5 Escape of Atmospheres to Space

So far, our discussion of atmospheric evolution has concentrated on atmosphere and climate fundamentals. Climate constrains possible life and, as we will see later in this book, the way that climate is thought to have evolved can explain many environmental differences between Earth, Venus, and Mars. Climate is closely tied to the composition of a planet's atmosphere, which determines the greenhouse effect. Consequently, to understand how climate has changed over time, we must consider how atmospheric composition has evolved. In turn, we must examine how atmospheric gases can be lost.

Gases are lost at an atmosphere's upper and lower boundaries: the planet's surface and interplanetary space. In this chapter, we consider the latter. Studies of the Solar System have shown that some bodies are vulnerable to atmospheric escape (Hunten, 1990). Indeed, many smaller objects, e.g., most moons and essentially all asteroids, are airless because of escape, making the theory of atmospheric escape crucial for explaining differences in surface volatiles. Escape processes can help us understand the lack of atmospheres on the Moon and Mercury, the barren nature of the Galilean satellites versus Titan (Griffith and Zahnle, 1995; Gross, 1974; Zahnle *et al.*, 1992), why the atmosphere of Mars is thin (Brain and Jakosky, 1998; Melosh and Vickery, 1989; Zahnle, 1993b), the red color of the Martian surface (Hartman and McKay, 1995; Hunten, 1979c), the lack of oceans on Venus (Kasting and Pollack, 1983) (see Ch. 13), and possibly the oxidizing nature of the Earth's atmosphere and surface (Catling *et al.*, 2001) (See Ch. 10).

We can group various types of atmospheric escape into three categories following Catling and Zahnle (2009). (i) *Thermal escape* is when irradiation from a parent star (or, less commonly, a very high heat flux from a planet or moon interior) heats an atmosphere, causing atmospheric molecules to escape to space. Two end-member approximations of thermal escape are appropriate under different circumstances: *Jeans' escape*, where individual molecules evaporate into a collisionless exosphere, and *hydrodynamic escape*, which is a bulk outflow with a velocity driven by atmospheric heating that induces an upward pressure gradient force (e.g., Johnson *et al.*, 2013d; Walker, 1982). (ii) *Suprathermal* (or *nonthermal*) *escape* is where individual atoms or molecules are boosted to escape velocity because of chemical reactions or ionic interactions. Finally, (iii) *impact erosion* is where atmospheric gases are expelled *en masse* as a result of large body impacts, such as the cumulative effect of asteroids hits. Of these three types, nonthermal escape is generally slow because if it were fast the molecules would collide and the escape would be in the thermal category. Theory suggests that the two mechanisms that can most efficiently cause substantial atmospheric loss are hydrodynamic escape driven by stellar irradiation (Lammer *et al.*, 2008; Sekiya *et al.*, 1981; Sekiya *et al.*, 1980a; Watson *et al.*, 1981; Zahnle *et al.*, 1990; Zahnle and Kasting, 1986) and impact erosion (Griffith and Zahnle, 1995; Melosh and Vickery, 1989; Walker, 1986; Zahnle *et al.*, 1992). In addition, hydrodynamic escape from early hydrogen-rich atmospheres on the terrestrial planets is relevant for observations of noble gases and their isotopes, as discussed in Ch. 6, because such escape can drag along heavier gases.

In this chapter, we focus particularly on the escape of hydrogen, for two reasons. First, hydrogen is the lightest gas and consequently the most prone to escape. Second, later in the book, we will see that substantial loss of hydrogen can affect the redox chemistry of a planet's atmosphere and surface, changing the chemical character of a planet. Rocky planets, as a whole, become more oxidized when hydrogen escapes to space. This oxidation occurs irrespective of whether the hydrogen is transported through the atmosphere as H_2, H_2O, CH_4, HCN, NH_3, or some other H-bearing compound. Oxidation occurs

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

because the hydrogen atom that escapes ultimately derives from some oxidized form of hydrogen such as water (H_2O), water of hydration in silicate rocks (–OH), or hydrocarbons (–CH). It was in these compounds that hydrogen was originally incorporated into planets like the Earth. Consequently, when hydrogen escapes, matter somewhere on a planet's surface or subsurface is irreversibly oxidized.

Oxidation is most obvious if we consider hydrogen that escapes after atmospheric water vapor undergoes photolysis. Consider water vapor photolysis and escape in the upper atmosphere of the Earth. In this case, the oxygen left behind can oxidize the Earth's surface so that any further oxygen produced (by photolysis and hydrogen escape) is less likely to be taken up by the crust and more liable to remain in the atmosphere. However, today's abiotic production rate of oxygen is ~10^2 times smaller than the rate of O_2 production from photosynthesis and, hence, plays a negligible role in the atmospheric oxygen budget. It is nonetheless important to understand such abiotic oxygen, both because of its possible effect on very early life on this planet and because of its future significance in interpreting spectra that may be obtained from exoplanets.

The effect of the escape of hydrogen in oxidizing surfaces is also widely considered to be responsible for the oxidized states of Venus and Mars, as illustrated by the red color of the Martian surface (Hartman and McKay, 1995; Hunten, 1979c). Ancient hydrogen escape has also been proposed as a means of oxidizing the Earth's atmosphere, crust, and mantle (Catling *et al.*, 2001; Kasting *et al.*, 1993a; Zahnle *et al.*, 2013) (see Ch. 10).

5.1 Historical Background to Atmospheric Escape

The idea of the escape of gases from the Earth's atmosphere is as old as kinetic theory and has an unusual history. A Scottish amateur scientist, John Waterston (1811–1883), first developed a theory of gases in which the mean kinetic energy of each species was proportional to temperature, and he also introduced the notion of atmospheric escape (Haldane, 1928, pp. 209–210). However, the Royal Society rejected Waterston's paper describing kinetic theory in 1845, and it remained unknown until Lord Rayleigh rediscovered the manuscript in 1891. By then, Waterston's ideas had been overtaken by the work of Clausius, Maxwell, and Boltzmann, while Waterston disappeared in 1883, presumed to have drowned near Edinburgh.

Later, the Irish physicist George Stoney (who gave us the term *electron*) understood that a few gas particles in the high-velocity tail of a Maxwell–Boltzmann distribution of velocities would have sufficient energy to escape from a planet's upper atmosphere even if an average particle did not (Stoney, 1898, 1900a, b, c, 1904). This process is nowadays called *Jeans' escape* after Sir James Jeans, who described its physics in *The Dynamical Theory of Gases* (1954, first edition 1904). At that time, in the early twentieth century, balloon soundings in Earth's lower atmosphere were extrapolated to the entire upper atmosphere, which was assumed to be isothermal at ~220 K. The hot, 1000–2000 K thermosphere was unknown. Consequently, Jeans incorrectly calculated an exceedingly low escape rate of hydrogen.

Later, the Space Age provided data from rocket soundings. As a result, in the next major treatment of atmospheric escape, Spitzer (1952) corrected Jeans' earlier mistake by using more realistic thermospheric temperatures. From the 1950s to the present day, data have become directly available on the number density of hydrogen and the temperature in the upper atmosphere. Measurements include satellite drag through the thermosphere, in situ mass spectrometer measurements, and images of the *geocorona*, which is a glow at the Lyman-α wavelength (121 nm) caused by resonant scattering of solar ultraviolet (UV) by a cloud of atomic hydrogen that surrounds the Earth. UV images taken by spacecraft show the hydrogen atoms. Atoms are on ballistic trajectories back to Earth, escaping, or in orbit (Fig. 5.1).

For astrobiology, we note that about half of the hydrogen atoms seen in Fig. 5.1 derive from decomposition of methane (CH_4), ~90% of which enters the atmosphere from the biosphere. Most of the other half of the H atoms originates from the photodissociation of water vapor. In Fig 5.1, we catch a glimpse of some of the 93 000 tonnes of hydrogen that escape each year (or 3 kg/s) from the Earth.

In the past 60 years, planetary exploration and astronomy have widened our perspective of both atmospheric escape and of *aeronomy*, the study of processes in the rarefied atmosphere from the stratosphere to interplanetary space. Space science led to the recognition of suprathermal escape, hydrodynamic escape, and impact erosion, as discussed in various reviews (e.g., Ahrens, 1993; Chamberlain, 1963; Hunten, 1990, 2002; Hunten and Donahue, 1976; Hunten *et al.*, 1989; Johnson *et al.*, 2008c; Lammer, 2013; Shizgal and Arkos, 1996; Strobel, 2002; Tinsley, 1974; Walker, 1977). Recently, the discovery of exoplanets has made atmospheric escape a fundamental consideration in understanding exoplanetary

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Table 5.1 Mechanisms for the escape of atmospheric gases and ions.

Impact erosion (different approximations)	Thermal escape (different end-member approximations)	Suprathermal (or nonthermal) escape (different mechanisms)
Walker "cookie cutter"	Jeans' escape	Photochemical escape
Ahrens"'bomb analogy"	Hydrodynamic escape	Charge exchange
Melosh "tangent plane"		Ion pickup
		Sputtering
		The polar wind
		Bulk removal

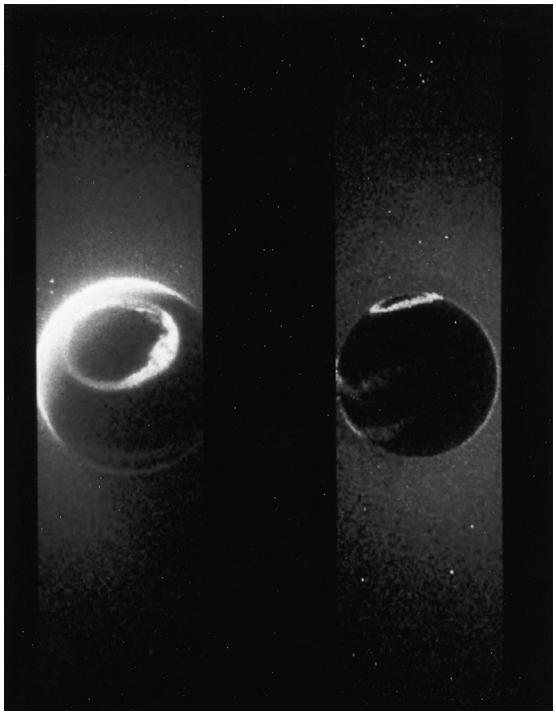

Figure 5.1 Earth imaged in the vacuum ultraviolet (VUV) by NASA's *Dynamics Explorer 1* (Rairden *et al.*, 1986). *Left panel*: View with the spacecraft at 16 500 km altitude above 67° N latitude at 2017 UT on October 14, 1981. Glow beyond the limb of the planet (red false color) is due to Lyman-α (121 nm) solar radiation resonantly scattered by Earth's extended hydrogen atmosphere or *geocorona*. Energetic hydrogen atoms in the geocorona are escaping to space. Features on the Earth's disk (dayglow from the sunlit atmosphere, a northern auroral oval, and equatorial airglow) are due to the emission of atomic oxygen at 130.4 and 135.6 nm and emission in the Lyman–Birge–Hopfield band of N_2 (140–170 nm). Isolated points of light are background stars that are bright in the VUV. *Right panel:* A view of Earth's dark hemisphere at 0222 UT on February 16, 1982, with the Sun behind Earth. Spacecraft altitude and latitude are 19,700 km and 13° N, respectively. Equatorial airglow straddles the magnetic equator in the pre-midnight sector. (Image credit: NASA.) (A black and white version of this figure will appear in some formats. For the color version, please refer to the plate section.)

atmospheres and whether they vanish, persist, or change composition (Koskinen *et al.*, 2014; Lammer *et al.*, 2003b; Luger and Barnes, 2015; Luger *et al.*, 2015; Owen and Jackson, 2012).

5.2 Overview of Atmospheric Escape Mechanisms

Table 5.1 summarizes the three principal categories of escape of atmospheric gases. Below, we give a brief overview of each type of escape. Then, the rest of the chapter examines the physics describing common ways that atmospheric gases escape, with particular emphasis on the two forms of thermal escape.

5.2.1 Thermal Escape Overview

Thermal escape is when heating of an atmosphere allows molecules to escape. In basic models, the theory assumes neutral species with a Maxwellian velocity distribution, which occurs when collisions between molecules are frequent. The "Jeans' escape" and "hydrodynamic escape" end-member approximations to thermal escape apply under different circumstances of atmospheric heating that we summarize below and discuss in further detail in Sec. 5.10.1.

Jeans' escape is when a relatively small number of high-energy molecules in the tail of the thermal distribution of velocities of molecules have sufficient kinetic energy to escape into a nearly collisionless exosphere from the collisional atmosphere below (see Sec. 5.6 for the physics). This process is important for the loss of hydrogen, a low-mass species that more easily attains escape speed at a given temperature. As such, Jeans' escape was likely influential in the atmospheric evolution of all the early terrestrial planets. Jeans' escape currently accounts for a non-negligible fraction of hydrogen escaping from Earth, Mars, and Titan, but it is negligible for

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Venus because of a cold upper atmosphere combined with relatively high gravity.

Hydrodynamic escape occurs when heating in the collisional region of an atmosphere causes an upward pressure gradient force that drives a bulk, radial outflow (see Sec. 5.10 for the physics). Under such collisional circumstances, the pressure force can remain active up to very high altitudes with the result that the whole upper atmosphere expands as a fluid into space and gases attain escape velocity.

Hydrogen-rich atmospheres on relatively low-gravity rocky planets or very hot hydrogen-rich atmospheres on bigger planets are susceptible to hydrodynamic escape, which can drag along heavier gases in a way that is moderately mass fractionating (see Sec. 5.11). Shklovskii (1951) and Öpik (1963) first discussed the concept that heavy gases might be dragged along by a large hydrogen escape flux from primitive atmospheres but, compared with other escape mechanisms, hydrodynamic escape only received limited attention prior to the 1980s (Gross, 1972; McGovern, 1973; Ziering and Hu, 1967; Ziering *et al.*, 1968). The lack of attention is probably because hydrodynamic escape was not thought to be active on any planet-sized body in the Solar System.

However, increasing evidence from within our Solar System and beyond suggests that warm hydrogen-rich atmospheres are prone to undergo hydrodynamic escape. The depletion of some light isotopes of noble gases in the atmospheres of Earth, Venus, and Mars, suggests that hydrodynamic escape may have operated very early in Solar System history if the very earliest atmospheres on these planets had been composed of a significant fraction of hydrogen in any chemical form. Beyond our Solar System, the gas giant HD 209458b, which orbits a Sun-like star at 0.05 AU, has hot H atoms beyond its Roche lobe, containing O, C^+, and Si^{2+}, presumably dragged there by hydrodynamic hydrogen flow (Linsky *et al.*, 2010; Vidal-Madjar *et al.*, 2003; Vidal-Madjar *et al.*, 2004; Vidal-Madjar *et al.*, 2008). Near-UV spectra suggest that hot Jupiter WASP-12B also has metals in its Roche lobe (Fossati *et al.*, 2013; Fossati *et al.*, 2010; Haswell *et al.*, 2012).

Hydrodynamic escape and Jeans' escape are both approximations to thermal escape in different ways. Hydrodynamic escape is approximate because it neglects the fact that at very high altitudes there will eventually be few collisions and so a fluid description of the flow becomes invalid. Classical Jeans' escape neglects distortions of the particle velocities away from a Maxwellian distribution because of mass motion.

Specific conditions indicate when it is appropriate to apply the approximations of classical Jeans escape or purely fluid flow for hydrodynamic escape. The Jeans case applies when the atmosphere is essentially hydrostatic and when stellar heating of a thermosphere results in production and loss of electrons and ions, which conduct heat down to the lower thermosphere and mesopause where heat is efficiently radiated away (see Sec 1.1.1). Such atmospheres have roughly isothermal upper thermospheres. Jeans evaporation of atoms or molecules occurs from a static atmosphere into an essentially collisionless exosphere.

A hydrodynamic situation can occur when the heating of an upper atmosphere is strong enough to drive a bulk outflow. The bulk upward flow can attain the speed of sound in the collisional domain at an altitude called the *sonic level*. The speed of sound u_{sound}, can be compared with the root mean square speed of the molecules from kinetic theory u_{rms}, as follows,

$$u_{\text{sound}} = \left(\frac{\gamma p}{\rho}\right)^{1/2} u_{\text{rms}} = \left(\frac{3p}{\rho}\right)^{1/2} \tag{5.1}$$

where p is pressure, ρ is density, and γ is the ratio of specific heats. Because $\gamma \sim 1.4$ for linear diatomic gases (e.g., H_2), we can see from comparing u_{sound} and u_{rms} that gas traveling at the speed of sound moves at a velocity similar to the mean thermal velocity of molecules, which is responsible in kinetic theory for providing the pressure of a gas (i.e., $p = (1/3)\rho u_{\text{rms}}^2$). In such a fast-moving fluid, a pressure gradient drives an upward bulk flow and the velocity increases above the sonic level to supersonic and then escape velocity. Because the density of the atmosphere decreases with altitude, and matter must be conserved, the flow velocity in such a case increases with altitude in order to maintain a constant mass flux [kg s^{-1}] through ever-larger planet-centered spheres. Under these circumstances, the fluid equations of hydrodynamic theory are reliable approximation, as noted by Walker (1977 pp. 149–151; 1982). The vertical profiles of density and velocity are relatively unaffected when the transition to the nearly collisionless domain occurs above the level from sonic to supersonic flow (Holzer *et al.*, 1971). Figure 5.2 shows a schematic diagram of these two end-member cases of thermal escape: Jeans' escape and transonic hydrodynamic escape.

In hydrodynamic escape, the temperature profile depends on the balance of adiabatic cooling from the expansion of the atmosphere and absorption of stellar radiation. If adiabatic cooling dominates, atmospheric temperature can decline with increasing altitude. However, temperature can also increase with altitude if

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

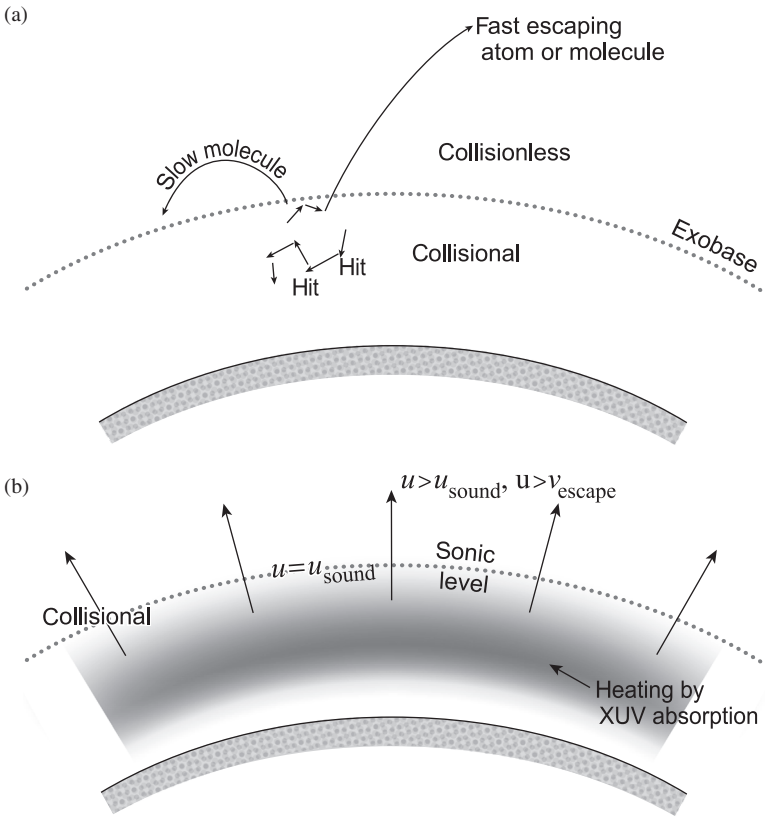

Figure 5.2 Schematic diagram of thermal escape end-members. (a) *Jeans' escape* is escape of molecules or atoms with an upward-directed component of velocity bigger than the escape velocity from the collisional part of the atmosphere into a virtually collisionless exosphere. (b) *Transonic hydrodynamic escape* is where the upper atmosphere has sufficient heating to produce a radial outward velocity u, at the speed of sound (u_{sound}) in the collisional part of the atmosphere at the radius labeled *sonic level*. Heating is typically by soft x-rays and extreme ultraviolet radiation (EUV). The atmosphere flows out to the vacuum of space driven by a pressure gradient (with the boundary condition set at the sonic level) and the supersonic flow reaches escape velocity, v_{escape}. At very high levels, the atmosphere will have very few collisions and the fluid assumption of hydrodynamic escape breaks down. However, under the transonic circumstances depicted, the fluid equations extrapolated to infinity provide a reasonable approximation to the density, temperature, and flow profiles.

absorption of stellar radiation is strong (e.g. Sekiya *et al.*, 1981; Sekiya *et al.*, 1980a).

When heating is smaller, or if there exists a sufficient backpressure at the top of the atmosphere, the outgoing flow may remain *subsonic* at all levels. An example of the latter might be escape in the direction towards the parent star, where eventually the escaping, and partly ionized, gas encounters the bow shock formed by interaction with a strong stellar wind. Perhaps more importantly, in an atmosphere that is weakly heated, or in which hydrogen is not sufficiently abundant, the atmosphere expands, but it is invalid to apply purely fluid hydrodynamic equations because the atmosphere becomes increasingly collisionless above some level. The application of hydrodynamic equations is a reasonable approximation if the mean free path remains smaller than the density scale height below the sonic level. But in the case when expanding gas becomes collisionless without reaching the speed of sound, it still exerts pressure, but that pressure cannot be calculated in the normal statistical way, which assumes a Maxwellian velocity distribution. Sophisticated models can be used to treat the transition from collisional to rarefied domains in order to calculate a realistic temperature density, temperature and flow structure. Such methods are based on the *Boltzmann equation*, which in its most general form is an equation for the time evolution of the velocity distribution function of species in a gas mixture as a result of external forces and collisions (see Ch. 3 of Schunk and Nagy, 2009).

Using the Boltzmann equation for a single-component atmosphere, Merryfield and Shizgal (1994) found that escape can be fractionally (~30%) greater than Jeans' escape due to streaming of particles from the heavy, denser regions below. Another such model of gas flow is "direct simulation Monte Carlo" (DSMC), in which a large set of particles is followed subject to collisions, heating and gravity (Bird, 1994; Volkov *et al.*, 2011a). DSMC models show that purely hydrodynamic models that were once applied to N_2 escape from Pluto (where a sonic level probably does not occur in the collisional domain) produce an erroneous atmospheric structure of temperature and density (Erwin *et al.*, 2013; Johnson *et al.*, 2013d). Another key prediction of DMSC models, at least for single component atmospheres, is a

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

sharp transition from "Jeans-like" escape to rapid, transonic hydrodynamic escape as the thermal energy of upper atmosphere gas molecules is increased relative to their gravitational binding energy (Volkov and Johnson, 2013; Volkov *et al.*, 2013; Volkov *et al.*, 2011b).

