud will be on the order of 5 pc to 500 pc.

From Eq. 12.40, the fractional change in the cloud size local to the values N_{HI} , f_{HI} , J_{ν_0} , F , and U , is

$$\frac{dL}{L} = \left(\frac{dN_{\text{HI}}}{N_{\text{HI}}} - \frac{df_{\text{HI}}}{f_{\text{HI}}} \right) + \frac{dU}{U} - \left(\frac{dJ_{\nu_0}}{J_{\nu_0}} + \frac{dF}{F} \right). \quad (12.46)$$

12.3.7.3 Cloud Masses

To infer the mass of a model cloud, a cloud geometry must be assumed. The model cloud geometry is a plane parallel slab, but to estimate masses, it is common to assume a spherical cloud. Even though this is a grand simplification, it provides a sensible expression for mass estimates. We begin with

$$M = \frac{4\pi}{3} \langle \rho \rangle R^3 \quad (12.47)$$

where $\langle \rho \rangle$ is the mean cloud mass density. Since model clouds have constant n_{H} , we have $\langle \rho \rangle = (A_{\text{H}} m_a / x_{\text{H}}) n_{\text{H}}$, where $A_{\text{H}} = 1.0079$ is the atomic weight of hydrogen in atomic mass units, m_a , and x_{H} is the mass fraction of hydrogen. Adopting $R \simeq L/2$ as an approximation for the cloud radius, we obtain,

$$M = \frac{4\pi}{3} \frac{A_{\text{k}} m_a}{x_{\text{H}}} n_{\text{H}} (L/2)^3. \quad (12.48)$$

Substituting Eqs. 12.41 for n_{H} and 12.44 for L into M gives

$$M = \frac{h^2 c^2}{96\pi} \frac{A_{\text{k}} m_a}{x_{\text{H}}} \frac{N_{\text{HI}}^3}{f_{\text{HI}}^3} \frac{U^2}{J_{\nu_0}^2} \frac{1}{F^2}. \quad (12.49)$$

For the above assumed typical ranges of values, the cloud mass is on the order of $10\text{--}10^5 M_{\odot}$,

$$M \simeq 150 \frac{U_{-2}^2}{J_{-22}^2} \frac{N_{17}^3}{f_{-2}^3} \frac{1}{(x_{\text{H}}/X_{\odot}) F^2} M_{\odot}, \quad (12.50)$$

where $X_{\odot} = 0.7381$ is the solar hydrogen mass fraction. As metallicity decreases, the ratio x_{H}/X_{\odot} increases, so the masses of lower metallicity clouds will be smaller than those of higher metallicity clouds. For $N_{\text{HI}} \simeq 10^{17} \text{ cm}^{-2}$, the scaling predicts cloud mass on the order $10^2 \leq M \leq 10^3 M_{\odot}$. Note, however, that the mass is highly sensitive to N_{HI} , so an order of magnitude change in the neutral hydrogen column density yields a factor of 1000 change in M . Thus, cloud masses are predicted to be on the order of 10 pc for $N_{\text{HI}} \simeq 10^{16} \text{ cm}^{-2}$, assuming they have the average conditions adopted here.

From Eq. 12.40, the fractional change in the cloud mass local to the values N_{HI} , f_{HI} , J_{ν_0} , F , U , and x_{H} is

$$\frac{dM}{M} = 3 \left(\frac{dN_{\text{HI}}}{N_{\text{HI}}} - \frac{df_{\text{HI}}}{f_{\text{HI}}} \right) + 2 \left(\frac{dU}{U} - \frac{dJ_{\nu_0}}{J_{\nu_0}} - \frac{dF}{F} \right) - \frac{dx_{\text{H}}}{x_{\text{H}}}. \quad (12.51)$$

12.3.8 Caveats and Complications of Ionization Modeling

In practice, how one determines the best-fit model cloud or range of model clouds is ostensibly a black art. There are as many steps to a “solution” as

there are absorption-line researchers. An ideal goal would be to develop a single definitive objective method that can effectively search a vast model grid for virtually every realization of observed data.

In practice, the data available for a given absorption system are limited. N_{HI} may or may not have been captured by the limited spectral region of the spectrograph used to obtain the data, and if it is observed it may have a large measurement uncertainty. Alternatively, key metal ion column densities may not be observed, or the data may be very noisy and only upper limits are available.

The point is that applying the data to constrain the chemical and ionization conditions of the observed gas using ionization models is most effective when physical insight and intuition is applied. Though some essential basic physical intuition was described in this chapter, there are deeper considerations from the ideal scenarios presented. For one, the above discussion has been based on the assumption that the measured metal ion column densities arise from a single phase of gas. For gas dominated by photoionization, low ionization threshold ions arise in cooler, higher density gas, whereas higher ionization threshold ions arise in warmer, lower density gas. If faced with a set of measurements from multiple gas phases, one will quickly find that a single model cloud, even allowing for reasonable abundance variations, cannot reproduce the data with a single ionization parameter or metallicity. One technique to tackle such data is to model each gas phase separately.

However, if a multiphase scenario applies, the HI absorption will reflect the sum of N_{HI} from the combined gas phases. Estimating the metallicity of each phase strongly depends on knowing what fraction of the neutral hydrogen absorption arises in each phase. It is nearly impossible to objectively determine the relative contribution. In practice, one must rely on a physical intuition that is a combination of the absorption line signatures themselves, astrophysical insight, and a thorough understanding of how ionization models behave and their limitations (for an excellent example, see Appendix A of Muzahid et al., 2015).

TODO: TALK ABOUT Non-CIE and Non-PIE models. Show differences.

References

- Abramowitz, M., & Stegun, I. A. 1972, *Handbook of Mathematical Functions*, New York: Dover Publications
- Aldenius, M., Johansson, S., & Murphy, M. T. 2006, Accurate laboratory ultraviolet wavelengths for quasar absorption-line constraints on varying fundamental constants, *Monthly Notices of the Royal Astronomical Society*, **370**, 444
- Aldrovandi, S. M. V., & Pequignot, D. 1973, Radiative and dielectronic recombination coefficients for complex ions, *Astronomy and Astrophysics*, **25**, 137
- Altun, Z., Yumak, A., Yavuz, I., Badnell, N. R., Loch, S. D., & Pindzola, M. S. 2007, Dielectronic recombination data for dynamic finite-density plasmas. XIII. The magnesium isoelectronic sequence, *Astronomy and Astrophysics*, **474**, 1051
- Aracil, B., Petitjean, P., Smette, A., et al. 2002, HST STIS observations of four QSO pairs, *Astronomy and Astrophysics*, **391**, 1
- Arnaud M., & J. Raymond R. 1992, Iron ionization and recombination rates and ionization equilibrium, *The Astrophysical Journal*, **398**, 394
- Arnaud, M., & Rothenflug, R. 1985, An updated evaluation of recombination and ionization rates, *Astronomy and Astrophysics Supplements Series*, **60**, 425
- Arons, J., & Wingert, D. W. 1972, Theoretical models of photoionized intergalactic hydrogen *The Astrophysical Journal*, **177**, 1
- Asplund, M., Grevesse, N., Sauval, A. J., & Scott, P. 2009, The chemical composition of the sun, *Annual Reviews of Astronomy & Astrophysics*, **47**, 481
- Barnett, R., Warren, S. J., Becker, G. D., et al. 2017, Observations of the Lyman series forest towards the redshift 7.1 quasar ULAS J1120+0641, *Astronomy and Astrophysics*, **601**, A16
- Badnell, N. R. 2003, Radiative recombination data for modeling dynamic finite-density plasmas, *The Astrophysical Journal Supplement Series*, **167**, 334
- Badnell, N. R., O'Mullane, M. G., Summers, H. P., et al. 2003, Dielectronic recombination data for dynamic finite-density plasmas. I. Goals and methodology, *Astronomy and Astrophysics*, **406**, 1151
- Bahcall, J. N. 1970, Multiple absorption redshifts in quasars, *Comments on Astrophysics and Space Physics*, **2**, 221
- Bahcall, J. N. 1971, Some unsolved problems in astrophysics, *The Astronomical Journal*, **76**, 283
- Bahcall, J. N., Bergeron, J., Boksenberg, A., et al. 1993, The Hubble Space Telescope quasar absorption line key project. I - First observational results, in-

- cluding Lyman-alpha and Lyman-limit systems, *The Astrophysical Journal Supplement Series*, **87**, 1
- Bahcall, J. N., Peterson, B. A., & Schmidt, M. 1966, On the absorption spectrum of 1116+12, *The Astrophysical Journal*, **145**, 369
- Bahcall, J. N., & Salpeter, E. E. 1965, On the interaction of radiation from distant sources with the intervening medium, *The Astrophysical Journal*, **142**, 1677
- Bahcall, J. N., & Salpeter, E. E. 1965, Absorption lines in the spectra of distance sources, *The Astrophysical Journal*, **144**, 847
- Bahcall, J. N., Sargent, W. L. W., & Schmidt, M. 1967, An analysis of the absorption spectrum of 3c 191, *The Astrophysical Journal*, **144**, 819
- Bahcall, J. N., & Spitzer, L. 1969, Absorption lines produced by galactic halos, *The Astrophysical Journal*, **156**, 63
- Bahcall, J. N., & Wolf, R. A. 1968, Fine-structure transitions, *The Astrophysical Journal*, **152**, 701
- Bajtlik, S., Duncan, R. C., & Ostriker, J. P. 1988, Quasar ionization of Lyman-alpha clouds - The proximity effect, a probe of the ultraviolet background at high redshift, *The Astrophysical Journal*, **327**, 570
- Barbieri, C., Capaccioli, M., & Zambon, M. 1975, Catalogue of quasi-stellar objects, *Memorie della Societa Astronomica Italiana*, **46**, 461
- Barlow, R. 2003, Asymmetric errors, in *Statistical Problems in Particle Physics, Astrophysics, and Cosmology*, eds. L.d Lyons, R. Mount, & R. Reitmeyer, 250, arXiv:physics/0401042
- Barlow, T. A., & Sargent, W. L. W. 1997, Keck high resolution spectroscopy of PKS 0123+257: Intrinsic absorption in a radio-loud quasar, *The Astronomical Journal*, **113**, 136
- Beaver, E. A., Burbidge, E. M., McIlwain, C. E., Epps, H. W., & Strittmatter, P. A. 1972, Digicon spectrophotometry of the quasi-stellar object PHL 957, *The Astrophysical Journal*, **178**, 95
- Bechtold, J., Dobrzycki, A., Wilden, B., et al. 2002, A uniform analysis of the Ly α forest at $z = 0-5$. III. Hubble Space Telescope Faint Object Spectrograph spectral atlas, *The Astrophysical Journal Supplement Series*, **140**, 143
- Becker, G. D., Rauch, M., & Sargent, W. L. W. 2007, The evolution of optical depth in the Ly α forest: Evidence against reionization at z 6, *The Astrophysical Journal*, **662**, 72
- Bergeron, J. 1986, The MgII absorption system in the QSO PKS 2128-12 - A galaxy disc/halo with a radius of 65 kpc, *Astronomy and Astrophysics*, **155**, L8
- Bergeron, J., & Boissé, P. 1991, A sample of galaxies giving rise to MgII quasar absorption systems, *Astronomy and Astrophysics*, **243**, 344
- Bergeron, J., Cristiani, S., & Shaver, P. A. 1992, Discovery of z of about 1 galaxies causing quasar absorption lines, *Astronomy and Astrophysics*, **257**, 417
- Bergeron, J., Petitjean, P., Sargent, W. L. W., et al. 1994, The Hubble Space Telescope quasar absorption line key project. 6: Properties of the metal-rich systems, *The Astrophysical Journal*, **436**, 33
- Bergeron, J., & Salpeter, E. E. 1970, Photoionization of intergalactic matter, *The Astrophysical Journal Letters*, **7**, 115
- Bergeron, J., & Stasińska, G. 1986, Absorption line systems in QSO spectra - Properties derived from observations and from photoionization models, *Astronomy and Astrophysics*, **169**, 1

