

# Atmospheric Composition and Vertical Structure

Thomas W. Schlatter

Retired Meteorologist  
Earth System Research Laboratory  
National Oceanic and Atmospheric Administration  
Boulder, CO, USA

**Keywords:** All the following words should be preceded by “atmospheric”: structure, composition, gases, properties, layers, hazards to aircraft and spacecraft. More keywords: troposphere, stratosphere, mesosphere, thermosphere, ionosphere

## Abstract

This article considers the gases constituting Earth’s atmosphere and the vertical structure of the atmosphere to roughly 1,000-km altitude. Turbulent mixing keeps the relative concentrations of gases nearly constant in the lowest 100 km. At higher altitudes, molecular diffusion controls the concentrations, with the lighter gases becoming relatively more abundant with increasing altitude. The atmosphere exerts drag on vehicles passing through it, by means of skin friction and larger-scale pressure forces controlled by vehicle shape. High-energy solar radiation dissociates atmospheric molecules and ionizes atoms and molecules at high altitudes. Ions and free electrons in the upper atmosphere interact with the Earth’s magnetic field. Solar and magnetic storms inject high-energy particles into the upper atmosphere that can be hazardous to humans, spacecraft, and the electrical components inside.

## 1. GUIDE TO THIS ARTICLE

This article reviews the composition and vertical structure of the atmosphere, the fluid medium in which aircraft and spacecraft fly. The upper boundary of the atmosphere is indefinite. The charged and electrically neutral particles comprising the very high atmosphere eventually merge with the solar wind, a stream of ionized gas flowing continually outward from the sun. That occurs at altitudes (tens of thousands of kilometers) much greater than considered here. To keep the scope of this article manageable, atmospheric structure is described from the surface of the Earth to roughly 1000-km altitude. Only the vertical dimension is considered, i.e., horizontal gradients of atmospheric properties and major atmospheric wind features are not discussed.

- Section 2 describes the most basic properties of atmospheric gases and the conditions under which they may be considered fluids.
- Section 3 explains why the mean molecular weight of air, a mixture of gases, is essentially constant up to an altitude of 85 km, why the mean molecular weight decreases above that altitude, and why charged particles become part of the atmospheric mix at high altitudes.
- Section 4 discusses the principal gases in the atmosphere and, where pertinent, their sources and sinks (the processes that control the addition of a gas to the atmosphere and its removal).

- Section 5 concentrates on the vertical temperature structure of the atmosphere and how it defines the various concentric shells comprising the atmosphere.
- Section 6 briefly discusses the ionosphere, where electrically charged particles exist in abundance, interact with neutral particles, and respond to the Earth’s magnetic field.
- Section 7 considers momentum transfer on the molecular scale, which relates directly to effects the atmosphere has on vehicles passing through it at high velocity, and turbulence, which can affect the attitude and stability of vehicles in flight.
- Section 8 is a brief summary of the hazards to air and space travel posed by the atmosphere.

Space does not permit a treatment of any of these topics in detail, but the intent is to provide the fundamental concepts without sacrificing scientific accuracy.

## 2. FUNDAMENTAL PROPERTIES OF ATMOSPHERIC GASES

Consider a single component gas of the Earth’s atmosphere. The most primitive form of the *ideal gas law* relates the number of molecules  $N$  of this gas contained in a volume  $V$  to the pressure  $p$  and absolute temperature  $T$  of the gas.

$$pV = NkT \quad (1)$$

$k$  is the Boltzmann constant; it has the value  $1.38 \times 10^{-23} \text{ J K}^{-1}$  (joules per degree Kelvin). If the molecular mass of this gas is  $m$ , the gas law may be rewritten as

$$p = m \frac{N}{V} \frac{k}{m} T = \rho RT . \quad (2)$$

The density  $\rho$  is the number of molecules times the mass of each divided by the volume occupied by the gas. The gas constant  $R = k/m$  is specific to the type of gas. A critical aspect of the ideal gas law is that absolute temperature is defined by

$$\frac{1}{2} \langle mv^2 \rangle = \frac{3}{2} kT , \quad (3)$$

where the brackets  $\langle \rangle$  imply an average,  $v$  is the molecular speed, and the left-hand side is the average kinetic energy of the molecules. In the lower atmosphere, individual molecules are in frenzied motion, traveling at hundreds of meters per second and colliding with each other at astronomically high rates within volumes smaller than a cubic centimeter. It is assumed that the individual molecules undergo elastic collisions and do not otherwise interact with one another. In any thermally insulated gas where speeds are high and collisions incessant, the temperature will quickly come to an equilibrium value and the speed distribution will approach the so-called Maxwell-Boltzmann speed distribution

$$f(v) = 4\pi \left( \frac{m}{2\pi kT} \right)^{3/2} v^2 \exp\left( -\frac{mv^2}{2kT} \right) , \quad (4)$$

where  $\pi$  is the constant 3.1416. Equation (4), whose derivation was a major accomplishment in gas kinetic theory, gives the distribution of molecular speeds for a given temperature and a given

gas. This distribution is valid at altitudes up to about 600 km for neutral gases. Above that altitude, the number density of neutral gases becomes quite low, and collisions become infrequent. On the other hand, the ratio of electrically charged particles to neutral particles increases with altitude, and the Coulomb forces between the charged particles (forces of electrical attraction and repulsion) cause their velocities to follow the Maxwell-Boltzmann distribution to altitudes much greater than 1000 km. Thus, when it is time to discuss temperature at altitudes between 600 and 1000 km, the speeds of ions and electrons, which can be measured, will determine the temperatures of each, and the two differ because their speeds differ.

Because it is a distribution function,  $f(v)$  has the property that

$$\int_0^{\infty} f(v)dv = 1 .$$

The mean speed is

$$\langle v \rangle = \int_0^{\infty} vf(v)dv = \left( \frac{8kT}{\pi m} \right)^{\frac{1}{2}} . \quad (5)$$

The root mean square of the speed  $v_{rms}$  is

$$v_{rms} = \sqrt{\langle v^2 \rangle} = \sqrt{\frac{3kT}{m}} \quad \text{where} \quad \langle v^2 \rangle = \int_0^{\infty} v^2 f(v)dv , \quad (6)$$

which is fully consistent with Equation (3).

One can also compute the most probable speed  $v_p$ , which makes  $f(v)$  a maximum, by setting the first derivative of  $f$  equal to zero. The result is

$$v_p = \left( \frac{2kT}{m} \right)^{\frac{1}{2}} . \quad (7)$$

Consider a column of gas of a specific type acted upon by gravity alone and not subject to vertical accelerations anywhere in the column. Such a gas is said to be in hydrostatic equilibrium: At any level in the column, the pressure (force per unit area exerted by the gas) is determined solely by the weight of the overlying gas in the column. The change in pressure with altitude is given by

$$\frac{\partial p}{\partial z} = -\rho g , \quad (8)$$

where  $g$  is the acceleration due to the Earth's gravity. At the Earth's surface,  $g$  is approximately  $9.807 \text{ m s}^{-2}$ . With the help of the ideal gas law, one can rewrite Equation (8) as

$$\frac{1}{p} \frac{\partial p}{\partial z} = -\frac{g}{RT}.$$

This expression can be integrated from sea level ( $z = 0$ ) to some higher altitude  $z$ , or equivalently from the pressure at sea level  $p_0$  to the pressure  $p$  at altitude  $z$ . The result is

$$p = p_0 \exp\left(-\int_0^z \frac{g}{RT} dz\right). \quad (9)$$

For a specific gas,  $R$  is constant with altitude,  $g$  decreases slowly with altitude, and  $T$  may increase or decrease. For determining pressure at various altitudes in a standard atmosphere (see section 5.5), Equation (9) may be integrated numerically. For now, it is useful to define a scale height  $H$  as

$$H = \frac{RT}{g} \quad (10)$$

$H$  has dimensions of meters, as it must. The physical significance of  $H$  is that, if it may be considered approximately constant, the pressure of a gas will decrease by the factor  $1/e$  as altitude increases by the scale height. Note that  $H$  depends upon the type of gas under consideration, in particular, upon its molecular mass (because  $R = k/m$ ). The implications of this are discussed in section 3.

Up to this point, discussion has centered on a specific gas, but the atmosphere is a mixture of gases, to be discussed specifically in section 4. Dalton's Law of Partial Pressures states that each component gas in a mixture exerts its own pressure as if it were alone in the volume containing the mixture at temperature  $T$ . The total pressure  $p$  exerted by all gases in a mixture is just the sum of the partial pressures  $p_i$ .

$$p = \sum_i p_i \quad (11)$$

The total density of a mixture of gases depends straightforwardly upon the number of molecules of each type  $n_i$  in a volume ( $n_i = N_i/V$  in the notation of Equation 1, which is the number density) and their molecular masses  $m_i$ .

$$\rho = \sum_i n_i m_i. \quad (12)$$

Applying the ideal gas law to the partial pressures in Equation (11), one obtains

$$p = kT \sum_i n_i. \quad (13)$$

$T$  has no subscript because the definition of temperature pertains to the average kinetic energy of all molecules in the gaseous mixture. Next, divide Equation (13) by Equation (12) to obtain

$$\frac{p}{\rho} = \frac{kT \sum_i n_i}{\sum_i n_i m_i}. \quad (14)$$

The mean molecular mass  $\langle m \rangle$ , weighted by the individual number densities, is

$$\langle m \rangle = \frac{\sum_i n_i m_i}{\sum_i n_i}.$$

With this definition, Equation (14) becomes

$$p = \frac{\rho k T}{\langle m \rangle}.$$

This restatement of the gas law applies to a mixture of gases with total pressure  $p$ , total density  $\rho$ , and temperature  $T$  defined as the average kinetic energy of molecules in the mixture of gases. It is often inconvenient to deal with the mass of individual molecules, so both  $k$  and  $m$  are scaled by Avogadro's number. Avogadro's Number  $A$  is the number of molecules in one mole of a substance. A mole is the mass  $M$  in grams of a substance that is numerically the same as the molecular weight of that substance.  $A = 6.022 \times 10^{23} \text{ mole}^{-1}$ . Thus,  $6.022 \times 10^{23}$  molecules of a gas has a mass of  $M$  grams. For oxygen ( $\text{O}_2$ ),  $M$  is very nearly 32 g. (Any departure from 32 occurs because of naturally occurring isotopes of oxygen.) With the definitions  $R^* = Ak$  and  $M = Am$ , the above equation becomes

$$p = \frac{\rho R^* T}{\langle M \rangle}. \quad (15)$$

$R^*$  is called the universal gas constant; it has the value  $8.314 \text{ J (gram-mole)}^{-1} \text{ K}^{-1}$ . The safe course of action is to use MKS units for all calculations; thus, if  $M$  is expressed in kilograms,  $R^*$  must be expressed as  $8314 \text{ J (kilogram-mole)}^{-1} \text{ K}^{-1}$ . A kilogram-mole of dry air at sea level has a mass  $M_d$  of 28.96 kg. Thus, one can define a gas constant for dry air as  $R_d = R^* / \langle M_d \rangle$ , and then the gas law for dry air may be written

$$p_d = \rho R_d T. \quad (16)$$

The numerical value of  $R_d$  is  $287 \text{ J kg}^{-1} \text{ K}^{-1}$ .

Standard atmospheres, to be discussed in section 5.5, assume dry air. The main reason is that the concentration of water vapor, the third most abundant atmospheric gas, is highly variable, mainly

because the maximum possible vapor pressure is a strong function of temperature. Once the air is saturated with water vapor, any reduction in temperature will cause condensation. No other constituent gas changes phase under normal atmospheric conditions. Water vapor is responsible for most of the weather people experience. Condensed into clouds and precipitation, water in liquid or ice phases can pose serious hazards to aircraft, but that is discussed in a companion article on meteorology in this volume.

A very good reference for the ideas discussed in this section is Bohren and Albrecht (1998, pp. 51-78).

### 3. THE HOMOSPHERE, HETEROSPHERE, AND EXOSPHERE

One might expect the hydrostatic relationship (Equation 9) to be valid separately for each gas comprising air. As will be noted, this is not generally true. With the subscript  $i$  denoting a particular gas,

$$p_i(z) = p_i(0) \exp\left(-\int_0^z \frac{g}{R_i T} dz\right) = p_i(0) \exp\left(-\int_0^z \frac{m_i g}{kT} dz\right). \quad (17)$$

Suppose that the molecules of gas 2 in a mixture are more massive than molecules of gas 1, i.e.,  $m_2 > m_1$ . Use the above equation to find the ratio of  $p_1(z)/p_2(z)$ . The result is

$$\frac{p_1(z)}{p_2(z)} = \frac{p_1(0)}{p_2(0)} \exp\left[-\frac{1}{k} \int_0^z \frac{g(m_2 - m_1)}{T} dz\right].$$

Because the integrand is positive, the argument of the exponent is negative, and so

$$\frac{p_1(z)}{p_2(z)} > \frac{p_1(0)}{p_2(0)} .$$

This relationship of the partial pressures of two gases at different altitudes says that the relative abundance of the “lighter” gas increases with altitude. This property is *not* observed below about 100-km altitude because of turbulence, the bulk motion of large volumes of air that effectively mixes the component gases of air so that their relative concentrations do not vary from the surface to about 100 km. For this reason the molecular weight of dry air does not vary in this altitude range, nor does the gas constant for dry air. That part of the atmosphere in which the gaseous constituents are mixed in constant ratios is called the *homosphere*. Equation (9) is valid in the homosphere for the mixture of gases in dry air but not for the individual constituent gases.

For  $z_{100} \approx 100$  km and  $z > z_{100}$

$$\frac{p_1(z)}{p_2(z)} > \frac{p_1(z_{100})}{p_2(z_{100})}, \quad (18)$$

Equation (18) is true. A transition region exists between roughly 80- and 120-km altitude in which bulk mixing of gases by turbulence gives way to molecular diffusion: a mixing process caused by molecules colliding with each other. Trails of rocket exhaust in this altitude range show a marked change in character. Up to 80 km, turbulence usually contorts the trails markedly within minutes. Above 120 km, the trails are much smoother. The region of the atmosphere in which molecular diffusion dominates mixing is also where gravitational settling results in the concentration of “lighter” gases growing relative to that of the heavier gases as altitude increases. This region is called the *heterosphere*. A good approximation to the number density of the  $i^{\text{th}}$  gas above 100 km is obtained from Equation (17), Equation (1) expressed as  $p_i = n_i kT$  and Equation (10) expressed as  $H_i = RT/g$ :

$$n_i(z) = n_i(z_{100}) \frac{T(z_{100})}{T(z)} \exp\left(-\int_{z_{100}}^z \frac{dz}{H_i}\right) \text{ for } z > 100 \text{ km.} \quad (19)$$

At this point, it is instructive to examine a table of data extracted from *The U.S. Standard Atmosphere, 1976*. The parameters to be examined are a function of geometric altitude, and they pertain to electrically neutral particles:

- Acceleration due to the Earth’s gravity:  $g(z) = g_0 \left(\frac{r_0}{r_0 + z}\right)^2$ , where  $g(z)$  is the acceleration at geometric altitude  $z$ ,  $g_0$  is the acceleration at sea level ( $9.807 \text{ m s}^{-2}$ ), and  $r_0$  is the radius of the Earth ( $6.357 \times 10^6 \text{ m}$ ).
- Number density of air molecules,  $n$ : Below 85 km,  $n$  is obtained from Equation (1) expressed as  $p = nkT$ . Above 85 km,  $n = \sum n_i$ , where the  $n_i$  are computed from an equation having the form of (19), but more complicated.

- Mean particle speed: computed from Equation (5) expressed as:  $\langle v \rangle = \left(\frac{8R^*T}{\pi \langle M \rangle}\right)^{1/2}$

- Mean collision frequency,  $\nu$ :  $\nu = 4A\sigma^2 p \left(\frac{\pi}{\langle M \rangle R^*T}\right)^{1/2}$ , where  $A$  is Avogadro’s

number, as before, and  $\sigma$  is called the effective collision diameter. To explain, if the distance between the centers of two molecules in random motion remains greater than  $\sigma$ , a collision will not occur.  $\sigma = 3.65 \times 10^{-10} \text{ m}$ .  $\nu$  gives the number of collisions per second experienced by a single molecule with its neighbors. To find the number of total collisions per second among all molecules within a unit volume, multiply  $\nu$  by  $n$ , the number density.

- The mean free path  $L$  is the mean distance traveled by a neutral particle between successive collisions with other particles. The volume in which mean free path is measured must be large enough to contain a very large number of particles.

$$L = \frac{\langle v \rangle}{\nu} = \frac{2^{1/2}}{2\pi\sigma^2 n}$$

- The mean molecular weight of air  $\langle M \rangle$

Table 1. Some properties of neutral atmospheric gas as a function of geometric altitude. A two-digit number preceded by the plus or minus sign indicates the power of ten by which the preceding number is to be multiplied. Data from the *U.S. Standard Atmosphere, 1976* for dry air.

Geometric Altitude	Acceleration due to gravity	Number density	Mean molecular speed	Collision frequency	Mean free path	Mean molecular weight	Scale Height
$z$ (km)	$g$ (m s <sup>-2</sup> )	$n$ (m <sup>-3</sup> )	$\langle v \rangle$ (m s <sup>-1</sup> )	$\nu$ (s <sup>-1</sup> )	$L$ (m)	$\langle M \rangle$ (kg kmol <sup>-1</sup> )	$H$ (km)
0	9.807	2.547 +25	458.9	6.919 +09	6.633 -08	28.96	8.434
20	9.745	1.849 +24	398.0	4.354 +08	9.139 -07	28.96	6.382
40	9.684	8.308 +22	427.8	2.104 +07	2.034 -05	28.96	7.421
70	9.594	1.722 +21	400.6	4.084 +05	9.810 -04	28.96	6.570
100	9.505	1.189 +19	381.4	2.68 +03	1.42 -01	28.40	6.009
150	9.360	5.186 +16	746.5	2.3 +01	3.3 +01	24.10	23.380
200	9.218	7.182 +15	921.6	3.9 +00	2.4 +02	21.30	36.183
300	8.943	6.509 +14	1,080	4.2 -01	2.6 +03	17.73	51.193
400	8.680	1.056 +14	1,149	7.2 -02	1.6 +04	15.98	59.678
600	8.188	5.950 +12	1,356	4.8 -03	2.8 +05	11.51	88.244
800	7.737	1.234 +12	1,954	1.4 -03	1.4 +06	5.54	193.86
1,000	7.322	5.442 +11	2,318	7.5 -04	3.1 +06	3.94	288.20

As indicated by Equation (3), the mean molecular speed defines temperature. The wide variation in  $\langle v \rangle$  with altitude mirrors corresponding wide excursions in temperature, as will be discussed in section 5. As noted earlier, the mean molecular weight remains constant until near 100-km altitude because of turbulent mixing. Above 100 km, however, molecular diffusion increasingly dominates the mixing process, and air flows become nearly laminar. Gravitational settling predominates above 120 km so that the relative concentration of the heavier gases shrinks with increasing altitude, and thus so does the mean molecular weight..