The physics of hydrodynamic escape of planetary atmospheres is analogous to the solar wind – the fully ionized, electrically neutral plasma that is a supersonic expansion of the solar corona as a result of a pressure difference between the corona and the interstellar medium. (See Ch. 4 of Kivelson (1995) for an introduction to the solar wind.) Essentially, the solar corona – the plasma that we see around the Sun in a total solar eclipse – is so hot that the Sun's gravity cannot hold on to it and it flows out as the solar wind. Consequently, hydrodynamic escape of a planetary atmosphere is sometimes called a *planetary wind*. However, there are differences. The key one is that the fluid description of solar wind in the vacuum of space works because of quasi-collisional effects caused by electromagnetic interactions between charged particles, whereas neutral species generally dominate planetary atmospheres and cannot be assumed to behave like that.

Finally, we note that some authors use the term *blow-off* interchangeably with "hydrodynamic escape" (e.g., Volkov *et al.* (2013); Hunten (1973)). Unfortunately, there are different definitions about the meaning of the term blowoff. Chamberlain and Hunten (1987, p. 377) state that blowoff is "when an escaping light gas is able to carry heavier constituents along with it." Lammer (2013, p.30) describes blowoff as when "the whole exosphere evaporates" because "the mean thermal energy...of gases at the exobase level exceeds their gravitational energy." Because of such difference in definitions, in the rest of the chapter, we avoid the term blowoff.

5.2.2 Suprathermal (or Nonthermal) Escape, in Brief

Suprathermal escape refers to loss processes that affect either neutral species or ions that attain a velocity significantly greater than that corresponding to the background neutral temperature. Consequently, suprathermal escape is also called *nonthermal escape* (Hunten, 2002), Most suprathermal processes involve ions, which may themselves have a Maxwellian velocity distribution but with a temperature exceeding that of the neutral population. Various types of suprathermal escape are as follows.

Photochemical escape occurs when atoms resulting from various photochemical reactions attain sufficient energy to escape to space. Such escape can happen when a neutral species is photoionized by solar EUV radiation and recombines with an electron to form a fast neutral atom. Photochemical escape is important for the loss of C, O, and N from Mars (Sec 12.2.4).

Charge exchange is where a fast ion can impart its charge to a neutral atom through collision, and become a fast neutral atom with escape velocity (Sec 5.7, below). In today's terrestrial atmosphere, charge exchange is usually the dominant mechanism for hydrogen escape, although it is exceeded by Jeans' escape at solar maximum.

Ion pickup occurs when atmospheric ions are exposed to an electric field from the magnetized solar wind. Atmospheric particles are ionized either by solar UV radiation (photoionization) or by charge exchange. Acceleration of ions due to the electric field can cause some ions to reach escape velocity, whereas others head into the atmosphere where they may cause sputtering, as we describe next.

Sputtering occurs when ions that have been picked up by the magnetic field embedded in the solar wind impact a planetary atmosphere and undergo charge exchange. Charge exchange neutralizes the ions, which can impart their large energies to surrounding particles by collision. Upward-directed energetic particles can then escape. This process may have been important on early Mars, after it lost its magnetic field and was no longer shielded from the solar wind.

The polar wind is a stream of hydrogen ions (protons) that flows upward near the poles where Earth's magnetic field lines are more or less vertical. These field lines do not necessarily reconnect, or they do so only sporadically, and so these hydrogen ions are eventually swept away by the solar wind.

Bulk removal is caused by instabilities at the solar wind–atmosphere interface that can strip away large portions of ionized atmosphere (Perez-de-Tejada, 1987) or cause ion outflow (Hartle and Grebowsky, 1990) from planets, such as Mars and Venus, that lack a protective magnetic field. This process is currently poorly understood.

5.2.3 Impact Erosion, in Brief

Impact erosion occurs when the hot vapor plume or high-speed ejecta associated with a large asteroid or comet impact imparts sufficient kinetic energy to atmospheric molecules for them to escape en masse (Sec. 5.12). The impactor is vaporized along with part of the target body. This erosion process affects smaller target bodies more strongly than larger ones and could have been important for removing virtually the entire early Martian

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

atmosphere (Sec. 12.3.3). Impact erosion may also explain why Titan has a thick atmosphere whereas the Galilean moons of Jupiter, which are subject to more energetic impacts, remain largely barren (Zahnle *et al.*, 1992).

5.2.4 The Upper Limit of Diffusion-Limited Escape, in Brief

Diffusion-limited escape of hydrogen is an *upper limit* on the escape rate set when the escape rate of hydrogen is not controlled by processes at high altitude but is regulated by the rate that hydrogen can diffuse up from the lower atmosphere (see Sec. 5.8.3, 5.8.4 and 5.8.9 for details). For example, in Earth's current atmosphere, hydrogen does not simply escape via Jeans' escape determined by the temperature of the exobase. Instead, both Jeans' escape and suprathermal processes remove hydrogen rapidly from the exobase, and the rate-limiting step is the relatively slow upward diffusion of hydrogen through the layer of background air between the homopause and exobase. Basically, the hydrogen escape rate is limited both by the supply of hydrogen from below and upward diffusion. Diffusion-limited escape is a remarkably successful theory that appears to apply to hydrogen escape from the current atmospheres of Earth, Venus, Mars, and Titan, as well as can be determined from the available data (Sec. 5.8.4 and Sec. 5.9).

Diffusion-limited escape can also apply as an *upper limit* to hydrodynamic escape. In this case, the upper limit on the rate of escape is set by diffusion of hydrogen through a layer of background air of heavier species beneath the level where hydrogen is accelerated radially outwards because of heating caused by the absorption of shortwave light from the parent star.

5.3 Breakdown of the Barometric Law

We begin our more detailed description of atmospheric escape by showing that the barometric law, which describes the vertical pressure distribution in the lower parts of Earth's atmosphere, must break down at some altitude above the surface. Our discussion follows that of Walker (1977), pp. 147–151.

In Ch. 1, we showed that (averaged over horizontal distances of several km), atmospheric pressure varies with altitude z according to the *hydrostatic equation* (eq. (1.12)):

$$p(z) = p_{surf} \exp\left(-z/H_a\right) \tag{1.12}$$

Here p_{surf} is the surface pressure and $H_a = kT/(mg)$ is the atmospheric scale height, where k is Boltzmann's constant, T is temperature, m is mean molecular mass, and g is gravitational acceleration.

We can write eq. (1.12) in differential form as follows:

$$\frac{1}{p}dp = -\frac{mg}{kT}dz = -\frac{1}{H_a}dz \tag{5.2}$$

To extend this relation high up in the atmosphere, the variation of g with altitude must be considered according to $g = GM/r^2$, where G is the universal gravitational constant (6.672×10^{-11} N m^2 kg^{-2}), M is the mass of the planet, and r is the radial distance from a planet's center. Then, eq. (5.2) may be rewritten as

$$\frac{1}{p}dp = -\frac{GMm}{r^2kT}dr \tag{5.3}$$

Integrating from the surface at radial distance r_{surf} up to radial distance r yields

$$p(r) = p_{surf} \exp\left[\frac{GMm}{kT}\left(\frac{1}{r} - \frac{1}{r_{surf}}\right)\right] \tag{5.4}$$

Now, consider what happens as $r \to \infty$. In this case, we obtain

$$p_\infty = p_{surf} \exp\left(-\frac{GMm}{kTr_{surf}}\right) \tag{5.5}$$

If this result were valid, it would imply that the pressure at infinity is small but finite; hence, the atmosphere would have infinite mass. For example, using m = 29 atomic mass units (the mean molecular mass in the Earth's lower atmosphere), $M = 5.97\times10^{24}$ kg, a temperature typical for the thermosphere of T = 1000 K, and $r_{surf} = 6371\times10^3$ m, eq. (5.5) predicts that $p_\infty/p_{surf} \sim 2\times10^{-95}$. If we take m = 1 a.m.u. instead, reflecting the fact that the uppermost atmosphere is composed mostly of atomic hydrogen, we get $p_\infty/p_{surf} \sim 5\times10^{-4}$. Neither result is physically realistic, but they demonstrate that the atmospheric pressure would be significant at high altitudes were the barometric law to apply in this way.

In fact, the actual pressure at the upper boundary of the Earth's atmosphere depends on location. It is highest in the sunward direction where the solar wind impinges on the magnetosphere, i.e., the region where the Earth's magnetic field dominates. At the subsolar *magnetopause*, which is at a distance of ~10 Earth radii, the ram pressure of the solar wind is ~3 nPa and balances plasma pressure within the magnetosphere. By comparison, the calculated pressures at infinity for the two cases above are ~10^{-90} Pa and 50 Pa, respectively. For the pure atomic H case, the

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

solar wind would not be able to supply the necessary backpressure even in the sunward direction. In the anti-sunward direction, the effective backpressure should be essentially zero. The barometric law obviously cannot apply at great distances from the Earth.

How can the breakdown of the barometric law be resolved? Two different possibilities exist: one that applies to the present terrestrial atmosphere and one that may have applied to the primitive atmospheres of the terrestrial planets.

For the present atmosphere, the key is that the atmosphere becomes virtually collisionless at some altitude. Once the atmosphere becomes collisionless, the atmosphere is no longer in a hydrodynamic regime of "continuum flow" but "free molecular flow." Equations (5.3)–(5.5) no longer apply once the velocity distribution in the collisionless region deviates from a Maxwellian distribution. The *exobase* or *critical level* is the altitude above which the virtually collisionless region occurs.

A different situation may have existed for Earth's primitive atmosphere shortly after it formed at 4.5 Ga and possibly for hundreds of millions of years afterwards. This very early atmosphere is thought to have been hydrogen-rich (see Sec. 6.5.2). In this case, absorption of EUV and x-rays in the upper atmosphere should have driven a bulk outflow from the upper atmosphere, which was not hydrostatic. Upwards-flowing hydrogen would have been pushed along by a pressure differential under hydrodynamic escape. The atmosphere literally should have expanded into the vacuum of space. We discuss the physics of *hydrodynamic escape* in detail in Sec. 5.10.

The fractional decrease of the upward flux of particles, Φ, from distance $r' = r$ to $r' = \infty$ due to collisions is given by

$$\frac{\Phi_\infty}{\Phi_r} = \exp\left[-\int_r^\infty n(r')\sigma_c dr'\right] \tag{5.6}$$

Here, σ_c is the collision cross-section of a molecule and $n(r')$ is the number density of *all* molecules at distance r'. The decrease of upward-directed particles caused by collisions is analogous to the Beer–Lambert–Bouguer Law in radiative transfer (Sec. 2.4.2.1). Thus, the term in the exponential is analogous to optical depth, except that the photon absorption cross-section has been replaced by a molecular collision cross-section and we deal with molecules instead of photons.

If the probability of escaping above the exobase is e^{-1} (i.e., $\Phi_\infty/\Phi_r = e^{-1}$) then the term inside the square brackets of eq. (5.6) must equal minus one for $r = r_{exob}$, i.e.,

$$\int_{r_{exob}}^\infty n(r')\sigma_c dr' = 1 \tag{5.7}$$

From eq. (1.21), the above integral is related to the atmospheric molecular column density, $N(r)$, overlying the exobase at radius r_{exob} by

$$N(r) = \int_{r_{exob}}^\infty n(r')dr' \cong n(r_{exob})H_a(r_{exob}) \tag{5.8}$$

Thus, at the exobase, the number density n and scale height H_a, respectively, are:

$$n(r_{exob}) \equiv \frac{1}{\sigma_c H_a(r_{exob})},\ H_a(r_{exob}) \equiv \frac{1}{\sigma_c n(r_{exob})} = \text{m.f.p.} \quad \text{for air of density } n(r_{exob}) \tag{5.9}$$

5.4 The Exobase or "Critical Level"

The *exosphere* is the uppermost layer of an atmosphere that is essentially collisionless. This means that the mean free path is so long that collisions can largely be neglected. We denote the height of the bottom of the exosphere, i.e., the exobase, as the radius r_{exob}, above a planet's center. The *exobase* is defined as the height where a proportion e^{-1} (~1/3) of fast, upward-directed particles experience no collisions and hence escape. Equivalently, the exobase can be defined as the altitude at which the mean free path of a molecule (in the horizontal direction) is equal to the local scale height. (The reason for stating "horizontal direction" is explained below.)

In the expression for the local scale height in eq. (5.9), the right-hand side is approximately the definition of mean free path ℓ_{mfp}, for air with number density $n(r_{exob})$, i.e., ℓ_{mfp} along a horizontal path at an altitude at radius r_{exob}. This explains the exobase definition given earlier. Strictly, for molecules with a Maxwellian velocity distribution, the relationship is $\ell_{\text{mfp}}(r) = 1/\left(\sqrt{2}n\sigma_c\right)$, but we ignore the $\sqrt{2}$ factor.

Another common way of quantifying the importance of collisions is the *Knudson number*, K_n, which is defined as follows:

$$K_n(r) = \frac{\ell_{\text{mfp}}(r)}{H_a(r)} = \frac{\text{mean free path}}{\text{local scale height}} \tag{5.10}$$

The Knudson number grows with altitude and the exobase occurs where $K_n = 1$.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Following Jeans (1954), we can combine eqs. (5.8) and (5.9) to write the condition for the exobase as

$$\sigma_c N(r_{exob}) = 1 \tag{5.11}$$

The thoughtful reader will reflect that the exobase where particles can escape is analogous in its definition to the level of unity optical depth where photons can escape (Sec. 2.4.4), merely by swapping photons for molecules, as noted above.

The number density at Earth's exobase can be estimated from eq. (5.9), taking $\sigma_c \sim 3\times10^{-15}$ cm^2. Gravity will be smaller at high altitude, i.e., at 500 km altitude, r_{exob} = 6870 km and g = 8.44 m s^{-2}. Thus, $H_a = RT/Mg$ = (8.314 J mol^{-1} K^{-1})(1500 K)/[(0.016 kg mol^{-1})(8.44 m s^{-2})] ≅ 92 km, or roughly 100 km (= 10^7 cm), where we use a molar mass appropriate for atomic oxygen, the dominant constituent of the thermosphere. Thus, $n_{exob} \cong 3.6\times10^7$ cm^{-3}, which is roughly similar to the value at the exobase on most planets and satellites. For this exobase number density, the corresponding altitude on Earth varies from solar minimum to solar maximum, but is typically around 450–500 km.

Of course, the exobase is an idealized concept. We have assumed that the transition from continuum flow to free molecular flow is sharp, but in reality it is gradual. However, we will show in Sec. 5.5.1 that defining the exobase in this way does not significantly affect calculations. It is also important to recognize that the dominant species at today's terrestrial exobase is atomic oxygen, which does not escape. The O atoms provide a static background from which hydrogen atoms evaporate off into space. The process of evaporation is described in Sec 5.6 below.

5.5 Escape Velocity

In Earth's present atmosphere, hydrogen atoms are lost to space by reaching escape velocity at the exobase. The escape velocity v_e, is attained when the kinetic energy of an atom of mass m equals its gravitational potential energy:

$$\frac{1}{2}mv_e^2 = \frac{GMm}{r} \tag{5.12}$$

or

$$v_e = \sqrt{\frac{2GM}{r}} \tag{5.13}$$

Here, M is the mass of the planet, G the universal gravitational constant, and r the radius from the center of the planet.

Atoms that have speeds in excess of v_e and whose velocities are directed upwards have a $1/e$ (i.e., 37%) probability of escaping at the exobase by avoiding collisions above it. For Earth, eq. (5.13) gives $v_e \sim 10.8$ km s^{-1} at the exobase. Escape velocities for other planets are 5.0 km s^{-1} for Mars, 10.4 km s^{-1} for Venus, and 60.2 km s^{-1} for Jupiter.

Two salient points arise from eq. (5.13). First, for planets of roughly similar mean density, we find $v_e \propto \sqrt{R_p{}^3/R_p} \propto R_p$, where R_p is the planetary radius. Thus, Mars' escape velocity is roughly half that of Earth or Venus because Mars has about half the diameter of these planets. But the escape energy is ~1/4 as high for Mars as for Earth or Venus because the energy is $\propto v_e{}^2$. Second, although eq. (5.13) shows that the escape velocity is independent of the mass of the escaping molecule, lighter molecules more easily attain escape velocity for a given kinetic temperature than heavier molecules. This is a consequence of the *equipartition theorem* in kinetic theory for an ideal gas with purely translational energy. This theorem states that molecules of different masses have the same average kinetic energy, given that the mixture of gases has a well-defined temperature.

The mean thermal velocity of a hydrogen atom at Earth's exobase can easily be calculated. If we assume that atoms at the top of the atmosphere are in thermal equilibrium, the mean thermal energy of particles is (3/2) kT, so that the kinetic energy is:

$$\frac{1}{2}mv^2 = \frac{3}{2}kT \tag{5.14}$$

Taking a thermosphere temperature of ~1000 K, the root mean square speed is $v = (3kT/m)^{1/2}$ = [3×(1.38×10^{-23} J K^{-1})×(1000 K)/(1.67×10^{-27} kg)]$^{1/2}$ ≈ 5 km s^{-1}. This speed is lower than the escape speed of 10.8 km s^{-1} at the exobase. Thus, even the lightest atom, hydrogen, does not have sufficient mean energy to escape from Earth. Instead, it is energetic atoms in the tail of the Maxwell–Boltzmann distribution of velocities that escape rather than those with typical speed, as illustrated schematically in Fig. 5.3.

Given that only the energetic atoms escape, integration over the Maxwell–Boltzmann velocity distribution is needed to calculate the escape flux of a particular gas, as discussed below in Sec. 5.6. A rule-of-thumb approach found in elementary textbooks is that if the average speed of a molecule or atom exceeds one-sixth of the escape speed, then escape is generally possible for that species. The average speed obviously depends on the temperature at the exobase. Consequently, Fig 5.4 illustrates the

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

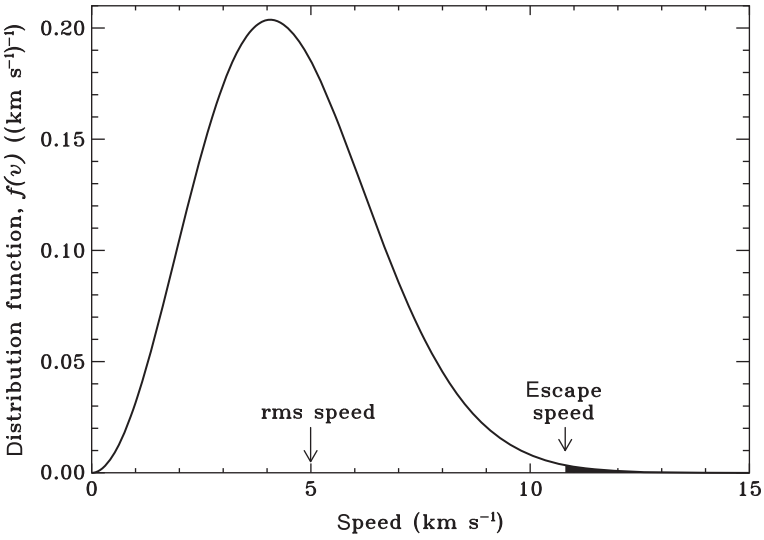

Figure 5.3 The Maxwell–Boltzmann speed distribution function $f(v)$, for hydrogen atoms at 1000 K. The escape speed at Earth's exobase and the root mean square (rms) speed of the atoms are indicated. Only those atoms in the energetic tail of the speed distribution with speeds exceeding the escape speed (shaded) are able to undergo *Jeans' thermal escape*. In contrast, atoms with the most probable speed at the peak of the distribution, or with the slightly higher rms speed, are unable to escape.

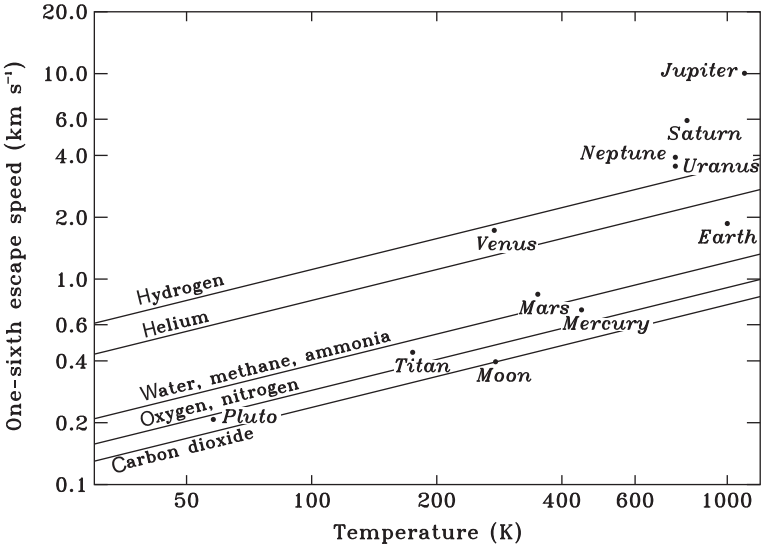

Figure 5.4 A schematic graph showing gas species that are prone to undergo thermal escape from planets in the Solar System on the basis of the "rule of thumb" approach discussed in the text. For bodies with substantial atmospheres, the temperature corresponds to the average exobase temperature. For the Moon and Mercury, the temperature is the mean surface temperature. The sloping lines correspond to the root mean square speed of the various gas molecules at the given temperature.

stability of gas species against Jeans' escape for planets in the Solar System using exobase temperatures. This diagram is a rough summary, but its general inferences are valid: Nothing escapes from the giant planets, whereas the smaller, rocky planets lose light gases.

In viewing Fig. 5.4, one should remember that *suprathermal* mechanisms that are not based on thermal equilibrium distributions of molecules allow light gases to escape. For example, O, C, and N can escape slowly from Mars via photochemical or ionic reactions; Fig 5.4 does not include such effects, which we discuss in Sec. 12.2.4.

5.6 Jeans' Thermal Escape of Hydrogen

5.6.1 Concept and Mathematical Derivation

Our purpose in what follows is to derive a general expression for the Jeans' thermal escape flux of hydrogen and other gases. Molecules in a collisional gas have a Maxwell–Boltzmann speed distribution (Box 5.1) and we have to integrate over the "tail" of the distribution above the escape speed to obtain the escape flux.

The number of molecules with speeds between v and $v+dv$ is given by

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Box 5.1 The Maxwell–Boltzmann Velocity Distribution

Derivation of the Maxwell–Boltzmann distribution starts with the translational kinetic energy of each molecule, $E = ½mv^2$. Then the probability that a molecule has a particular speed is proportional to the Boltzmann factor, i.e.,

$$\text{probability of particular speed} \propto \exp(-E/kT) = \exp\left(-mv^2/2kT\right)$$

For the distribution of velocities, imagine a sphere representing velocity parameter space, centered on v_x, v_y, v_z axes. Within a shell of thickness Δv at radius v, the volume is proportional to the square of the radius times the thickness, i.e. $v^2\Delta v$. So the number of speed states between v and $v+\Delta v$ is $\propto v^2\Delta v$. Thus, the probability of finding a given molecule in the speed range v to $v+\Delta v$ follows the proportionality

$$\begin{pmatrix}\text{probability of finding a}\\ \text{given molecule in speed}\\ \text{range } v \text{ to } v+dv\end{pmatrix} \propto \begin{pmatrix}\text{no. of microstates}\\ \text{in the speed range}\end{pmatrix} \times \begin{pmatrix}\text{probability of finding}\\ \text{the molecule in a given}\\ \text{microstate in the speed range}\end{pmatrix}$$

$$f(v)dv \propto v^2 \exp\left(-\frac{mv^2}{2kT}\right)dv$$

A proportionality constant is then selected to satisfy the condition $\int_0^\infty f(v)dv = 1$, i.e. that all molecules must be in some state. This results in eq. (5.15).

