الجامعة التكنولوجية

قسم الهندسة الكيمياوية

المرحلة الاولى

فيزياء البيئة

أ.م. د. فلك اسامة

Environmental physics:

By dr. falak O. Abas

Essential of environmental physics:

Physics has always been concerned with understanding the natural environment, and, in its early days, was often referred to as "Natural Philosophy." Environmental Physics, as we choose to define it, is the measurement and analysis of interactions between organisms and their environments.

To grow and reproduce successfully, organisms must come to terms with the state of their environment. Some microorganisms can grow at temperatures between −6 and 120 ◦C and, when they are desiccated, can survive even down to −272 ◦C.

Higher forms of life on the other hand have adapted to a relatively narrow range of environments by evolving sensitive physiological responses to external physical stimuli. When environments change, for example because of natural variation or because of human activity, organisms may, or may not, have sufficiently flexible responses to survive.

The physical environment of plants and animals has five main components which determine the survival of the species:

(i) The environment is a source of radiant energy which is trapped by the process of photosynthesis in green cells and stored in the form of carbohydrates, proteins, and fats. These materials are the primary source of metabolic energy for all forms of life on land and in the oceans;

(ii) The environment is a source of the water, carbon, nitrogen, other minerals, and trace elements needed to form the components of living cells;

(iii) factors such as temperature and day length determine the rates at which plants grow and develop, the demand of animals for food, and the onset of reproductive cycles in both plants and animals;

(iv) The environment provides stimuli, notably in the form of light or gravity, which are perceived by plants and animals and provide frames of reference both in time and in space. These stimuli are essential for resetting biological clocks, providing a sense of balance, etc.;

(v) The environment determines the distribution and viability of pathogens and parasites which attack living organisms, and the susceptibility of organisms to attack.

To understand and explore relationships between organisms and their environment, the biologist should be familiar with the main concepts of the environmental sciences. He or she must search for links between physiology, biochemistry, and molecular biology on the one hand and atmospheric science, soil science, and hydrology on the other. One of these links is environmental physics. The presence of an organism modifies the environment to which it is exposed, so that the physical stimulus received *from* the environment is partly determined by the physiological response *to* the environment. When an organism interacts with its environment, the physical processes involved are rarely simple and the physiological mechanisms are often imperfectly understood.

Fortunately, physicists are trained to use Occam's Razor when they interpret natural phenomena in terms of cause and effect: i.e. they observe the behavior of a system and then seek the simplest way of describing it in terms of governing variables. Boyle's Law and Newton's Laws ofMotion are classic examples of this attitude. More complex relations are avoided until theweight of experimental evidence shows they are essential. Many of the equations discussed in this book are approximations to reality which have been found useful to establish and explore ideas. The art of environmental physics lies in choosing robust approximations which maintain the principles of conservation for mass, momentum, and energy.

Properties of Gases and Liquids

The physical properties of gases influence many of the exchanges that take place between organisms and their environment. The relevant equations for air therefore form an appropriate starting point for an environmental physics text. They also provide a basis for discussing the behavior of water vapor, a gas whose significance in meteorology, hydrology, and ecology is out of all proportion to its relatively small concentration in the atmosphere. Because the evaporation of water from soils, plants, and animals is also an important process in environmental physics, this chapter reviews the principles by which the state of liquid water can be described in organisms and soil, and by which exchange occurs between liquid and vapor phases of water.

Gases and Water Vapor

Pressure, Volume, and Temperature

The observable properties of a gas such as temperature and pressure can be related to the mass and velocity of its constituent molecules by the Kinetic Theory of Gases which is based on Newton's Laws of Motion. Newton established the principle that when force is applied to a body, its momentum, the product of mass and velocity, changes at a rate proportional to the magnitude of the force. Appropriately, the unit of force in the Système Internationale is the Newton and the unit of pressure (force per unit area) is the Pascal—from the name of another famous natural philosopher.

The pressure p which a gas exerts on the surface of a liquid or solid is a measure of the rate at which momentum is transferred to the surface from molecules which strike it and rebound. Assuming that the kinetic energy of all the molecules in an enclosed space is constant and by making further assumptions about the nature of a *perfect* gas, a simple relation can be established between pressure and kinetic energy per unit volume. When the density of the gas is ρ and the mean square molecular velocity is $v2$, the kinetic energy per unit volume is $\rho v2/2$ and Equation (2.3), which is a statement of the *Ideal Gas Law*, is sometimes used in the Form

$$p = \rho RT / M$$

Obtained by writing the density of a gas as its molecular mass divided by its molecular volume, i.e.

$$\rho = M / V_{\mathrm{m}}.$$

For unit mass of any gas with volume V, $\rho = 1/V$ so Eq. (2.5) can also be written in the form

$$pV = RT / M.$$

Equation (2.7) provides a general basis for exploring the relation between pressure, volume, and temperature in unit mass of gas and is particularly useful in four cases:

1. Constant volume—p proportional to T ,

2. Constant pressure (isobaric)—V proportional to T ,

3. Constant temperature (isothermal)—V inversely proportional to p,

4. Constant energy (adiabatic)—p, V, and T may all change.

When the molecular weight of a gas is known, its density at STP can be calculated from Eq. (2.6) and its density at any other temperature and pressure from Eq. (2.5).

Table 2.1 contains the molecular weights and densities at STP of the main constituents of dry air. Multiplying each density by the appropriate volume fraction gives the mass concentration of each component and the sum of these concentrations is the density of dry air. From a density of 1.292 kg m−3 and from Eq. (2.5) the effective molecular weight of dry air (in g) is 28.96 or 29 within 0.1%.

Since air is a mixture of gases, it obeys Dalton's Law, which states that the total pressure of a mixture of gases that do not react with each other is given by the sum of the partial pressures. *Partial pressure* is the pressure that a gas would exert at the same temperature as the mixture if it alone occupied the volume that the mixture occupies.

Water Vapor and its Specification

The evaporation of water at the earth's surface to form water vapor in the atmosphere is a process of major physical and biological importance because the latent heat of vaporization is large in relation to the specific heat of air. The heat released by condensing 1 g of water vapor is enough to raise the temperature of 1 kg of air by 2.5 K. Water vapor has been called the "working substance" of the atmospheric heat engine because of its role in global heat transport. The total mass of water vapor in the air at any moment is enough to supply only 1 week of the world's precipitation, so the process of evaporation must be very efficient in replenishing the atmospheric reservoir. On a much smaller scale, it is the amount of latent heat removed by the evaporation of sweat that allows man and many other mammals to survive in hot climates. Sections which follow describe the physical significance of different ways of specifying the amount of vapor in a sample of air and relations between them.

Transport of Heat, Mass, and Momentum

The last chapter was concerned primarily with ways of specifying the state of the atmosphere in terms of properties such as pressure, temperature, and gas concentration.

To continue this introduction to some of the major concepts and principles on which environmental physics depends, we now consider how the transport of entities such as heat, mass, and momentum is determined by the state of the atmosphere and the corresponding state of the surface involved in the exchange, whether soil, vegetation, the coat of an animal, or the integument of an insect or seed.

General Transfer Equation

A simple general equation can be derived for transport within a gas by "carriers," which may be molecules or particles or eddies, capable of transporting units of a property P such as heat, water vapor, or a gas. Even when the carriers are moving randomly, net transport may occur in any direction provided that the concentration of P decreases with distance in that direction. The carrier can then "unload" its excess of P at a point where the local value is less than at the starting point.

To evaluate the net flow of P in one dimension, consider a volume of gas with unit horizontal cross-section and a vertical height l assumed to be the mean distance for unloading a property of the carrier (Figure 3.1). Over the plane defining the base of the volume, P has a uniform value $P(0)$, and if the vertical gradient (change with height) of P is dP/dz, the value at height l will be $P(0)+l\ dP/dz$. Carriers which originate from a height l will therefore have a load corresponding to $P(0)+l\ dP/dz$, and if they move a distance l vertically downwards to the plane where the standard load corresponds to $P(0)$, they will be able to unload an excess of $l\ dP/dz$.

To find the rate of transport equivalent to this excess, assume that n carriers per unit volume move with a random *root mean square velocity* v so that the number moving toward one face of a cube at any instant is $nv/6$ per unit area. The downward flux of P (quantity per unit area and per unit time) is therefore given by $(nvl/6)dP/dz$. However, there is a corresponding upward flux of carriers reaching the same plane from below after setting off with a load given by $P(0)-l\ dP/dz$. Mathematically, the upward flow of a deficit is equivalent to the downward flow of an excess, so the net *downward* flux of P is

$$\mathbf{F} = (nvl/3)\, dP/dz.$$

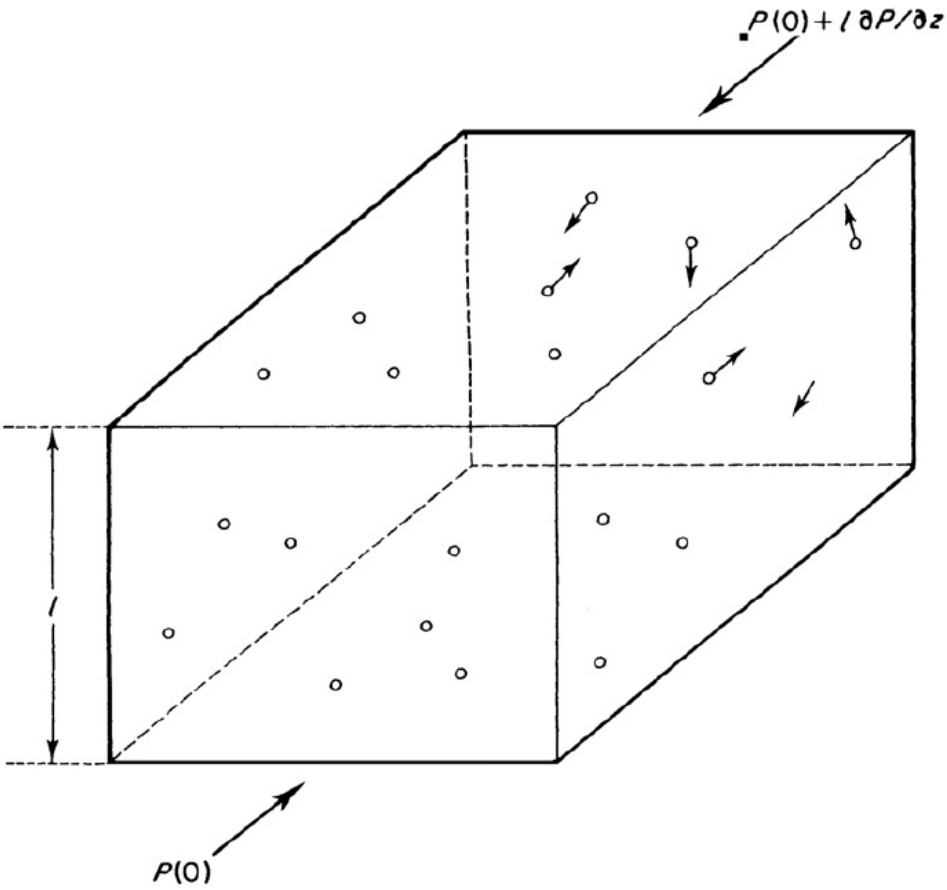


Figure 1 Volume of air with unit cross-section and height l containing n carriers per unit volume moving with random velocity v.

For the application of Eq. (3.1) to the transport of entities by molecular movement, the mean velocity in the z direction, w, is often related to the root mean square molecular velocity v by assuming that, at any instant, one-third of the molecules in the system are moving in the z direction so that $w = v/3$.

On the other hand, micrometeorologists studying transfer by turbulent eddies are concerned with a form of Eq. (3.1) in which P is replaced by the amount of an entity per unit mass of air (the *specific concentration*, q) rather than per unit volume (since the latter depends on temperature); the two quantities are related by $nP = \rho q$, where ρ is air density. The flux equation therefore becomes

$$\mathbf{F} = -\overline{\rho w l}\,(d\bar{q}/dz).$$

The minus sign is needed to indicate that the flux is downwards if q increases upwards; averaging bars are a reminder that both w and q fluctuate over a wide range

of time scales as a consequence of turbulence. In this context, the quantity *l* is known as the "*mixing length*" for turbulent transport.

It is also possible to write the instantaneous values of *q* and *w*as the sum of mean values $\overline{q}$ or $\overline{w}$ and corresponding deviations from the mean *q*_ and *w* _. The net flux across a plane then becomes

$$\overline{(\rho\bar{w} + \rho w')(\bar{q} + q')} = \overline{\rho w' q'},$$

Where $\overline{q}$_ and $\overline{w}$

_ are zero by definition and $\overline{w}$ is assumed to be zero near the ground when averaging is performed over a period long compared with the lifetime of the largest eddy (say 10 min). This relation provides the "eddy covariance" method of measuring Vertical fluxes discussed in Chapter 16.

Molecular Transfer Processes

According to Eq. (2.1), the mean square velocity of molecular motion in an ideal gas is v2 = 3p/ρ. Substituting p = 105 N m−2 and ρ = 1.29 kg m−3 for air at STP gives the root mean square velocity as *(v*2*)*1/2 = 480 m s−1, and the kinetic theory of gases may be used to show that the mean free path (the average distance a molecule travels between collisions) is 63 nm. Molecular motion in air is therefore extremely rapid over the whole range of temperatures found in nature, and collisions are frequent. This motion is responsible for a number of processes fundamental to micrometeorology: the transfer of momentum in moving air responsible for the phenomenon of viscosity; the transfer of heat by the process of conduction; and the transfer of mass by the diffusion of water vapor, carbon dioxide, and other gases. Because all three forms of transfer are a direct consequence of molecular agitation, they are described by similar relationships which will be considered for the simplest possible case of diffusion in one dimension only.

Momentum and Viscosity

When a stream of air flows over a solid surface, its velocity increases with distance from the surface. For a simple discussion of viscosity, the velocity gradient *du/dz* will be assumed linear as shown in Figure 3.2. (Amore realistic velocity profile will be

considered in Chapter 7.) Provided the air is isothermal, the velocity of molecular agitation will be the same at all distances from the surface but the horizontal component of bulk velocity in the *x* direction increases with vertical distance *z*. As a direct consequence of molecular agitation, there is a constant interchange of molecules between adjacent horizontal layers with a corresponding vertical exchange of horizontal momentum. The horizontal momentum of a molecule attributable to bulk motion of the gas as distinct from random motion is *mu*, so from Eq. (3.1) the rate of transfer of momentum, otherwise known as the *shearing stress*, can be written

$$\tau = (nvl/3)\, d\,(mu)/dz = (vl/3)\, d\,(\rho u)/dz,$$

as the density of the gas is $\rho = mn$. This is formally identical to the empirical equation defining the *kinematic viscosity* v of a gas, viz.

$$\tau = vd(\rho u)/dz,$$

showing that v is a function of molecular velocity and mean free path. Where the change of ρ with distance is small, it is more convenient to write

$$\tau = \mu du/dz,$$

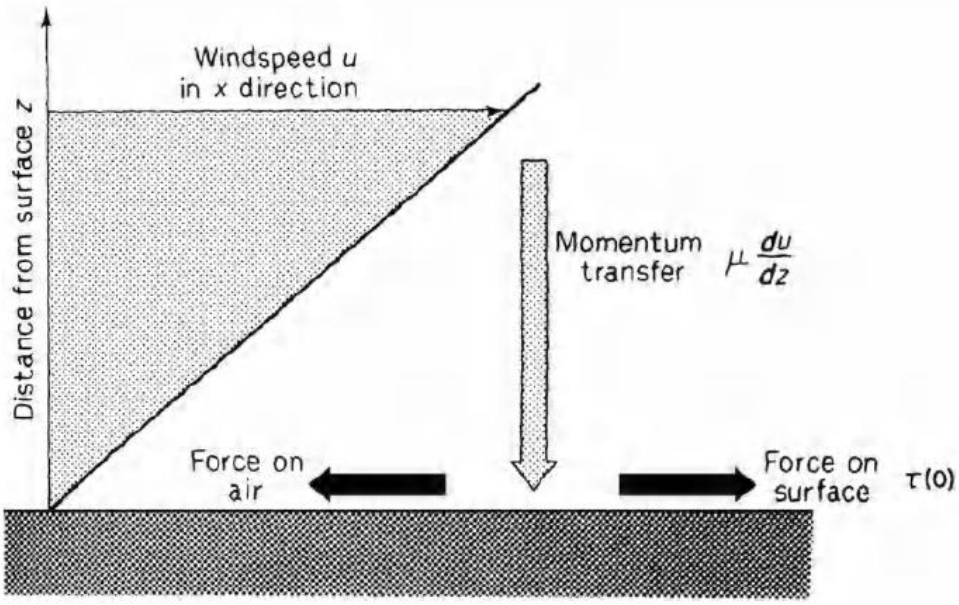


Figure 2 Transfer of momentum from moving air (moving left to right) to a stationary surface, showing related forces.

Where $\mu = \bar{\rho}\nu$ is the coefficient of *dynamic viscosity* and $\bar{\rho}$ is a mean density. By convention, the flux of momentum is taken as positive when it is directed *toward* a surface and it therefore has the same sign as the velocity gradient (see Figure 3.2). The momentum transferred layer by layer through the gas is finally absorbed by the surface which therefore experiences frictional force acting in the direction of the flow. The reaction to this force required by Newton's Third Law is the frictional drag exerted on the gas by the surface in a direction opposite to the flow.

Heat and Thermal Conductivity

The conduction of heat in still air is analogous to the transfer of momentum. In Figure 3.3, a layer of warm air makes contact with a cooler surface. The velocity of molecules therefore increases with distance from the surface and the exchange of molecules between adjacent layers of air is responsible for a net transfer of molecular energy and hence of heat. The rate of transfer of heat is proportional to the gradient of heat content per unit volume of the air and may therefore be written

$$\mathbf{C} = -\kappa d(\rho c_{\mathrm{p}} T)/dz,$$

where κ, the *thermal diffusivity* of air, has the same dimensions (L2 T−1) as the kinematic viscosity and ρc_{p} is the heat content per unit volume of air. As in the treatment of momentum, it is convenient to assume that ρ has a constant value of $\bar{\rho}$ over the distance considered and to define a *thermal conductivity* as $k = \bar{\rho} c_{\mathrm{p}} \kappa$ so that

$$\mathbf{C} = -k\, dT/dz$$

Identical to the equation for the conduction of heat in solids.

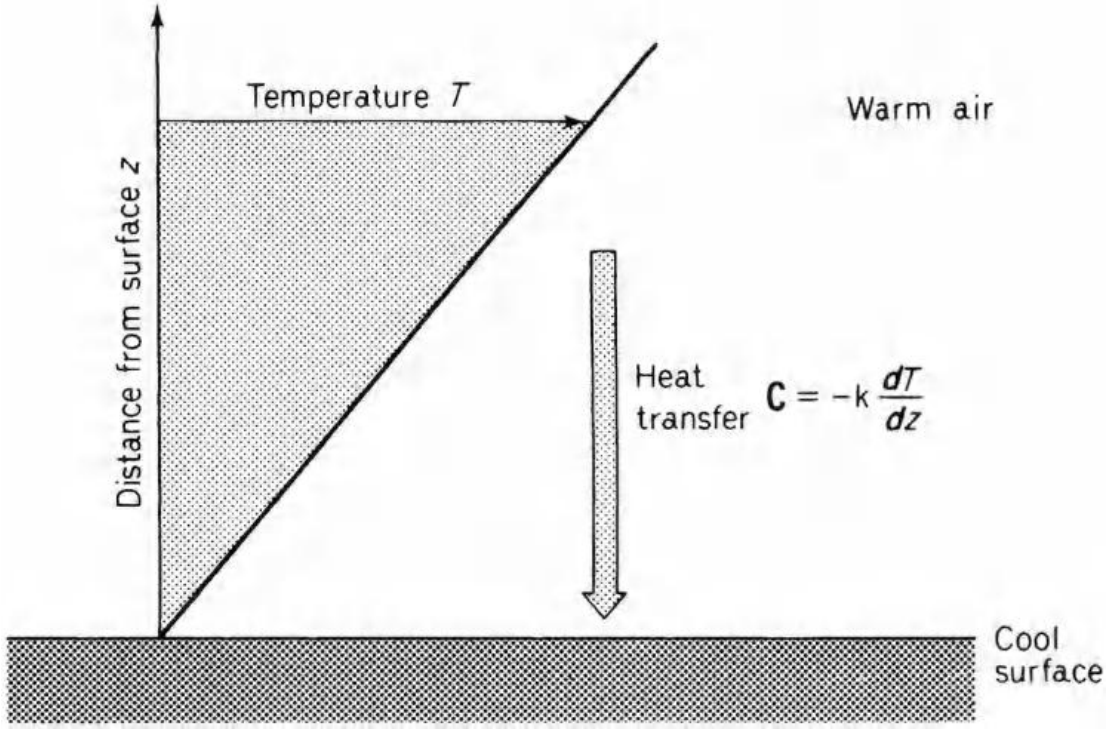


Figure 3 Transfer of heat from still, warm air to a cool surface.

In contrast to the convention for momentum, **C** is conventionally taken as positive when the flux of heat is *away* from the surface in which case *dT/dz* is negative. The equation therefore contains a minus sign.

Mass Transfer and Diffusivity

In the presence of a gradient of gas concentration, molecular agitation is responsible for a transfer of mass, generally referred to as "diffusion" although this word can also be applied to momentum and heat. In Figure 3.4, a layer of still air containing water vapor makes contact with a hygroscopic surface where water is absorbed. The number of molecules of vapor per unit volume increases with distance from the surface and the exchange of molecules between adjacent layers produces a net movement toward the surface. The transfer of molecules expressed as a mass flux per unit area ***(E)*** is proportional to the gradient of concentration and the transport equation analogous to Eqs. (3.6) and (3.8) is

$$\mathbf{E} = -D\,d\chi/dz = -Dd(\rho q)/dz = -\bar{\rho}D\,dq/dz,$$

where $\overline{\rho}$ is an appropriate mean density, D (dimensions L2 T−1) is the molecular diffusion coefficient forwater vapor, and q is the specific humidity. The sign convention in this equation is the same as for heat.

Diffusion Coefficients

Because the same process of molecular agitation is responsible for all three types of transfer, the diffusion coefficients for momentum, heat, water vapor, and other gases

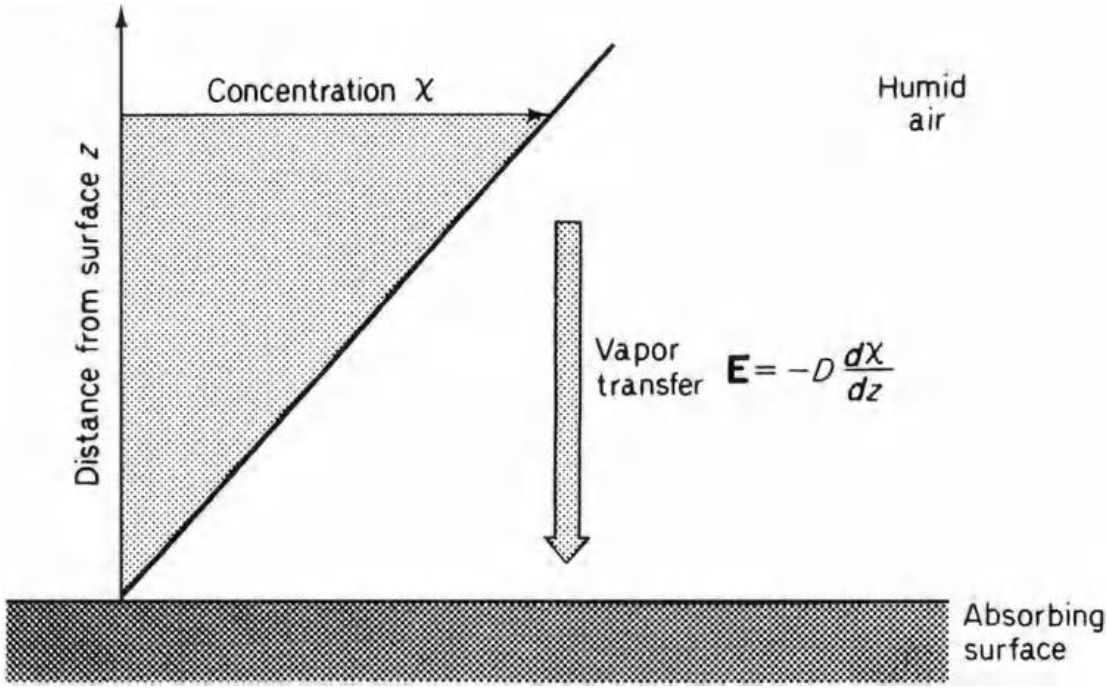


Figure 4 Transfer of vapor from humid air to an absorbing surface.

are similar in size and in their dependence on temperature. Values of the coefficients at different temperatures calculated from the Chapman-Eskog kinetic theory of gases agree well with measurements and are given in the Appendix, Table A.3. The temperature and pressure dependence of the diffusion coefficients is usually expressed by a power law, e.g.

$$D(T)/D(0) = \{T/T(0)\}^n \{P(0)/P\},$$

where $D(0)$ is the coefficient at base temperature and pressure $T(0)$ (K) and $P(0)$ respectively, and n is an index between 1.5 and 2.0. Within the limited range of temperatures relevant to environmental physics, say −10 to 50 ◦C, a simple temperature coefficient of 0.007 is accurate enough for practical purposes, i.e.

$$D(T)/D(0) = \kappa(T)/\kappa(0) = \nu(T)/\nu(0) = (1 + 0.007T),$$

where T is the temperature in ◦C and the coefficients in units of m2 s−1 are

$$\begin{aligned}
\nu(0) &= 13.3 \times 10^{-6} \text{ (momentum)},\\
\kappa(0) &= 18.9 \times 10^{-6} \text{ (heat)},\\
D(0) &= 21.2 \times 10^{-6} \text{ (water vapor)}\\
&= 12.9 \times 10^{-6} \text{ (carbon dioxide)}.
\end{aligned}$$

Graham's Law states that the diffusion coefficients of gases are inversely proportional to the square roots of their densities, i.e. $D \propto (M)$ −0.5 since density is proportional to molecular weight. Consequently, the diffusion coefficient *Dx* for an unknown gas of molecular weight *Mx* can be estimated from values for a known gas (*Dy* , *My*) using the relation *Dx* = *Dy(My/Mx)*0.5.

1 Resistances to Transfer

Equations (3.5)–(3.7) have the same form flux = diffusion coefficient × gradient, which is a general statement of Fick's Law of Diffusion. This law can be applied to problems in which diffusion is a one-, two-, or three-dimensional process but only one dimensional cases will be considered here. Because the gradient of a quantity at a point is often difficult to estimate accurately, Fick's law is generally applied in an integrated form. The integration is very straightforward in cases where the (one-dimensional) flux can be treated as constant in the direction specified by the coordinate *z*, e.g. at right angles to a surface. Then the integration of (3.9) for example gives

$$\mathbf{E} = -\frac{\int_{z_1}^{z_2} d(\rho q)}{\int_{z_1}^{z_2} dz/D} = \frac{\rho q(z_1) - \rho q(z_2)}{\int_{z_1}^{z_2} dz/D},$$

where $\rho q(z1)$ and $\rho q(z2)$ are concentrations of water vapor at distances $z1$ and $z2$ from a surface absorbing or releasing water vapor at a rate **E**. Usually $\rho q(z1)$ is taken as the concentration at the surface so that $z1 = 0$. Equation (3.11) and similar equations derived by integrating Eqs. (3.6) and (3.8) are analogous to Ohm's Law in electrical circuits, i.e. current through resistance = potential difference across resistance. Equivalent expressions for diffusion can be written as:

rate of transfer of entity = potential difference resistance

$$\text{rate of momentum transfer } \boldsymbol{\tau} = \rho u / \int dz/\nu,$$

$$\text{rate of heat transfer } \mathbf{C} = -\rho c_p T / \int dz/\kappa,$$

$$\text{rate of mass transfer } \mathbf{E} = -\rho q / \int dz/D.$$

The definition of *resistance* r to mass transfer is therefore

$$r = \int dz/D$$

and similar equations define resistances to heat and momentum transfer. Diffusion coefficients have dimensions of (length) 2 × (time) so the corresponding resistances have dimensions of (time)/(length) or 1/(velocity). In a system where rates of diffusion are governed purely by molecular processes, the coefficients can usually be assumed independent of z so that $\int_{z1}^{z2} dz/D$, for example, becomes simply $(z2-z1)/D$ or (diffusion path length)/(diffusion coefficient).

Example 1. What is the resistance to water vapor diffusion by molecular agitation for a path length of 1 mm of air at 20 ◦C and 101.3 kPa?

Solution: The molecular diffusion coefficient of water vapor in air for the specified temperature and pressure is given in Appendix A.3, and is 24.9 × 10−6 m2 s−1.

Consequently the resistance for a pathlength of 1 mm (1 × 10−3 m) is 1 × 10−3/24.9 × 10−6 = 40 s m−1.

It is often convenient to treat the process of diffusion in laminar boundary layers (layers where transfer is only by molecular motion) in terms of resistances, and in the remainder of this book the following symbols are used:

*r*M resistance for momentum transfer at the surface of a body

*r*H resistance for convective heat transfer

*r*V resistance for water vapor transfer

*r*C resistance for CO2 transfer

The concept of resistance is not limited to molecular diffusion but is applicable to any system in which fluxes are uniquely related to gradients. In the atmosphere, where turbulence is the dominant mechanism of diffusion, diffusion coefficients are several orders of magnitude larger than the corresponding molecular values and increase with height above the ground (Chapter 16). Diffusion resistances for momentum, heat, water vapor and carbon dioxide in the atmosphere will be distinguished by the symbols *r*aM, *r*aH, *r*aV, and *r*aC;

In studies of the deposition of radioactive material and pollutant gases from the atmosphere to the surface, the rate of transfer is sometimes expressed as a *deposition velocity*, which is the reciprocal of a diffusion resistance. In this case, the surface concentration is often assumed to be zero and the deposition velocity is found by dividing the rate of deposition of the material by its concentration at an arbitrary height.

Plant physiologists also frequently use the reciprocal of resistance (in this context termed *conductance*) to describe transfer between leaves and the atmosphere, arguing that the direct proportionality between *flux* and *conductance* is a more intuitive concept than the inverse relationship between flux and resistance. In this book we generally prefer the resistance formulation because of its familiarity to physicists, particularly when combinations of resistances in parallel and series must be calculated.

1 Alternative Units for Resistance and Conductance

Units of sm−1 for resistance andms−1 for conductance are the result of expressing mass fluxes as mass flux density (e.g. kg m−2 s−1) and driving potentials as concentrations (e.g. kg m−3). The forms of the equations for momentum, heat, and

mass transfer (3.12) ensure that resistance units for these variables are also s m−1. A criticism of this convention is that, when resistance is defined as (diffusion path length)/(diffusion coefficient), r is proportional to pressure P and inversely approximately proportional to T 2 (see Eq. (3.10)). Thus, for example, analyzing the effects of altitude on fluxes can be confusing. Alternative definitions of resistance units, less sensitive to temperature and pressure, are sometimes used, particularly by plant physiologists, as follows.

Since biochemical reactions concern numbers of molecules reacting, rather than the mass of substances, it is convenient to express fluxes in *mole flux density*, J (mol m−2 s−1). Similarly, amount of the substance can be expressed as the *mole fraction* x, i.e. the number of moles of the substance as a fraction of the total number of moles in the mixture (mol mol−1). Using the Gas Laws, it is readily shown that mass concentration χ (kg m−3), or its equivalent ρq, is related to x by

$$\chi = \rho q = xP/RT,$$

so Eq. (3.9) may be written

$$J = \frac{x(z_1) - x(z_2)}{\frac{RT}{P}\int dz/D}$$

or, since $x = p/P$, where p is the partial pressure,

$$J = \frac{p(z_1) - p(z_2)}{RT\int dz/D}.$$

The molar resistance r_m (m^2 s mol^{-1}) is defined as

$$r_m = \frac{RT}{P}\int \frac{dz}{D} = \frac{RT}{P}r$$

demonstrating that rm is independent of pressure and less dependent on temperature than r . pointed out that it is important to use partial pressure or mole fraction to describe potentials that drive diffusion when systems are not isothermal. At 20 ◦C and 101.3 k Pa, the approximate conversion between resistance units is r m*(*m2 s mol−1*)* = 0.024r *(*s m−1*)*.

Diffusion of Particles (Brownian Motion)

The random motion of particles suspended in a fluid or gas was first described by the English botanist Brown in 1827, but it was nearly 80 years before Einstein used the kinetic theory of gases to show that the motion was the result of multiple collisions with the surrounding molecules. He found that the mean square displacement $x2$ of a particle in time t is given by

$x2 = 2Dt$, (3.16)

where D is a diffusion coefficient (dimensions L2 T−1) for the particle, analogous to the coefficient for gas molecules. The quantity D depends on the intensity of molecular bombardment (a function of absolute temperature), and on the viscosity of the fluid, as follows.

Suppose that particles, each with mass m (kg), are dispersed in a container where they neither stick to the walls nor coagulate. The Boltzmann statistical description of concentration, derived from kinetic theory, requires that, as a consequence of the earth's gravitational field, the particle concentration n should decrease exponentially with height z(m) according to the relation

$n = n(0)$ exp $(-mgz/kT)$, (3.17)

where

$n = n(0)$ at $z = 0$,

g = gravitational acceleration(m s−2),

T = absolute temperature (K),

k = Boltzmann's constant (J K−1).

Across a horizontal area at height z within the container, the flux of particles upwards by diffusion (cf. diffusion of gases) is:

$$\mathbf{F}_1 = -D\frac{dn}{dz} = n\frac{Dmg}{kT}$$

from Eq. (3.17). Because all particles tend to move downwards in response to gravity, there must be a downward flux through the area of

F2 = *n*V*s*,

where *V*s is the "sedimentation velocity"

Since the system is in equilibrium, **F**1 = **F**2 and so

$D = kT\ Vs/mg$. (3.19)

For spherical particles, radius r , obeying Stokes' Law, the downward force mg due to gravity is balanced by a drag force $6\pi\rho g v r Vs$, where ρg is the gas density and v is the kinematic viscosity (see Chapter 9). Consequently

$D = kT/6\pi\rho g v r$. (3.20)

Thus D depends inversely on particle radius (see Appendix, Table A.6); its dependence on temperature is dominated by the dependence of the kinematic viscosity v on temperature. Equations (3.16) and (3.17) show that the root mean square displacement of a particle by Brownian motion is proportional to T 0.5 and to $r-0.5$. Surprisingly, $\overline{\ \ } x2$ does not depend on the mass of the particle, an inference confirmed by experiment.

Problems

1. Calculate the resistance (in s m−1) for carbon dioxide diffusion in air over a path length of 1 mm, assuming air temperature is 20 ◦C and atmospheric pressure is 101 kPa. Recalculate the resistance assuming that atmospheric pressure decreased to 70 kPa (keeping temperature constant). Repeat the two calculations using molar units.

2. Use the data in the Appendix Table A.6 to investigate how the diffusion coefficient of particles depends on temperature.

The Origin and Nature of Radiation

Electromagnetic radiation is a form of energy derived from oscillating magnetic and electrostatic fields and is capable of transmission through empty space where its velocity is c = 3.0 × 108 m s−1. The frequency of oscillation v is related to the wavelength λ by the standard wave equation $c = \lambda v$ and the wave number $1/\lambda = v/c$ is sometimes used as an index of frequency.

The ability to emit and absorb radiation is an intrinsic property of solids, liquids, and gases and is always associated with changes in the energy state of atoms and

molecules. Changes in the energy state of atomic electrons are associated with line spectra confined to a specific frequency or set of frequencies. In molecules, the energy of radiation is derived from the vibration and rotation of individual atoms within the molecular structure. The principle of energy conservation is fundamental to the material origin of radiation. The amount of radiant energy emitted by an individual atom or molecule is equal to the decrease in the potential energy of its constituents.

Absorption and Emission of Radiation

All molecules possess a certain amount of "internal" energy (i.e. not associated with their motion in the atmosphere).Most of the energy is associated with electrons orbiting around the nucleus, but part is related to vibration of atoms in the molecular structure and to rotation of the molecule. Quantum physics predicts that only certain electron orbits, vibration frequencies, and rotation rates are allowed for a particular molecule, and each combination of orbits, vibrations, and rotations corresponds to a particular amount of energy associated with the three features. Molecules may make a transition to a higher or lower energy level by absorbing or emitting electromagnetic radiation respectively.

Quantum theory allows only certain discrete changes in energy levels, which are the same whether the energy is being absorbed or emitted. Since the energy associated with a photon is related to its wavelength by $E = hc/\lambda$, where h is Planck's constant, it follows that molecules can interact only with certain wavelengths of radiation. Thus the variation of absorption and emission of molecules with wavelength takes the form of a *line spectrum*, consisting of a finite number of wavelengths where interaction is allowed, interspaced by gaps where there is no interaction.