Even at sea level, air molecules occupy only about one thousandth of the volume containing them. Yet their high number density and furious agitation result in mean free paths less than a micrometer and billions of collisions per second for each molecule. The short mean free path and the astronomical number of collisions per second in the homosphere are the fundamental reasons why large volumes of air on the scale of centimeters to many meters can move

cohesively (turbulence). At an altitude of 100 km, the fuzzy boundary between the homosphere and heterosphere, the mean free path is close to 10 cm, and the collision rate is reduced to a few thousand per second. Above this level, turbulence is no longer effective, and the diffusion of molecules becomes the dominant process in transferring heat and momentum.

At 500 km, the mean free path (not listed in the table above) is  $7.70 \times 10^4$  m. Molecules that travel 77 km without colliding with a neighbor are effectively on ballistic trajectories, that is, they are projectiles acted upon only by the force of gravity. The top of the heterosphere is often defined as that altitude where the scale height is equal to the mean free path. At 500 km,  $\langle M \rangle$  is 14.33 kg,  $g$  is  $8.43 \text{ m s}^{-2}$ , and the temperature is  $\sim 1000$  K. With these values, the scale height (Equation 9) expressed as  $H = R^*T/(\langle M \rangle g)$  is  $6.88 \times 10^4$  m. 69 km is not far different from 77 km; in fact, in the *U.S. Standard Atmosphere, 1976*,  $H = L$  at 492 km altitude. That portion of the atmosphere above about 500 km is called the *exosphere*, because molecules on an upward trajectory, provided their speeds are sufficiently high, have a chance to escape the Earth's atmosphere without colliding with another molecule.

A molecule traveling straight up at the base of the exosphere (altitude  $z_e$ ) attains the *escape speed*  $v_e$  if its kinetic energy is sufficient to overcome the potential energy acquired between  $z_e$  and outer space.

$$\frac{1}{2}mv_e^2 = \int_{z_e}^{\infty} mgdz = \int_{z_e}^{\infty} mg_0 \left( \frac{r_0}{r_0 + z} \right)^2 dz$$

After integration, the escape speed is found to be

$$v_e = \left( \frac{2g_0r_0}{1 + z_e/r_0} \right)^{\frac{1}{2}}. \quad (20)$$

With  $z_e = 500$  km and the values of other constants given earlier, the escape speed is calculated to be about  $10.8 \text{ km s}^{-1}$ . This result is independent of the type of molecule. It is also considerably higher than the mean molecular velocities listed in the Table 1.

The Maxwell-Boltzmann distribution of molecular *speeds* (Equation 4) is derived from the more fundamental Maxwell-Boltzmann *velocity* distribution:

$$F(\vec{v}) = \left( \frac{m}{2\pi kT} \right)^{\frac{3}{2}} \exp \left( -\frac{m(\vec{v})^2}{2kT} \right). \quad (21)$$

The distinction between  $f(v)$  and  $F(\vec{v})$  is discussed in Schunk and Nagy (2000, pp. 513-517). Think of all possible velocity vectors with their origin at the center of a sphere, their magnitude  $v$  representing the radius, and their direction given by the colatitude  $\theta$  and the longitude  $\phi$ . The

vertical direction (against gravity) is at  $\theta = 0$ . An element of volume in this space is  $dV = (v \sin \theta d\phi)(v d\theta)dv$ . If one is interested in the distribution function for speed alone, then  $v$  is constant and it is appropriate to integrate over all directions

$$f(v) = \int_0^{2\pi} \int_{-\pi/2}^{\pi/2} F(\vec{v})(v d\theta)(v \sin \theta d\phi) = 4\pi v^2 F(\vec{v}), \quad (22)$$

which is recognizable as Equation (4).

The distribution  $F(\vec{v})$  is useful in calculating how fast the atmosphere is losing its gases due to random thermal motion. More precisely, one wants to know the vertical flux  $\Phi_e$  of gas atoms or molecules having the escape speed or greater at the base of the exosphere. We are interested only in the spherical shell beyond  $v = v_e$  and in  $0 \leq \theta \leq \pi/2$ , because this range of values includes vectors with an outbound component and excludes those with a downward component. The appropriate multiple integral is similar to that in Equation (22) except for the addition of integration over  $v$ :

$$\Phi_e = \int_0^{2\pi} \int_0^{\pi/2} \int_{v_e}^{\infty} (v \cos \theta) n_e F(\vec{v})(v d\theta)(v \sin \theta d\phi) dv \quad (23)$$

This integral was originally proposed by James H. Jeans in 1916.  $n_e$ , a constant, is the number density of the gas in question at the base of the exosphere (500 km altitude).  $n_e F(\vec{v})dv$  gives the number density of particles traveling between speeds  $v$  and  $v + dv$ .  $v \cos \theta$  is the vertical component of particle velocity, where, again,  $\theta$  is the angle that the velocity vector makes with local vertical. The product of  $v \cos \theta$  and  $n_e F(\vec{v})dv$  gives a vertical particle flux. Substitute Equation (7) for the most probable velocity into Equation (21) to obtain

$$F(v) = \frac{1}{\pi^{3/2}} \frac{1}{v_p^3} \exp\left(-\frac{v^2}{v_p^2}\right). \quad (24)$$

Then use Equation (24) in Equation (23).

$$\Phi_e = \int_0^{2\pi} d\phi \int_0^{\pi/2} \sin \theta \cos \theta d\theta \left[ \frac{1}{\pi^{3/2}} n_e \int_{v_e}^{\infty} \frac{v^3}{v_p^3} \exp\left(-\frac{v^2}{v_p^2}\right) dv \right]$$

With the change in variable  $x = v/v_p$ , the above equation becomes

$$\Phi_e = \frac{n_e v_p}{\pi^{3/2}} \int_0^{2\pi} d\phi \int_0^{\pi/2} \sin \theta \cos \theta d\theta \left[ \int_{v_e/v_p}^{\infty} x^3 e^{-x^2} dx \right].$$

The integral in square brackets yields to integration by parts. The final result is

$$\Phi_e = \frac{n_e v_p}{2\pi^{1/2}} \left( 1 + \frac{v_e^2}{v_p^2} \right) \exp\left( -\frac{v_e^2}{v_p^2} \right). \quad (25)$$

This is called the Jeans thermal escape flux. We will calculate the thermal escape flux for a relatively heavy gas, nitrogen, and the lightest of all neutral gases, atomic hydrogen.

Nitrogen, the most common gas in the atmosphere, has a molecular weight  $\langle M \rangle$  of about 28 kg kmol<sup>-1</sup>. Using this value in Equation (7) expressed as  $v_p = (2R^*T / \langle M \rangle)^{1/2}$  and  $T \approx 1000$  K at the base of the exosphere, one finds that the most probable velocity is about 0.77 km s<sup>-1</sup>. In Equation (25),  $v_e / v_p$  is then 10.8/0.77 = 14.0, and the number of escaping nitrogen molecules is about  $n_e \times 3.2 \times 10^{-81}$  m<sup>-2</sup> s<sup>-1</sup>. Even if we use the sea level number density of all molecules from Table 1 ( $2.5 \times 10^{25}$ ) and multiply the escape flux times the surface area of a sphere of radius  $r_0 + 500$  km ( $5.9 \times 10^{14}$  m<sup>2</sup>), the number of nitrogen molecules departing the entire atmosphere per second is still vanishingly small:  $1.5 \times 10^{-41}$ . We conclude that gravity keeps a tight lid on the escape of the heavier atmospheric gases, oxygen included.

Hydrogen is present only in trace amounts in the lower atmosphere, but its relative abundance among atmospheric gases becomes substantial above 500-km altitude because high-energy photons from the sun dissociate water vapor, methane molecules, and molecular hydrogen to produce atomic hydrogen. The probability that a hydrogen atom (molecular weight 1.0 kg kmol<sup>-1</sup>) reaches the escape velocity is calculated in the same way as for the nitrogen molecule. The most probable velocity is 4.08 km s<sup>-1</sup>, and  $v_e / v_p$  is 2.65 in Equation (25). With this information, we find that the flux of escaping hydrogen molecules is about  $5.3n_e$ . From the *U.S. Standard Atmosphere, 1976*, one finds that the number density of hydrogen atoms at 500 km altitude is  $8.0 \times 10^{10}$  m<sup>-3</sup>; thus the thermal escape flux of hydrogen atoms is roughly  $4.2 \times 10^{11}$  m<sup>-2</sup> s<sup>-1</sup>. The size of this flux is sensitive to temperature, here assumed to be 1000K. During solar storms, to be discussed later, the temperature could be substantially higher and the escape flux much larger.

If there were no source of atomic hydrogen in the atmosphere, how long would it take for this gas to escape entirely? A rough estimate comes from the fact that the average number density of atomic hydrogen from 200- to 1000-km altitude, where much of it is concentrated, is  $8.2 \times 10^{11}$  m<sup>-3</sup> (*U.S. Standard Atmosphere, 1976*). The time for depletion is thus approximately the number in a square meter column 800 km high ( $6.6 \times 10^{17}$  m<sup>-2</sup>) divided by the escape flux above. The answer is  $1.9 \times 10^6$  s or about 22 days. Evidently, the thermal escape of atomic hydrogen from the atmosphere is significant, which explains why so little of it is left in the atmosphere and most of that is in the heterosphere, where production occurs in the presence of sunlight. Thermal escape of helium ( $\langle M \rangle = 4$  kg kmol<sup>-1</sup>,  $v_p = 2.04$  km s<sup>-1</sup>) is orders of magnitude less than for hydrogen but still significant.

Figure 1 illustrates the three atmospheric layers described in this section, the homosphere, the heterosphere, and the exosphere. A few molecular trajectories are shown in the exosphere.

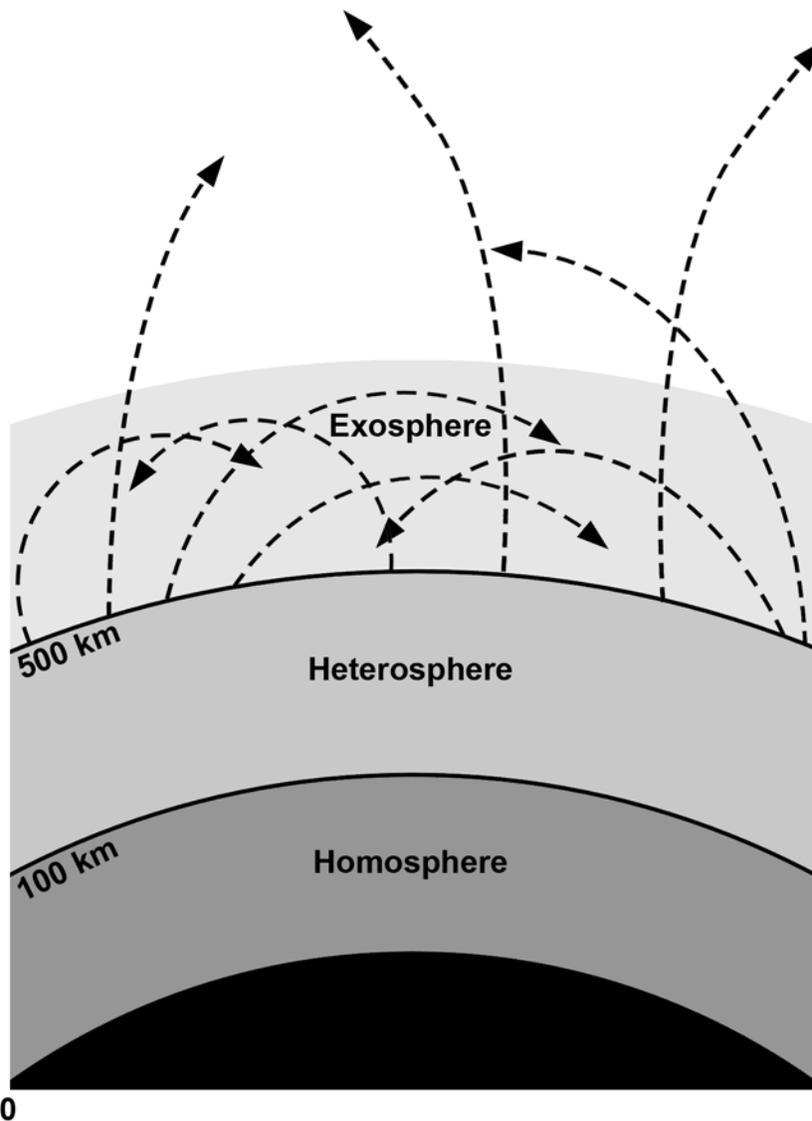


Figure 1. A cross section through the homosphere, heterosphere, and exosphere. From Salby (1996, p. 13).

From Fig. 1.5, p. 13 of Salby, 1996. HomoHeteroExo.tif

Before moving on to other subjects, this is an appropriate place to note two exceptions to previous statements. Already noted is that water vapor is not uniformly mixed in the homosphere. Another gas not uniformly mixed there is ozone, whose formation and destruction have to do with sunlight and chemical reactions in the atmosphere between roughly 15- and 50-km altitude. Second, because of energetic radiation from the sun in the high atmosphere, some

gases with very low concentrations in the lower atmosphere have appreciable concentrations higher up, notably, atomic hydrogen, already mentioned, atomic oxygen, and ozone. The principle atmospheric gases will be described in the next section.

So far, the discussion has dealt exclusively with electrically neutral gases, but not all the atmosphere is electrically neutral. Because of energetic radiation from the sun absorbed in the high atmosphere, charged particles abound, and this has far-reaching consequences that will be discussed in section 6 on the ionosphere.

#### **4. THE PRINCIPAL ATMOSPHERIC GASES**

A few atmospheric gases have already been mentioned in passing; here we shall discuss many important neutral atmospheric gases, their concentrations, sources and sinks, and longevity in the atmosphere. Table 2 lists the atmospheric gases in the order of concentration in parts per million by volume of dry air. For an ideal gas, parts per million by volume (ppmv) is equivalent to mole fraction, which is the same as fraction of total molecules. Another way of expressing it is the number density of a specific gas divided by the number density of all gases in the mixture (here, dry air). If one were to add up all the ppmv values in the table, the result would be slightly more than 1,000,000 primarily because the concentrations of carbon dioxide and methane have been growing, and the ppmv values for other less variable gases have not been adjusted for this. The mean molecular weight is not an integer because of the natural occurrence of isotopes, atoms with the same number of protons but differing numbers of neutrons.

Table 2. Concentration of gases constituting dry air in parts per million by volume (ppmv). Concentration pertains to near sea level. Gases are listed in order of decreasing concentration.

Gas	Chemical Symbol	Mean Molecular Weight ( $\text{m mol}^{-1}$ )	Concentration (ppmv)
Nitrogen	$\text{N}_2$	28.013	780,840
Oxygen	$\text{O}_2$	31.999	209,460
Argon	Ar	39.948	9,340
Carbon dioxide	$\text{CO}_2$	44.010	384
Neon	Ne	20.180	18.18
Helium	He	4.003	5.24
Methane	$\text{CH}_4$	16.043	1.774
Krypton	Kr	83.798	1.14
Hydrogen	$\text{H}_2$	2.016	0.56
Nitrous oxide	$\text{N}_2\text{O}$	44.012	0.320
Xenon	Xe	131.293	0.09
Ozone	$\text{O}_3$	47.998	0.01-0.10

#### 4.1 Nitrogen

As Table 2 shows, gaseous nitrogen in the atmosphere is almost exclusively in the form of  $\text{N}_2$ ; there are nearly  $3.9 \times 10^{18}$  kg of it in the atmosphere. The concentration of the next most plentiful nitrogen-containing gas, nitrous oxide, is more than six orders of magnitude less. Molecular nitrogen is very stable. Its sources and sinks are thought to be in balance. The annual flux of  $\text{N}_2$  into and out of the atmosphere is on the order of  $10^{11}$  kg. The residence time of  $\text{N}_2$  in the atmosphere, the time for all of it to be depleted if there are no sources to replenish it, is therefore more than  $10^7$  years.

Nitrogen fixation occurs whenever atmospheric  $\text{N}_2$  reacts to form any other nitrogen compound. The greatest loss of  $\text{N}_2$  occurs when specialized nitrogen-fixing bacteria, living in soil and water,

incorporate molecular nitrogen into biomass. This is important for plant growth. The manufacture of fertilizers and explosives, and the combustion of fossil fuels in power plants and road vehicles are two anthropogenic causes of nitrogen fixation. Finally, lightning removes small amounts of N<sub>2</sub> from the atmosphere.

Denitrification, the reduction of nitrates back to gaseous nitrogen, primarily N<sub>2</sub>, completes what is called the nitrogen cycle. This is carried out mostly by bacteria in oxygen-depleted environments. Biomass burning also returns nitrogen to the atmosphere, as does the use of nitrogen fertilizers.

## 4.2 Oxygen

Today's atmosphere contains about  $1.2 \times 10^{18}$  kg of O<sub>2</sub>, a level probably reached not later than 25 million years ago. The main source of atmospheric oxygen is photosynthesis, accomplished by plants on land and phytoplankton, a kind of algae that live near the surface of the ocean. In the presence of sunlight, gaseous carbon dioxide and liquid water react chemically to form carbohydrate, a solid, and oxygen gas. The reaction has the form



A secondary, but much smaller (at the 1-2% level) source of O<sub>2</sub> is the photodissociation of water vapor and nitrous oxide (N<sub>2</sub>O) by high energy ultraviolet radiation from the sun. Photodissociation of O<sub>2</sub> in the high atmosphere yields significant amounts of atomic oxygen (O). In fact, the number density of O exceeds that of O<sub>2</sub> above 120-km altitude.

Respiration and decay, in which animal life and bacteria consume oxygen and release carbon dioxide, are responsible for over 93% of O<sub>2</sub> loss from the atmosphere. Consumption of biomass fuels and biomass burning are responsible for another 4%.

The residence time of O<sub>2</sub> in the atmosphere is roughly 5000 years.