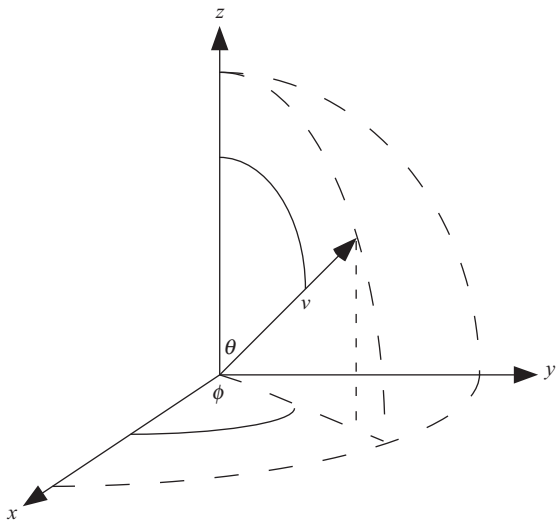

Figure 5.5 The geometry for Jeans' escape of a molecule with velocity v, in spherical polar coordinates, where θ is zenith angle and ϕ is azimuthal angle.

$$nf(v)dv = \frac{4n}{\sqrt{\pi}}\left(\frac{m}{2kT}\right)^{3/2} v^2 \exp\left(-\frac{mv^2}{2kT}\right)dv \tag{5.15}$$

Here, n represents the total number density of the constituent being considered.

Now we make the further assumption that the velocity distribution is isotropic, so that the same number of molecules travels in every direction. In spherical polar coordinates, with azimuth angle ϕ and polar angle θ, an element of solid angle $d\Omega$, is given by the following (see Fig. 5.5):

$$d\Omega = \sin\theta d\theta d\phi \tag{5.16}$$

Integrating around azimuth angle ϕ, gives $d\Omega = 2\pi \sin\theta d\theta$. Thus, the fraction of upward traveling molecules is $2\pi \sin\theta d\theta/4\pi$. In turn, the number of molecules with velocities between v and $v+dv$ travelling at an angle between θ and $\theta+d\theta$ from the vertical is given by

$$\frac{nf(v)dv\cancel{2\pi}\sin\theta d\theta}{\cancel{4\pi}} = \frac{1}{2}nf(v)dv\sin\theta d\theta \tag{5.17}$$

The vertical flux of molecules, Φ, is found by multiplying by the vertical component of velocity, $v\cos\theta$, and integrating over the upwards hemisphere from $\theta = 0$ to $\pi/2$:

$$\text{vertical flux with speed } v = \Phi dv = \int_0^{\pi/2} \frac{1}{2}nf(v)vdv\cos\theta\sin\theta d\theta \tag{5.18}$$

Noting that $d(\sin^2\theta)/d\theta = 2\sin\theta\cos\theta$, we can evaluate the integral in eq. (5.18) using

$$\int_0^{\pi/2} \cos\theta\sin\theta d\theta = \frac{1}{2}\left[\sin^2\theta\right]_0^{\pi/2} = \frac{1}{2} \tag{5.19}$$

Thus, the vertical flux of molecules with velocity v is given by

$$\Phi dv = \frac{1}{4}nf(v)vdv \tag{5.20}$$

The total escape flux (with SI units of particles m^{-2} s^{-1}) is found by setting $n = n_{exob}$, the number density at the

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

exobase, substituting for $nf(v)\mathrm{d}v$ from eq. (5.15), and integrating over all velocities exceeding the escape velocity, v_e

$$\Phi_{esc} = \int_{v_e}^{\infty} \Phi dv = \frac{n_{exob}}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} \int_{v_e}^{\infty} v^3 \exp\left(-\frac{mv^2}{2kT}\right) dv \tag{5.21}$$

To evaluate the integral, we substitute $x = v^2$, $dx = 2vdv$, and integrate by parts, giving,

$$\text{escape flux} = \Phi_{esc} = \frac{n_{exob}}{\sqrt{\pi}} \left(\frac{m}{2kT}\right)^{3/2} \left(\frac{kT}{m}\right) \left(v_e^2 + \frac{2kT}{m}\right) \exp\left(-\frac{mv_e^2}{2kT}\right) \tag{5.22}$$

One can simplify this expression by using eq. (5.13) to substitute for the square of the escape velocity at level r_{exob},

$$v_e^2 = \frac{2GM}{r_{exob}} \tag{5.23}$$

Also, let us define, v_s, which is the most probable speed in the Maxwell–Boltzmann speed distribution function

$$v_s \equiv \left(\frac{2kT}{m}\right)^{1/2} \tag{5.24}$$

We also introduce the (*Jeans*) *escape parameter*, λ_J, which is the ratio of gravitational potential energy GMm/r to thermal energy $\sim kT$ (Chamberlain, 1963),

$$\lambda_J \equiv \frac{GMm}{kTr}; \text{ at exobase, } \lambda_{Jexo} \equiv \frac{GMm}{kTr_{exob}} \equiv \frac{r_{exob}}{H_{exob}} \equiv \frac{{v_e}^2}{{v_s}^2} \tag{5.25}$$

Then, eq. (5.22) becomes

$$\text{escape flux} = \Phi_{esc} = \frac{n_{exob}}{2\sqrt{\pi}} \frac{1}{v_s} \left(v_e^2 + v_s^2\right) \exp\left(-\frac{v_e^2}{v_s^2}\right)$$

or

$$\Phi_{esc} = \frac{1}{2\sqrt{\pi}} n_{exob} v_s (1 + \lambda_{Jexo}) e^{-\lambda_{Jexo}} \tag{5.26}$$

This is a convenient expression by which one may evaluate the Jeans escape flux.

Equation (5.26) for the Jeans escape rate can differ from the actual rate of escape because of evaporative cooling of the exobase and distortions in the velocity distribution function. Several papers have suggested that the expression *overestimates* the number of high-velocity molecules at the exobase that escape because escape depletes the high-speed molecules and can cool the background gas. The overestimate is ~20%–30% when hydrogen escapes from a two- or multi-component gas model (Brinkmann, 1971; Pierrard, 2003; Shizgal and Blackmore, 1986) for Earth or Titan (Tucker *et al.*, 2013). Figure 5.6 shows a variety of estimates of the correction factor for the Earth in the range ~0.6-0.8. Fortunately, we can often ignore such discrepancies in atmospheric evolution studies given that other uncertainties dominate, such as the bulk composition of ancient atmospheres.

Volkov *et al.* (2011a; 2011b), who considered the angular distribution of the velocity distribution at the exobase for a single-component gas, found that Jeans escape can *underestimate* the actual thermal escape in that case. The ratio of actual escape to the Jeans escape rate at the exobase in their model was ~1.4–1.7 for an atmosphere with Jeans' parameter ranging from 6 to 15. This arose because some molecules gained energy to escape from

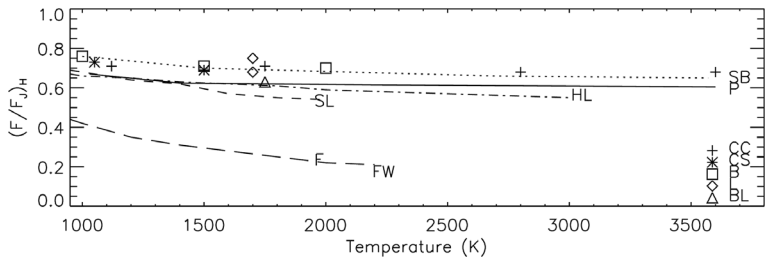

Figure 5.6 The ratio of the non-Maxwellian escape flux of hydrogen, F, to the Maxwellian Jeans' escape flux, F_J, versus the temperature at the Earth's exobase (from Pierrard, 2003). The solid line (P) shows the results of Pierrard (2003). Results are compared with those obtained with Monte Carlo simulations represented by different symbols: CC (Chamberlain and Campbell, 1967); L (Lew, 1967); B (Brinkmann, 1970); CS (Chamberlain and Smith, 1971); BL (Barakat and Lemaire, 1990). The other lines correspond to analytic solutions: HL (Hays and Liu, 1965); F (Fahr, 1976); FW (Fahr and Weidner, 1977); SL (Shizgal and Lindenfeld, 1980), SB (Shizgal and Blackmore, 1986).

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

collisions above the exobase. See Sec. 5.10.1 for further discussion of such enhanced Jeans-like escape.

The value of the escape parameter, λ_{Jexo}, is important in eq. (5.26). When the gravitational binding energy (GMm/r_{exob}) is much larger than kT, λ_{Jexo} is large, causing a small exponential factor $e^{-\lambda_{Jexo}}$ in eq. (5.26) and a small escape rate. In contrast, if the gravitational binding energy is much smaller than kT and λ_{Jexo} is small, the exponential factor in eq. (5.26) approaches unity, gas expands into the vacuum of space, and an atmosphere is unstable such as on the Moon.

One criticism of the above derivation is that the definition of the exobase is somewhat arbitrary. Why should it be the height where $\sigma N(r) = 1$? After all, only H atoms whose velocities are directed precisely upwards will have a $1/e$ chance of escaping. Those headed off at some other angle will have a somewhat lower probability of doing so. Fortunately, it can be shown that the escape flux is only weakly sensitive to the exact value of r_{exob} (c.f., Walker (1977)). Suppose, for example, that we picked some other altitude $r_{exob}{}'$ as the location of the exobase. Let the number density at that altitude be $n_{exob}{}'$. Then, using the analog to eq. (5.4), but picking the reference point to be r_{exob} rather than r_{surf}, and taking T to be constant, allows us to write

$$n'_{exob} = n_{exob} \exp\left[\frac{GMm}{kT}\left(\frac{1}{r'_{exob}} - \frac{1}{r_{exob}}\right)\right] \tag{5.27}$$

Using the definition of λ_{Jexo} above allows us to rewrite this as

$$n'_{exob} e^{-\lambda'_{Jexo}} = n_{exob} e^{-\lambda_{Jexo}} \tag{5.28}$$

The Jeans escape flux evaluated at distance $r_{exob}{}'$ would be related to the flux at distance r_{exob} by

$$\frac{\Phi'}{\Phi} = \frac{n'_{exob}(1+\lambda'_{Jexo})e^{-\lambda'_{Jexo}}}{n_{exob}(1+\lambda_{Jexo})e^{-\lambda_{Jexo}}} = \frac{1+\lambda'_{Jexo}}{1+\lambda_{Jexo}} \approx 1 \tag{5.29}$$

Because λ_J is only a slowly varying function of r, the Jeans escape flux is relatively insensitive to the exact altitude at which it is evaluated.

5.6.2 Effusion Velocity

To provide some physical insight into the implications of eq. (5.26), let us calculate the effective *effusion velocity* (m s^{-1}) for hydrogen escaping from Earth's atmosphere, which is the average rate at which hydrogen atoms or molecules are drifting upwards. Using eq. (5.26), we can write

$$\text{effusion velocity}, v_J = \frac{\Phi_J}{n_{exob}} = \frac{1}{2\sqrt{\pi}} v_s (1+\lambda_{Jexo}) e^{-\lambda_{Jexo}} \tag{5.30}$$

In Earth's upper atmosphere, the temperature varies from ~1000 K at solar minimum to 1500 K or more at solar maximum. As a result of these hot temperatures, molecular hydrogen is broken down efficiently into atomic hydrogen by the reaction

$$H_2 + O \rightarrow H + OH \tag{5.31}$$

Thus, the dominant form of hydrogen at the exobase on Earth is H, rather than H_2. This is not true on Mars or Titan, where the upper atmosphere is much colder. There, H_2 and H are both important hydrogen-bearing constituents at high altitudes.

If we calculate the numbers for atomic H at Earth's exobase, we find $\lambda_c = 7$ and $v_J = 0.87$ m s^{-1} for $T_\infty = 1000$ K, and $\lambda_{Jexo} = 4.7$ and $v_J = 110$ m s^{-1} for $T_\infty = 1500$ K. This tells us two things. First, it shows why elements heavier than He do not escape from Earth's atmosphere. The next lightest gas-forming element, carbon (C), has a mass number of 12 and, thus, $\lambda_{Jexo} > 50$. For example, at 1500 K, λ_{Jexo} for C is 12 times the value for atomic hydrogen, giving $12 \times 4.7 = 56.4$. Because λ_{Jexo} appears as a negative exponential in eq. (5.26), this effectively precludes thermal loss of C or any heavier element from Earth's atmosphere. Second, it shows that effusion velocities for atomic H vary widely from solar minimum to solar maximum. As we will see in the next section, Jeans' escape is the dominant escape mechanism for hydrogen at solar maximum, but is outweighed by suprathermal hydrogen loss processes at solar minimum. In fact, overall suprathermal (or nonthermal) processes dominate the time-averaged loss rate.

5.7 Suprathermal (Nonthermal) Escape of Hydrogen

Thermal loss is only one of several possible mechanisms by which gases can escape from atmospheres. Various *suprathermal loss processes* (also called *nonthermal*) dominate Earth's current hydrogen escape. Suprathermal molecules or atoms are particles whose velocities exceed the expected values from the Maxwellian distribution because they acquire kinetic energy in ways other than purely thermal collisions. What suprathermal processes have in common is that a boost from a chemical reaction or electrical or magnetic acceleration imparts escape velocity to single particles.

The two most important suprathermal hydrogen loss mechanisms for Earth are as follows.

(a) H–H$^+$ charge exchange. In this process, neutral H atoms in the upper atmosphere exchange charge with fast-moving ("hot") H$^+$ ions in Earth's plasmasphere

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

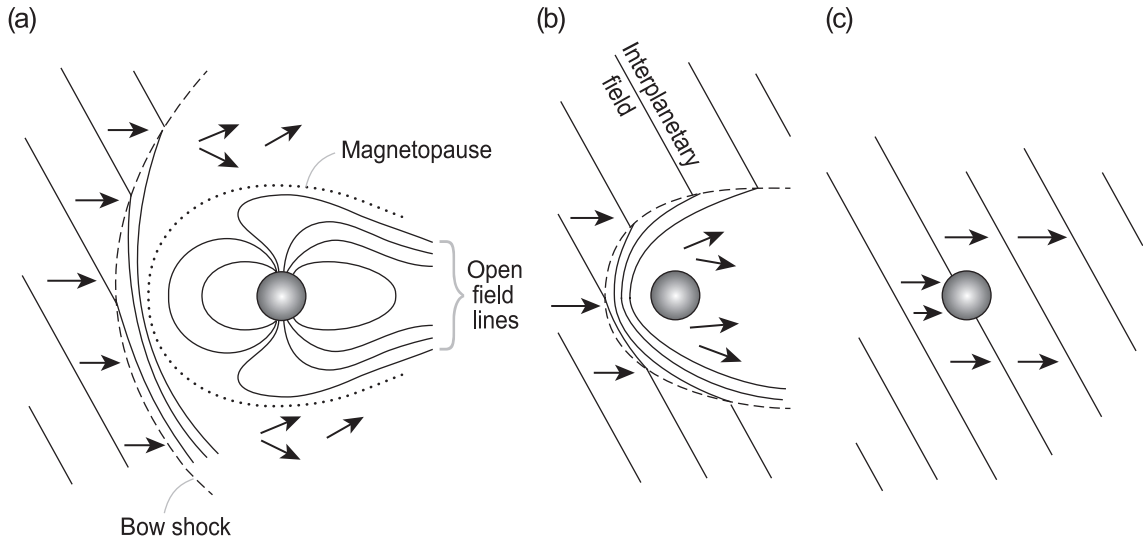

Figure 5.7 (a) Interaction of the solar wind with a planet that has a significant magnetic field, e.g. Earth (and similar also for Jupiter, Mercury, Saturn, and Uranus). Inside the magnetopause, the magnetic field is dominated by that of the planet and solar wind particles that are thermalized at the bow shock (where the wind is brought to rest) flow around the magnetopause and have little influence on the atmospheric evolution. Open field lines, however, allow ions to escape from the poles if they exceed the escape velocity. (b) Interaction of the solar wind with a planet without a significant magnetic field, e.g. present-day Mars or Venus. A bow shock is formed much closer to the planet due to an induced field in the ionosphere. Thermalized solar wind particles can interact directly with the atmosphere. (c) The solar wind collides directly with a body that has low electrical conductivity, no ionosphere/atmosphere, and no magnetic field, e.g. the Moon. (Adapted from Fig. 3.8 in Lewis and Prinn (1984).)

$$H + H^+(\text{hot}) \rightarrow H^+ + H(\text{hot}) \tag{5.32}$$

The *plasmasphere* is the region, including the ionosphere (60–3000 km) and magnetosphere, where large numbers of ionized species are present. The ion temperature in Earth's plasmasphere is ~ 5000–20 000 K. It is much hotter than the neutral temperature because of heating by Coulomb collisions, i.e. acceleration due to ion charge attraction and repulsion. When these hot ions exchange charge and become neutral hydrogen atoms, they retain their original high velocities, which can be approximated by a Maxwellian distribution at a much higher temperature. Hence, the fraction of these hydrogen atoms that exceed escape velocity is far higher than for the neutral background population. According to model calculations by Yung *et al.* (1989), about 75% of Earth's hydrogen escapes suprathermally when averaged over time. Charge exchange reactions account for ~50% of the escape. At solar maximum, however, the majority of the hydrogen follows the Jeans escape path.

(b) The polar wind. Near the magnetic poles of the Earth, the magnetic field lines are open over about 1/40 of the Earth's surface, that is, they do not connect to field lines that re-enter the Earth's surface (see Fig. 5.7(a)). Thus, hydrogen ions that are accelerated upwards along these field lines can escape. A well-known mechanism exists to produce such upward acceleration. When solar extreme ultraviolet radiation and x-rays ionize neutral atoms and molecules in the upper atmosphere, the lighter (and more mobile) electrons congregate higher up than the heavier, less mobile ions. This produces a *charge separation electric field* that has the effect of accelerating positively charged ions in the upward direction. The dominant ion in the F region (150–800 km) of the ionosphere, where most of atomic ion production occurs, is O^+. These oxygen ions are too heavy to escape. Rather, they remain fixed and sustain the electric field, while hydrogen atoms that become ionized in Earth's polar regions are accelerated out to space. According to Yung *et al.* (1989), about 15% of the hydrogen that escapes from Earth does so by this process. Details of the polar wind process are reviewed by Schunk (1988) and Ganguli (1996).

For present-day Earth, the loss mechanisms for hydrogen in order of contribution are 60%–90% for charge

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

exchange, 10%–40% for Jeans' escape, and ~10%–15% for the polar wind (Liu and Donahue, 1974; Maher and Tinsley, 1977; Yung *et al.*, 1989). In fact, hydrogen escape occurs efficiently by a variety of mechanisms once hydrogen reaches the exobase. As we shall see below, upward diffusion through the static background thermosphere below the exobase limits the rate of hydrogen escape on Earth today.

It is worth noting that solar wind particles have a more direct interaction with the atmospheres on planets that do not have global dipole magnetic fields, such as Mars or Venus (Fig. 5.7(b)). In such cases, the interaction is primarily with an ionosphere rather than a magnetosphere. For small planets, such as Mars, this can lead to substantial atmospheric loss over time (Sec. 12.2.4). Of course, on bodies without atmospheres and little electrical conductivity, e.g., the Moon, the solar wind impinges directly on the surface (Fig. 5.7(c)).

5.8 Upwards Diffusion and the "Diffusion-Limited Escape" Concept

Under a variety of circumstances (covering bodies as diverse as Venus, Earth, Mars, and Titan), the flux of hydrogen escaping to space is not constrained by the rate of removal of hydrogen at the exobase but is limited by the slower, upwards supply of hydrogen through the atmosphere below the exobase. This so-called *limiting flux* will be shown to be linearly proportional to the hydrogen mixing ratio at the homopause. To derive this result, we must consider upward diffusion of hydrogen through an atmosphere. Such diffusion can take place either by molecular diffusion or by eddy diffusion, which we consider in turn.

5.8.1 Molecular Diffusion

To understand how hydrogen diffuses upward through an atmosphere, we start by considering a binary mixture of two gases, denoted 1 and 2. The relative diffusion velocity of gas 1 with respect to gas 2 in its most general 3-D form is given by Banks and Kockarts (1973 Part B, p. 33 ff.) as:

$$\begin{aligned}\vec{v} &= \vec{v}_1 - \vec{v}_2 \\ &= -D_{12}\left[\underbrace{\frac{n^2}{n_1 n_2}\nabla\left(\frac{n_1}{n}\right)}_{\text{conc.gradient term}} + \underbrace{\frac{m_2 - m_1}{m}\nabla(\ln p)}_{\text{mass gradient term}} + \underbrace{\alpha_T \nabla(\ln T)}_{\text{temp.gradient term}} - \underbrace{\frac{m_1 m_2}{mkT}\left(\vec{F}_1 - \vec{F}_2\right)}_{\text{force gradient term}}\right]\end{aligned} \tag{5.33}$$

Terms on the right-hand side of this equation account for diffusion because of gradients in concentration, mass and temperature, respectively, as indicated. The last right-hand side term is diffusion caused by differential forces. Here,

n_1, n_2 = number densities of gases 1 and 2 (in m^{-3} or cm^{-3})
$n = n_1 + n_2$
m_1, m_2 = molecular masses of gases 1 and 2
$m = (n_1 m_1 + n_2 m_2)/(n_1 + n_2)$
k = Boltzmann's constant (1.38×10^{-23} J K^{-1})
T = temperature (K)
$\vec{F}_1, \vec{F}_2$ = accelerations acting on particles of gas 1 and 2 from external forces, e.g. gravity or an electric field
D_{12} = binary diffusion coefficient
α_T = thermal diffusivity ($\cong -0.25$ for H or H_2 in terrestrial air, ~0 for gases of comparable molecular mass)
∇ = the gradient operator, or $\vec{i}\frac{\partial}{\partial x} + \vec{j}\frac{\partial}{\partial y} + \vec{k}\frac{\partial}{\partial z}$ in Cartesian coordinates.

The binary diffusion coefficient, D_{12}, can be written as

$$D_{12} = \frac{b_{12}}{n} \tag{5.34}$$

where b_{12} is a binary diffusion parameter that is found empirically to vary as $b = AT^s$ where A and s are constants for particular binary gases mixtures. Diffusion parameters for H and H_2 in terrestrial air at 300 K are given in the Table 5.2.

The vector equation (5.33) describes a 3-D distribution of particles. To apply this to a plane-parallel planetary atmosphere, we simplify to just the vertical component. For neutral gases in a gravitational field, the external force term is zero because the gravitational acceleration, g, is independent of a particle's mass, i.e.,

Table 5.2 Binary diffusion parameters for H and H_2 in terrestrial air. (Data from Banks and Kockarts (1973 Part B, pp.40–41).)

Gas 1	Gas 2	A	s	b_{12} ($cm^{-1}s^{-1}$)
H_2	Air	2.7×10^{17}	0.75	1.95×10^{19}
H	Air	4.8×10^{17}	0.70	2.60×10^{19}

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

$\vec{F}_1 = \vec{F}_2 = g$. This would not be true for an ionized gas in the presence of an electric field where the opposite forces acting on positively charged ions and on electrons give rise to a phenomenon referred to as *ambipolar diffusion*.