- Bergvall, N., Marquart, T., Way, M. J., et al. 2016, Local starburst galaxies and their descendants. Statistics from the Sloan Digital Sky Survey, *Astronomy and Astrophysics*, **587**, A72
- Bethe, H. A., & Salpeter, E. E. 1957, *Quantum Mechanics of One and Two Electron Atoms*, New York: Academic Press
- Beare, R., Brown, M. J. I., Pimblet, K., Bian, F., & Lin, Y.-T. 2015, The $z < 1.2$ optical luminosity function from a sample of 410,000 galaxies in Boötes, *The Astrophysical Journal*, **815**, 94
- Bird, S., Garnett, R., & Ho, S. 2017, Statistical properties of damped Lyman-alpha systems from Sloan Digital Sky Survey DR12, *Monthly Notices of the Royal Astronomical Society*, **466**, 2111
- Boggess, N. W., Mather, J. C., Weiss, R., et al. 1992, The COBE mission - Its design and performance two years after launch, *The Astrophysical Journal*, **397**, 420
- Bohlin, R. C., Dickinson, M. E., & Calzetti, D. 2001, Spectrophotometric standards from the far-ultraviolet to the near-infrared: STIS and NICMOS fluxes, *The Astronomical Journal*, **122**, 2118
- Bohr, N. 1913, On the constitution of atoms and molecules. Part I., *Philosophical Magazine*, **26**, 1
- Boksenberg, A. 1978, The physics of QSO absorption line regions, *Physica Scripta*, **17**, 205
- Boksenberg, A., & Burgess, D. E. 1972, An image photon counting systems for optical astronomy, *Photo-electric Image Devices*, 835
- Boksenberg, A., Carswell, R. F., & Sargent, W. L. W. 1979, Multiple components in the $z = 0.424$ absorption system in the spectrum of the BL Lacertae object PKS 0735+178, *The Astrophysical Journal*, **227**, 370
- Boksenberg, A., & Sargent, W. L. W. 1975, Doublet structure in the absorption redshifts in the spectrum of PKS 0237-23, *The Astrophysical Journal*, **198**, 31
- Boksenberg, A., & Sargent, W. L. W. 1978, The existence of CaII absorption lines in the spectrum of the quasar 3c 232 due to the Galaxy NGC 3067, *The Astrophysical Journal*, **220**, 42
- Boksenberg, A., & Sargent, W. L. W. 2015, Properties of QSO metal-line absorption systems at high redshifts: Nature and evolution of the absorbers and new evidence on escape of ionizing radiation from galaxies, *The Astrophysical Journal Supplement Series*, **218**, 7
- Bolzonella, M., Miralles, J.-M., & Pelló, R. 2000, Photometric redshifts based on standard SED fitting procedures, *Astronomy and Astrophysics*, **363**, 476
- Bordoloi, R., Tumlinson, J., Werk, J. K., et al. 2014, The COS-Dwarfs Survey: The carbon reservoir around sub- L_* galaxies, *The Astrophysical Journal*, **796**, 136
- Bowen, D. V., Blades, J. C., & Pettini, M. 1995, Interstellar MgII absorption lines from low-redshift galaxies *The Astrophysical Journal*, **448**, 634
- Boyer, W., & Lynas-Gray, A. E. 2014, Evaluation of the Voigt function to arbitrary precision, *Monthly Notices of the Royal Astronomical Society*, **444**, 2555
- Bransden, B. H., & Joachain, C. J. 2003, *Physics of Atoms and Molecules*, New York: John Wiley & Sons, Inc.
- Bruzual A., G., & Charlot, S. 1993, Spectral evolution of stellar populations using isochrone synthesis, *The Astrophysical Journal*, **405**, 538
- Bunn, E. F., & Hogg, D. W. 2009, The kinematic origin of the cosmological redshift, *American Journal of Physics*, **77**, 688

- Burbidge, E. M. 1967, Quasi-stellar objects, *Annual Reviews of Astronomy & Astrophysics*, **5**, 399
- Burbidge, E. M., Lynds, C. R., & Burbidge, G. R. 1966, On the measurement and interpretation of absorption features in the spectrum of the quasi-stellar object 3c 191, *The Astrophysical Journal*, **144**, 447
- Burbidge, E. M., Lynds, C. R., & Stockton, A. N. 1968, Further observations of quasi-stellar objects with absorption-line spectra: Ton 1530, PKS 0237–23, and PHL 938, *The Astrophysical Journal*, **144**, 447
- Burbidge, G. R., Crowne, A. H., & Smith, H. E. 1977, An optical catalog of quasi-stellar objects, *The Astrophysical Journal Supplement Series*, **33**, 113
- Burbidge, G. R., Odell, S. L., Roberts, D. H., & Smith, H. E. 1977, On the origin of the absorption spectra of quasistellar and BL Lacertae objects, *The Astrophysical Journal*, **218**, 33
- Burchett, J. N., Tripp, T. M., Bordoloi, R., et al. 2016, A deep search for faint galaxies associated with very low redshift CIV absorbers. III. The mass- and environment-dependent circumgalactic medium, *The Astrophysical Journal*, **832**, 124
- Burles, S., Nollett, K. M., & Turner, M. S. 2001, Big bang nucleosynthesis predictions for precision cosmology, *The Astrophysical Journal Letters*, **552**, L1
- Carroll, S. M., Press, W. M., & Turner, E. L. 1992, The cosmological constant, *Annual Reviews of Astronomy & Astrophysics*, **30**, 499
- Carswell, R. F., Hilliard, R. L., Strittmatter, P. A., Taylor, D. J., & Weymann, R. J. 1975a, A cross-dispersed spectrograph and a study of the spectrum of the QSO 1331+170, *The Astrophysical Journal*, **196**, 351
- Carswell, R. F., Williams, R. E., Beaver, E. A., Harms, R., & Strittmatter, P. A. 1975, The spectrum of OH 471 (0642+44), *The Astrophysical Journal*, **195**, 269
- Cashman, F. H., Kulkarni, V. P., Kisieliu, R., Ferland, G. J., & Bogdanovich, P. 2017, Atomic data revisions for transitions relevant to observations of interstellar, circumgalactic, and intergalactic matter, *The Astrophysical Journal Supplement Series*, **230**, 8
- Cassata, P., Tasca, L. A. M., Le Fèvre, O., et al. 2015, The VIMOS Ultra-Deep Survey (VUDS): fast increase in the fraction of strong Lyman- α emitters from $z = 2$ to $z = 6$, *Astronomy and Astrophysics*, **573**, A24
- Cen, R., & Fang, T. 2006, Where are the baryons? III. Nonequilibrium effects and observables, *The Astrophysical Journal*, **650**, 573
- Cen, R., Miralda-Escudé, J., Ostriker, J. P., & Rauch, M. 1994, Gravitational collapse of small-scale structure as the origin of the Lyman-alpha forest, *The Astrophysical Journal Letters*, **437**, L9
- Charlton, J. C., & Churchill, C. W. 1996, MgII absorbing galaxies: Halos or disks?, *The Astrophysical Journal*, **465**, 631
- Charlton, J. C., & Churchill, C. W. 1998, The kinematic composition of MgII absorbers, *The Astrophysical Journal*, **499**, 181
- Charlton, J. C., Ding, J., Zonak, S. G., et al. 2003, High-resolution STIS/Hubble Space Telescope and HIRES/Keck spectra of three weak MgII absorbers toward PG 1634+706, *The Astrophysical Journal*, **589**, 111
- Chelouche, D., & Bowen, D. V. 2010, A Kennicutt-Schmidt law for intervening absorption line systems, *The Astrophysical Journal*, **722**, 1821

- Chen, H.-W., Helsby, J. E., Gauthier, J.-R., et al. 2010, An empirical characterization of extended cool gas around galaxies using MgII absorption features, *The Astrophysical Journal*, **714**, 1521
- Chen, H.-W., Lanzetta, K. M., & Webb, J. K. 2001a, The origin of CIV absorption systems at redshifts $z < 1$: Discovery of extended CIV envelopes around galaxies, *The Astrophysical Journal*, **556**, 158
- Chen, H.-W., Lanzetta, K. M., Webb, J. K., & Barcons, X. 2001b, The gaseous extent of galaxies and the origin of Ly α absorption systems. V. Optical and near-infrared photometry of Ly α -absorbing galaxies at $z < 1$, *The Astrophysical Journal*, **559**, 654
- Chen, S.-F. S., Simcoe, R. A., Torrey, P., et al. 2017, MgII absorption at $2 < z < 7$ with Magellan/Fire. III. Full statistics of absorption toward 100 high-redshift QSOs, *The Astrophysical Journal*, **850**, 188
- Churchill, C. W. 2001, MgII absorber number density at $z \simeq 0.05$: Implications for Ω_{DLA} evolution, *The Astrophysical Journal*, **560**, 92
- Churchill, C. W. 1995, Introduction to echelle data reduction using the Image Reduction Analysis Facility: emphasizing the Hamilton echelle spectrograph, *Lick Observatory Technical Report*, **74**
- Churchill, C. W., & Allen, S. L. 1995, A treatment for background correction on the Hamilton echelle spectrograph, *Publications of the Astronomical Society of the Pacific*, **107**, 193
- Churchill, C. W., Klimek, E., Medina, A., & Vander Vliet, J. R. 2014, Ionization modeling of astrophysical gaseous structures. I. The optically thin regime, *The Astrophysical Journal*, arXiv:1409.0916
- Churchill, C. W., & Le Brun, V. 1998, High-metallicity MgII absorbers in the $z < 1$ Ly α forest of PKS 0454+039: Giant low surface brightness galaxies?, *The Astrophysical Journal*, **499**, 677
- Churchill, C. W., Mellon, R. R., Charlton, J. C., et al. 2000, Low- and high-ionization absorption properties of MgII absorption-selected galaxies at intermediate redshifts. II. Taxonomy, kinematics, and galaxies, *The Astrophysical Journal*, **543**, 577
- Churchill, C. W., Rigby, J. R., Charlton, J. C., & Vogt, S. S. 1999, The population of weak MgII absorbers. I. A survey of 26 QSO HIRES/Keck spectra, *The Astrophysical Journal Supplement Series*, **120**, 51
- Churchill, C. W., Steidel, C. C., & Vogt, S. S. 1996, On the spatial and kinematic distributions of MgII absorbing gas in $z \simeq 0.7$ galaxies, *The Astrophysical Journal*, **471**, 164
- Churchill, C. W., & Vogt, S. S. 2001, The kinematics of intermediate-redshift MgII absorbers, *The Astronomical Journal*, **122**, 679
- Churchill, C. W., Vogt, S. S., & Charlton, J. C. 2003, The physical conditions of intermediate-redshift MgII absorbing clouds from Voigt profile analysis, *The Astronomical Journal*, **125**, 98
- Codoreanu, A., Ryan-Weber, E. V., Crighton, N. H. M., et al. 2017, The comoving mass density of MgII from $z \simeq 2$ –5.5, *Monthly Notices of the Royal Astronomical Society*, **472**, 1023
- Compton, A. H. 1923, A quantum theory of the scattering of X-rays by light elements, *Physical Review*, **21**, 483
- Cooke, J., Wolfe, A. M., Gawiser, E., & Prochaska, J. X. 2006, Survey for galaxies associated with $z \sim 3$ damped Ly α systems. II. Galaxy-absorber correlation functions, *The Astrophysical Journal*, **652**, 994