Most absorption lines associated with orbital changes are in the X-ray, UV, and visible spectrum. Vibrational changes are associated with absorption at infrared wavelengths, and rotational changes correspond to lines at even longer, microwave wave lengths. Molecules such as$CO2$,$H2O$, and$O3$ have structures that allow vibration rotational transitions simultaneously, and these correspond to clusters of very closely spaced lines in the infrared region. Molecules such as $O2$ do not interact this way, so have only small numbers of absorption lines.

When large numbers of molecules are present in a gas, the width of their absorption and emission lines is greatly enhanced by broadening associated with random

molecular motions (*Doppler broadening*—depending on the square root of absolute temperature) and with interactions during collisions (*collision broadening*—depending on the frequency of collisions, which is proportional to gas pressure). Collision broadening is most important for atmospheric molecules below about 30 km and results in overlapping of lines in the clusters associated with vibration-rotational transitions in CO2 and H2O, creating *absorption bands* in the infrared for these gases. Since pressure decreases with increasing height, the absorptivity and emissivity of a gas distributed in the lower atmosphere with a constant mixing ratio changes with height, making calculations of radiative transfer complex.

Full or Black Body Radiation

Relations between radiation absorbed and emitted by matter were examined by Kirchhoff. He defined the absorptivity of a surface $\alpha(\lambda)$ as the fraction of incident radiation absorbed at a specific wavelength λ. The emissivity $\varepsilon(\lambda)$ was defined as the ratio of the actual radiation emitted at thewavelength λ to a hypothetical amount of radiant flux $\mathbf{B}(\lambda)$. By considering the thermal equilibrium of an object inside an enclosure at a uniform temperature, he showed that $\alpha(\lambda)$ is always equal to $\varepsilon(\lambda)$. For an object completely absorbing radiation at wavelength λ, $\alpha(\lambda) = 1$, $\varepsilon(\lambda) = 1$, and the emitted radiation is $\mathbf{B}(\lambda)$. In the special case of an object with $\varepsilon = 1$ at *all* wavelengths, the spectrum of emitted radiation is known as the "full" or *black body spectrum*. Within the range of temperatures prevailing at the earth's surface, nearly all the radiation emitted by full radiators is confined to the waveband 3–100 μm, and most natural objects—soil, vegetation, water—behave radiatively almost like full radiators in this restricted region of the spectrum (but not in the visible spectrum). Even fresh snow, one of the whitest surfaces in nature, emits radiation like a black body between 3 and 100 μm. The statement "snow behaves like a black body" refers therefore to the radiation *emitted* by a snow surface and not to solar radiation *reflected* by snow. The semantic confusion inherent in the term "black body" can be avoided by referring to "*full radiation*" and to a "*full radiator*."

After Kirchhoff's work was published in 1859, the emission of radiation by matter was investigated by a number of experimental and theoretical physicists. By combining a spectrometer with a sensitive thermopile, it was established that the spectral distribution of radiation from a full radiator resembles the curve in Figure 4.1 in which the chosen temperatures of 6000 and 300 K correspond approximately to the

mean full radiation temperatures of the sun and the earth's surface. A theoretical explanation of

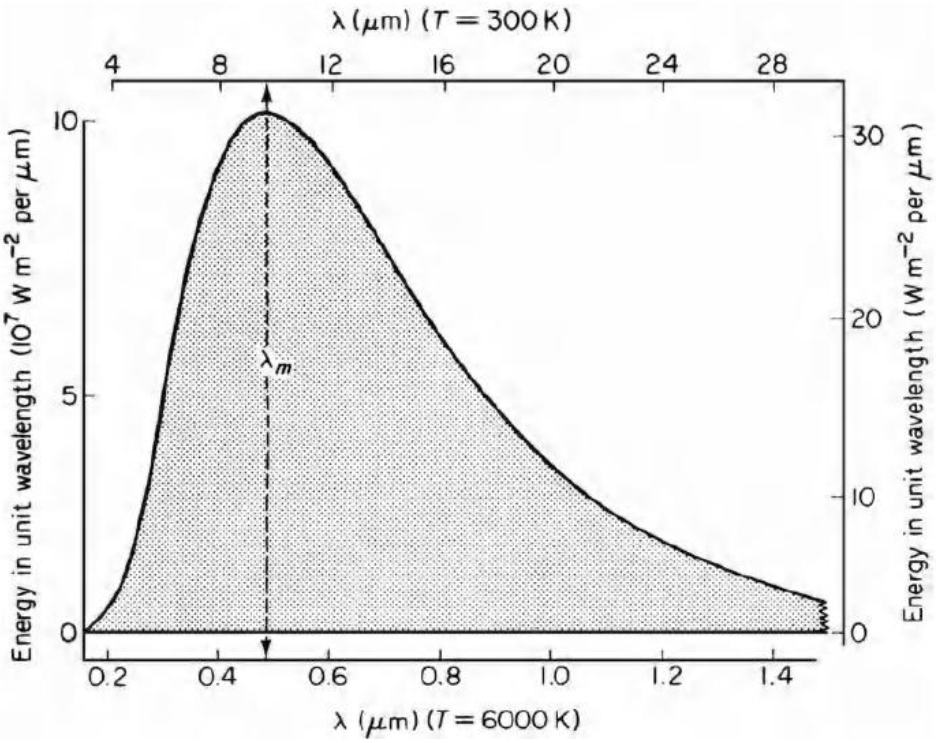
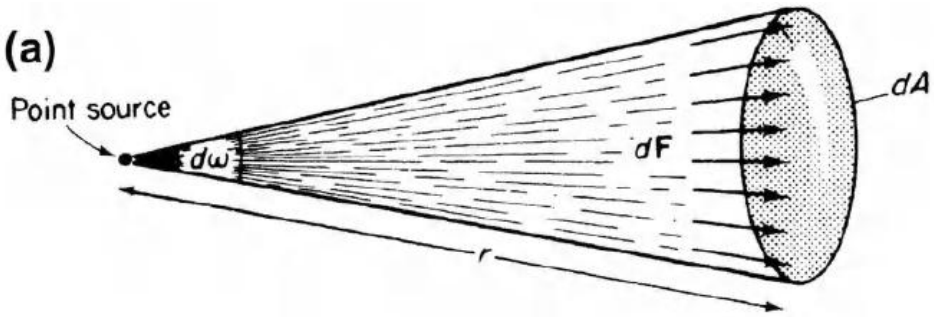


Figure 1 Spectral distribution of radiant energy from a full radiator at a temperature of (a) 6000 K, left-hand vertical and lower horizontal axis and (b) 300 K, right-hand vertical and upper horizontal axis. Note that the scales on the left and right vertical axes differ by more than six orders of magnitude. About 10% of the energy is emitted at wavelengths longer than those shown in the diagram. If this tail were included, the total area under the curve would be proportional to $\sigma T 4$ (Wm−2). λm is the wavelength at which the energy per unit wavelength is maximal. The distribution eluded physicists until Plank's quantum hypothesis emerged (Section 4.1.5).

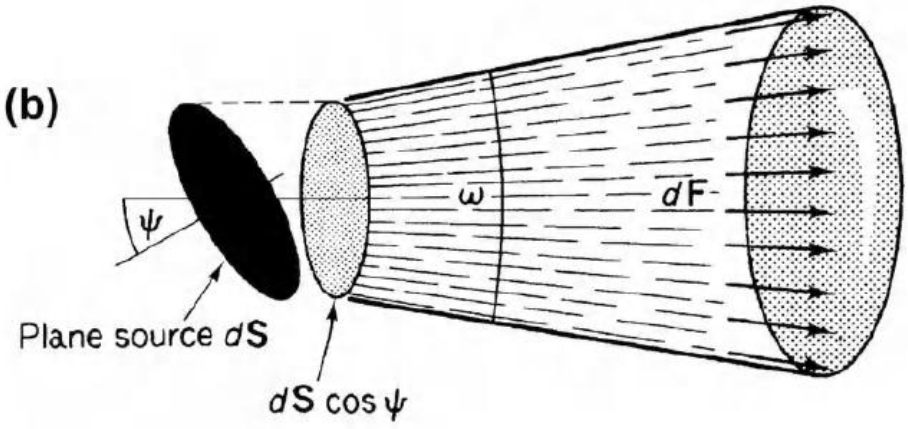
Cosine Law for Emission and Absorption

The concept of radiance is linked to an important law describing the spatial distribution of radiation emitted by a full radiator which has a uniform surface temperature *T*. This temperature determines the total flux of energy emitted by the surface ($\sigma T 4$) and can be estimated by measuring the radiance of the surface with a radiometer. As the surface of a full radiator must appear to have the same temperature whatever angle ψ it is viewed from, the intensity of radiation emitted from a point on the surface and the radiance of an element of surface must both be independent of ψ.

On the other hand, the flux per unit solid angle divided by the *true* area of the surface must be proportional to cosψ. Figure 3 makes this point diagrammatically. A radiometer R mounted vertically above an extended horizontal surface *XY* "sees" an area *dA* and measures a flux which is proportional to *dA*. When the surface is tilted through an angle ψ, the radiometer now sees a larger surface *dA*/ cosψ, but provided the temperature of the surface stays the same, its radiance will be constant and the flux recorded by the radiometer will also be constant. It follows that the flux emitted per unit area (the emittance of the surface) at an angle ψ must be proportional to cosψ so that the product of emittance



Solid angle $d\omega = dA/r^2$
Intensity $I = dF/d\omega$



Intensity $dI = dF/\omega$
Radiance $= (dF/\omega) \div dS\cos\psi$
$= dI/(dS\cos\psi)$

Figure 2 (a) Geometry of radiation emitted by a point source. (b) Geometry of radiation emitted by a surface element. In both diagrams a portion of a spherical surface receives radiation at normal incidence, but when the distance between the source and the receiving surface is large, it can be treated as a plane.

($\propto \cos\psi$) and the area emitting to the instrument ($\propto 1/\cos\psi$) stay the same for all values of ψ. This argument is the basis of *Lambert's Cosine Law* which states that when radiation is emitted by a full radiator at an angle ψ to the normal, the flux per unit solid angle emitted by unit surface is proportional to $\cos\psi$. As a corollary to Lambert's Law, it can be shown by simple geometry that when a full radiator is exposed to a beam of radiant energy at an angle ψ to the normal, the flux density of the absorbed radiation is proportional to $\cos\psi$. In remote sensing it is often necessary to specify the directions of both incident and reflected radiation, and reflectivity is then described as "bi-directional."

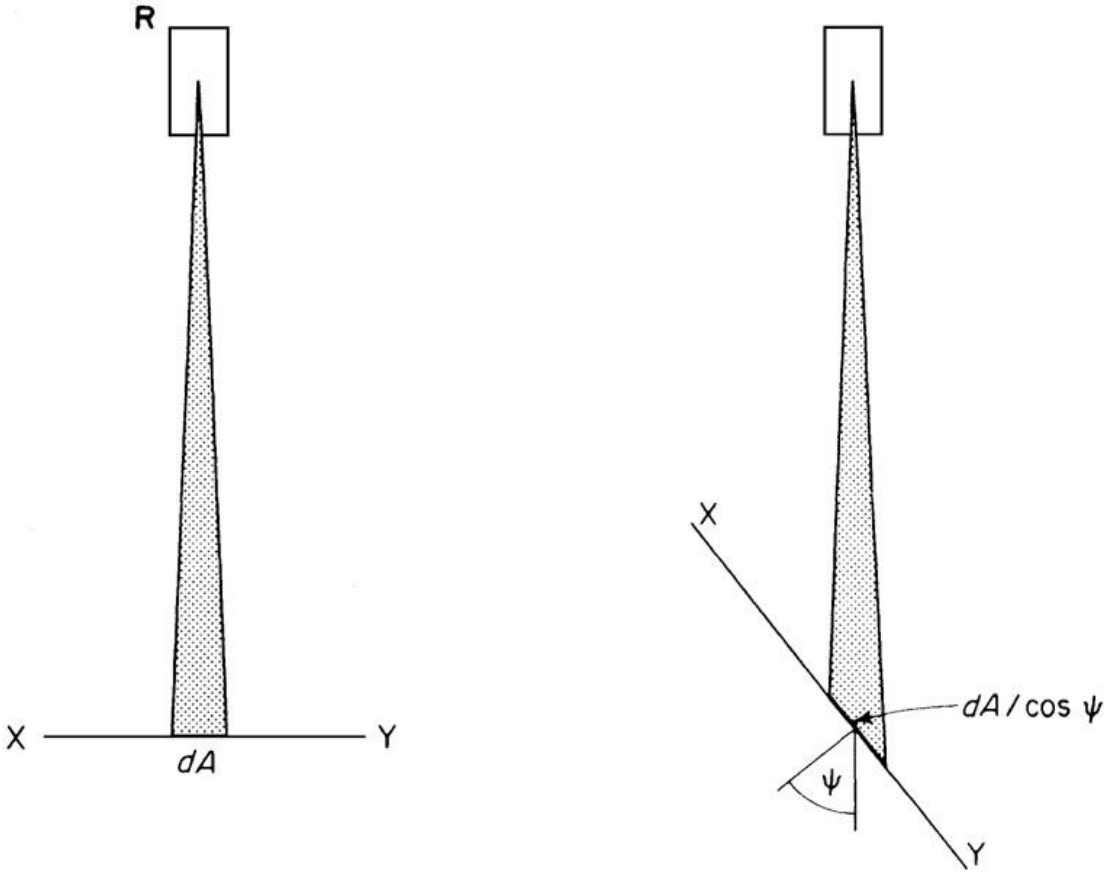


Figure 3 The amount of radiation received by a radiometer from the surface *XY* is independent of the angle of emission, but the flux emitted per unit area is proportional to cosψ.

Reflection

The reflectivity of a surface $\rho(\lambda)$ is defined as the ratio of the incident flux to the reflected flux at the same wavelength. Two extreme types of behavior can be distinguished. For surfaces exhibiting specular or mirror-like reflection, a beam of radiation incident at an angle ψ to the normal is reflected at the complementary angle $(-\psi)$. On the other hand, the radiation scattered by a perfectly diffuse reflector (also called a *Lambertian* surface) is distributed in all directions according to the Cosine Law, i.e. the intensity of the scattered radiation is independent of the angle of reflection but the flux scattered from a specific area is proportional to cosψ. The nature of reflection from the surface of an object depends in a complex way on its electrical properties and on the structure of the surface. In general, specular reflection assumes increasing importance as the angle of incidence increases, and surfaces

acting as specular reflectors absorb less radiation than diffuse reflectors made of the same material.

Most natural surfaces act as diffuse reflectors when ψ is less than 60◦ or 70◦, but as ψ approaches 90◦, a condition known as *grazing incidence*, the reflection from open water, waxy leaves, and other smooth surfaces becomes dominantly specular and there is a corresponding increase in reflectivity. The effect is often visible at sunrise and sunset over an extensive water surface, or a lawn, or a field of barley in ear.

When surfaces are observed by techniques of remote sensing, the direction of the radiation received by the radiometer is significant, and several additional definitions are necessary:

Bi-directional reflectance (sr−1) is the ratio of the radiation *reflected* in a specific direction of view to the radiation *incident* in that direction. The *Bi-directional reflectance factor* of a surface (BRF) is the ratio of the reflected radiance from a specific view direction to the radiance that would be observed from a perfectly diffuse surface at the same location. Most examples in this book deal with a plane surface below a uniform hemispherical source of radiation (e.g. an overcast sky). The fraction of incident radiation reflected from such a surface is sometimes referred to as the *bi-hemispherical reflectance* or simply the *reflection coefficient*. The reflection coefficient for solar radiation is commonly known as the *albedo.*

Radiance and Irradiance

When a plane surface is surrounded by a uniform source of radiant energy, a simple relation exists betweenthe irradiance of the surface (the flux incident per unit area) and the radiance of the source. Figure 4.4 displays a surface of unit area surrounded by a radiating hemispherical shell so large that the surface can be treated as a point at the center of the hemisphere. The shaded area *dS* is a small element of radiating surface and the radiation reaching the center of the hemisphere from *dS* makes an angle β with the normal to the plane. As the projection of unit area in the direction of the radiation is $1 \times \cos \beta$, the solid angle which the area subtends at *dS* is $\omega = \cos \beta / r$

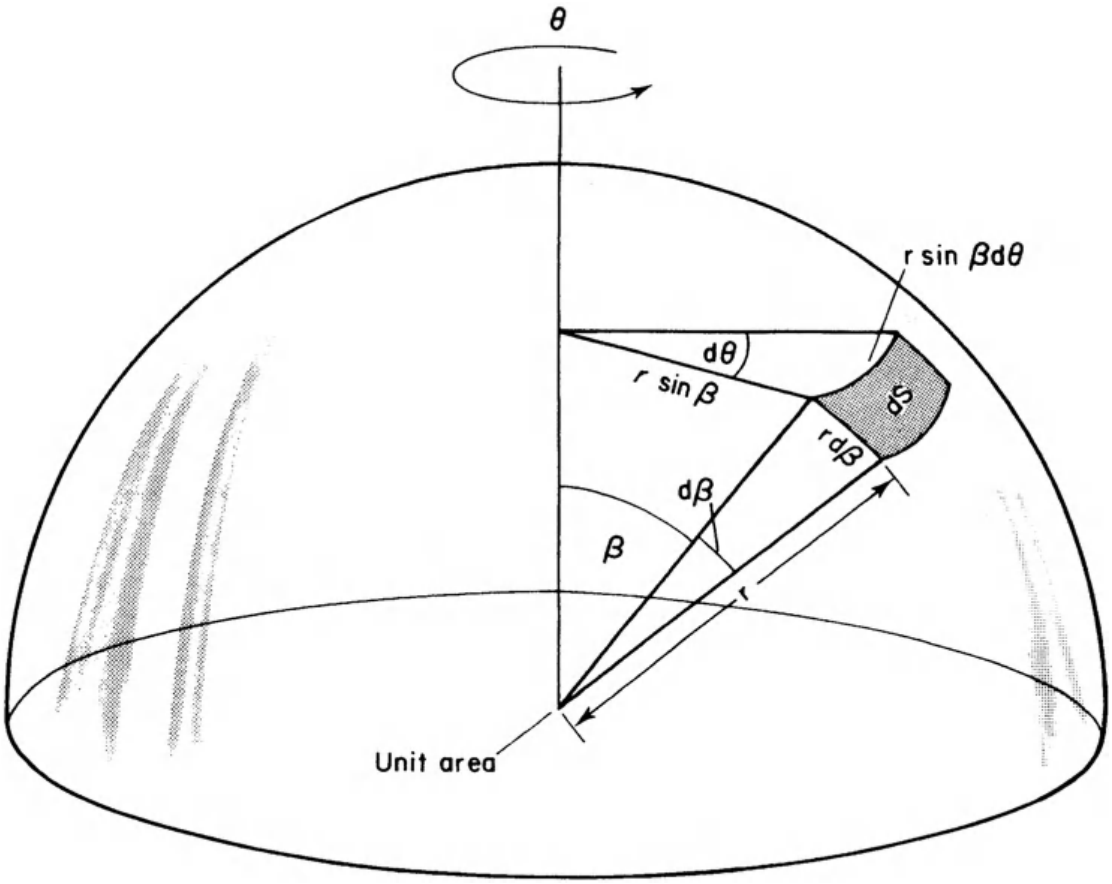


Figure 4 Method for calculating irradiance at the center of an equatorial plane from a surface element *dS* at angle *β* to vertical axis.

If the element *dS* has a radiance **N**, the flux emitted by *dS* in the direction of the plane must be **N** × *dS* × *ω* = **N***dS* cos *β/r* 2. To find the total irradiance of the plane, this quantity must be integrated over the whole hemisphere, but if the radiance is uniform, conventional calculus can be avoided by noting that *dS* cos *β* is the area *dS* projected on the equatorial plane. It follows that _ *dS* cos *β* is the area of the whole plane or *πr* 2, so that the total irradiance at the center of the plane becomes

$$(\mathbf{N}/r^2)\int \cos\beta\, dS = \pi\mathbf{N}.$$

The irradiance expressed in Wm−2 is therefore found by multiplying the radiance in Wm−2 sr−1 by the solid angle *π*. A more rigorous treatment is needed if the radiance depends on the position of *dS* with respect to the surface receiving radiation. It is necessary to treat *dS* as a rectangle whose sides are *rdβ* and *r* sin *βdθ* where *θ* is an

azimuth angle with respect to the axis of the hemisphere, radius r . Given that $dS = r\,2 \sin \beta d\beta d\theta$, the integral becomes

$$\int_{\theta=0}^{2\pi} \int_{\beta=0}^{\pi/2} N(\beta, \theta) \left(\frac{\cos \beta}{r^2} \right) r^2 \sin \beta \, d\beta \, d\theta.$$

If N is independent of azimuth (i.e. only a function of β), Eq. (4.9) simplifies to

$$= 2\pi \int_{\beta=0}^{\pi/2} N(\beta) \sin \beta \cos \beta \, d\beta$$

$$= \pi \int_{\beta=0}^{\pi/2} N(\beta) \sin 2\beta \, d\beta.$$

Attenuation of a Parallel Beam

When a beam of radiation consisting of parallel rays of radiation passes through a gas or liquid, quanta encounter molecules of the medium or particles in suspension. After interacting with the molecule or particle, a quantum may suffer one of two fates: it may be absorbed, thereby increasing the energy of the absorbing molecule or particle; or it may be scattered, i.e. diverted from its previous course either forwards (within 90◦ of the beam) or backwards in a process akin to reflection from a solid. After transmission through the medium, the beam is said to be "attenuated" by losses caused by absorption and scattering.

Beer's Law, frequently invoked in environmental physics, describes attenuation in a very simple system where radiation of a single wavelength is absorbed but not scattered when it passes through a homogeneous medium. Suppose that at some distance x into the medium the flux density of radiation is $\_(x)$ (Figure 4.5).

Absorption in a thin layer dx, assumed proportional to dx and to $\_(x)$, may be written

$$d\_ = -k\_(x)dx,$$

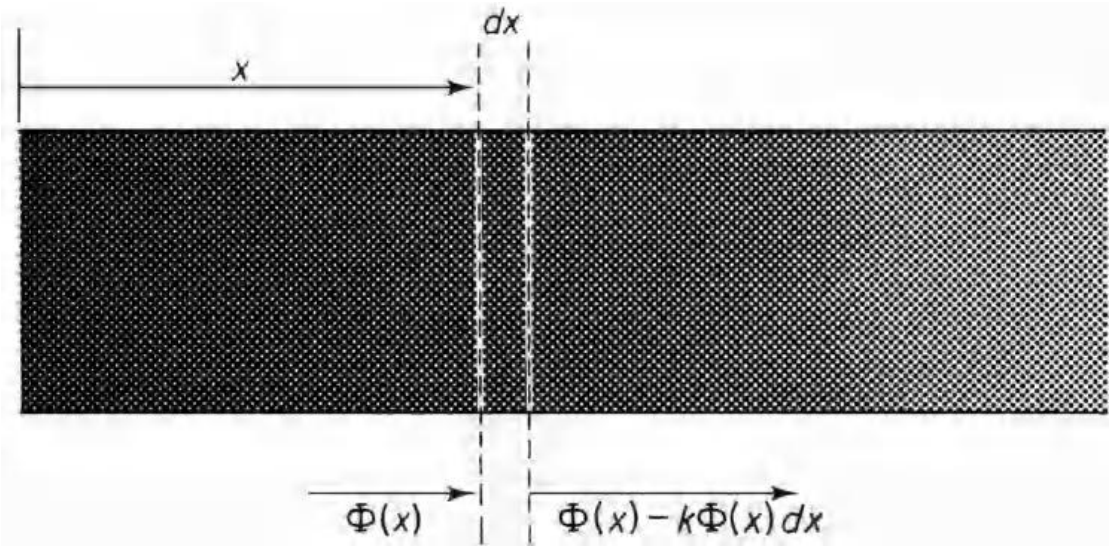


Figure 5 The absorption of a parallel beam of monochromatic radiation in a homogeneous medium with an absorption coefficient *k*. *_(0)* is the incident flux, *_(x)* is the flux at depth *x*, and the flux absorbed in a thin layer *dx* is *k_(x)dx*.

where the constant of proportionality *k*, described as an "*attenuation coefficient*," is the probability of a quantum being intercepted within the small distance *dx*. Integration gives

_(x) = *_(0)* exp *(−kx)*, (4.12)

where *_(0)* is the flux density at $x = 0$.

Beer's Law can also be applied to radiation in a waveband over which *k* is constant; or to a system in which the concentration of scattering centers is so small that a quantum is unlikely to interact more than once ("*single scattering*"). The treatment of *multiple scattering* is much more complex for two main reasons:

(a) radiation in a beam scattered backwards must be considered as well as the forward beam; (b) if *k* depends on beam direction, the angular distribution of scattering must be taken into account. For the simplest case where *k* is independent of beam direction, equations of a type developed by Kubelka and Munk are valid. They allow the attenuation coefficient to be expressed as a function of a reflection coefficient ρ which is the probability of an interacting quantum being reflected backwards and τ which is the probability of forward scattering, implying that the probability of absorption is

$\alpha = 1 - \rho - \tau$.

In a system with multiple scattering, it is necessary to distinguish two streams of radiation. One moves *into* the medium, and at a distance *x* from the boundary has a flux density *_+(x)*.

The other, generated by scattering, moves *out* of the medium with a flux density $\_^{-}(x)$. The inward flux is depleted by absorption and reflection but is augmented by reflection of a fraction of the outward stream. The *net* loss of inward flux at a depth x and in a distance dx is therefore

$$d\_^{+}(x) = (-(\alpha + \rho)\_^{+}(x) + \rho\_^{-}(x))dx, \quad (4.13)$$

where $(\alpha + \rho)$ is a probability of interception.

The outward stream is also weakened by absorption and reflection but is augmented by reflection of the inward stream to give a net outward flux of

$$d\_^{-}(x) = -((\alpha + \rho)\_^{-}(x) - \rho\_^{+}(x))dx, \quad (4.14)$$

where the minus sign in front of the brackets is a reminder that the outward flux is moving in a negative direction with respect to the x axis.

For the special case of uniform scattering in all directions (isotropic scattering), $\rho = \tau = \alpha/2$. It may then be shown that the bulk reflection coefficient $\rho\_$ is given by

$$\rho\_ = (1 - \alpha 0.5)/(1 + \alpha 0.5), \quad (4.15)$$

and the bulk attenuation coefficient $k\_$ is given by

$$k\_ = \alpha 0.5. \quad (4.16)$$

(These equations are not relevant to the limiting case $\alpha = 1$ when Beer's Law applies.)

When the forward beam strikes a boundary before it is completely attenuated, a fraction ρb may be reflected. The fluxes of radiation in the medium, both forwards and backwards, can then be expressed as functions of ρb, ρ, τ and of the concentration of the medium. Complex numerical methods must be deployed to obtain fluxes when condition

(b) is not satisfied, i.e. when k is a function of the direction of scattering examples of the application of Beer's Law to the atmosphere (where the assumption of single scattering is usually valid). application of the Kubelka-Munk equations is discussed with reference to crop canopies and animal coats.

Problems

1. Ultra-violet radiation in the waveband 280–320 nm incident at the top of the earth's atmosphere supplies about 20Wm−2 of radiant energy at normal incidence. Assuming a mean wavelength of 300 nm, calculate the photon flux at normal incidence.

2. At what wavelength does the peak emission from an oxyacetylene welding torch burning at 3800 K occur?

Radiation Environment

Almost all the energy for physical and biological processes at the earth's surface comes from the sun and much of environmental physics is concerned with ways in which this energy is dispersed or stored in thermal, mechanical, or chemical form. considers the quantity and quality of solar (short-wave) radiation received at the ground and the exchange of terrestrial (long-wave) radiation between the ground and the atmosphere.

Solar Radiation

1 Solar Constant

At themean distance *R* of the earth from the sun, which is 1.50×108 km, the irradiance of a surface perpendicular to the solar beam and just outside the earth's atmosphere is known as the *Solar Constant*. The name is somewhat misleading because this quantity is known to change by small amounts over periods of weeks to years in response to changes within the sun, and the preferred term to describe the irradiance at mean earth-sun distance is the *Total Solar Irradiance*, TSI.

Increasingly precise determinations of the TSI have been made from mountain tops, balloons, rocket aircraft flying above the stratosphere and, since the late 1970s, from satellites. Satellite observations have clearly demonstrated that the annual mean TSI varies by about 1.6 W m−2 between the minimum and maximum of the 11-year cycle of solar activity (Figure 1) (Frohlich and Lean, 1998; Kopp and Lean, 2011). Although satellite radiometers are capable of great precision, their absolute accuracy is much poorer. Kopp and Lean (2011) used data from an improved satellite radiometer to conclude that the most probable absolute value of the TSI representative of solar minimum activity is 1361 ± 0.5 W m−2, about 5 W m−2 lower than the value recommended in about 2000. The downward revision is attributed to improved measurement accuracy, not a change in solar activity. Indirect proxies that vary with TSI (e.g. sunspot number) suggest that the total solar irradiance (Solar Constant) may have increased by about 0.3 W m−2 since 1750, thus making a minor contribution to global warming, but reliability of this estimate is poor (IPCC, 2007). The spectrum of radiation from the sun closely resembles that of a full (black body) radiator, and the temperature of the equivalent full radiator may be readily estimated from knowledge of the TSI as follows. Assuming that the sun radiates uniformly in all directions, the

earth intercepts only a small fraction of the radiated solar energy passing across the surface of an imaginary

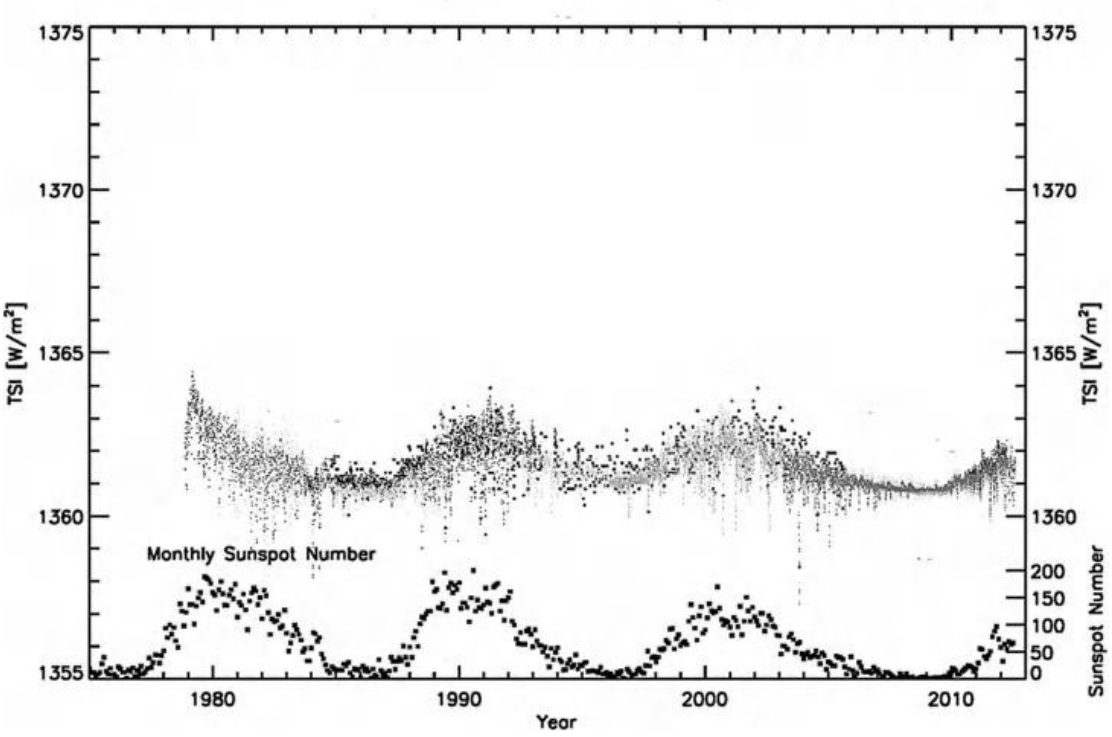


Figure 1 Composite record of total solar irradiance, TSI (the "Solar Constant"), from 1975 to 2012 compiled from multiple satellite radiometric measurements adjusted to a standard reference scale. Also shown is the monthly mean sunspot number, illustrating the strong correlation with TSI. (data courtesy of Greg Kopp, NASA.)

sphere centered at the sun and with radius *R*, the earth's mean distance from the sun *(*149.6×109 m*)*. The total rate at which energy is emitted from the sun may therefore be calculated by multiplying the TSI by the surface area of the sphere, i.e.

$E = 4\pi R2 \times 1361 = 3.83 \times 1026$ W.

The radius of the sun is $r = 6.96 \times 108$ m. Hence, assuming the surface behaves like a full radiator, its effective temperature *T* is given by

$\sigma T 4 = 3.83 \times 1026/(4\pi r 2)$, from which the (rounded) value of *T* is 5771 K.

2 Sun-Earth Geometry

Major features of radiation at the surface of the earth are determined by the earth's rotation about its own axis and by its elliptical orbit around the sun. The polar axis

about which the earth rotates is fixed in space (pointing at the Pole Star) at a mean angle of 66.5◦ to the plane of the earth's orbit (termed the *obliquity*) but with a small top-like wobble (the *precession* of the axis). The angle between the orbital plane and the earth's equatorial plane therefore oscillates between a maximum of 90–66.5 = 23.5◦ in midsummer and a minimum of −23.5◦ in midwinter with small deviations attributable to the wobble. This angle is known as the *solar declination (δ)* and its value for any date and year can be found from astronomical tables.

The shape of the earth's orbit (the *eccentricity*), obliquity, and precession each vary over millenia, with cycles ranging from about 100,000 to 23,000 years. The Russian astronomer Milutin Milankovitch proposed in the 1930s that these orbital variations were responsible for long-term cyclical changes in the earth's climate. Modern evidence from the analysis of layers of deep ocean sediments has confirmed that the *Milankovitch theory* explains part, but not all, of the variations in climate over the past few hundred thousand years.

Currently, the eccentricity of the earth's orbit is relatively small. The earth is about 3%closer to the sun in January than in July, so the irradiance at the top of the atmosphere is almost 7% larger in January (irradiance is proportional to the inverse square of the sun-earth radius). At any point on the earth's surface, the angle between the direction of the sun and a vertical axis depends on the latitude of the site, and on time t (hours), most conveniently referred to the time when the sun reaches its zenith. The hour angle θ (radians) of the sun is the fraction of 2π through which the earth has turned after local solar noon, i.e. $\theta = 2\pi t/24$. Since $2\pi \equiv 360^\circ$, each hour corresponds to 15◦ rotation.

Three-dimensional geometry is needed to show that the *zenith angle* ψ of the sun at latitude ϕ is given by:

$$\cos\psi = \sin\phi \sin\delta + \cos\phi \cos\delta \cos\theta \quad (5.1)$$

and that the azimuth angle A of the sun with respect to south is given by:

$$\sin A = -\cos\delta \sin\theta / \sin\psi. \quad (5.2)$$

The minimum zenith angle of the sun ψn occurs at local solar noon and may be found by putting $\theta = 0$ in Eq. (5.1) to give

$$\cos\psi n = \sin\phi \sin\delta + \cos\phi \cos\delta$$

$$= \cos(\phi - \delta)$$

so

$$\psi n = \phi - \delta. \quad (5.3)$$

To find the *daylength*, defined as the period for which the sun is above the horizon, the hour angle θs at sunset (i.e. the half-day length t) is first found by putting $\psi = \pi/2$ radians (90◦) in Eq. (5.1), and rearranging to give:

$$\cos\theta_s = -\frac{\sin\varphi}{\cos\varphi}\frac{\sin\delta}{\cos\delta} = -\tan\varphi\tan\delta,$$

$$\theta_s = 2\pi t/24 = \cos^{-1}(-\tan\varphi\tan\delta),$$

from which the daylength $2t$ in hours is

$$2t = (24/\pi)\theta_s = (24/\pi)\cos^{-1}(-\tan\varphi\tan\delta).$$

Some developmental processes in plants and activity in animals occur at the very weak levels of radiation received during twilight before sunrise and after sunset. The biological length of a day may therefore exceed the daylength given by Eq. (5.6) to an extent which can be estimated from the time of *civil twilight* (sun dropping to 6◦ below horizon with a correction for refraction, i.e. ψ = 96◦ or 1.68 radians) or *astronomical twilight* (down to 18◦ below horizon, ψ = 108◦ or 1.89 radians). The beginning of civil and astronomical twilight can be calculated from Eq. (5.1) by substituting these values of ψ. At the equator, the interval between civil twilight and sunrise is close to 22 min throughout the year but at latitude 50◦ it ranges from 45 min at midsummer to 32 min at the equinoxes.

For the application of Eqs. (5.1)–(5.6), values of δ may be found in astronomical tables (e.g. List, 1966) or δ may be calculated from empirical expressions such as sin $\delta = a \sin[b + cJ + d \sin(e + cJ)]$, (5.7)

where J is the calendar day with J = 1 on January 1, and the constants are a = 0.39785; b = 278.97; c = 0.9856; d = 1.9165; e = 356.6 (Campbell and Norman, 1998).