With the above simplifications, let the vertical component of velocity be denoted by w. Then, eq. (5.33) becomes

$$w_1 - w_2 = -D_{12}\left[\frac{n^2}{n_1 n_2}\frac{d(n_1/n)}{dz} + \frac{m_2 - m_1}{m}\frac{1}{p}\frac{dp}{dz} + \frac{\alpha_T}{T}\frac{dT}{dz}\right] \tag{5.35}$$

The differential in the first term in the square brackets can be expanded as follows,

$$\frac{d}{dz}\left(\frac{n_1}{n}\right) = \frac{1}{n}\frac{dn_1}{dz} - \frac{n_1}{n^2}\frac{dn}{dz}$$

so that the first term in the brackets becomes

$$\frac{n^2}{n_1 n_2}\frac{d(n_1/n)}{dz} = \frac{n}{n_1 n_2}\frac{dn_1}{dz} - \frac{1}{n_2}\frac{dn_2}{dz} \tag{5.36}$$

We apply these relations to a light gas moving through a stationary, heavier background gas, so species 1 = H or H_2, and species 2 = air. Then, by assumption, $w_2 = 0$ for the static air. Let us further assume that species 1, which we will henceforth denote by the subscript 'i', is a minor constituent. That is,

$$\begin{aligned} &n_1 \equiv n_i << n_2 \\ &n \equiv n_1 + n_2 \cong n_2 \\ &m \cong m_2 \end{aligned} \tag{5.37}$$

Then, by using eq. (5.36) and relationships (5.37), we can rewrite eq. (5.35) as

$$w_i = -D_i\left[\frac{1}{n_i}\frac{dn_i}{dz} - \frac{1}{n}\frac{dn}{dz} + \left(1 - \frac{m_i}{m}\right)\frac{1}{p}\frac{dp}{dz} + \frac{\alpha_T}{T}\frac{dT}{dz}\right] \tag{5.38}$$

The ideal gas law, $p = nkT$ gives $n = p/kT$, so $\ln n = \ln p - \ln T - \ln k$. Taking d/dz of this last expression, gives

$$\frac{1}{n}\frac{dn}{dz} = \frac{1}{p}\frac{dp}{dz} - \frac{1}{T}\frac{dT}{dz} \tag{5.39}$$

Substituting $(1/n)(dn/dz)$ from eq. (5.39) into eq. (5.38) yields

$$w_i = -D_i\left[\frac{1}{n_i}\frac{dn_i}{dz} - \frac{m_i}{m}\frac{1}{p}\frac{dp}{dz} + \left(\frac{1+\alpha_T}{T}\right)\frac{dT}{dz}\right] \tag{5.40}$$

where we have used the fact that two of the terms containing dp/dz cancel. Finally, we use the barometric law (eq. (1.10)) to write

$$\frac{1}{p}\frac{dp}{dz} = -\frac{1}{H_a} = -\frac{mg}{kT} \tag{5.41}$$

Consequently, the second term in the square brackets of eq. (5.40) can be written,

$$\frac{m_i}{m}\frac{1}{p}\frac{dp}{dz} = \frac{m_i g}{kT} \equiv \frac{1}{H_i}$$

where H_i represents the scale height of species "i". Then, eq. (5.40) becomes

$$w_i = -D_i\left[\frac{1}{n_i}\frac{dn_i}{dz} + \frac{1}{H_i} + \left(\frac{1+\alpha_T}{T}\right)\frac{dT}{dz}\right] \tag{5.42}$$

The *flux* associated with molecular diffusion (also called "Fickian diffusion"), which we will denote by using a superscript "*mol*", is then

$$\Phi_i^{mol} = n_i w_i = -D_i n_i\left[\frac{1}{n_i}\frac{dn_i}{dz} + \frac{1}{H_i} + \left(\frac{1+\alpha_T}{T}\right)\frac{dT}{dz}\right] \tag{5.43}$$

Here the first term ($D_i\, dn_i/dz$) on the right-hand side is the familiar form of *Fick's First Law*, which expresses how the flux of molecules of species i across unit area in unit time is proportional to the concentration gradient of that species, dn_i/dz.

5.8.2 Eddy Diffusion

Thus far, we have described the individual motions of gas molecules of trace species "i" with respect to molecules of a static background atmosphere. In the lower atmosphere, though, most of the mass transport occurs not by diffusion of individual molecules but, rather, by turbulent, macroscopic eddies or by advection of air parcels, again of macroscopic scale. For convenience, aeronomers lump all such transport into the single process of *eddy diffusion*. Its magnitude is parameterized by an *eddy diffusion coefficient, K*. Eddy diffusion, by its very nature, acts so as to reduce gradients in relative species concentrations. Thus, if we define the *volume mixing ratio* of species i as $f_i \equiv n_i/n$, then we can write the flux due to eddy diffusion as

$$\Phi_i^{eddy} = -Kn\frac{df_i}{dz} \tag{5.44}$$

The total flux due to both molecular and eddy diffusion is their sum,

$$\Phi_i = \Phi_i^{mol} + \Phi_i^{eddy} \tag{5.45}$$

The value of the eddy diffusion coefficient, K, is not precisely defined, unlike the molecular diffusion coefficient,

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

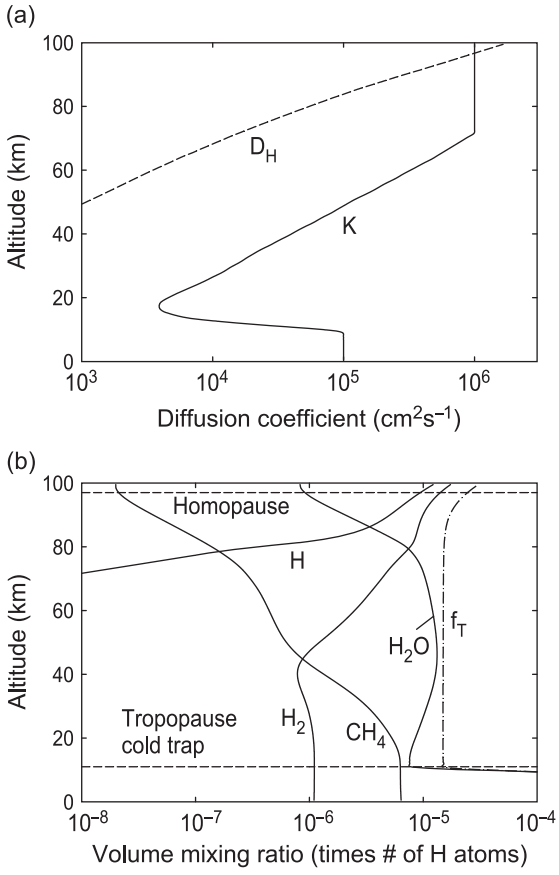

Figure 5.8 (a) Vertical profiles of the eddy diffusion coefficient, K, and the molecular diffusion coefficient for atomic hydrogen, D_H. (b) Vertical mixing ratio profiles of major hydrogen-bearing species, weighted by the number of H atoms in each species, e.g., methane at ground level is 1.8 ppmv × 4 = 7.2 ppmv H. The curve labeled f_T represents the total hydrogen mixing ratio defined in the text. The value of f_T begins to increase near the homopause because of diffusive separation of species that favors H and H_2.

D_i. No theory yet gives us the exact magnitude of K; rather, K is determined empirically using tracer studies (Hunten, 1975; Massie and Hunten, 1981; Gutowsky, 1976; National Research Council, 1979). A typical example is the eddy diffusion profile shown in Fig. 5.8(a).

Theory provides some guidance on how K should vary in certain altitude regimes. For example, in the upper stratosphere and mesosphere, most of the mass transport is caused by turbulent eddies generated by the breaking of upwards-propagating gravity waves (Garcia and Solomon, 1985; Leovy, 1964). The amount of energy dissipated by such waves is predicted to vary as the inverse square root of density (see Sec. 4.3.4 and 4.4.1), and so in this region it is assumed that $K \propto n^{-1/2}$. Because the spatial scale of the eddies is relatively small, parameterizing mass transport as "diffusion" is a good approximation in this region. By contrast, in the lower stratosphere and troposphere, much of the mass transport occurs as a result of large-scale advection, so the eddy diffusion approximation is less well justified.

At some altitude (~100 km on Earth), D_i becomes greater than K. This altitude is the *homopause* (as defined in Sec. 1.1.1). Although K increases as $n^{-1/2}$ in the mesosphere, the molecular diffusion coefficient, D_i, is proportional to $1/n$ (eq. (5.35)) and, hence, increases more rapidly with height than K. The region below the homopause, where eddy diffusion dominates, is the *homosphere* where air is mixed, or homogenized, by the processes that we have just discussed. Species that lack strong chemical sources or sinks, e.g., N_2, have constant mixing ratios in the homosphere. This can be easily demonstrated by setting $\Phi_i^{eddy} = 0$ in eq. (5.44). Then, $df_i/dz = 0$ as well, which implies that f_i is constant with altitude.

The region above the homopause, where molecular diffusion dominates, is the *heterosphere*. Here, mixing ratios of lighter species increase with altitude. This can be easily demonstrated by considering a non-reactive species such as N_2 for which the vertical flux is essentially zero. Then, eq. (5.43) says that

$$\frac{1}{n_i}\frac{dn_i}{dz} = -\frac{1}{H_i} - \left(\frac{1+\alpha_T}{T}\right)\frac{dT}{dz} \tag{5.46}$$

If we neglect thermal diffusion (i.e., set $\alpha_T = 0$), this implies that the partial pressure, $p_i = n_i kT$, of each individual species varies with its own scale height, i.e.

$$p_i = p_0 \exp\left(-\frac{z - z_0}{H_i}\right) \tag{5.47}$$

where p_0 is the partial pressure at some reference height z_0. Partial pressure $p_i(z)$ is the analog of the barometric law (eq. (1.12)) for the total atmospheric pressure, $p(z)$.

5.8.3 Diffusion-Limited Escape of Hydrogen

We now derive a very useful expression for the maximum upward flux of an escaping gas, such as hydrogen. As shown below, the escape of hydrogen from Earth's atmosphere is limited by the rate at which it can diffuse upwards through the homopause. Physically, one may think of the static background atmosphere, N_2 and O_2 in Earth's case, as providing a frictional resistance that retards the upward flow of hydrogen. The significance of this phenomenon was first pointed out by Donald Hunten (1973) in a paper that was motivated by his work on the escape of hydrogen

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

from Saturn's moon, Titan. It was also well described by Walker (1977) and provided the basis for his pioneering work on predicting Earth's prebiotic O_2 concentrations (Walker, 1978). The limiting flux concept has proven extremely valuable in understanding the behavior of atmospheres on rocky planets because, to the degree that data are available, diffusion-limited escape explains the H escape rate from Earth, Venus, Mars, and Titan (Sec. 5.9).

We begin by recasting the expression for the molecular diffusion flux (eq. (5.43)) in a form similar to that of the eddy diffusion flux (eq. (5.44)). First we note that

$$\frac{df_i}{dz} = \frac{d}{dz}\left(\frac{n_i}{n}\right) = \frac{1}{n}\frac{dn_i}{dz} - \frac{n_i}{n^2}\frac{dn}{dz}$$

Hence,

$$n\frac{df_i}{dz} = \frac{dn_i}{dz} - \frac{n_i}{n}\frac{dn}{dz} = n_i\left(\frac{1}{n_i}\frac{dn_i}{dz} + \frac{1}{H_a} + \frac{1}{T}\frac{dT}{dz}\right) \qquad (5.48)$$

where we have used eq. (5.39) and hydrostatic eq. (5.41) to eliminate dn/dz. By comparing eq. (5.48) with the molecular diffusion flux (eq. (5.43)), we can write

$$\Phi_i^{mol} = -D_i n\frac{df_i}{dz} + D_i n_i\left[\frac{1}{H_a} - \frac{1}{H_i} - \frac{\alpha_T}{T}\frac{dT}{dz}\right] \qquad (5.49)$$

If we now combine eq. (5.49) with eddy diffusion (eq. (5.44)) and total flux (eq. (5.45)) equations, we can write the total flux of species "i" as

$$\Phi_i = \underbrace{-(K + D_i)n\frac{df_i}{dz}}_{\text{counter-gradient flux term}} + \underbrace{D_i n_i\left(\frac{1}{H_a} - \frac{1}{H_i} - \frac{\alpha_T}{T}\frac{dT}{dz}\right)}_{\text{limiting flux term}} \qquad (5.50)$$

The first term on the right-hand side of eq. (5.50) is called the *counter-gradient flux* and is denoted by Φ_c

$$\Phi_c = -(K + D_i)n\frac{df_i}{dz} \qquad (5.51)$$

The magnitude of this term is proportional to the gradient in species mixing ratio, df_i/dz. It turns out that this term is unable to sustain any net upwards diffusion of hydrogen. Physically, the reason is because f_i must decrease with altitude in order to drive upwards diffusion via this term. But if f_i decreases with height, then n_i will decrease even more rapidly. The upward flux is equal to $n_i w_i$, where w_i is a velocity. Rapidly decreasing n_i would require rapidly increasing w_i, which is not physically possible. (See Walker, 1977, for a mathematical derivation of this result.) Thus, a maximum must occur in Φ_i (eq. (5.50)) when $df_i/dz = 0$ and $\Phi_c = 0$.

The second term on the right-hand side of eq. (5.50) is the *limiting flux* or the *diffusion-limited flux* and is denoted by Φ_l, as follows:

$$\begin{aligned}\Phi_l &= D_i n_i\left(\frac{1}{H_a} - \frac{1}{H_i} - \frac{\alpha_T}{T}\frac{dT}{dz}\right) \\ &\approx D_i n_i\left(\frac{(m_a - \cancel{m_i})g}{kT} - \frac{\alpha_T}{\cancel{T}}\frac{\cancel{dT}}{dz}\right)\end{aligned} \qquad (5.52)$$

As may be ascertained from the form of this equation, this term is entirely due to the difference in molecular weight between the escaping gas (presumed to be hydrogen) and the background atmosphere and to the thermal diffusivity of hydrogen. This expression is typically applied either at the homopause or in the lower stratosphere.

We can simplify the limiting flux equation as indicated by the strike-throughs in eq. (5.52). Temperature gradients are small, so $dT/dz \sim 0$ and the thermal diffusion term is generally neglected. Furthermore, for a light gas (H or H_2) diffusing through air (an N_2–O_2 mixture for Earth, or CO_2 for Mars and Venus), $H_i >> H_a$, so eq. (5.52) simplifies to

$$\boxed{\text{diffusion limited flux, } \Phi_l \cong \frac{D_i n_i}{H_a} = \frac{b_i f_i}{H_a} \propto f_i} \qquad (5.53)$$

where we have used eq. (5.34), which relates D_i to the binary diffusion parameter b_i.

It should be remembered that eq. (5.53) was derived for a minor constituent. It can, however, be applied to a major constituent as well (c.f. Walker, 1977) if one replaces the term f_i with $f_i/(1+f_i)$. This form should be used in cases where f_i exceeds a few percent.

5.8.4 Application of Diffusion-Limited Hydrogen Escape to Earth's Atmosphere

To apply the limiting flux equation to the Earth, we begin by evaluating some of the parameters. At the Earth's homopause, the temperature is ~208 K, $b_H \cong 2.73\times 10^{19}$ cm^{-1}s^{-1} and $b_{H2} \cong 1.46\times 10^{19}$ cm^{-1}s^{-1}. H_2 is several times more abundant that H at the homopause, 5.2×10^7 cm^{-3} vs. 1.8×10^7 cm^{-3} (Liu and Donahue, 1974). Rather than calculating the flux of each species separately, let us combine their mixing ratios and use a weighted average value for b_i of 1.8×10^{19} cm^{-1}s^{-1}. At the homopause, the scale height, H_a, is ~6.36 km, so $b_i/H_a \cong 1.8\times 10^{19}$ cm^{-1}s^{-1}/6.36×10^5 cm = 2.8×10^{13} cm^{-2} s^{-1}.

We now define the *total hydrogen mixing ratio*, f_T(H), as the sum of the mixing ratios of hydrogen in all of its chemical forms, weighted by the number of hydrogen atoms each species contains. Thus,

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

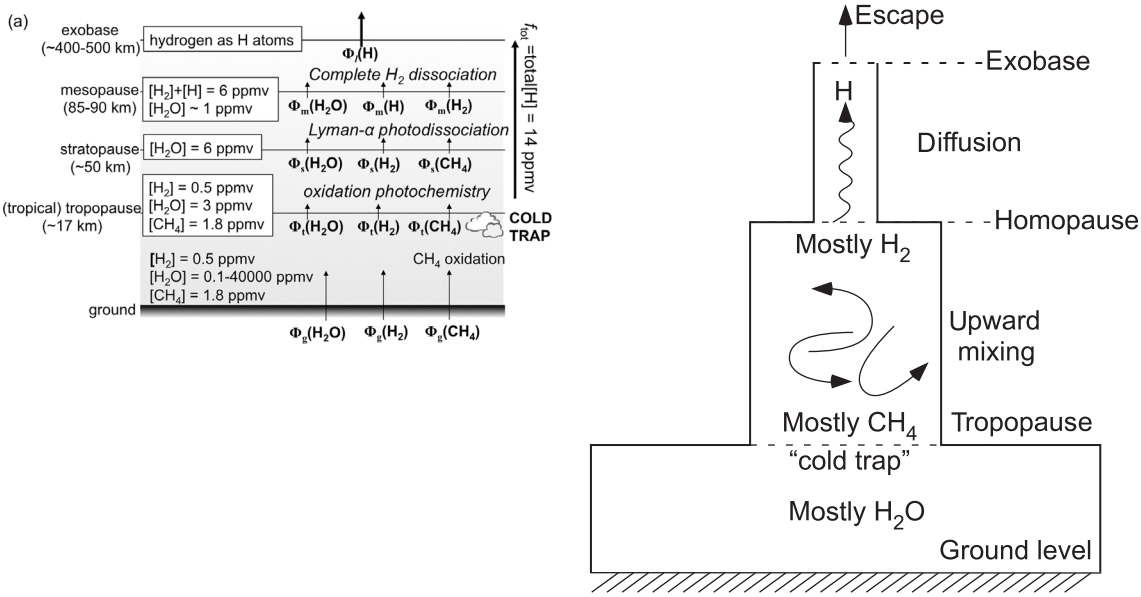

Figure 5.9 (a) A schematic diagram of hydrogen-bearing species in Earth's atmosphere showing processes responsible for the profiles shown in Fig. 5.8(b). Hydrogen is input from the ground in fluxes (Φ) of water, hydrogen and methane. At the tropical tropopause, a cold trap limits the flux of hydrogen from water vapor into the stratosphere. The total H mixing ratio f_T, just below the homopause sets the diffusion-limited escape flux of hydrogen to space (Φ_l), which is reflected by f_T in the lower stratosphere above the tropical tropopause cold trap. (b) A schematic diagram showing the two principal "bottlenecks" for hydrogen escape: the tropical tropopause cold trap and diffusion above the homopause.

$$f_T(\mathrm{H}) = f_\mathrm{H} + 2f_{\mathrm{H_2}} + 2f_{\mathrm{H_2O}} + 4f_{\mathrm{CH_4}} + \cdots \tag{5.54}$$

These terms, along with f_T(H) itself, are shown in Fig. 5.8 (b). At the Earth's homopause, almost all hydrogen exists as either H or H_2, so only the first two terms are important. We could evaluate f_T(H) there by using rocket measurements of the concentrations of these two species. However, there is an easier way to evaluate f_T(H) that is also much more useful in studying atmospheres in general, which we now examine.

To find a simple way to estimate the total hydrogen mixing ratio, f_T(H), we start with fundamentals by writing the *continuity equation* for species i as

$$\frac{d\Phi_i}{dz} = P_i - L_i \tag{5.55}$$

Here, P_i and L_i represent chemical production and loss rates per unit volume. For a species with no chemistry, $d\Phi_i/dz = 0$, so Φ_i is constant with altitude. But we know that $\Phi_i = \Phi_l$ at the homopause. Consequently, the total flux at all altitudes below and down to the lower stratosphere must be equal to the diffusion-limited flux. Equation (5.50) therefore implies that $df_i/dz = 0$, that is, the total hydrogen mixing ratio should be constant with altitude. This result is not exact because the species scale height, H_i, and thermal diffusivity, α_T, change as one hydrogen-containing compound is transformed into another. However, these changes can be compensated by gradients in f_i that are so small as to be negligible. Thus, despite all the complexity of atmospheric photochemistry, we arrive at the simple result:

The rate of escape of hydrogen is proportional to the concentration of hydrogen compounds in the lower stratosphere.

We can see how this works by looking at a schematic diagram showing the abundance of hydrogen in all its forms in the Earth's atmosphere as a function of height shown in Fig. 5.8(b) and schematically in Fig. 5.9(a). The total hydrogen mixing ratio f_T is essentially preserved above the tropopause, as verified by satellite measurements of hydrogen-bearing species (Harries *et al.*, 1996).

We can write the diffusion limited escape flux as a simple linear equation. Using eq. (5.53), we can represent the escape flux as

$$\Phi_l \cong (\text{constant}) f_T(\mathrm{H}) \tag{5.56}$$

We deduced a constant of 2.8×10^{13} cm^{-2} s^{-1} from our previous calculation. However, a more detailed calculation that takes account of other terms in eq. (5.52) suggests that the constant is 2.5×10^{13} cm^{-2} s^{-1} (Hunten and Strobel 1974), i.e.,

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

$$\Phi_l \cong \left(2.5 \times 10^{13}\right) f_T(\mathrm{H}) \left[\text{atoms cm}^{-2}\,\text{s}^{-1}\right] \tag{5.57}$$

Thus, provided we know f_T(H) we can easily calculate the diffusion-limited escape flux. Here, f_T(H) is simply the total hydrogen mixing ratio, in all its forms, above the cold trap at the tropopause, which can be readily measured from balloon-borne instruments in Earth's present atmosphere or computed theoretically for other atmospheres.

Diffusion-limited escape is an *upper limit*. The hydrogen cannot escape any faster than the rate of upwards diffusion through lower levels of the atmosphere by eddy and molecular means. Thus, conceptually, the Earth has two "bottlenecks" that limit the upward supply of hydrogen (Fig. 5.9(b)): the cold-trap of the tropical tropopause where water freezes out and the region above the homopause where diffusion is rate limiting.

In Earth's present atmosphere, the concentrations of hydrogen-bearing gases in the lower stratosphere are 1.8 ppmv CH_4, ~3 ppmv H_2O, and 0.55 ppmv H_2. From these measurements, we can calculate the diffusion-limited escape rate for hydrogen from the Earth. First we calculate f_T(H), which is 14×10^{-6} (= [2(0.55) + 4(1.8) + 2(3)]$\times10^{-6}$).