- Cooksey, K. L., Kao, M. M., Simcoe, R. A., O'Meara, J. M., & Prochaska, J. X. 2013, Precious metals in SDSS quasar spectra. I. Tracking the evolution of strong, $1.5 < z < 4.5$ CIV absorbers with thousands of systems, *The Astrophysical Journal*, **763**, 37
- Cowan, R. D. 1981, *The Theory of Atomic Structure and Spectra*, Berkeley: University of California Press
- Coleman, G. D., Wu, C.-C., & Weedman, D. W. 1980, Colors and magnitudes predicted for high redshift galaxies, *The Astrophysical Journal Supplement Series*, **43**, 393
- Crenshaw, D. M., Kraemer, S. B., Boggess, A., et al. 1999, Intrinsic absorption lines in Seyfert 1 galaxies. I. Ultraviolet spectra from the Hubble Space Telescope, *The Astrophysical Journal*, **516**, 750
- Croom, S. M., Richards, G. T., Shanks, T., et al. 2009, The 2dF-SDSS LRG and QSO Survey: the spectroscopic QSO catalogue, *Monthly Notices of the Royal Astronomical Society*, **392**, 19
- D'Agostini, G., & Raso, M. 2000, Uncertainties due to imperfect knowledge of systematic effects: general considerations and approximation formulae, arXiv:hep-ex/0002056
- Danforth, C. W., Keeney, B. A., Tilton, E. M., et al. 2016, An HST/COS survey of the low-redshift intergalactic medium. I. Survey, methodology, and overall results, *The Astrophysical Journal*, **817**, 111
- Davé, R., Hernquist, L., Katz, N., & Weinberg, D. H. 1999, The low-redshift Ly α forest in cold dark matter cosmologies, *The Astrophysical Journal*, **511**, 521
- Davé, R., Hernquist, L., Weinberg, D. H., & Katz, N. 1997, Voigt-profile analysis of the Ly α forest in a cold dark matter universe, *The Astrophysical Journal*, **477**, 21
- Davé, R., & Tripp, T. M. 2001, The statistical and physical properties of the low-redshift Ly α forest observed with the Hubble Space Telescope/STIS, *The Astrophysical Journal*, **553**, 528
- Davis, T. M., & Lineweaver, C. H. 2004, Expanding confusion: Common misconceptions of cosmological horizons and the superluminal expansion of the universe, *Publications of the Astronomical Society of Australia*, **21**, 97
- De Barros, S., Pentericci, L., Vanzella, E., et al. 2017, VLT/FORS2 view at $z \sim 6$: Lyman- α emitter fraction and galaxy physical properties at the edge of the epoch of cosmic reionization, *Astronomy and Astrophysics*, **608**, A123
- de Veny, J. B., Osborn, W. H., & Janes, K. 1971, A catalogue of quasars, *Publications of the Astronomical Society of the Pacific*, **83**, 611
- Dekker, H., D'Odorico, S., Kaufer, A., Delabre, B., & Kotzlowski, H. 2000, Design, construction, and performance of UVES, the echelle spectrograph for the UT2 Kueyen Telescope at the ESO Paranal Observatory, *Proceedings of the Society for Optical Engineers*, **4008**, 534
- Dittmann, O. J., & Koeppen, J. 1995, Quasar absorption lines. I. The chemical composition of the absorbing clouds, *Astronomy and Astrophysics*, **297**, 671
- Dopita, M. A., & Sutherland, R. S. 2003, *Astrophysics of the Diffuse Universe*, Berlin, New York: Springer
- Draine, B. T. 2011, *Physics of the Interstellar and Intergalactic Medium*, Princeton University Press
- Ebbets, D. C., & Brandt, J. C. 1983, Ultraviolet high-resolution spectroscopy from the Space Telescope, *Publications of the Astronomical Society of the Pacific*, **95**, 543

- Eggen, O. J., Lynden-Bell, D., & Sandage, A. R. 1962, Evidence from the motions of old stars that the Galaxy collapsed, *The Astrophysical Journal*, **136**, 748
- Ellison, S. L., Yan, L., Hook, I. M., et al. 2001, The CORALS survey I: New estimates of the number density and gas content of damped Lyman alpha systems free from dust bias, *Astronomy and Astrophysics*, **379**, 393
- Emerson, D. 1996, *Interpreting Astronomical Spectra*, Wiley-VCH
- Evans, J. L. 2011, MgII quasar absorption lines systems as a probe of galaxy halo kinematic evolution, PhD, New Mexico State University
- Faber, S. M., Willmer, C. N. A., Wolf, C., et al. 2007, Galaxy luminosity functions to $z = 1$ from DEEP2 and COMBO-17: Implications for red galaxy formation, *The Astrophysical Journal*, **665**, 265
- Fan, X., White, R. L., Davis, M., et al. 2000, The discovery of a luminous $z = 5.80$ quasar from the Sloan Digital Sky Survey, *The Astronomical Journal*, **120**, 1167
- Fathivavsari, H., Petitjean, P., Jamialahmadi, N., et al. 2018, Eclipsing damped Ly α systems in the Sloan Digital Sky Survey Data Release 12, *Monthly Notices of the Royal Astronomical Society*, arXiv:1804.07222
- Ferland, G. J. 2002, Hazy: A brief introduction to Cloudy 96, *University of Kentucky Department of Physics and Astronomy Internal Report*
- Ferland, G. J. 2003, Quantitative spectroscopy of photoionized clouds, *Annual Reviews of Astronomy & Astrophysics*, **41**, 517
- Ferland, G. J., Korista, K. T., Verner, D. A., et al. 1998, Cloudy 90: Numerical simulation of plasmas and their spectra, *Publications of the Astronomical Society of the Pacific*, **110**, 761
- Ferland, G. J., Porter, R. L., van Hoof, P. A. M., et al. 2013, The 2013 release of Cloudy, *Revista Mexicana de Astronomia y Astrofísica*, **49**, 137
- Ferland, G. J., Chatzikos, M., Guzmán, F., et al. 2017, The 2017 release of Cloudy, *Revista Mexicana de Astronomia y Astrofísica*, **53**, 385
- Filippenko, A. V. 1982, The importance of atmospheric differential refraction in spectrophotometry, *Publications of the Astronomical Society of the Pacific*, **94**, 715
- Foltz, C. B., Weymann, R. J., Peterson, B. M., et al. 1986, CIV absorption systems in QSO spectra - Is the character of systems with $z_{\text{abs}} \simeq z_{\text{em}}$ different from those with $z_{\text{abs}} \ll z_{\text{em}}$?, *The Astrophysical Journal*, **307**, 504
- Ford, W. K. J., & Rubin, V. C. 1966, Quasi-stellar objects with small redshifts: 1217+02, 3c 249.1, and 3c 263, *The Astrophysical Journal*, **145**, 357
- Fox, A. J., Savage, B. D., & Wakker, B. P. 2006, A survey of OVI, CIII, and HI in highly ionized high-velocity clouds, *The Astrophysical Journal Supplement Series*, **165**, 229
- Fox, A. J., Wakker, B. P., Barger, K. A., et al. 2014, The COS/UVES absorption survey of the Magellanic Stream. III. Ionization, total mass, and inflow rate onto the Milky Way, *The Astrophysical Journal*, **787**, 147
- Freedman, W. L., Madore, B. F., Gibson, B. K., et al. 2001, Final results from the Hubble space telescope key project to measure the Hubble constant, *The Astrophysical Journal*, **553**, 47
- Freedman, W. L., Madore, B. F., Scowcroft, V., et al. 2012, Carnegie Hubble Program: A mid-infrared calibration of the Hubble constant, *The Astrophysical Journal*, **758**, 24
- Frieman, J. A., Turner, M. S., & Huterer, D. 2008, Dark energy and the accelerating universe, *Annual Reviews of Astronomy & Astrophysics*, **46**, 385

- Fumagalli, M., O'Meara, J. M., Prochaska, J. X., Rafelski, M., & Kanekar, N. 2015, Directly imaging damped Ly α galaxies at $z > 2$ - III. The star formation rates of neutral gas reservoirs at $z \sim 2.7$, *Monthly Notices of the Royal Astronomical Society*, **446**, 3178
- Ganguly, R., Eracleous, M., Charlton, J. C., & Churchill, C. W. 1999, Intrinsic narrow absorption lines in Keck HIRES spectra of a sample of six quasars, *The Astronomical Journal*, **117**, 2594
- Garnett, R., Ho, S., Bird, S., & Schneider, J. 2017, Detecting damped Ly α absorbers with Gaussian processes, *Monthly Notices of the Royal Astronomical Society*, **472**, 1850
- Gaunt, J. A. 1930, Continuous Absorption, *Philosophical Transactions of the Royal Society of London Series A*, **229**, 163
- Gauss, C. F. 1823, *Theoria combinationis obsevationum erroribus minimis obnoxiae, Werke*, **Vol. 4**, Göttingen: Germany
- Gehren, T., & Ponz, D. 1986, Echelle background correction, *Astronomy and Astrophysics*, **168**, 386
- Ghavamian, P., Aloisi, D., Lennon, D. et al. 2009, Preliminary characterization of the post-launch line spread functions of COS, *Cosmic Origins Spectrograph Instrument Science Report, 2009-01* Version 1.0, Baltimore: Space Telescope Science Institute
- Gehrels, N. 1986, Confidence limits for small numbers of events in astrophysical data, *The Astrophysical Journal*, **303**, 336
- Gilliland, R. L. 1992, Details of noise sources and reduction processes, *Astronomical CCD observing and reduction techniques, Publications of the Astronomical Society of the Pacific Conference Series*, **23**, 68
- Gnat, O., & Sternberg, A. 2007, Time-dependent ionization in radiatively cooling gas, *The Astrophysical Journal Supplement Series*, **168**, 213
- Gnedin, N. Y., & Hollon, N. 2012, Cooling and heating funtions of photoionized gas, *The Astrophysical Journal Supplement Series*, **202**, 13
- Gordon, W. 1929, Zur berechnung der matrizen bein wasserstoff atom, *Annalen der Physik*, **2**, 1031
- Gould, R. J. 1968, Intergalactic matter, *Annual Reviews of Astronomy & Astrophysics*, **6**, 195
- Green, J. C., Froning, C. S., Osterman, S., et al. 2012, The Cosmic Origins Spectrograph, *The Astrophysical Journal*, **744**, 60
- Green, L. C., Matsushima, S., & Kolchin, E. K. 1958, Tables of the continuum wave function for hydrogen, *The Astrophysical Journal Supplement Series*, **3**, 459
- Greenstein, J. L., & Schmidt, M. 1964, The quasi-stellar radio sources 3c 48 and 3c 237, *The Astrophysical Journal*, **140**, 1
- Griesmann, U., & Kling, R. 2000, Interferometric measurement of resonance transition wavelengths in CIV, SiIV, AlIII, AlII, and SiII, *The Astrophysical Journal*, **563**, 113
- Gray, D. F. 1974, On scattered light corrections for stellar spectrographs, *Publications of the Astronomical Society of the Pacific*, **86**, 526
- Gray, D. F. 2005, *The Observational and Analysis of Stellar Photospheres*, Cambridge Astrophysics Series, Vol. 20
- Guillemin, P., & Bergeron, J. 1997, Evolution of quasar absorption-selected galaxies, *Astronomy and Astrophysics*, **328**, 499