Example.

Find the solar zenith angle at local solar noon, and the day length at Edinburgh, Scotland (55.0◦N) on June 21 (calendar day 172).

Solution. Using astronomical tables or Eq. (5.7), the solar declination δ on June 21 is 23.5◦. Equation (5.1) could be used to calculate the solar zenith angle ψ, but since the time is solar noon, when θ = 0, the much simpler Eq. (5.3) applies, and

$\psi = \varphi - \delta = 55.0 - 23.5 = 31.5^\circ$

The day length is given by Eq. (5.6),

$2t = (24/\pi) \cos{-1} (-\tan \phi \tan \delta),$

$= (24/\pi) \cos{-1} (-\tan 55.0 \tan 23.5),$

$= (24/\pi) \cos{-1} (-0.621) = 24 \times 2.24/\pi = 17.1$ h

(note that $\cos{-1} (-0.621)$ is 2.24 radians).

3 Spectral Quality

For biological work, the spectrum of solar radiation can be divided into three major wavebands shown in Table 5.1 with the corresponding fraction of the TSI. Most measurements of solar energy at the ground are confined to the visible and near-infra-red wavebands which contain energy in almost equal proportions. **Table 5.1** Distribution of Energy in the Spectrum Emitted by the Sun.

Waveband (nm)	Energy (%)
0–200	0.7
200–280 (UV-C)	0.5
280–320 (UV-B)	1.5
320–400 (UV-A)	6.3
400–700 (visible/PAR)	39.8
700–1500 (near-infra-red)	38.8
$1500 - \infty$	12.4

Ultra-violet radiation contains sufficient energy per quantum to damage living cells. The ultra-violet spectrum is divided into UVA (320–400 nm) responsible for tanning the skin; UVB (280–320 nm) responsible for skin cancer and Vitamin D synthesis; and UVC (200–280 nm), potentially the most harmful waveband but absorbed almost completely by molecular oxygen in the stratosphere. Human skin is about 1000 times more sensitive to the UVB range than to the UVA. The waveband to which the eyes of humans and most terrestrial animals are sensitive ranges from blue (400 nm) through green (550 nm) to red (700 nm), with maximum sensitivity at around 500 nm. Eyes of aquatic mammals have peak sensitivity at slightly shorter wavelengths,

around 488 nm, perhaps a consequence of the “blueness” of the ocean habitat (Mcfarland and Munz, 1975). Eyes of fish have similar peak sensitivity.

Photosynthesis is stimulated by radiation in the samewaveband as human vision, and this waveband is referred to as *Photosynthetically Active Radiation* (PAR), a misnomer because it is green cells which are active, not radiation. Initially, the term PAR was applied to radiation measured in units of energy flux density (W m−2) but for two reasons it is more appropriately expressed as quantum flux density (mol m−2 s−1):

(i) when photosynthesis rates are compared for light of different quality (e.g. from the sun and from lamps), they are more closely related to the quantum content than to the energy content of the radiation (Jones, 1992) and (ii) because the number of moles of carbon dioxide fixed in photosynthesis is closely proportional to the number of moles of photons absorbed in the PAR waveband. The fraction of PAR to total energy in the extraterrestrial solar spectrum is about 0.40 (Table 5.1) but it is closer to 0.50 for solar radiation at the earth’s surface (p. 68) because the atmosphere absorbs almost all the ultra-violet wavelengths and a significant part of the solar infra-red spectrum. Many developmental processes in green plants have been found to depend on the state of the pigment *phytochrome* which exists in two photo-interconvertible forms that absorb radiation in wavebands centered at 660 nm (red light—the Pr form) and 730 nm (far-red light—the Pfr form). The ratio of the two forms present in plant tissue changes in response to the ratio of spectral irradiance at these wavelengths, known as the red:far-red ratio (R:FR), so the phytochrome is an effective detector of the quality of radiation, for example detecting shading caused by other plants. *Shade avoidance* is one of the most important competitive strategies that plants possess. It consists of a range of physiological and developmental responses by plants that provides competitive advantages over other species when growing in the shade of other plants, e.g. weeds in crops or understory species in forest canopies. The responses may influence germination, stem elongation, leaf development, flowering, and reproduction. Shade avoidance responses are all initiated by a single environmental signal, a reduction in the ratio of red (R) to far-red (FR) (i.e. R:FR) radiation that occurs in crowded plant communities (Smith and Whitelam, 1997).

Attenuation of Solar Radiation in the Atmosphere

As the solar beam passes through the earth's atmosphere, it is modified in quantity, quality, and direction by processes of scattering and absorption (Figure 5.2).

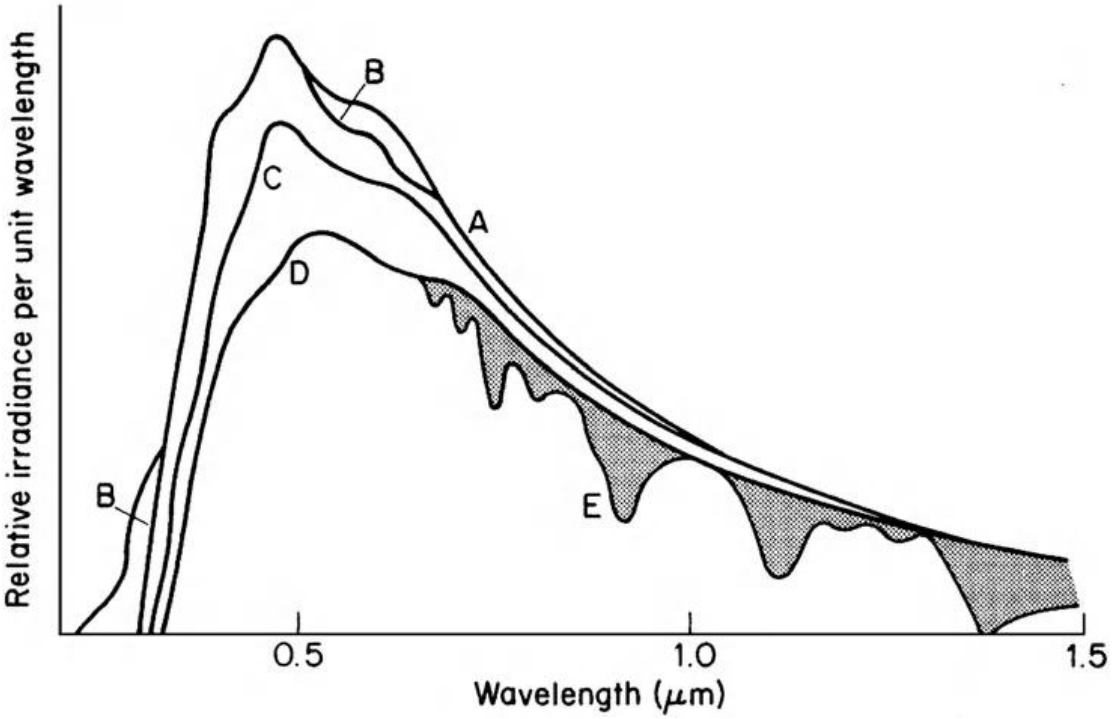


Figure 2 Successive processes attenuating the solar beam as it penetrates the atmosphere. A— extraterrestrial radiation, B—after ozone absorption, C—after molecular scattering, D—after aerosol scattering, E—after water vapor and oxygen absorption (from Henderson, 1977).

Scattering has two main forms. First, individual quanta striking molecules of any gas in the atmosphere are diverted in all directions, a process known as *Rayleigh scattering* after the British physicist who showed theoretically that the effectiveness of molecular scattering is proportional to the inverse fourth power of the wavelength. The scattering of blue light *(*λ = 400 nm*)* therefore exceeds the scattering of red (700 nm) by a factor of (7/4)4 or about 9. This is the physical basis of the blue color of the sky as seen from the ground and the blue haze surrounding the Earth when viewed from space. The redness of the sun's disk near sunrise or sunset is further evidence that blue light has been removed from the beam preferentially. Rayleigh also showed that the spatial distribution of scattered radiation was proportional to *(*1 + cos2 θ*)*, where

θ is the angle between the initial and scattered directions of the radiation. Thus the probability of forward- and backscattering is twice that at 90◦. Rayleigh scattering is confined to systems in which the diameter *(d)* of the scatterer is much smaller than the wavelength of the radiation λ. This condition is not met for particles of dust, smoke, pollen, etc., in the atmosphere, referred to as "aerosol," which often have diameters d in the range $0.1\lambda < d < 25\lambda$. The theory developed by Mie predicts that the wavelength dependence of scattering by aerosol particles should be a function of d/λ, and that, for some values of the ratio, longer wavelengths should be scattered more efficiently than short—the reverse of Rayleigh scattering. This happens rarely, but smoke with the appropriate narrow range of particle sizes, e.g. from forest fires, can occasionally cause the sun and moon viewed from the earth to appear blue! Usually, aerosol contains such a wide range of particle sizes that scattering is not strongly dependent on λ. Angstrom (1929) proposed that the dependence could often be described by a power law, i.e. $\propto \lambda^{-\alpha}$ where α had an average value of 1.3, and many investigators have confirmed the power law dependence. For example, a set of measurements in the English Midlands corresponded to α between 1.3 and 2.0 (McCartney and Unsworth, 1978). When particles are large and sufficiently dense for multiple scattering, there is essentially no wavelength dependence and the scattered light appears white, for example with very hazy skies in summer. Aerosol scattering is usually predominantly "forwards," i.e. in a narrow cone around the direction of the incident radiation.

Sources and Radioactive Properties of Aerosols

Aerosols are solid or liquid particles small enough to remain suspended in the atmosphere for long periods. They have an important *direct* influence on radiation reaching the ground because they scatter and absorb solar and long-wave radiation in the atmosphere, thus contributing to *radiative forcing* of the earth's energy budget. Aerosols also influence cloud albedo and duration by increasing the number of droplets in clouds and altering the efficiency of precipitation production, thus also having *indirect* effect on radiation (IPCC, 2007). These indirect effects are hard to isolate in the free atmosphere, but Coakley et al. (1987) demonstrated that they can be seen in "ship tracks" observed from space (Figure 3). When there is a thin layer of marine stratus cloud, the aerosols emitted from ships smokestacks increase the cloud density behind the vessel.

Aerosols in the lower atmosphere have relatively short lifetimes before they are removed by precipitation and turbulent transfer. Aerosols in the stratosphere, for example as a result of explosive volcanic eruptions, have much longer lifetimes and may be dispersed around the globe, causing observable effects on the earth's climate (Hansen et al., 2011).

The size distribution of aerosols is critical for their radiative effects. Particles in the "accumulation size range" (Chapter 12) (i.e. with diameters between about 0.1 and 1.0 μm) scatter more light per unit mass than larger particles, and have longer lifetimes in the atmosphere, so are particularly important in influencing the irradiance at the ground. Some aerosols are *hygroscopic*, i.e. they absorb water depending on atmospheric humidity, thus changing the aerosol size distribution. Examples are sea-salt particles and ammonium sulfate.

Primary aerosols are generated at the surface by natural processes and human activity. Desert dust storms generate aerosols that can be transported across the Atlantic and Pacific oceans (Kaufman et al., 2002). Inefficient combustion of wood or fossil fuels releases organic and black carbon aerosols.

Secondary aerosols are created in the atmosphere by chemical reactions. An ubiquitous example is ammonium sulfate aerosol, formed by reactions that include the gases ammonia, sulfur dioxide, and dimethyl sulfide, which have natural sources (volcanoes, ocean plankton) and human sources (fossil fuel, animal production). Other secondary aerosols are created during photochemical smog episodes.

Absorption of radiation by aerosols is very variable. For example, desert dust may absorb little, but black carbon aerosols from wildfires and human activity are much stronger absorbers (Hansen et al., 2004). Absorption by secondary aerosols can be greatly increased when black carbon particles are incorporated into the aerosol. For several reasons, radiative effects of aerosols are generally much more difficult to assess than those of atmospheric gases: aerosols are not uniformly distributed, and may be formed and transformed in the atmosphere; some aerosol types (e.g. dust and sea salt) consist of particles whose physical properties that influence scattering and absorption have wide ranges; and aerosol species may combine to form mixed particles with optical properties different from their precursors. IPCC (2007) included a good review of progress in understanding aerosol effects on the earth's energy balance.

The second process of attenuation is *absorption* by atmospheric gases and aerosols. The most important gases absorbing solar radiation are ozone and oxygen (particularly in the ultra-violet spectrum), and water vapor and carbon dioxide (with strong bands in the infra-red). Absorption by aerosols is very variable, depending on the source of the material.

In contrast to scattering, which simply changes the direction of radiation, absorption removes energy from the beam so that the aerosol, and atmosphere containing it, is heated. In the ultra-violet, absorption by oxygen and ozone in the stratosphere removes.

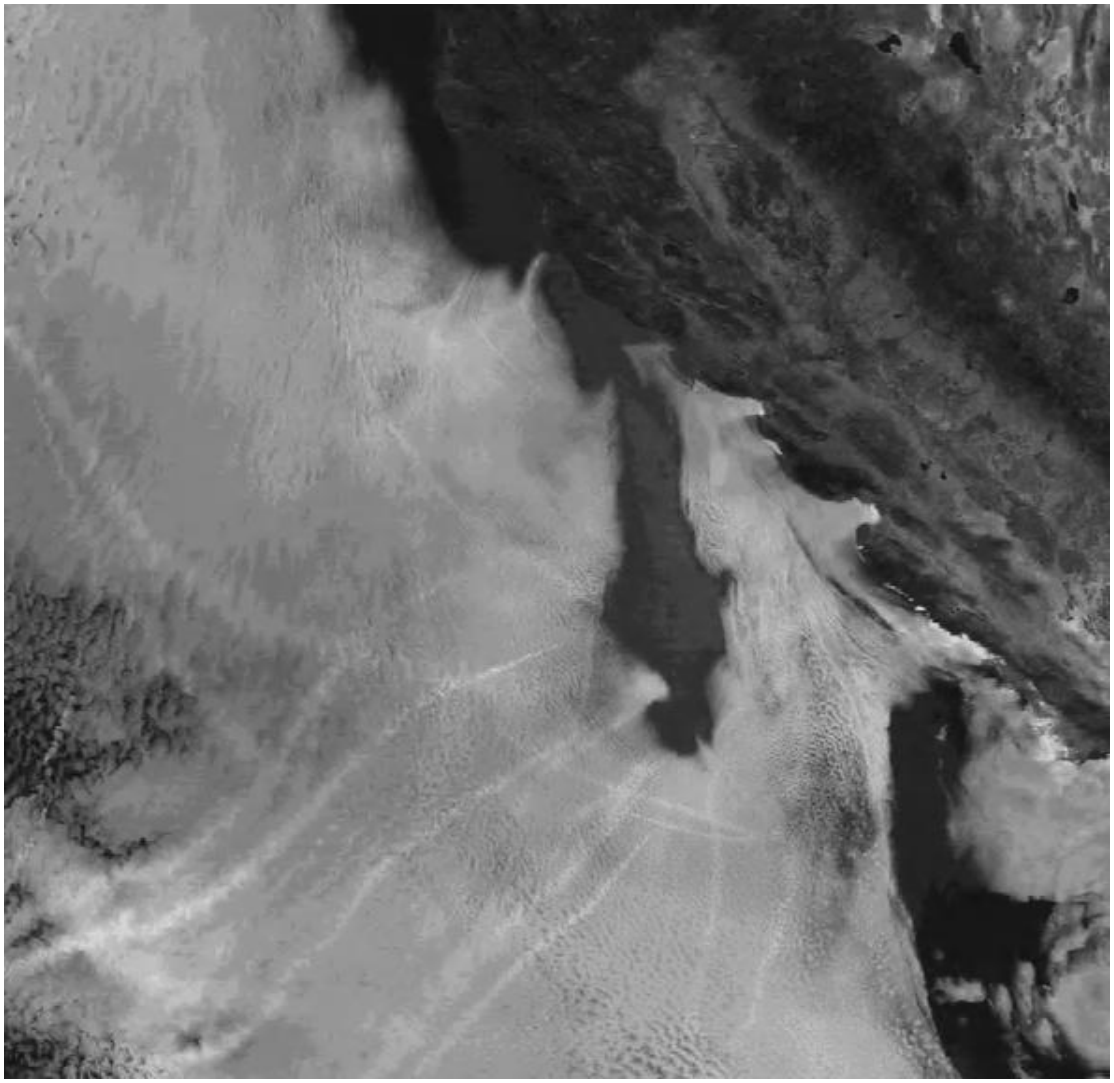


Figure 3 Infra-red satellite image of reflected solar radiation at 2.1μm off the coast of California near San Francisco. Pollution from ship smoke emissions increases the reflectivity of marine stratus clouds by supporting the formation of larger numbers of

smaller droplets than in unpolluted cloud nearby. This enables ship tracks to be seen clearly in the image. (Courtesy of Dr. J. A. Coakley)

All UVC and most UVB radiation before it reach the ground, resulting in stratospheric heating. The remaining UVB and UVA radiation is scattered very effectively however, so that it is possible to suffer serious sunburn beneath a cloudless sky even if not exposed to the direct solar beam.

In the visible region of the spectrum, absorption by atmospheric gases is much *less* important than molecular scattering in determining the spectral distribution of solar energy at the ground. In the infra-red spectrum, however, absorption is much *more* important than scattering, because several atmospheric constituents absorb strongly, notably water vapor with absorption bands between 0.9 and 3 μm. The presence of water vapor in the atmosphere thus increases the amount of visible radiation relative to infra-red radiation.

The scale of absorption and scattering in the atmosphere depends partly on the path length of the solar beam and partly on the amount of the attenuating constituent present in the path. The path length is usually specified in terms of an *air mass number m*, which is the length of the path relative to the vertical depth of the atmosphere. Air mass number therefore depends on altitude (represented by the pressure exerted by the atmospheric column above the site) and on zenith angle ψ. For values of zenith angle ψ less than 80◦, the air mass number at a location where atmospheric pressure is P is simply

$m = (P/P0)\sec\psi$, where $P0$ is standard atmospheric pressure at sea level (101.3 kPa), but for values between 80◦ and 90◦ ,m is smaller than $\sec\psi$, because of the earth's curvature. Values, corrected for refraction, can be obtained from tables (e.g. List, 1966).

The most variable absorbing gas in the atmosphere is water vapor, the amount of which can be specified by a depth of *precipitable water u*, defined as the depth of water that would be formed if all the vapor were condensed (u is typically between 5 and 50 mm at most stations). If the precipitable water is u, the path length for water vapor is um. Similarly, the total amount of ozone in the atmosphere (the *ozone column*) is specified by an equivalent depth of the pure gas at a standard pressure of one atmosphere (101.3 kPa). At mid-latitudes, the ozone column is typically about 3 mm and varies only slightly with season. However over some parts of the Antarctic,

up to 60% of the ozone column is lost during the Antarctic Spring (September-October). When this phenomenon was first reported by Joe Farman of the British Antarctic Survey, who analyzed surface observations of irradiance in the UVB waveband (Farman et al., 1985), it could not be explained by atmospheric chemists and had not been detected by satellite monitoring. In retrospect, it turned out that a computer program had caused the anomalous Antarctic Spring data from the satellite to be ignored, and chemists traced the cause of the alarming decrease in ozone to the presence in the Antarctic stratosphere of frozen particles of nitric acid that served as catalytic sites for reactions destroying ozone. Figure 5.4 shows the decline of ozone column thickness from the mid-1950s in the Antarctic Spring. The extremely cold stratospheric conditions necessary for ozone destruction by this mechanism are less common over the Arctic, and fortunately do not exist over the large parts of the earth where plant and animal populations would be vulnerable to the extra UVB radiation that would reach the surface if ozone were depleted.

In contrast with the mainly regional effects of ozone depletion, the consequences of radiative absorption by the steadily increasing amount of carbon dioxide in the earth's atmosphere are apparent on a global scale as discussed in Chapter 2 and extensively reviewed by in Assessment Reports by the IPCC (2007).

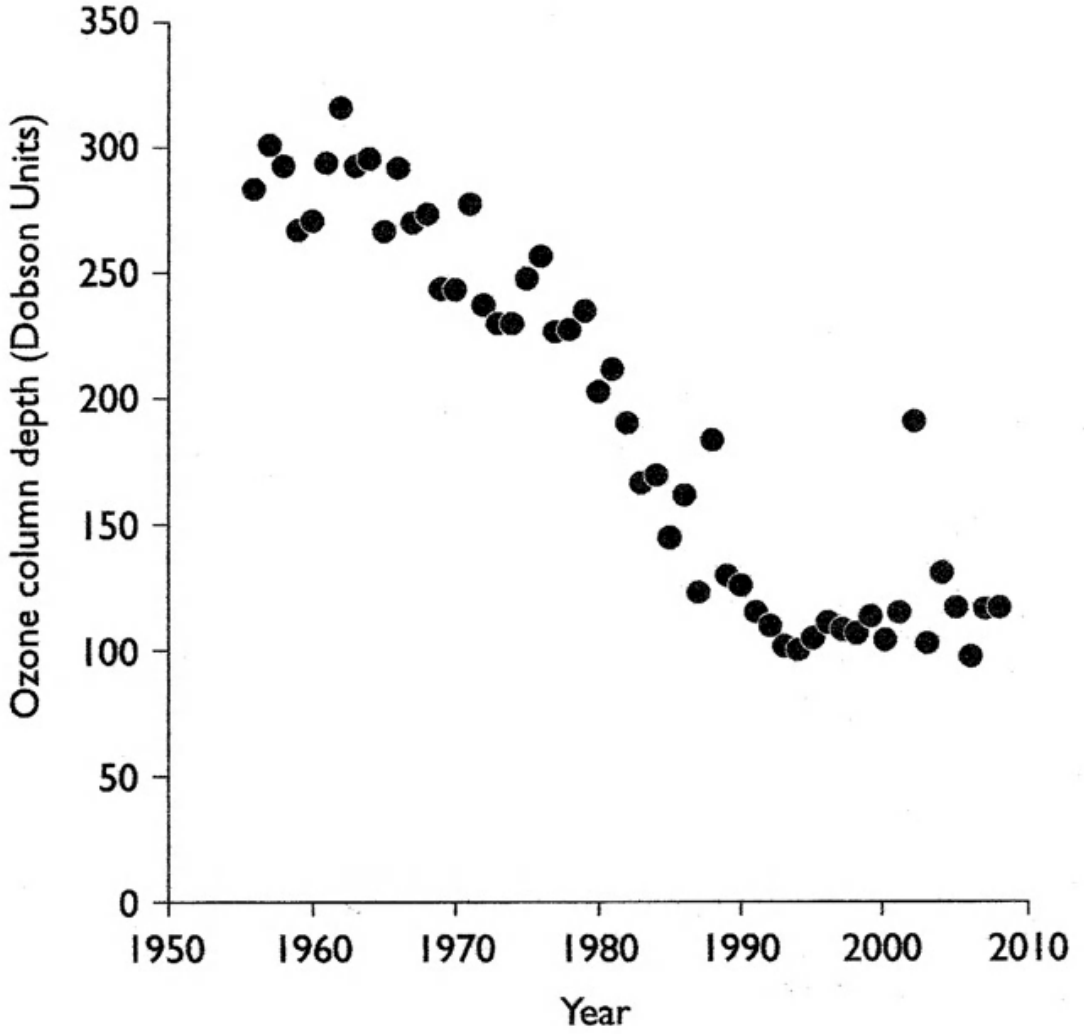


Figure 4 Minimum October ozone column depth at the British Antarctic Survey research station, Halley Bay, Antarctica from 1956 to 2009. (One Dobson Unit corresponds to a layer of gaseous ozone 10 μm thick at STP) (data courtesy of the British Antarctic Survey.)

Clouds, consisting of water droplets or ice crystals, scatter radiation both forwards and backwards, but when the depth of cloud is substantial, back-scattering predominates and thick stratus can reflect up to 70% of incident radiation, appearing snow-white from an aircraft flying above it. About 20% of the radiation may be absorbed, leaving only 10% for transmission, so that the base of such a cloud seems gray. However, at the edge of a cumulus cloud, where the concentration of droplets is small, forward scattering is strong—the silver lining effect—and under a thin sheet of cirrus the reduction of irradiance can be less than 30%, see Figure 10.

5.3 Solar Radiation at the Ground

As a consequence of attenuation, radiation has two distinct directional properties when it reaches the ground. *Direct* radiation arrives from the direction of the solar disk and includes a small component scattered directly forward. The term *diffuse* describes all other scattered radiation received from the blue sky (including the very bright aureole surrounding the sun) and from clouds, either by reflection or by transmission. The sum of the energy flux densities for direct and diffuse radiation is known as *total* or *global* radiation and for climatological purposes is measured on a horizontal surface. The symbols **S**b, **S**d, and **S**t describe direct, diffuse, and total irradiance, respectively, on a horizontal surface, and **S**p signifies direct irradiance measured at right angles to the solar beam.

Direct Radiation

At sea level, direct radiation **S**p rarely exceeds 75% of the Solar Constant i.e. about 1000 Wm−2. The minimal loss of 25% is attributable to molecular scattering and absorption in almost equal proportions with a negligible contribution from aerosol when the air mass is clean. Aerosol increases the ratio of diffuse to global radiation by forward scattering and changes the spectral composition. Several expressions are available to describe atmospheric transmissivity as affected by molecular and aerosol components. A simple practical relation is

$$\mathbf{S}p = \mathbf{S}*T^{m}, \quad (5.8)$$

where $\mathbf{S}*$ is the Solar Constant (TSI), T is the *atmospheric transmissivity*, and m is the air mass number. Liu and Jordan (1960) found that T ranged from about 0.75 to 0.45 on cloudless days implying, as above, that aerosol attenuation was insignificant on the clearest days.

To illustrate more directly the combined impact of attenuation by aerosols and molecules, Beer's Law may be applied to give

$$\mathbf{S}p(\tau) = \mathbf{S}* \exp(-\tau m), \quad (5.9)$$

where τ is an *extinction coefficient* or *optical thickness* and m is the air mass number. When the value of τ is expressed as the sum of molecular extinction (τm) and aerosol extinction (τa), Eq. (5.9) may be written in the form

$$\mathbf{S}p(\tau) = \mathbf{S}* \exp(-\tau m m) \exp(-\tau a m) = \mathbf{S}p(0) \exp(-\tau a m), \quad (5.10)$$

Where $\mathbf{S}p(0)$ is the irradiance of the direct beam below an atmosphere free of aerosol. Comparison of Eqs. (5.8), (5.9), and (5.10) yields

$T = \exp(-\tau) = \exp(-(\tau m + \tau a))$ (5.11)

implying that τm was about 0.3 in Liu and Jordan's measurements (assuming $\tau a = 0$ when $T = 0.75$), and τa was about 0.5 on the most turbid days that they recorded. The value of τa at a site can be determined from Eq. (5.10) by measuring $\mathbf{S}p(\tau)$ as a function of $m(=\sec \psi)$ and by calculating $\mathbf{S}p(0)$ (also a function of m) from the properties of a clean atmosphere containing appropriate amounts of gases and water vapor. For example, a series of measurements in Britain gave values of τa ranging from 0.05 for very clear air of Arctic origin to 0.6 for very polluted air in the English Midlands during a stagnant anticylone (Unsworth and Monteith, 1972). Corresponding values of $\exp(-\tau am)$ for $\psi = 30^\circ$ are 0.92 and 0.50, indicating radiant energy losses of up to 50% from the direct solar beam due to aerosol scattering and absorption.

Using a bulk aerosol optical thickness such as τa, determined from measurements of a broad waveband of solar radiation, is convenient when only standard radiometers are available. However, values of aerosol optical thickness derived from more sophisticated spectral observations in narrow wavebands are more commonly reported, e.g. for wavebands around 0.55 μm in the mid-visible spectrum (IPCC, 2007). Such values are generally consistent with τa and are similar in magnitude because the solar spectrum peaks around 0.5 μm.

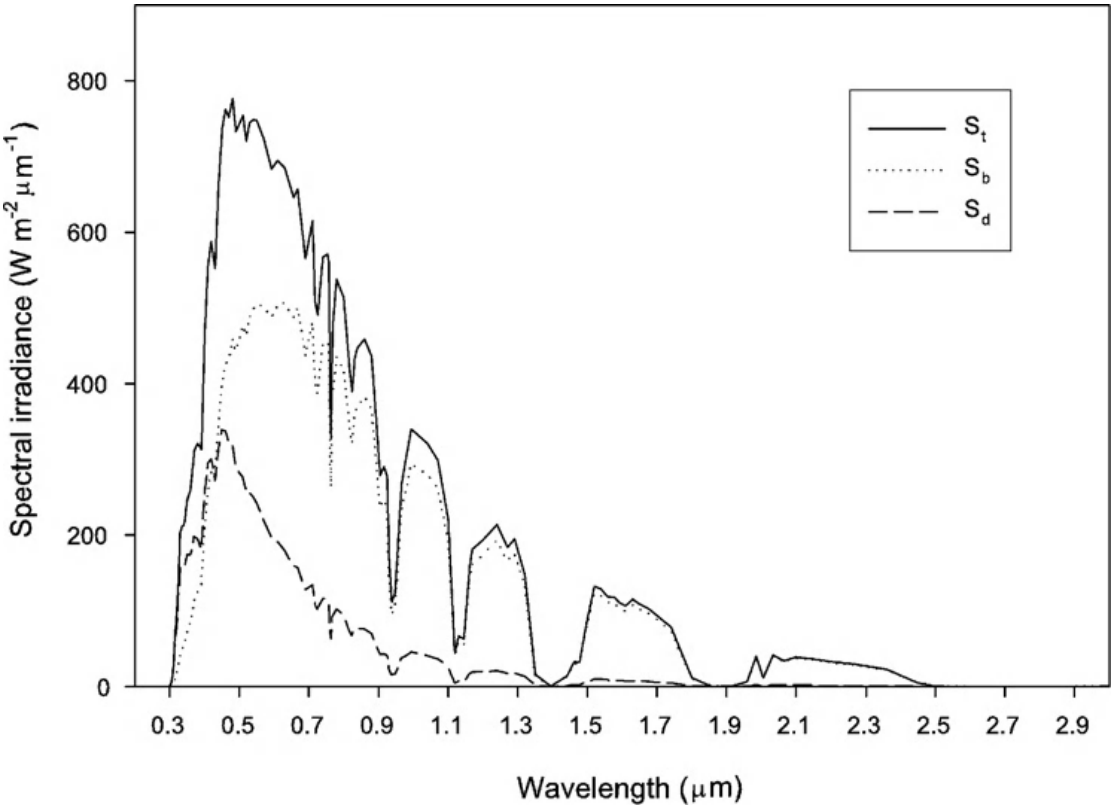


Figure 5 Spectral distribution of direct, diffuse, and total solar radiation calculated using a simple model of a cloudless atmosphere (courtesy of the Solar Energy Research Institute; see Bird and Riordan, 1984). Solar zenith angle is 60◦ *(m* = 2*)*, precipitable water is 20 mm, ozone thickness is 3 mm, and aerosol optical depth is 0.2. Note that the diffuse flux has maximum energy per unit wavelength at about 0.46 μm.

The spectrum of direct radiation depends strongly on the path length of the beam and therefore on solar zenith angle. A spreadsheet model for calculating direct and diffuse spectral irradiance (based on work by Bird and Riordan (1986)) is available from the Solar Energy Research Institute, Golden, Colorado (http://rredc.nrel.gov/solar/models/spectral/). Figure 5 shows results at sea level calculated from the model, indicating that the spectral irradiance calculated for the solar beam is almost constant between 500 and 700 nm whereas, in the corresponding extraterrestrial spectrum, irradiance decreases markedly as wavelength increases beyond $\lambda m \approx 500$ nm (Figure2). The difference is mainly a consequence of energy removed from the beam by Rayleigh scattering which increases as wavelength decreases; ozone absorption is implicated too. As zenith angle increases, attenuation by scattering becomes very pronounced, and the wavelength for maximum direct solar

irradiance moves into the infra-red waveband when the sun is less than 20◦ above the horizon (see Figure 5).

The measurements by Unsworth and Monteith (1972) also showed that, for zenith angles between 40◦ and 60◦, the ratio of visible to all-wavelength radiation in the direct solar beam decreased from a maximum of about 0.5 in clean air to about 0.4 in very turbid air. The maximum ratio exceeds the figure of 0.40 for extraterrestrial radiation (Table 5.1) because losses of visible radiation by scattering are more than offset by losses of infra-red radiation absorbed by water vapor and oxygen (Figure 2). McCartney (1978) found that the quantum content of *direct* radiation increased with turbidity from a minimum of about 2.7 μmol J−1 total radiation in clean air to about 2.8 μmol J−1 in turbid air. Theoretical values can be calculated using the spectral irradiance model referred to above (Confusion has arisen in the literature Radiation where quantum content values for *direct* radiation are compared with values referring to *global* radiation, as given later.)

2 Diffuse Radiation

Beneath a clean, cloudless atmosphere, the diffuse irradiance **S**d reaches a broad maximum somewhat less than 200Wm−2 when ψ is less than about 50◦, and the ratio of diffuse to global irradiance then falls between 0.1 and 0.15. As turbidity increases, so does **S**d/**S**t, and forψ <60◦, observations in the English Midlands fit the relation

Sd/**S**t = 0.68τa + 0.10. (5.12)

Forψ > 60◦

, **S**d/**S**t is also a function of ψ and is larger than Eq. (5.12) predicts.

With increasing cloud amount also, **S**d/**S**t increases and reaches unity when the sun is obscured by dense cloud: but the absolute level of **S**d is maximal when cloud cover is about 50%. The spectral composition of diffuse radiation is also strongly influenced by cloudiness. Beneath a cloudless sky, diffuse radiation is predominantly within the visible spectrum (Figure 5.4) but as cloud increases, the ratio of visible/all-wavelength radiation decreases toward the value of about 0.5 characteristic of global radiation.

3 Angular Distribution of Diffuse Radiation

Under an overcast sky, the flux of solar radiation received at the ground is almost completely diffuse. If it were perfectly diffuse, the radiance of the cloud base

observed from the ground would be uniform and would therefore be equal to $\mathbf{S}d/\boldsymbol{\pi}$ from Eq. (4.8).

The source providing this distribution is known as a *Uniform Overcast Sky* (UOS). In practice, the average radiance of a heavily overcast sky is between two and three times greater at the zenith than the horizon (because multiple-scattered radiation is attenuated by an air mass depending on its perceived direction and so regions near the horizon appear relatively depleted). To allow for this variation, ambitious architects and pedantic professors describe the radiance distribution of overcast skies as a function of zenith angle given by

$$\mathbf{N}(\psi) = \mathbf{N}(0)(1 + b \cos \psi)/(1 + b). \quad (5.13)$$

This distribution defines a *Standard Overcast Sky*. Measurements indicate that the number $(1 + b)$ which is the ratio of radiance at the zenith to that at the horizon is typically in the range 2.1–2.4 (Steven and Unsworth, 1979), values supported by theoretical analysis (Goudriaan, 1977) which also shows the dependence on surface reflectivity. The value $(1 + b) = 3$, in common use, is based on photometric studies and significantly overestimates the diffuse irradiance of surfaces.

Under a cloudless sky, the angular distribution of skylight depends on the position of the sun and cannot be described by any simple relation. In general, the sky round the sun is much brighter than elsewhere because there is a preponderance of scattering in a forward direction but there is a sector of sky about 90_ from the sun where the intensity of skylight is below the average for the hemisphere (Figure 6). On average, the diffuse radiation from a blue sky tends to be stronger nearer the horizon than at the zenith. As the atmosphere becomes more dusty, the general effect is to reduce the radiance of the circumsolar region and to increase the relative radiance of the upper part of the sky at the expense of regions near the horizon. Consequently the angular distribution of radiance becomes more uniform as turbidity increases.