Then the escape rate from eq. (4.51) is

$$\Phi_l \cong \left(2.5 \times 10^{13}\right) f_T(\mathrm{H}) = \left(2.5 \times 10^{13}\,\text{cm}^{-2}\text{s}^{-1}\right)\left(14 \times 10^{-6}\right) = 3.5 \times 10^{8}\,\text{atoms}\,\text{cm}^{-2}\text{s}^{-1}.$$

Given that the area of the Earth is 5.1×10^{18} cm^2, the area-integrated escape rate is (3.5×10^8 atoms cm^{-2} s^{-1}) × (5.1×10^{18} cm^2) = 1.8×10^{27} atoms s^{-1} = 5.6×10^{34} atoms yr^{-1} = 9.3×10^{10} moles H atoms yr^{-1} = 93 000 tonnes of H yr^{-1}.

Measurements of exospheric temperatures and hydrogen densities prove that the limiting flux concept works well (Bertaux, 1975). Essentially, the total flux consists of the sum of suprathermal and Jeans' escape fluxes. The sum remains constant at about the limiting flux value at the homopause. As the Jeans flux varies due to changing exospheric temperature with the solar cycle, the suprathermal flux component changes to make up the balance.

The limiting flux concept does not work for terrestrial helium. Helium escape is rate-limited by the removal process at the top of the atmosphere to a flux ~100 times lower than diffusion-limited escape. Instead helium escapes efficiently as an ion along the open magnetic field lines at high latitudes (Axford, 1968; Johnson and Axford, 1969), with the result that the lifetime of helium in Earth's atmosphere is ~10^6 years.

5.9 Diffusion-Limited Hydrogen Escape Applied to Mars, Titan, and Venus

Now that we have discussed the theory for hydrogen escape and the limiting flux concept, it is interesting to apply it to other planetary atmospheres.

5.9.1 Mars

On Mars, simulated homopause temperatures vary as a function of season and range from ~150 K to ~200 K (e.g., Bougher *et al.*, 2000). Let us take T ~ 180 K, corresponding to the dayside homopause at ~135 km altitude. Here, the most abundant hydrogen species is H_2 (Nair *et al.*, 1994). The binary diffusion coefficient of H_2 in CO_2 is approximated a function of temperature (Marrero and Mason, 1972),

$$b(\mathrm{H_2\,in\,CO_2}) = \frac{3.1 \times 10^{-6}}{k} T^{0.75} e^{-11.7/T} \left[\text{cm}^{-1}\,\text{s}^{-1}\right] \tag{5.58}$$

A temperature of 180 K yields $b = 1.0\times10^{19}$ cm^{-1} s^{-1} (for H_2 in CO_2), and a scale height H_a = 9.3 km. Consequently, the limiting flux from eq. (5.53) is

$$\Phi_l \cong \left(1.1 \times 10^{13}\right) f_T(\mathrm{H}) \tag{5.59}$$

Spectroscopy suggests that the H_2 mixing ratio on Mars is ~15±5 ppmv in the lower atmosphere (Krasnopolsky and Feldman, 2001). Insertion of f_T(H)= (30±10)$\times10^{-6}$ in eq. (5.59) gives a diffusion-limited escape rate of (3.3±1.1)$\times$ 10^8 H atoms cm^{-2} s^{-1} .

Note that the low gravity on Mars leads, counter-intuitively, to a *smaller* diffusion-limited escape flux, according to eq. (5.53). This is because the atmospheric scale height, $H_a = kT/mg$, appears in the denominator of the expression, making $\Phi_\ell \propto g$. Essentially, the rate of diffusion of the light escaping gas relative to the heavier background gas is enhanced when the planet's gravitational attraction is stronger. This scaling will break down for very large planets when gravity-dependent removal processes at the exobase become inefficient and the limiting flux no longer applies to hydrogen.

Hydrogen escape fluxes cannot be observed directly but are inferred from the vertical profile of hydrogen. *Mariner 6* and *7* Lyman-α observations imply an escape flux of (1–2)$\times10^8$ H atoms cm^{-2} s^{-1} if a

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Maxwellian velocity distribution is assumed (Anderson, 1974; Anderson and Hord, 1971). *Mars Express* data have been interpreted to indicate that atomic hydrogen in Mars' exosphere has two populations: a suprathermal one (~1000 K) with a low number density and a colder one with a greater number density. The suprathermal population may arise from ionic reactions or charge exchange reactions with the solar wind (Galli *et al.*, 2006a; Galli *et al.*, 2006b). Such data imply a Jeans escape rate of $4.3^{+5.6}_{-2.5}\times10^{8}$ H atoms cm^{-2} s^{-1} (Zahnle *et al.*, 2008). Lyman-α and -β data from *Rosetta* were acquired during a gravity assist swing-by of Mars (while en route to comet 67P/Churyumov–Gerasimenko) and used to infer a somewhat lower escape rate (no error bar) of ~0.8×10^{8} H atoms cm^{-2} s^{-1} (Feldman *et al.*, 2011).

Overall, the limiting flux concept appears to work within uncertainty in explaining the current hydrogen escape rate for Mars. Since diffusion-limited escape depletes the typical abundance of water vapor in the atmosphere (~5×10^{19} cm^{-2} H atoms) in ~5 kyr, atmospheric water vapor must be replenished. One nuance is that *Mars Express* and *Hubble Space Telescope* data suggest that the H escape rate, although consistent with limiting flux, can decline when there is a decrease in the source of water vapor during large dust storms (Chaffin *et al.*, 2014; Clarke *et al.*, 2014).

5.9.2 Titan

On Titan, the homopause altitude is constrained by data from the *Cassini* spacecraft to be ~800–900 km (Strobel *et al.*, 2009) while the exobase is around 1400–1500 km altitude (see Sec. 12.4.3). Photochemical models suggest that most hydrogen escapes as H_2, with ~30% as H (Yung *et al.*, 1984).

The *Huygens* probe measured hydrogen-bearing species in Titan's atmosphere, with CH_4 ~1.48±0.09% in the stratosphere and H_2 ~1010±160 ppmv (Niemann *et al.*, 2010). At the homopause, the temperature is T~150 K, which gives a scale height $H_a = RT/Mg$ = [(8.314 J mol^{-1} K^{-1})(150 K)]/[0.0286 kg mol^{-1}) (0.733 m s^{-2})] = 59.5 km. The binary diffusion parameter for H_2 in N_2 at 150 K is $b = 1.1\times10^{19}$ cm^{-1} s^{-1}. Hence the limiting flux obtained from eq. (5.53) is

$$\Phi_l \cong \frac{b_i}{H_a} f_T = \frac{1.1\times10^{19}}{5.95\times10^{6}} f_T = 1.8\times10^{12} f_T \qquad (5.60)$$

Since methane is the dominant hydrogen-bearing species we must take into account the overall chemistry of methane photolysis and the fate of its products. Methane can be destroyed through different photolysis paths with differing yields of hydrogen:

$$2CH_4 \rightarrow C_2H_2 + 3H_2 \text{ (ionospheric)} \qquad (5.61)$$

$$2CH_4 \rightarrow C_2H_4 + 2H_2 \text{ (direct)} \qquad (5.62)$$

$$2CH_4 \rightarrow C_2H_6 + H_2 \text{ (catalytic)} \qquad (5.63)$$

Hydrogen atoms contained in the hydrocarbon products are lost as rainout to the surface whereas the H_2 is subject to escape. If we assume that every two CH_4 molecules produce an H_2 molecule (eq. (5.63)), then the total mixing ratio of H atoms will be f_T = (0.0148/2)2 + 2(0.001) = 0.0168, giving a limiting flux of 3×10^{10} atoms cm^{-2} s^{-1}. More detailed model calculations suggest an escape flux of 2.0×10^{10} atoms cm^{-2} s^{-1} (Table 1 in Lebonnois *et al.* (2003)), which is similar to the column-integrated destruction rate of CH_4 ~1.5×10^{10} cm^{-2} s^{-1} due to photolysis (Yung *et al.*, 1984).

The limiting flux for H escape from Titan is supported by observations. Utilizing *Cassini* ion-neutral mass spectrometer (INMS) data, Bell *et al.* (2010a; 2010b) find that hydrogen escapes at close to the Jeans rate, while others have suggested that the rate is sometimes enhanced with energy input by Saturn's magnetospheric particles (Cui *et al.*, 2011). The inferred escape rate is (2.0–2.1)$\times10^{10}$ H atoms cm^{-2} s^{-1} (Bell *et al.*, 2010a; Cui *et al.*, 2008), which is essentially the diffusion-limited rate evaluated at the homopause (Bell *et al.*, 2014; Strobel, 2012).

5.9.3 Venus

We can calculate a diffusion-limited H escape flux for Venus of ~3×10^{7} H atoms cm^{-2} s^{-1}, based on spectroscopic evidence for a total atomic hydrogen abundance of a few ppmv at the homopause. This total hydrogen mixing ratio comes from summing HCl ~0.5 ppmv (Bézard *et al.*, 1990) and H_2O ~ 1 ppmv (Fink *et al.*, 1972) above the cloud tops. This diffusion-limited escape flux agrees reasonably well with estimates of the globally averaged escape flux based on *Pioneer Venus* measurements of upper atmospheric composition and temperature, which in units of ~10^{7} atoms cm^{-2} s^{-1} are 1.7 (Hodges and Tinsley, 1981), 0.2 (McElroy *et al.*, 1982) and 2.7 (Kumar *et al.*, 1983).

On Venus, Jeans' escape of hydrogen is negligible because of the relatively high, Earth-like gravity and low exospheric temperature. The upper atmosphere of Venus is cold, with a ~300±25 K dayside exobase temperature, because of strong radiative cooling via decay of vibrationally excited CO_2 (see Sec. 3.5). At 275 K, the thermal

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

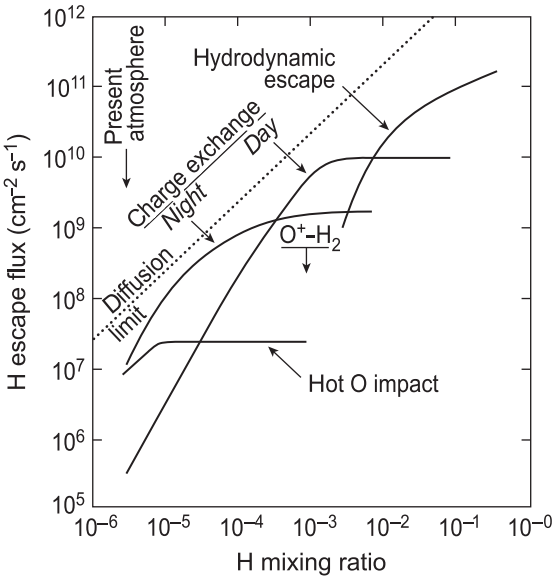

Figure 5.10 Calculated hydrogen escape fluxes from Venus. From left to right: increasing H concentrations to ~1% eventually permit hydrodynamic escape. At lower concentrations of H, charge exchange processes are the dominant source of hot H atoms. However, in all cases, the rate of escape is limited by diffusion through the atmosphere from lower layers (dashed line). (From Kumar *et al.* (1983).)

escape rate is only ~10^4 atoms cm^{-2} s^{-1}. On Venus, escape of suprathermal hydrogen completely dominates.

Mariner 5 and *10*, *Venera 9–12* and *Venus Express* experiments all found evidence for a suprathermal population of hydrogen in Venus' exosphere with a temperature of ~1000±500 K and exobase density of ~10^3 cm^{-3}, compared to a cold, background hydrogen density of ~10^5 cm^{-3} at the exobase and temperature 300±25 K (dayside) or 150 K (nightside) (Bertaux *et al.*, 1978; Bertaux *et al.*, 1982; Chaufray *et al.*, 2012; Kumar *et al.*, 1983). Suprathermal escape mechanisms that are considered to be important for Venus include (Hodges, 1999; Lammer *et al.*, 2006): (1) charge exchange between neutral hydrogen atoms and hot H^+ or O^+ ions, (2) charge exchange with hot oxygen atoms produced by CO_2 photolysis or from electron impact dissociation of CO_2, and (3) dissociative recombination of H-containing ions. Protons (H^+), produced by photoionization or solar wind interaction, can also escape, as measured at solar minimum at a rate of 3×10^6 cm^{-2} s^{-1} by an ion spectrometer on *Venus Express* (Barabash *et al.*, 2007).

Figure 5.10 is a summary plot of the various processes responsible for H escape from Venus, including on early Venus when hydrodynamic escape likely operated. The dashed line shows the diffusion-limited flux, which represents the maximum escape rate from the sum of these processes.

5.10 Hydrodynamic Escape

5.10.1 Conditions for Hydrodynamic Escape

In Sec. 5.3, we saw how the barometric law breaks down at high altitude and how a gravitationally bound atmosphere at a finite temperature must lose mass to the vacuum of space. We noted the two end-member cases of thermal escape processes associated with this breakdown, namely, Jeans' escape and hydrodynamic escape. Jeans' escape is the more accurate description when dealing with escape of a light gas, e.g., hydrogen, from a heavier and static background atmosphere. Hydrodynamic escape is a better approximation when the background atmosphere itself is escaping and driven by a pressure gradient force between the dense atmosphere below and vacuum above, or, equivalently, when the upper atmosphere is hydrogen-dominated and sufficiently heated. An important aspect of hydrodynamic hydrogen escape for atmospheric evolution is that heavy molecules can achieve escape velocity through collisions with hydrogen, and be dragged upward. Thus, rapid flow of hydrogen will carry away heavy gases into space, even though such gases would be too heavy to undergo Jeans' escape.

The classical Jeans' approximation treats the transition from collisional to collisionless as a discontinuity, but this concept begins to break down once the height of the exobase becomes defined by the escaping gas, typically hydrogen, rather than by some static background gas. As we noted earlier (Sec. 5.6), in a multi-component atmosphere, the classical Jeans' formula can overestimate the escape flux of hydrogen because it fails to account for depletion of molecules in the high-energy tail of the Maxwellian velocity distribution. Also, the Jeans approximation ignores the reduction in temperature caused by the loss of fast, escaping particles. This error of 20%–30% for Earth is not large because that tail is replenished through collisions with non-escaping molecules. In the transitional region from collisional to non-collisional gas, collisions between the atoms or molecules of the minor escaping gas are unimportant.

However, once the escaping gas becomes the dominant constituent in the thermosphere, the situation is more akin to a single-component atmosphere and the net result can be an anisotropic velocity distribution rather than an isotropic Maxwellian distribution. By integrating the nonlinear Boltzmann equation for an atmosphere of atomic hydrogen, Merryfield and Shizgal (1994) found that, at high altitudes, a population of escaping hydrogen atoms

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

was enhanced by streaming of particles from warmer, denser air below. There was also upward conduction of heat due to the escape of the particles from the tail of the distribution and some particles above the exobase gained escape velocity due to occasional collisions. Consequently, in such a single component model, Jean's escape was found to *underestimate* actual escape by a factor of ~1.3 for a Jeans' parameter of ~6.5 (eq. (5.25)). Using a direct simulation Monte Carlo model (DSMC), Volkov *et al.* (2011a; 2011b) similarly found that Jeans escape in a single-component atmosphere *underestimate*s actual escape at the exobase by a factor of ~1.4–1.7 for a Jeans' parameter ranging from 6 to 15.

Sophisticated models, such as DSMC, are required for precise estimates of strong "Jeans-like" escape to bridge the collisionless and hydrodynamic regimes when the Jeans parameter is small but escape is not strong enough for a purely fluid description of hydrodynamic escape to be justified. Apart from DSMC models, another method that has been used extensively to simulate the polar wind (Lemaire *et al.*, 2007; Tam *et al.*, 2007) is to make an approximation to the Boltzmann equation with a number of moments, e.g., Grad's 13-moment approximation (Cui *et al.*, 2008; Grad, 1949; Schunk, 1977; Ch. 3 of Schunk and Nagy, 2009).

The theory of hydrodynamic escape for a planetary atmosphere is best put in context by referring back to the old literature concerning the hydrodynamic nature of the solar wind. In the 1960s, Joseph Chamberlain and Eugene Parker debated the question of whether the solar wind was subsonic or transonic. Parker thought that it was transonic while Chamberlain thought that it was not. Ultimately, Parker was proved correct.

The debate happened because both subsonic and transonic solutions exist for the equations of hydrodynamic outflow of the solar wind. The wind, which is an expansion of the Sun's extremely hot corona, is fully ionized. The particles moving within it are charged and are subject to long-range electrostatic forces; hence, the wind is always in the collisional hydrodynamic regime. More precisely, the dynamics of the solar wind are described by magneto-hydrodynamics, which includes the influence of the magnetic field, but early treatments of the solar wind ignored this complication. In general, the solutions of the full equations of hydrodynamic outflow are difficult to obtain. If, however, one ignores the energy equation and assumes isothermal outflow, then the solution is analytic (Box 5.2, eq. (B5.11)). These isothermal outflow solutions are shown in Fig. 5.11.

The isothermal outflow solutions for the solar wind fall into six categories, but only three are physically plausible. Two solutions pass through the *critical point* or *sonic level*, r_0, which is the radius at which the flow velocity is equal to the isothermal sound speed, $u_0 = \sqrt{kT/m}$. Variables are defined in Box 5.2. Solution IV in Fig. 5.11, which starts from low velocity near the Sun and becomes supersonic at large distances, is the transonic escape solution. This solution was eventually shown by spacecraft data to be correct for the solar wind. Solution VI has low velocity at large distances and high velocity near the Sun. It represents infall of material, called *Bondi accretion* (Bondi, 1952; Shu, 1991 pp. 77–81). It could apply, for example, to capture of gas from a surrounding solar nebula, although the isothermal assumption would need to be abandoned in this case. This solution is thus more illustrative than practical.

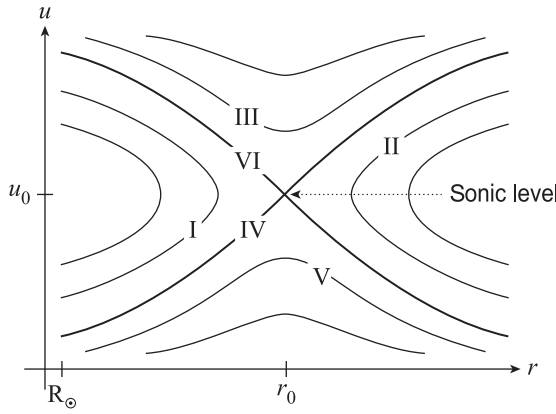

Figure 5.11 Solutions to the isothermal, time-independent, hydrodynamic escape equations (Box 5.2, eqs. (B5.11) or (B5.12)), but for an ionized gas. Six different classes of solution are shown (I to VI). Solution IV, which increases in speed and passes through the *sonic level* (also called the *critical point*) and becomes supersonic, is the solar wind solution. (Following Parker (1963).)

The other categories of solutions (I, II, III, V) fall into the four quadrants delineated on the graph by solutions IV and VI. Double-valued solutions in quadrants I and II do not connect the surface of the Sun ($r << r_0$) to locations far from the Sun ($r >> r_0$) and are physically meaningless. Quadrant III solutions have supersonic speeds at the Sun and are also unphysical. Apart from solutions IV and VI, only solutions in the bottom center quadrant (V) are potentially meaningful. These subsonic solutions have low velocity near the Sun, a peak velocity at distance r_0, and return to low velocities at large distances.

For an escaping fluid like the solar wind, which is collisional, the question of whether the flow is subsonic or transonic depends on the boundary conditions at large distances from the Sun. If the pressure of the interstellar medium is low, which it is, then the escape will be

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Box 5.2 Equations of Hydrodynamic Escape and Isothermal Approximations Pertaining to Them

The equations describing hydrodynamic escape, and fluid dynamics in general, tend to be rather opaque to those who do not deal with them on a regular basis. For many of us, there is much to be learned from a simplified description of a problem because one can gain physical insight from analytic approximations. Hydrodynamic escape is amenable to an analytic simplification with certain assumptions. In general, one needs to deal with three equations: conservation of mass, momentum, and energy. In spherical geometry, these equations can be expressed as follows (using equations of Chamberlain and Hunten (1987), p. 71–73, in spherical geometry).

Conservation of mass

$$\underbrace{\frac{\partial \rho}{\partial t}}_{\text{density change}} = \underbrace{-\frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \rho u\right)}_{\text{radial mass inflow or outflow (divergence)}} \tag{B5.1}$$

Conservation of momentum

$$\underbrace{\frac{\partial (\rho u)}{\partial t}}_{\text{momentum change}} = \underbrace{-\rho u \frac{\partial u}{\partial r}}_{\text{momentum inflow or outflow}} \underbrace{-\frac{\partial p}{\partial r}}_{\text{pressure gradient force}} \underbrace{-\rho \frac{GM}{r^2}}_{\text{gravitational force}} \tag{B5.2}$$

Conservation of energy

$$\rho c_{\mathrm{p}} \frac{\partial T}{\partial t} = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2 \kappa \frac{\partial T}{\partial r}\right) - \frac{kTu}{m}\frac{\partial \rho}{\partial r} - \rho c_{\mathrm{v}} u \frac{\partial T}{\partial r} + q \tag{B5.3}$$

Here, t is time, r is radial distance from the planet's center, ρ is mass density, u is radial velocity, p is pressure, G is the universal gravitational constant, M is the planet's mass, c_{p} and c_{v} are the specific heats at constant pressure and volume, κ is the thermal conductivity, k is Boltzmann's constant, and q is the specific heating (less cooling) rate per unit volume. Note that if there is no radial flow then $u(r) = 0$ and eq. (B5.2) reduces to the hydrostatic equation, $\partial p/\partial r = -g\rho$.

The last of the equations above can be expressed in a number of forms. If one neglects the term in eq. (B5.3) involving $\partial T/\partial r$, the three equations are often referred to as *the Euler equations* of fluid dynamics. Techniques for solving these equations, e.g., Godunov's method, are described in Toro (1999) and LeVeque (2002). In general, numerical solution is complex, partly as a consequence of the existence of shock waves in the solutions and partly because of possible transitions from subsonic to supersonic flow. Simpler numerical techniques, specifically the Lax–Friedrichs method, have been used to find transonic solutions to these equations (Tian *et al.*, 2005), but artificial numerical diffusion in such techniques can violate mass and energy conservation and cause order of magnitude underestimation of the escape rate compared to a more accurate "constrained interpolation profile" (CIP) scheme (Kuramoto *et al.*, 2013). More elementary methods can be used to find steady-state (time-independent) solutions (Watson *et al.*, 1981).