- Gunn, J. E., & Peterson, B. A. 1965, On the density of neutral hydrogen in intergalactic space, *The Astrophysical Journal*, **142**, 1633
- Haardt, F., & Madau, P. 1996, Radiative transfer in a clumpy universe. II. The ultraviolet extragalactic background, *The Astrophysical Journal*, **461**, 20
- Haardt, F., & Madau, P. 2001, Modelling the UV/X-ray cosmic background with CUBA, in *Clusters of Galaxies and the High Redshift Universe Observed in X-rays*, XXIst Moriond Astrophysics Meeting, eds. D. M. Neumann & J. T. V. Tran, 64
- Haardt, F., & Madau, P. 2012, Radiative transfer in a clumpy universe. IV. New synthesis models of the cosmic UV/x-ray background, *The Astrophysical Journal*, **746**, 125
- Hallstadius, L. 1979, Extended measurements of isotope shifts in MgI, *Zeitschrift für Physik A*, **291**, 1220
- Halverson, N. W., Leitch, E. M., Pryke, C., et al. 2002, Degree angular scale interferometer first results: a measurement of the cosmic microwave background angular power spectrum, *The Astrophysical Journal*, **568**, 38
- Hamann, F., Barlow, T. A., Junkkarinen, V., & Burbidge, E. M. 1997, High-resolution spectra of intrinsic absorption lines in the quasi-stellar object UM 675, *The Astrophysical Journal*, **478**, 80
- Hamann, F., & Ferland, G. 1999, Elemental abundances in quasistellar objects: Star formation and galactic nuclear evolution at high redshifts, *Annual Reviews of Astronomy & Astrophysics*, **37**, 487
- Hamann, F., & Sabra, B. 2004, The diverse nature of intrinsic absorbers in AGNs, in *AGN Physics with the Sloan Digital Sky Survey*, **311**, 203
- Hamuy, M., Walker, A. R., Suntzeff, N. B., Gigoux, P., Heathcote, S. R., & Phillips, M. M. 1992, Southern spectrophotometric standards, *Publications of the Astronomical Society of the Pacific*, **104**, 533
- Hanany, S., Ade, P., Balbi, A., et al. 2000, Maxima-1: A measurement of the cosmic microwave background anisotropy on angular scales of $10'$ – 5° , *The Astrophysical Journal Letters*, **545**, L5
- Harms, R., & Fitch, J. 1991, Faint Object Spectrograph (FOS) early performance, *Proceedings of the Society for Optical Engineers*, **1494**, 49
- Harrison, E. 1993, The redshift-distance and velocity-distance law, *The Astrophysical Journal*, **403**, 28
- Haschick, A. D., & Burke, B. F. 1975, Neutral hydrogen absorption in the spectrum of the quasar-galaxy pair 4c 32.33/NGC 3067, *The Astrophysical Journal*, **200**, 137
- Heap, S. R., Williger, G. M., Smette, A., et al. 2000, STIS observations of HeII Gunn-Peterson absorption toward Q0302–003, *The Astrophysical Journal*, **534**, 69
- Hebb, M. H., & Menzel, D. H. 1940, Physical processes in gaseous nebulae. X. Collisional excitation of nebularium, *The Astrophysical Journal*, **92**, 408
- Hecht, E. & Zajac, A. 1974, *Optics*, Addison Wesley
- Hennawi, J. F., Prochaska, J. X., Burles, S., et al. 2006, Quasars probing quasars. I. Optically thick absorbers near luminous quasars, *The Astrophysical Journal*, **651**, 61
- Henry, R. B. C., & Worthey, G. 1999, The distribution of heavy elements in spiral and elliptical galaxies, *Publications of the Astronomical Society of the Pacific*, **111**, 919

- Herenz, P., Richter, P., Charlton, J. C., & Masiero, J. R. 2013, The Milky Way halo as a QSO absorption-line system. New results from an HST/STIS absorption-line catalogue of Galactic high-velocity clouds, *Astronomy and Astrophysics*, **550**, A87
- Herzberg, H. 1944, *Atomic Spectra and Atomic Structure*, New York: Dover Publications
- Hewitt, A., & Burbidge, G. R. 1993, A revised and updated catalog of quasi-stellar objects, *The Astrophysical Journal Supplement Series*, **87**, 451
- Hinshaw, G., Larson, D., Komatsu, E., et al. 2013, Nine-year Wilkinson Microwave Anisotropy Probe (WMAP) observations: cosmological parameter results, *The Astrophysical Journal Supplement Series*, **208**, 19
- Holmberg, E. 1976, in *Galaxies and the Universe: Volume 9, Stars and Stellar Systems*, eds. A. Sandage & J. Kristian, Chicago: Chicago University Press, 123
- Hogg, D. W. 1999, Distance measures in cosmology, arXiv:astro-ph/9905116
- Horne, K. 1986, An optimal extraction algorithm for CCD spectroscopy, *Publications of the Astronomical Society of the Pacific*, **98**, 609
- Howell, S., B. 2000, *Handbook of CCD astronomy*, Cambridge: Cambridge University Press
- Howk, J. C., & Sembach, K. R. 2000, Background and scattered-light subtraction in the high-resolution echelle modes of the Space Telescope imaging spectrograph, *The Astronomical Journal*, **119**, 2481
- Hoyle, F. 1948, A new model for the expanding universe, *Monthly Notices of the Royal Astronomical Society*, **108**, 372
- Hu, E. M., Kim, T.-S., Cowie, L. L., Songaila, A., & Rauch, M. 1995, The distribution of column densities and b -values in the Lyman-alpha forest, *The Astronomical Journal*, **110**, 1526
- Hu, E. M., McMahon, R. G., & Cowie, L. L. 1999, An Extremely Luminous Galaxy at $Z = 5.74$, *The Astrophysical Journal Letters*, **522**, L9
- Hu, W., Chen, C., Fang, D., Wang, Y., Lu, F., & Yang, F. 1996, Systematic study on ionization cross sections of electron-ion collisions for the Li-like isoelectronic sequence, *Journal of Physics B: Atomic and Optical Physics*, **29**, 2887
- Hubble, E. 1929, A relation between distance and radial velocity among extragalactic nebulae, *Proceedings of the National Academy of Science*, **15**, 168
- Humlíček, J. 1979, An efficient method for evaluation of the complex probability function: The Voigt function and its derivatives, *Journal of Quantitative and Radiative Transfer*, **21**, 309
- Hummer, D. G., & Seaton, M. J. 1963, The ionization structure of planetary nebulae, *Monthly Notices of the Royal Astronomical Society*, **125**, 437
- Huterer, D., & Shafer, D. L. 2018, Dark energy two decades after: observables, probes, consistency tests, *Reports on Progress in Physics*, **81**, 016901
- Jannuzi, B. T., Bahcall, J. N., Bergeron, J., et al. 1998, The Hubble Space Telescope quasar absorption line key project. XIII. A census of absorption-line systems at low redshift, *The Astrophysical Journal Supplement Series*, **118**, 1
- Jauncy, D. L., Wright, A. E., Peterson, B. A., & Condon, J. J. 1975, PKS 0528–250: A neutral stellar object at $z = 2.812$, with no emission lines and a rich absorption-line spectrum, *The Astrophysical Journal*, **221**, 109
- Jenkins, E. B. 1996, A procedure for correcting the apparent optical depths of moderately saturated interstellar absorption lines, *The Astrophysical Journal*, **471**, 292

- Kaastra, J. S., & Mewe, R. 1993, X-ray emission from thin plasmas. I - Multiple Auger ionisation and fluorescence processes for Be to Zn, *Astronomy and Astrophysics Supplements Series*, **97**, 443
- Kacprzak, G. G., & Churchill, C. W. 2011, The H I mass density in galactic halos, winds, and cold accretion as traced by Mg II absorption, *apj*, **743**, 34
- Kacprzak, G. G., Churchill, C. W., Barton, E. J., & Cooke, J. 2011, Halo gas and galaxy disk kinematics of a volume-limited sample of Mg II absorption-selected galaxies at $z \sim 0.1$, *apj*, **733**, 105
- Kacprzak, G. G., Churchill, C. W., Ceverino, D., et al. 2010, Halo gas and galaxy disk kinematics derived from observations and Λ CDM simulations of Mg II absorption-selected galaxies at intermediate redshift, *The Astrophysical Journal*, **711**, 533
- Kacprzak, G. G., Churchill, C. W., & Nielsen, N. M. 2012, Tracing outflows and accretion: A bimodal azimuthal dependence of Mg II absorption, *The Astrophysical Journal Letters*, **760**, L7
- Kacprzak, G. G., Muzahid, S., Churchill, C. W., Nielsen, N. M., & Charlton, J. C. 2015, The azimuthal dependence of outflows and accretion detected using O VI absorption, *The Astrophysical Journal*, **815**, 22
- Keeney, B. A., Stocke, J. T., Danforth, C. W., et al. 2017, Characterizing the circumgalactic medium of nearby galaxies with HST/COS and HST/STIS absorption-line spectroscopy. II. Methods and models, *The Astrophysical Journal Supplement Series*, **230**, 6
- Kim, T.-S., Hu, E. M., Cowie, L. L., & Songaila, A. 1997, The redshift evolution of the Ly- α forest, *The Astronomical Journal*, **114**, 1
- Kingdon, J. B., & Ferland, G. J. 1996, Rate coefficients for charge transfer between hydrogen and the first 30 elements, *The Astrophysical Journal Supplement Series*, **106**, 205
- Kinman, T. D. 1966, Object PHL 938 in the Haro-Luyten catalogue of blue stars, *The Astrophysical Journal*, **144**, 1232
- Klein, O., & Nashina, Y. 1929 On the scattering of radiation by free electrons according to the new relativistic quantum dynamics of Dirac, *Zeitschrift für Physik*, **52**, 853
- Kobayashi, N., Tokunaga, A. T., Terada, H., et al. 2000, IRCS: infrared camera and spectrograph for the Subaru Telescope, *Proceedings of the Society for Optical Engineers*, **4008**, 1056
- Kochanoc, V. P. 2011 Efficient approximations to the Voigt and Rautian-Sobelman profiles, *Atmospheric and Oceanic Optics*, **V24**, No. 5, 432
- Komatsu, E., Dunkley, J., Nolte, M. R., et al. 2009 Five-year Wilkinson microwave anisotropy probe (WMAP) observations: cosmological interpretation, *The Astrophysical Journal Supplement Series*, **180**, 330
- Kompaneets, A. S. 1957, The establishment of the thermal equilibrium between quanta and electrons, *Journal of Experimental and Theoretical Physics (USSR)*, **4**, 730
- Kramida, A. E. 2010, A critical compilation of experimental data on spectral lines and energy levels of hydrogen, deuterium, and tritium, *Data Nuclear Tables*, **96**, 586
- Kramida, A. E., Ralchenko, Y., Reader, J., & The NIST ASD Team 2014, NIST atomic spectra database, **V5.2**, 2015-Jun19 (<http://physics.nist.gov/asd>)
- Kriss, G. A. 2011, Improved medium resolution line spread functions for COS FUV