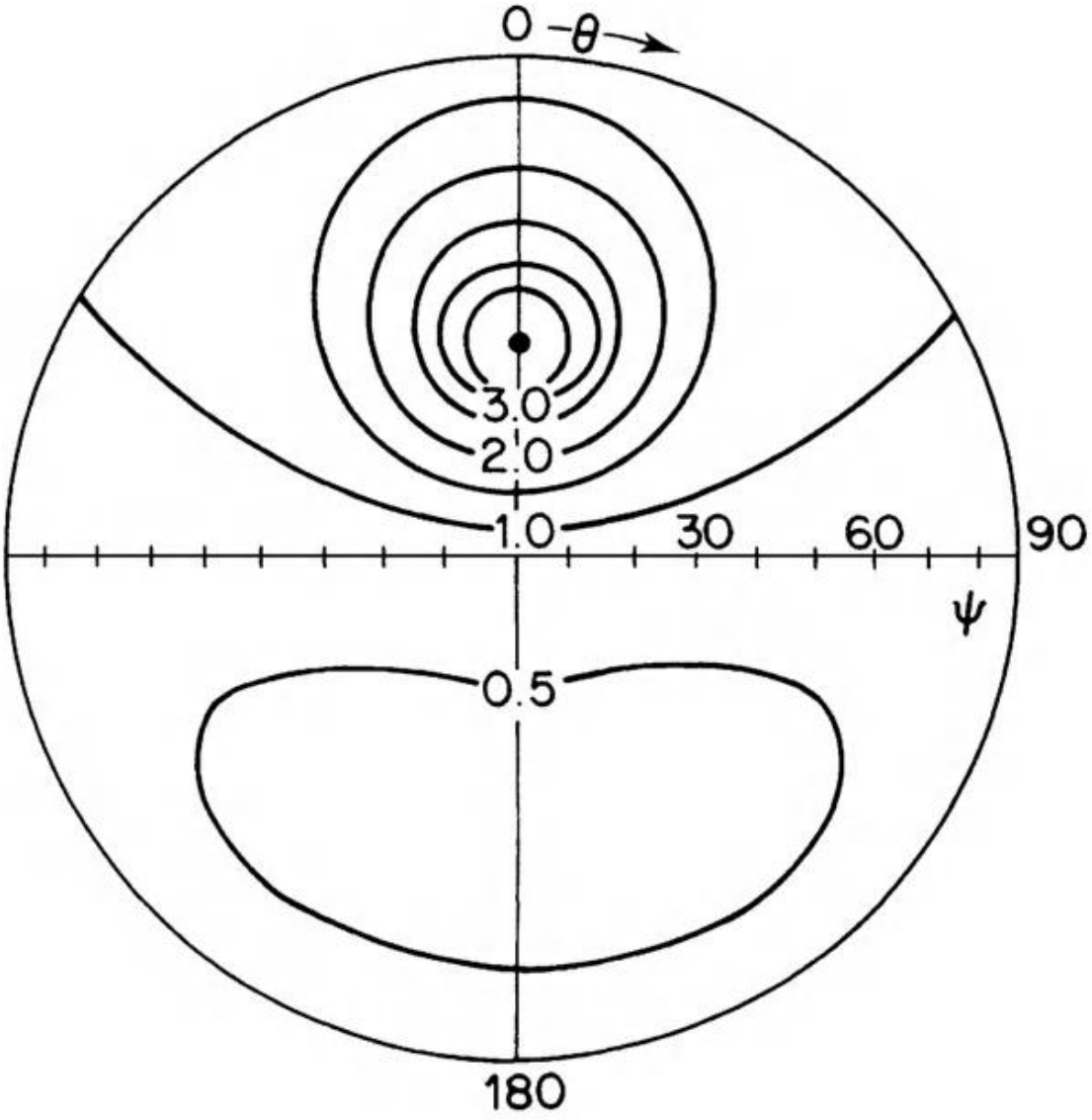


Figure 6 Standard distribution of normalized sky radiance π**N**/**S**d for solar zenith angle 35◦, where **N** is the value of sky radiance at a point and π**N** is the diffuse flux which the surface would receive if the whole sky were uniformly bright (see p. 46) (from Steven, 1977).

5.3.4 Total (Global) Radiation

The total (global) radiation on a horizontal surface is given formally by

St = **S**p cosψ + **S**d

= **S**b + **S**d, (5.14)

where **S**b = **S**p cosψ is the contribution from the direct beam.

Figure 7 contains an example of measured values of **S**t, **S**d, and **S**b as a function of solar zenith angle at Sutton Bonington, England, and Figure 5.8 shows similar data (with the addition of **S**p) from a more southerly site at Eugene, Oregon. **Estimating**

Irradiance Under Cloudless Skies A simple spreadsheet model that enables calculations of **S**p, **S**d, and **S**t under cloudless skies at any location and date for specified values of water vapor, ozone

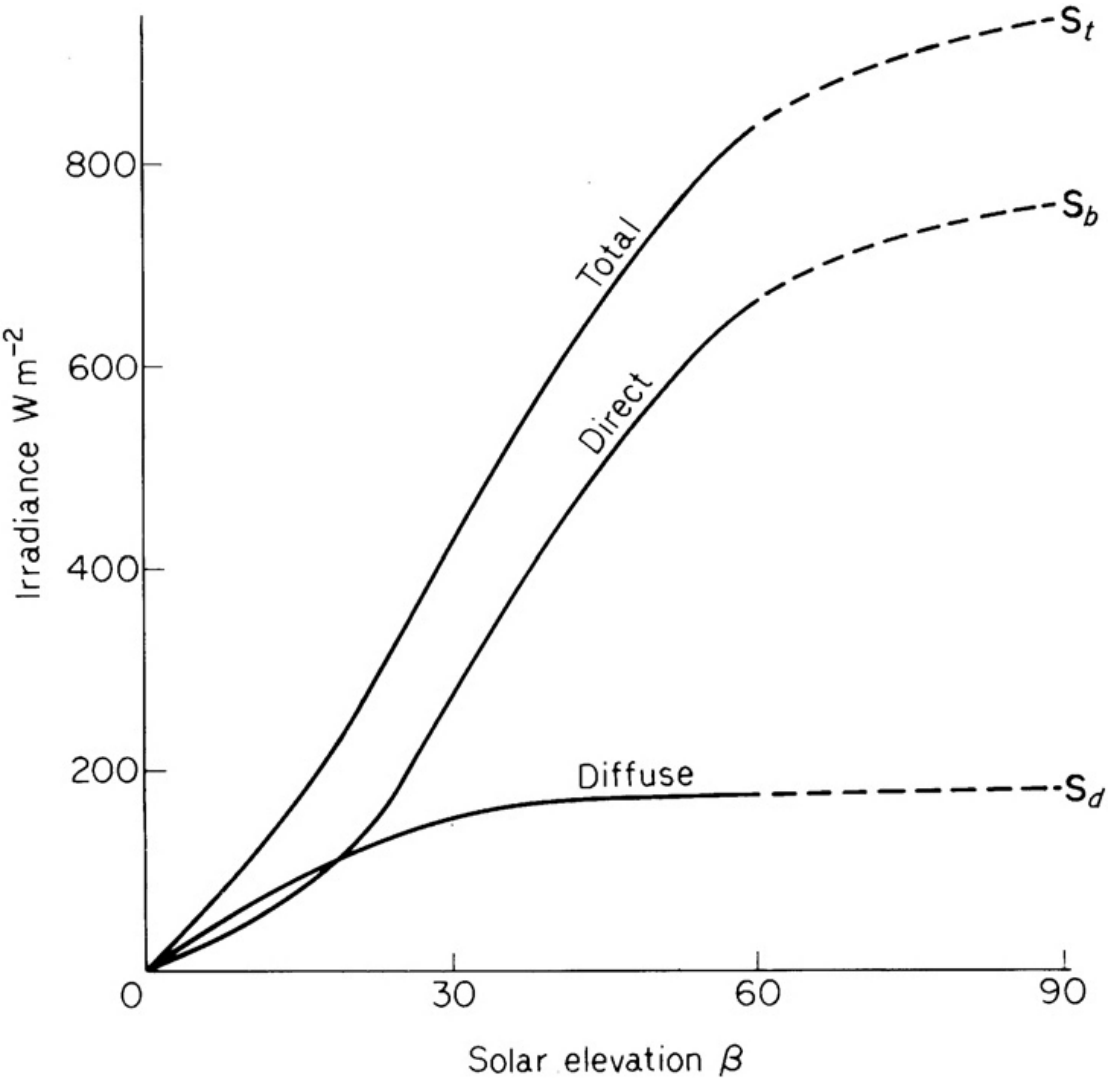


Figure 7 Solar irradiance on a cloudless day (16 July 1969) at Sutton Bonington (53◦N, 1◦W): **S**t total flux; **S**d diffuse flux, **S**b direct flux on a horizontal surface.

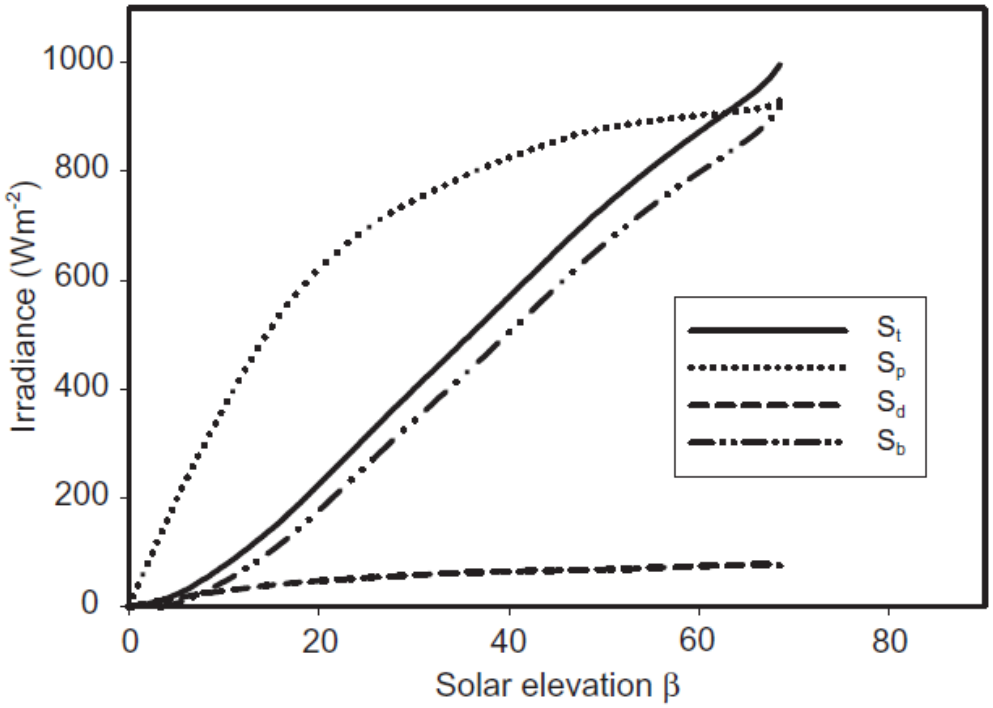


Figure 8 Solar irradiance on a cloudless day (12 June 2002) at Eugene,Oregon (44◦N, 123◦W): **S**t total flux; **S**p direct flux at normal incidence; **S**d diffuse flux; **S**b direct flux on a horizontal surface. (data courtesy of Frank Vignola, University of Oregon.)

and aerosol optical thickness is based on the work of Bird and Hulstrom (1981)

and is available from the Solar Energy Research Institute, Golden, Colorado (http://rredc.nrel.gov/solar/models/clearsky/). The following simpler approach for direct and diffuse (and hence total) radiation gives estimates of appropriate accuracy for biological calculations. Using Eq. (5.10), it can be shown that

$$\mathbf{S}p = \mathbf{S}* \exp(-\tau m m) \exp(-\tau a m). \quad (5.15)$$

The Solar Constant $\mathbf{S}*$ is 1361Wm−2 (Section 5.1.1), the optical thickness for molecular attenuation τm is typically about 0.3 (but changes with the amount of water vapor and other absorbing gases in the atmosphere), and the aerosol optical thickness τa is typically in the range 0.05–0.50. Given appropriate values of these parameters, **S**p can be calculated for a specific air mass number m (where the dependence of m on solar zenith angle ψ is given by $m = (P/P0) \sec\psi$ (p. 56).

Approximate values of **S**d as a function of m on cloudless days can be estimated from an empirical equation based on measurements by Liu and Jordan (1960), i.e.

$$\mathbf{S}d = 0.3S*[1 - \exp(-(\tau m + \tau a)m] \cos\psi. \quad (5.16)$$

The total irradiance is then given by

St = **S**p cosψ + **S**d.

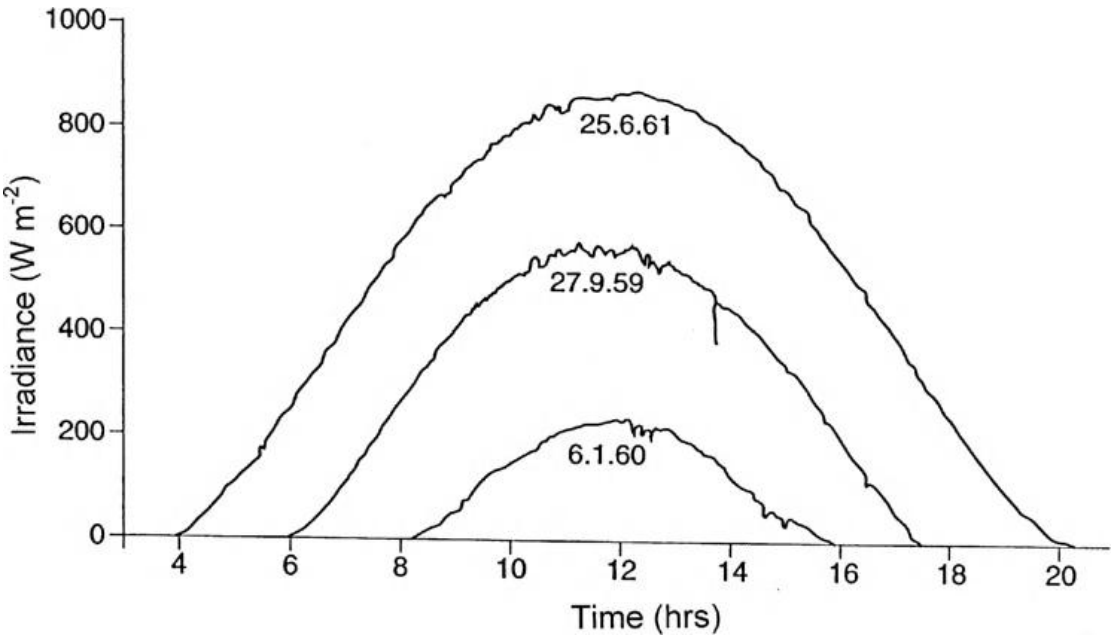


Figure 9 Solar radiation on three cloudless days at Rothamsted, England (52◦N, 0◦W).During the middle of the day, the record tends to fluctuate more than in the morning and evening, suggesting a diurnal change in the amount of dust in the lower atmosphere, at least in summer and autumn.

On cloudless days, illustrated in Figure 5.9, the change of **S**t with time is approximately sinusoidal. This form is distorted by cloud, but in many climates the degree of cloud cover, averaged over a period of a month, is almost constant throughout the day so the monthly average variation of irradiance over a day is again sinusoidal. In both cases, the irradiance at t hours after sunrise can be expressed as

St = **S**tm sin *(πt/n),* (5.17)

where **S**tm is the maximum irradiance at solar noon and *n* is the day length in hours. This equation can be integrated to give an approximate relation between maximum irradiance and the daily integral of irradiation (the *insolation*) by writing

$$\int_0^n \mathbf{S}_t dt \approx 2\mathbf{S}_{tm} \int_0^{n/2} \sin(\pi t/n) dt = (2n/\pi)\mathbf{S}_{tm}.$$

For example, over southern England in summer, **S**tm may reach 900Wm−2 on a clear cloudless day and with n = 16 h = 58×103 s, the insolation calculated from equation (5.18) is 33 MJ m−2 compared with a measured maximum of about 30 MJ m−2. In Israel, where **S**tm reaches 1050Wm−2 in summer for a day length of 14 h, the equation gives an insolation of 34MJ m−2 compared with 32 MJ m−2 by measurement.

At higher latitudes in summer when dawn and dusk are prolonged, a full sine wave may be more appropriate than Eq. (5.18). If **S**t is given by

$\mathbf{S}t \approx \mathbf{S}tm(1 - \cos 2\pi t/n) = \mathbf{S}tm \sin 2 (\pi t/n)$ (5.19)

integration yields

$$\int_0^n \mathbf{S}_t dt = \mathbf{S}_{tm} n/2.$$

showed that the radiation regime at Aberdeen (57◦N) was described best by the average of values given by Eqs. (5.18) and (5.20).

In most climates, the daily receipt of total solar radiation is greatly reduced by cloud for at least part of the year. Figure 5.10 shows the extent to which the total irradiance beneath continuous cloud depends on cloud type and solar elevation $\beta(=\pi/2 - \psi)$.

The fraction of extraterrestrial radiation can be read from the full lines and the corresponding irradiance by interpolation between the dashed lines.

The formation of a small amount of cloud in an otherwise clear sky always increases the diffuse flux but the direct component remains unchanged provided neither the sun's disk nor its aureole is obscured. With a few isolated cumuli, the total irradiance can therefore exceed the total irradiance beneath a cloudless sky by 5–10%. On a day of broken cloud (Figure 5.11), the temporal distribution of radiation is strongly bimodal:

the irradiance is very weak when the sun is completely occluded and strong when it is exposed. For a few minutes before and after occlusion, the irradiance commonly reaches 1000Wm−2 in temperate latitudes and even exceeds the Solar Constant in the tropics. This effect is a consequence of strong forward scattering by the small concentration of water droplets present at the edge of a cloud.

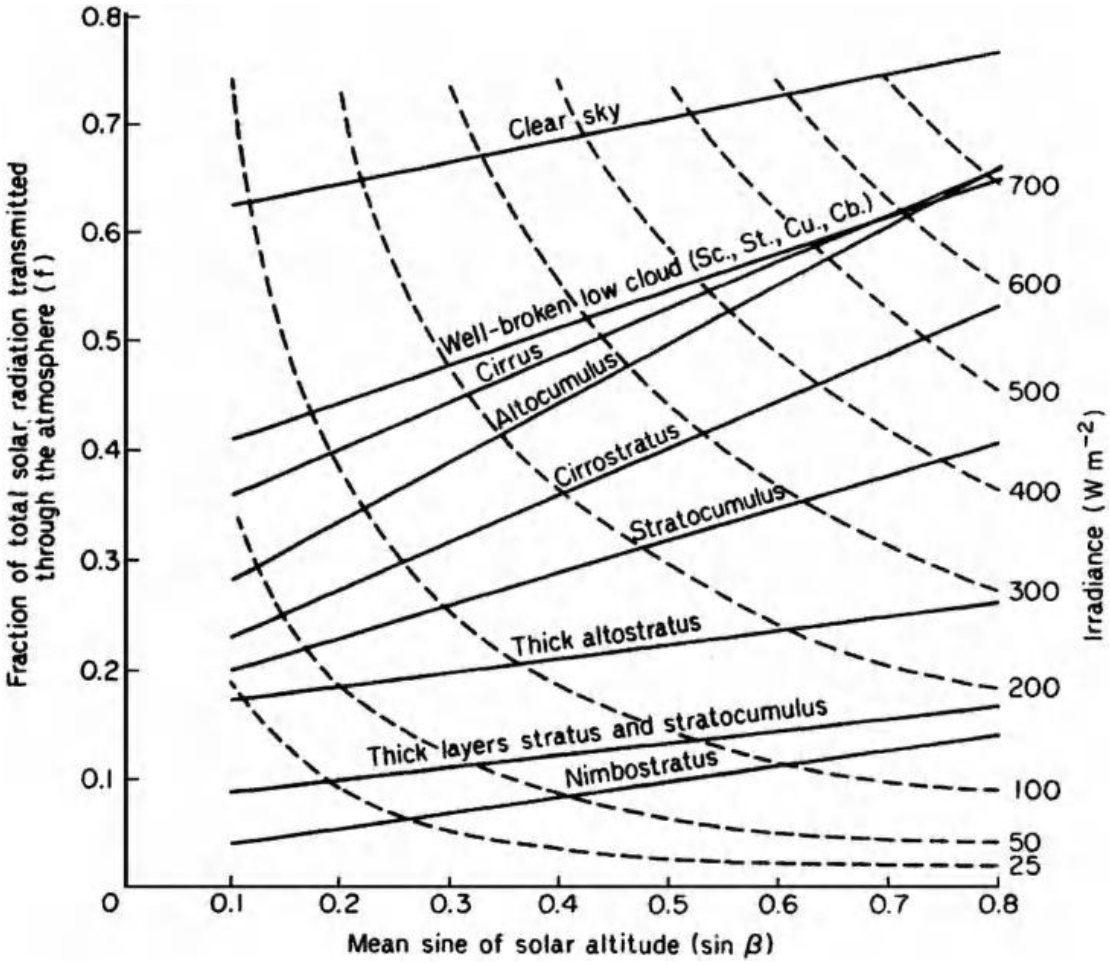


Figure 10 Empirical relations between solar radiation and solar angle for different cloud types from measurements in the North Atlantic (52◦N, 20◦W). The curves are isopleths of irradiance (from Lumb, 1964). Sc, stratocumulus; St, stratus; Cu, cumulus; Cb, cumulonimbus.

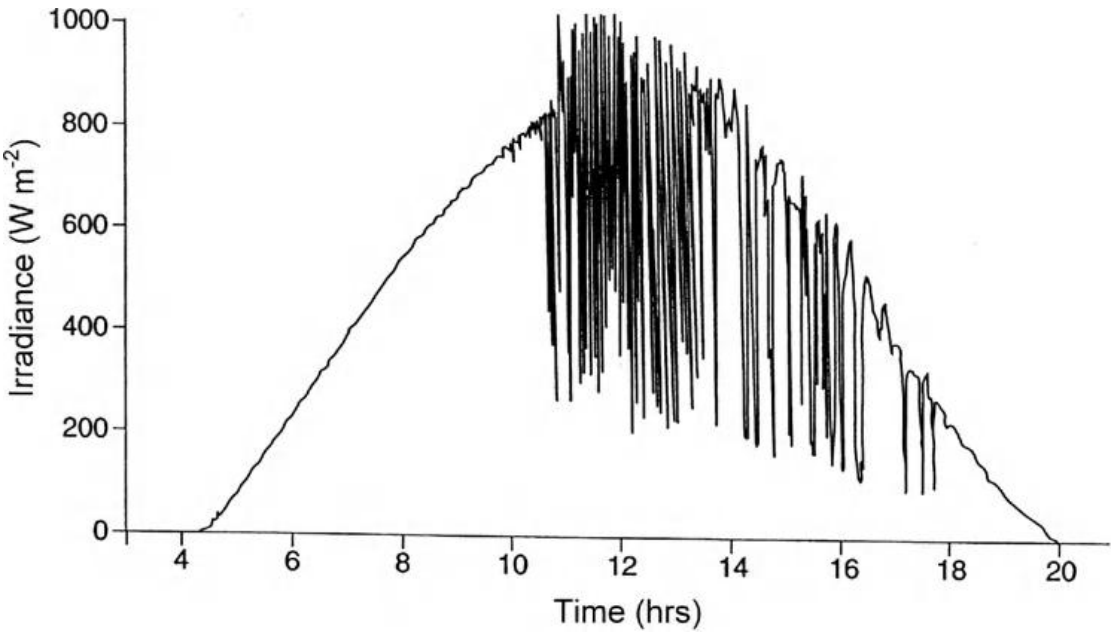


Figure 11 Solar radiation on a day of broken cloud (11 June 1969) at Rothamsted, England (52◦N, 0◦W). Note very high values of irradiance immediately before and after occlusion of the sun by cloud and the regular succession of minimum values when the sun is completely obscured.

As a consequence of cloud, the average insolation over most of Europe in summer is restricted to between 15 and 25 MJ m−2, about 50–80% of the insolation on cloudless days. Comparable figures in the USA range from 23 MJ m−2 round the Great Lakes to 31 MJ m−2 under the almost cloudless skies of the Sacramento and San Joaquin valleys

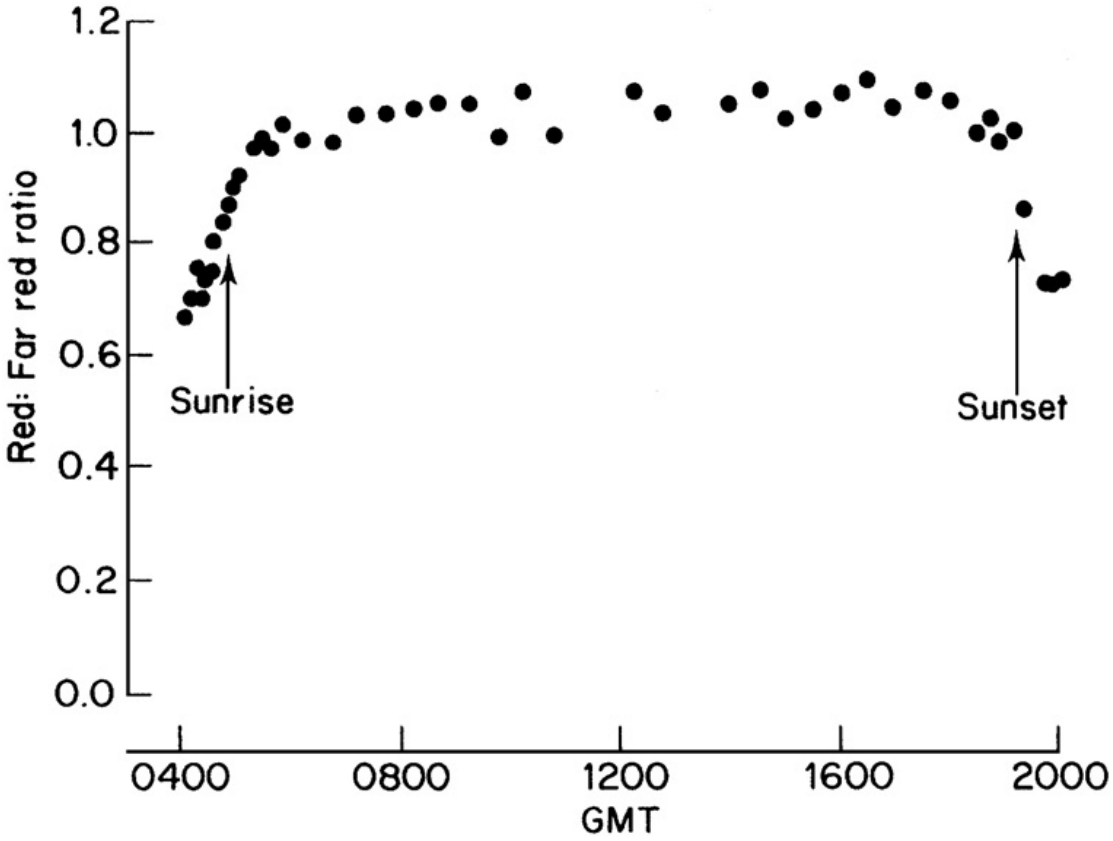


Figure 12 Ratio of spectral irradiance at 660 nm (red) to irradiance at 730 nm (far-red) on an overcast day (25 August 1980) near Leicester in the English Midlands (from Smith and Morgan, 1981).

Winter values range from 1 to 5MJ m−2 over most of Europe and from 6 MJ m−2 in the northern USA to 12 MJ m−2 in the south. Australian stations record a range of values similar to those of the USA.

The energy efficiency of photovoltaic solar panels is defined as the ratio of the electrical power produced by the panel to the radiative power falling on the panel. Commercial flat panels typically achieve a maximum efficiency of about 20%. That is, if the incident irradiance on a panel of area 1 m2 is 1 kW the maximum power output would be 200W.

Although the difference of irradiance with and without cloud is roughly an order of magnitude, the radiation to which plants are exposed covers a much wider range. In units of micromoles of photons m−2 in the PAR waveband, full summer sunshine is approximately 2200, shade on a forest floor 20, twilight 1, moonlight 3×10−4, and radiation from an overcast sky on a moonless night about 10−7 (Smith and Morgan, 1981).

At any location, annual changes of insolation depend in a complex way on seasonal changes in the water vapor and aerosol content of the atmosphere and on the seasonal distribution of cloud. Table 5.2 shows the main components of attenuation for four "seasons" at Kew Observatory, a suburban site 10 miles (16 km) west of the center of London. The data were collected in the 1950s when London air was more heavily polluted with smoke particles from inefficiently burned coal than it is now, so the "dust and smoke" losses for winter are relatively large. For the annual average, roughly a third of the radiation received outside the atmosphere is scattered back to space, a third is absorbed, and a third is transmitted to the surface. The flux at the surface is 20–25% less than it would be in a perfectly clear atmosphere. Because the climate at Kew is relatively cloudy, the diffuse component is larger than the direct component throughout the year.

Table 5.2 Short-Wave Radiation Balance of Atmosphere and Surface at Kew Observatory (51.5°N) for 1956–1960 Expressed as a Percentage of Extraterrestrial Flux

	Winter (Nov–Jan)	**Spring (Feb–Apr)**	**Summer (May–Jul)**	**Autumn (Aug–Oct)**	**Year**
Extraterrestrial radiation (ET)					
Seasonal total (MJ m^{-2})	800	2050	3720	2340	8910
Daily mean (MJ m^{-2} day^{-1})	8.7	22.3	40.4	25.4	24.4
Losses in the atmosphere (%ET)					
(a) Absorption					
Water vapor	15	12	13	15	13
Cloud	8	9	9	9	9
Dust and smoke	15	10	5	8	8
Total	38%	31%	27%	32%	30%
(b) Scattering (away from surface)	37%	35%	33%	34%	34%
Radiation at surface (%ET)					
Direct	8	14	18	14	15
Diffuse	17	20	22	20	21
Total	25%	34%	40%	34%	36%
	100%	100%	100%	100%	100%
Total as MJ m^{-2} day^{-1}	2.2	7.6	16.2	8.7	8.8

Spectrum of Total Solar Radiation

The spectrum of global solar radiation depends, in principle, on solar zenith angle, cloudiness, and turbidity and the interaction of these three factors limits the usefulness of generalizations. As zenith angle increases beyond 60◦, so does the proportion of scattered radiation and therefore the ratio of visible to all-wavelength radiation. In one record from Cambridge, England, this ratio increased from about 0.49 at ψ = 60◦ to 0.52 at ψ = 10◦ (Szeicz, 1974). Cloud droplets absorb radiation in the infra-red spectrum, so with increasing cloud the fraction of visible radiation should increase.

Again at Cambridge, the range was between 0.48 in summer and 0.50 in winter. Finally, with increasing turbidity, shorter wavelengths are scattered preferentially, depleting the direct beam but contributing to the diffuse flux so that the change in global radiation is relatively small. In another set of measurements at a site close to Cambridge, the visible: all-wavelength ratio decreased from 0.53 at τa = 0.1 to 0.48 at τa = 0.6 (McCartney, 1978). Elsewhere, a smaller value of the ratio, 0.44, with little seasonal variation was reported for a Californian site by Howell et al. (1983) and, in the tropics, Stigter and Musabilha (1982) found that the ratio increased from 0.51 with clear skies to 0.63 with overcast. It is probable that some of the apparent differences between sites reflect differences or errors in instrumentation and technique rather than in the behavior of the atmosphere.

At a given site, the relation between quantum content and energy appears to be conservative. McCartney (1978) reported a value of 4.56 ± 0.05 μmol J−1 (PAR) in the English Midlands. This is equivalent to about 2.3 μmol J−1 total (global) radiation but other values reported range from 2.1 μmol J−1 in California to 2.9 μmol J−1 in Texas (Howell et al., 1983). The ratio of quantum content to energy in the PAR wave band for diffuse radiation from a cloudless sky is smaller than that for total radiation, about 4.25 μmol J−1 PAR, because quanta at shorter wavelengths carry more energy than those at longer wavelengths.

The ratio of energy per unit wavelength in red and far-red wavebands is also conservative and has a value of about 1.1 during the day when solar elevation exceeds 10◦ (Figure 11). Thus the ratio of the red to far-red forms of phytochrome in plants is constant through the day unless shading by other vegetation occurs (see p. 53). As the sun approaches the horizon, the ratio decreases because red light is scattered more than far-red and because only a small fraction of the forward-scattered light reaches the ground. There is some evidence that the ratio starts to increase and returns to

values greater than 1 when the solar disk falls below the horizon, presumably because skylight alone is relatively rich in shorter wavelengths (Figure 5).

Terrestrial Radiation

At the wavelengths associated with solar radiation, emission from radiatively active gases in the earth's atmosphere, and from the earth's surface is negligible, so the previous section considered only absorption and scattering. In contrast, at wavelengths of terrestrial radiation (i.e. long-wave radiation originating in the earth's atmosphere and at its surface), both absorption and emission are important and will be considered in this section.

Most natural surfaces can be treated as "full" radiators which emit long-wave radiation, in contrast to the short-wave solar radiation emitted by the sun. At a surface temperature of 288 K, the energy per unit wavelength of terrestrial radiation (based on Wien's Law, p. 39) reaches a maximum at 2897/288 or 10 μm, and arbitrary limits of 3 and 100 μm are usually set for the long-wave spectrum. Figure 5.13 illustrates the spectrum of radiation that would be emitted to the atmosphere from a surface that was a full radiator at 288 K.

In the absence of cloud, most of the radiation emitted by the earth's surface is absorbed within the atmosphere in specific wavebands by radiatively active atmospheric gases, mainly water vapor and carbon dioxide. A small fraction of radiation from the surface escapes to space, mostly through the *atmospheric window* between 8 and 12 μm. The energy absorbed in atmospheric gases is re-radiated (emitted) in all directions. Atmospheric gases do not emit like full radiators: rather, they have an emission spectrum similar to their absorption spectrum (Kirchhoff's principle, p. 38). Figure 13 shows the approximate spectral distribution of the downward flux of atmospheric radiation that would be received at the earth's surface from a cloudless atmosphere at 263 K. In reality, much of the atmospheric radiation that reaches the surface arises from gases close to the surface, and consequently close to surface temperature: atmospheric

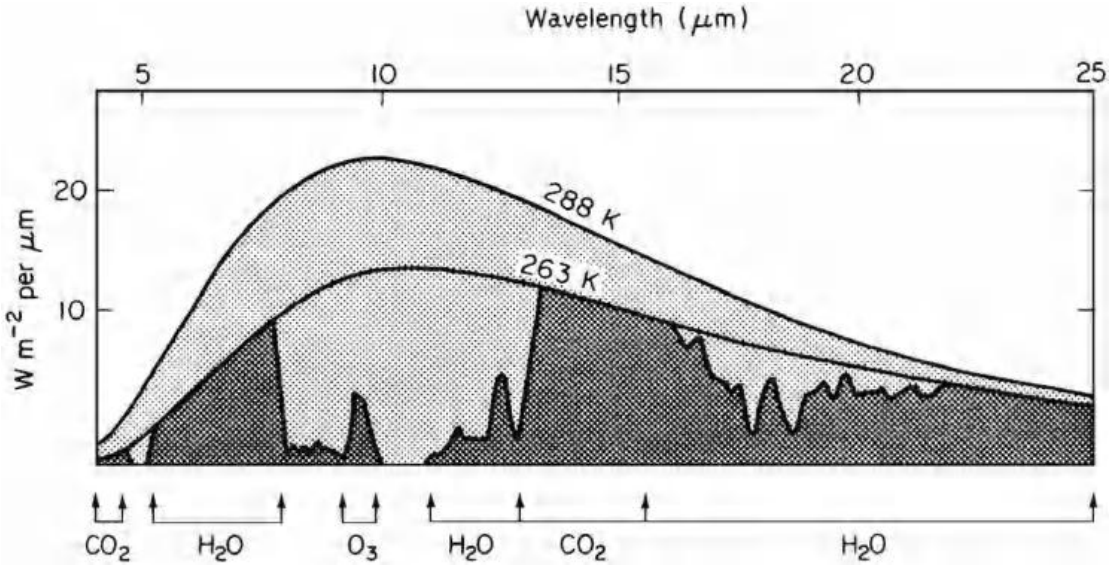


Figure 5.13 Spectral distribution of long-wave radiation for black bodies at 298 K and 263 K. Dark gray areas show the emission from atmospheric gases at 263 K. The light gray area therefore shows the net loss of radiation from a surface at 288 K to a cloudless atmosphere at a uniform temperature of 263 K (after Gates, 1980).

radiation that is lost to space is emitted mainly from gases higher in the troposphere where temperatures are less. Radiation emitted to space is therefore partly surface emission, escaping through the atmospheric window, and partly atmospheric emission from the upper troposphere and stratosphere.

To satisfy the First Law of Thermodynamics for the earth as a planet, assuming that the earth is in equilibrium, the average annual loss of long-wave radiative energy to space must balance the average net gain from solar radiation. If rE is the radius of the earth, $\mathbf{S}*$ is the Solar Constant, ρE is the planetary albedo (the fraction of solar radiation scattered to space from clouds and the surface), and **L** is the emitted radiative flux density (emittance) to space, this balance may be expressed as

$$(1 - \rho\mathrm{E})\mathbf{S}* \pi r2\mathrm{E} = 4\pi r2\mathrm{E}\ \mathbf{L} \quad (5.21)$$

or

$$\mathbf{L} = (1 - \rho\mathrm{E})\mathbf{S}* /4.$$

Taking ρE = 0.30 and $\mathbf{S}*$ = 1361Wm−2 yields **L** = 238Wm−2, and using the Stefan-Boltzmann Law(Eq. 4.2) this corresponds to an equivalent black-body temperature of the earth viewed from space of 254 K (−19 ◦C). The low value in comparison to the mean temperature at the surface (about 288 K, 15 ◦C) is an indication of the extent to

which atmospheric gases and cloud create a favorable climate for life on earth. This phenomenon is commonly called the *greenhouse effect*, though the term is a poor one, as real greenhouses become warm by reducing heat loss by the wind and convection rather than primarily by radiative effects. In fact, the assumption of radiative equilibrium for the earth is incorrect: the earth is currently experiencing additional *radiative forcing* (i.e. absorption of radiation) of about 1.6Wm−2 relative to pre-industrial times as a consequence of human activities, primarily emissions of greenhouse gases (IPCC, 2007), and this is causing *global warming.* If human perturbations of the atmosphere were to cease, eventually the earth would come to equilibrium at a new higher temperature satisfying Eq. (5.21).