Simplification is possible if one ignores both time dependence and the energy equation (B5.3) and assumes steady-state, isothermal expansion. Under these assumptions, eqns. (B5.1) and (B5.2) can be rewritten as

$$\frac{1}{r^2}\frac{d}{dr}\left(r^2 \rho u\right) = 0 \quad \text{mass conservation} \tag{B5.4}$$

$$\rho u \frac{du}{dr} = -\frac{dp}{dr} - \rho \frac{GM}{r^2} \quad \text{momentum conservation} \tag{B5.5}$$

Parker (1963) first presented a solution of these coupled equations, as follows. First, we eliminate dp/dr from equation (B4.5) using the ideal gas

$$p = \rho kT/m \tag{B5.6}$$

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Differentiation of eq. B5.6 with respect to r gives,

$$\frac{dp}{dr} = \frac{kT}{m}\frac{d\rho}{dr} \tag{B5.7}$$

where m is molecular (or atomic) mass, and we have used an isothermal approximation by setting $dT/dr = 0$. Substituting equation (B5.7) back into (B5.5) and dividing by ρ yields

$$u\frac{du}{dr} = -\frac{kT}{m}\frac{1}{\rho}\frac{d\rho}{dr} - \frac{GM}{r^2} \tag{B5.8}$$

Now, we use mass conservation equation (B5.4). Integrating with respect to r yields

$$r^2\rho u = \text{constant} \equiv C \tag{B5.9}$$

Here the constant C is related to the escape flux. The physical meaning of eq. (B5.9) is seen if we multiply by 4π, i.e., $4\pi r^2\rho u = F$, where F is the mass flux [kg s^{-1}] through the surface area of a sphere. Now take logarithms of eq. (B5.9) (i.e., $2\ln r + \ln\rho + \ln u = \ln C$) and differentiate with respect to r to get

$$\frac{2}{r} + \frac{1}{\rho}\frac{d\rho}{dr} + \frac{1}{u}\frac{du}{dr} = 0 \Rightarrow \frac{1}{\rho}\frac{d\rho}{dr} = -\frac{1}{u}\frac{du}{dr} - \frac{2}{r} \tag{B5.10}$$

We will use eq. (B5.10) to eliminate $(1/\rho)d\rho/dr$ in equation (B5.8). For convenience, we also define

$$u_0^2 = \frac{kT}{m}\; r_0 = \frac{GMm}{2kT}$$

Here, u_0 is the isothermal sound speed, while r_0 is related to the Jeans escape parameter (eq. (5.25)), $r_0 = r(\lambda_J/2)$. Dividing equation (B5.8) by kT/m, making the above substitutions, and using eq. B5.10, gives

$$\begin{aligned}\frac{u}{u_0^2}\frac{du}{dr} &= -\frac{1}{\rho}\frac{d\rho}{dr} - \frac{2r_0}{r^2} = \left(\frac{1}{u}\frac{du}{dr} + \frac{2}{r}\right) - \frac{2r_0}{r^2} \\ &\Rightarrow \frac{1}{u}\frac{du}{dr}\left(1 - \frac{u^2}{u_0^2}\right) = \frac{2r_0}{r^2} - \frac{2}{r}\end{aligned} \tag{B5.11}$$

This is a differential form of *Bernoulli's equation*, named after Daniel Bernoulli (1700–1782). In essence, the meaning of Bernoulli's equation is that absent any input or output of energy, when fluid is accelerated, the pressure drops. We will see how this applies to hydrodynamic escape shortly.

Bernoulli's equation has a wide range of mathematical solutions, some physical and some unphysical (see Fig. 5.11). As discussed further in the main text, the particular solution that is of physical interest to the hydrodynamic escape problem is the *transonic* solution that starts at low velocities near the planet and accelerates to high velocities at great distance. The distance, $r = r_0$, at which the flow goes supersonic, $u = u_0$, is termed the *critical point*. To avoid confusion with the critical level or exobase for Jeans' escape, we prefer to call it the *sonic level*. Note that both sides of equation (B5.11) vanish at the sonic level. This is what leads to the mathematical complexity of Bernoulli's equation.

Bernoulli's equation also has an integral form. Equation (B5.11) can be integrated term by term. We do this taking the limits from r to r_0 and u to u_0, which gives us the transonic solution:

$$\begin{aligned}\frac{1}{u}\frac{du}{dr} - \frac{1}{u_0^2}u\frac{du}{dr} = -\frac{2r_0}{r^2} - \frac{2}{r} \;&\Rightarrow \int_{u_0}^{u}\frac{1}{u}du - \frac{1}{u_0^2}\int_{u_0}^{u}udu = \int_{r_0}^{r}\frac{2r_0}{r^2}dr - \int_{r_0}^{r}\frac{2}{r}dr \\ &\Rightarrow \ln\left(\frac{u}{u_0}\right) - \frac{1}{2}\left[\frac{u^2}{u_0{}^2} - 1\right] = -2\left[\frac{r_0}{r} - 1\right] - 2\ln\left(\frac{r}{r_0}\right) \\ &\Rightarrow \ln\left(\frac{u}{u_0}\right) - \frac{1}{2}\left(\frac{u}{u_0}\right)^2 = -\frac{2r_0}{r} - 2\ln\frac{r}{r_0} + \frac{3}{2}\end{aligned} \tag{B5.12}$$

Consider the solution at large distances from the planet. Recall that the concept of hydrodynamic escape was motivated by the fact that the mass of a static atmosphere is infinite if the barometric law remains valid (Sec. 5.8). For

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

the case of transonic escape, though, the solution at large distances is quite different. As $r \to \infty$, the two largest terms in equation (B5.12) give

$$\frac{1}{2}\left(\frac{u}{u_0}\right)^2 \approx 2\ln\frac{r}{r_0}$$

or

$$u \approx 2u_0\left(\ln\frac{r}{r_0}\right)^{\frac{1}{2}} \tag{B5.14}$$

Inserting this back into equation (B5.9) shows that the mass density decreases at large distances as

$$\rho \approx \frac{C}{r^2 u} \propto \frac{1}{r^2 \ln r} \tag{B5.15}$$

The total atmospheric mass M_{atm} is given by

$$M_{atm} = 4\pi \int_{r_0}^{\infty} \rho r^2 dr \tag{B5.16}$$

Thus, M_{atm} is bounded because, in the integral, ρ decreases faster than $1/r^2$ for this transonic solution, i.e., as $1/(r^2 \ln r)$ according to eq. (B5.15). Thus, the transonic hydrodynamic escape solution is the one of physical interest for a planet like Earth that is embedded in a tenuous interplanetary medium.

transonic, and a termination shock will be created near the boundary with the interstellar medium where the solar wind slows down. The heliosphere is the region around the Sun dominated by the solar wind and its edge with the interstellar medium is the heliopause, which is somewhat more distant than the termination shock. Voyager 1 (launched in 1977) reached the heliopause at 122 AU distance in 2012, by measuring a sudden decrease in heliospheric ions by a factor $>10^3$ and increase in the intensity of galactic cosmic ray nuclei (Krimigis *et al.*, 2013; Stone *et al.*, 2013). Theoretically, higher background pressures could lead to subsonic outflow, or even to inflow if the background pressure were high enough.

For planetary winds, the background pressure would only be important for extremely strong stellar winds. As mentioned in Sec. 5.3, the ram pressure of the current solar wind on Earth is small. A planet has its host star and the impinging stellar wind on one side as compared to nearly empty interplanetary space on the other. Possibly, hydrodynamic flow far from a planet might be bent around in the anti-stellar direction by strong stellar wind pressure. Meanwhile, escape from the opposite side of the planet, like a comet's tail, would still be possible.

To assess whether an assumption of spherically symmetric hydrodynamic flow is reasonable, we can compare the planetary wind pressure at the sonic level with the stellar wind pressure. The planetary wind pressure is the static pressure plus the dynamic pressure, $p + \rho u^2$, where p is static pressure, ρ is density, and u is the flow speed. Calculations show that generally the planetary wind pressure exceeds the present solar wind pressure by orders of magnitude (Fig. 5.12), although we must bear in mind that stronger winds are possible from young stars.

Another complication is that planetary winds are expected to be largely neutral, and hence become collisionless at some great height. But if the exobase is beyond the sonic level, then the flow should become transonic and fluid equations are a good approximation. It can be demonstrated mathematically that the flow at the sonic level is independent of anything that occurs beyond that distance, so that the sonic level is a boundary condition.

If the exobase is below the sonic level, then the pressure force will be weak, the flow will remain subsonic, and flow velocity should eventually decrease at very high altitudes. We can quantity the speed as the Mach number M, the ratio of the flow speed, u, to the speed of sound. As shown by Walker (1977 p. 149), a subsonic flow with $M << 1$ causes expansion of an atmosphere and an outward velocity, but the inertial term, $\rho u(du/dr)$, in the equation of motion (eq. (B5.5)) is negligible. Consequently, the density profile is unaffected by the expansion and has an exponential, barometric form, as one can deduce from eq. (B5.5) with a negligible inertial term. Such flow should not be treated with purely

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

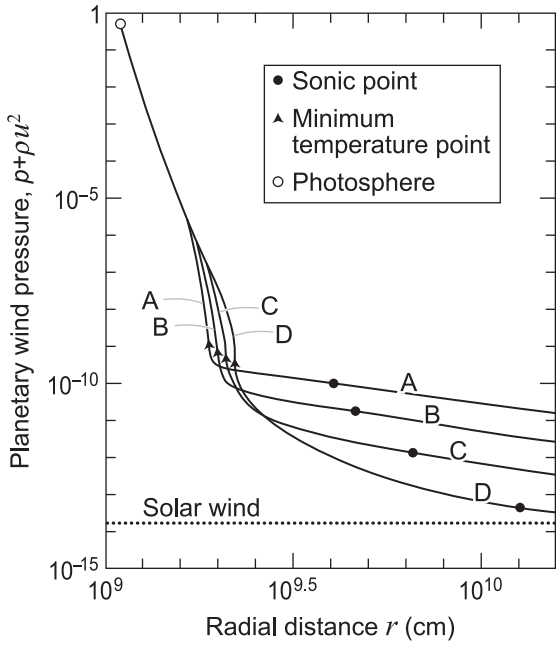

Figure 5.12 Transonic hydrodynamic escape model results from Sekiya, M. *et al.* (1980a, b) for a primordial H_2-rich atmosphere showing the pressure ($p + \rho u^2$) of the planetary wind as a function of distance. The four cases A–D have a net absorbed solar EUV flux of 1, 0.1, 0.01, and 0.001 W m^{-2}, respectively. This flux is absorbed in the outer atmosphere through ionization of H_2 where it heats the gas and drives hydrodynamic escape. The photosphere is a much lower altitude level where solar visible light is assumed to be absorbed and mostly reemitted in thermal infrared, so that it is not available to drive escape. Expansion of the atmosphere with increasing flow velocity typically causes a minimum temperature point, marked by triangle (▲). The sonic level (●) typically has a higher pressure than today's solar wind pressure, which is marked with a dashed horizontal line. (From Sekiya, M. *et al.* (1980). Reproduced with permission. Copyright 1980, The Physical Society of Japan.)

hydrodynamic fluid equations but kinetic models, as discussed above. Such models show that a subsonic regime in a single-component fluid produces an escape rate that is enhanced up to a factor of ~1.3–2 compared to the Jeans' escape rate (Merryfield and Shizgal, 1994; Volkov *et al.*, 2011a), which can reasonably be called Jeans-like because of the small factor (Johnson *et al.*, 2013d).

The question of subsonic flow arose in the context of old suggestions that escape of N_2 might be hydrodynamic on Pluto (Hunten and Watson, 1982; McNutt, 1989; Trafton, 1980; Trafton *et al.*, 1997), and rather than transonic, that escape was in a so-called "slow hydrodynamic escape" regime where fluid equations might be appropriate (Krasnopolsky, 1999; Strobel, 2008a; Tian and Toon, 2005). Pluto is a low-gravity body so that even for N_2 at once presumed temperatures ~80–90 K, the Jeans parameter (eq. (5.25)) at the exobase is fairly small, ranging ~4 to 6 depending on model type and solar UV variability (Erwin *et al.*, 2013).

If a sonic level does not lie in the collisional regime, we would *not* expect fluid equations to be accurate and, indeed, purely fluid equations applied to Pluto produce erroneous profiles compared to DSMC calculations (Erwin *et al.*, 2013; Johnson *et al.*, 2013d). While the purely fluid assumption underestimates the escape rate only slightly, it produces a very inaccurate temperature–density structure compared to more complete calculations (Fig. 5.13), which is important because temperature and density profiles are observable with remote sensing, whereas escape fluxes have to be inferred. In particular, the exobase in the purely fluid model is calculated to be at a much lower altitude and far colder than in more realistic models. When Jeans escape is evaluated at such an exobase of such a model, it leads to the erroneous conclusion that the escape rate greatly exceeds Jeans' escape. In fact, Pluto's escape is in a Jeans regime from a ~70 K exobase due to cooling from HCN and C_2H_2; and escape rates are only ~10^{23} N_2 molecules s^{-1} (Gladstone *et al.*, 2016). Thus, Fig. 5.13 is merely illustrative.

Subsonic hydrodynamic escape models have also been discussed for Titan (Strobel, 2008b) but, again, given that no sonic level lies below Titan's exobase, such an approach is physically problematic, as demonstrated by comparison to more sophisticated kinetic models (Bell *et al.*, 2014; Cui *et al.*, 2008) and DSMC models (Tucker and Johnson, 2009), including three-component ones (Tucker *et al.*, 2013). DSMC models show that hydrogen escapes from Titan at about the diffusion-limited rate, as mentioned previously (Sec. 5.9.2), while methane escapes at a negligible Jeans rate. Consequently, Titan's escape of hydrogen is ultimately limited by condensation of methane at the tropopause cold-trap.

5.10.2 Energy-Limited Escape

Particles in the solar wind obtain the energy needed to escape from heating in the solar corona at the base of the flow. That energy is transported outwards by thermal conduction, which is efficient in a fully ionized wind. This process does not work for planetary winds, because thermal conduction is inefficient for neutral particles. Instead, planetary winds are powered by absorption of extreme ultraviolet (EUV) radiation from a host star. EUV nominally spans wavelengths from 10 nm to 100 nm although the lower bound is indistinct and extends into soft x-rays. All EUV wavelengths below 91.2 nm can be directly absorbed by atomic hydrogen because this *Lyman limit* wavelength is where a photon has enough energy to ionize the H atom,

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

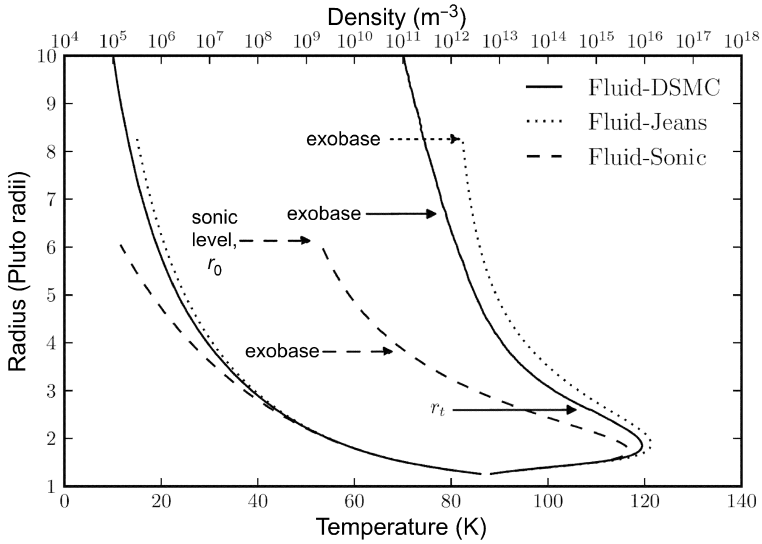

Figure 5.13 Comparison of simulations of atmospheric escape from Pluto with models using a single component, N_2, for solar-medium conditions at 32 AU. Curves on the left map to the upper axis of number density. Curves on the right map to the left vertical axis, which shows radial distance in number of Pluto radii. The radius of Pluto is 1153 km. *Solid lines* show a fluid-"direct simulation Monte Carlo" model, where fluid and DSMC methods are coupled at radius r_t, at a Knudsen number of 0.1. This model has an escape rate of 2.6×10^{27} N_2 molecules s^{-1}. *Dashed lines* show a hydrodynamic escape model, where fluid equations have been applied despite a sonic level above the exobase (Strobel, 2008a). This model has an escape rate of 2.5×10^{27} N_2 molecules s^{-1} but produces inaccurate number density and temperature profiles. *Dotted lines* show a model with an upper boundary of classical Jeans escape, which has an escape rate of 2.6×10^{27} N_2 molecules s^{-1}. (Modified with permission from Johnson *et al.* (2013d).)

temporarily creating a free proton and free electron. Indeed, most elements strongly absorb EUV because the outer electron binding energy (or, equivalently, the ionization potential) is typically smaller than EUV photon energy.

The maximum rate at which a gas can escape from an atmosphere in the Solar System can be calculated by equating the globally averaged flux of incoming solar EUV radiation $S_{\rm EUV}$, with the energy carried out by the escaping particles. Assume for now that the escaping gas is atomic hydrogen. Each hydrogen atom has gravitational energy GMm/r, where G is the universal gravitational constant, m is the hydrogen atom mass, and r is the radius of the relevant escape level. Consequently, the energy-limited flux of hydrogen, Φ_{el}, can be approximated by the following expression (Watson *et al.*, 1981)

$$\Phi_{el} = \frac{S_{\rm EUV}}{(GMm/r)} \tag{5.64}$$

To evaluate expression (5.64), we need an estimate for the solar EUV flux in the past. Despite being fainter in the visible (see Sec. 11.1), data from other stars suggest that the young Sun was significantly brighter at UV and EUV wavelengths (Claire *et al.*, 2012; Ribas *et al.*, 2005; Zahnle and Walker, 1982). The reason is that the Sun, like other stars, should have rotated more rapidly in its youth, before being slowed by torques exerted by its magnetic field as it interacted with the escaping solar wind. Faster rotation increases the strength of the solar magnetic dynamo. This increase causes increased flare activity and heats up the layers of the Sun's atmosphere, the chromosphere and corona. The Sun's photosphere is where most light comes from, but short-wavelength radiation is emitted from the hot chromosphere and corona.

Ribas *et al.* (2005) report short wavelength fluxes from 0.1 nm to 118 nm for the Sun and other young, solar-type stars based on data from the Far UV Spectroscopic Explorer (FUSE) satellite. Modifying their eq. (1) to include only wavelengths below 92 nm gives

$$S(<92\ \text{nm}) = 23.3{\tau_{\rm age}}^{-1.23}\ \text{erg cm}^{-2}\text{s}^{-1} = 23.3\times10^{-3}{\tau_{\rm age}}^{-1.23}\ \text{W m}^{-2} \tag{5.65}$$

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

This yields S = 3.6 erg cm^{-2} s^{-1} (=3.6 mW m^{-2}) at a current solar age, τ_{age}, of 4.56 billion years. Estimating the EUV flux that is available to drive escape requires division of this number by 4 to account for the ratio of the planet's surface area to its cross section. The flux should also be scaled upward to account for the larger effective cross-section of the atmosphere compared with the planet and downward to account for inefficiency of EUV heating. Not all EUV can drive escape because some absorbed energy is lost by radiation to space. Following Watson *et al.* (1981), we'll assume a heating efficiency of 0.15 and a geometric enhancement factor of 2; thus, S_{EUV}= (3.6 × 0.15 × 2)/4 = 0.3 erg cm^{-2} s^{-1}.

Now, let's use these numbers to calculate the *energy-limited escape rate* of hydrogen from the modern Earth. Plugging in values, and setting r equal to the radius of the Earth, yields $\Phi_{el} \cong 3 \times 10^{11}$ H atoms cm^{-2}s^{-1}. This is a very large number, nearly 1000 times greater than the diffusion-limited escape rate for hydrogen calculated in Sec. 5.8.4 (3.5×10^{8} cm^{-2} s^{-1}). It says nothing about how fast hydrogen escapes today, as the modern Earth, with its hydrogen-poor atmosphere, is *not* in the energy-limited escape regime. It shows instead that hydrogen escape could conceivably have been a very important process early in Earth's history if the atmosphere was more hydrogen-rich.

For illustrative purposes, let's calculate the amount of time that it would take for the hydrogen in Earth's oceans to escape at this rate. The oceans would be ~3 km deep if they were spread evenly over the globe, which is equivalent to a column mass of 3×10^{5} g cm^{-2}. Only 2/18 of this is hydrogen, so the hydrogen column mass is 3.3×10^{4} g cm^{-2}, and its column density is ~2×10^{28} H atoms cm^{-2}. The lifetime of this hydrogen is thus 2×10^{28} H atoms cm^{-2}/3×10^{11} H atoms cm^{-2} s^{-1} ≈ 6.7×10^{16} s or ~2 billion years. Equivalently, over two oceans' worth of hydrogen could have been lost over the Earth's history if hydrogen escaped at the energy-limited rate. Indeed, the actual number is more than five times higher than this if one accounts for the high early EUV flux predicted by eq. (5.65) by integrating that equation. Clearly, hydrogen escape has the potential to alter the water inventory of the Earth or other Earth-like planets, given the right conditions.

One caveat is that escape of hot hydrogen can reach a limit less than the energy limit called *radiation-recombination-limited escape*. With a high EUV flux, a hydrogen-rich upper atmosphere, e.g., on a hot Jupiter, can thermostat to a temperature ~10^{4} K because the energy input is balanced by radiative recombination and Lyman-α cooling rather than adiabatic cooling of the expanding gas through "*pdV*" work (Murray-Clay *et al.*, 2009). Consequently, a radiation-recombination-limited escape rate is less than the energy-limited escape limit, and is found to vary as $\sqrt{S_{EUV}}$ rather than linearly as in eq. (5.64) (Murray-Clay *et al.*, 2009; Owen and Jackson, 2012).

5.10.3 Density-Limited Hydrodynamic Escape

The energy-limited escape flux predicted by eq. (5.64) is also not likely to be achieved when availability of hydrogen is limiting, particularly in a multi-component atmosphere on a rocky world. If more hydrogen is available at the base of the expansion, then a greater percentage of the absorbed EUV energy can be utilized to drive escape. When hydrogen is scarce, much of this EUV energy is either absorbed and radiated to space or conducted downwards through the expanding thermosphere. Higher hydrogen densities also increase the spatial extent of the atmosphere, thereby increasing the total absorbed energy that powers escape.

The importance of hydrogen density is illustrated by the calculations of Watson *et al.* (1981), shown in Fig. 5.14. Transonic solutions to the hydrodynamic escape equations are labeled A to E in Fig. 5.14 and correspond to a progressive increase of the number density at the lower boundary. The solutions were found using the *shooting method*, whereby one integrates both outward and inward from the sonic level, attempting to match the boundary conditions on each side. Figure 5.14(b) shows how the number density increases from case A to case E. Case E has an enormous number density at the bottom of the model, 120 km, and is not physically plausible. The case E escape flux equals, or even slightly exceeds, the energy-limited flux because energy limit is not absolute given that the effective cross-section of the atmosphere increases with increasing number density. The large escape flux in case E causes strong adiabatic cooling within the flow, resulting in a deep temperature minimum near 2000 km altitude (Fig. 5.14(a)). This also is physically unreasonable, but it demonstrates that this solution is indeed near the energy-limited escape rate. The low-hydrogen cases, A and B, are more physically realistic and have escape rates that are 20%–50% of the energy-limited escape rate.