## 4.3 Water Vapor

Because the concentration of water vapor in the atmosphere is temperature limited, and winds can easily transport vapor thousands of kilometers, this gas is highly variable in space and time. Its concentration is 0-4% by volume. Almost all water vapor in the atmosphere is confined to the troposphere, where clouds and storms occur. Low temperatures at the top of the troposphere (-50 to -70°C) assure that condensation will remove all but trace amounts of vapor before it can reach the stratosphere, the layer of atmosphere immediately above the troposphere.

The principal source of water vapor is evaporation from the oceans, mostly in the tropics where the temperature is relatively high. Evaporation from lakes and soils, and transpiration from plants are other important sources of vapor. Precipitation removes water vapor from the atmosphere.

Water vapor is a naturally occurring greenhouse gas, that is, it absorbs little incoming solar radiation (visible wavelengths), but absorbs significant outgoing longwave radiation (infrared wavelengths).

Oceans contain about  $1.35 \times 10^9 \text{ km}^3$  of water, about 97% of all the Earth's water. In contrast, the atmosphere contains only  $1.3 \times 10^4 \text{ km}^3$  of liquid equivalent water. The volume of global daily precipitation is roughly  $1,400 \text{ km}^3$ , leading to the estimate that the residence time of water vapor in the atmosphere is between nine and ten days.

#### 4.4 Carbon Dioxide

At a 2008 concentration of 384 ppmv and growing,  $\text{CO}_2$  can be considered a trace gas, yet it is important for at least three reasons: 1) it is involved in photosynthesis, 2) it is a potent greenhouse gas, and 3) it is a major player in the carbon cycle.

The major reservoirs of carbon are in the atmosphere (750 GtC: 1 GtC is  $10^9$  metric tons of carbon); the land surface, in soils and vegetation, alive and decaying (2,200 GtC); and the ocean (40,000 GtC). The net annual exchanges of carbon per year among these reservoirs are quite small compared to the stored carbon, and so there are substantial uncertainties in the carbon budget. The main sources of  $\text{CO}_2$  are fossil fuel combustion, cement manufacture, decaying vegetation, and deforestation (less  $\text{CO}_2$  uptake by plants). The main sinks are photosynthesis and absorption within the ocean, leading to measurable acidification. The Intergovernmental Panel on Climate Change (IPCC, 2007) estimates that from 2000-2005, the average annual net flux of  $\text{CO}_2$  from the Earth's surface to the atmosphere was approximately 7.2 GtC. Of that amount, 4.1 GtC remained in the atmosphere, 2.2 GtC was absorbed by the ocean, and 0.9 GtC went into the land.

The numbers immediately above were *net* fluxes. The actual fluxes in each direction are much larger (see Fig. 7.3, p. 515, IPCC, 2007). For example, photosynthesis and absorption of  $\text{CO}_2$  in the ocean remove roughly 200 GtC per year from the atmosphere. That loss rate, compared with the atmospheric reservoir of 750 GtC, leads to a residence time of 3.8 years, about the time it would take for an atmospheric  $\text{CO}_2$  molecule to be dissolved in the ocean or taken up by plants. This is far different from the atmospheric lifetime of  $\text{CO}_2$ , the time it takes for a sudden increase in  $\text{CO}_2$  to decay to  $1/e$  (.368) of its original value. The latter is thought to be about 100 years.

#### 4.5 The Noble Gases

The noble gases, in the order of increasing atomic number, are helium, neon, argon, krypton, xenon, and radon. They were originally called noble because they were considered inert, i.e., chemically unreactive. All six gases have the maximum possible number of electrons in their outer shells, two for helium and eight for the other gases, making them very stable. All these gases except radon are present in sufficient concentrations to merit inclusion in Table 2.

After hydrogen, *helium* is the second most abundant element in the universe, but it is relatively rare in the Earth's atmosphere. Its principal earthly source is radioactive decay in minerals containing uranium and thorium, which produces helium nuclei ( $\text{He}^{2+}$ ), with which electrons readily combine. About 3,000 metric tons of helium are generated each year in the Earth's crust. Helium is extracted from natural gas wells. Most atmospheric helium eventually escapes to space, as noted in section 3. Nonetheless, its residence time in the atmosphere is about a million years. Above 700-km altitude, helium is the most abundant atmospheric gas.

*Neon* is the least reactive of all elements. Though it is the fifth most abundant chemical element in the universe (after hydrogen, helium, oxygen, and carbon), it is rare on Earth. Some neon is locked in the Earth's crust, and is found in samples of volcanic gas.

As Table 2 indicates, the concentration of *argon* in the atmosphere is more than two orders of magnitude greater than that of the next most plentiful noble gas. The mass of atmospheric argon is estimated to be  $6.6 \times 10^{16}$  kg. It is the product of radioactive decay of Potassium-40, and reaches the atmosphere through outgassing from the lithosphere at a rate of  $2.2 \times 10^6$  kg yr<sup>1</sup>.

*Krypton* may be present in trace amounts in the Earth's crust. One known source is the fission of uranium and plutonium in nuclear reactors.

*Xenon* has been found in gases emitted from some mineral springs. Small amounts are locked within quartz rock. Radioactive xenon is produced by fission in fuel rods.

All the noble gases are thought to have been present in the Earth's atmosphere for much of its geologic history. Except for helium atoms, which can attain the escape velocity, the noble gases suffer no detectable losses, mainly because they react with very few substances. Their residence time in the atmosphere far exceeds a million years.

#### 4.6 Methane

The concentration of *methane* in the atmosphere has more than doubled since pre-industrial times, but is no longer growing. It is the most abundant organic gas in the atmosphere. Anthropogenic sources are agriculture, mainly from livestock and rice paddies; landfills; coal mining; natural gas distribution; and biomass burning. Natural sources are comparable in size. Methane is produced in oxygen-poor habitats such as tropical wetlands, swamps, tundra, and high-latitude marshes mainly as a result of decomposition of organic mater. Methane hydrates are an aggregate of methane and water, in which water molecules cluster around one or more methane molecules. They exist in the cold depths of the ocean and in permafrost. There is concern that melting of permafrost could release considerable methane into the atmosphere. Per unit mass, methane is 22 times more effective than CO<sub>2</sub> in absorbing infrared radiation over a span of 100 years, but its residence time in the atmosphere is comparatively short, about nine years.

Oxidation of methane by the hydroxyl radical, primarily in the stratosphere, accounts for more than 90% of methane loss. In chemistry, "hydroxyl" refers to a molecule consisting of a hydrogen atom and an oxygen atom. The neutral form is called the hydroxyl radical (OH). Its concentrations are very low in the troposphere (parts per trillion by volume—pptv), rising to nearly 1 ppbv at 50-km altitude. OH is highly reactive; its lifetime is measured in tens of minutes. Oxidation of methane by the hydroxyl radical eventually results in the production of water vapor (the major stratospheric source) and molecular hydrogen. The remaining 10% of methane loss occurs through uptake by the soil and destruction in the stratosphere.

#### 4.7 Nitrous oxide

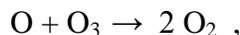
*Nitrous oxide* (N<sub>2</sub>O) is an important greenhouse gas, ranking fourth behind CO<sub>2</sub>, CH<sub>4</sub>, and CFC-12\* in its ability to change the Earth's infrared radiative budget. Its current concentration in the atmosphere of 0.32 ppmv is about 18% more than the pre-industrial value, after having varied by only several percent for the 10,000 years before that. The natural sources of N<sub>2</sub>O are soils, primarily in the tropics but also in mid-latitudes, and the oceans. Natural sources outweigh anthropogenic sources by nearly 60%. The latter sources are, in order of size, cultivated soils; rivers, estuaries, and coastal zones; fossil fuel combustion and industrial emissions; biomass burning; and cattle and feed lots. The stratosphere is the principal sink for N<sub>2</sub>O. In a reaction caused by the absorption of solar ultraviolet radiation, N<sub>2</sub>O is converted to N<sub>2</sub> and an oxygen atom (O). A small amount may also be removed in soils. The residence time of N<sub>2</sub>O in the atmosphere is estimated at 110-150 years.

#### 4.8 Ozone

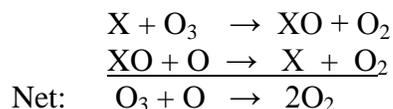
Ozone is an important atmospheric gas for several reasons. Its absorption of solar ultraviolet (UV) radiation in the stratosphere, particularly at wavelengths between 0.23 and 0.32 μm, greatly reduces the UV incident at the Earth's surface. Reductions in stratospheric ozone are linked to increasing incidence of skin cancer. As discussed in Section 5, absorption of UV radiation by O<sub>3</sub> causes heating and essentially determines the vertical temperature profile in the stratosphere. Ozone participates in many chemical reactions. For example, it is product of reactions that form photochemical smog near the Earth's surface when sunlight interacts with oxides of nitrogen from fossil fuel combustion and volatile organic compounds. In the lower troposphere, O<sub>3</sub> is a health hazard. Finally, like water vapor, O<sub>3</sub> is a naturally occurring greenhouse gas. It absorbs infrared radiation strongly at 9.6 μm.

Sea-level values of O<sub>3</sub> vary widely, from 10-500 ppbv. In the stratosphere, where ozone is produced, the concentration is much greater, reaching a maximum of 8-10 ppmv at 35 km. The production of stratospheric ozone begins with the dissociation of molecular oxygen by UV radiation into atomic oxygen: O<sub>2</sub> + sunlight → O + O. This reaction was first mentioned in section 4.2. In the presence of N<sub>2</sub> or O<sub>2</sub>, here represented by M, atomic and molecular oxygen combine to form ozone: O + O<sub>2</sub> + M → O<sub>3</sub> + M.

The destruction of stratospheric O<sub>3</sub> occurs by several pathways. The most direct is the reaction



which is fairly slow, but several catalytic reactions of the form



can speed the destruction of O<sub>3</sub>, where X is the catalyst. The most important catalysts are hydrogen (H), the hydroxyl radical (OH), nitric oxide (NO), and chlorine (Cl). All these

---

\* CFC-12 (CCl<sub>2</sub>F<sub>2</sub>) is one of several chlorofluorocarbon gases used in refrigeration, as a propellant in aerosol spray cans, and in cleaning solvents. Atmospheric CFCs are present in minute concentrations (pptv) but are radiatively very important, having figured in stratospheric ozone depletion. These compounds are not discussed further here.

reactions are effective at altitudes above 30 km. Below 30 km, other reactions not involving atomic oxygen, dominate. The residence time of O<sub>3</sub> in the atmosphere is only about 100 days.

The foregoing describes the natural production and destruction of ozone. Chlorofluorocarbons, manufactured chemicals not present in the atmosphere 100 years ago, have been implicated in the destruction of ozone at polar latitudes in early spring, when stratospheric air is still very cold, 190K or less. When sun shines on polar stratospheric clouds, composed chiefly of water ice, chemical reactions occurring on the surface of cloud particles result in the liberation of chlorine, which then destroys O<sub>3</sub> in the catalytic reaction just described and other reactions. This process has resulted in substantial seasonal ozone depletion, commonly called the “ozone hole.”

## 4.9 Hydrogen

The total mass of molecular hydrogen (H<sub>2</sub>) in the Earth’s atmosphere is very roughly  $2 \times 10^{11}$  kg. Estimates of the H<sub>2</sub> fluxes into and out of the atmosphere vary by a factor of three from 0.4 to more than  $1.3 \times 10^{11}$  kg yr<sup>-1</sup>, giving a residence time of a few years.

Natural sources of H<sub>2</sub> are emission from the oceans, biological processes, and perhaps geothermal steam and volcanoes. Photochemical oxidation of methane in the stratosphere, mentioned earlier, is a substantial source, as is the release of isoprene and terpenes (hydrocarbons) from plants, accelerated by photosynthesis and when leaf temperatures rise. Photochemical reactions produce H<sub>2</sub> from these substances. Anthropogenic sources are internal combustion engines and biomass burning.

Soils extract the bulk of H<sub>2</sub> from the atmosphere. Other substantial losses occur in the troposphere and stratosphere in two ways. 1) The reaction of molecular hydrogen with the hydroxyl radical forms water and atomic hydrogen:  $\text{OH} + \text{H}_2 \rightarrow \text{H}_2\text{O} + \text{H}$ . 2) Molecular hydrogen reacts with an electronically excited oxygen atom O(<sup>1</sup>D) to form the hydroxyl radical and atomic hydrogen:  $\text{O}(\text{}^1\text{D}) + \text{H}_2 \rightarrow \text{OH} + \text{H}$ . The loss of molecular hydrogen thus results in a gain of atomic hydrogen, which becomes the dominant species of hydrogen beyond the top of the homosphere (100-km altitude). As discussed in section 3, the escape of atomic hydrogen from the exosphere is significant.

Two excellent references for much of the material in this section are Hobbs (2000, pp. 143-152) and Brasseur *et al.* (1999, pp.159-203). Figure 2 is from the latter reference. It shows the mixing ratios of many of the gases discussed above from the surface to 120-km altitude. The mixing ratio is the number density of the gas under consideration divided by number density of all gases in dry air. Not depicted are N<sub>2</sub> and the rare noble gases, whose mixing ratios are nearly constant up to 100 km because of their chemical stability. The N<sub>2</sub> mixing ratio curve would parallel the O<sub>2</sub> curve and lie to its right, starting at a sea-level value of 0.78. As noted above, the stratosphere is a major sink for methane and nitrous oxide, and so their mixing ratios begin to decrease above 20-km altitude. Water vapor concentration is temperature limited, so its mixing ratio decreases rapidly with decreasing temperature in the troposphere. Finally, as just noted, O<sub>3</sub> has a distinct stratospheric maximum at 35 km.

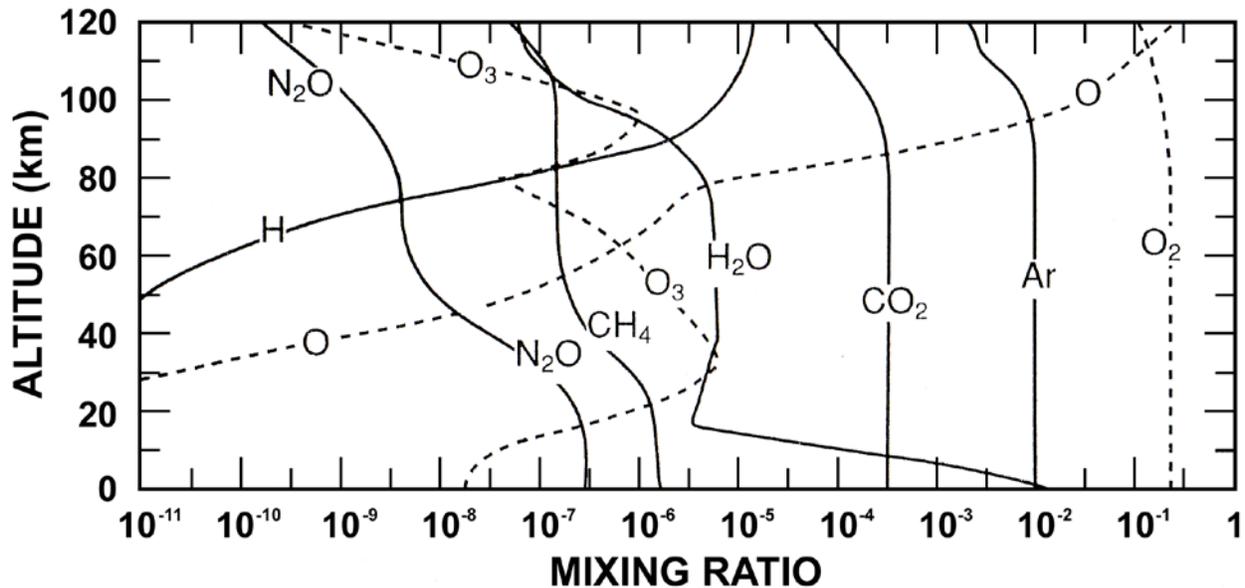


Figure 2. The mixing ratio of some of the atmospheric gases listed in Table 2. Neither nitrogen nor the rare noble gases are shown because they are all well mixed in the homosphere (up to about 100 km). Adapted from Brasseur *et al.* (1999; p. 9)

GasMixingRatioVsAlt.tif

## 5. ATMOSPHERIC SHELLS DEFINED BY TEMPERATURE VARIATION WITH ALTITUDE

Section 3 categorized the atmosphere in terms of three layers, the homosphere, heterosphere, and exosphere, primarily on the basis of number densities of gas molecules and the relative importance of turbulence versus molecular diffusion. This section categorizes the atmosphere in terms of vertical temperature structure. Four main “shells” are defined and the boundaries between them. In order of increasing altitude, they are: the *troposphere*, whose top boundary is called the *tropopause*; the *stratosphere*, whose top boundary is the *stratopause*; the *mesosphere*, whose top boundary is the *mesopause*; and the *thermosphere*. We consider each of these layers in turn and explain the temperature variation within them.

A *standard atmosphere* is an idealized representation of the atmosphere based upon long-term average conditions. Figure 3 shows the vertical temperature structure of the U.S. Standard Atmosphere (1976), representative of mid-latitude conditions, to just beyond 100-km altitude, along with the atmospheric layers and interfaces. The temperature decreases with altitude in the troposphere, increases in the stratosphere, decreases again in the mesosphere, then increases for good in the thermosphere, which extends far beyond the upper limit of Fig. 3. More information on standard atmospheres follows in section 5.5.