Analysis of the exchange and transfer of long-wave radiation throughout the atmosphere is one of the main problems of physical meteorology but micrometeorologists are concerned primarily with the simpler problem of measuring or estimating fluxes at the surface. The upward radiative flux **L**u from a surface can be measured with a radiometer or from a knowledge of the surface temperature and emissivity. The downward flux from the atmosphere **L**d can also be measured radiometrically, calculated from a knowledge of the temperature and water vapor distribution in the atmosphere, or estimated from empirical formulae.

1 Terrestrial Radiation from Cloudless Skies

The radiance of a cloudless sky in the long-wave spectrum (or the effective radiative temperature) is least at the zenith and greatest near the horizon. This variation is a direct consequence of the increase in the path length of water vapor and carbon dioxide, the main emitting gases. In general, more than half the radiant flux received at the ground from a cloudless atmosphere comes from gases in the lowest 100 m and roughly 90% from the lowest kilometer. Themagnitude of the flux received at the surface is therefore strongly determined by temperature gradients near the ground.

It is convenient to define the apparent emissivity of the atmosphere εa as the flux density of downward radiation divided by full radiation at air temperature Ta measured near the ground, i.e.

$$\mathbf{L}\mathrm{d} = \varepsilon \mathrm{a} \sigma T \mathrm{a}^4 . \quad (5.22)$$

Similarly, the apparent emissivity at a zenith angle ψ or $\varepsilon a(\psi)$ can be taken as the flux density of downward radiation at ψ divided by σT_a^4. Many measurements show that the dependence of $\varepsilon a(\psi)$ on ψ over short periods can be expressed as

$$\varepsilon a(\psi) = a + b \ln (u \sec \psi), \quad (5.23)$$

where u is precipitable water (corrected for the pressure dependence of radiative emission) and a and b are empirical constants that change with the vertical gradient of temperature and with the distribution of aerosol (Unsworth and Monteith, 1975). Integration of this equation over a hemisphere using Eq. (4.9) gives the effective (hemispherical emissivity as

$$\varepsilon a = a + b (\ln u + 0.5). \quad (5.24)$$

Comparing Eqs. (5.23) and (5.24) shows that the hemispherical emissivity is identical to the emissivity at a representative angle $\psi\_$ such that

$$\ln (u \sec \psi\_) = \ln u + \ln \sec \psi\_ = \ln u + 0.5.$$

It follows that $\ln \sec \psi\_ = 0.5$, giving $\psi\_ = 52.5^\circ$ irrespective of the values of a and b. Hence a directional radiometer recording the radiance at 52.5◦ could be used tomonitor the value of εa for cloudless skies, and Ld would be given by Eq. (5.22).

Formulae for estimating εa for cloudless skies were reviewed by Prata (1996) and Niemela et al. (2001). The most successful equation was

$$\varepsilon a = 1 - (1 + aw) \exp -[(b + cw)0.5],$$

where$w$is the precipitablewater content of the atmosphere in kgm−2, and the empirical constants are a = 0.10 kg−1 m−2, b = 1.2, c = 0.30 kg−1 m2. Values of w were given approximately by

$$w = 4.65 ea/Ta,$$

where e_a is water vapor pressure near the ground (Pa) and T_a is air temperature (K). An even simpler formula for estimating **L**d was developed by Unsworth and Monteith (1975), viz.

$$\mathbf{L}d = c + d\sigma T_a^4. \quad (5.25)$$

For measurements in the English Midlands, which covered a temperature range from −6 to 26 ◦C, the empirical constants were c = −119±16Wm−2 and d = 1.06±0.04. The uncertainty of a single estimate of **L**d was ±30Wm−2. Measurements in Australia (Swinbank, 1963) gave similar values of c and d but with much less scatter. The lack of an explicit dependence of **L**d on humidity in expressions such as Eq. (5.25) is probably because there is often a strong correlation between air temperature and humidity in the lower atmospheric layers responsible for most of the radiant emission.

Using a linear approximation to the dependence of full radiation on temperature above a base of 283 K allows Eq. (5.25) to be written in the form

Ld = 213 + 5.5T _ a , (5.26)

where T _ a is air temperature in ◦C. Outward long-wave radiation, assumed to be σT _ a 4 (i.e. $\varepsilon = 1$), is given by a similar approximation as

Lu = 320 + 5.2T _ a (5.27)

The net loss of long-wave radiation is therefore

Lu − **L**d = 107 − 0.3T _ a (5.28)

implying that 100Wm−2 is a good average figure for the net loss to a clear sky (see Figure 14).

If the cloudless atmosphere emitted like a full radiator (an incorrect but commonly applied simplification in climatology texts), an expression for the *effective radiative temperature* T _ b of the atmosphere could be obtained by writing

Ld = 320 + 5.2T _ b = 213 + 5.5T _ a ,

so that

T _ b = _ 5.5T _ a – 107 _ /5.2 = _ T _ a – 21 _ + 0.06T _ a

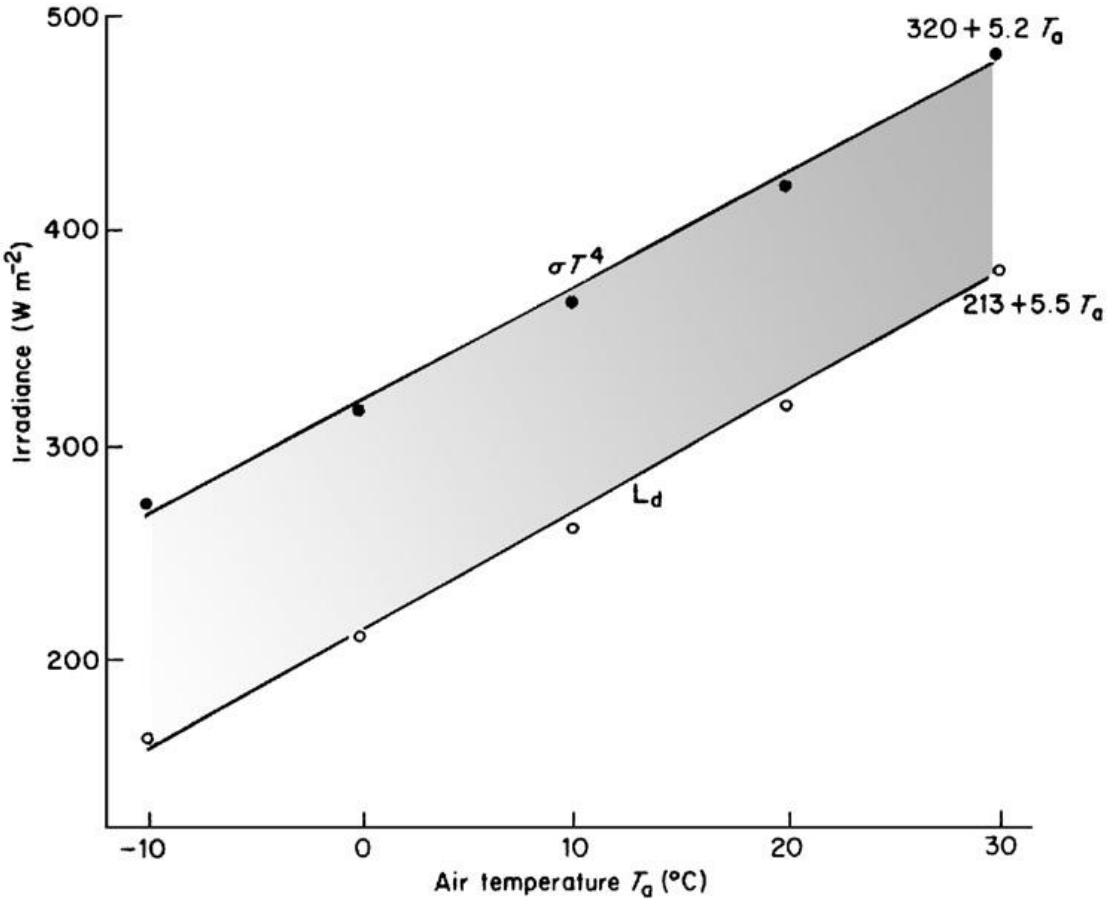


Figure 14 Black-body radiation at *T*a (full circles) and long-wave radiation from clear sky (open circles) from Eq. (5.25). Straight lines are approximations from Eqs. (5.26) and (5.27), respectively.

This relation shows that when T _ a is between 0 and 20◦C, the mean effective radiative temperature of a cloudless sky is usually about 21–19 ◦Cbelowthemean air temperature near the ground.

In general, for climatological work, the more complex formulae for estimating **L**d that require humidity or precipitable water data have little additional merit over the simpler expressions above since the main uncertainty lies in the influence of cloud—the next topic.

2 Terrestrial Radiation from Cloudy Skies

Clouds dense enough to cast a shadow on the ground emit long-wave radiation like full radiators at the cloud base temperature of the water droplets or ice crystals from which they are formed. The presence of cloud increases the flux of atmospheric radiation received at the surface because the radiation from water vapor and carbon dioxide in the lower atmosphere is supplemented by emission from clouds in the waveband which the gaseous emission lacks, particularly from 8 to 13 μm (see Figure 12). Because most of the gaseous component of atmospheric radiation reaching the surface originates below the base of clouds, the gaseous component of the downward flux can be treated as if the sky was cloudless with an apparent emissivity εa. From Kirchhoff's principle, the transmissivity of radiation from the cloud through the air layer beneath cloud base is 1−εa, and if the cloud base temperature is Tc the downward radiation received at the surface from a fully overcast sky will be

$$\mathbf{L}\mathrm{d} = \varepsilon\mathrm{a}\sigma T\,4\,\mathrm{a} + (1 - \varepsilon\mathrm{a})\sigma T\,4\,\mathrm{c}\,. \quad (5.29)$$

Using the linear approximation Eq. (4.5)with $\delta T = T\mathrm{a} - T\mathrm{c}$, and ignoring second-order terms in δT, Tc can be eliminated by writing

$$\sigma T\,4\,\mathrm{c} = \sigma(T\mathrm{a} - \delta T)4 \approx \sigma T\,4\,\mathrm{a} - 4\sigma T\,3\,\mathrm{a}\,\delta T = \sigma T\,4\,\mathrm{a}\,(1 - 4\delta T/T\mathrm{a}).$$

Hence Eq. (5.29) may be written

$$\mathbf{L}\mathrm{d} = \sigma T\,4\,\mathrm{a}\,\{1 - 4(1 - \varepsilon\mathrm{a})\delta T/T\mathrm{a}\} \quad (5.30)$$

and the emissivity is

$$\varepsilon\mathrm{a} = \mathbf{L}\mathrm{d}/\sigma T\,4\,\mathrm{a} = 1 - 4(1 - \varepsilon\mathrm{a})\delta T/T\mathrm{a}.$$

The analysis of a series of measurements near Oxford, England (Unsworth and Monteith, 1975) gave an annual mean of δT = 11 K with a seasonal variation of ±2 K, figures consistent with a mean cloud base at about 1 km, higher in summer than in winter. Taking 283 K as a mean value of Ta for this location gives $4\delta T / T\mathrm{a} = 0.16$, so that the emissivity of a completely overcast sky (cloud fraction c = 1) at this site would be

$$\varepsilon\mathrm{a}(1) = \mathbf{L}\mathrm{d}/\sigma T\,4\,\mathrm{a} = 1 - 0.16\,\{1 - \varepsilon\mathrm{a}\} = 0.84 + 0.16\varepsilon\mathrm{a}. \quad (5.31)$$

For a sky covered with a fraction c of cloud, interpolation gives

$$\varepsilon\mathrm{a}(c) = c\varepsilon\mathrm{a}(1) + (1 - c)\varepsilon\mathrm{a} = (1 - 0.84c)\varepsilon\mathrm{a} + 0.84c. \quad (5.32)$$

The main limitation to this formula lies in the choice of appropriate values for cloud temperature and for δT which depend on base height and therefore on cloud type. It is important to remember that the formulae presented in this section are statistical correlations of radiative fluxes with weather variables at particular sites and do not describe direct functional relationships. For prediction, they are most accurate when the air temperature does not increase or decrease rapidly with height near the surface and when the air is not exceptionally dry or humid. They are therefore appropriate for climatological studies of radiation balance but are often not accurate enough for micrometeorological analyses over periods of a few hours. In particular, the simple equations cannot be used to investigate the diurnal variation of **L**d. At most sites, the amplitude of **L**d in cloudless weather is much smaller than the amplitude of **L**u, behavior to be expected because changes of atmospheric temperatures are governed by, and follow, changes of surface temperature.

Net Radiation

All surfaces receive short-wave radiation during daylight and exchange long-wave radiation continuously with the atmosphere. The net rate at which a surface receives radiation of long and short wavelengths is called the *net radiation balance*, or *net radiation*. (The term *balance* is used in the same sense as in a bank balance, which may be positive, negative, or zero depending on gains and losses.) Thus the net radiation received by unit area of a horizontal surface with reflection coefficient ρ is defined by the equation

$\mathbf{R}n = (1 - \rho)\mathbf{S}t + \mathbf{L}d - \mathbf{L}u.$ (5.33)

The concept of the net radiation balance can also be applied at the global scale: maps and animations of the seasonal variation of the earth's net radiation balance may be found at

http://earthobservatory.nasa.gov/GlobalMaps/view.php?d1=CERES_NETFLUX_M.

*Micro*climatological applications of Eq. (5.33) are considered. Here, a discussion of the net receipt of radiation by a standard horizontal surface is needed to round off the chapter, although net radiation is not strictly a *macro*climatological quantity: it depends on the temperature, emissivity, and reflectivity of a surface. Net radiation is measured routinely at only a few climatological stations partly because of the problem of providing a standard surface but also because instruments in the past were difficult to maintain. More robust instruments consisting of four independent sensors for the components of **R**n are now available from several manufacturers (e.g. from Kipp and Zonen B.V., Delft, The Netherlands) which make the task of observing and interpreting Rn at climatological sites more feasible. Examples of annual and daily measurements follow.

Figure 15 shows the annual change of components in the net radiation balance of a short grass surface at Hamburg, Germany (54◦N, 10◦E) from February 1954 to January 1955. Each entry in the graph represents the gain or loss of radiation for a period of 24 h. The largest term in the balance is **L**u, the long-wave emission from the grass surface, ranging between winter and summer from about 23 to about 37 MJ m−2 day−1. The equivalent mean radiative flux density, found by dividing by the number of seconds in a day (86,400), is 270–430Wm−2. The minimum values of downward atmospheric radiation **L**d (corresponding to about 230Wm−2) were recorded in spring, presumably in cloudless anticyclonic conditions bringing very cold dry airmasses; and maximum values *(* ≈ 380 W m−2*)* were recorded duringwarm humid weather in the autumn. The net loss of long-wave radiation was about 60Wm−2 on average (cf. 100Wm−2 for cloudless skies on p. 72) andwas almost zero on a few, very foggy days in autumn and winter.

In the lower half of the diagram, the income of short-wave radiation forms a Manhattan skyline with much larger day-to-day changes and a much larger seasonal amplitude than the income of long-wave radiation. The maximum value of **S**t is about 28 MJ m−2 day−1 (320Wm−2) and **S**t is smaller than Ld on every day of the year .The reflected radiation is about 0.25**S**t except on a few days in January and February

when snow increased the reflection coefficient to between 0.6 and 0.8. The net radiation **R**n given by $(1 - \rho)$**S**t + **L**d − **L**u is shown in the top half of the graph. During summer at this northerly site (54◦N), the ratio of daily **R**n/**S**t was almost constant from day to day at about 0.57, but the ratio decreased during the autumn and

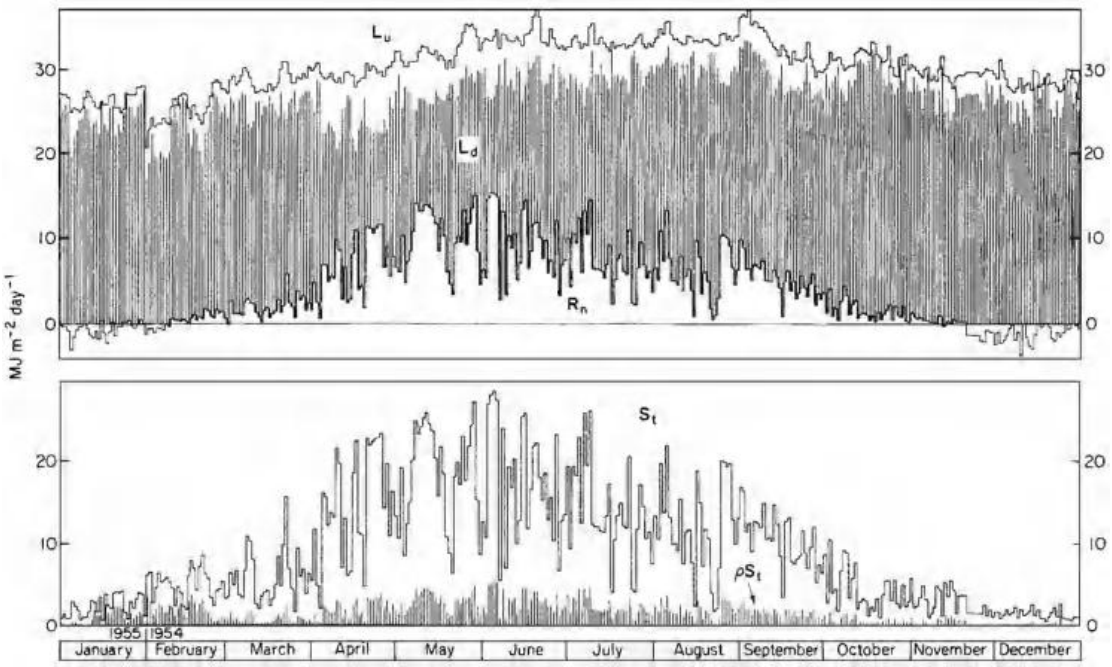


Figure 15 Annual radiation balance at Hamburg, Germany, 1954/55: **S**t, total solar radiation; ρ**S**t, reflected solar radiation; **L**u, upward long-wave radiation; **L**d, downward long-wave radiation; **R**n, net radiation (after Fleischer, 1955).

reached zero in November. From November until the beginning of February daily **R**n was negative on most days. In summer, net radiation and mean air temperature were positively correlated with sunshine. In winter, with the sun low in the sky and the days short, the correlation was negative: sunny cloudless days were days of minimum net radiation with the mean air temperature below average.

Staff at the University of Bergen maintained records of incoming short- and long wave radiation for many years and overcame the problem of maintaining a standard surface by using black-body radiation at air temperature for **L**u and by adopting a reflection coefficient of 0.2 (grass) or 0.7 (snow) as appropriate (The device of referring net radiation to a surface at air temperature is convenient in *micro*climatology too (see p. 229) and deserves to be more widely adopted.) The

shaded part of Figure 16 represents the difference between the fluxes **L**d and **L**u i.e. the net loss of long-wave radiation; the bold line is **R**n. In cloudless weather, the diurnal change in the two long-wave components is much smaller than the change of short-wave radiation which follows an almost sinusoidal curve. The curve for net radiation is therefore almost parallel to the **S**t curve during the day, decreases to a minimum value in the early evening, and then increases very slowly for the rest of the night (because the lower atmosphere is cooled by radiative exchange with the earth's surface). In summer, the period during which **R**n is positive is usually about 2 or 3 h shorter than the period during which **S**t is positive. Comparison of the curves for clear spring and winter days (Figure 15a and b) shows that the seasonal change of **R**n/**S**t noted in Figure 14 is a consequence of

(i) the shorter period of daylight in winter and (ii) much smaller maximum values of **S**t, in winter, unmatched by an equivalent decrease in the net long-wave loss. Figures 17 and 18 show the net radiation balances during several cloudless days of short grass (Corvallis, Oregon), and of an old-growth Douglas fir/Western Hemlock forest(Wind River, Washington State) in the Pacific Northwest of the USA. By day, solar radiation is the dominant influence on the radiation balance. The reflection coefficient for solar radiation is about 0.23 for the grass and 0.07 for the forest, a consequence of the optical properties of the foliage and the canopy structure.

The upward long-wave radiation varies less over the forest than over the grass because the tree foliage remains close to air temperature during the day whereas the short grass gets much warmer than the air. At night, the net radiation over both surfaces is most strongly negative shortly after sunset when the canopies are warmer than they are later in the night, and consequently lose more radiation to the cloudless sky. Unusually, net radiation at night is similar in magnitude over both canopies. More typically, if wind speeds had been similar over the grass and the forest, heat stored in the soil would have flowed upwards to keep the grass canopy warmer than the more isolated forest canopy. It is likely that the dry soil in summer did not conduct heat effectively in this example. Figures 17 and 18 demonstrate a good correlation between **R**n and **S**t which could be used to estimate **R**n from total solar radiation records (a method that has commonly been used (e.g. Kaminsky and Dubayah, 1997)). However, the principles

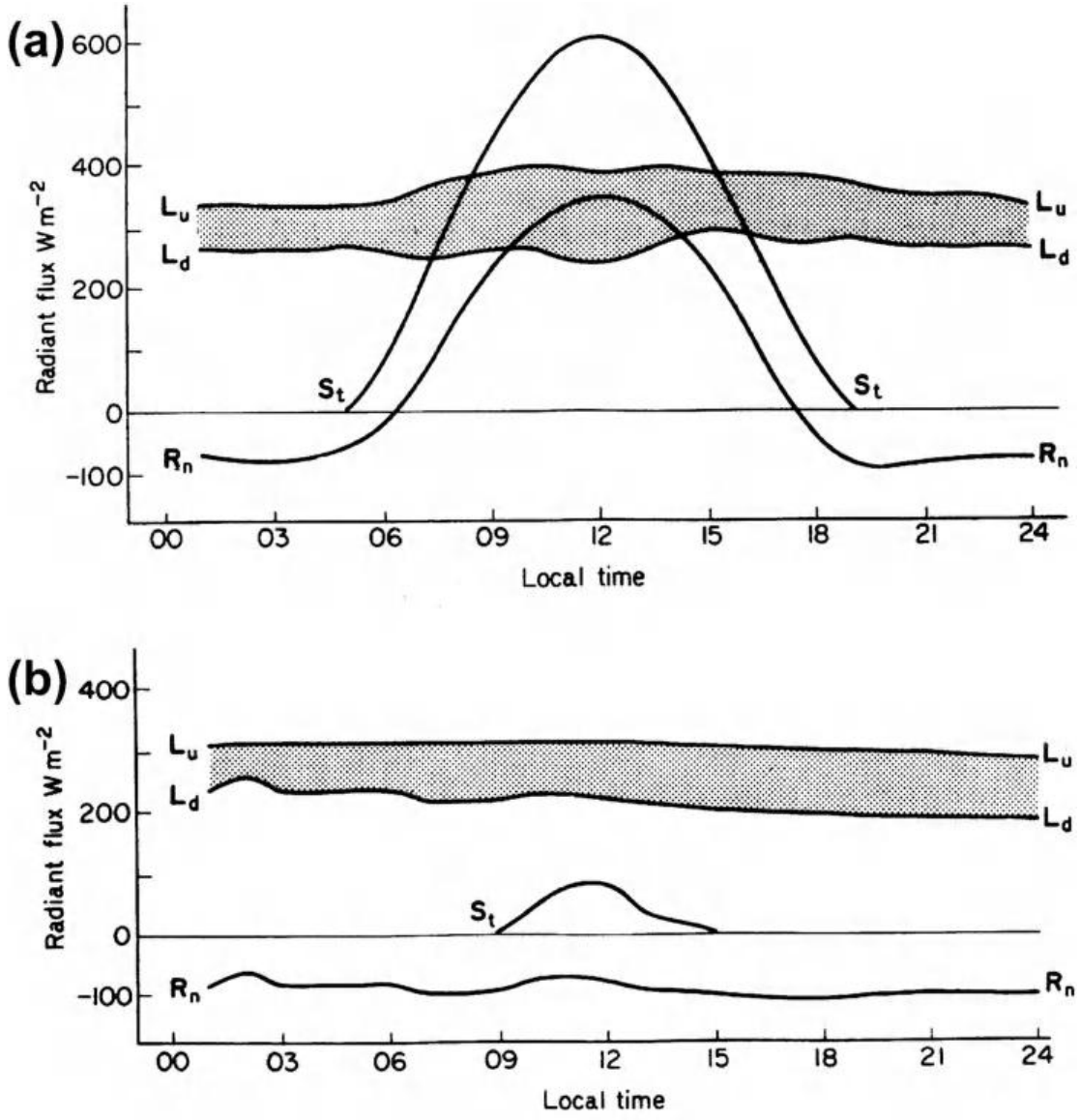


Figure 16 Radiation balance at Bergen, Norway (60◦N, 5◦E): (a) on 13 April 1968, (b) on 11 January 1968. The gray area shows the net long-wave loss and the line **R**n is net radiation.

Note that net radiation was calculated from measured fluxes of incoming short- and long-wave radiation, assuming that the reflectivity of the surface was 0.20 in April (e.g. vegetation) and 0.70 in January (e.g. snow). The radiative temperature of the surface was assumed equal to the measured air temperature.

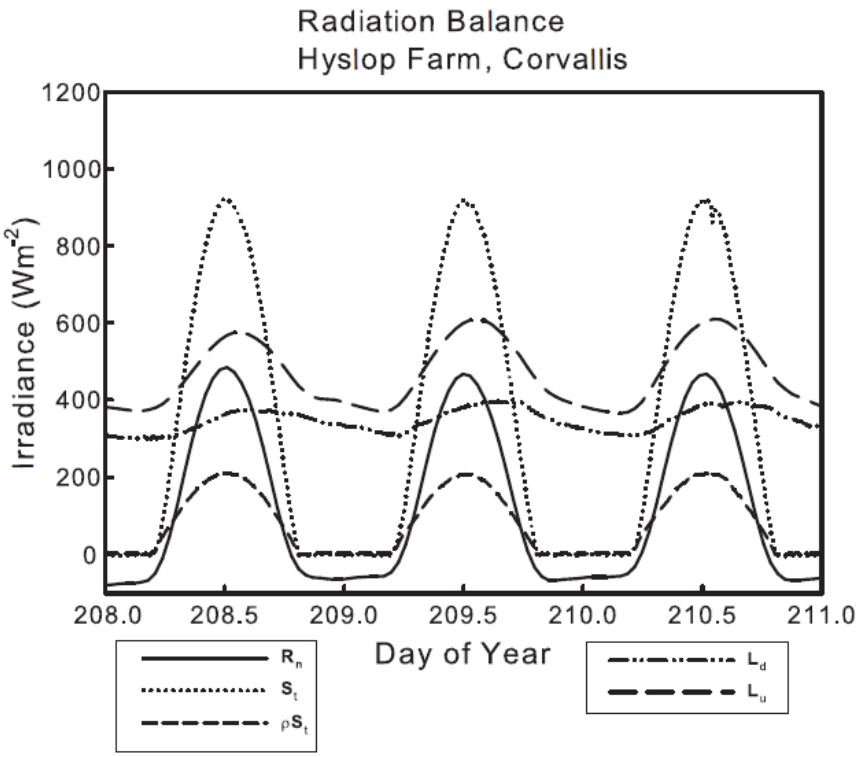


Figure 17 Components of the net radiation balance during three cloudless days over short grass near Corvallis Oregon (45◦N, 123◦W). (data courtesy of Reina Nakamura, Oregon State University.)

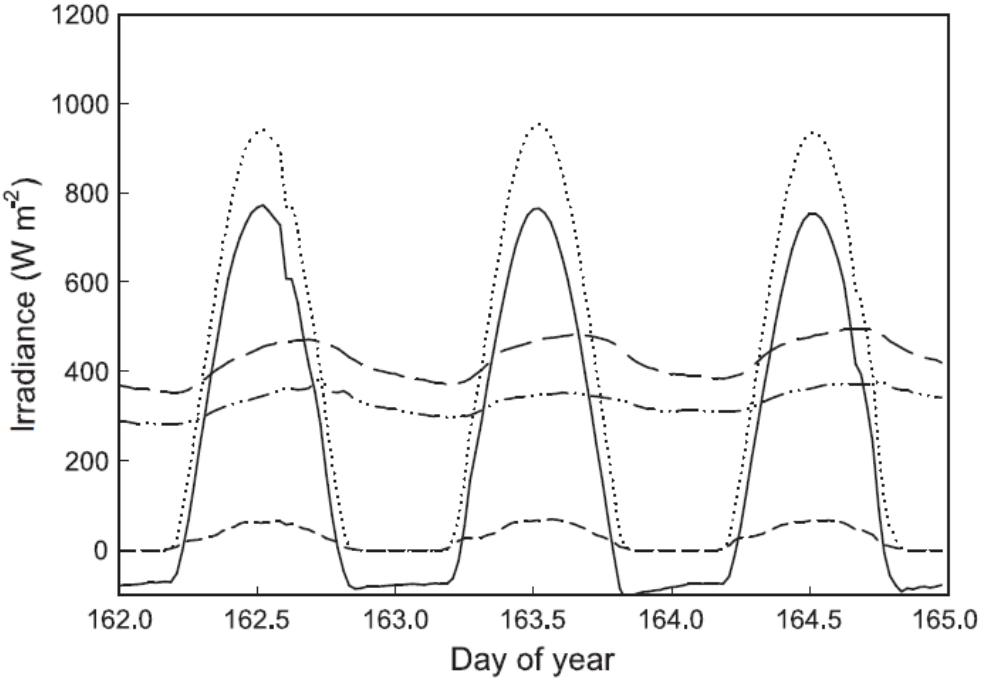


Figure 18 Components of the net radiation balance during three cloudless days over an old growth Douglas fir/Western Hemlock forest at Wind River Experimental

Forest, Washington, (46◦N, 120◦W). Legends and line type are as in Fig. 5.17. (data courtesy of Ken Bible, University of Washington.)

summarized in this chapter show that such an approach is not only site dependent, but also surface dependent, so that methods for estimating **R**n from its components are preferable (Offerle et al., 2003).

In the material just presented we have concentrated primarily on radiation fluxes when skies are cloudless because these conditions provide maximum daily energy to the surface and are most amenable to analysis. In overcast weather, with low cloud, **L**d becomes almost equal to $\sigma T 4$ a , so the net long-wave exchange is close to zero and **R**n is almost zero at night. During the day under overcast skies **R**n ≈ $(1 - \rho)$**S**t.

Problems

1. Estimate the time difference between sunset and the beginning of civil twilight in Greenwich, England (51.5◦N, 0.0◦W) on 21 June (solar declination 23.5◦).

2. Use the simple model for solar radiation under cloudless skies (p. 64) to estimate the direct, diffuse, and total solar irradiance on a horizontal surface at 45◦N at local solar noon on 21 June for three values of aerosol optical thickness, 0.05, 0.20, 0.40. Assume that the optical thickness for molecular attenuation is 0.30.

3. Using the results of Question 2, estimate the daily insolation at 45◦N at solar noon on 21 June for the three aerosol loadings assuming cloudless sky.

4. Assuming a cloudless sky, estimate the downward long-wave irradiance and the atmospheric emissivity at the ground when the air temperature is 25 ◦C. If the sky was covered by 50% cloud, estimate the long-wave irradiance in this case. Why are these values only climatological approximations?

5. Four solar radiometers are exposed side by side. A and B have clear glass domes transmitting all wave lengths in the solar spectrum; C and D have domes that transmit 95% of the radiant energy from 700 to 3000 nm and no radiation below 700 nm. A and C receive global (total) radiation whereas B and D have a shade ring that intercepts 10% of radiation from the sky as well as all the direct solar beam. On a cloudless summer day the instruments give the following outputs: A 11.00 mV; B 1.30 mV; C 5.30 mV; and D 0.25 mV. Assuming that all instruments have the same sensitivity, 12.0 μVW−1 m−2, calculate

(i) the ratio of diffuse to global radiation,

(ii) the fraction of photosynthetically active radiation (0.4–0.7 μm) in the diffuse component,

(iii) the fraction of visible radiation in the direct solar beam, and

(iv) the global irradiance in the photosynthetically active waveband.

6. A farmer plans to plant seeds in a field of soil with a solar radiation reflection coefficient $\rho s = 0.20$. She decides that the soil will be warmer (and germination more rapid) if she spreads a thin layer of black soot over it, decreasing the solar radiation reflection coefficient to $\rho\_s = 0.05$. To assess the effect of the change, she prepares adjacent plots and records the following data around noon on a cloudless day: Incident total solar irradiance **S**t = 900Wm−2 (same for both plots). Difference in net radiation **R**n(soot-covered)− **R**n(bare soil) = 69Wm−2. Radiative temperature of bare soil *T*s = 303 K.

Estimate the radiative temperature of the soot-covered surface, stating any assumptions that you need to make.

Steady-State Heat Balance

(i) Water Surfaces, Soil, and Vegetation The heat budgets of plants and animals will now be examined in the light of the principles and processes considered in previous chapters. The First Law of Thermodynamics states that when a balance sheet is drawn up for the flow of heat in any physical or biological system, income and expenditure must be exactly equal. In environmental physics, radiation and metabolism are the main sources of heat income; radiation, convection, and evaporation are methods of expenditure.

For any component of a system, physical or biological, a balance between the income and expenditure of heat is achieved by adjustments of temperature. If, for example, the income of radiant heat received by a leaf began to decrease because the sun was obscured by cloud, leaf temperature would fall, reducing expenditure on convection and evaporation. If the leaf had no mass and therefore no heat capacity, the decrease in expenditure would exactly balance the decrease in income, second by second. For a real leaf with a finite heat capacity, the decrease in temperature would lag behind the decrease in radiation and so would the decrease in expenditure on convection and evaporation, but the First Law of Thermodynamics would still be satisfied because the

reduced income from radiant heat would be supplemented by the heat given up by the leaf as it cooled. This chapter is concerned with the heat balance of relatively simple systems in which

(a) temperature is constant so that changes in heat storage are zero and

(b) metabolic heat is a negligible term in the heat budget. The heat budget of warm blooded animals, controlled by metabolism.

1 Heat Balance Equation

The heat balance of any organism can be expressed by an equation with the form

$$\mathbf{R}n + \mathbf{M} = \mathbf{C} + \lambda\mathbf{E} + \mathbf{G}. \quad (13.1)$$

The individual terms are

Rn = net gain of heat from radiation,

M = net gain of heat from metabolism,

C = loss of sensible heat by convection,

$\lambda\mathbf{E}$ = loss of latent heat by evaporation,

G = loss of heat by conduction to environment.

The over bars in Eq. (13.1) indicate that each term is an average heat flux per unit surface area. (In the rest of this chapter, they are implied but not printed.) In this context, it is convenient to define surface area as the area from which heat is lost by convection although this is not necessarily identical to the area from which heat is gained or lost by radiation. The conduction term **G** is included for completeness but is negligible for plants and has rarely been measured for animals. An equation similar to (13.1) applies to bare soil surfaces or water bodies but without the term **M**.

The grouping of terms in the heat balance equation is dictated by the arbitrary sign convention that fluxes directed away from a surface are positive. (When temperature decreases with distance z from a surface so that $\partial T/\partial z < 0$, the outward flux of heat $\mathbf{C} \propto -\partial T/\partial z$ is a positive quantity, see p. 29.) The sensible and latent heat fluxes **C** and $\lambda\mathbf{E}$ are therefore taken as *positive* when they represent losses of heat from the surface to the atmosphere, and as *negative* when they represent gains. On the left-hand side of the equation, **R** and **M** are positive when they represent gains and negative when they represent losses of heat. When both sides of a heat balance equation are positive, the equation is a statement of how the total supply of heat available from sources is

divided between individual sinks. When both sides are negative, the equation shows how the total demand for heat from sinks is divided between available sources.

The sections which follow deal with the size and manipulation of individual terms in the heat balance equation (13.1), with some fundamental physical implications of the equation, and with several examples of biological applications.

1.1 Convection and Long-Wave Radiation

When the surface of an organism loses heat by convection, the rate of loss per unit area is determined by the scale of the system as well as by its geometry, by wind speed, and by temperature gradients. Convection is usually accompanied by an exchange of long-wave radiation between the organism and its environment at a rate which depends on geometry and on differences of radiative temperature but is independent of scale. The significance of scale can be demonstrated by comparing convective and radiative losses from an object such as a cylinder with diameter d and uniform surface temperature $T0$ exposed in a wind tunnel whose internal walls are kept at the temperature T of the air flowing through the tunnel with velocity u.When Re exceeds 103, the resistance to heat transfer by convection increases with d according to the relation $rH = d/(\kappa Nu) \propto d0.4V-0.6$

In contrast, the corresponding resistance to heat transfer by radiation rR (p. 41) is independent of d. Figure 13.1 compares rH and rR for cylinders of different diameters at wind speeds of 1 and 10 m s−1 chosen to represent outdoor conditions. Corresponding rates of heat loss are shown in the right-hand axis for a surface temperature excess $(T0 - T)$ of 1 K.