- spectra, *Cosmic Origins Spectrograph Instrument Science Report, 2011-01* Version 1.0, Baltimore: Space Telescope Science Institute
- Krolik, J. H., McKee, C. F., & Tarter, C. B. 1981, Two-phase models of quasar emission line regions, *The Astrophysical Journal*, **249**, 422
- Lehner, N., O'Meara, J. M., Howk, J. C., Prochaska, J. X., & Fumagalli, M. 2016, The cosmic evolution of the metallicity distribution of ionized gas traced by Lyman limit systems, *The Astrophysical Journal*, **833**, 283
- Lehner, N., Savage, B. D., Richter, P., et al. 2007, Physical properties, baryon content, and evolution of the Ly α forest: New insights from high-resolution observations at $z \leq 0.4$, *The Astrophysical Journal*, **658**, 680
- Lanzetta, K. M. 1993, QSO absorption lines: Implications for galaxy formation and evolution, in *The Environment and Evolution of Galaxies*, **188**, Dordrecht: Kluwer, 237
- Lanzetta, K. M., & Bowen, D. 1990, Intermediate-redshift galaxy halos - Results from QSO absorption lines, *The Astrophysical Journal*, **357**, 321
- Lanzetta, K. M., & Bowen, D. V. 1992, The kinematics of intermediate-redshift gaseous galaxy halos, *The Astrophysical Journal*, **391**, 48
- Lanzetta, K. M., Bowen, D. V., Tytler, D., & Webb, J. K. 1995a, The gaseous extent of galaxies and the origin of Lyman-alpha absorption systems: A survey of galaxies in the fields of Hubble Space Telescope spectroscopic target QSOs, *apj*, **442**, 538
- Lanzetta, K. M., Wolfe, A. M., & Turnshek, D. A. 1995, The IUE survey for Damped Lyman-alpha and Lyman-Limit absorption systems: Evolution of the gaseous content of the universe, *The Astrophysical Journal*, **440**, 435
- Lanzetta, K. M., Wolfe, A. M., & Turnshek, D. A. 1987, An absorption-line survey of 32 QSOs at red wavelengths - properties of the MgII absorbers, *The Astrophysical Journal*, **322**, 739
- Lauroesch, J. T., Truran, J. W., Welty, D. E., & York, D. G. 1996, QSO absorption line systems and early chemical evolution, *Publications of the Astronomical Society of the Pacific*, **108**, 641
- Le Brun, V., Bergeron, J., Boisse, P., & Christian, C. 1993, A deep imaging survey of fields around quasars with z less than 1.2 MgII absorption systems, *Astronomy and Astrophysics*, **279**, 33
- Le Brun, V., Bergeron, J., Boisse, P., & Deharveng, J. M. 1997, The nature of intermediate-redshift damped Ly α absorbers, *Astronomy and Astrophysics*, **321**, 733
- Leitherer, C., Schaerer, D., Goldader, J. D., et al. 1999, Starburst99: Synthesis models for galaxies with active star formation, *The Astrophysical Journal Supplement Series*, **123**, 3
- Leitherer, C., Ortiz Otálvaro, P. A., Bresolin, F., et al. 2010, A library of theoretical ultraviolet spectra of massive, hot stars for evolutionary synthesis, *The Astrophysical Journal Supplement Series*, **189**, 309
- Leitherer, C., Ekström, S., Meynet, G., et al. 2014, The effects of stellar rotation. II. A comprehensive set of Starburst99 models, *The Astrophysical Journal Supplement Series*, **212**, 14
- Levich, E. V., & Sunyaev, R. A. 1970, The heating of gas in the vicinity of quasars, nuclei of Seyfert galaxies, and pulsars by the induced Compton effect, *The Astrophysical Journal Letters*, **7**, 69
- Li, W.-K., & Blinder, S. M. 2014, Atomic states: The L-S and j - j coupling schemes and their correlation, arXiv:1409.2032

- Loveday, J., Norberg, P., Baldry, I. K., et al. 2012, Galaxy and Mass Assembly (GAMA): ugriz galaxy luminosity functions, *Monthly Notices of the Royal Astronomical Society*, **420**, 1239
- Lowrance, J. L., Morton, D. C., Zucchini, P., Oke, J. B., & Schmidt, M. 1972, The spectrum of the quasi-stellar object PHL 957, *The Astrophysical Journal*, **171**, 233
- Lu, L., Sargent, W. L. W., Womble, D. S., & Takada-Hidai, M. 1996, The Lyman-alpha forest at $z \approx 4$: Keck HIRES observations of Q0000-26, *The Astrophysical Journal*, **472**, 509
- Lynds, C. R. 1971, The absorption-line spectrum of 4c 05.34, *The Astrophysical Journal Letters*, **164**, 73
- Lynds, C. R., Hill, S. J., Heere, K., & Stockton, A. N. 1966, New spectroscopic observations of fourteen quasi-stellar sources, *The Astrophysical Journal*, **144**, 1244
- Madau, P. 1995, Radiative transfer in a clumpy universe: The colors of high-redshift galaxies, *The Astrophysical Journal*, **441**, 18
- Madau, P., Haardt, F., & Rees, M. J. 1999, Radiative transfer in a clumpy universe. III. The nature of cosmological ionizing sources, *The Astrophysical Journal*, **514**, 648
- Maller, A. H., & Bullock, J. S. 2004, Multiphase galaxy formation: high-velocity clouds and the missing baryon problem, *Monthly Notices of the Royal Astronomical Society*, **355**, 694
- Marsh, T. R. 1989, The extraction of highly distorted spectra, *Publications of the Astronomical Society of the Pacific*, **101**, 1032
- Martinez, P., & Klotz, A. 1998, *A practical guide to CCD astronomy*, Cambridge: Cambridge University Press
- Massey, P., & Jacoby, G. H. 1992, CCD data: the good, the bad, and the ugly, *Astronomical CCD observing and reduction techniques, Publications of the Astronomical Society of the Pacific Conference Series*, **23**, 240
- Massey, P., Strobel, K., Barnes, J. V., & Anderson, E. 1988, Spectrophotometric standards, *The Astrophysical Journal*, **328**, 315
- Matejek, M. S., & Simcoe, R. A. 2012, A Survey of MgII absorption at $2 < z < 6$ with Magellan/FIRE. I. Sample and evolution of the MgII frequency, *The Astrophysical Journal*, **761**, 112
- Meiring, J. D., Tripp, T. M., Werk, J. K., et al. 2013, QSO absorption systems detected in NeVIII: High-metallicity clouds with a large effective cross section, *The Astrophysical Journal*, **767**, 49
- Menzel, D. 1930, Atomic coefficients of general absorption, *Publications of Lick Observatory*, **17**, 222
- Menzel, D., & Pekeris, C. 1935, Absorption coefficients and H-line intensities, *Monthly Notices of the Royal Astronomical Society*, **96**, 77
- Meylan, G. 1995, *QSO Absorption Lines*, Berlin: Springer-Verlag
- Mihalas, D. 1978, *Stellar Atmospheres*, San Francisco: W. H. Freeman & Company
- Mihalas, D., & Mihalas, B. D. 1984, *Foundation of Radiation Hydrodynamics*, New York: Oxford University Press
- Misawa, T., Saez, C., Charlton, J. C., et al. 2016, Multi-sightline observation of narrow absorption lines in lensed quasar SDSS J1029+2623, *The Astrophysical Journal*, **825**, 25
- Mo, H. J., & Miralda-Escude, J. 1996, Gaseous galactic halos and quasi-stellar object absorption-line Systems, *The Astrophysical Journal*, **469**, 589

- Moore, C. E. 1970, Selected tables of atomic spectra, *National Standard Reference Series*, National Bureau of Standards, Washington, DC: U.S. Department of Commerce
- Moore, C. E., & Merrill, P. W. 1968, Partial Grotrian diagrams of astrophysical interest, *National Standard Reference Data Series*, National Bureau of Standards, **23**
- Morton, D. C., & Morton, W. A. 1972, Absorption-line profiles in the quasi-stellar object PHL 957, *The Astrophysical Journal*, **174**, 237
- Murdoch, H. S., Hunstead, R. W., Pettini, M., & Blades, J. C. 1986, Absorption spectrum of the $z = 3.78$ QSO 2000-330. II - The redshift and equivalent width distributions of primordial hydrogen clouds, *The Astrophysical Journal*, **309**, 19
- Murphy, M. T., Webb, J. K., Flambaum, V. V., Churchill, C. W., & Prochaska, J. X. 2001, Possible evidence for a variable fine-structure constant from QSO absorption lines: systematic errors, *Monthly Notices of the Royal Astronomical Society*, **327**, 1223
- Muzahid, S., Kacprzak, G. G., Churchill, C. W., et al. 2015, An extreme metallicity, large-scale outflow from a star-forming galaxy at $z \sim 0.4$, *The Astrophysical Journal*, **811**, 132
- Narayanan, A., Misawa, T., Charlton, J. C., & Kim, T.-S. 2007, A survey of weak MgII absorbers at $0.4 < z < 2.4$, *The Astrophysical Journal*, **660**, 1093
- Narayanan, A., Wakker, B. P., & Savage, B. D. 2009, Detection of NeVIII in an intervening multiphase absorption system toward 3C 263, *The Astrophysical Journal*, **703**, 74
- Nestor, D. B., Turnshek, D. A., & Rao, S. M. 2005, MgII absorption systems in Sloan Digital Sky Survey QSO spectra, *The Astrophysical Journal*, **628**, 637
- Netterfield, C. B., Ade, P. A. R., Bock, J. J., et al. 2002, A measurement by Boomerang of multiple peaks in the angular power spectrum of the cosmic microwave background, *The Astrophysical Journal*, **571**, 604
- Nielsen, N. M., Churchill, C. W., Kacprzak, G. G., & Murphy, M. T. 2013a, MAGI-CAT I. The Mg II Absorber-Galaxy Catalog, *The Astrophysical Journal*, **776**, 114
- Nielsen, N. M., Churchill, C. W., Kacprzak, G. G., & Murphy, M. T. 2013b, MAGI-CAT II. General characteristics of the MgII absorbing circumgalactic medium, *The Astrophysical Journal*, **776**, 115
- Nielsen, N. M., Churchill, C. W., Kacprzak, G. G., Murphy, M. T., & Evans, J. L. 2015, MAGI-CAT V. Orientation of outflows and accretion determine the kinematics and column densities of the circumgalactic medium, *The Astrophysical Journal*, **812**, 83
- Nielsen, N. M., Churchill, C. W., Kacprzak, G. G., Murphy, M. T., & Evans, J. L. 2016, MAGI-CAT IV. Kinematics of the circumgalactic medium and evidence for quiescent evolution around red galaxies, *The Astrophysical Journal*, **818**, 171
- Nielsen, N. M., Kacprzak, G. G., Muzahid, S., et al. 2017, The highly ionized circumgalactic medium is kinematically uniform around galaxies, *The Astrophysical Journal*, **834**, 148
- Noguchi, K., Ando, H., Izumiura, H., Kawanomoto, S., Tanaka, W., & Aoki, W. 1998, High Dispersion Spectrograph (HDS) for the Subaru Telescope, *Proceedings of the Society for Optical Engineers*, **3355**, 454