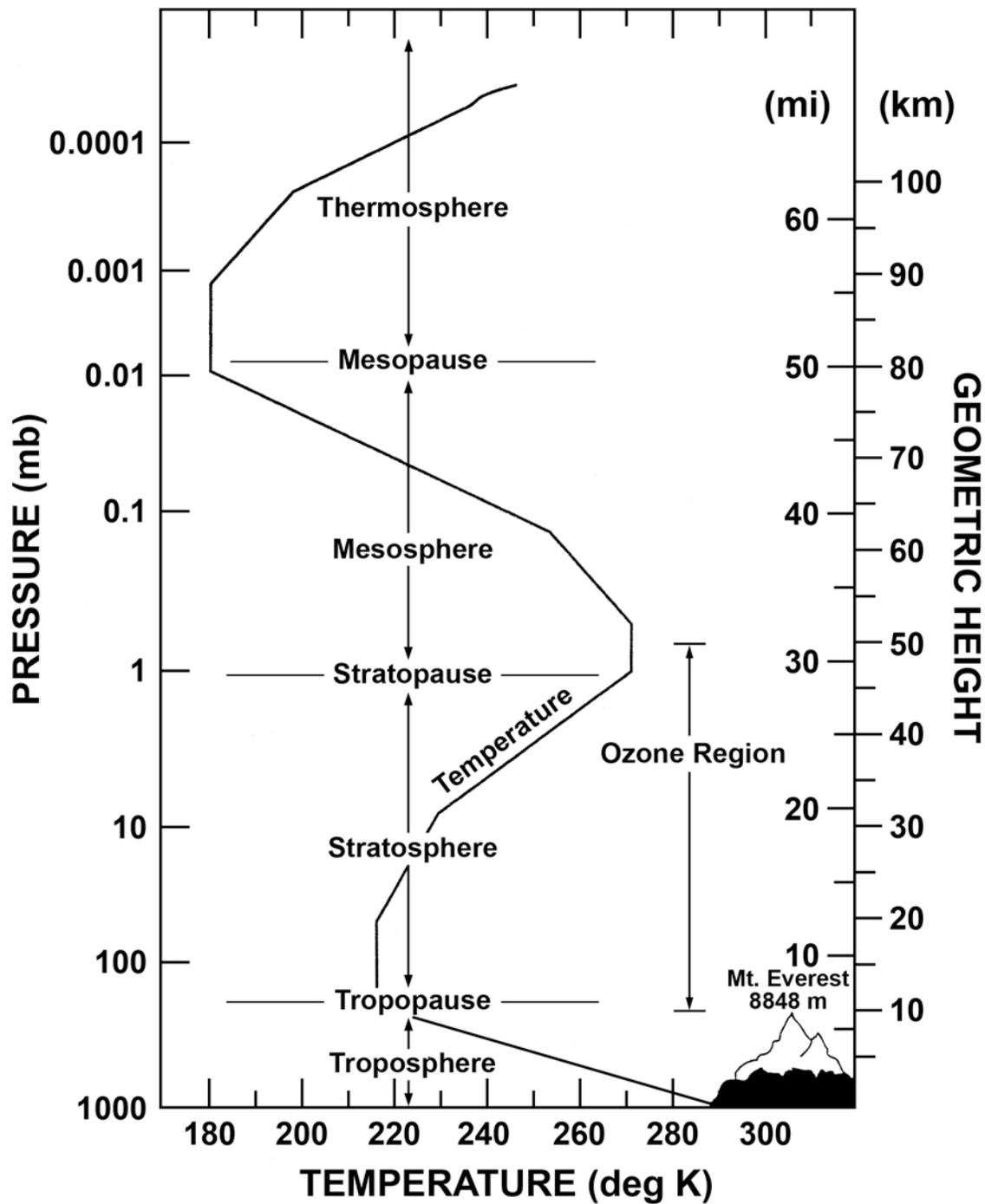


Figure 3. Vertical temperature structure of the atmosphere extending from the surface of the Earth to approximately 110-km altitude as given in the *U.S. Standard Atmosphere, 1976*. The principal layers defined by this temperature structure and the interfaces between them are labeled. As indicated, ozone is found principally in the stratosphere. Two vertical coordinates are given: pressure on the left in millibars (1 mb = 100 Pascal) and geometric altitude on the right (km).

Fig. 3 is adapted from Fig. 1.5 from Brasseur et al, p. 7. VertTempStruc.jpg

## 5.1 Troposphere

The troposphere is the lowest layer of the atmosphere, typically ranging from 9 to 17 km in thickness, in which temperature generally decreases with height. Clouds, precipitation, and storms inhabit the troposphere. This section explains why temperature decreases with altitude in the troposphere and how an equilibrium established between radiative and convective processes determines the height of the tropopause.

Nearly half (47%) the amount of solar radiant energy incident at the top of the atmosphere is absorbed at the ground or in the oceans. Much of the absorbed energy goes back into the atmosphere in the form of sensible heat, latent heat acquired through the evaporation of water, and infrared radiation. Convection explains why tropospheric temperature, on average, decreases with altitude. Tropospheric convection would still occur without water vapor but it would not be nearly so vigorous.

The decrease of temperature with height is called the *lapse rate*, expressed as  $(-dT/dz)$ . The atmospheric lapse rate seldom exceeds 9.8°C cooling per kilometer of altitude. This is called the *dry adiabatic lapse rate*  $\Gamma_d$ .

$$\Gamma_d = g / c_p, \quad (26)$$

where  $c_p$  is the specific heat of air at constant pressure ( $\sim 1005 \text{ J kg}^{-1} \text{ K}^{-1}$ ). Note that  $dT/dz = -\Gamma_d$ . The nomenclature arises because this is the rate at which a volume of cloud-free air, hypothetically isolated from its surroundings (no heat exchange), will cool if forced upward or warm if forced downward.

It is important to distinguish between a measured temperature profile (by an instrumented balloon or aircraft) and the parcel trajectory, the temperature change a volume of air would experience if forced upward or downward. If an air parcel is warmer than its surroundings at the same pressure, it is less dense (buoyant) and accelerates upward. Conversely, if it is cooler than its surroundings, it is more dense and accelerates downward.

Stability in the atmosphere is judged as follows. The atmosphere is considered *stable* if an air parcel, displaced upward or downward, experiences a restoring force and returns to its original position. It is considered *unstable* if a displacement results in a buoyancy force and acceleration away from its initial position. It is considered *neutral* if the initial displacement results in no force on the parcel.

The dry adiabatic lapse rate characterizes *neutral* stability because, if a volume of cloud-free air is displaced upward or downward under this condition, it continues on its way without acceleration. Rising or sinking parcels maintain the same temperature as their surroundings. Dry adiabatic lapse rates are common in the lowest few kilometers of the atmosphere on sunny summer afternoons. Thermals are common, and this is when soaring pilots like to fly.

The lapse rate is seldom greater than  $\Gamma_d$  except within centimeters of intensely heated ground. On the other hand, the lapse rate is often less than  $\Gamma_d$ . In fact, the temperature sometimes increases with height. This is called an *inversion*. Inversions represent very stable stratification of the air because they strongly inhibit vertical motion. An air parcel rising through an inversion quickly finds itself cooler than its surroundings and sinks back toward its origin. The *tropopause* is marked by a semi-permanent stable layer, either an isothermal layer or an inversion, at altitudes of 10-17 km (higher in summer than winter, and higher near the equator than near the poles). This stable layer invariably limits the vertical development of clouds.

Water vapor is the fuel for strong tropospheric convection. When water vapor condenses within a rising air parcel, heat is released. This boost of temperature within the air parcel can cause it to become warmer than its surroundings and remain buoyant, even as it rises many kilometers. When this happens, a tall cloud (often a thunderstorm) forms.

When the measured temperature and moisture profiles favor parcel buoyancy (due to condensation heating), the atmosphere is said to be *conditionally unstable*. Conditional instability occurs when the environmental lapse rate  $\Gamma$  (positive when temperature decreases with height) satisfies the following inequality:

$$\Gamma_m < \Gamma < \Gamma_d , \quad (27)$$

where  $\Gamma_m$  is called the *moist adiabatic lapse rate*, the rate at which saturated (cloudy) air cools as it rises, and any condensate that forms does not fall out. The expression for  $\Gamma_m$  is considerably more complicated than that for  $\Gamma_d$ . It is derived in Bohren and Albrecht (1998, pp. 287-292). As noted earlier, the amount of vapor in saturated air (air on the verge of condensation) increases exponentially with temperature. In saturated ascent (for which  $\Gamma_m$  describes how the parcel temperature changes with altitude), the condensation rate and the heating within the parcel are both higher at higher temperatures, because more moisture is present. In the lower troposphere,  $\Gamma_m$  is about  $5^\circ\text{C km}^{-1}$ . High in the troposphere, where the air is very cold and moisture is minimal,  $\Gamma_m$  approaches  $\Gamma_d$ .

To fix ideas, see Figure 4, which illustrates the temperature (thick, solid, light grey curve) and dewpoint (thick, solid, dark grey curve) at each altitude, as measured by a sounding balloon. This sounding gives a picture of environmental conditions from the surface to 16-km altitude. The dashed curve indicates the path a parcel would follow if it were lifted from the surface, starting at the same temperature and dewpoint as measured by the balloon. In the first small segment, from the surface to about 2-km altitude, the parcel temperature is the same as the environment temperature because the environmental lapse rate is dry adiabatic, and rising parcels of air cool at this rate until condensation occurs. (The thin straight line indicates the dry adiabatic lapse rate.) From 2- to about 4-km altitude, the parcel is cooler than its surroundings. If a thunderstorm is to form, some means must be present to force the parcel upward until it becomes warmer than its surroundings. This could be as simple as converging airflows along a ridgetop.

The water vapor mixing ratio that the parcel had when it began its ascent from the surface remains constant in dry adiabatic ascent. The dewpoint within the parcel decreases as it rises, but not as fast as the temperature. The parcel temperature and dewpoint converge. They meet at cloud base, where condensation begins. At this point, internal heating warms the parcel so that it no longer cools at the dry adiabatic but at the moist adiabatic lapse rate. The abrupt change in parcel trajectory at cloud base indicates this. Because of condensation heating, the parcel soon becomes warmer than its surroundings (at the point where the dashed curve crosses the solid curve). This is called the level of free convection, at which point, the parcel begins rising like a hot-air balloon.

In this example, the parcel remains warmer than its surroundings until nearly 13-km altitude. By then the temperature stratification outside the cloud has become more stable. The tropopause is at 14 km. Above 13 km, the parcel rapidly becomes cooler and denser than its surroundings, and the updraft building the cloud quickly drops to zero. If the cloud overshoots the tropopause, it will do so only briefly, then, within minutes, will sink back to its equilibrium level.

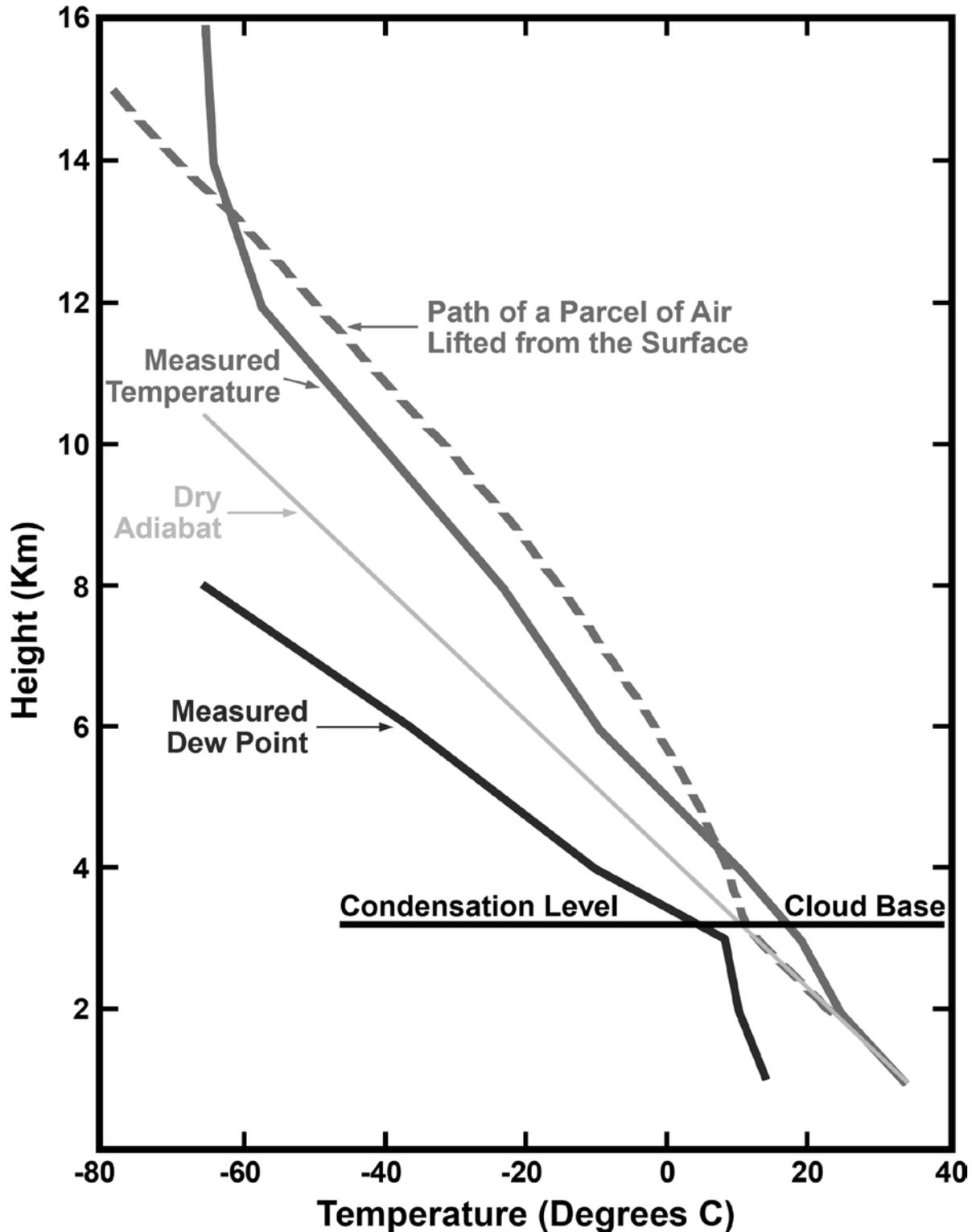


Figure 4. The thick grey and the thick black curves show measured temperature and dewpoint, respectively. The dashed grey curve is the trajectory a parcel would follow, if its starting temperature and dewpoint at the surface were the same as observed. A relevant dry adiabat is shown as the thin straight line. [RaobSoundingwCAPE.tif](#)

In Fig. 3, the standard tropospheric lapse rate is  $0.65^{\circ}\text{C}$ , less than  $\Gamma_d$  and within the normal range of  $\Gamma_m$ . We have already noted that the tropopause is highest over the equator and slopes downward toward the poles. Why should this be? The answer is that deep convection adjusts some very steep lapse rates that would occur if the troposphere were in pure radiative equilibrium.

Radiative equilibrium is said to occur if the upward and downward fluxes of radiation at each wavelength and at each atmospheric level are in balance. The simplest models of radiative equilibrium consider two broad wavelength intervals, which are very nearly separated: solar radiation up to  $4\text{-}\mu\text{m}$  wavelength, and infrared radiation, emanating from the Earth's surface and atmosphere at wavelengths greater than  $4\text{ }\mu\text{m}$ . The atmosphere is considered transparent to solar radiation, both incident from the sun and reflected, and so does not affect the radiative balance. Equilibrium thus requires equal upward and downward fluxes of only infrared radiation at each level. The calculations require knowledge of the absorbing and emitting properties of atmospheric gases.

The result of these simple models is a nearly constant temperature profile above 15 km, then gradually steepening lapse rates at lower altitudes. The lapse rate for radiative equilibrium typically exceeds  $\Gamma_m$  by 6-km altitude and  $\Gamma_d$  by 5 km, becoming absolutely unstable below 5 km. For reasons described above, the absolute instability cannot endure, and moist convection would quickly reduce the tropospheric lapse rate to a value below  $\Gamma_d$ . *Radiative-convective equilibrium* is the result. The point where the environmental temperature profile established by convection intersects at higher altitudes the nearly isothermal profile consistent with pure radiative equilibrium is effectively the tropopause.

Surface temperature is positively correlated with the height of the troposphere. In the tropics, where surface temperature and moisture supply are both higher than at most other latitudes, convection tends to be deeper and the troposphere higher. With increasing latitude, surface temperature lowers or, if not (for example, in the subtropical deserts), moisture supply is usually limited. Convection, though sometimes more vigorous than in the tropics, is not as deep. In midlatitudes, seasonal swings in surface temperature bring corresponding swings in available moisture, frequency of convection, and tropopause height. During the polar winter, with little or no sunlight for weeks to months, surface inversions can become pronounced due to infrared radiation losses to the atmosphere and space. The tropopause can become ill-defined under such conditions.

## 5.2 Stratosphere

The stratosphere extends from the tropopause to approximately 50-km altitude. Temperature increases throughout its depth, slowly near the bottom, more rapidly near the top. The stratosphere is therefore a very stable layer, not subject to convection. It is also dry because most of its water vapor enters through the tropopause, which, by virtue of its low temperature, is often called a “moisture trap.” Polar stratospheric clouds are occasionally observed, but only at exceptionally low temperatures. If fine particulate matter from an explosive volcanic eruption enters the stratosphere, the residence time is long, often more than a year, because there is no precipitation to wash the ash out.

The stratosphere contains 90% of atmospheric ozone. The ozone mixing ratio is typically 3 ppmv at 20-km altitude, rises to a maximum of 8-10 ppmv at 35 km, then decreases to 2 ppmv at the stratopause. Conditions in the stratosphere are dominated by radiative processes. The increase of temperature with altitude results from the absorption of solar ultraviolet radiation by ozone.

The absorption of solar ultraviolet radiation at stratospheric altitudes and above is fortuitous because unattenuated ultraviolet radiation at the Earth's surface would be harmful to plant and animal life alike. Figure 5 shows the altitude at which absorption lowers the intensity of ultraviolet radiation to  $1/e$  (37%) of its value at the top of the atmosphere. Shorter wavelength radiation is generally absorbed at higher altitudes. UVA and UVB light, which most sun blocks screen effectively, occupy wavelengths of 400-320 nm and 320-280 nm, respectively. The principal gaseous absorbers at different altitudes are shown. As noted earlier, ozone is the principal absorber in the stratosphere.