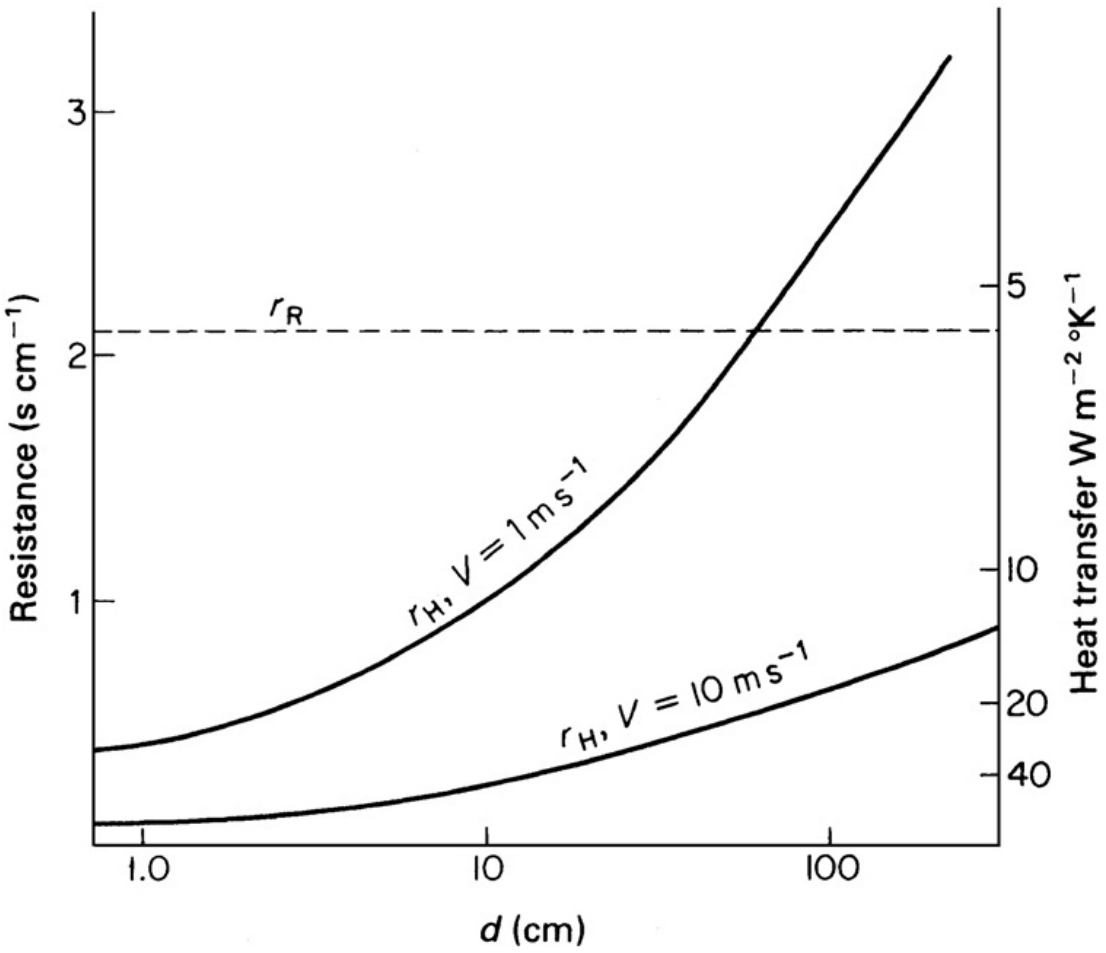


Figure 1 Dependence of resistance to convective heat transfer *r*H and resistance to radiative heat exchange *r*R as a function of body size represented by the characteristic dimension of a cylinder *d*.

The cylinder is assumed exposed in a wind tunnel where air and wall radiative temperatures are identical. The curves are calculated for wind speeds, *V*, of 1 m s−1 and 10 m s−1. The right-hand axis shows the rate of heat transfer assuming a difference of 1 K between cylinder surface temperature and air/wall radiative temperature. Because the dependence of *r*H on scale and windspeed is similar for planes, cylinders, and spheres provided that the appropriate dimension is used to calculate Nusselt numbers, a number of generalizations may be based on Figure 13.1. For organisms on the scale of a small insect or leaf ($0.1 < d < 1$ cm), *r*H is much smaller than *r*R, implying that convection is a much more effective mechanism of heat transfer than long-wave radiation. The organism is tightly coupled to air temperature but not to the radiative temperature of the environment. For organisms on

the scale of a farm animal or a man *(*10 < *d* < 100 cm*)* *r*H and *r*R are of comparable importance at low wind speeds.

For very large mammals *(d* > 100 cm*),* *r*H can exceed *r*R at low wind speeds, and in this case the surface temperature will be coupled more closely to the radiative temperature of the environment than to the air temperature. These predictions are consistent with measurements on locusts and on piglets, for example, and they emphasize the importance of wall temperature as distinct from air temperature in determining the thermal balance of large farm animals in buildings with little ventilation. They also explain why warming the air in a cold room does not induce a feeling of comfort for people occupying the room if the radiative temperature of the walls remains low. When an organism with an emissivity of unity and a surface temperature of *T*0 exchanges heat (a) by convection to air at temperature *T* and (b) by radiation to an environment with a mean radiative temperature equal to air temperature, the net rate at which heat is gained or lost is

$$\rho c_p \left\{(T - T_0)/r_H + \left((T - T_0)/r_R\right)\right\} = \rho c_p \left\{(T - T_0)/r_{HR}\right\}, \qquad (13.2)$$

where *r*HR = _ *r*−1 H + *r*−1 R _−1 is a combined resistance for convection (p. 31) and long-wave radiation formed by grouping the component resistances in parallel because the fluxes are in parallel.

Heat Balance of Thermometers

1 Dry-Bulb

As an introduction to the more relevant physics of the wet-bulb thermometer, it is worth considering the implications of the general heat balance equation for a dry-bulb thermometer (i.e. a thermometer with a dry sensing element). For measurements in the open, it is essential to avoid heating a thermometer by direct exposure to sunlight, so screening is employed. In several common designs used for precise measurements of air temperature, a thermometer with a cylindrical bulb is housed in a tube through which air is drawn rapidly, and for the sake of the following discussion we assume that the tube completely surrounds the bulb. If the tube itself is exposed to sunshine, its temperature, *T*s, may be somewhat above the temperature of the air (*T*) and of the thermometer (*T*t).

The net long-wave radiation received by the thermometer from the housing (assuming emissivity $\varepsilon = 1$) is

$$\mathbf{R}_\mathrm{n} = \sigma \left(T_\mathrm{s}^4 - T_\mathrm{t}^4 \right) = \rho c_\mathrm{p} \left(T_\mathrm{s} - T_\mathrm{t} \right) / r_\mathrm{R}$$

using the definition of rR on p. 41 with the assumption that Ts − Tt is small. The loss of heat by convection from the bulb to the air is

$$\mathbf{C} = \rho c_\mathrm{p} \left(T_\mathrm{t} - T \right) / r_\mathrm{H}.$$

In equilibrium, heat balance equation (13.1) reduces to **R**n = **C** and rearrangement of terms in Eq. (13.4) gives

$$T_\mathrm{t} = \frac{r_\mathrm{H} T_\mathrm{s} + r_\mathrm{R} T}{r_\mathrm{R} + r_\mathrm{H}}$$

indicating that the temperature recorded by the thermometer (the apparent temperature) is a weighted mean between the true temperature of the air and the temperature of the thermometer housing. There are two main ways in which the difference between apparent and true air temperature can be minimized:

i. by making rH much smaller than rR (i.e. decoupling the sensor from the radiative environment), either by adequate ventilation or by choosing a thermometer with very small diameter (see Figure 13.1). The numerator in Eq. (13.5) then tends to rRT and the denominator to rR.

ii. by making Ts very close to T , e.g. by using a reflective metal or white painted screen, by introducing insulation between outer and inner surfaces, or by increasing ventilation on both sides of the screen.

In the standard instrument known as an *Assmann psychrometer*, the screen is a double-walled cylinder, nickel-plated on the outer surface, and aspirated at about 3 m s−1. Since the diameter of the mercury-in-glass thermometer it contains is about 3 mm, Figure 1 illustrates that it is effectively decoupled from its radiative environment

(rH _ rR).

2 Wet-Bulb

The concept of a "wet-bulb temperature" is central to the environmental physics of systems in which latent heat is a major heat balance component, and it has two distinct connotations: the *thermodynamic wet-bulb temperature*, which is a theoretical abstraction and the temperature of a thermometer with its sensing element covered with a wet sleeve, which, at best, is a close approximation to the thermodynamic wet-bulb temperature.

A value for the thermodynamic wet-bulb temperature can be derived by considering the behavior of a sample of air enclosed with a quantity of pure water in a container with perfectly insulating walls. This is an *adiabatic system* within which the sum of sensible and latent heat must remain constant. The initial state of the air can be specified by its temperature T , vapor pressure e, and total pressure p. Provided e is smaller than $e_s(T)$, the saturated vapor pressure at T , water will evaporate and both e and p will increase. Because the system is adiabatic, the increase of latent heat represented by the increase in water vapor concentration must be balanced by a decrease in the amount of sensible heat which is realized by cooling the air. The process of humidifying and cooling continues until the cooling air becomes saturated at a temperature T_w which, by definition, is the thermodynamic wet-bulb temperature i.e. the thermodynamic wet bulb temperature is the minimum temperature which may be achieved by bringing an air parcel to saturation by evaporation in adiabatic conditions. The corresponding saturated vapor pressure is $e_s(T_w)$.

To relate T_w and $e_s(T_w)$ to the initial state of the air, the initial water vapor concentration is approximately $\rho\varepsilon e/p$ when p is much larger than e (see p. 15). When the vapor pressure rises from e to $e_s(T_w)$, the total change in latent heat content per unit volume is $\lambda\rho\varepsilon_w e_s(T_w) - e/p$. The corresponding amount of heat supplied by cooling unit volume of air from T to T_w is $\rho c_p(T - T_w)$. (A small change in the heat content of water vapor is included in more rigorous treatments but is usually unimportant in micrometeorological problems.) Equating latent and sensible heat

$$\lambda\rho\varepsilon\left[e_{\mathrm{s}}(T')-e\right]/p=\rho c_{\mathrm{p}}\left(T-T'\right)$$

and rearranging terms gives

$$e=e_s(T')-(c_{\mathrm{p}}p/\lambda\varepsilon)(T-T').$$

The group of terms *(*cp *p/λε)* is often called the "*psychrometer constant*" for reasons explained shortly, but it is neither constant (because *p* is atmospheric pressure and *λ* changes somewhat with temperature) nor exact (because of the approximations made).

The psychrometer constant is often assigned the symbol *γ* and, at a standard pressure of 101.3 kPa, has a value of about 66 Pa K−1 at 0 ◦C increasing to 67 Pa K−1 at 20 ◦C. Thus

$$e=e_s(T')-\gamma(T-T').$$

Another useful quantity which has the same dimensions as *γ* is the change of saturation vapor pressure with temperature or *∂*es*(T)/∂T* , usually given the symbol (or *s*) . This quantity can be used to obtain a simple (but approximate) relation between the saturation vapor pressure deficit *D* = *e*s*(T)* − *e* and the wet-bulb depression *B* = *T* − *T* _. When the saturation vapor pressure at wet-bulb temperature *T* _ is written as

$$e_{\mathrm{s}}(T')\approx e_{\mathrm{s}}(T)-\Delta\left(T-T'\right),$$

where is evaluated at a mean temperature of *(T* +*T* _ *)*/2, the psychrometer equation (13.7) becomes

$$e \approx e_s(T) - (\Delta + \gamma)(T - T')$$

or

$$D \approx (\Delta + \gamma) B.$$

Equation (13.7) can be represented graphically by plotting *es(T)* against *T* (Figure 13.2). The curve QYP represents the relation between saturation vapor pressure and temperature and the point X represents the state of any sample of air in terms of *e* and *T* . Suppose the wet-bulb temperature of the air is *T* _ and that the point Y represents the state of air saturated at this temperature. The equation of the straight line XY joining the points *(T, e), (T* _ *, es(T* _ *))* is

$$e - e_s(T') = \text{slope} \times (T - T').$$

Comparison of Eqs. (13.7) and (13.11) shows that the slope of XY is $-\gamma$. The wet-bulb temperature of any sample of air can therefore be obtained graphically by drawing a line with slope $-\gamma$ through the appropriate coordinates *T* and *e* to intercept the saturation curve at a point whose abscissa is *T* _.

If a sample of air in the state given by X was cooled toward the state represented by the point Y, the path XY shows how temperature and vapor pressure would change in adiabatic evaporation, i.e. with the total heat content of the system constant. Similarly, starting from Y and moving to X, the path YX shows how *T* and *e* would change

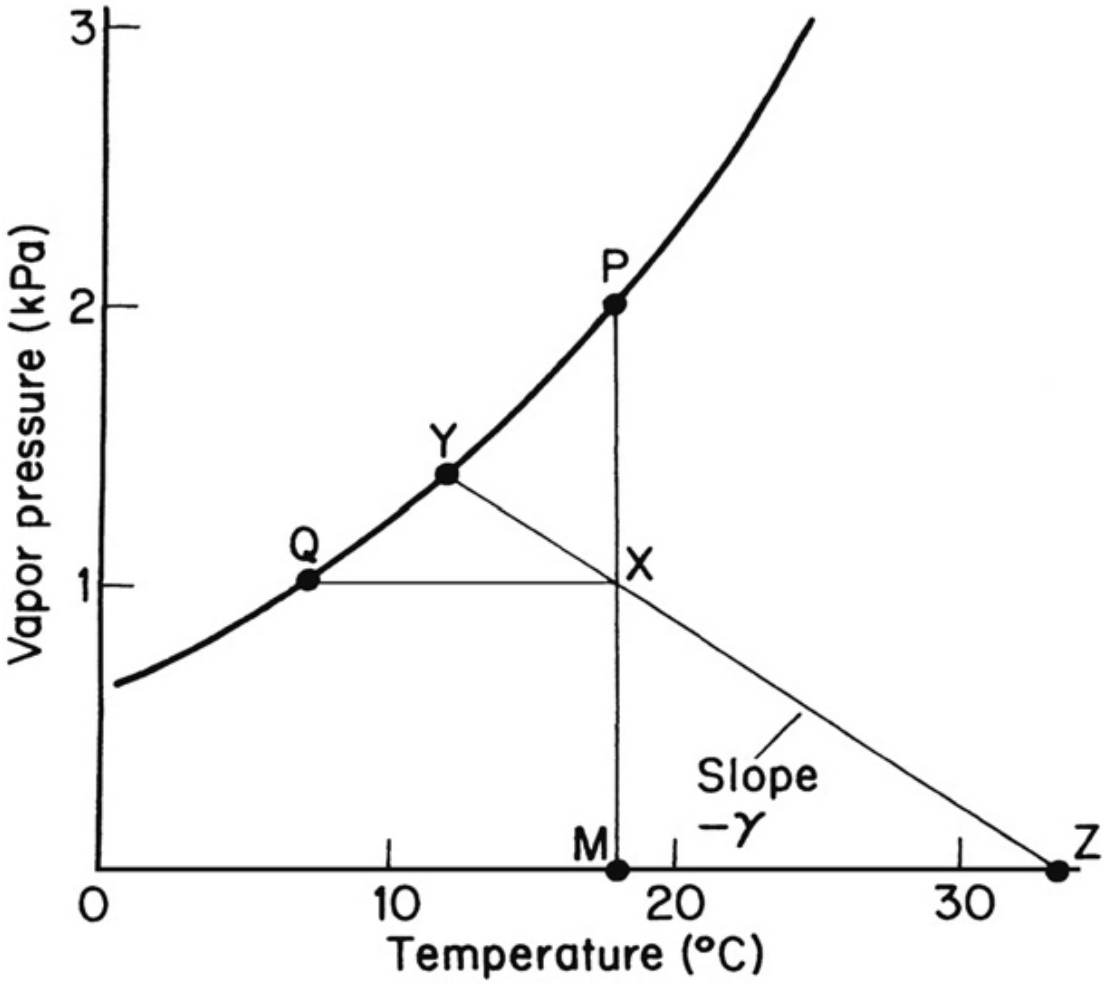


Figure 2 The relation between dry-bulb temperature, wet-bulb temperature, equivalent temperature, vapor pressure, and dew point. The point X represents air at 18 ◦C and 1 kPa vapor pressure. The line YXZ with a slope of $-\gamma$ gives the wet-bulb temperature from Y (12 ◦C) and the equivalent temperature from Z (33.3 ◦C). The line QX gives the dew-point temperature from Q (7.1 ◦C). The line XP gives the saturation vapor pressure from P (2.1 kPa).

if water vapor were condensed adiabatically from air that was initially saturated. As condensation proceeded, the temperature of the air would rise until all the vapor had condensed. This state is represented by the point Z at which $e = 0$. The corresponding temperature Te is called the "*equivalent temperature*" of the air. As Z has coordinates (Te, 0), the equation of the line ZX can be written in the form

$$T\text{e} = T + e/\gamma. \quad (13.12)$$

Alternatively, the equation of YZ can be written

$Te = T\_ + es\_ T\_\_/\gamma$, (13.13)

showing that the equivalent and wet-bulb temperatures are uniquely related. Both $T\_$ and Te remain unchanged when water is evaporated or condensed adiabatically within a sample of air.

Moving from theoretical principles to a real wet-bulb thermometer, it is necessary to account for the finite rate at which heat is lost by evaporation and gained by convection and radiation.

Suppose that a thermometer bulb covered with a wet sleeve has a temperature Tw when it is exposed to air at temperature T and surrounded by a screen that is also at air temperature. The rate at which heat is gained by convection and radiation is

$\mathbf{C} + \mathbf{R}n = \rho cp\_ T - Tw\_ /rHR$. ((cf. Eq. (13.2)). The rate at which latent heat is lost may be found by applying the principles of equation (11.3) to give

$\boldsymbol{\lambda}\mathbf{E} = \lambda \{\chi s(Tw) - \chi\} /rv$. (13.15)

where χ is absolute humidity. Using the relations in Eq. (2.28), Eq. (13.15) may be written

$\boldsymbol{\lambda}\mathbf{E} = \_ \lambda\rho\varepsilon/p\_\_ es\_ Tw\_ - e\_ /rv = \rho cp\_ es\_ Tw\_ - e\_ /\gamma\, rv$. (13.16)

In equilibrium, $\boldsymbol{\lambda}\mathbf{E} = \mathbf{R}n + \mathbf{C}$ from which

$e = es\_ Tw\_ - \gamma\_ rv/rHR\_\_ T - Tw\_$. (13.17)

It is often convenient to regard ($\gamma\, rv/rHR$) (or simply ($\gamma\, rv/rH$) in some circumstances explained below) as a *modified psychrometer constant*, written $\gamma *$. That is

$e = es\_ Tw\_ - \gamma *\_ T - Tw\_$. (13.17a)

Comparing Eqs. (13.7a) and (13.17), it is clear that the measured wet-bulb temperature will not be identical to the thermodynamic wet-bulb temperature unless $rv = rHR$. Because $rv = (\kappa/D)0.67rH$, which may be written $rv = 0.93rH$ (p. 181), this condition implies that

$0.93rH = \_ r{-}1\ H + r{-}1\ R\ \_{-}1$ (13.18)

from which $rH = 0.075rR$. At 20 ◦C, $rR = 2.1$ s cm−1, so the thermodynamic and measured wet-bulb temperatures would be identical when $rH = 0.17$ s cm−1. A wet-bulb thermometer would therefore record a temperature above or below the thermodynamic wet-bulb temperature depending on whether rH was greater or less than this value.

Because both rv and rH are functions of windspeed and rR is not, $\gamma *$ decreases with increasing wind speed and, when rv is much less than rR, tends to a constant value independent of wind speed, viz.

$\gamma * = \gamma \_ r\text{v}/r\text{H} \_ = 0.93\gamma.$ (13.19)

In the Assmann psychrometer, regarded as a standard for measuring vapor pressure in the field, the resistances corresponding to the instrument specification already given are rv = 0.149 s cm−1, rH = 0.156 s cm−1 giving $\gamma *$ = 63 Pa K−1. A much more detailed discussion of psychrometry leading to a standard value of 62 Pa K−1 for the Assman psychrometer was given by Wylie (1979).With this and similar instruments, the error involved in using γ instead of $\gamma *$ is often negligible in micrometeorological work.

A further source of psychrometer error not considered here (but treated by Wylie, 1979) is the conduction of heat along the stem of the thermometer which can be minimized by using a long sleeve and/or a thermometer with very small diameter.

Problems

1. A cylindrical thermometer element, diameter 3 mm is enclosed in a radiation shield that excludes all solar radiation but is 5.0 ◦C warmer than the true air temperature which is 20 ◦C. If the thermometer is to record a temperature within 0.1 ◦C of true air temperature, at what wind speed must the element be ventilated? (Assume long-wave emissivities of 1.0.)

2. Awet-bulb thermometer in a radiation shield has a cylindrical element 4 mm diameter. Estimate the radiative resistance assuming that the temperatures of the wet-bulb and the shield are about 10 ◦C. Hence, plot a graph to show how the difference between measured and thermodynamic wet-bulb temperatures would vary with ventilation speed over the range 0.5–5 m s−1 .

3. A leaf has a boundary layer resistance for heat transfer (both surfaces in parallel) of 40 s m−1, and a combined stomatal and boundary layer resistance of 110 s m−1. Taking air temperature as 22 ◦C and vapor pressure as 1.0 kPa, set up a spread sheet to calculate the rates of sensible heat transfer **C**, latent heat transfer $\boldsymbol{\lambda}$**E**, and the sum **C**+$\boldsymbol{\lambda}$**E** as a function of leaf temperature over the range 20–26 ◦C. Hence find

graphically or otherwise the leaf temperature when the net radiation absorption by the leaf is 300 W m−2.

4. For the data in question 3, use the Penman-Monteith Equation to determine the latent heat flux from the leaf and hence determine the sensible heat flux and leaf temperature.

5. Show, using the Penman-Monteith Equation, that the necessary condition for dewfall is − **R**n > ρcp {*es (T)* − *e*} *r*−1 H , and that the maximum rate of dew-fall is **E**max = **R**n/λ(+ γ ∗). On a cloudless night, the net radiation flux density of an isolated leaf at air temperature *T* (K) is given by the empirical expression **R**n = (0.206*e*0.5−0.47)σ *T* 4, where *e* (kPa) is the vapor pressure. Neglecting other sources of heat, plot a graph to show how the maximum rate of dew-fall on the leaf varies with air temperature on the range 0–25 ◦C, and explain your results.

6. A marathon runner, treated as a cylinder with diameter 33 cm moving at 19 km h−1 relative to the surrounding air, has a net radiation load of 300 W m−2. The air temperature and vapor pressure are 30 ◦C and 2.40 kPa, respectively. Assuming that the runner's skin is covered with sweat that is a saturated salt solution for which the relative humidity of air in contact with the solution is 75%. Estimate the rate of latent heat loss. If all the salt could be washed off when the runner was sprayed with water, what would be the new rate of latent heat loss?

7. On the Oregon coast, Sitka spruce is frequently exposed to wind-driven sea spray. In the extreme case where water on needle surfaces is saturated with sea salt (equilibrium relative humidity 75%), by how much would the needle decoupling coefficient be changed relative to needles with pure water on their surfaces?
Assume a wind speed of 5 m s−1 and needle dimension of 1 mm.

Heat Balance Components

Metabolism (M)

Standard measurements of the *basal metabolic rate* (BMR) are made when an animal has been deprived of food and is resting in an environment wheremetabolic rate is independent of external temperature. It has been widely accepted by animal physiologists that the basal metabolic rate of an animal **M**b (expressed here in watts and not watts per unit surface area as elsewhere) can be related to simply body mass *W* by the empirical *allometric equation*

$\mathbf{M}b = BW^n$, (14.2)

where B is a constant, implying that ln$\mathbf{M}$b is a linear function of ln$W$. Based on statistical analysis of many sets of measurements on animals of a wide range of sizes, animal physiologists concluded in the mid-20th century that n = 0.75 in Eq. (14.2), which became known as the 3⁄4 power law and has been widely used by animal ecologists to model metabolism in relation to body size in ecology. Kleiber (1965) suggested that, for *intra*-specific work, B = 3.4 (with units of W per kg0.75); Hemmingsen (1960) found that B = 1.8 was a better value for inter-specific comparisons over a range of mass from 0.01 to 10 kg. For application to humans, where knowledge of BMR is important in exercise physiology and dietetics, more complicated expressions are often used, taking into account gender, height, weight, and age (Frankenfield et al., 2005); they generally agree with Eq. (14.2) within 10–20%. Hemmingsen's (1960) review also showed that the BMR of a poikilotherm kept at a body temperature of 20 ◦C is about 5% of the value for a homeotherm of the same mass with a deep body temperature of 39 ◦C. Hibernating mammals metabolize energy at about the same rate as poikilotherms of the same mass. Figure 14.1 plots the relations between BMR and body weight proposed by Hemmingsen for homeotherms and poikilotherms, and includes observed data from animals working at metabolic rates up to their maximum.

Referring basal metabolic rates to the 3⁄4 power of body mass is inconsistent with the physics of heat transfer since most of the energy produced in metabolism is lost from

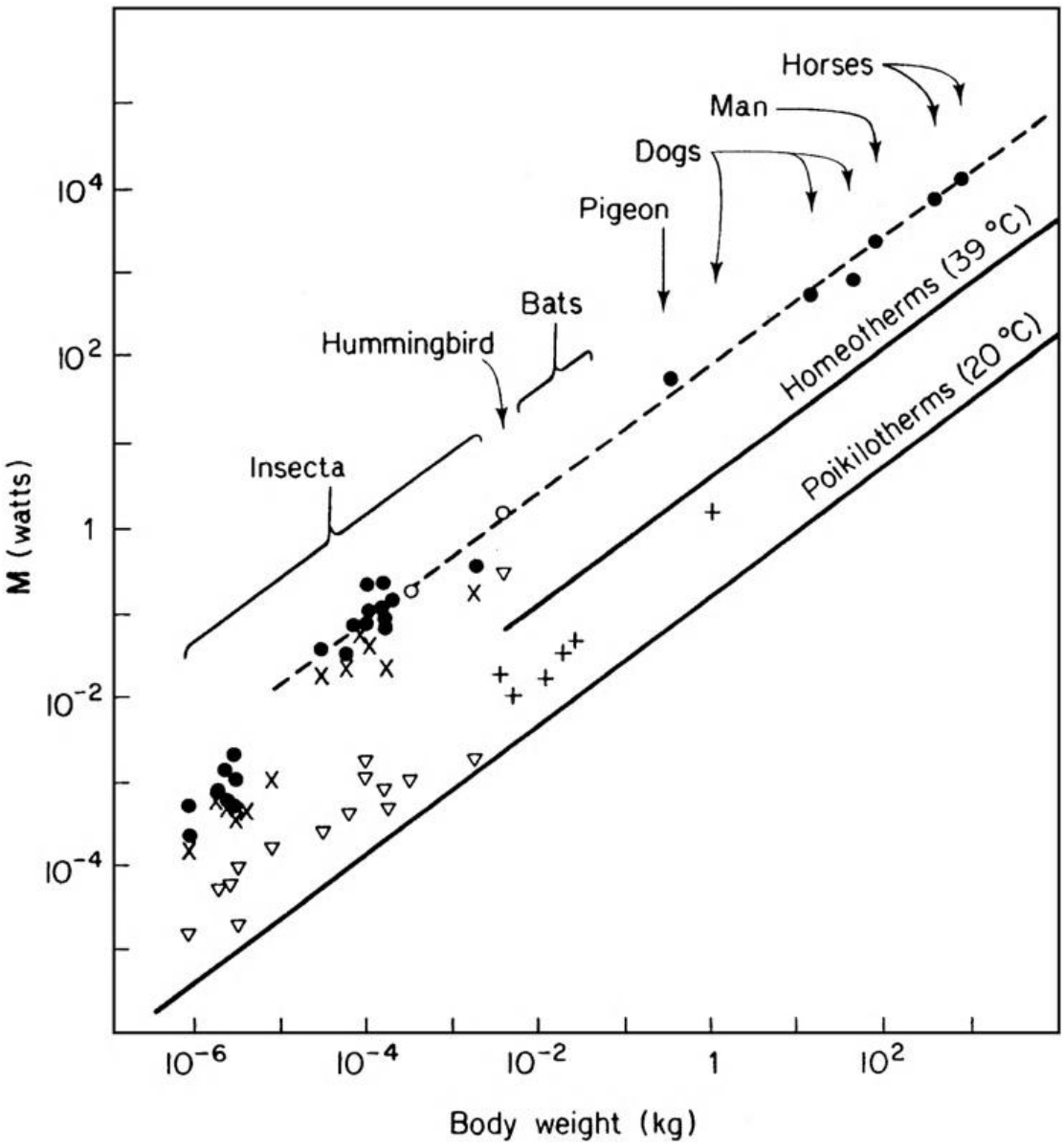


Figure 1 Relation between body weight and basal metabolic rate (BMR) of homeotherms (upper continuous line), maximum metabolic rate for sustained work by homeotherms (upper pecked line), and basal rate for poikilotherms at 20 ◦C (lower continuous line) (from Hemmingsen, 1960).

the surface of the animal. For any set of objects with identical geometry but different in size, surface area is proportional to the 2⁄3 power of mass, assuming constant mass per unit volume. It follows that if basal metabolic rates were proportional to surface

area, they would be proportional to the 2/3 power of mass, at least when animals of like geometry are compared. In a careful re-examination of experimental data to retain only metabolic measurements taken in conditions conducive to the basal rate, Prothero (1984) found that n was closer to 2/3 than to 3/4. More recently, Roberts et al. (2010) further criticized the empirical basis of the 3/4 power law and developed a physically realistic analysis relating BMR to body size that included heat flow from the core to the surface of an animal and heat loss from the surface. Assuming that mammals of sizes ranging from a mouse to a man could be described by an ellipsoidal shape, they used experimental data in their analysis to show that Eq. (14.2) fitted their data with constants $n = 2/3$ and $B = 4.9$ with **M**b in Watts and Win kg. Figure 2 shows their equation and the observed BMR data that they used. The figure also shows the relation proposed by Kleiber (1965) with $n = 3/4$ and $B = 3.4$

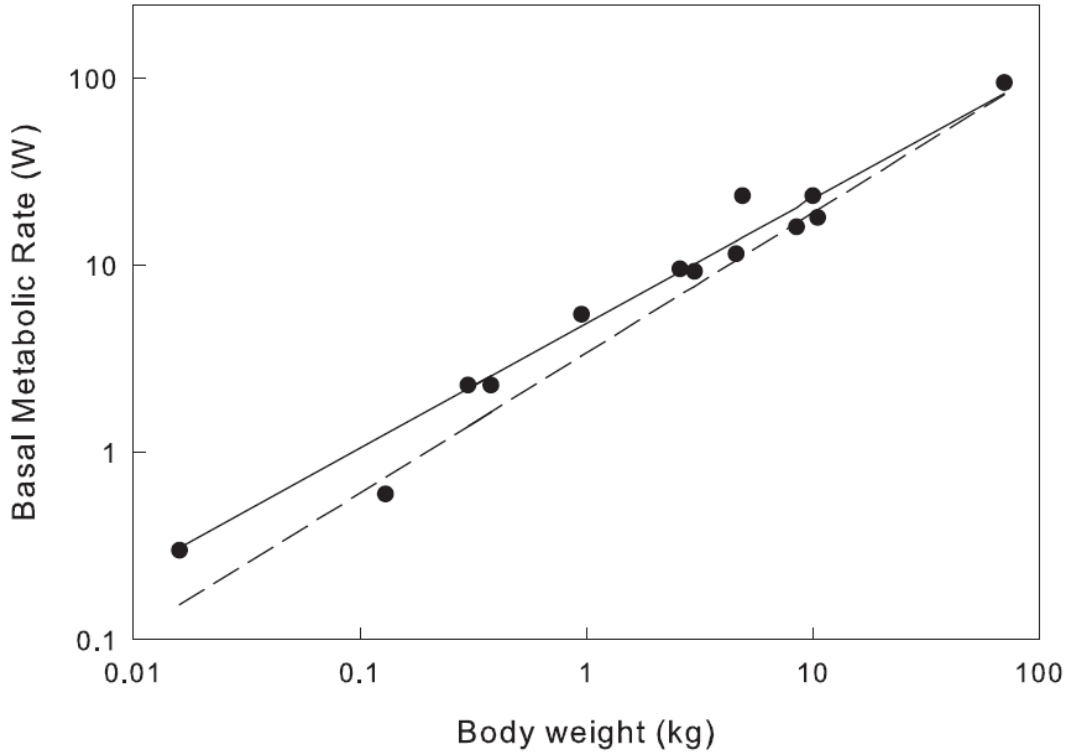


Figure 2 Relation between basal metabolic rate (BMR) of mammals and body weight. Data points represent BMR for animals ranging in size from a mouse to a human. The continuous line is the relation with $\mathbf{M} \propto W2/3$ proposed by Roberts et al. (2010); the dashed line is the relation with $\mathbf{M} \propto W3/4$ as proposed by Kleiber (1965).

Figure 14.2 shows that the proportionality between BMR and metabolic rate holds for mammals ranging in size from a mouse (0.02 kg) to a human (70 kg). The average basal metabolic rate for a mammal per unit of body surface area is about 50 W m−2. This is much smaller than the flux of short-wave radiation absorbed by a dark-coated animal in bright sunshine (say 300 W m−2) but is comparable with the energy that might be absorbed under shade. On a cloudless night, the basal metabolic rate is comparable to the net loss of heat by long-wave radiation from a surface close to air temperature. For walking or climbing, the efficiency with which man and domestic animals use additional metabolic energy is about 30%. To work against gravity at a rate of 20 W m−2, for example, metabolism must increase by about 60 W m−2. The power expended in walking increases with velocity and, in man, reaches 700 W at 2 m s−1 (about 7 km h−1 or 4.3 mph).

For rapid forms of locomotion such as running or flying, the work done against wind resistance is proportional to the wind force times the distance traveled. When the drag coefficient of a moving body is independent of velocity, the drag force should be proportional to $V2$ (p. 139), and as the distance traveled is proportional to V, the rate of energy dissipation should increase with $V3$. However, studies on birds by Tucker (1969) revealed a range of velocities (e.g. 7–12 m s−1 for gulls) in which the rate of energy production was almost independent of V, implying a remarkable decrease of drag coefficient with increasing velocity. To reduce the energy required to overcome drag forces, athletes such as sprinters and swimmers wear skin-tight suits of special fabrics, but the benefits may be more psychological than physiological. However, a tail-wind definitely helps sprinters; the maximum permitted tail-wind of 2 m s−1 could reduce the 100 mrunning time for a champion sprinter by more than 0.1 s by decreasing the energy expended on overcoming form drag (Barrow, 2012).

Latent Heat (λE)

In the absence of sweating or panting, the loss of latent heat from an animal is usually a small fraction of metabolic heat production and takes place both from the lungs during breathing (*respiratory evaporation*) and from the skin as a result of the diffusion of water vapor, sometimes politely referred to as “insensible perspiration.” For a man inhaling completely dry air, the vapor pressure difference between inhaled and exhaled air is about 5.2 kPa and the heat used for respiratory evaporation λE_r is the latent heat equivalent of about 0.8 mg of water vapor per ml of absorbed oxygen

(Burton and Edholm, 1955). With round figures of 2.4 J mg−1 for the latent heat of vaporization of water and 21 J ml−1 oxygen for the heat of oxidation (respiration), λ**E**r/**M** is about 10%.When airwith a vapor pressure of 1.2 kPa is breathed (e.g. air at about50%relative humidity and 20 ◦C), the difference of vapor pressure decreases to 4 kPa and λ**E**r/**M**is about 7%.

In the absence of sweating, the latent heat loss from human skin (λ**E**s) is roughly twice the loss from respiratory evaporation, implying a total evaporative loss of the order of 25–30% of **M** depending on vapor pressure. When sweating, however, a human can produce about 1.5 kg of fluid per hour, equivalent to 600 W m−2 if the environment allows sweat to evaporate as fast as it is produced. More commonly, the rate of evaporation is restricted to a value determined by resistances and vapor pressure differences. Excess sweat then drips off the body or soaks into hair and clothing.