- Novotny, E. 1973, *Introduction to Stellar Atmospheres and Interiors*, New York: Oxford University Press
- Nussbaumer, H., & Storey, P. J. 1983, Dielectronic recombination at low temperatures, *Astronomy and Astrophysics*, **126**, 75
- Nussbaumer, H., & Storey, P. J. 1986, Dielectronic recombination at low temperatures. III - Recombination coefficients for Mg, Al, Si, *Astronomy and Astrophysics Supplements Series*, **64**, 545
- Nussbaumer, H., & Storey, P. J. 1987, Dielectronic recombination at low temperatures. IV - Recombination coefficients for neon, *Astronomy and Astrophysics Supplements Series*, **69**, 123
- Oke, J. B. 1990, Faint spectrophotometric standard stars, *The Astronomical Journal*, **99**, 1621
- Oliveira, C. M., Hébrard, G., Howk, J. C., et al. 2003, Interstellar deuterium, nitrogen, and oxygen abundances toward GD 246, WD 2331-475, HZ 21, and Lanning 23: Results from the FUSE Mission, *The Astrophysical Journal*, **587**, 235
- O'Meara, J. M., Chen, H.-W., & Kaplan, D. L. 2006, A shot in the dark: A technique for locating the stellar counterparts of Damped Ly α Absorbers, *The Astrophysical Journal Letters*, **642**, L9
- O'Meara, J. M., Lehner, N., Howk, J. C., et al. 2015, The first data release of the KODIAQ survey, *The Astronomical Journal*, **150**, 111
- O'Meara, J. M., Lehner, N., Howk, J. C., et al. 2017, The second data release of the KODIAQ survey, *The Astronomical Journal*, **154**, 114
- O'Meara, J. M., Prochaska, J. X., Burles, S., et al. 2007, The Keck+Magellan survey for Lyman limit absorption. I. The frequency distribution of super Lyman limit systems, *The Astrophysical Journal*, **656**, 666
- Oppenheimer, B. D., & Schaye, J. 2013, Non-equilibrium ionization and cooling of metal-enriched gas in the presence of a photoionization background, *Monthly Notices of the Royal Astronomical Society*, **434**, 1043
- Osmer, P. S. 1979, Spectrophotometry of eight optically selected quasars with z between 3.0 and 3.5: A low-resolution approach to the absorption-line problem, *The Astrophysical Journal*, **227**, 18
- Osterbrock, D. E., & Ferland, G. J. 2006, *Astrophysics of Gaseous Nebulae and Active Galactic Nuclei*, Sausalito, CA: University Science Books
- Pâris, I., Petitjean, P., Ross, N. P., et al. 2017, The Sloan Digital Sky Survey Quasar Catalog: Twelfth data release, *Astronomy and Astrophysics*, **597**, A79
- Peacock, J. A. 1999, *Cosmological Physics*, Cambridge: Cambridge University Press
- Peebles, J. 1996, *Principles of Physical Cosmology*, Princeton University Press
- Penton, S. V., Shull, J. M., & Stocke, J. T. 2000, The local Ly α forest. II. Distribution of H I absorbers, Doppler widths, and baryon content, *The Astrophysical Journal*, **544**, 150
- Perlmutter, S., Aldering, G., Goldhaber, G., et al. 1999a, Measurements of Ω and Λ from 42 high-redshift supernovae, *The Astrophysical Journal*, **517**, 565
- Perlmutter, S., Turner, M. S., & White, M. 1999b, Constraining dark energy with Type Ia supernovae and large-scale structure, *Physical Review Letters*, **83**, 670
- Perry, J. J., Burbidge, E. M., & Burbidge, G. R. 1978, Absorption on the spectra of quasi-stellar objects and BL Lacertae objects, *Publications of the Astronomical Society of the Pacific*, **90**, 337
- Peterson, B. A. 1978, QSO absorption lines and intergalactic hydrogen clouds, *Proceedings of the International Astronomical Union*, **79**, 389

- Péroux, C., Bouché, N., Kulkarni, V. P., York, D. G., & Vladilo, G. 2011, A SINFONI integral field spectroscopy survey for galaxy counterparts to damped Lyman- α systems - I. New detections and limits for intervening and associated absorbers, *Monthly Notices of the Royal Astronomical Society*, **410**, 2237
- Petitjean, P., & Bergeron, J. 1990, MgII quasar absorption systems and properties of gaseous haloes, *Astronomy and Astrophysics*, **231**, 309
- Petitjean, P., & Bergeron, J. 1994, CIV QSO absorption systems and properties of galactic haloes at high redshift, *Astronomy and Astrophysics*, **283**, 759
- Pickering, J. C., Thorne, A. P., Murray, J. E., et al. 2000, Accurate laboratory wavelengths of some ultraviolet lines of Cr, Zn, and Ni relevant to time variations of the fine structure constant, *Monthly Notices of the Royal Astronomical Society*, **319**, 163
- Pieri, M. M., Frank, S., Mathur, S., et al. 2010, A search for oxygen in the low-density Ly α forest using the Sloan Digital Sky Survey, *The Astrophysical Journal*, **716**, 1084
- Planck Collaboration, Adam, R., Ade, P. A. R., et al. 2016a, Planck 2015 results. I. Overview of products and scientific results, *Astronomy and Astrophysics*, **594**, A1
- Planck Collaboration, Ade, P. A. R., Aghanim, N., et al. 2016b, Planck 2015 results. XIII. Cosmological parameters, *Astronomy and Astrophysics*, **594**, A13
- Pomraning, G. C. 1973, *The Equations of Radiation Hydrodynamics* International Series of Monographs in Natural Philosophy, Oxford: Pergamon Press
- Pradhan, A. K., & Nahar, S. N. 2011, *Atomic Astrophysics and Spectroscopy*, Cambridge: Cambridge University Press
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. 2007, *Numerical Recipes: The Art of Scientific Computing*, New York: Cambridge University Press
- Prochaska, J. X., & Herbert-Fort, S. 2004, The Sloan Digital Sky Survey damped Ly α survey: Data release 1, *Publications of the Astronomical Society of the Pacific*, **116**, 622
- Prochaska, J. X., O'Meara, J. M., Fumagalli, M., Bernstein, R. A., & Burles, S. M. 2015, The Keck + Magellan survey for Lyman limit absorption. III. Sample definition and column density measurements, *The Astrophysical Journal Supplement Series*, **221**, 2
- Prochaska, J. X., & Wolfe, A. M. 1996, A Keck HIRES Investigation of the metal abundances and kinematics of the $z = 2.46$ damped Lyman alpha system toward Q0201+365, *The Astrophysical Journal*, **470**, 403
- Prochter, G. E., Prochaska, J. X., & Burles, S. M. 2006, On the incidence and kinematics of strong MgII absorbers, *The Astrophysical Journal*, **639**, 766
- Rao, S. M., Belfort-Mihalyi, M., Turnshek, D. A., et al. 2011, A ground-based imaging study of galaxies causing damped Lyman- α (DLA), sub-DLA and Lyman limit system absorption in quasar spectra, *Monthly Notices of the Royal Astronomical Society*, **416**, 1215
- Rao, S. M., Nestor, D. B., Turnshek, D. A., et al. 2003, Low-redshift damped Ly α galaxies toward the quasars B2 0827+243, PKS 0952+179, PKS 1127-145, and PKS 1629+120, *The Astrophysical Journal*, **595**, 94
- Rao, S. M., & Turnshek, D. A. 2000, Discovery of Damped Ly α Systems at $z < 1.65$ and results on their incidence and cosmological mass density *The Astrophysical Journal Supplement Series*, **130**, 1

- Rao, S. M., Turnshek, D. A., Sardane, G. M., & Monier, E. M. 2017, The statistical properties of neutral gas at $z < 1.65$ from UV measurements of Damped Lyman Alpha systems, *Monthly Notices of the Royal Astronomical Society*, **471**, 3428
- Rao, S. M., Turnshek, D. A., & Nestor, D. B. 2006, Damped Ly α systems at $z < 1.65$: The expanded Sloan Digital Sky Survey Hubble Space Telescope sample, *The Astrophysical Journal*, **636**, 610
- Rauch, M. 1998, *Annual Reviews of Astronomy & Astrophysics*, The Lyman alpha forest in the spectra of QSOs, **36**, 267
- Rauch, M., Sargent, W. L. W., Womble, D. S., & Barlow, T. A. 1996, Temperature and kinematics of CIV absorption systems, *The Astrophysical Journal Letters*, **467**, L5
- Rees, M. J., & Ostriker, J. P. 1977, Cooling, dynamics and fragmentation of massive gas clouds - Clues to the masses and radii of galaxies and clusters, *Monthly Notices of the Royal Astronomical Society*, **179**, 541
- Ribaudo, J., Lehner, N., & Howk, J. C. 2011, A Hubble Space Telescope study of Lyman limit systems: Census and evolution, *The Astrophysical Journal*, **736**, 42
- Richter, P., Nuza, S. E., Fox, A. J., et al. 2017, An HST/COS legacy survey of high-velocity ultraviolet absorption in the Milky Way's circumgalactic medium and the Local Group, *Astronomy and Astrophysics*, **607**, A48
- Richter, P., Savage, B. D., Tripp, T. M., & Sembach, K. R. 2004, FUSE and STIS observations of the warm-hot intergalactic medium toward PG 1259+593, *The Astrophysical Journal Supplement Series*, **153**, 165
- Riess, A. G., Casertano, S., Yuan, W., et al. 2018, New parallaxes of galactic Cepheids from spatially scanning the Hubble Space Telescope: Implications for the Hubble constant, *The Astrophysical Journal*, **855**, 136
- Riess, A. G., Filippenko, A. V., Challis, P., et al. 1998, Observational evidence from supernovae for an accelerating universe and a cosmological Constant, *The Astronomical Journal*, **116**, 1009
- Riess, A. G., Macri, L. M., Hoffmann, S. L., et al. 2016, A 2.4% determination of the local value of the Hubble constant, *The Astrophysical Journal*, **826**, 56
- Rubin, K. H. R., Diamond-Stanic, A. M., Coil, A. L., Crighton, N. H. M., & Moustakas, J. 2018, Galaxies probing galaxies in PRIMUS. I. Sample, spectroscopy, and characteristics of the $z \sim 0.5$ MgIIabsorbing circumgalactic medium, *The Astrophysical Journal*, **853**, 95
- Rubin, K. H. R., Prochaska, J. X., Koo, D. C., et al. 2014, Evidence for ubiquitous collimated galactic-scale outflows along the star-forming sequence at $z \sim 0.5$, *The Astrophysical Journal*, **794**, 156
- Rybicki, G. B., & Lightman, A. P. 2004, *Radiative Processes in Astrophysics*, Wiley-VCH: Verlag GmbH & Company
- Sandage, A. R., 1965, The existence of a major new constituent of the universe: the quasi-stellar galaxies, *The Astrophysical Journal*, **141**, 1560
- Sargent, W. L. W., Boksenberg, A., & Steidel, C. C. 1988a, CIV absorption in a new sample of 55 QSOs - Evolution and clustering of the heavy-element absorption redshifts, *The Astrophysical Journal Supplement Series*, **68**, 539
- Sargent, W. L. W., Steidel, C. C., & Boksenberg, A. 1988b, MgII absorption in the spectra of high and low redshift QSOs, *The Astrophysical Journal*, **334**, 22
- Sargent, W. L. W., Steidel, C. C., & Boksenberg, A. 1989, A survey of Lyman limit