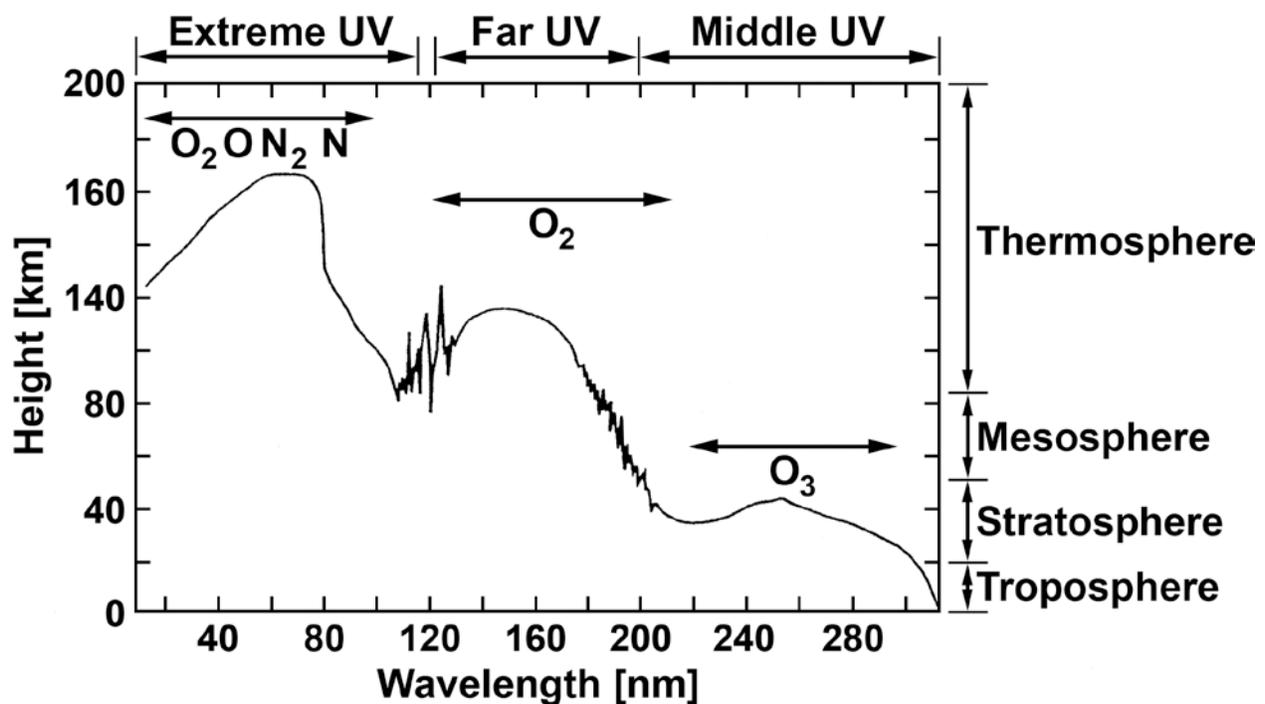


Figure 5. Absorption of ultraviolet radiation by neutral atmospheric gases as a function of altitude and wavelength. Absorption results both in dissociation, the formation of atomic oxygen and nitrogen from the corresponding molecules, and ionization, stripping an electron from any of these species.

Fig. 5 is from Ondoh and Marubashi (2000), p. 8, UVAbsorb.jpg

### 5.3 Mesosphere

The mesosphere extends from the stratopause up to 80- to 90-km altitude. Molecular oxygen and nitrogen still comprise 99% of air, having virtually the same mixing ratios as at sea level. Tiny

meteors burn up in the mesosphere regularly. High-speed collisions between meteorites and gas molecules generate enough heat to vaporize all but the largest of them. The temperature decreases throughout the mesosphere, reaching a minimum of about 185 K at the mesopause. There is less UV absorption in the mesosphere than in the stratosphere, primarily because the concentration of ozone decreases with altitude (Fig. 2). Ozone is still the most important absorber, followed by molecular oxygen. These are the two main sources of radiative warming. CO<sub>2</sub> emits infrared radiation in the mesosphere, and is the major source of radiative cooling. The summer mesopause is the coldest region in the atmosphere. Temperatures as low as 170 K have been observed there. *Noctilucent clouds* of tiny ice particles occasionally form near the summer mesopause at high latitudes.

Radiative equilibrium alone does not explain the very low temperatures observed at the summer mesopause. In fact, despite the absence of sunlight at high latitudes of the winter hemisphere, the winter mesopause is not as cold. Large-scale upward motion and adiabatic cooling play a role in maintaining the low temperatures at the summer mesopause. Gravity waves are thought to be partly responsible for this; they influence, as well, the mean flow from west to east in the mesosphere. A gravity wave is a wave disturbance in which buoyancy acts as the restoring force on parcels of air displaced from hydrostatic equilibrium.

Most gravity waves emanate from the troposphere or lower stratosphere. Common sources are mountain ranges, vigorous convection, atmospheric fronts, and vertical shear of the horizontal wind. As they propagate upward from their source regions with a characteristic period of 1-2 h, gravity waves encounter less dense air and grow in amplitude, transporting energy and horizontal momentum vertically. Their energy per unit mass increases by one to two orders of magnitude as they propagate from troposphere to mesosphere. When the amplitude becomes sufficiently large, as often occurs in the mesosphere, the wave breaks, resulting in turbulence and energy dissipation. These breaking waves have a major effect on the mean mesospheric wind, causing local accelerations, referred to as *gravity wave drag*. They influence both the mean wind and the thermal structure at mesospheric altitudes (Fritts, 1995).

## 5.4 Thermosphere

The thermosphere extends upward from roughly 86 km and gradually merges with the cloud of neutral hydrogen that surrounds Earth. The temperature increases rapidly with altitude in the lower thermosphere up to about 120 km, then increases more and more slowly, approaching 1000 K by 700-km altitude. Because the temperature is so strongly affected by the absorption of ultraviolet radiation, it changes rapidly between day and night. The temperature is also strongly affected by solar storms. The atmosphere is thin enough in the upper thermosphere that many satellites orbit the Earth at altitudes between 400 and 1000 km. Atmospheric density here is quite low but still sufficient to impose a weak frictional drag on spacecraft that affects orbital prediction. The aurora occurs at altitudes from 80 to 300 km. The ionosphere refers to the layer of free electrons and positively charged ions that overlaps the neutral thermosphere, and extends from the upper mesosphere to beyond 1000 km. It will be discussed in the next section, but suffice it to say here that interactions among neutral gases and charged particles play a big role in the thermospheric structure.

The boundary between the homosphere and heterosphere lies near the bottom of the thermosphere, near 100-km altitude. Below this altitude, eddy diffusion (bulk motion of large volumes of air) dominates atmospheric transport. Above this altitude, molecular diffusion controls the relative concentration of gases and the temperature profile. Absorption of extreme ultraviolet (EUV) at wavelengths less than 103 nm ionizes (strips electrons from) molecular and atomic oxygen and nitrogen. UV radiation also dissociates O<sub>2</sub> and N<sub>2</sub> into O and N so that, above 200-km altitude, the atomic species outnumber the molecular species.

According to Roble (2003), about one-third of the UV energy absorbed by the neutral gases of the thermosphere is used for local heating. The rest can be radiated away or transformed into chemical energy. Infrared cooling in the thermosphere is relatively weak, through emissions by CO<sub>2</sub>, O, and nitric oxide (NO), but this is not enough to balance the UV heating, which is greatest in the high thermosphere because that is where the most energetic radiation is absorbed. Thus the thermal stratification is controlled by downward molecular diffusion of heat to at least 120 km. The diffusion is particularly efficient above 500 km because the thermal diffusivity is so high; it is inversely proportional to the number density, which decreases by eight orders of magnitude between the mesopause and 1000-km altitude. That is why the temperature profile becomes isothermal high in the thermosphere.

## 5.5 Standard Atmospheres

*Standard atmospheres* are prescriptions of atmospheric properties, primarily as a function of altitude, useful for engineering applications. One very widely used standard atmosphere is the *U.S. Standard Atmosphere, 1976*, already referenced several times. Table 3 is constructed from this source and summarizes the foregoing discussion of atmospheric shells.

Table 3. Temperature, pressure, and density as a function of geometric altitude from the *U.S. Standard Atmosphere, 1976* for each of the main atmospheric layers. Dry air is assumed. Levels where the lapse rate changes are included, except in the thermosphere. A two-digit number preceded by the plus or minus sign indicates the power of ten by which the preceding number is to be multiplied.

Geometric Altitude (m)	Temperature (K)	Pressure (hPa) <sup>*</sup>	Density (kg m <sup>-3</sup> )
<b>Troposphere – Constant lapse rate 0.65 K km<sup>-1</sup> (cooling with altitude) from 0.0 to 11.0 km.</b>			
0	288.150	1.01325 +03	1.2250 +00
1,000	281.651	8.9876 +02	1.1117 +00
2,000	275.154	7.9501 +02	1.0066 +00
4,000	262.166	6.1660 +02	8.1935 -01
6,000	249.187	4.7217 +02	6.6011 -01
8,000	236.215	3.5651 +02	5.2579 -01

10,000	223.252	2.6499 +02	4.1351 -01
<b>Tropopause</b> 11,000	216.774	2.2699 +02	3.6480 -01
<b>Stratosphere – Isothermal (216.65) K from 11.1 to 20.0 km, then warming with altitude. Lapse rate <math>-1.0 \text{ K km}^{-1}</math> from 20 to 32 km, then <math>-2.8 \text{ K km}^{-1}</math> from 32.0 to 47.4 km.</b>			
15,000	216.650	1.2111 +02	1.9476 -01
20,000	216.650	5.5293 +01	8.8910 -02
25,000	221.552	2.5492 +01	4.0084 -02
32,000	228.490	8.8906 +00	1.3555 -02
37,000	242.050	4.3324 +00	6.2355 -03
<b>Stratopause</b> 47,400	270.650	1.1022 +00	1.4187 -03
<b>Mesosphere – Isothermal (270.650 K) from 47.4 to 51.0 km, then cooling with altitude. Lapse rate <math>2.8 \text{ K km}^{-1}</math> from 51 to 71 km, then <math>2.0 \text{ K km}^{-1}</math> from 71 to 86 km.</b>			
48,000	270.650	1.0229 +00	1.3167 -03
51,000	270.650	7.0458 -01	9.0690 -04
61,000	244.274	1.9157 -01	2.7321 -04
71,000	216.846	4.4795 -02	7.1966 -04
<b>Mesopause</b> 86,000	186.87	3.7338 -03	6.958 -06
<b>Thermosphere – Isothermal (186.87 K) from 86 to 92 km, then warming with altitude. Lapse rate decreases until it reaches <math>-12 \text{ K km}^{-1}</math> at 120 km, then increases toward zero. Temperature asymptotically approaches 1000 K, and reaches that point at 815-km altitude.</b>			
92,000	186.96	1.2887 -03	2.393 -06
100,000	195.08	3.2011 -04	5.604 -07
120,000	360.00	2.5382 -05	2.222 -08
150,000	634.39	4.5422 -06	2.076 -09
200,000	854.56	8.4736 -07	2.541 -10
300,000	976.01	8.7704 -08	1.916 -11
500,000	999.24	3.0236 -09	5.215 -13
700,000	999.97	3.1908 -10	3.070 -14
1,000,000	1000.00	7.5138 -11	3.561 -15

\* hPa = hectoPascal = 100 Pascals = 100 Newtons per square meter = 1 millibar

Other standard atmospheres have been developed in recent years. Among them are the International Civil Aviation Organization (ICAO) Standard Atmosphere (ICAO, 1993), which provides standard values of temperature, pressure, and density for levels up to 80 km and the underlying equations used to calculate them. This atmosphere contains no water vapor. It is used for calibration of pressure altimeters, evaluation of aircraft performance, and engineering design.

The International Standard Atmosphere (ISA) is published by the International Organization for Standardization. The ISA model for vertical profiles of pressure, density, temperature, and viscosity is based on average conditions at midlatitudes. The *U.S. Standard Atmosphere, 1976* is identical to the ICAO Standard Atmosphere up to 32 km and to the International Standard Atmosphere up to 50 km. All are dry atmospheres.

The COSPAR International Reference Atmosphere (CIRA-86) (<http://badc.nerc.ac.uk/data/cira/>) is more detailed than other standard atmospheres in that it provides mean zonal wind (component of wind from the west) in addition to temperature, pressure and density, at altitudes up to 120 km and at latitudes between 80 S and 80 N. At altitudes from 120 to 2000 km, CIRA-86 merges with the MSIS-86 standard atmosphere, mentioned in the next paragraph.

*MSIS* stands for Mass Spectrometer and Incoherent Scatter radar, the two primary sources of atmospheric data for development of MSIS standard atmospheres. MSIS-86 is an early version. The latest version, NRLMSISE-00, has been developed by the US Naval Research Laboratory (NRL) and has become a standard for international space research. It has been calibrated with actual satellite drag data. The *E* indicates that this standard atmosphere extends from the surface well into the exosphere; the *00* gives the year of release, 2000. The inputs to NRLMSISE-00 include year, day, time of day, latitude, longitude, altitude, local solar time, and information about recent solar activity and the magnetic index. The output is number density for primary atmospheric gases, total mass density, and temperature. For more information, go to <http://www.nrl.navy.mil/content.php?P=03REVIEW105>.

One of the most important applications of the standard atmosphere is to keep aircraft vertically separated by a safe distance. Above what is called the *transition altitude* (18,000 ft in the U.S., generally 3,000 to 14,000 ft elsewhere), aircraft under any kind of flight control fly on constant pressure surfaces. This is easy because their altimeters are really pressure sensors. An aircraft reporting a flight level ( $z_{FL}$ ) of 35,000 ft is flying on the pressure surface ( $p_{FL}$ ) that corresponds to that altitude in the ICAO standard atmosphere. Note that  $z_{FL}$  is not the true altitude, but rather what the altitude would be in the standard atmosphere, given the pressure. At 10,000 m, the difference between  $z_{FL}$  and true altitude could be as much as 500 m. In practice, standard flight levels are separated by 1000 ft (305 m), except in China, where the separation is 300 m. Headings are based on ground-track magnetic directions, with eastbound defined as  $0^\circ$  through  $179^\circ$  (including N, NE, E, SE) and westbound defined as  $180^\circ$  through  $359^\circ$  (including S, SW, W, NW). Eastbound and westbound aircraft occupy alternating flight levels for obvious reasons. The relation between  $p_{FL}$  and  $z_{FL}$  is given by

$$p_{FL} = p_0 \left( \frac{T_0 - \gamma z_{FL}}{T_0} \right)^{\frac{g}{\gamma R_d}}, \quad z_{FL} \leq 11,000 \text{ m}, \quad (28)$$

where  $p_{FL}$  and  $p_0$  are expressed in hPa,  $p_0 = 1013.25$  hPa (standard sea-level pressure),  $T_0 = 288.15$  K (standard sea-level temperature), and  $\gamma = 0.0065$  K m<sup>-1</sup> is the tropospheric lapse rate in the ICAO standard atmosphere. See

<http://hurri.kean.edu/~yoh/calculations/standatm/StdAtm.html> for a derivation. Note that  $z_{FL}$  must be expressed in meters. For aircraft above a flight level of 11 km,

$$p_{FL} = p_1 \exp \left[ -\frac{g}{R_d T_1} (z_{FL} - 11,000) \right], \quad 11,000 < z_{FL} \leq 20,000 \text{ m}, \quad (29)$$

where  $p_1 = 226.32$  hPa is the pressure at the tropopause and  $T_1 = 216.65$  K is the constant temperature of the lower stratosphere in the ICAO standard atmosphere (Stull, 2000, p. 13). Note that both Equations (28) and (29) give the same answer when  $z_{FL} = 11,000$  m.

If an aircraft were to descend all the way to sea level, its altimeter would not read zero unless the sea-level pressure happened to be 1013.25 hPa. For the same reason, if an aircraft is landing at an airport whose runway elevation is  $z_{sta}$ , its altimeter will not read correctly when the wheels contact the runway unless the real atmosphere happens to match the ICAO standard atmosphere. For this reason, when flying below the transition altitude, pilots set their altimeter to the *altimeter setting*  $p_{as}$  for the appropriate airport, which is derived from the station pressure  $p_{sta}$  (true pressure as would be measured by a mercury barometer at the runway elevation) under the assumption of ICAO standard atmospheric conditions between the airport elevation and sea level. They will still be flying on surfaces of constant pressure below the transition altitude, but their altimeter will read the correct elevation when they land, even if the atmosphere both above and below the transition altitude is non-standard. This is critical, especially when runway visibility is low. The altimeter setting is calculated from

$$p_{as} = \left[ (p_{sta} - 0.3)^{\frac{R_d \gamma}{g}} + \frac{\gamma z_{sta}}{T_0} p_0^{\frac{R_d \gamma}{g}} \right]^{\frac{g}{R_d \gamma}}. \quad (30)$$

This equation is closely related to Equation (28) (see <http://hurri.kean.edu/~yoh/calculations/altimeter/>), and they share the same constants. The pressure should be expressed in hPa because the constant 0.3 is in hPa. This constant accounts for the fact that the altimeter is mounted approximately 3 m above the runway (at a pressure about 0.3 hPa less than that on the runway) on many commercial aircraft. The altimeter setting is sometimes given in hPa and sometimes in inches of mercury (in Hg), depending upon the country, and so a units conversion may be appropriate after applying Equation (30): 1 hPa = 0.02953 in Hg.

## 6. THE IONOSPHERE AND THE EARTH'S ELECTRICAL CIRCUIT

### 6.1 Formation of the Ionosphere

The ionosphere is a region of the upper atmosphere containing significant concentrations of ions (almost all carrying a positive charge) and free electrons. It extends from about 60-km altitude in the upper mesosphere to well beyond the 1000-km altitude considered in this essay. The cloud of charged particles in the ionosphere is called a *plasma*, a mixture of nearly equal numbers of positively and negatively charged particles, in this case, within a medium of neutral gases, the thermosphere. Even where the number density of electrons (similar to the number density of positive ions) peaks, near 300-km altitude, the number density of neutral gases is still more than 100 times greater. Still, the plasma is highly conductive. The response of the plasma to electric and magnetic fields, and the interactions between charged and neutral particles strongly influence the properties of the ionosphere.

For this discussion, the following nomenclature is useful: x-ray corresponds to the wavelength interval 0.01-1.0 nm; soft x-ray, 1.0 to 10 nm; extreme ultraviolet (EUV) 10 to 100 nm, and ultraviolet (UV), 100-300 nm. The shorter the wavelength, the more energetic the radiation. Physicists denote the energy carried by photons as  $h\nu$ , where  $h$  is the Planck constant ( $6.6261 \times 10^{-34} \text{ J s}^{-1}$ ) and  $\nu = c/\lambda$  is the frequency of the radiation.  $c$  is the speed of light ( $2.998 \times 10^8 \text{ m s}^{-1}$ ) and  $\lambda$  is the wavelength in meters. Radiation from the sun at UV and shorter wavelengths is energetic enough to strip electrons from neutral atmospheric gases, especially  $\text{N}_2$ ,  $\text{O}_2$ , and  $\text{O}$ , thereby ionizing them. It also causes dissociation: splitting molecules into component parts without producing free electrons.

#### Reactions

As short-wave radiation penetrates deeper into the atmosphere, the fraction absorbed increases, but so does the number density of atmospheric gases capable of absorbing it. The maximum electron/ion density occurs near 300 km. Below this level, photochemical equilibrium tends to prevail, whereby ion production is nearly balanced by recombination of the ionized species with electrons. The number density of particles and the high collision rate ensure this. Above 300 km, diffusive equilibrium prevails; as is the case with the neutral gases, the concentration of heavier ions decreases with altitude more rapidly than that of the lighter ions and much lighter electrons.