Sheep and pigs lack glands of the type that allow man to sweat profusely but species such as cattle and horses can lose substantial amounts of water by sweating. Dogs and cats have sweat glands only on their foot pads so lose very little heat by this mode. Sheep, dogs, and cats compensate for their inability to sweat by panting in hot environments. For cattle exposed to heat stress, the respiratory system can account for 30% of total evaporative heat loss, the remaining 70% coming from the evaporation of sweat from the skin surface and from wetted hair. The maximum evaporative heat loss from ruminants is only a little larger than metabolic heat production, whereas a sweating man can lose far more heat by evaporation than he produces metabolically if he is exposed to strong sunlight.

Inter-specific differences in λ**E**/**M** and in mechanisms of evaporation may play an important part in adaptation to dry environments. Relatively small values of λ**E**/**M** have been reported for a number of desert rodents which appear to conserve water evaporated from the lungs by condensation in the nasal passages where the temperature is about 25 ◦C (Schmidt-Nielsen, 1965). The respiratory system operates as a form of counter-current heat exchanger. In contrast, measurements of total evaporation from a number of reptiles yield figures ranging from 4 to 9 mg of water per ml of oxygen.

These figures imply that 50−100% of the heat generated by metabolism was dissipated by the evaporation of water, lost mainly through the cuticle. For differing animal species, the amount of body water available for evaporation depends on body

volume ,whereas the maximum rate of cutaneous evaporation depends on surface area. At one extreme, insects have such a large surface: volume ratio that evaporative cooling is an impossible luxury. Humans and larger animals can use water for limited periods to dissipate heat during stress and, in general, the larger the animal the longer it can survive without an external water supply. The camel, for example, can lose up to 30% of its body mass without harm when deprived of drinking water for long periods because it has several physiological adaptations to survive dehydration. But humans become critically dehydrated after a water loss of only about 12% of their body weight—typically about 8 l of water.

Convection (C)

Convective heat loss from animals can be an important component in the energy balance. Using principles discussed in previous chapters we can investigate the relation between skin temperature, wind speed, net radiation absorbed, and animal size. Figure 3 illustrates this analysis graphically. The left-hand side of Figure.3 demonstrates how resistance to convective heat transfer between an organism and its environment decreases as body size decreases, a significant relation in animal ecology (see also Figure 13.1). The right-hand side of the diagram, for which the theoretical basis will be discussed later (p. 262), shows the linear relationship between **R**ni and r^{-1} HR for specified values of $(T_f - T)$, the temperature difference between the organism and its environment when convection and isothermal net radiation are equated (see Eq. (13.2)). In the example given, an organism with characteristic dimension 2 cm has r^{-1} H = 6.7 cm s^{-1}. Transferring this to the right-hand side of the figure, we see that,

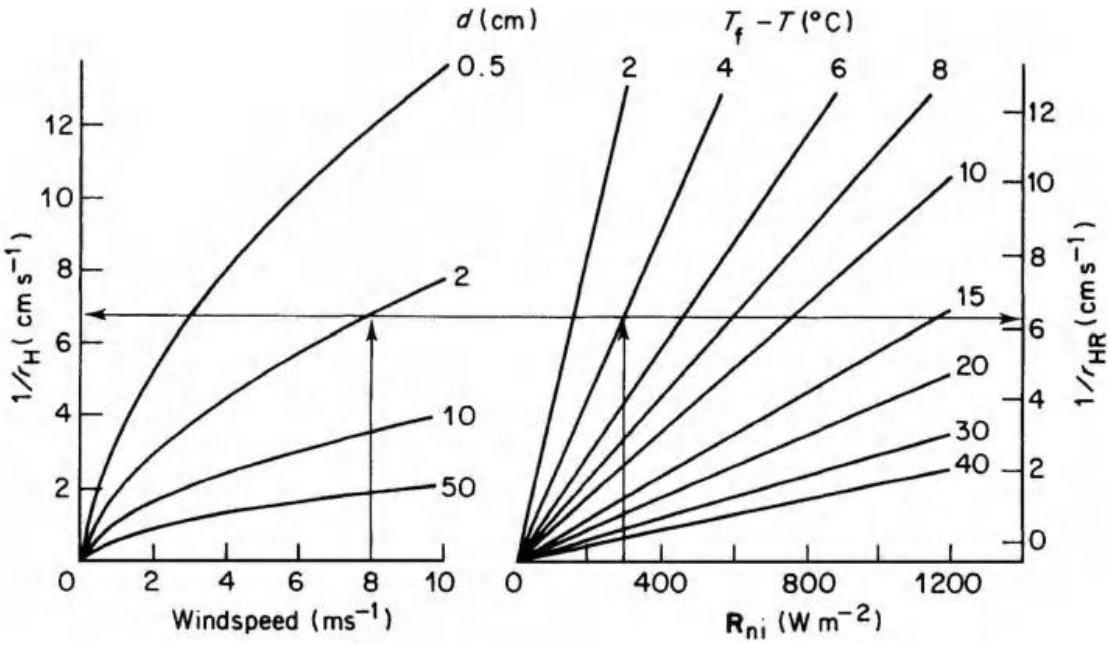


Figure 3 Diagram for estimating the thermal radiation increment (see text and p. 262) when wind speed, body size, and net radiation are known. For example, at 8 m s−1 an animal with *d* = 2 cmhas *r*−1 H = 6.7 cm s−1 and *r*−1 HR = 6.2 cm s−1. When **R**ni = 300 W m−2, *T*f − *T* is 4 ◦C.

when **R**ni = 300 W m−2, *(T*f − *T)*, the amount by which the effective air temperature exceeds the actual air temperature because of radiative loading (see p. 262), is 4 ◦C. Schmidt-Nielsen (1965) regarded both convective and radiative exchanges as proportional to surface area and concluded that, in the desert, "the small animal with its relatively larger relative surface area is in a much less favourable position for maintaining a tolerably low body temperature." It is true that small animals are unable to keep cool by evaporating body water, but when convection is the dominant mechanism of heat loss they can lose heat more rapidly (per unit surface area) than large animals exposed to the same wind speed. In this context, the main disadvantage of smallness is microclimatic: because wind speed increases with height above the ground, small animals moving close to the surface are exposed to slower wind speeds than larger mammals. As *r*H is approximately proportional to *(d/V)*0.5, an animal with *d* = 5 cm exposed to a wind of 0.1 m s−1 will be coupled to air temperature in the same way as a much larger animal with *d* = 50 cm exposed to wind at 1 m s−1. Insects and birds in flight and tree-climbing animals escape this limitation.

Conduction (G)

Few attempts have been made to measure the conduction of heat from an animal to the surface on which it is lying. Mount (1968) measured the heat lost by young pigs to different types of floor materials and found that the rate of conduction was strongly affected by posture as well as by the temperature difference between the body core and the substrate. When the temperature of the floor (and air) was low, the animals assumed a tense posture and supported their trunks off the floor, but as the temperature was raised, they relaxed and stretched to increase contact with the floor. Figure 4 shows the heat loss per unit area, recalculated from Mount's data, for newborn pigs in a relaxed posture. As the heat flow was approximately proportional to the temperature difference, the thermal resistance for each type of floor can be calculated from the slope of the lines. The resistances are about 8, 17, and 58 s cm−1 for concrete, wood, and polystyrene, respectively. As the corresponding resistances for convective and radiative transfer are usually around 1–2 s cm−1 (Figure 13.1), it follows that heat losses by conduction to the floor of an animal house are likely to be significant only when it is made of a relatively good heat conductor such as concrete. Conduction will be negligible when the floor is wood or concrete covered with a thick layer of straw.

Gatenby (1977) measured the conduction of heat beneath a sheep with a fleece length of about 2 cm lying on grass in the open. When the deep soil temperature was 10 ◦C, downward fluxes of about 160 W/m2 of contact surface were measured when the sheep lay down, equivalent to about 40 W/m2 of body surface and therefore comparable with the loss by convection. Energy requirements for free-ranging animals therefore depend to some extent on the time spent lying, on soil temperature, and on the thermal properties of the ground.

Storage (S)

As discussed in more detail below, most homeotherms have mechanisms that maintain their core body temperature almost constant over a wide range of

environmental

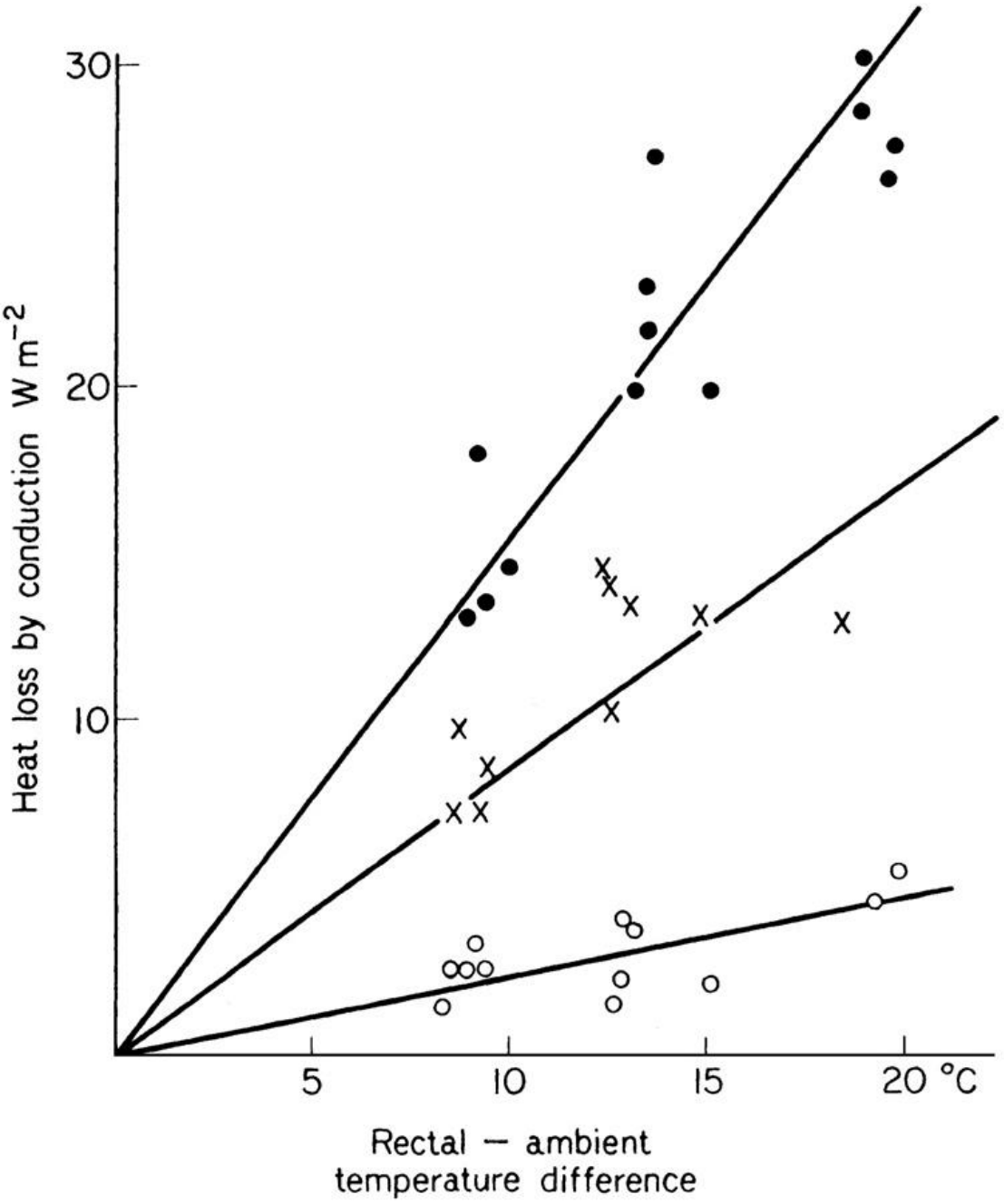


Figure 4 Measurements of heat lost by conduction from a pig to different types of floor covering expressed aswatts per square meter of total body area (after Mount, 1967).

•—concrete; X—wood; _—polystyrene. conditions. On the other hand, the body temperatures of poikilotherms vary in response to external heat inputs. All animals alter their behavior to avoid extreme conditions, e.g. seeking shade, sheltering from wind, or adopting a nocturnal lifestyle. In desert environments where day-time air temperatures may be much larger than body temperature and minimizing evaporative

loss is important, some homeotherms allow their body temperature to increase during the day, thus reducing the need for evaporative cooling not only by storing heat but also by reducing sensible heat gain (i.e. reducing the air-skin temperature gradient) and increasing long-wave radiation loss. The stored heat is dissipated at night by sensible and radiative heat loss. Schmidt-Nielsen et al. (1956) studied the body temperature of dromedary camels kept in full desert sun. When the camels were allowed to drink water at the beginning of each day, the body temperature varied by about 2.1 ◦C between day and night, but when the camels were eprived of water their body temperatures during the day were up to 6.5 ◦C greater

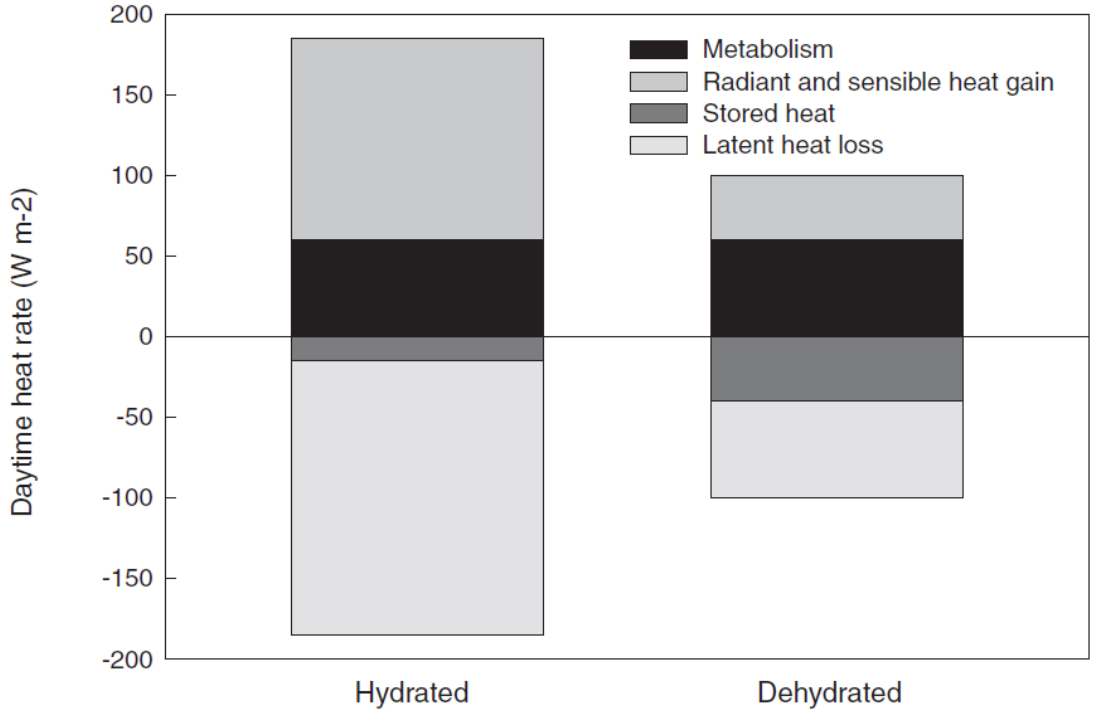


Figure 5 Heat balance components of a camel in a desert environment when hydrated and dehydrated, expressed as mean heat rates *(*W m−2*)* over the 10 h during which external heat loads were largest (based on data from Schmidt-Neilsen et al. 1956).

than at night. Such a range would be life-threatening for many mammals. For a camel weighing 260 kg, assuming a body specific heat of 3.6 kJ kg−1 K, the daytime heat storage rate **S** averaged 15 W m−2 for the hydrated and 42 W m−2 for the dehydrated animals. Schmidt-Neilsen also measured evaporation rate **E** from the skin (ignoring the small respiratory evaporation rate), and estimated metabolic heat production **M**. Using the heat balance Eq. (14.1), he interpreted the residual **E** + **S** − **M** as the heat

gained from net radiation and sensible heat. Figure 14.5 illustrates the heat balances over the hottest 10 h of the day for the hydrated and dehydrated camels. When water was available, the animals balanced their energy gains principally by evaporation of sweat; when dehydrated, the sweating rate was much decreased, and heat storage was similar in magnitude to the loss of heat by evaporation. Some of the large reduction in the radiant and sensible heat gain is probably due to postural changes that the camels adopted to reduce radiation interception when short of water.

Although heat storage is not viable as a long-term component of the heat balance of most homeotherms, it is important during short bursts of activity. For example, Taylor and Lyman (1972) found that running antelopes generated heat at about 40 times their resting metabolic rate, and the increase of their body temperature of up to 6 ◦C over 5– 15 min (depending on running speed) accounted for 80–90% of this heat. For humans, body temperature Tb also increases during exercise in direct proportion to the metabolic rate **M** until the limits for thermoregulation are reached. A convenient approximation for humans (Kerslake, 1972) is

Tb = 36.5 + 4.3 × 10−3 **M**,

where Tb is in Celsius when **M** is in W m−2.

Specification of the Environment—The Effective Temperature

The thermo-neutral diagram has been used mainly to summarize observations of metabolic rate in livestock or humans exposed to the relatively simple environment of a calorimeter. In the real world, metabolic rate depends on several microclimatic factors in addition to air temperature, notably radiation, wind speed, and vapor pressure.

Attempts have therefore been made to replace the simple measure of actual air temperature used in the previous discussion with an *effective temperature* which incorporates the major elements of microclimate. Two examples are now given, the first dealing with the radiative component of the heat balance equation.

1 Radiation Increment

When the radiant flux is expressed in terms of net isothermal radiation, the heat balance equation may be written

Rni + *(***M**− **λE***)* = ρcp*(T*0 − *T)/r*HR, (14.13)

where **λE** is here assumed to be a relatively small term, independent of T . Suppose

that the thermal effect of radiation on an organism can be handled by substituting for air temperature an *effective* temperature *T*f such that

*(***M**− ***λE**) = *ρc*p _*T*0 − *T*f_ /*r*HR. (14.14)

Then eliminating **M**− ***λE*** from Eqs. (14.13) and (14.14) gives

*T*f = *T* + {**R**ni*r*HR/*ρc*p}, (14.15)

where the term in curly brackets is similar to the radiation increment used by Burton and Edholm (1955), Mahoney and King (1977), and many others. Figure 3 provides a graphical method of estimating the temperature increment *T*f−*T* when windspeed and net radiation are known.Given the characteristic dimension of an organism *d* and the velocity of the surrounding air *V*, the corresponding value of *r*H can be read from the left-hand vertical axis and *r*HR derived from Equation (13.2) can be read from the right-hand axis. From the right-hand section of the figure, the coordinates of *r*HR and **R**ni define a unique value of *T*f − *T* . An example is shown for *V* = 8 m s−1, *d* = 2 cm,which gives 1/*r*H = 6.8 cm s−1 from the left-hand axis and 1/*r*HR = 6.3 cm s−1 from the right-hand axis. At **R**ni = 300 W m−2, *T*f − *T* = 4 K. In a system where the resistances are fixed, the relation between temperature gradients and heat fluxes can be displayed by plotting temperature against flux as in Figure 7. By definition, a resistance is proportional to a temperature difference divided by a flux and is therefore represented by the slope of a line in the figure. From a start at the bottom left-hand corner, *T*f is determined by drawing a line (1) with slope *r*HR/*ρc*p to intercept the line *x* = **R**ni at *y* = *T*f. Rearranging Equation (14.15), the equation of this line is

*T*f − *T* = *r*HR**R**ni/*ρc*p.

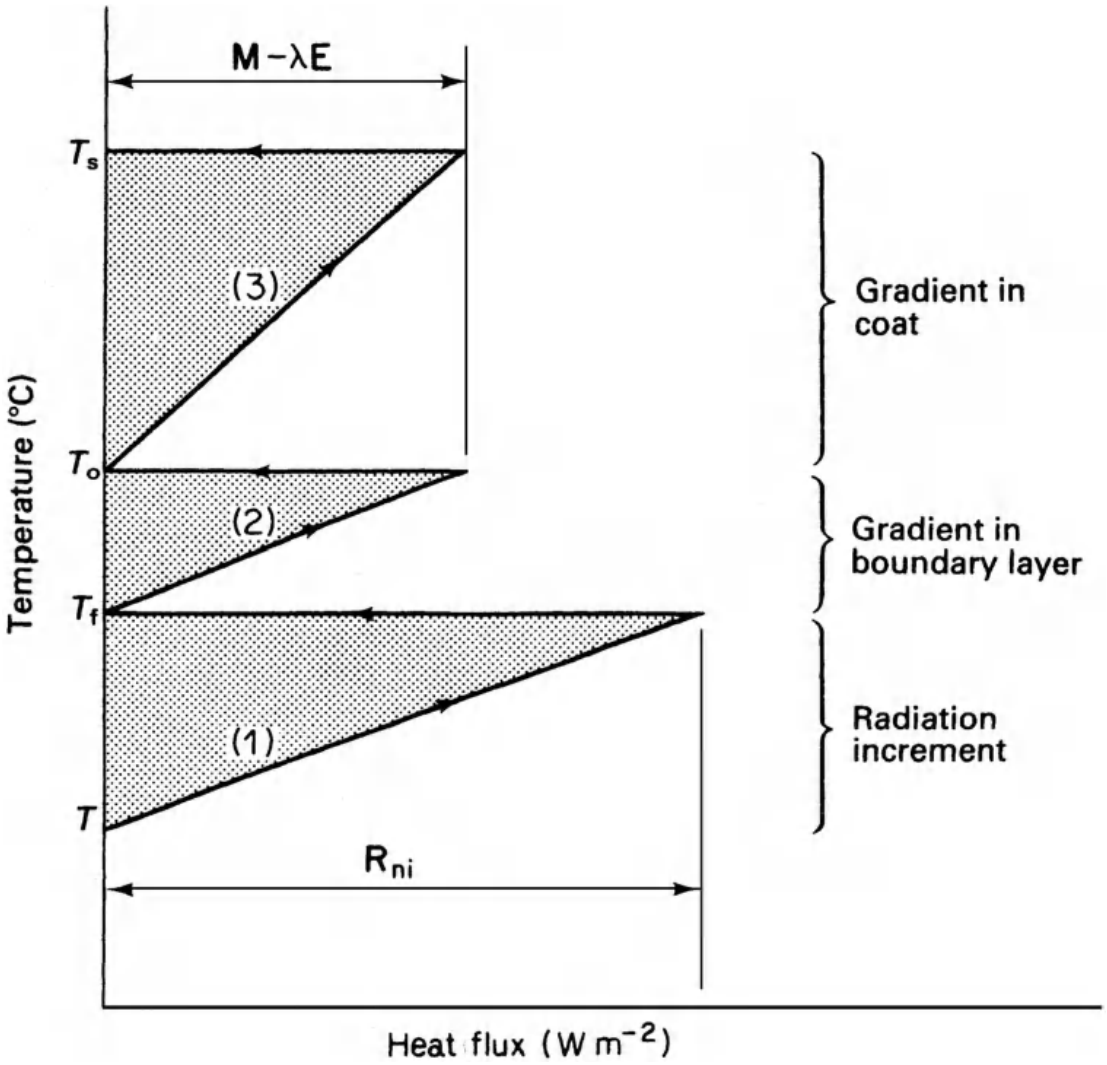


Figure 7 Main features of temperature/heat-flux diagram for dry systems. *T*s is skin temperature, *T*0 is coat surface temperature, *T*f is effective environment temperature, and *T* is air temperature.

The temperature of the surface *T*0 is now found by drawing a second line (2) with the same slope as (1) to intersect $x = \mathbf{M}-\boldsymbol{\lambda}\mathbf{E}$ at $y = T0$. The equation of this line (see Equation (14.14)) is

$T0 - Tf = rHR(\mathbf{M} - \boldsymbol{\lambda}\mathbf{E})/\rho cp$. (14.17)

Finally, for an animal covered with a layer of hair, a mean skin temperature *T*s can be determined if the mean coat resistance *r*c is known. Provided that evaporation is confined to the surface of the skin and the respiratory system, the increase of temperature through the coat is represented by the line (3) whose equation is

$Ts - T0 = rc(\mathbf{M} - \boldsymbol{\lambda}\mathbf{E})/\rho cp$. (14.18)

This form of analysis can be used to solve two types of problems. When the environment of an animal is prescribed in terms of wind speed, temperature, and net radiation, it is possible to use a thermo-neutral diagram to explore the physiological states in which the animal can survive. Conversely, when physiological conditions are specified, a corresponding range of environmental conditions can be established, forming an ecological "niche" or "climate space" (Gates, 1980).

4 Apparent Equivalent Temperature

A second type of environmental index is needed when the loss of latent heat from an organism is predominantly by the evaporation of sweat from the skin. When the losses of sensible and latent heat occur from the same surface, temperature and vapor pressure can be combined in a single variable which may be called the *apparent equivalent temperature*. To derive this quantity, the heat balance equation is written,

$$\mathbf{R}\text{ni} + \mathbf{M} = \rho c\text{p}(T0 - T)r\text{HR} + \rho c\text{p}(e0 - e)\,\gamma\, r\text{v}\,, \quad (14.20)$$

where $T0$ and $e0$ are mean values for the skin surface and T and e refer to the air.

If γ is replaced by $\gamma * = \gamma\,(r\text{v}/r\text{HR})$, Eq. (14.20) can be written

$$\mathbf{R}\text{ni} + \mathbf{M} = \rho c\text{p}\,\_T * \text{eo} - T * \text{e}\,\_/r\text{HR}, \quad (14.21)$$

where $T * \text{e}$ is the apparent equivalent temperature of ambient air given by $T + e/\gamma *$ and therefore equal to the equivalent temperature derived on p. 223 when $r\text{v} = r\text{HR}$. By analogy with the radiation increment, $e/\gamma *$ may be regarded as a humidity increment. The mean value of the apparent equivalent temperature at the surface is $T * \text{eo}$. In principle, the apparent equivalent temperature should be used in a thermo-neutral diagram in place of conventional temperature when the metabolic rate of an animal is measured in an environment where vapor pressure changes as well as temperature. Extending this process a stage further, the apparent equivalent temperature of the environment can be modified to take account of the radiation increment by writing

$$T * \text{eR} = T * \text{e} + \mathbf{R}\text{ni}r\text{HR}/\rho c\text{p}. \quad (14.22)$$

The quantity $T * \text{eR}$ is an index of the thermal environment which allows the heat balance equation to be reduced to the form

$$(1 - x)\mathbf{M} = \rho c\text{p}\,\_T * \text{eo} - T * \text{eR}\_\_\, r\text{HR}, \quad (14.23)$$

where x is the fraction of metabolic heat dissipated by respiration. A thermo-neutral

diagram using $T * eR$ as a thermal index rather than T or $T * e$ would be valid for changes in radiant heat load as well as in vapor pressure and temperature.

This type of formulation provides a relatively straightforward way of investigating the heat balance of a naked animal whose skin is either dry or entirely wetted by sweat.

The intermediate case of partial wetting is more difficult to handle because the value of $T * e$ depends on fractional wetness which is not known *a priori*. A complete solution of the relevant equations for an animal leads to a complex expression within which the structure of the original Penman equation can be identified (McArthur, 1987).

Problems

1. The metabolic heat production **M** of a cow standing in strong sunshine at air temperature 22.0 ◦C is 140 W m−2. The average solar irradiance of the animal is 300 W m−2 and the effective radiative temperature of the surroundings is equal to air temperature. The mean temperatures of the skin and coat surface are 34.0 and 31.6 ◦C, respectively, and the rates of evaporative loss from the respiratory system and the skin surface are 4.5 × 10−3 g m−2 s−1 and 40 × 10−3 g m−2 s−1, respectively. Calculate the thermal resistances of (i) the boundary layer around the cow (*r*HR) and (ii) the coat (*r*c). If shading reduces the average solar irradiance of the animal by two-thirds, and other environmental variables remain unchanged, calculate the corresponding reduction in the rate of evaporation from the skin necessary to maintain a steady heat balance if **M** remains constant. (Assume that changes in respiratory evaporation can be neglected and that the absorption of solar radiation occurs close to the outer surface of the coat, which has a reflection coefficient of 0.40.)

2. A pig, trunk diameter 0.50 m, with its skin covered in wet mud at 30◦C is exposed in still air when air temperature is 25 ◦C and relative humidity is 30%. Calculate the resistance *r*v for water vapor transfer from the skin to the atmosphere, (i) ignoring humidity gradients and (ii) taking humidity into account. Calculate the evaporative flux of water per unit skin area. What percentage error would you have made if you had ignored effects of humidity gradients on *r*v?What other factorswould determine the evaporation rate in practice?

3. Pigs wallow in wet mud to increase body cooling by evaporation. Calculate the maximum air temperature at which a pig could maintain a state of thermal equilibrium

(i) if the skin was completely dry and (ii) if the skin was completely covered in wet mud. Assume the following:

Minimum metabolic heat production rate 60 W m−2

Rate of latent heat loss by respiration 10 W m−2

Radiative heat load from the environment 240 W m−2

Skin temperature 33 ◦C

Vapor pressure of air 1.0 kPa

Combined resistance to radiative and convective heat transfer 80 s m−1

Mean resistance to conduction through mud layer 8 s m−1

x-rays, microwaves

Introduction

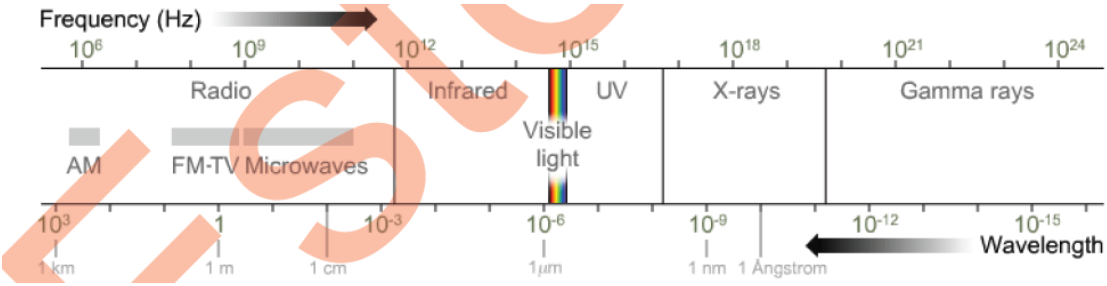
Radio and television signals, x-rays, microwaves each is a form of electromagnetic radiation. If steam and internal combustion engines symbolize the Industrial Revolution, and microprocessors and memory chips now power the Information Revolution, it almost seems that we have neglected to recognize the "Electromagnetic Revolution." Think about it: Can you imagine life without television sets or cell phones? You may long for such a life, or wonder how people ever survived without these devices! These examples are from the world of engineered electromagnetic radiation. Even if you think we might all prosper without such technologies to entertain us, do our cooking, carry our messages, and diagnose our illnesses, you would be hard-pressed to survive without light. This form of electromagnetic radiation brings the Sun's energy to the Earth, warming the planet and supplying energy to plants, and in turn to creatures like us that depend on them. There are primitive forms of life that do not depend on the Sun's energy, but without light there would be no seeing, no room with a view, no sunsets, and no Rembrandts.

Some of the electromagnetic radiation that reaches your eyes was created mere nanoseconds earlier, like the light from a lamp. Other electromagnetic radiation is still propagating at its original speed through the cosmos, ten billion years or more after its birth. An example of this is the microwave background radiation, a pervasive remnant of the creation of the universe that is widely studied by astrophysicists.

Back here on Earth, this chapter covers the fundamental physical theory of electromagnetic radiation. Much of it builds on other topics, particularly the studies of

waves, electric fields and magnetic fields. **Electromagnetic radiation: Rainbows and radios. Sundazzled reflections. Shadowlamps and lampshadows. Red, white, and blue.**

1 - The electromagnetic spectrum

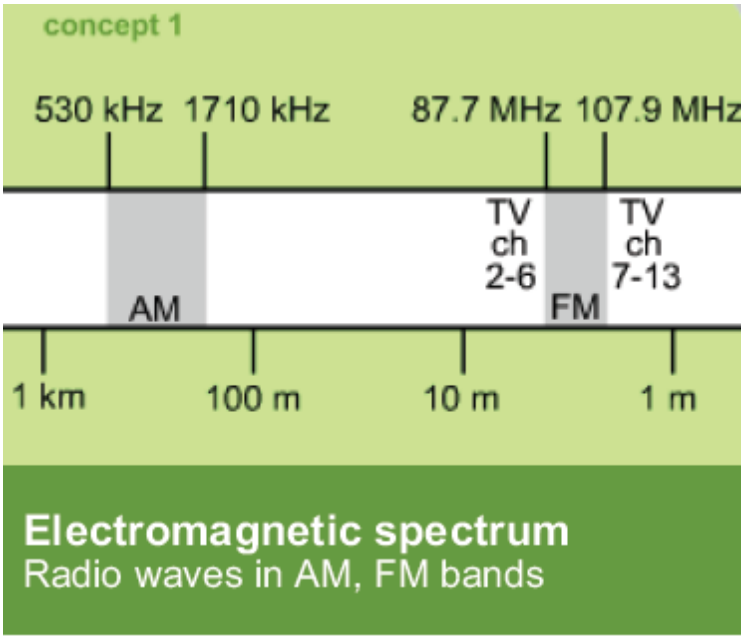


Electromagnetic spectrum: Electromagnetic radiation ordered by frequency or wavelength.

Electromagnetic radiation is a traveling wave that consists of electric and magnetic fields. Before delving into the details of such waves, we will discuss the electromagnetic spectrum, a system by which the types of electromagnetic radiation are classified. The illustration of the electromagnetic spectrum above orders electromagnetic waves by frequency and by wavelength. In the diagram, frequency **increases** and wavelength **decreases** as you move from the left to the right. The chart's scale is based on powers of 10. Wavelengths range from more than 100 meters for AM radio signals to as small as 10□16 meters for gamma rays.

All electromagnetic waves travel at the same speed in a vacuum. This speed is designated by the letter *c* and is called the speed of light. (The letter *c* comes from *celeritas*, the Latin word for speed. It might be more accurate to refer to it as the speed of electromagnetic radiation.) The speed of light in a vacuum is exactly 299,792,458 m/s, and it is only slightly less in air.

The unvarying nature of this speed has an important implication: The wavelength of electromagnetic radiation is inversely proportional to its frequency. As you may recall, the speed of a wave equals the product of its frequency and wavelength. This means that if you know the wavelength of the wave, you can determine its frequency (and vice versa). For instance, an electromagnetic wave with a wavelength of 300 meters, in the middle of the AM radio band, has a frequency of 1×106 Hz. This equals 3×108 m/s, the speed of light, divided by 300 m. The frequencies of electromagnetic waves range from less than one megahertz, or 106 Hz, for long radio waves to over 1024 Hz for gamma rays. We will now review some of the bands of electromagnetic radiation and their manifestations. The lowest frequencies are often utilized for radio



Electromagnetic wave: A wave consisting of electric and magnetic fields oscillating transversely to the direction of propagation.

Physicist James Clerk Maxwell's brilliant studies pioneered research into the nature of electromagnetic waves. He correctly concluded that oscillating electric and magnetic fields can constitute a self-propagating wave that he called electromagnetic radiation. His law of induction (a changing electric field causes a magnetic field) combined with Faraday's law (a changing magnetic field causes an electric field) supplies the basis for understanding this kind of wave.

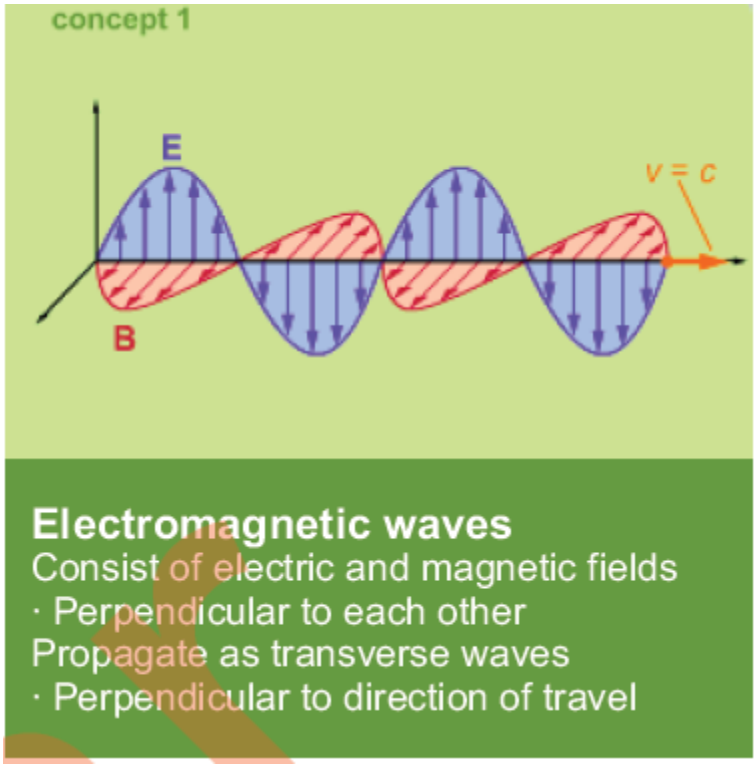
As the diagrams to the right show, the electric and magnetic fields in an electromagnetic wave are perpendicular to each other and to the direction of propagation of the wave. These illustrations also show the amplitudes of the fields varying sinusoidally as functions of position and time. Electromagnetic waves are an example of *transverse waves*. The fields can propagate outward from a source in all directions at the speed of light; for the sake of visual clarity, we have chosen to show them moving only along the *x* axis.