A similar dependence of the escape rate on hydrogen density has been demonstrated for the process of water loss from Venus during a runaway greenhouse (Kasting and Pollack, 1983). Their calculations are like those shown by the transonic curve IV in Fig. 5.11. In Kasting

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

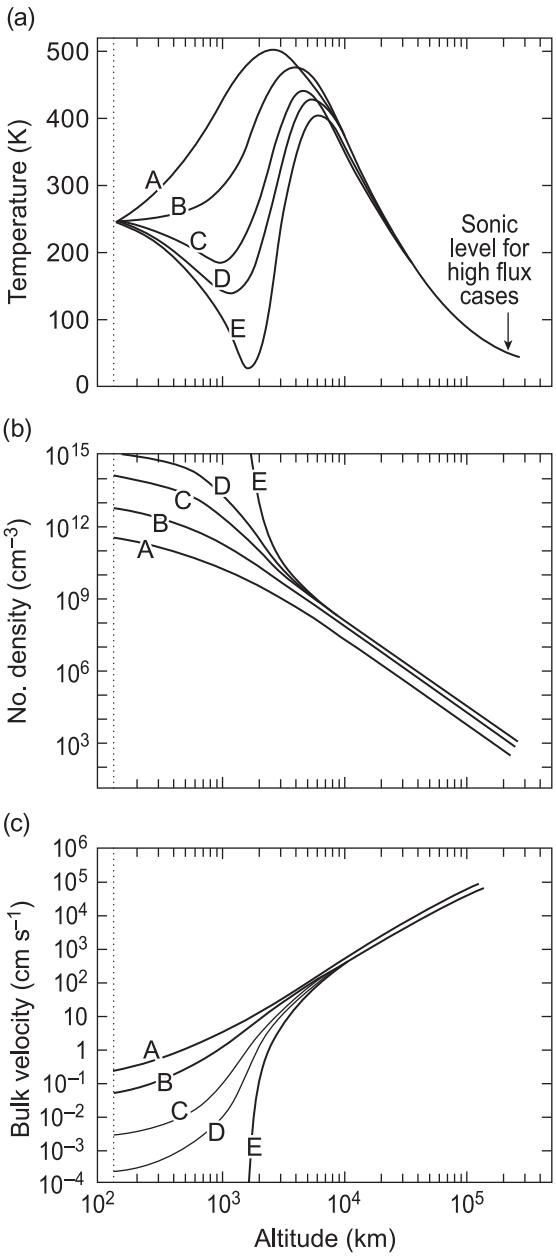

Figure 5.14 Temperature (a), number density (b), and upward velocity (c) versus altitude for a transonically escaping pure atomic H terrestrial atmosphere. (After Watson *et al.* (1981).). The sonic level is shown by the arrow in panel (a) at 2×10^5 km or ~30 planet radii. An EUV heating efficiency of 0.15 was assumed in the calculations. The escape fluxes (H atoms cm^{-2} s^{-1}) for the five cases, normalized to the Earth's surface, are: (A) 5.7×10^{10}, (B) 2.2×10^{11}, (C) 2.9×10^{11}, (D) 3.1×10^{11}, and (E) 3.5×10^{11}. The escape flux for case E is close to the energy-limited escape flux.

and Pollack, the factor that was varied was the H_2O mixing ratio at the lower boundary. Chemistry included within the model converted this to atomic H within the expanding upper atmosphere. The solutions generated in this model were subsonic, but the authors argued that the escape rates were close to the transonic escape rate because the peak Mach number in the flow (the velocity divided by the sound velocity) exceeded ~0.7.

5.10.4 Maximum Molecular Mass Carried Away in Hydrodynamic Escape

Jeans' escape is an exponential function of molecular mass (Sec. 5.6), and consequently is important only for hydrogen on Earth, but this is not necessarily true in hydrodynamic escape because heavier species can be dragged along.

Consider an atmosphere composed largely of H_2 that flows out into space. (This may be more realistic for the early Earth than the pure atomic H atmosphere modeled by Watson *et al.* (1981).) The escaping hydrogen will drag along well-mixed minor gases. If there were no diffusion, there would be no separation by mass and the mixing ratios of the minor constituents would remain constant. Diffusion, however, allows the heavier gases to flow downwards under the action of gravity and this, in principle, means that heavier species will be retained on the planet to a degree depending on their mass. Whether the discrimination in mass is important or not depends on the relative magnitudes of the hydrogen outflow and the diffusion velocity of the minor constituent (Hunten, 1979a; Hunten *et al.*, 1987; Sekiya *et al.*, 1980a, b).

In a full treatment of outflow in a hydrodynamically escaping atmosphere, an equation of motion is used that includes acceleration terms (Zahnle and Kasting, 1986). These terms have an important influence on the escape flux of the light, major gas constituent. However, their effect is small for the escape of heavy, minor gases. This is because diffusion processes occurring between the homopause and the sonic level where the outflow velocity is subsonic determine the flux of a heavy gas molecule. At these relatively low altitudes, the ambient density closely follows the barometric equation and acceleration terms in the momentum equation are negligible. Once the fluxes of heavy constituents are established they must obey the continuity equation at higher altitudes.

Consider the diffusion of a heavy, minor constituent relative to the ambient light gas. Our treatment follows Walker (1982) and Hunten *et al.* (1987). See also Chamberlain and Hunten (1987), Ch. 7. We denote the masses, fluxes, vertical velocities relative to the planet, and number densities of the light gas 1 and heavy gas 2 by

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

m_1 and m_2, F_1 and F_2, w_1 and w_2, and n_1 and n_2, respectively. In an isothermal atmosphere, the relative velocities are determined by diffusion. Following eq. (5.35) and its manipulation in Sec. 5.8.1, and neglecting the minor terms containing thermal diffusivity, the relative velocities can be written as

$$w_1 - w_2 = \frac{n_1 w_1}{n_1} - \frac{n_2 w_2}{n_2} = \frac{F_1}{n_1} - \frac{F_2}{n_2} = -\frac{b}{n_2}\left[\frac{1}{n_1}\frac{dn_1}{dr} + \frac{m_1 g}{kT}\right] \tag{5.66}$$

$$w_2 - w_1 = \frac{F_2}{n_2} - \frac{F_1}{n_1} = -\frac{b}{n_1}\left[\frac{1}{n_2}\frac{dn_2}{dr} + \frac{m_2 g}{kT}\right] \tag{5.67}$$

In these equations, r indicates the distance from the center of the planet. Also we have used the relationship $F = nw$ between vertical flux F, number density n, and vertical velocity w, and we express the diffusion coefficient as $D = b/n$. Each of these equations can be rearranged to put the number density gradient on the left-hand side:

$$\frac{dn_1}{dr} = -\frac{m_1 g}{kT} n_1 + \frac{1}{b}(n_1 F_2 - n_2 F_1) \tag{5.68}$$

$$\frac{dn_2}{dr} = -\frac{m_2 g}{kT} n_2 + \frac{1}{b}(n_2 F_1 - n_1 F_2) \tag{5.69}$$

If we add eqs. (5.68) and (5.69), the last terms on the right-hand side cancel, and we get

$$\frac{d}{dr}(n_1 + n_2) = -(n_1 m_1 + n_2 m_2)\frac{g}{kT} \tag{5.70}$$

This equation gives the variation of the total number density ($n = n_1 + n_2$) with altitude, and is a differential form of the barometric law. We define the mole fraction, or mixing ratio, of heavy gas 2 as

$$X_2 = \frac{n_2}{n_1 + n_2} = \frac{n_2}{n} \tag{5.71}$$

The logarithm of eq. (5.71) is $\ln X_2 = \ln n_2 - \ln n$, which we can differentiate with respect to radial distance r, to give

$$\frac{1}{X_2}\frac{dX_2}{dr} = \frac{1}{n_2}\frac{dn_2}{dr} - \frac{1}{n}\frac{dn}{dr} \tag{5.72}$$

If we now substitute from eq. (5.69) for the first term on the right-hand side of eq. (5.72) and from eq. (5.70) for the second term, we get

$$\begin{aligned}\frac{1}{X_2}\frac{dX_2}{dr} &= \frac{1}{n_2}\left(-\frac{m_2 g}{kT} n_2 + \frac{1}{b}(n_2 F_1 - n_1 F_2)\right) - \frac{1}{n}\left(-(n_1 m_1 + n_2 m_2)\frac{g}{kT}\right) \\ &= -\frac{m_2 g}{kT} + \frac{1}{b}(F_1 - \frac{X_1}{X_2} F_2) + (X_1 m_1 + X_2 m_2)\frac{g}{kT}\end{aligned} \tag{5.73}$$

We note that $1 - X_2 = X_1$, or $-(m_2 - X_2 m_2) = -X_1 m_2$, which we can apply to collected terms in $m_2 g/kT$. Thus, eq. (5.73) rearranges to

$$\frac{1}{X_2}\frac{dX_2}{dr} = \frac{1}{b}\left(F_1 - \frac{X_1}{X_2} F_2\right) - (m_2 - m_1)\frac{X_1 g}{kT} \tag{5.74}$$

If heavy gas 2 is carried along efficiently by light gas 1, then we can assume that the mole fraction X_2 will be constant with height, so the left-hand side will be zero, giving

$$\begin{aligned}(m_2 - m_1)g &= \frac{kT}{b}\left(\frac{F_1}{X_1} - \frac{F_2}{X_2}\right) \\ \Rightarrow (m_2 - m_1)g &= \frac{nkT}{b}(w_1 - w_2)\end{aligned} \tag{5.75}$$

This equation expresses a balance of forces. A molecule of gas 2 will be subject to a downward gravitational force of $m_2 g$ and an upward buoyancy force of $m_1 g$, resulting in a net downward force of $(m_2 - m_1)g$, given by the left-hand side of eq. (5.75). This net downward force will be balanced by an upward viscous drag, proportional to the difference in velocities of the molecules, $w_1 - w_2$, given by the right-hand side of eq. (5.75). This is illustrated in Fig. 5.15.

Equation (5.75) can be interpreted in terms of fluxes of the two gases. Consider a mass of gas 2 sufficiently heavy that it is not dragged along out into space by the lighter gas 1. Putting $F_2 = 0$ in eq. (5.75), we obtain the required mass:

$$m_{crossover} = m_1 + \frac{kTF_1}{bgX_1} \tag{5.76}$$

This mass is called the *crossover mass* and is the smallest mass for which the flux of constituent 2 is zero. It can be interpreted thus.

- If $m_2 > m_{crossover}$ then the buoyancy force is not enough to compensate for the gravitational force and viscous drag acting on molecules of gas 2, and gas 2 will not be lifted out of the atmosphere. The mole fraction X_2 will decrease with altitude with a scale height that is the diffusive equilibrium value augmented by an amount depending on the flux of gas 1. Meanwhile, the mole fraction X_1 will approach a value of 1 at high altitude.
- If $m_2 = m_{crossover}$ then the drag force is just sufficient to balance the net downward force on molecules of gas 2, but gas 2 will not be lifted out of the atmosphere.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

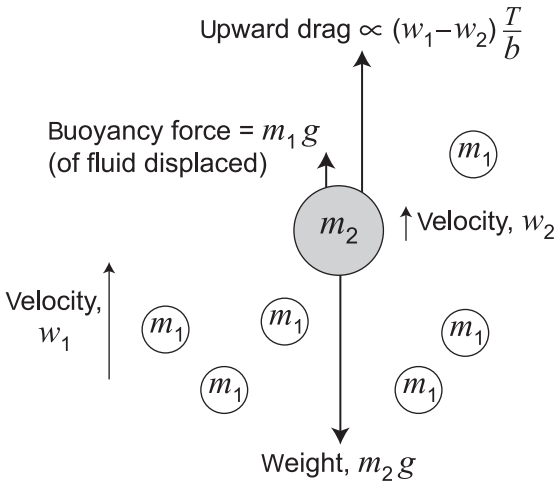

Figure 5.15 The forces acting on a heavy gas molecule of mass m_2 moving at an upward velocity w_2 immersed in a light gas flowing upwards in the diffusively separated upper atmosphere. The light gas has molecules of mass m_1 that move at an upward velocity of w_1. Absolute temperature is T and $b = Dn$ where D is the diffusion coefficient and n is total number density. If the initial total upward force on molecule 2 is greater than the downward force of its weight minus buoyancy, molecule 2 will be accelerated until the upward and downward forces come into balance.

- If $m_2 < m_{crossover}$ then constituent 2 will be carried along by the flux of light gas and swept out to space. Mole fractions X_1 and X_2 will be independent of altitude.

We can derive an expression for the flux of heavy gas 2. Rearranging eq. (5.75), we get

$$F_2 = \frac{X_2}{X_1} F_1 \left(1 - (m_2 - m_1)\frac{bgX_1}{kTF_1}\right) \tag{5.77}$$

If we then note that $bgX_1/kTF_1 = 1/(m_{crossover} - m_1)$ from eq. (5.76), we can substitute in eq. (5.77) to give

$$\begin{aligned} F_2 &= \frac{X_2}{X_1} F_1 \left(1 - \frac{m_2 - m_1}{m_{crossover} - m_1}\right) \\ &= \frac{X_2}{X_1} F_1 \left(\frac{m_{crossover} - m_2}{m_{crossover} - m_1}\right) \end{aligned} \tag{5.78}$$

The important point from eq. (5.78) is that if heavy constituent 2 is swept out of the atmosphere into space, then its flux, F_2, will vary linearly with its mole fraction X_2 and molecular mass m_2. Note that the crossover mass is different for different gases, however, because the binary diffusion coefficient, b, in eq. (5.76) varies with species. The linear dependence on molecular mass contrasts with the exponential dependence on molecular mass in Jeans' escape described earlier in Sec. 5.6.2.

In the concept of crossover mass, there is a maximum molecular mass than can be carried away. Gases with masses exceeding $m_{crossover}$ are not affected. Actually, numerical results that incorporate nonlinear terms show a very slow loss of gases with molecular masses exceeding $m_{crossover}$ (Zahnle and Kasting, 1986) but the same results also show that eq. (5.78) is a good approximation for the escape of a trace constituent provided that

$$\left(\frac{m_{crossover} - m_2}{m_{crossover} - m_1}\right) > \frac{m_1}{m_2} \tag{5.79}$$

For noble gases escaping in hydrogen, eq. (5.77) is a good approximation for Ne, Ar, Kr, and Xe, but is not so good for He.

The crossover mass can be calculated as a function of the hydrodynamic escape flux. Consider the energy-limited escape flux that was calculated for atomic hydrogen on Earth in Sec. 5.10.2. Let's divide that flux by 2 to convert it to an H_2 flux, yielding $F_1 = 1.5\times10^{11}$ H_2 molecule cm^{-2} s^{-1}. Substituting this value of F_1 into the equation for the crossover mass, eq. (5.76), and dividing by the mass of an H atom, m_H, to convert to atomic mass units yields

$$M_{crossover} - M_1 = \frac{kTF_1}{bgm_H X_1} \tag{5.80}$$

Here, $M_{crossover}$ is the molecular mass of the heavy species in a.m.u. and M_1 (= 2 a.m.u.) is the molecular mass of H_2. Assume T = 400 K (from Fig. 5.14) and binary diffusion parameter $b = 2\times10^{19}$ cm^{-1} s^{-1} from Table 5.2. Taking g = 980 cm s^{-2} and X_1 = 1 yields $M_{crossover} - M_1 = 0.25$. In other words, escape of H_2 at the energy-limited rate from the modern Earth would be incapable of dragging along any heavier gas, including He.

Now consider gases that might have been dragged away earlier in Earth's history when the solar EUV flux was higher. Assume that the escape flux was energy-limited, i.e., eq. (5.65). H_2 escape fluxes and corresponding crossover masses are listed in Table 5.3. Evidently, gases as heavy as neon (molecular mass 20 or 22), N_2 or CO, might have escaped during the first 100–200 million years of Earth's history.

This same analysis, based on eq. (5.80) can be applied to other planets if one scales the EUV flux by orbital distance and adjusts for the planet's gravity. For Mars, interestingly, these two factors almost cancel: the solar flux is lower by a factor of 2.3, whereas gravity is lower by a factor of 2.6. Hence, the predicted crossover masses for Mars are nearly the same as for Earth.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Table 5.3 Energy-limited escape fluxes of H_2 and the corresponding maximum mass of a molecule (the crossover mass) that can escape by being dragged along by the hydrogen to space at different times in Earth's history.

Time after Earth's formation (billions of years)	EUV enhancement compared to today	Energy-limited escape rate of H_2 (cm^{-2} s^{-1})	Crossover mass, $M_{crossover}$ (a.m.u.)
0.1	110	1.7×10^{13}	30
0.2	47	7.0×10^{12}	14
0.5	15	2.3×10^{12}	5.8
1.0	6.5	9.8×10^{11}	3.6
2.0	2.8	4.2×10^{11}	2.7
4.56	1	1.5×10^{11}	2.25

5.11 Mass Fractionation by Hydrodynamic Escape

5.11.1 Fractionation Theory

We now consider how the quantity of heavy gases will be fractionated by mass when hydrodynamic escape is integrated over time because then we can see if predictions are consistent with data, such as those of noble gas isotopes. In the previous section, we saw how the flux of a heavy constituent varies linearly with its mass in eq. (5.78). Let us assume that the inventory of light gas is denoted by N_1 and that the inventory of heavy gas is denoted by N_2. If we substitute N_2/N_1 for X_2/X_1 in eq. (5.78), we get

$$\frac{F_1}{F_2} = \left(\frac{m_{crossover} - m_1}{m_{crossover} - m_2}\right)\frac{N_1}{N_2} \tag{5.81}$$

Because escape fluxes are proportional to their reservoirs, the evolution of the reservoirs can be treated as a *Rayleigh fractionation* process, which can be thought of as analogous to a distillation process where the heavy isotope concentration depends increasingly on the depletion of the light component.

The fundamental equation for Rayleigh fractionation is as follows, where dN is an infinitesimal number of particles removed per unit time,

$$\frac{dN_1}{dN_2} = (\text{fractionation factor}) \times \frac{N_1}{N_2} = (1+y)\frac{N_1}{N_2} \tag{5.82}$$

Here, we write the fractionation factor as 1+ y, which is usually slightly larger than unity, so that y is a very small value. Defining the fractionation factor this way gives the degree to which one gas escapes relative to another. Other symbols are sometimes employed for this "1+ y" factor in the literature, such as R (Yung *et al.*, 1988) or x (Zahnle and Kasting, 1986). The present notation is convenient for dealing with cases where the mass difference between species 1 and species 2 is small compared to their total mass, as is the case for most noble gas isotopes. Bearing in mind that fluxes are $F_1 = dN_1/dt$ and $F_2 = dN_2/dt$, comparison of eq. (5.82) with eq. (5.81) shows that

$$(1+y) = \left(\frac{m_{crossover} - m_1}{m_{crossover} - m_2}\right) \tag{5.83}$$

We proceed from the basic Rayleigh fractionation equation, eq. (5.82), by integrating. We assume initial inventories indicated by a superscript of 0, as follows:

$$\int_{N_1^0}^{N_1}\frac{dN_1}{N_1} = (1+y)\int_{N_2^0}^{N_2}\frac{dN_2}{N_2} \Rightarrow \ln\left(\frac{N_1}{N_1^0}\right) = \ln\left(\frac{N_2}{N_2^0}\right)^{(1+y)}$$
$$\Rightarrow \left(\frac{N_2}{N_2^0}\right) = \left(\frac{N_1}{N_1^0}\right)^{1/(1+y)} \Rightarrow \left(\frac{N_2}{N_2^0}\right) = \left(\frac{N_1}{N_1^0}\right)^{\frac{(m_{crossover}-m_2)}{(m_{crossover}-m_1)}} \tag{5.84}$$

Figure 5.16 shows a plot of eq. (5.84) for an example crossover mass, of 100 a.m.u. The graph shows how the depletion of the heavier gas (N/N_0) increases with the depletion of the lighter gas, where the lines of increasing slope indicate the latter. The depletion of the heavy gas also depends on its particular mass, m_2, shown on the horizontal axis. However, eq. (5.84) is a simplification because we are assuming that y, and by implication the crossover mass, is constant in time. In reality, as the solar EUV flux decreases over time, the crossover mass decreases and also the hydrogen escape flux (Table 5.3). This would cause the slope of lines in Fig. 5.16 to decrease with time because heavier gases would cease to evolve while lighter gases would continue to change. This would cause curved lines in Fig. 5.16, concave downwards.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

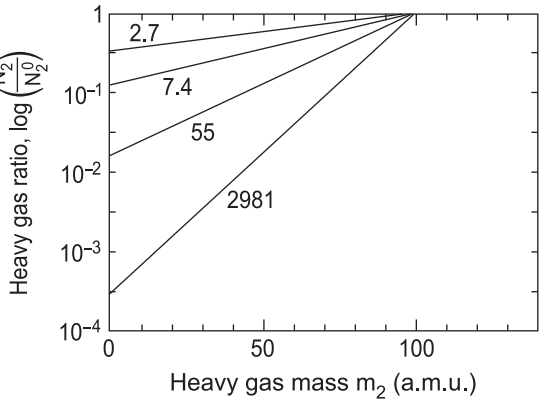

Figure 5.16 The evolution of the inventory of heavy gas 2 relative to its initial inventory as a function of molecular mass, m_2. For this graph, a constant crossover mass of 100 is assumed and a constant hydrogen escape flux. The numbers on the diagonal lines indicate indicate values of N_0^1/N_1, which shows the depletion of the lighter gas inventory. Thus the lines at the bottom of the plot correspond to later times in an evolutionary history. (Adapted from Hunten *et al.* (1987).)

5.11.2 Applications of Mass Fractionation in Hydrodynamic Escape: Noble Gas Isotopes

In Ch. 6, we discuss how elemental and isotopic abundances of gases provide constraints on the origin and evolution of the atmospheres of the rocky planets. In particular, the noble gases provide good tracers of atmospheric evolution for three reasons. First, noble gases tend to reside in the atmosphere because of their (near) chemical inertness. Second, the fractionation of different isotopes of non-radiogenic noble gases tells us about atmospheric evolution because lighter isotopes are lost preferentially to a degree that depends on ancient atmospheric conditions. Third, radiogenic noble gases – those derived from radioactive decay of other elements – act as chronometers of planetary evolution. Of course, all of these inferential principles are tempered by the fact that the patterns of abundance and fractionation in noble gases are complex and presently not fully understood.

Nonetheless, hydrodynamic escape could account for the isotopic fractionation of some of the noble gases found in planetary atmospheres. On the Earth, hydrodynamic escape could explain the isotopic differences between the $^{20}Ne/^{22}Ne$ ratio of ~9.8 in the atmosphere versus that in the Earth's upper mantle, ~10–13 (Pepin, 1991; Sasaki and Nakazawa, 1988; Zahnle *et al.*, 1990). Hydrodynamic escape allows the lighter neon isotope to escape preferentially. Hydrodynamic escape has also been invoked to explain the Martian $^{36}Ar/^{38}Ar$ ratio (Bogard, 1997), which is isotopically heavy (4.2±0.1 (Atreya *et al.*, 2013)) compared with a terrestrial ratio of 5.32 and the average carbonaceous chondrite value of ~5.3 (Pepin, 1989). The fractionation of xenon on Earth (Hunten *et al.*, 1987; Pepin, 1991, 2000; Pepin and Porcelli, 2006; Sasaki and Nakazawa, 1988) and Mars (Pepin, 1991) has also been attributed to hydrodynamic escape. These applications of hydrodynamic escape are discussed below.

Terrestrial neon. The Earth's mantle is enriched in light neon isotopes relative to the atmosphere, and escape could have made the atmosphere isotopically heavy. Both $^{20}Ne/^{22}Ne$ and $^{21}Ne/^{22}Ne$ ratios are higher in the mantle. $^{20}Ne/^{22}Ne$ in mid-ocean ridge basalts ranges from near the atmospheric value (9.8) up to 13, while $^{21}Ne/^{22}Ne$ ranges from near air (0.029) to 0.07 (Farley and Neroda, 1998). The $^{20}Ne/^{22}Ne$ ratio of the material from which Earth accreted was probably ~12.5–13.6 (Farley and Poreda, 1993).