Before discussion of various regions of the ionosphere, it is instructive to examine Figure 6 from Hobbs (2000, p. 27), which summarizes much of the material presented so far. At left is a global mean temperature profile, similar to that in Fig. 3 except that it extends to 1500-km instead of 110-km altitude. Note the major difference in thermospheric temperatures depending upon whether the sun is quiet or active. Large swings in temperature also occur between day and night. The middle part of the figure illustrates the gradual change in the mean molecular weight of dry air, constant at  $28.9 \text{ kg kmol}^{-1}$  in the homosphere, but gradually decreasing in the heterosphere. The square brackets refer to the number densities of the gases whose chemical

symbols are inside. Thus the number density for O becomes greater than that for O<sub>2</sub> at about 110-km altitude, and greater than that for N<sub>2</sub> near 180 km. At right, the ozone layer and various regions of the ionosphere are depicted along with representative electron densities. The vertical scale in kilometers is logarithmic.

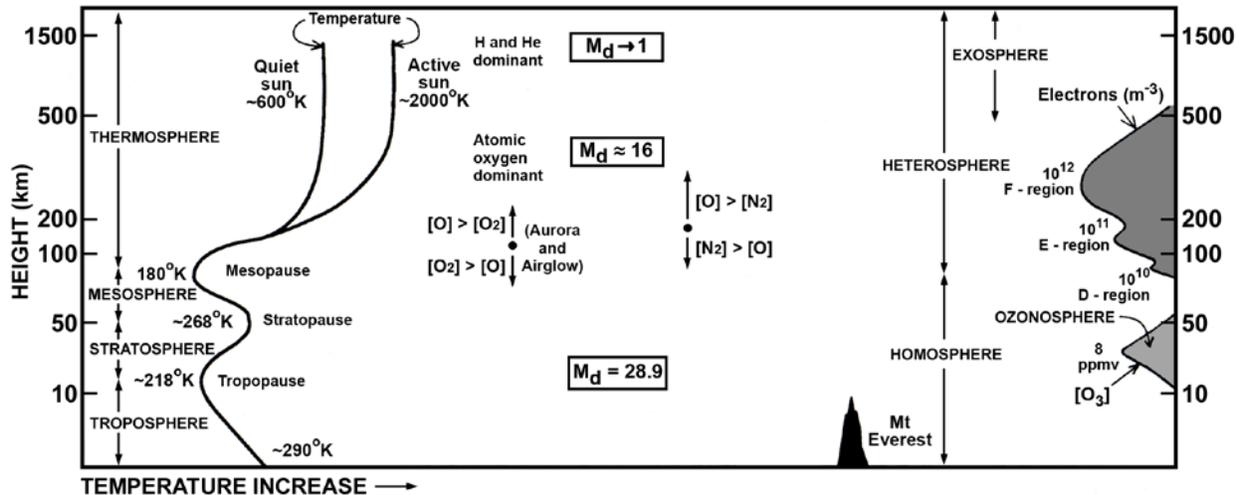


Figure 6. Left: representation of the vertical temperature structure in the atmosphere; center: variation of the mean molecular weight of dry air and a comparison of number densities of three atmospheric gases; right: the ionosphere and its relationship to the homosphere, heterosphere, and exosphere. Representative number densities of electrons are shown for three regions in the ionosphere. From Hobbs (2000, p. 27). [AtmosTo1500km.tif](#)

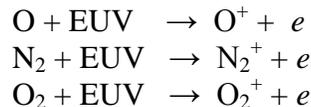
## 6.2 Ionospheric Regions

Several regions in the ionosphere are usually distinguished, the D-, E-, and F-regions. They are characterized by the photochemistry that occurs within them. The boundaries between them are not sharply defined.

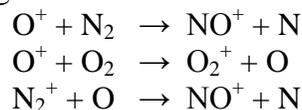
The *D-region* extends from approximately 60- to 90-km altitude. At this altitude range, the shorter wavelengths of EUV radiation have been greatly depleted (Fig. 5), but ionization of O<sub>2</sub>, N<sub>2</sub> and nitric oxide NO occurs at wavelengths > 112 nm. So-called Lyman-alpha radiation at 121.6 nm ionizes NO. Water cluster ions, both positive and negative, that incorporate various numbers of H<sub>2</sub>O molecules, form from the primary ionized species, NO<sup>+</sup> and O<sub>2</sub><sup>+</sup>, and these dominate below about 85 km. The daytime ion number density approaches 10<sup>10</sup> m<sup>-3</sup>, ten orders of magnitude less than the neutral gas number density. At night, ionization essentially ceases, the free electrons are quickly captured, and the D-region almost disappears. The most common manifestation of the D-region is the absorption of AM (amplitude modulated) radio waves during the day. After the sun sets, AM radio waves (frequencies 540-1700 kHz) can propagate to the higher-altitude F-region and bounce back to the surface, allowing long-range transmission of the signal, impossible by day.

The *E-region* extends from approximately 90- to 150-km altitude. Ion production rates are greatest for  $O_2^+$  and  $N_2^+$  in this altitude range, less for  $O^+$ , for which the maximum production rate occurs close to 180-km altitude. The ionizing radiation is principally Lyman-beta (102.5 nm), EUV radiation at less than 100 nm, and soft x-rays.  $NO^+$  is produced in two reactions facilitated by collisions,  $O^+ + N_2 \rightarrow NO^+ + N$  and  $N_2^+ + O \rightarrow NO^+ + N$ . A byproduct of ionization is so-called “hot” (highly energetic) photoelectrons, which can subsequently ionize more neutral species. After equilibrium is reached between production and loss (through recombination),  $NO^+$  and  $O_2^+$  become the dominant ions. Like the D-layer, the E-layer mostly disappears at night. Occasionally, a so-called *Sporadic E-layer*, a thin layer of high-density electrons will form in the same altitude range as the E-region. It is patchy and lasts from minutes to hours. It supports anomalous reflection of radio waves from 25-225 MHz.

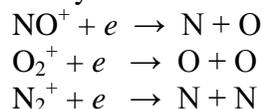
The *F-region* extends from approximately 150- to 400-km altitude. It makes possible short-wave radio communications over long distances. Atomic oxygen is the dominant neutral gas, followed by  $N_2$  and  $O_2$ , but all three gases are ionized by EUV radiation, primarily at wavelengths  $< 91.1$  nm, which is at the short-wave end of the Lyman Series. As detailed in Ondoh and Marubashi (2000, p. 78), these photoionization reactions are



EUV indicates absorption of extreme ultraviolet radiation. The symbol  $e$  represents a free electron. The primary ions and neutral particles then react through collisions to form new ions and atomic oxygen and nitrogen. This is called the ion-atom interchange.



Note that the first and last of the above set of three also occur in the E-region. Ion loss occurs by dissociative recombination, which yields neutral atomic species.



During the daytime, the F-region divides into two regions, the *F1* (150- to 250-km altitude) and the *F2* (250- to 400-km altitude). Electron density increases steadily with altitude in the F1 region. Below about 180-km altitude, the ion-atom interchange quickly consumes the  $O^+$  to make  $NO^+$  and  $O_2^+$  ions. The loss of these molecular ions occurs via recombination, and the rate of loss depends upon the square of the electron (ion) density. The peak in the  $O^+$  production rate is near 180 km, and so during the day a bulge in the electron density profile appears near this altitude. Well above 180 km, the number densities of  $N_2$  and  $O_2$  are much less than that of  $O$ , so that the third ion-interchange reaction above proceeds much more quickly than the other two, leaving  $O^+$  as the dominant ion. The loss rate well above 180 km is proportional to the electron density, as opposed to the square of the density. At night, the F1 region virtually disappears because the electron loss rate is high and the ion production rate is almost nil.

The F2 region (250- to 400-km altitude) features a peak in electron density at about 300 km. If photochemical equilibrium prevailed above this altitude, the electron density would continue to

increase, but plasma transport becomes important either by means of vertical diffusion or horizontal displacement by the neutral wind. The F2 region is in the heterosphere, where heavier particles seek hydrostatic equilibrium via diffusion. As a result, the number density of lighter particles decreases with altitude more slowly than that of heavier particles. This has an interesting effect on the ionospheric plasma, in that the heavier ions tend to settle downward while the much lighter electrons migrate upward. This generates an electrical field in the vertical, which tends to draw the electrons downward and the positive ions upward, an effect called *ambipolar diffusion*. The charges on these particles cause their number densities to decrease with altitude only half as fast as neutral species with the effect that Maxwell-Boltzmann velocity distributions remain valid for the plasma to much greater altitudes than for neutral particles.

The F2 region survives the night. With a collision frequency of less than one per second and a mean free path exceeding 1 km at 300 km, the electron density takes nearly 10 h to fall to 1/e of its sunset value, by which time the night is mostly over at low and midlatitudes. In addition, plasma diffuses downward at night into the F2 region from above.

The *topside ionosphere* is that region above the 300-km peak in electron density, where  $O^+$  is the dominant ion. It extends to above 1000-km altitude during the day but only to 600 km at night. Above that lies the *protonosphere*, where hydrogen and helium atomic ions ( $H^+$  and  $He^+$ ) dominate.

Figure 7 summarizes much of the preceding discussion. It displays a vertical profile of ion concentration, taken from the International Reference Ionosphere (IRI, see <http://modelweb.gsfc.nasa.gov/ionos/iri.html>). These examples, for solar noon and solar midnight, are typical during solar maximum. All ion (and electron) number densities are greater during daylight than at night, in the lower ionosphere by one or two orders of magnitude. The sum of ion number densities approximately equals the electron density at each altitude. Near 300 km, where electron density peaks, and for nearly 100 km above that,  $O^+$  is the only ion of consequence. However, it is not the dominant ion in the E region nor the D region just below. The day-night difference in the thickness of the topside ionosphere is evident. Its upper boundary, where the number density of  $H^+$  exceeds that of  $O^+$  is near 1000 km during the day and 600 km at night in this illustration.

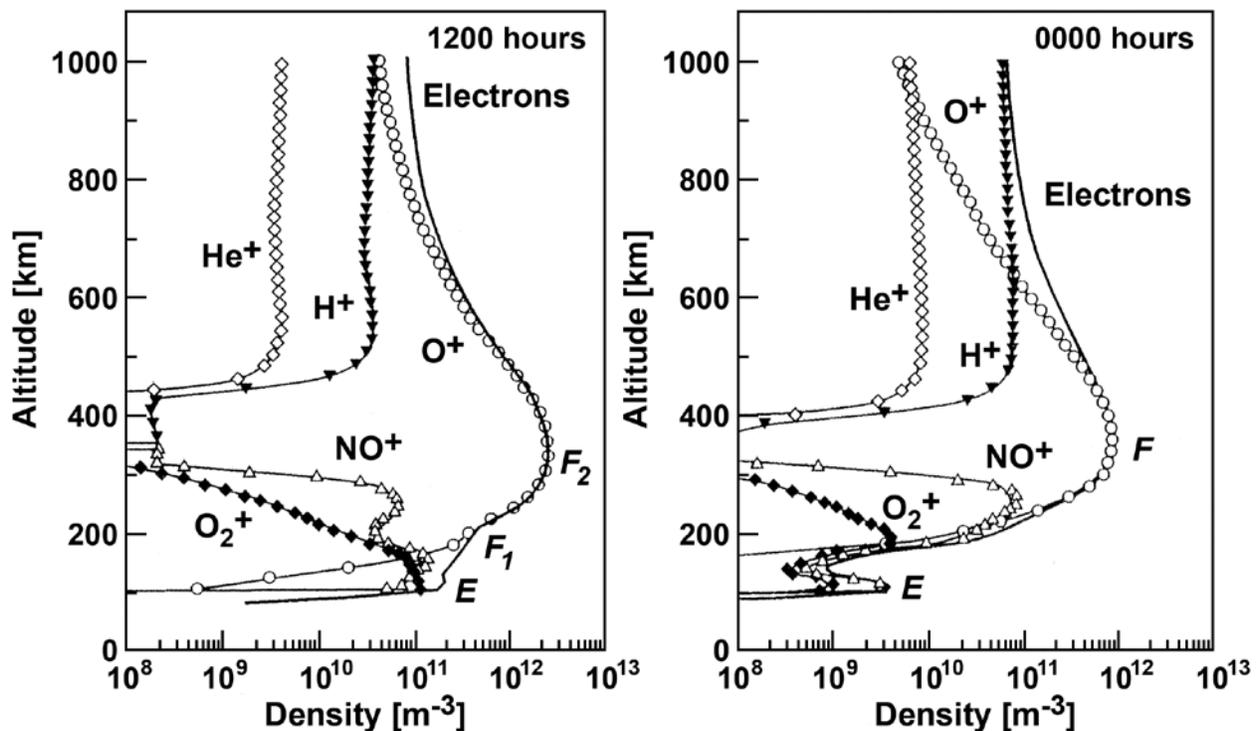


Figure 7. Vertical profile of ion and electron number density for solar noon (1200 hours, local time) and solar midnight (0000 hours). These conditions, typical of solar maximum, were extracted from the International Reference Ionosphere. The most important ions discussed in the text are labeled, as are the E, F1, and F2 regions. From Ondoh and Marubashi (2000, p. 89).

From Fig. 3.10, p. 89, *Science of Space Environment*, filename: tom/IonNumDensVsAlt.tif

### 6.3 Variability in the Ionosphere Caused by Solar Heating, Geomagnetic Activity, and Solar Variability

Before proceeding, it is helpful to review a few properties of the Earth's magnetic field and its effect on the atmosphere. The magnetic field arises within the Earth's core, apparently driven by the circulation of liquid metal. The Earth's magnetic poles are moving at speeds exceeding 40 km per year. In 2005 the magnetic poles were estimated by the Canadian Geological Survey to be near 82.7°N, 114.4°W and 63.1°S, 137.5°E. As these positions indicate, the two poles are not opposite each other, and their locations wander independently. At its simplest, the Earth's magnetic field resembles a dipole (Fig. 8), whose axis is tilted with respect to the Earth's rotation axis.

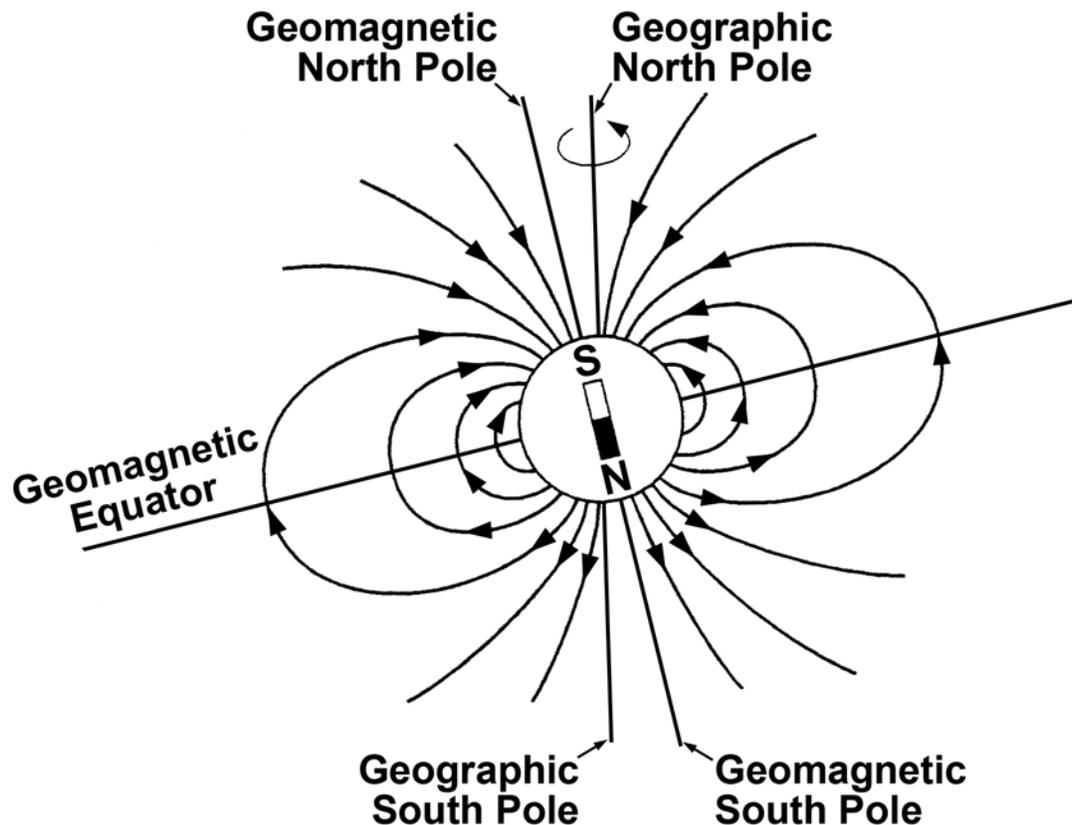


Figure 8. The Earth's magnetic field resembles a simple magnetic dipole. Lines of magnetic flux are shown by the arrows. The Earth's magnetic poles are currently offset about  $13^\circ$  from its axis of rotation.

From Fig. 1.7, p. 14 from *Science of Space Environment*. Filename: tom/GeomagFieldEarth.tif  
This figure is so common and so simple, it's probably not copyrighted.

The influence of the Earth's magnetic field extends throughout the *magnetosphere*, a region confined by the solar wind, a plasma flowing rapidly outward from the sun in all directions at typical speeds of  $400 \text{ km s}^{-1}$ . The *magnetopause* marks the boundary between the Earth's magnetic field and that carried by the solar wind. In the direction of the sun, the magnetopause lies at about 9-10 Earth radii ( $R_E$ ). Above the day-night boundary, the magnetopause extends to  $15 R_E$ . In the direction away from the sun, the magnetosphere extends in a long tail, to hundreds of  $R_E$ . Except during geomagnetic storms caused by solar activity, the magnetic field may be considered fixed below 600-km altitude.

Large changes in ion/electron densities, temperature, and wind occur daily and even hourly in the ionosphere. Such changes are caused by solar variability, in the form of both radiative and plasma output of the sun. The radiative output has a direct impact on the upper atmosphere, heating and ionizing the neutral gas. The plasma output from the sun first interacts with Earth's magnetosphere, drives currents in the magnetosphere and upper atmosphere, and appears to change Earth's magnetic field. This magnetospheric source of energy to the upper atmosphere is often referred to as geomagnetic activity because of the apparent change in Earth's magnetic field, but the latter is actually a response to, rather than the driver of ionospheric variability. The

sources of upper atmosphere change caused by solar variability are hardly noticed in the troposphere and stratosphere.