- absorption in the spectra of 59 high-redshift QSOs, *The Astrophysical Journal Supplement Series*, **69**, 703
- Sargent, W. L. W., Young, P. J., Boksenberg, A., & Tytler, D. 1980, The distribution of Lyman-alpha absorption lines in the spectra of six QSOs - Evidence for an intergalactic origin, *The Astrophysical Journal Supplement Series*, **42**, 41
- Savage, B. D., Lehner, N., & Narayanan, A. 2011, COS observations of metal line and broad Ly α absorption in the multi-phase OVI and NeVIII system at $z = 0.20701$ toward HE 0226-4110, *The Astrophysical Journal*, **743**, 180
- Savage, B. D., Lu, L., Bahcall, J. N., et al. 1993, The Hubble Space Telescope quasar absorption line key project. III - First observational results on Milky Way gas, *The Astrophysical Journal*, **413**, 116
- Savage, B. D., & Sembach, K. R. 1991, The analysis of apparent optical depth profiles for interstellar absorption lines, *The Astrophysical Journal*, **379**, 245
- Savage, B. D., Sembach, K. R., Wakker, B. P., et al. 2003, Distribution and kinematics of OVI in the galactic halo, *The Astrophysical Journal Supplement Series*, **146**, 125
- Savage, B. D., Wakker, B., Jannuzi, B. T., et al. 2000, The Hubble Space Telescope quasar absorption line key project. XV. Milky Way absorption lines, *The Astrophysical Journal Supplement Series*, **129**, 563
- Savedoff, M. P. 1956, Physical constants in extragalactic nebulae, *Nature*, **178**, 688
- Schaye, J. 2001, Model-independent Insights into the nature of the Ly α forest and the distribution of matter in the universe, *The Astrophysical Journal*, **559**, 507
- Schechter, P. 1976, An analytic expression for the luminosity function for galaxies, *The Astrophysical Journal*, **203**, 297
- Schmidt, B. P., Suntzeff, N. B., Phillips, M. M., et al. 1998, The high-Z supernova search: Measuring cosmic deceleration and global curvature of the universe using type Ia supernovae, *The Astrophysical Journal*, **507**, 46
- Schmidt, M. 1963, 3c 273: A star-like object with large red-shift, *nature*, **197**, 1040
- Schmidt, M. 1965, Large redshifts of five quasi-stellar radio sources, *The Astrophysical Journal*, **141**, 1295
- Schmidt, M. 1966, Redshifts of fourteen quasi-stellar radio sources, *The Astrophysical Journal*, **144**, 443
- Schmidt, M. 1969, Quasistellar objects, *Annual Reviews of Astronomy & Astrophysics*, **7**, 527
- Schmidt, M., & Green, R. F. 1983, Quasar evolution derived from the Palomar bright quasar survey and other complete surveys, *The Astrophysical Journal*, **269**, 352
- Schmidt, M., & Matthews, T. A. 1964, Redshift of the quasi-stellar radio sources 3c 47 and 3c 147, *The Astrophysical Journal*, **139**, 781
- Schneider, D. P., Hartig, G. F., Jannuzi, B. T., et al. 1993, The Hubble Space Telescope quasar absorption line key project. II - Data calibration and absorption-line selection, *The Astrophysical Journal Supplement Series*, **87**, 45
- Schneider, D. P., Schmidt, M., & Gunn, J. E. 1989, PC 1158 + 4635 - An optically selected quasar with a redshift of 4.73, *The Astronomical Journal*, **98**, 1951
- Schroeder, D. J. 1978, *Astronomical Optics*, San Diego: Academic Press
- Schure, K. M., Kosenko, D., Kaastra, J. S., Keppens, R., & Vink, J. 2009, A new radiative cooling curve based on an up-to-date plasma emission code, *Astronomy and Astrophysics*, **508**, 751

- Scott, J., Bechtold, J., Dobrzycki, A., & Kulkarni, V. P. 2000, A uniform analysis of the Ly α forest at $z = 0 - 5$. II. Measuring the mean intensity of the extragalactic ionizing background using the proximity effect, *The Astrophysical Journal Supplement Series*, **130**, 67
- Seaton, M. J. 1958, Thermal inelastic collision processes, *Reviews of Modern Physics*, **30**, 979
- Seaton, M. J. 1959, Radiative recombination of hydrogenic ions, *Monthly Notices of the Royal Astronomical Society*, **119**, 81
- Sembach, K. R., & Savage, B. D. 1992, Observations of highly ionized gas in the Galactic halo, *The Astrophysical Journal Supplement Series*, **83**, 147
- Shenstone, A. G. 1970, The second spectrum of nickel (NiII), *Journal of Research of the National Bureau of Standards, Sect. A*, **74(6)**, 801
- Scheuer, P. A. G. 1965, A sensitive test for the presence of atomic hydrogen in intergalactic space, *Nature*, **207**, 963
- Shklovsky, I. S. 1964, Remarks on radio stars, *Astronomicheskii Zhurnal*, **41**, 176
- Shklovsky, J. 1967, On the nature of “standard” absorption spectra of the quasi-stellar objects, *The Astrophysical Journal*, **150**, 1
- Shull, J. M., Danforth, C. W., Tilton, E. M., Moloney, J., & Stevans, M. L. 2017, An ultraviolet survey of low-redshift partial Lyman-limit systems with the *HST* Cosmic Origins Spectrograph, *The Astrophysical Journal*, **849**, 106
- Shull, J. M., & Van Steenberg, M. 1982, The ionization equilibrium of astrophysically abundant elements, *The Astrophysical Journal Supplement Series*, **48**, 95
- Sievers, J. L., Bond, J. R., Cartwright, J. K., et al. 2003, Cosmological parameters from Cosmic Background Imager observations and comparisons with Boomerang, Dasi, and Maxima, *The Astrophysical Journal*, **591**, 599
- Simcoe, R. A., Burgasser, A. J., Schechter, P. L., et al. 2013, FIRE: A facility class near-infrared echelle spectrometer for the Magellan telescopes, *Publications of the Astronomical Society of the Pacific*, **125**, 270
- Simcoe, R. A., Sargent, W. L. W., & Rauch, M. 2002, Characterizing the warm-hot intergalactic medium at high redshift: A high-resolution survey for OVI at $z = 2.5$, *The Astrophysical Journal*, **578**, 737
- Simcoe, R. A., Sargent, W. L. W., Rauch, M., & Becker, G. 2006, Observations of chemically enriched QSO absorbers near $z \sim 2.3$ galaxies: Galaxy formation feedback signatures in the intergalactic medium, *The Astrophysical Journal*, **637**, 648
- Smette, A., Surdej, J., Shaver, P. A., et al. 1992, A spectroscopic study of UM 673 A and B: On the size of Ly α clouds, *The Astrophysical Journal*, **389**, 39
- Smith, B., Sigurdsson, S., & Abel, T. 2008, Metal cooling in simulations of cosmic structure formation, *Monthly Notices of the Royal Astronomical Society*, **385**, 1443
- Songaila, A., & Cowie, L. L. 2010, The evolution of Lyman-limit absorption systems to redshift six, *The Astrophysical Journal*, **721**, 1448
- Spergel, D. N., Verde, L., Peiris, H. V., et al. 2003, First-year Wilkinson microwave anisotropy probe (WMAP) observations: determination of cosmological parameters, *The Astrophysical Journal Supplement Series*, **148**, 175
- Spitzer, L., Jr., & Fitzpatrick, E. L. 1993, Composition of interstellar clouds in the disk and halo. I - HD 93521, *The Astrophysical Journal*, **409**, 299

- Steidel, C. C. 1990a, The high redshift extension of the survey for CIV absorption in the spectra of QSOs: The redshift evolution of heavy element absorbers, *The Astrophysical Journal Supplement Series*, **72**, 1
- Steidel, C. C. 1990b, The properties of Lyman limit absorbing clouds at $z = 3$: Physical conditions in the extended gaseous halos of high-redshift galaxies, *The Astrophysical Journal Supplement Series*, **74**, 37
- Steidel, C. C. 1993a, The History of Halo Gas in Normal Galaxies, in *Galaxy Evolution. The Milky Way Perspective*, **49**, 227
- Stiedel, C. C. 1993b, The properties of absorption-line selected high-redshift galaxies, in *The Environment and Evolution of Galaxies*, **188**, Dordrecht: Kluwer, 263
- Steidel, C. C. 1995, The nature and evolution of absorption-selected galaxies, in *QSO Absorption Lines*, ed. G. Meylan, Berlin: Springer-Verlag, 139
- Steidel, C. C., & Dickinson, M. 1992, The unusual field of the quasar 3C 336 - Identification of three foreground MgII absorbing galaxies, *The Astrophysical Journal*, **394**, 81
- Steidel, C. C., Dickinson, M., & Bowen, D. V. 1993, A dwarf galaxy near the sight line to PKS 0454+0356 - A fading 'faint blue galaxy', *The Astrophysical Journal Letters*, **413**, L77
- Steidel, C. C., Dickinson, M., Meyer, D. M., Adelberger, K. L., & Sembach, K. R. 1997, Quasar absorbing galaxies at $z \leq 1$: Deep imaging and spectroscopy in the field of 3C 336, *The Astrophysical Journal*, **480**, 568
- Steidel, C. C., Dickinson, M., & Persson, S. E. 1994, Field galaxy evolution since z approximately 1 from a sample of QSO absorption-selected galaxies, *The Astrophysical Journal Letters*, **437**, L75
- Steidel, C. C., Kollmeier, J. A., Shapley, A. E., et al. 2002, The kinematic connection between absorbing gas toward QSOs and galaxies at intermediate redshift, *The Astrophysical Journal*, **570**, 526
- Steidel, C. C., & Sargent, W. L. W. 1992, MgII absorption in the spectra of 103 QSOs - Implications for the evolution of gas in high-redshift galaxies, *The Astrophysical Journal Supplement Series*, **80**, 1
- Stengler-Larrea, E. A., Boksenberg, A., Steidel, C. C., et al. 1995, The Hubble Space Telescope quasar absorption line key project. V. Redshift evolution of Lyman-limit absorption in the spectra of a large sample of quasars, *The Astrophysical Journal*, **444**, 64
- Strittmatter, P. A., & Williams, R. E. 1976, The line spectra of quasi-stellar objects, *Annual Reviews of Astronomy & Astrophysics*, **14**, 307
- Stocke, J. T., Keeney, B. A., Danforth, C. W., et al. 2013, Characterizing the circumgalactic medium of nearby galaxies with HST/COS and HST/STIS absorption-line spectroscopy, *The Astrophysical Journal*, **763**, 148
- Stocke, J. T., Penton, S. V., Danforth, C. W., et al. 2006, The galaxy environment of OVI absorption systems, *The Astrophysical Journal*, **641**, 217
- Stockton, A. N., & Lynds, C. R. 1966, The remarkable absorption spectrum of 3c 191, *The Astrophysical Journal*, **144**, 451
- Storrie-Lombardi, L. J., McMahon, R. G., Irwin, M. J., & Hazard, C. 1994, Evolution of Lyman-limit absorption systems over the redshift range $0.40 < z < 4.69$, *The Astrophysical Journal Letters*, **427**, L13
- Storrie-Lombardi, L. J., & Wolfe, A. M. 2000, Surveys for $z > 3$ damped Ly α absorption systems: The evolution of neutral gas, *The Astrophysical Journal*, **543**, 552