Why is atmospheric $^{20}Ne/^{22}Ne$ smaller than in the mantle? Either the atmosphere was partially derived from an external isotopically light neon source during late bombardment or the atmosphere has been modified by escape. In the latter case, models show that hydrodynamic escape can drag off ^{20}Ne in preference to ^{22}Ne, and reduce the solar $^{20}Ne/^{22}Ne$ ratio to the observed atmospheric value (Hunten *et al.*, 1987; Zahnle *et al.*, 1990). A hydrogen-rich upper atmosphere is required for this to have occurred. Also the escape must have happened early in Earth's history when the solar EUV flux was much higher than it is today (Claire *et al.*, 2012; Ribas *et al.*, 2005; Walter and Barry, 1991; Zahnle and Walker, 1982). Both conditions would have been met in an impact-produced steam atmosphere that occurred continuously during the main accretion period, and intermittently thereafter (Matsui and Abe, 1986a, b; Zahnle *et al.*, 1988).

Zahnle *et al.* (1990) showed that fractionation of neon would have occurred in a steam atmosphere as a byproduct of hydrodynamic hydrogen escape regulated at the diffusion-limit through an atmosphere of a major background constituent, such as CO_2, N_2, or CO. Unlike the $^{20}Ne/^{22}Ne$ ratio, there is no clear distinction between mantle and atmospheric $^{36}Ar/^{38}Ar$, which suggests that Earth's argon was unaffected by the hydrodynamic escape. The reason neon can escape while argon cannot is that neon is less massive than the likely background gases in the atmosphere (CO_2, N_2, or CO). Figure 5.17 shows that it would have taken only ~10 m.y. to produce

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

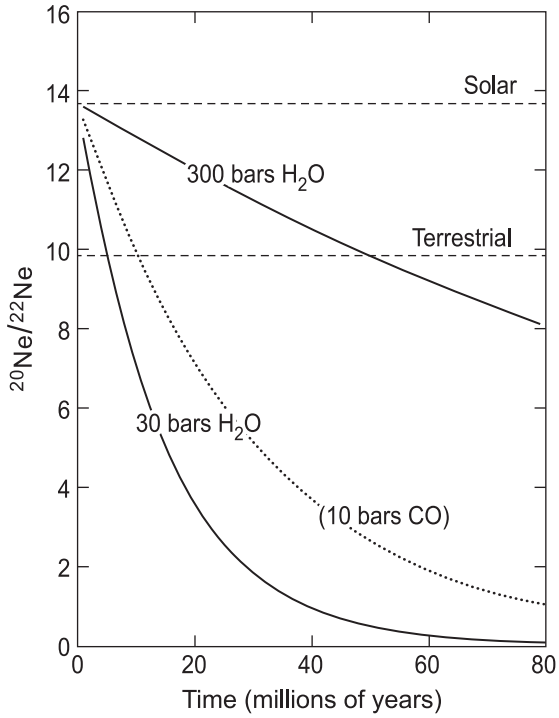

Figure 5.17 Neon isotope fractionation produced by hydrodynamic escape of hydrogen in various steam atmospheres during accretion of the Earth. Three cases are shown. The calculated timescale for neon fractionation is consistent with the expected lifetime of impact-induced steam atmospheres on early Earth. (From Fig. 15 of Zahnle *et al.* (1990).)

the observed neon fractionation with a 30 bar steam atmosphere and 10 bars of CO. Atmospheres with less CO take shorter times, while thicker atmospheres take longer.

Martian argon. From analysis of trapped pockets of Martian air in the impact glass of the EET79001, Wiens *et al.* (1986) deduced a $^{36}Ar/^{38}Ar$ value of 4.1 ± 0.2 within the uncertainty of 3.6 ± 0.44 obtained by Swindle *et al.* (1986). The ratio in the Martian atmosphere has been measured by the *Curiosity Rover* as 4.2 ± 0.1 (Atreya *et al.*, 2013). This ratio is considerably less than 5.305 ± 0.008 in the Earth's atmosphere (Lee *et al.*, 2006) or 5.50 ± 0.01 in the solar wind derived from samples collected by the *Genesis* mission (Pepin *et al.*, 2012; Vogel *et al.*, 2011). Thus, light argon isotopes have been preferentially lost from Martian air relative to heavy argon.

Atreya *et al.* (2013) argue that solar wind sputtering since the end of heavy bombardment accounts for the $^{36}Ar/^{38}Ar$ ratio, but in our opinion it is possible that ancient hydrodynamic escape was the main fractionation mechanism. In sputtering, the solar wind picks up and accelerates ions, and a fraction of the energetic ions or neutrals impacts the exobase, causing Ar escape (Hutchins and Jakosky, 1996; Hutchins *et al.*, 1997; Jakosky *et al.*, 1994; Jakosky and Phillips, 2001). Sputtering is fractionating because argon isotopes are diffusively separated above the homopause. Models estimate that 75%–99% of ^{36}Ar is lost (Hutchins *et al.*, 1997). But if this much Ar is lost, then it needs to be replenished from volcanism to be consistent with the Ar/Kr ratio on Mars because Kr is not subject to sputtering. Moreover, neon, which is even more prone to sputtering than Ar, needs even more replenishment. However, the estimated volcanic outgassing on Mars is too small by one or two orders of magnitude to do the job (Hutchins and Jakosky, 1996).

Early hydrodynamic escape provides an alternative for the fractionation of Martian argon (Pepin, 1991; Zahnle, 1993a; Zahnle *et al.*, 1990). Of course, if argon escapes and fractionates, neon must also. Martian atmospheric $^{20}Ne/^{22}Ne$ appears to be ~10, somewhat similar to the terrestrial atmospheric ratio, though some data are consistent with lower values for Mars (Bogard *et al.*, 2001; Bogard and Garrison, 1998). If the original Martian ratios of $^{36}Ar/^{38}Ar$ and $^{20}Ne/^{22}Ne$ were 5.35 and 13.7, respectively, then diffusion-limited hydrodynamic escape results in a $^{20}Ne/^{22}Ne$ ratio no greater than 9.5 ± 1.3, consistent with observation (Fig. 5.18). The presence of abundant CO_2 or a hydrogen escape flux sufficient to drag away neon but not argon would result in a yet lower $^{20}Ne/^{22}Ne$ ratio.

Terrestrial xenon. Interpretation of xenon is complicated because xenon has nine stable isotopes, several of which have been affected by the decay of extinct radionuclides. Also, xenon, with atomic weight 131.3, should be less depleted and less fractionated than krypton with atomic weight 83.8. But the opposite is observed. Krypton is depleted in the terrestrial atmosphere by a factor of 3.3×10^4 relative to solar composition while xenon is depleted by a factor of 4.8×10^4. Nonradiogenic xenon isotopes are also much more strongly fractionated compared to krypton isotopes. The unexpected paucity of xenon is known as the *missing xenon paradox* (Ojima and Podosek, 2002; Pepin, 1991; Tolstikhin and O'Nions, 1994).

Vigorous hydrodynamic escape could produce the observed fractionation pattern in xenon (Hunten *et al.*, 1987; Pepin, 1991, 2006; Sasaki and Nakazawa, 1988) but additional circumstances must have led to no correspondingly large fractionation in krypton isotopes, which are less massive. There are three possible solutions. First,

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

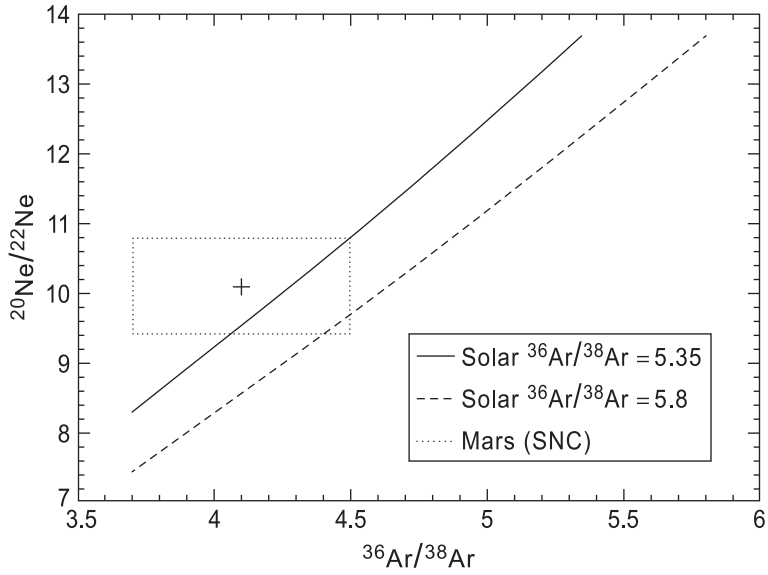

Figure 5.18 The $^{20}Ne/^{22}Ne$ ratio produced in fractionating argon by hydrodynamic escape from Mars, assuming an initial $^{20}Ne/^{22}Ne$ of 13.7. (From Fig. 7 of Zahnle (1993).)

Pepin (1991) suggested that xenon fractionation is a significant remnant of an early escaped atmosphere, whereas neon, argon and krypton were later outgassed from the mantle while xenon was not because it was incorporated into the core. Pepin (1991) assumed that xenon behaves as a siderophile at high pressure to justify why it partitions into the core. Second, xenon's low solubility in silicate melts could be used to argue that it was the most strongly partitioned into the earliest atmosphere. Third, xenon may have escaped as an ion during hydrodynamic escape of hydrogen ions along the open magnetic field lines at high latitudes (Zahnle, 2000). Ions interact strongly with each other and consequently cross-sections for ion–ion interactions are large. In contrast to xenon, krypton is extremely difficult to ionize and would not be subject to such ionic escape.

A very intriguing report is that the pattern of xenon's nine stable isotopes was lighter in the Archean than today, based on analysis of fluid inclusions in Archean barites and quartz (Hebrard and Marty, 2014; Pujol *et al.*, 2011). This relationship would require xenon to escape to space during the Archean and mass fractionate, long after the early period of very high solar EUV. If xenon escaped as an ion, a plausible explanation is that it was dragged by rapid hydrogen escape in a *polar wind* (Sec. 5.7). Such hydrogen escape would be expected from an anoxic Archean atmosphere relatively rich in H_2 and CH_4 (e.g., Catling *et al.*, 2001). However, whether the xenon data provide evidence for such theories is uncertain because the ancient xenon might be a mixture of modern air with an unfractionated mantle component (Pepin, 2013).

Martian xenon. Like Earth, Mars also has *missing xenon*, except more severely. The nonradiogenic isotopes are ~80 times less abundant and have a fractionation pattern generally similar to the Earth's. On the other hand, the ^{129}Xe derived from the decay of ^{129}I (half-life 17 m.y.) is about one third that of Earth. The high ratio of radiogenic to nonradiogenic xenon implies that escape took place very early before ^{129}I had undergone several half lives. Fractionation of the nonradiogenic isotopes on Mars can also plausibly be explained by mass fractionation during hydrodynamic escape of hydrogen (Pepin, 1991). However, explaining why krypton is not also strongly fractionated must be considered, as for the Earth. The same kind of explanations for the Earth can be proffered for Mars.

We have not discussed Venus above because Venus is anomalous in its noble gas abundances. On a gram per gram of planet basis, Venus is remarkably well endowed with nonradiogenic argon ($^{36,38}Ar$) and neon. It has ~60 times more ^{36}Ar than Earth, for example. It is plausible that Venus stochastically accreted a large (>600 km) comet from the outer Solar System, where temperatures would have been cold enough for argon to condense (Owen and Bar-Nun, 1995). The chance of such a single event happening is about 25% (Zahnle, 1998). This probability is large enough for plausibility and on the other hand small enough that Earth need not have suffered a similar fate.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

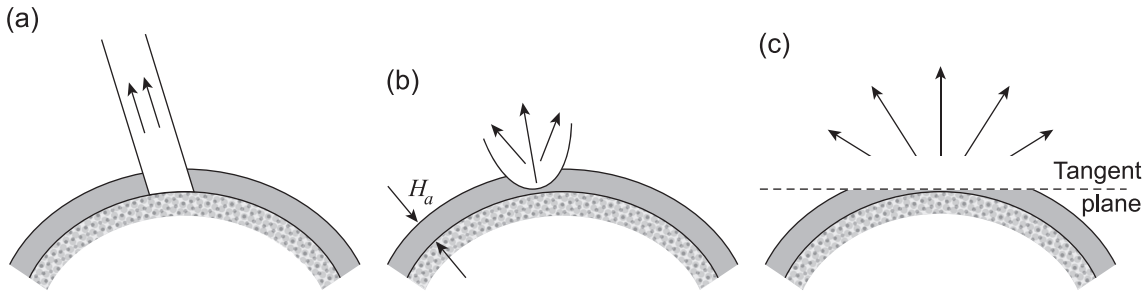

Figure 5.19 Various semi-analytical approximations to atmospheric impact erosion. (a) In the *cookie-cutter* approximation, the mass of gas that escapes is similar to that intercepted by the impactor multiplied by a factor close to unity (Walker, 1986). (b) In the *massless point explosion approximation*, an accelerating shock wave drives off the atmosphere of scale height (Ahrens, 1993). (c) In the *tangent plane model* of atmospheric impact erosion, a sufficiently massive impactor removes the entire atmosphere above a plane tangent to the planet. The impactor is required to have an impact velocity above a threshold of twice the escape velocity of the target (Melosh and Vickery, 1989).

5.12 Impact Erosion of Planetary Atmospheres

A large impact on a planet is very unlike an impact in everyday experience, such as throwing a stone into mud, because extraterrestrial impactors possess enormous kinetic energy and they vaporize in a process akin to a thermonuclear explosion. Consequently, a sufficiently large and energetic impactor can heat atmospheric gases to escape velocity while very high-speed ejecta can accelerate atmospheric gases to escape speed.

Early in a planet's history, there are many large impacts, as witnessed by the craters on Mars, the Moon, Mercury, and other bodies, so impact erosion could have been an important process for early loss of atmospheres on vulnerable bodies. Whether a body is subject to atmospheric impact erosion depends primarily on an object's escape velocity and whether it experiences a high impact velocity regime by virtue of its orbital position in a planetary system.

The velocity of an impactor depends on the escape velocity, v_e, which sets a minimum impact velocity by energy conservation for an object falling in from infinity, and a median encounter velocity, v_{enc}, which depends on the type and origin of the impactor:

$$v_{impact}^2 = v_e^2 + v_{enc}^2 \tag{5.85}$$

As a rough guideline, if the Keplerian orbital velocity of a planet around the Sun is v_{orb}, asteroids and Kuiper Belt comets tend to hit planets with moderate median encounter velocity, $v_{enc} \approx 0.5 v_{orb}$, because these intruders are prograde and in the ecliptic plane (Bottke *et al.*, 1995; Zahnle *et al.*, 1992). In contrast, Oort Cloud comets have typical $v_{enc} \approx 1.7 v_{orb}$. Thus, in general, planets close to their host stars that move at higher orbital velocities compared to those farther away tend to suffer energetic impacts (Lissauer, 2007). Of course, such planets also receive higher stellar irradiation and are more prone to thermal escape too, so their atmospheres are doubly vulnerable to escape.

Four concepts have been used to estimate the atmospheric erosion effect of impact. Walker (1986) suggested that an atmosphere between an impactor and its target is heated by multiple shocks with the net effect that an atmospheric cross-section comparable to the impactor is lost: a fraction $\sim {r_{imp}}^2 / {R_\mathrm{p}}^2$, where r_{imp} is the radius of the impactor and R_p is the planet's radius (Fig. 5.19(a)). This is a so-called *cookie-cutter approximation* to impact erosion. Second, others have suggested that impacts behave like massless point explosions (Ahrens, 1993). The shock wave from the explosion accelerates through the escape velocity as it propagates up into exponentially thinner gas (Fig. 5.19(b)). The fraction of atmosphere to escape in this approximation scales as $\sim 10 {H_a}^2 / R_\mathrm{p}^2$, where H_a is the atmospheric scale height.

A third and popular model for estimating cumulative impact erosion is that of Melosh and Vickery (1989) who deduced that escape driven by high-speed impact ejecta makes erosion more efficient and that a sufficiently large and energetic impact can erode all of the atmosphere above a plane tangent to the planet (Fig. 5.19(c)). This model is called the *tangent plane approximation*. A much larger amount of atmosphere escapes when the momentum of the ejecta is not much impeded by the inertia of the intervening atmosphere, and the ejecta move faster than the escape speed. Tangent plane erosion requires impact velocities that Melosh and Vickery (1989) estimated as $>2v_e$, cautioning that this threshold was uncertain. The

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

mass of gas to escape relative to the total atmospheric mass in a qualifying impact is $\sim H_a/2R_p$. Many models of impact erosion for Mars and icy satellites have used tangent-plane erosion. They assume that loss of atmospheric mass, m_{tangent}, above a plane tangent to the planet occurs if $v_{impact}/v_e > 2$ and the mass of the impactor, m_{impactor}, exceeds a critical mass that was originally proposed to be m_{tangent} (Melosh and Vickery, 1989) but later revised to a multiple of m_{tangent} (Pham *et al.*, 2011; Vickery and Melosh, 1990).

A fourth form of impact erosion could occur when an impactor is so big and fast that a shock wave propagates through a planet and erupts at the surface, particularly the antipode (Chen and Ahrens, 1997; Genda and Abe, 2003, 2005). The expelled solid surface can carry along much atmosphere. For example, on a planet with a deep, thick gas envelope, air above a critical isobar might be accelerated to the escape velocity, given that shock waves accelerate as they move into thinner gas. Currently, research on this mechanism is limited. However, simulations of late stage terrestrial planet accretion suggest that roughly half of the collisions between planets strip off the outer mantle of the larger planet while obliterating the smaller planet (Agnor and Asphaug, 2004), so the effect on atmospheres ought to be important.

The rate of net change of volatiles $\dot{M}_{atm}$ for a planet will be the difference between the mass rate delivered ($\dot{M}_{deliv}$) and eroded ($\dot{M}_{erode}$), i.e.,

$$\dot{M}_{atm} = \dot{M}_{deliv} - \dot{M}_{erode} \tag{5.86}$$

The delivery rate of volatiles $\dot{M}_{deliv}$, depends on the volatile mass fraction in impactors, which is typically ~0.01–0.1, and how much impactor mass escapes relative to the mass of the impactor. The erosion rate $\dot{M}_{erode}$ depends on the fraction of mass that escapes that is atmospheric gas or volatiles from impactor or target. Both $\dot{M}_{deliv}$ and $\dot{M}_{erode}$ have to be estimated from a double integral of the impactor velocity distribution over all velocities and the mass density of the flux of impactors over all masses. In turn, $\dot{M}_{atm}$ must be integrated over time. The literature gives details for such models using the tangent plane approximation (Manning *et al.*, 2006b; Pham *et al.*, 2011; Schlichting *et al.*, 2015; Zahnle *et al.*, 1992; Zahnle, 1993b).

In the Solar System, the mass density of the flux of impactors can be derived from the derivative with respect to mass of the time-dependent cumulative number flux of impactors, which varies as $\sim m^{-b}$. Parameter b is the spectral slope of the cumulative mass distribution, often estimated as ~0.5–0.8. It is likely that b is a natural outcome of collisional cascades (Dohnanyi, 1972), and so such distributions may also apply to atmospheric impact erosion for exoplanets.

The most sophisticated impact erosion models are 3-D numerical models. Hydrocodes are numerical models that deal with shock physics and solve the mass, momentum and energy conservation equations as a function of time a grid (e.g., Barr and Canup, 2010; Barr and Citron, 2011; Kraus *et al.*, 2011; Pierazzo *et al.*, 2008; Senft and Stewart, 2007, 2008, 2011). Another numerical approach, *smoothed particle hydrodynamics* (SPH), models bodies as a large number of discrete of discrete, often spherically symmetric particles that are sometimes fuzzy (i.e., with spatial kernels), whose individual dynamics and compositional identities are followed in time.

Such numerical models have been applied to impact erosion for terrestrial planets (Maindl *et al.*, 2015; Shuvalov, 2009; Shuvalov *et al.*, 2014) and Titan (Artemieva and Lunine, 2005; Korycansky and Zahnle, 2011).

5.13 Summary of the Fundamental Nature of Atmospheric Escape

In this chapter, we have discussed various mechanisms for the escape of gases from planetary atmospheres, noting how thermal escape, suprathermal (nonthermal) escape, and impact erosion are three basic categories.

Today in the Solar System, no gases attain escape velocity from the gas giants, but the rocky planets cannot hold on to light gases such as hydrogen. On Venus, Earth, Mars, and Titan, the escape of hydrogen from current atmospheres is described well by the diffusion-limited flux. In this limit, the escape rate is set by two factors: the amount of hydrogen in all its chemical forms at the homopause, and the diffusion of hydrogen above that level to the exobase, the bottom of the exosphere. For Earth, the total mixing ratio of hydrogen in all its forms in the lower stratosphere above the "cold trap" at the tropopause can be used to calculate the diffusion-limited flux because this mixing ratio is similar to that at the homopause. Jeans' escape is often a minor component of the time-average escape flux from Earth's exobase. Suprathermal escape mechanisms dominate on Earth in the time-average and contribute essentially the entire hydrogen flux from the cold upper atmosphere of Venus.

The early terrestrial planets may have had more hydrogen-rich atmospheres after they formed for ~10^7–10^8 years. Also on Earth, hydrogen-rich steam atmospheres would have formed intermittently because of large, ocean-vaporizing impacts during heavy bombardment. Under such circumstances, the absorption of higher

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

ultraviolet radiation flux from the youthful Sun likely drove a bulk hydrodynamic outflow of hydrogen-rich upper atmospheres.

In hydrodynamic escape, heavy atoms can be dragged along when collisions with hydrogen push the heavy atoms upward faster than gravity pulls them downward. This can lead to loss of heavy gases and mass fractionation of different isotopes of noble gases. The isotopic patterns of noble gases on Earth and Mars are consistent with fractionation by early hydrodynamic hydrogen escape. It is possible that major gases, such as carbon dioxide and nitrogen, were also lost from Mars during hydrodynamic escape.

Looking ahead to later chapters, we will see that hydrogen escape has had an important influence on the chemical evolution of the atmospheres and surfaces of Venus and Mars. In Ch. 10, we will see that hydrogen escape may also have affected the evolution of the oxidation state of the Earth's atmosphere and, as a consequence, biological evolution.

Finally, impact erosion can be effective early in a planetary system's history, particularly for small bodies. In our own Solar System, apart from Mercury, Mars was the planet most prone to impact erosion. Moons around Jupiter were also vulnerable. For exoplanets, small bodies close to parent stars will be vulnerable because their large orbital velocities imply bigger impact speeds. In conclusion, escape processes are fundamental for understanding the existence and evolution of planetary atmospheres.

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006

Downloaded from https://www.cambridge.org/core. University of Chicago, on 19 Apr 2018 at 03:28:44, subject to the Cambridge Core terms of use, available at https://www.cambridge.org/core/terms. https://doi.org/10.1017/9781139020558.006