In the neutral atmosphere, the major forces acting on volumes of air are those of gravity, pressure gradient, forces arising from the Earth's rotation, and friction. In the ionosphere, the force acting on charged particles in electric and magnetic fields must be added, namely,  $q(\mathbf{E} + \mathbf{v} \times \mathbf{B})$ , where  $q$  denotes the charge, positive or negative,  $\mathbf{E}$  the electric field vector,  $\mathbf{B}$  the magnetic field vector, and  $\mathbf{v}$  the velocity vector of the charged particle. Yet another force must be considered. Because neutral particles are not affected by  $\mathbf{E}$  and  $\mathbf{B}$ , their velocities are different from those of ions and electrons. Collisions among charged and neutral particles result in *ion drag* and affect the velocities of both. One type of ion drag force is given by  $m_i v_{in} (\mathbf{v}_i - \mathbf{v}_n)$ , where  $m_i$  is the mass of the ion,  $v_{in}$  is the frequency of collisions between the ion and neutral particles,  $\mathbf{v}_i$  is the bulk ion velocity, and  $\mathbf{v}_n$  is the bulk velocity of neutral particles. At ionospheric altitudes from 100 to 300 km, this term is the important one because momentum exchange between neutrals and ions is significant. Electrons are not massive enough to exert significant drag. In the topside ionosphere, however, collisions among charged particles must be taken into account with drag terms similar to the one for ions and neutrals.

When the sun is quiet, the solar heating cycle drives upper atmospheric motions. The atmosphere on the sunlit side of the Earth expands, generating pressure gradient forces that move the air away from the sub-solar point toward the opposite location on the night side. The Earth's rotation deflects this flow in both hemispheres, and heating due to electrical fields in the magnetosphere and the rain of energetic electrons into the ionosphere at high latitudes can either retard or augment the basic anti-solar flow. According to Schunk and Nagy (2000, p. 28), the net effect of these processes is to decrease the anti-solar flow on the dayside and increase it both over the poles and on the night side. Wind speeds in the topside ionosphere commonly reach 100-300 m s<sup>-1</sup> during undisturbed solar conditions.

The sun follows an 11-year cycle of activity. An active sun occasionally produces a coronal mass ejection, resulting in huge clouds of energetic plasma with embedded sharp variations in the solar magnetic field propagating earthward in the solar wind. At the magnetopause, where the magnetic fields of sun and Earth meet, and especially if the solar magnetic field is directed southward, the two fields connect efficiently. Rapid fluctuations in electrical fields and currents in the magnetosphere and upper atmosphere are seen as changes in the Earth's magnetic field, and drive large variations in electron and ion number densities in the ionosphere, causing it to heat and expand upward on the side facing the sun. A strong solar storm of this type can cause havoc, as will be discussed in section 8. It also often enhances the northern and southern lights known as the Aurora Borealis and Aurora Australis, respectively, by injecting solar wind particles, channeled by the Earth's magnetic field, into the ionosphere at high latitudes. These energetic particles dissociate air molecules, and excite atoms and molecules to higher energy states. When these atoms and molecules relax back to their ground states, they emit colored light.

## 6.4 The Earth's Electrical Circuit and the Role of Thunderstorms

Measurements have established that a potential difference of 240 to 260 kV (kilovolts) exists between the ground and about 60-km altitude. The potential gradient defining this difference is  $100\text{-}150\text{ V m}^{-1}$  at the ground, decreases rapidly to a few  $\text{V m}^{-1}$  by 8 km and to near zero in the stratosphere. In fair weather, the ground is negatively charged (roughly  $5 \times 10^5\text{ C}$ ) and the atmosphere positively charged by an equal amount. Most of the positive charge is in the troposphere. The lower atmosphere is weakly conducting due to the presence of positive and negative ions created by cosmic rays and, near the Earth's surface, by the release of radioactive gases from the soil. Conductivity allows a small leakage current, one that would neutralize the positive charge in the atmosphere and the negative charge at the surface within tens of minutes. The leakage current is continuous, however, suggesting that the net positive atmospheric charge must somehow be replenished.

As described in MacGorman and Rust (1998, pp. 29-40), thunderstorms, 1500-2000 of them in progress at all times, generate a conduction current that completes the global electrical circuit. This current has been measured by high-flying aircraft above thunderstorms. It arises because the net positive charge in the upper parts of the storm drives positive ions upward in the clear air above the thundercloud. A simplified view of the global electrical circuit is shown in Fig. 9.

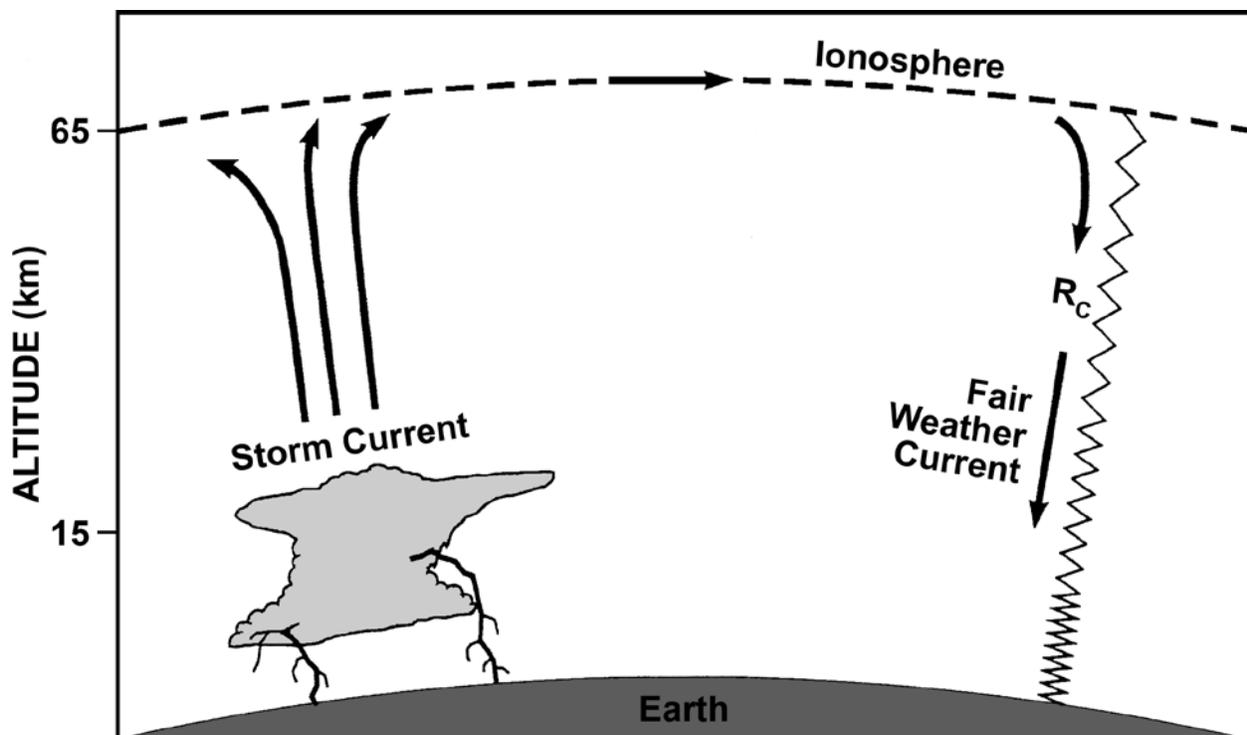


Figure 9. Simple conceptual model of the global electrical circuit. Thunderstorms drive positive charge upward toward the highly conductive ionosphere. The positive space charge in the fair-weather atmosphere then leaks back toward the surface against columnar resistance  $R_c$ , which increases toward the surface. Adapted from MacGorman and Rust (1998).

Insert Fig. 1.11, p. 31 from MacGorman and Rust. Filename: tom/ElectCircuitEarth.tif

## 7. MOMENTUM TRANSFER AT MOLECULAR AND LARGER SCALES

As a vehicle passes through the atmosphere, air molecules in its vicinity are disturbed. Some air molecules briefly stick to the object while many others pass around it, exerting aerodynamic forces opposing the motion. The size of these forces depends upon the shape and speed of the object, and the density, viscosity, and compressibility of the air. This section begins with a discussion of molecular viscosity, then defines two similarity parameters, Reynolds number and Mach number, both important in modeling aerodynamic forces and assessing their magnitudes. Drag refers to forces that resist the passage of a vehicle through air, and it changes dramatically as vehicle speed increases from tens to thousands of meters per second. The section concludes with a discussion of turbulence, the chaotic motion of volumes of air large enough to affect vehicle orientation and motion.

### 7.1 Viscosity

When air flows parallel to a smooth surface, its velocity increases from zero at the surface itself to a limiting value, representative of undisturbed flow, at some distance  $z$  above. Evidently, the surface is retarding the air flow in its immediate vicinity. For ideal gases like air, it is experimentally observed that the shearing stress  $\tau$  (force per unit area) exerted by the surface on the air is proportional to the change in air speed  $v$  observed over distance  $z$ .

$$\tau = \mu \frac{dv}{dz} \quad (28)$$

The constant of proportionality  $\mu$  is called the *dynamic viscosity*, a measure of the internal friction of gases and liquids, or, equivalently, a measure of resistance to internal momentum transfer. For ideal gases, dynamic viscosity satisfies the following proportionality:

$$\mu \propto \rho L \langle v \rangle. \quad (29)$$

The variables on the right have appeared before: density, the mean-free path between molecular collisions, and mean molecular speed. From Eq. (5), we recall that  $\langle v \rangle$  is proportional to the square root of absolute temperature. The expression just before Table 1 indicates that  $L$  is inversely proportional to number density, and hence mass density, so that the product  $\rho L$  depends only upon the intrinsic properties of air molecules (e.g., effective collision diameter) but not the density or temperature. A formula by D. M. Sutherland makes the temperature dependence explicit:

$$\mu = \mu_0 \left( \frac{T}{T_0} \right)^{\frac{3}{2}} \left( \frac{T_0 + c}{T + c} \right). \quad (30)$$

For dry air,  $\mu_0 = 1.827 \times 10^{-5} \text{ kg m}^{-1}\text{s}^{-1}$ ,  $T_0 = 291.15\text{K}$ , and  $c = 120.0\text{K}$ . At the freezing point of water,  $273.15 \text{ K}$ ,  $\mu = 1.736 \times 10^{-5} \text{ kg m}^{-1}\text{s}^{-1}$ . The dynamic viscosity of air increases with temperature. This formula has good accuracy for temperatures  $< 500\text{K}$ , and so can be applied throughout the homosphere.

## 7.2 Similarity Parameters: Reynolds Number and Mach Number

Aerodynamicists use *similarity parameters* to model airflow around various objects. Because it is both difficult and expensive to examine properties of airflow around full-sized objects, they resort to building much smaller objects of the same shape, and they test them in wind tunnels. If the similarity parameters characterizing flow around the object are nearly the same for the scale model as for the full-sized object, then aerodynamicists have confidence that the relative importance of the various forces are being correctly modeled. Two similarity parameters of interest here are the Reynolds number and the Mach number.

The dimensionless Reynolds number  $Re$  is essentially the ratio of the inertial force (involved in accelerating or decelerating air) to the viscous force (arising from spatial gradients in the shearing stress, the latter defined by Eq. 28). Its definition is  $Re = \rho v \ell / \mu$ , where  $\rho$  is the density,  $v$  is the speed of the air,  $\ell$  is the characteristic dimension of the object over which the air passes, and  $\mu$  is the dynamic viscosity. For air passing around a sphere,  $\ell$  is the diameter of the sphere. The Reynolds number is important because the character of the flow around an object can change many times as this number increases from less than unity to more than  $10^5$ . For  $Re < 1$ , the flow is laminar; it passes smoothly around the object and reconnects smoothly downstream. At sea level, this is observed for small objects (millimeter-sized) and very low airspeeds ( $1 \text{ cm s}^{-1}$ ), for example fog droplets falling in still air. Both the size and speed of aircraft and spacecraft guarantee that  $Re \gg 1$ , and the flow of air around these vehicles is turbulent, containing many eddies and vortices. As a similarity parameter, the Reynolds number is relevant for steady, incompressible flows.

The *boundary layer* is a useful concept for describing the flow of air over a rigid surface. Consider high-speed air meeting a stationary wing in a wind tunnel. As the air comes in contact with the wing, some of the molecules temporarily stick to the surface. As these molecules rejoin the air stream, they retard the motion of air molecules very close to the wing. The exchange of momentum among molecules very close to the wing sets up a boundary layer, within which the air velocity goes from zero at the wing surface to its undisturbed value. This causes a shearing stress (Eq. 28), which exerts drag on the wing. The thickness of the boundary layer increases from front to back of the wing, ranging from perhaps a millimeter near the forward edge to centimeters toward the trailing edge.

Boundary-layer flow is initially laminar but quickly becomes turbulent with increasing distance from the forward edge of the wing. Over a smooth surface such as a flat plate or a thin wing, the Reynolds number at which the transition from laminar to turbulent flow takes place is roughly  $10^5$ , and the location of the transition may be a few centimeters from the forward edge. The drag on the wing increases in the boundary layer as the transition from laminar to turbulent flow occurs. Over a wing, the boundary layer may separate from the surface, thus effectively

changing the shape over which the undisturbed airstream passes. Reversed flow (backwash) between the separated boundary layer and the wing surface adds to the drag.

The other similarity parameter is Mach number  $M$ , the ratio of vehicle speed  $v$  to the speed of sound  $c$ :  $M = v/c$ . The speed of sound in dry air is  $c = (\gamma R_d T)^{1/2}$ , where  $\gamma = c_p/c_v$  is the ratio of specific heats at constant pressure and constant volume, respectively, for dry air ( $\gamma = 1.4$ );  $R_d$  is the gas constant for dry air ( $287 \text{ J kg}^{-1} \text{ K}^{-1}$ ); and  $T$  is absolute temperature [K]. As a similarity parameter, Mach number is relevant for compressible flows, which occur when a vehicle approaches or exceeds the speed of sound.

### 7.3 Drag

Aircraft experience a drag force, which must be overcome by a propulsive force to keep the vehicle in flight. Spacecraft entering the atmosphere use the drag force to decelerate. The drag force is complicated. Computer models are used to calculate drag forces accurately. Only the most fundamental concepts are considered here.

At high Reynolds numbers ( $Re > 1000$ ), a rough estimate of the drag force  $\mathbf{F}_d$  exerted on a body moving through a fluid (e.g., a Boeing 737 moving through air) is given by

$$\mathbf{F}_d = -\frac{1}{2} \rho v^2 A C_d \hat{\mathbf{v}} \quad (31)$$

where  $\rho$  is the air density,  $v$  is the air speed,  $A$  is an area representing the size of the obstacle presented to the oncoming air,  $C_d$  is the dimensionless drag coefficient, and  $\hat{\mathbf{v}}$  is the unit vector indicating the direction of the air flow. The estimate is rough because of the ambiguity in specifying the representative area and the fact that the drag coefficient often depends upon the Reynolds number. For aircraft,  $A$  is often taken to be the square of the chord of the airfoil (the distance between the leading and trailing edges of the wing), and a typical range of  $C_d$  is 0.02 to 0.04.

Several kinds of drag affect vehicles in flight. One category is called *parasitic drag*, which has three major components: 1) *Skin friction* has already been discussed; it arises from the viscosity of air acting within the thin boundary layer close to the airframe. 2) *Form or pressure drag* occurs because the aircraft generates sharp velocity gradients of size comparable to parts of the aircraft itself and corresponding local pressure gradients that retard aircraft motion. Boundary layer separation (see preceding section) contributes to form drag. This type of drag depends essentially upon the shape of the aircraft. 3) *Interference drag* occurs at the junction of airframe structures, for example, the junction of wings with fuselage or jet engines with wings.

Turbulence generated at such junctions increases drag.

Another category of drag is *lift-induced drag*. During aircraft flight, counter-rotating vortices are produced at the wingtips. These vortices produce a downwash of air behind the wings, strongest near the wingtips, decreasing toward the fuselage. The downwash increases the angle of attack of the wings, thereby inducing additional drag. Lift-induced drag is greatest at lower speeds and a higher angle of attack; it is least at high speeds and low angle of attack.

Induced drag tends to dominate other forms of drag at low airspeeds. Form drag increases steadily with airspeed and eventually dominates at higher subsonic speeds.

When airspeed approaches the speed of sound, *wave drag* must be considered. Wave drag is substantial; it arises from shock waves (compression waves) that form on surfaces of the aircraft at speeds approaching or exceeding the speed of sound. In supersonic flight, these waves contain considerable acoustic energy trailing tens of kilometers behind the aircraft—energy extracted from the aircraft in the form of drag. Wave drag depends upon Mach number  $M$ . For high supersonic speeds,  $3 < M < 5$ , aerodynamic heating becomes important for aircraft design. For *hypersonic speeds* ( $M > 5$ ), air chemistry must be considered because the energy imparted to the air at these speeds can break chemical bonds, creating free electrons and ions. Manned spacecraft entering the atmosphere at  $M \sim 25$  heat the surrounding air enough to produce a plasma, and the spacecraft skin must be insulated against high temperatures.

A comprehensive textbook on the concepts covered in sections 7.2 and 7.3 is Anderson (2007). Volume 1 of this encyclopedia also covers these concepts in much greater detail.

#### **7.4 Atmospheric Turbulence**

Though the passage of a vehicle rapidly through air generates turbulence in the immediate vicinity, with drag being the inevitable result, we have so far assumed that the air is initially in smooth and uniform flow. But what if the air is turbulent, full of eddies and swirls at scales matching the size of the vehicle? This can cause sudden, seemingly random vehicle accelerations. Airline passengers experience this as “bumpy air” or “clear-air turbulence.” In addition, torques could change the vehicle’s orientation.

As noted earlier, turbulence is confined to the homosphere. Turbulence is caused by strong shear in the velocity vector of air, that is, significant changes in the direction and/or speed of the wind over short distances. The shear is generated in three ways: mechanically, by convection, and by the constant cascade of kinetic energy from larger- to smaller-scale eddies.