- Stumpff, P. 1977, On the computation of barycentric radial velocities with classical perturbation theories, *Astronomy and Astrophysics*, **56**, 13
- Stumpff, P. 1980, Two self-consistent fortran subroutines for the computation of the Earth's motion, *Astronomy and Astrophysics Supplements Series*, **41**, 1
- Sutherland, R. S. 1998, Accurate free-free Gaunt factors for astrophysical plasmas, *Monthly Notices of the Royal Astronomical Society*, **300**, 321
- Sutherland, R. S., & Dopita, M. A. 1993, Cooling functions for low-density astrophysical plasmas, *The Astrophysical Journal Supplement Series*, **88**, 253
- Suzuki, N., Tytler, D., Kirkman, D., O'Meara, J. M., & Lubin, D. 2003, Relative flux calibration of Keck HIRES echelle spectra, *Publications of the Astronomical Society of the Pacific*, **115**, 1050
- Tennyson, J. 2011, *Astronomical Spectroscopy: An Introduction to the Atomic and Molecular Physics of Astronomical Spectra*, London: World Scientific
- Thom, C., & Chen, H.-W. 2008, A Space Telescope Imaging Spectrograph survey for OVI absorption systems at $0.12 < z \leq 0.5$. II. Physical conditions of the ionized gas, *The Astrophysical Journal Supplement Series*, **179**, 37
- Thompson, T. A., Quataert, E., Zhang, D., & Weinberg, D. H. 2016, An origin for multiphase gas in galactic winds and haloes, *Monthly Notices of the Royal Astronomical Society*, **455**, 1830
- Tripp, T. M., Giroux, M. L., Stocke, J. T., Tumlinson, J., & Oegerle, W. R. 2001, The ionization and metallicity of the intervening OVI absorber at $z = 0.1212$ in the spectrum of H1821+643, *The Astrophysical Journal*, **563**, 724
- Tripp, T. M., Lu, L., & Savage, B. D. 1998, The relationship between galaxies and low-redshift weak Ly α absorbers in the directions of H1821+643 and PG 1116+215, *The Astrophysical Journal*, **508**, 200
- Tripp, T. M., Meiring, J. D., Prochaska, J. X., et al. 2011, The hidden mass and large spatial extent of a post-starburst galaxy outflow, *Science*, **334**, 952
- Tripp, T. M., Savage, B. D., & Jenkins, E. B. 2000, Intervening OVI quasar absorption systems at low redshift: A significant baryon reservoir, *The Astrophysical Journal Letters*, **534**, L1
- Tripp, T. M., Sembach, K. R., Bowen, D. V., et al. 2008, A high-resolution survey of low-redshift QSO absorption lines: statistics and physical conditions of OVI absorbers, *The Astrophysical Journal Supplement Series*, **177**, 39
- Tumlinson, J., Thom, C., Werk, J. K., et al. 2011, The large, oxygen-rich halos of star-forming galaxies are a major reservoir of galactic metals *Science*, **334**, 948
- Tumlinson, J., Thom, C., Werk, J. K., et al. 2013, The COS-Halos survey: Rationale, design, and a census of circumgalactic central hydrogen, *The Astrophysical Journal*, **777**, 59
- Turner, M. S. 2002, Making sense of the new cosmology, *International Journal of Modern Physics A*, **17**, 180
- Turnshek, D. A. 1986, Broad absorption lines QSOs, in *Quasars*, Proceedings of the IAU Symposium 119, Dordrecht: D. Reidel Publishing Co., 317
- Turnshek, D. A., Bohlin, R. C., Williamson, R. L., et al. 1990, An atlas of Hubble Space Telescope photometric, spectrophotometric, and polarimetric calibration objects, *The Astronomical Journal*, **99**, 1243
- Tytler, D. 1982, QSO Lyman limit absorption, *Nature*, **298**, 427
- Tytler, D., Boksenberg, A., Sargent, W. L. W., Young, P., & Kunth, D. 1987, High-resolution spectra of 24 low-redshift QSOs - The properties of MgII absorption systems, *The Astrophysical Journal Supplement Series*, **64**, 667

- Tytler, D., Fan, X.-M., & Burles, S. 1996, Cosmological baryon density derived from the deuterium abundance at redshift $z = 3.57$, *Nature*, **381**, 207
- Vander Vliet, J. R. 2017, Observing the baryon cycle in hydrodynamic cosmological simulations, PhD, New Mexico State University
- Vázquez, G. A., & Leitherer, C. 2005, Optimization of Starburst99 for intermediate-age and old stellar populations, *The Astrophysical Journal*, **621**, 695
- Verner, D. A., & Yakovlev, D. G. 1995, Analytic fits for partial photoionization cross sections, *Astronomy and Astrophysics Supplements Series*, **109**, 125
- Verner, D. A., Ferland, G. J., Korista, K. T., & Yakovlev, D. G. 1996, Atomic data for astrophysics. II. New analytic fits for photoionization cross sections of atoms and ions, *The Astrophysical Journal*, **465**, 487
- Vilkoviskij, E. Y., Yefimov, S. N., Pavlova, L. A., & Baturina, E. B. 2003, Radiation pressure and mass outflows from hot stars and quasars, *Astronomical and Astrophysical Transactions*, **22**, 219
- Vogt, S. S., Allen, S. L., Bigelow, B. C., et al. 1994, HIRES: the high-resolution echelle spectrometer on the Keck 10-m Telescope, *Proceedings of the Society for Optical Engineers*, **2198**, 362
- Wagner, R. M. 1992, Point source spectroscopy, *Astronomical CCD observing and reduction techniques, Publications of the Astronomical Society of the Pacific Conference Series*, **23**, 160
- Wagoner, R. V. 1967, Some effects of an intervening galaxy on the radiation from very distant objects, *The Astrophysical Journal*, **149**, 465
- Wakker, B. P., Savage, B. D., Sembach, K. R., et al. 2003, The Far Ultraviolet Spectroscopic Explorer survey of OVI absorption in and near the Galaxy, *The Astrophysical Journal Supplement Series*, **146**, 1
- Wang, Y., Ferland, G. J., Lykins, M. L., et al. 2014, Radiative cooling II: effects of density and metallicity, *Monthly Notices of the Royal Astronomical Society*, **440**, 3100
- Weedman, D. W. 1986, *Quasar astronomy*, Cambridge: Cambridge University Press
- Weisheit, J. C. 1978, On the use of line shapes in the analysis of QSO absorption spectra, *The Astrophysical Journal*, **219**, 829
- Weymann, R. J., Butcher, H. R., Boroson, T. A., & Peterson, B. M. 1978, An attempt to detect faint objects near quasi-stellar objects with low-redshift absorption systems, *The Astrophysical Journal*, **226**, 603
- Weymann, R. J., Carswell, R. F., & Smith, M. G. 1981, Absorption lines in the spectra of quasistellar objects, *Annual Reviews of Astronomy & Astrophysics*, **19**, 41
- Weymann, R. J., Jannuzi, B. T., Lu, L., et al. 1998, The Hubble Space Telescope quasar absorption line key project. XIV. The evolution of Ly α absorption lines in the redshift interval $z = 0-1.5$, *The Astrophysical Journal*, **506**, 1
- Weiner, B. J., Coil, A. L., Prochaska, J. X., et al. 2009, *The Astrophysical Journal*, Ubiquitous outflows in DEEP2 spectra of star-forming galaxies at $z = 1.4$, **692**, 187
- Wiersma, R. P. C., Schaye, J., & Smith, B. D. 2009, The effect of photoionization on the cooling rates of enriched, astrophysical plasmas, *Monthly Notices of the Royal Astronomical Society*, **393**, 99
- Wingert, D. W. 1975, The absorption-line spectrum of the quasi-stellar object PHL 957, *The Astrophysical Journal*, **198**, 267

- Woodgate, B. E., Kimble, R. A., Bowers, C. W., et al. 1998, The Space Telescope Imaging Spectrograph design, *Publications of the Astronomical Society of the Pacific*, **110**, 1183
- Wolfe, A. M., Broderick, J. J., Condon, J. J., & Johnston, K. J. 1976, 3c 286: A cosmological quasar, *The Astrophysical Journal*, **208**, 47
- Wolfe, A. M., Broderick, J. J., Johnston, K. J., & Condon, J. J. 1978, A comprehensive radio study of the $z = 0.524$ absorption system in AO 0235+164, *The Astrophysical Journal*, **222**, 752
- Wolfe, A. M. & Davis, M. M. 1979, Detection of 21-cm absorption at $z \sim 1.8$ in the quasi-stellar object 1331+170, *The Astronomical Journal*, **84**, 699
- Wolfe, A. M., Turnshek, D. A., Smith, H. E., & Cohen, R. D. 1986, Damped Lyman-alpha absorption by disk galaxies with large redshifts. I - The Lick survey, *The Astrophysical Journal Supplement Series*, **61**, 249
- Wood, B. E., & Linsky, J. L. 1997, A new measurement of the electron density in the local interstellar medium, *The Astrophysical Journal Letters*, **474**, L39
- Yanny, B., & York, D. G. 1992, Emission-line objects near quasi-stellar object absorbers. III - Clustering and colors of moderate-redshift HII regions, *The Astrophysical Journal*, **391**, 569
- York, D. G., Dopita, M., Green, R., & Bechtold, J. 1986, On the origin of some QSO absorption lines, *The Astrophysical Journal*, **311**, 610
- Young, P. J., Sargent, W. L. W., & Boksenberg, A. 1982a, A high resolution study of the absorption spectra of three QSOs: Evidence for cosmological evolution in the Lyman-alpha lines, *The Astrophysical Journal*, **252**, 10
- Young, P. J., Sargent, W. L. W., & Boksenberg, A. 1982b, CIV absorption in an unbiased sample of 33 QSOs: Evidence for the intervening galaxy hypothesis, *The Astrophysical Journal Supplement Series*, **48**, 455
- Young, P. J., Sargent, W. L. W., Boksenberg, A., Carswell, R. F., & Whelan, J. A. J. 1979, A high-resolution study of the absorption spectrum of PKS 2126-158, *The Astrophysical Journal*, **229**, 891
- Zafar, T., Popping, A., & Péroux, C. 2013, The ESO UVES advanced data products quasar sample. I. Dataset and new N_{HI} measurements of damped absorbers, *Astronomy and Astrophysics*, **556**, A140
- Zhang, Y., Anninos, P., & Norman, M. L. 1995, A multispecies model for hydrogen and helium absorbers in Lyman-alpha forest clouds, *The Astrophysical Journal Letters*, **453**, L57
- Zhu, G., & Ménard, B. 2013, The JHU-SDSS Metal Absorption Line Catalog: Redshift evolution and properties of MgII absorbers, *The Astrophysical Journal*, **770**, 130
- Zsargó, J., Sembach, K. R., Howk, J. C., & Savage, B. D. 2003, Highly ionized gas in the Galactic Halo: A FUSE survey of OVI absorption toward 22 halo stars, *The Astrophysical Journal*, **586**, 1019
- Zwicky, F. 1937, On the masses of nebulae and of clusters of nebulae, *The Astrophysical Journal*, **86**, 217