Mechanical turbulence arises from shear in the mean wind. Wind swirling around fixed obstacles, such as tall buildings, a thicket of trees, a snow fence, or rough terrain, generates shear. Any surface roughness retards the wind flow near the ground and generates shear in the vertical direction. Vertical shear also arises in the free atmosphere, far removed from the effects of surface friction. It can become pronounced when the temperature stratification is stable, particularly within inversions. The stronger the inversion, the greater the suppression of vertical motion, as noted in section 5.1. When vertical shear is substantial, vertical motion can bring air with higher momentum downward to altitudes where air momentum is less. This mixing of air with different momentums causes turbulence. At the ground this is experienced as wind gusts.

Convective turbulence arises in thermals, bubbles of buoyant air rising from heated patches of ground, and in cumulus clouds from small to storm-size. Section 5.1 described how the release of latent heat within a convective cloud can easily accelerate parcels of air to tens of meters per second. Aircraft passing from clear air that has negligible vertical speed into a towering cumulus

cloud where the vertical speed is  $10 \text{ m s}^{-1}$  will receive a sudden jolt because of the shear in the vertical wind along a horizontal direction.

The cascade of energy from larger to smaller scales—all the way down to molecular scales—is never-ending in the atmosphere, and the energy associated with turbulence lies in between. Shear generates turbulence when small instabilities develop in the flow, grow into waves, and eventually break. Buoyant energy is created at cloud scales through the release of latent heat in updrafts or evaporative cooling in downdrafts. Shear at the edges of convective updrafts and downdrafts generates turbulence. Section 5.3 mentioned how gravity waves, generated in the troposphere or lower stratosphere at scales tens of kilometers across, propagate upward while growing in amplitude, finally breaking and causing turbulence at high altitudes.

A useful parameter for determining whether the vertical shear in the horizontal wind will generate turbulence is the Richardson number  $Ri$ , defined for dry air by

$$Ri = \frac{\frac{g}{\bar{\theta}} \frac{\partial \bar{\theta}}{\partial z}}{\left( \frac{\partial \bar{v}}{\partial z} \right)^2} \quad (32)$$

where  $\bar{v}$  is the average horizontal wind speed, and  $\bar{\theta}$  is a variable not previously introduced, the mean potential temperature:

$$\theta = T \left( \frac{p_0}{p} \right)^{R_d/c_p} \quad (33)$$

$p_0$  is a reference pressure (1000 hPa), and the ratio  $R_d/c_p = 0.286$  is dimensionless. Air parcels undergoing dry adiabatic ascent or descent conserve their potential temperature. To see how this relates to the discussion of section 5.1, take the material derivative of Eq. (33) with respect to

altitude, set  $\frac{d\theta}{dz} = 0$ , and then solve for  $\frac{dT}{dz}$  while making use of the ideal gas law (Eq. 16),

and the assumption of hydrostatic equilibrium (Eq. 8). The result is  $\frac{dT}{dz} = -\frac{g}{c_p}$ , which is the

definition of the dry adiabatic lapse rate seen before and also a statement about how the temperature of a parcel forced upward or downward will change with altitude. For dry air, if  $\frac{\partial \bar{\theta}}{\partial z} > 0$ , the temperature stratification is said to be stable. This quantity can increase to  $0.1 \text{ K m}^{-1}$  in strong inversions.

The interpretation of the Richardson number is straightforward. A measure of stability is in the numerator, and a measure of vertical shear is in the denominator. If the atmosphere is very

stable (large numerator), substantial shear can be maintained without dynamical instability causing wave development along the inversion, thereby hindering the development of turbulence. If the atmosphere is less stable, that is, less able to damp vertical motions (smaller numerator), and the shear is very large (large denominator), then instabilities develop and turbulence follows. If the Richardson number is small ( $< 0.25$ ), shear carries the day, and turbulence is likely. If  $Ri > 1$ , stability dominates, and turbulence is unlikely. Laminar flow usually becomes turbulent when  $Ri$  drops below 0.25, but turbulent flow may remain turbulent as  $Ri$  rises above 0.25, almost as high as 1.0.

## **8. HAZARDS TO AIR AND SPACE TRAVEL POSED BY ATMOSPHERIC COMPOSITION AND STRUCTURE**

This section summarizes a number of hazards to people and equipment aboard aircraft and spacecraft. The hazards stem directly from the atmospheric properties and processes discussed in previous sections. Except for shear-induced turbulence, we shall not discuss hazards posed by tropospheric weather; that is the province of a companion article on meteorology.

### **8.1 Turbulence**

Travelers on commercial airlines inevitably experience turbulence. Usually it is a nuisance. It rarely causes structural damage to aircraft, but personal injuries occur every year, mostly to passengers not wearing their seatbelts or flight attendants working in the cabin. Commercial pilots often alter flight altitude in order to avoid bumpy air, usually resulting in greater fuel consumption. Turbulence can occur anywhere within the troposphere, stratosphere, or mesosphere, but the associated kinetic energy is greatest in the upper troposphere and lower stratosphere, where relatively high-density air and large shears combine. Spacecraft launched vertically must have precise attitude control for successful insertion into orbit. Launches are delayed if vertical shear of the horizontal wind threatens to affect spacecraft attitude after liftoff.

### **8.2 Hazards Posed by Solar Activity and Changing Magnetic Fields**

All other hazards discussed in this section are related in one way or another to the influx of high-energy particles from the sun during various types of solar activity, and from the magnetosphere during geomagnetic storms and substorms. Geomagnetic storms can last for days; they occur when the sun's magnetic field, as carried by the solar wind, turns southward and connects with the Earth's magnetic field, which points northward. Substorms are more common. On average, they occur several times a day, last two to three hours, and affect auroral activity. They are more intense during geomagnetic storms.

#### **8.2.1 The Electric Power Industry**

Even though the electric power industry is not directly related to air or space travel, its vulnerability to geomagnetic storms is so large that it deserves at least brief mention. Sudden changes in the Earth's magnetic field induce currents in long conductors: railroad tracks, gas and oil pipelines, and transmission lines. These geomagnetically induced currents resemble direct

currents. They interfere with transmission of electrical power (60 cycles per second alternating current). These current surges occasionally damage or destroy high-voltage transformers that serve large areas, cost many millions of dollars, and take many months to replace. The March 1989 solar storm caused widespread power outages in Quebec. Because Hydro-Quebec was exporting 1325 MW of power to the U.S. at the time, the sudden loss of power in the Northeast States put severe strain on the electrical grid, almost to the point of catastrophic failure. The interconnectedness of the power grid and increasing long-distance transmission of power between different parts of the country make the U.S. more vulnerable to a similar storm now. A magnetic storm equal to the most severe experienced in the past century could cause failure of, or severely damage, over a hundred high-voltage transformers, leading to widespread and lengthy power outages. The ramifications could be catastrophic: immediately no street or home lighting, phones, radio, TV, heating or air conditioning and no cooking on electric ranges; within a day or two no potable water due to pump failure, loss of perishable foods and medications, no home natural gas supply; still later, no sewage disposal, transportation, or fuel resupply (NRC, 2008, p. 77).

The remaining hazards all pertain to air and space travel.

### 8.2.2 Spacecraft

The influx of high-energy plasmas into the magnetosphere and current flow during solar storms cause rapid heating of hundreds of degrees Kelvin. In response, the atmosphere expands upward. At a given altitude, particle number densities may increase by a factor of five or more, causing significant drag not only on satellites in low earth orbit (LEOs at 300- to 500-km altitude) but also on the thousands of space debris objects larger than a baseball. The U.S. Space Command tracks these objects. Because of their high velocity ( $7\text{-}8\text{ km s}^{-1}$ ), they pose a hazard to operational satellites and manned spacecraft. Severe solar storms perturb orbits enough that the Space Command can temporarily lose track of more than a thousand of these objects. Orbits must be recalculated after the atmosphere settles back to its undisturbed state.

Static discharges from both the exterior skin of spacecraft and insulating materials inside the spacecraft can severely damage mission components. Magnetic substorms, episodic injections of energetic plasma into the lower magnetosphere from the tail of the magnetosphere (on the side of the Earth opposite the sun), occur more frequently (several per day) than solar magnetic storms, often independently of them, and subside more quickly (usually in less than an hour). During both magnetic substorms and the less frequent solar geomagnetic storms, the outer skin of spacecraft can accumulate electrons more quickly than they can bleed off. The result is an electrostatic (spark) discharge.

The energy carried by electrons and ions is measured in electron volts (eV). High-energy particles (MeV for electrons, tens of MeV for ions) can penetrate spacecraft. For example, a 3-MeV electron can penetrate 6 mm of aluminum, a material commonly used in spacecraft construction. When energetic electrons accumulate in dielectric (insulating) materials inside the spacecraft, such as those used around coaxial cables, charge buildup can eventually trigger electrical discharge. Though rare, these can render electronic components inoperable.

Penetrating high-energy ions can cause on-board computer memory problems. They can also shorten the lifetime of solar panels, essential for the power generation, by several years.

### **8.2.3 Radiation Hazards to Humans**

In keeping with the 1000-km altitude limit for this article, this section presumes that astronaut exposure to radiation lies within the magnetosphere. This is certainly true for manned, orbiting space stations. Hazards within the magnetosphere are great enough. Protons from the sun at energies above 10 MeV can penetrate space suits and reach the skin or the lens of the eye. Prolonged exposure can lead to erythema or cataracts. Since the beginning of the space age, five solar storms have occurred with particle fluxes sufficiently energetic to jeopardize the health of crews, even behind normal spacecraft shielding. Because the flux of high-energy particles can change by many orders of magnitude within minutes, advanced warning of solar storms is vital, especially if crews are working outside the spacecraft. Because of the configuration of the Earth's magnetic field, near-space operations below 1000-km altitude are less vulnerable to harmful radiation at lower than at higher latitudes.

The Van Allen radiation belts consist of two torus-shaped belts of high energy protons and electrons. The inner belt of the two stretches from 700- to 10,000-km altitude. Typical proton energies are 100 MeV; electron energies are hundreds of keV. The Van Allen belts are discussed in a separate article. [\[Where?\]](#)

### **8.2.4 Commercial Aviation**

Commercial aviation is occasionally affected by magnetic storms but only at high latitudes. The most serious effect is prolonged blackouts of high-frequency radio communications essential for message passing among pilots and air traffic controllers. GPS navigation systems can also be affected (next section). Normally, only over-the-pole routes are affected by magnetic storms to the extent that commercial airlines will change flight routes. Radiation exposure on a round-trip, over-the-pole flight is the equivalent to two x-rays—a minimal exposure perhaps, but more than a person on the ground would experience. Finally, flights above 10,000 m at high latitudes are very likely to be in the lower stratosphere, where ozone concentrations are greater. Ozone irritates the lungs and causes respiratory symptoms.

### **8.2.5 Industries Based on the Global Positioning System**

The Global Positioning System is one of several global navigation satellite systems in operation or planned. A constellation of satellites in 12-h orbits at a radius of 26,600 km transmits signals at two L-band frequencies, 1.5754 GHz (L1 signal) and 1.2276 GHz (L2 signal) with precise timing information. Receivers on the ground, which always have multiple GPS satellites in view, process this information, thereby obtaining precise three-dimensional location. Computation of position is complicated because the atmosphere refracts the signal (slows it down) for a variety of reasons. The most difficult correction relates to rapid fluctuations in electron density in the ionosphere occasioned by magnetic disturbances. Position errors are considerably greater in the vertical than the horizontal, as large as 50 m, which is unacceptable if one is trying to land an airplane in the fog.

The U.S. Federal Aviation Administration (FAA) operates a Wide Area Augmentation System, which covers almost all U.S. airspace. Relying on GPS, the WAAS allows reduction in vertical and horizontal separation between aircraft that have GPS navigation systems. Ionospheric disturbances, which perturb the electron content, degrade the accuracy of navigation information. The FAA disables WAAS during severe disturbances and for up to 8 h after they cease if wide areas are affected. During October and November 2003, this occurred for a total of 30 h. Dual-frequency GPS avionics would avoid this limitation, but airlines won't have this equipment for another decade.

### 8.2.6 Communications

Radio waves (3 Hz to 300 GHz) are at the heart of wireless communications. They carry the signal for AM and FM radio and television; and messages from air to ground, ship to shore, and satellite to ground. They are vital in national defense applications and civilian emergency response. The ionosphere affects radio signals in different ways. The *plasma frequency* is proportional to the square root of the number density of electrons. For an unmagnetized plasma and at frequencies below the plasma frequency, the ionosphere strongly refracts incident radio signals, enabling over-the-horizon communications, but not without damping of the incident signal. At frequencies above the plasma frequency, the ionosphere transmits radio waves but not without some refraction, less at higher frequencies than at lower frequencies. Ionospheric refraction causes tracking radars to “see” space objects in other than their true positions. The ionosphere is also a dispersive medium, in which signals at different frequencies travel at slightly different velocities. In wideband communications (covering many frequencies), differential delay caused by dispersion can be a problem.

As Section 6.3 emphasized, the ionosphere is subject to large and sudden changes in electron density during solar and magnetic storms. This affects the plasma frequency and, in turn, ionospheric reflection, transmission, and attenuation of radio waves. Blackouts in high-frequency communications at high latitudes during solar storms have already been mentioned. Currents induced by magnetic field variations in long conductors have caused service disruptions in Atlantic fiber-voice cables.

Plasma bubbles with anomalously low or high number densities cause scintillation in satellite-to-ground signals, especially between  $\pm 20^\circ$  of magnetic latitude and around the auroral oval. Equatorial scintillation begins around sunset and continues until an hour or two past local midnight. Auroral scintillation can occur at any time. Scintillation becomes more intense when solar activity rises. Scintillation affects both signal strength (amplitude) and phase. Reduction in signal strength can occur from destructive interference at wave fronts when the signal takes multiple paths. Very high frequencies (VHF, 30-300 MHz) suffer the most from scintillation; L-band (1-2 GHz), employed by GPS, is moderately affected. This is of special concern to the military because guided missile systems cannot afford to lose their lock on GPS signals, even momentarily.

The hazards discussed in this section are summarized in Table 4, adapted from Lanzerotti, (2007, p. 256).

Table 4. Adverse effects of solar and magnetic storms

---

Ionospheric Variations

Induction of electrical currents in the Earth

Long communication cables

Gas and oil pipelines

Wireless signal reflection, propagation, and attenuation

Commercial radio and television

Local and national safety and security entities

Aircraft communications

Communication satellite signal interference, scintillation

Commercial telecommunications and broadcast

Solar Radio Bursts

Excess noise in wireless communications systems

Interference with radar and radio receivers

Charged Particle Radiation

Solar cell damage

Damage and failure of semiconductor devices

Faulty operation of semiconductor devices

Spacecraft charging, surface and interior materials

Aircraft communications avionics

Atmosphere

Drag on satellites in low-Earth orbit (LEOs)

---

**ACKNOWLEDGMENTS**

I thank John C. Osborn for formatting, drafting, or redrafting most of the figures. Tim Fuller-Rowell, George C. Reid, Raymond G. Roble, and Susan Solomon, helped in suggesting content

for this article and pointed me to many primary references. **Credit those who reviewed the manuscript or parts of it.**

## REFERENCES

- Anderson, JD, Jr. *Fundamentals of Aerodynamics*, 4<sup>th</sup> edition, McGraw-Hill Higher Education, New York, 2007; 1008 pp.
- Bohren CF and Albrecht BA. *Atmospheric Thermodynamics*. Oxford University Press, New York, 1998; 402 pp.
- Brasseur GP, Orland JJ, and Tyndall GS (eds.). *Atmospheric Chemistry and Global Change*. Oxford University Press, New York, Chapter 5, “Trace Gas Exchanges and Biogeochemical Cycles,” 1999; 654 pp.
- Hobbs PV. *Introduction to Atmospheric Chemistry*. Cambridge University Press, New York, Chapter 8, “Tropospheric Chemical Cycles,” 2000; 143-152.
- Fritts DC. Gravity wave forcing and effects in the mesosphere and lower thermosphere. In *The Upper Mesosphere and Lower Thermosphere: A Review of Experiment and Theory*, Johnson RM and Killeen TL, eds. American Geophysical Union, Washington, D.C., 1995; 89-100.
- ICAO. *Manual of the ICAO Standard Atmosphere: Extended to 80 kilometres*. International Civil Aviation Organization, CD-ROM (7488-CD), third edition, 1993; ISBN 92-9194-004-6.
- IPCC. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* [Solomon S, Qin D, Manning M, Chen Z, Marquis M, Averyt KB, Tignor M, and Miller HL, eds]. Cambridge University Press, New York, 2007; 996 pp.
- Lanzerotti LJ: Space weather effects on communications. In *Space Weather–Physics and Effects*, Bothmer V and Daglis IA, eds., Springer-Verlag, New York, and Praxis Publishing, Ltd, Chichester, UK, 2007: 247-265.
- MacGorman DR and Rust WD. *The Electrical Nature of Storms*. Oxford University Press, New York, 1998; 422 pp.
- NRC: *Severe Space Weather Events. Understanding Societal and Economic Impacts. A workshop report*. Space Studies Board, Division on Engineering and Physical Sciences, National Research Council, National Academies Press, Washington, D.C.; 2008, 131 pp.
- Ondoh T and Marubashi K (eds.). *Science of Space Environment*. Ohmsha/ IOS Press Inc., Burke, Virginia, 2000; 302 pp.
- Roble RG. Thermosphere. In *Encyclopedia of Atmospheric Sciences* Holton J, Pyle J, and Curry J, eds. Academic Press, New York, 2003; 2282-2290.

Salby ML. *Fundamentals of Atmospheric Physics*. Academic Press, New York, 1996; 8-15.

Schunk RW and Nagy AF. *Ionospheres. Physics, Plasma Physics, and Chemistry*. Cambridge University Press, New York, 2000; 554 pp.

Stull RB. *Meteorology for Scientists and Engineers*. Brooks/Cole (Division of Thompson Learning), Pacific Grove, California.

*U.S. Standard Atmosphere, 1976*, prepared jointly by the National Oceanic and Atmospheric Administration, National Aeronautics and Space Administration, and the U.S. Air Force, U.S. Government Printing Office, Washington, D.C., 1976; 227 pp.

Version of 5:00 p.m., Thu 23 Jul 2009. 18,658 words.

Article ID: eae31

Thomas W. Schlatter

NOAA Earth Systems Research Laboratory

David Skaggs Research Center

Mail code: R/GSD

325 Broadway

Boulder, CO 80305 USA

Telephone: 303 497-6938

Fax: 303 497-6821

Email: [Tom.Schlatter@noaa.gov](mailto:Tom.Schlatter@noaa.gov)

Filename for this draft: eae31MS.doc (Microsoft Word document)