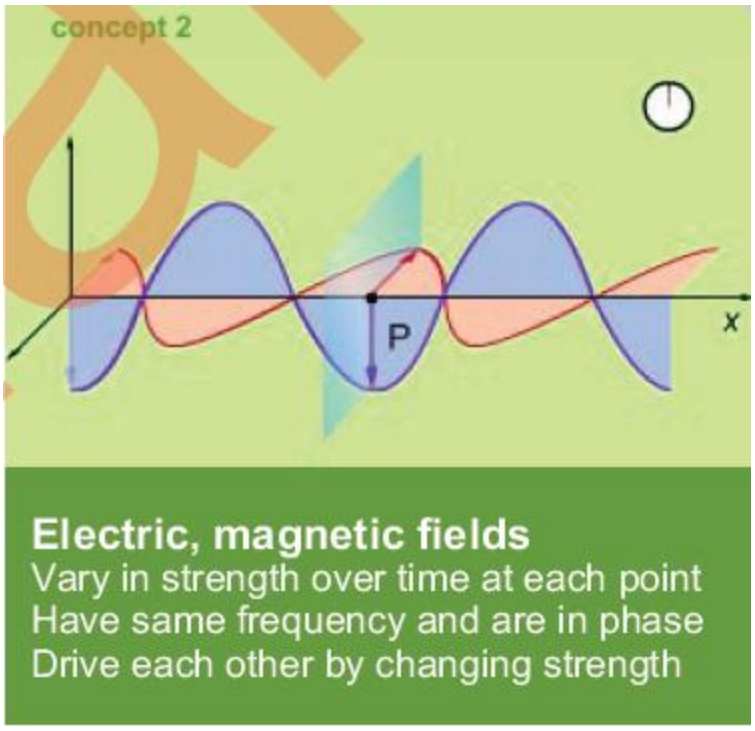
The animated diagram in Concept 2 and the illustrations below are used to emphasize three points. First, the depicted wave moves away from the source. For example, if you push the "transmit" button on a walkie-talkie, a wave is initiated that travels away from the walkie-talkie. Second, at any fixed location in the path of the wave, both fields change over time. The wave below is drawn at intervals that are fractions $T/4$ of the period T. Look at the point P below, on the light blue vertical plane. The vectors from point P represent the direction and strength of the electric and magnetic fields at this point. As you can see, the vectors, and the fields they represent, change over time at P. Concept 2 shows them varying continuously with time at the point P.

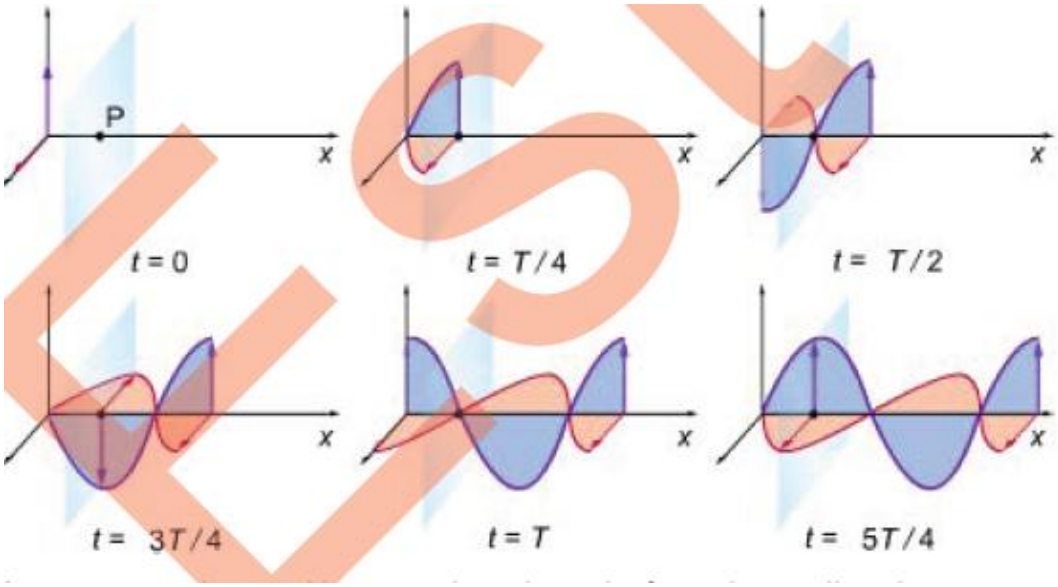
Third, the diagrams reflect an important fact: The electric and magnetic fields have the same frequency and phase. That is, they reach their peaks and troughs simultaneously. A wave on a string provides a good starting point for understanding electromagnetic waves. Both electromagnetic radiation and a wave on a string are transverse waves. The strengths of the two fields constituting the radiation can be described using sinusoidal functions, just as we can use a sinusoidal function to calculate the transverse displacement of a particle in a string through which a wave is moving. There is a crucial difference, though: Electromagnetic radiation consists of electric and magnetic fields, and does not require a medium like a string for its propagation. Electromagnetic waves can travel in a vacuum. If this is troubling to you, you are in good company. It took some brilliant physicists a great deal of hard work to convince the world that light and other electromagnetic waves do not require a medium of transmission.

Furthermore, when electromagnetic waves radiate in all directions from a compact source like an antenna or a lamp, the radiation emitted at a particular instant travels outward on the surface of an expanding sphere, and its strength diminishes with distance from the source. The waves cannot be truly sinusoidal, since the amplitude of a sinusoidal function never diminishes. In the sections that follow we will analyze

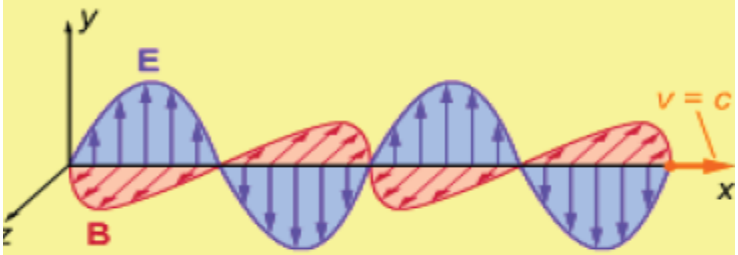
plane waves, which propagate through space, say in the positive x direction, in parallel planar wave fronts rather than expanding spherical ones. They are good approximations to physical waves over small regions that are distant from the source of the waves. Plane waves never diminish in strength; they can be accurately modeled using sinusoidal functions, and we will do so.







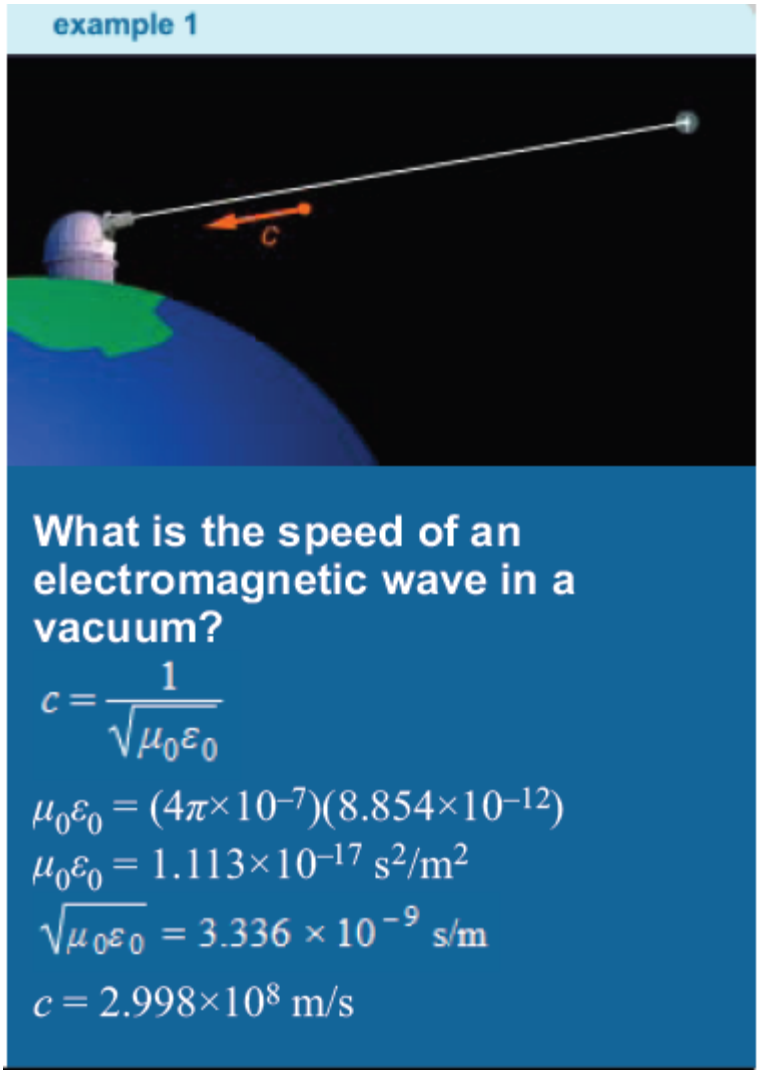
equation 1



Speed of an electromagnetic wave

$$c = \frac{\omega}{k} = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

c = speed of electromagnetic wave
ω = angular frequency of wave
k = angular wave number of wave
μ_0 = permeability of free space
Constant $\mu_0 = 4\pi \times 10^{-7}$ T·m/A
ε_0 = permittivity of free space
Constant $\varepsilon_0 = 8.854 \times 10^{-12}$ C²/N·m²



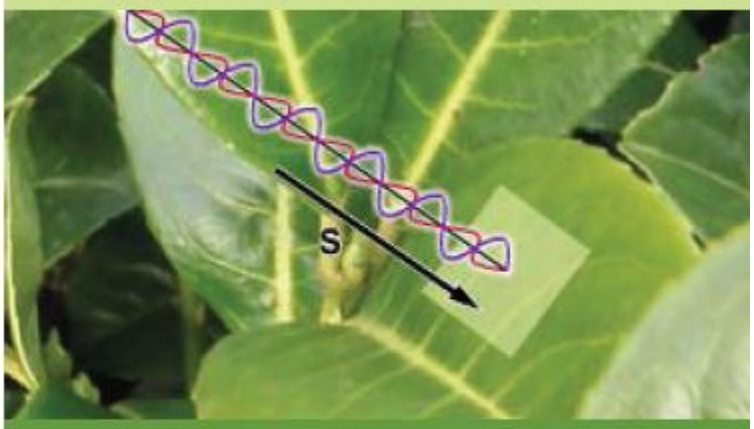
Creating electromagnetic waves: antennas Radio antennas create electromagnetic waves. A radio antenna is part of an overall system called a radio transmitter that converts the information contained in sound waves into electromagnetic waves. A radio receiver then reverses the process, converting the signals from electromagnetic waves back to sound waves. The system depicted to the right shows the fundamentals of a radio transmitter. In the illustrations, the terminals of an AC generator are connected to two rods of conducting material: an antenna. The AC generator produces an emf □ that varies sinusoidally over time. The emf drives positive

through the surface by the wave, per unit area, is called the *area power density* of the wave. The area power density is equal to the magnitude *S* of the Poynting vector. The surface area through which the instantaneous power density is measured is perpendicular to the direction of the wave's propagation. When radiation reaches a physical surface obliquely, the cosine of its angle with the area vector can be used to calculate the power conveyed to the surface. This is analogous to the calculation of electric or magnetic flux.

As Equation 1 shows, the Poynting vector equals the cross product of the vectors representing the electric and magnetic fields of the electromagnetic radiation, divided by the permeability constant. Since these fields are always perpendicular to one another, the sine of the angle between them, used to evaluate the magnitude of the cross product, always equals one, and can be effectively ignored when calculating the instantaneous area power density *S*. The units of the Poynting vector are watts per square meter. The direction of **S** is determined by the right-hand rule. If you apply the rule, wrapping your fingers from **E** to **B** and noting the direction of your thumb, you can correctly determine that it is parallel to the direction of propagation of the wave. When **E** reverses its direction, so does **B**, and the direction of **S** remains the same, "pointing" (heh, heh) in the direction of the wave's motion.

As an electromagnetic wave passes through a surface, the strengths of its electric and magnetic fields there change sinusoidally with time. Since the Poynting vector is the product of these fields, it changes sinusoidally over time, as well. In fact, it varies with values between zero and $E\text{max}\ B\text{max}/\mu 0$, with a frequency twice that of the fields. If you are curious why it has this frequency, recall from the field equations that *E* and *B* are both cosine functions of time at a fixed point. Then use the trigonometric identity $\cos 2\ t = [1 + \cos 2t]/2$.

concept 1

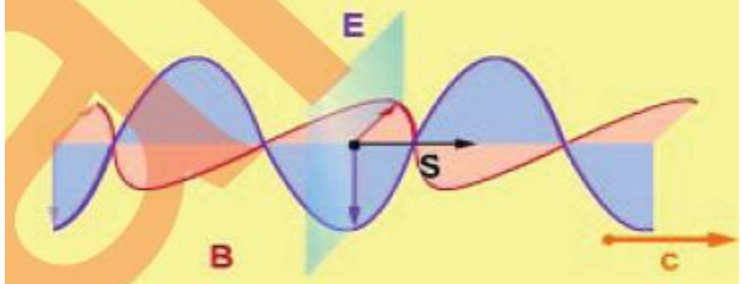


Poynting vector

Power per unit surface area

Surface perpendicular to wave direction

equation 1



Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}$$

$\mathbf{S}$ = Poynting vector

S = instantaneous area power density

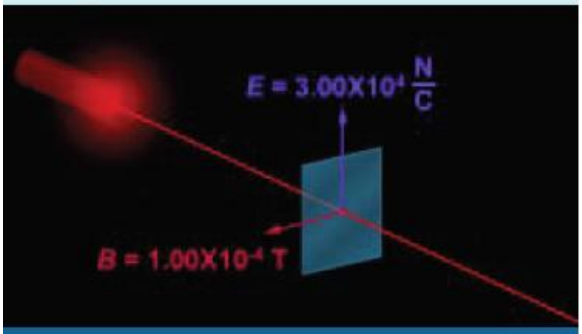
μ_0 = permeability of free space

$\mathbf{E}$ = electric field

$\mathbf{B}$ = magnetic field

Units: watts per square meter (W/m^2)

example 1



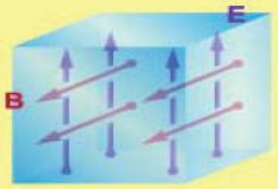
At this instant, what is the area power density of the ruby laser light?

$S = EB / \mu_0$

$$S = \frac{(3.00 \times 10^4 \frac{\text{N}}{\text{C}})(1.00 \times 10^{-4}\ \text{T})}{4\pi \times 10^{-7} \frac{\text{T} \cdot \text{m}}{\text{A}}}$$

$S = 2.39 \times 10^6\ \text{W/m}^2$

equation 2



Electric, magnetic energy densities

$$u_E = \frac{\varepsilon_0 E^2}{2} \qquad u_B = \frac{B^2}{2\mu_0}$$

Since $E^2/B^2 = c^2 = 1/\mu_0\varepsilon_0$, then

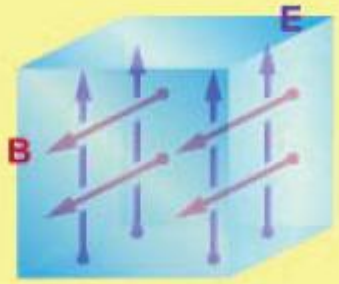
$$u_E = u_B$$

The total energy density is

$$u = u_E + u_B = 2u_E = 2u_B$$

u_E = electric field energy density
u_B = magnetic field energy density
u = total energy density

equation 3



Average energy per unit volume

Average value of u over time

$$u_{avg} = \frac{\varepsilon_0 E_{max}^2}{2}$$

u_{avg} = average total energy density

Units: joules per cubic meter (J/m^3)

How electromagnetic waves travel through matter Light and other forms of electromagnetic radiation can travel through a vacuum, and it is often simplest to study them in that setting. However, radiation can also pass through matter: If you look through a glass window, you are viewing light that has passed through the Earth's atmosphere and the glass. Other forms of radiation such as radio waves pass through matter, as well.

This section focuses on how such transmission occurs. It relies on a classical model of electrons and atoms that predates quantum theory. In this model, electrons orbit an atom. They have a resonant frequency that depends on the kind of atom. On a larger scale, atoms themselves and the molecules composed of them also have resonant thermal frequencies at which they can vibrate or rotate.

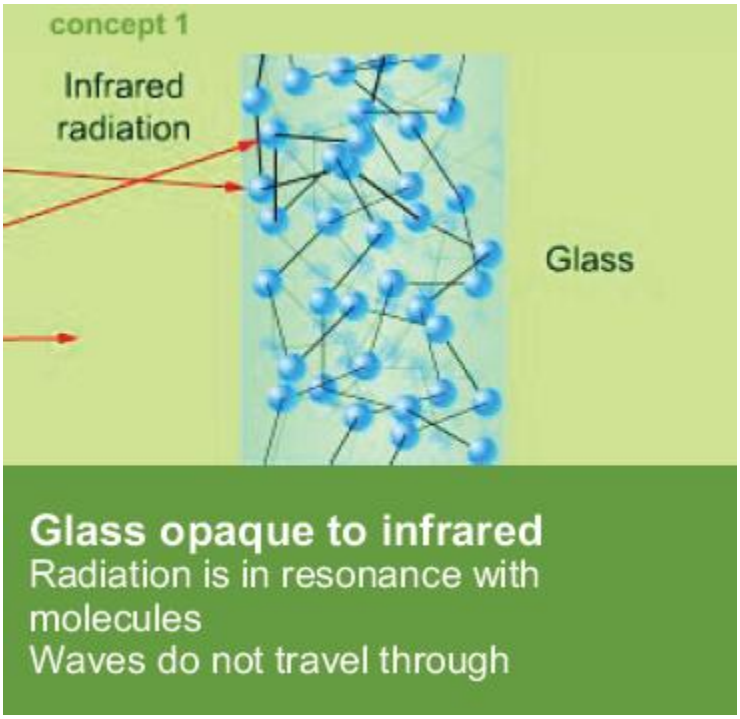
We will use the example of light striking the glass in a window to discuss how substances transmit (or do not transmit) electromagnetic radiation. When an electromagnetic wave encounters a window, it collides with the molecules that make up the glass. If the frequency of the wave is near the resonant thermal frequency of

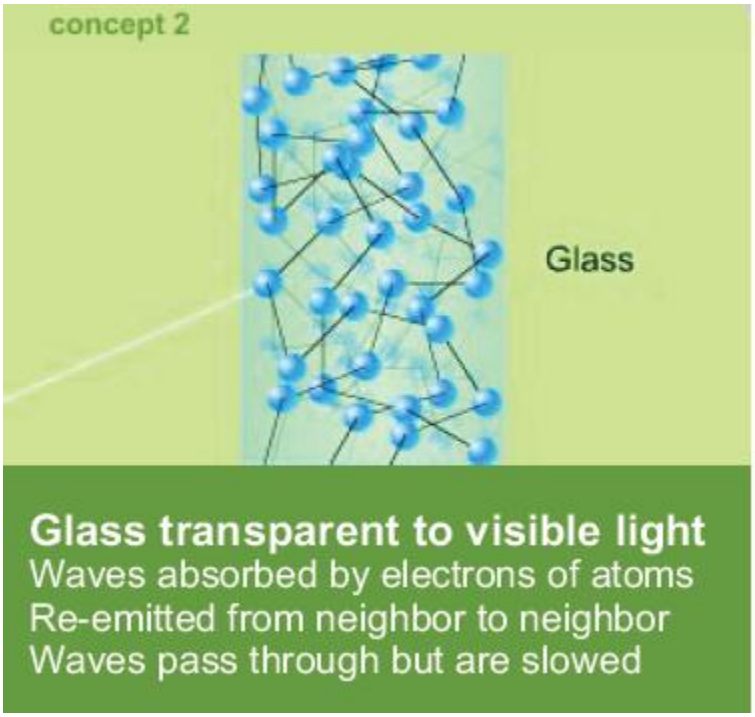
the glass molecules, which is true for infrared radiation, the amplitude of the molecules' vibrations increases. They absorb the energy transported by the wave, and dissipate it throughout the glass by colliding with other molecules and heating up the window. Because it absorbs so much infrared energy, the glass is opaque to radiation of this frequency, preventing its transmission.

Scientists in the 19th century noted a phenomenon in greenhouses caused by the opaqueness of glass to infrared radiation, which they called the *greenhouse effect*. The glass in a greenhouse admits visible light from the Sun, which is then absorbed by the soil and plants inside. They reradiate the solar energy as longer infrared waves, which cannot pass back out through the glass and so help warm up the greenhouse. The same phenomenon occurs on a vaster scale in the atmosphere as gases like methane and carbon dioxide trap solar energy near the Earth's surface.

In contrast to infrared radiation, higher frequency radiation such as visible light does not resonate thermally with atoms or molecules, but may resonate with the electrons of the atoms of a substance. In glass, visible light experiences much less reduction in the amplitude of its waves than infrared radiation does, and most of its energy passes through the glass quite easily. Atoms with resonant electrons that do absorb energy from a light wave quickly pass on that energy by re-emitting it as radiation of the same frequency to other atoms, which in turn pass it on to their neighbors.

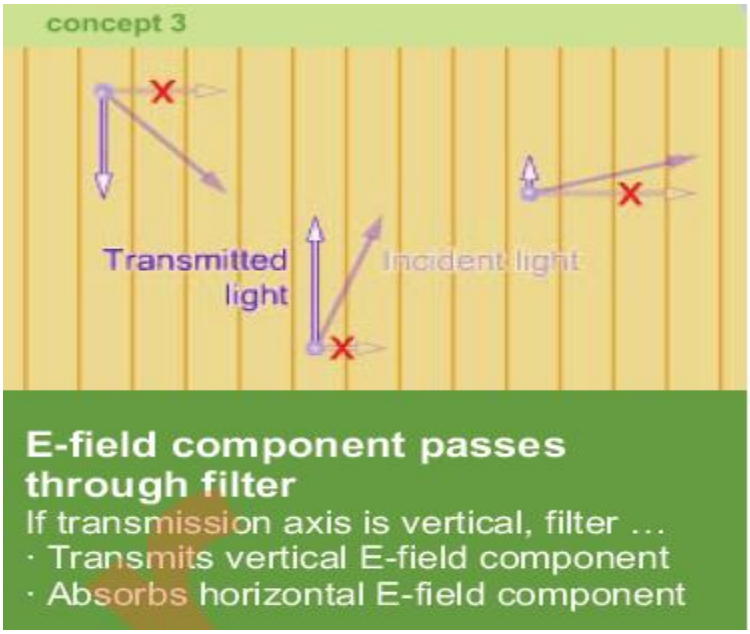
This chain of absorptions and re-emissions, called *forward scattering*, follows a path close to the light's original direction of travel. A beam of light that strikes a pane of glass will reach the "last atom" on the far side of the pane in an extremely short time. We see the light after it emerges, and think of glass as transparent.





Radiation also can be *partially polarized*, having a few waves oscillating in all planes, but with most of its waves concentrated in a single plane. This is true of sunlight scattered by the atmosphere. As the photo above shows, the sky in certain directions is partially polarized in a vertical plane so that most of its light can pass through a pair of sunglasses whose transmission axis is vertical. Less light (but still some) passes through the rotated sunglasses. (Polarizing sunglasses are specifically intended to

reduce horizontally polarized glare reflected from roadways and water, not skylight.) Many forms of artificial electromagnetic radiation are polarized. A radio transmitter emits polarized radiation. If the rods of its antenna are vertical, then so is the electric field of every radio wave it creates. In this case, the most efficient receiving antenna is also vertically oriented; a horizontal receiving antenna would absorb radio waves much less efficiently. You may be familiar with this fact if you have ever tried to maneuver a radio antenna wire or a set of television "rabbit ears" to get the best reception. (If you do not know what "rabbit ears" are for television, well, before there was cable television, there was….)



Scattering of light *Scattering:* Absorption and re-emission of light by electrons, resulting in dispersion and some polarization.

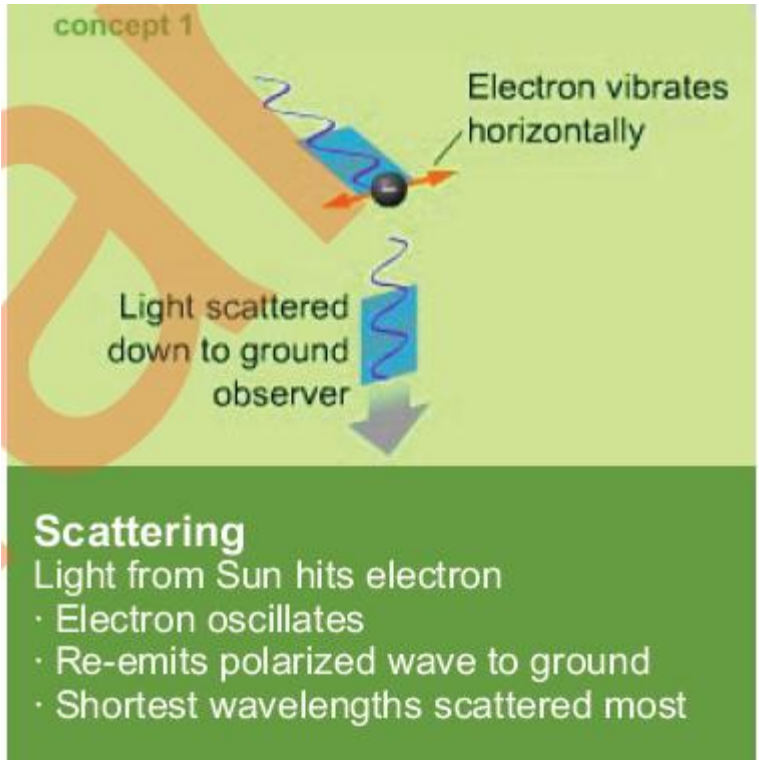
The answer to a classic question Why is the sky blue? rests in a phenomenon called scattering. In this section, we give a classical (as opposed to quantum mechanical) explanation of how scattering occurs. When light from the Sun strikes the electrons of various atoms in the Earth's atmosphere, the electrons can absorb the light's energy, oscillating and increasing their own energy. The electrons in turn re-emit this energy as light of the same wavelength. In effect, the oscillating electrons act like tiny antennas, emitting electromagnetic radiation in the frequency range of light.

An electron oscillates in a direction parallel to the electric field of the wave that energizes it, as shown in Concept 1. The electron then emits light polarized in a plane parallel to its vibration. We show a particular polarized wave that is re-emitted downward toward the ground, since we are concerned with what an observer on the surface of the Earth sees. Other light is scattered in other directions, including light scattered upward and light scattered forward in its original direction of travel. Scattering explains why we see the sky: Light passing through the atmosphere is redirected due to scattering toward the surface of the Earth. In contrast, for an astronaut observer in the vacuum of space, sunlight is not scattered at all so there is no sky glow:

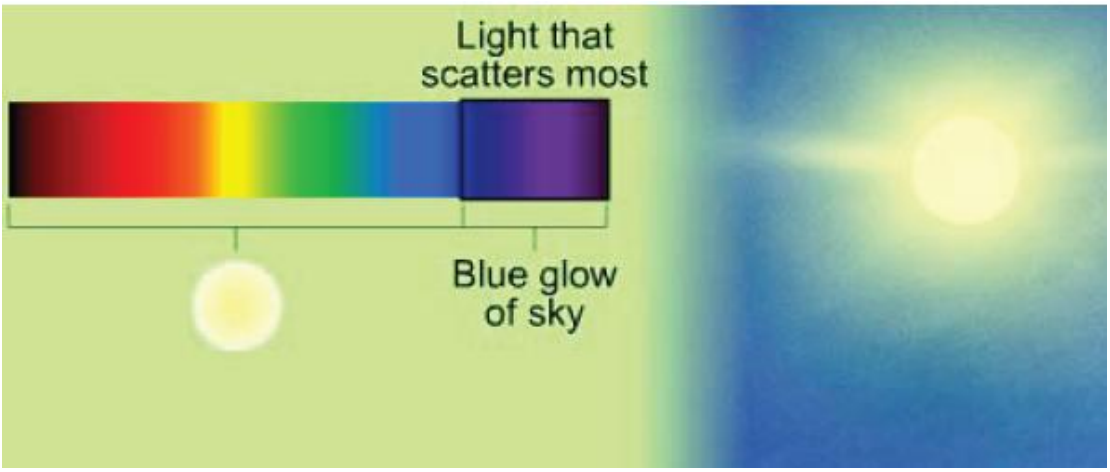
Except for the stars, the sky appears black. To the astronaut, the disk of the Sun, a combination of all colors, looks white. We illustrate this below: The full spectrum combines to form white light. The question remains, why is our sky blue rather than some other color? Light at the blue end of the visible spectrum, which has the shortest wavelength, is 10 times more resonant with the electrons of atmospheric atoms than red light. This means blue light is scattered more than red, so that more of it is redirected toward the ground.

Scattering also explains why we see the Sun as yellow rather than white. When you look up at the disk of the Sun from the Earth's surface, the bluest portion of its light has been scattered away to the sides. The remaining part of the Sun's direct light appears somewhat yellowish. You may also have noted how the Sun appears to change color when it sets. As the Sun's disk descends toward the horizon, its light must pass through a greater and greater thickness of atmosphere in order to reach you. Since a certain amount of sunlight is scattered aside for each kilometer of atmosphere

it passes through, its position at sunset causes it to lose large amounts of light at the blue end and even toward



View from space
No scattering: sky is black
Sun appears white



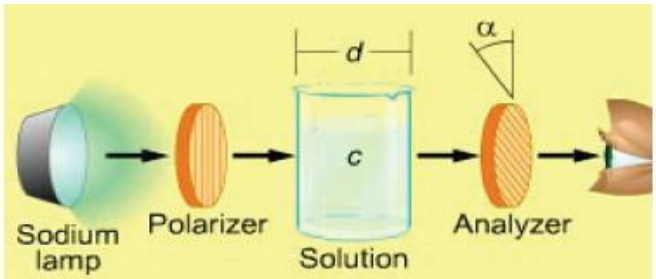
Why the sky is blue (and the Sun is yellow)
Shortest waves are scattered: sky is blue
With blue scattered, Sun appears yellowish white

the concentration of the dissolved substance. The rotation is also proportional to a constant □0 called the *specific rotation* of the substance, which reflects the rotating power of its molecules. These relationships are summarized to the right in the *polarimeter equation*. Note that (and this is unusual for a physics equation) the rotation angle □ is measured in degrees rather than radians, and the clockwise direction is considered positive.

The *polarimeter* is a device that can be used to measure the net rotation of polarized light passing through an optically active solution. An experimenter directs polarized light through a container of the solution to be analyzed. The analyzer, which starts out parallel to the polarizer, does not transmit all the light from the polarizer because the light's plane of polarization has been rotated by the solution. The experimenter turns the analyzer to one side or the other until the transmitted light has maximum

brightness. Then she knows that the analyzer's transmission axis matches the rotated polarized light, and she can measure the angle through which the analyzer has turned. The polarimeter equation gives an expression for the angle of the analyzer at which the transmitted light will be the brightest. If the polarized light encounters more molecules of the optically active substance, either because the solution is more concentrated or because the immersed light path is longer, the rotation will be greater. Since the amount of rotation also depends on the wavelength of the light, the specific rotations 0 given in tables for particular dissolved substances are based on a polarimeter employing the 589 nm light that is emitted by a sodium vapor lamp. Dextrose and fructose molecules are chemically identical (they have the same atoms arranged in the same pattern) but they are mirror images of each other. Because of this they rotate polarized light by the same amount in opposite directions. Organic molecules such as *carvone* may exist in two mirror image forms; you smell carvone as caraway or spearmint, depending on which way the molecule twists. The scents are different because the smell receptors in the nose react differently to the mirror image forms. Using a polarimeter is one way to distinguish between the two forms of mirror image compounds. Also, if the specific rotation of a particular substance is known, the device can be used together with the polarimeter equation to determine the concentration of the substance in a solution. You are asked to perform such an analysis in the example problem to the right.

The diagrams below show the mirror image molecular forms of the citrus oil *limonene,* which is the essence of either orange or lemon, depending on the orientation of its molecules! (The gray spheres represent carbon atoms, and the blue spheres are hydrogen atoms.)

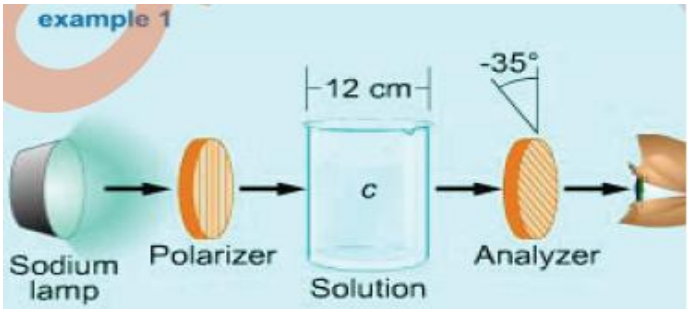


Polarimeter equation

$$\alpha = dc\alpha_0 / 100$$

α = rotation of light (° clockwise)
d = length of immersed light path (m)
c = concentration of substance (kg/m^3)
α_0 = specific rotation of substance
Units of α_0: °m^2/kg

example 1



Carvone's specific rotation is +62.5 (caraway) or –62.5 (spearmint). What is the concentration of the carvone in this beaker?

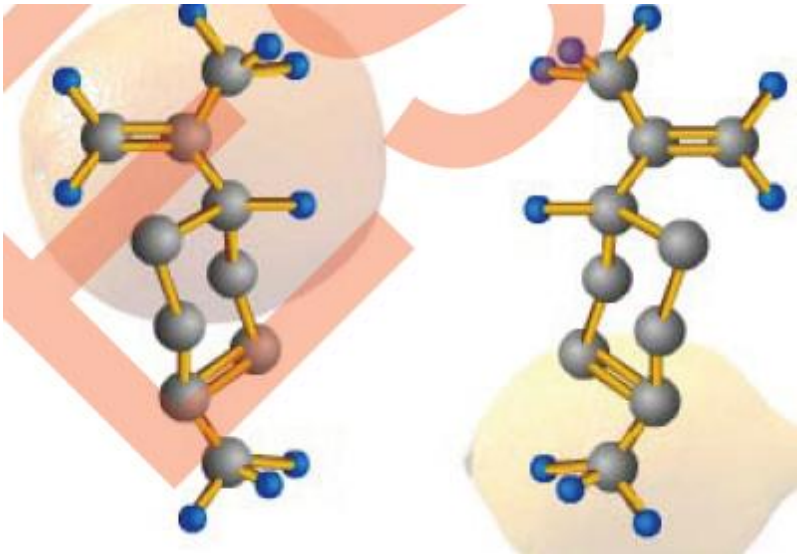
$\alpha = dc\alpha_0 / 100$

$$c = \frac{100\alpha}{d\alpha_0}$$

Counterclockwise rotation means spearmint, so we use $\alpha_0 = -62.5$:

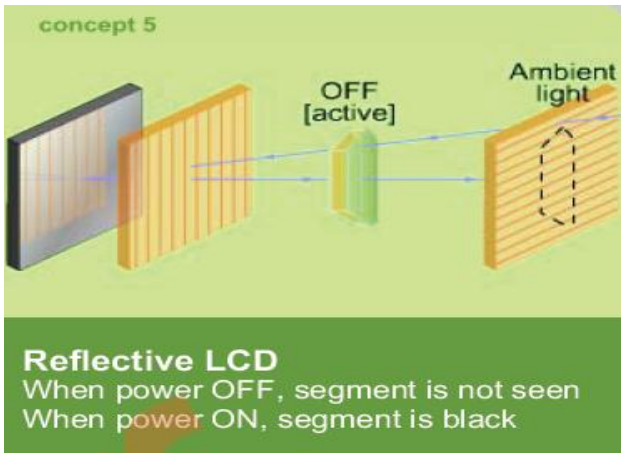
$$c = \frac{(100)(-35°)}{(0.12\ \text{m})(-62.5\ °\text{m}^2/\text{kg})}$$

$c = 470$ kg/m^3



34.23 - Physics at work: liquid crystal displays (LCDs) The liquid crystal display (LCD) in the watch face above demonstrates *variable optical activity* at work. LCDs are found in many common devices, including calculators, cellular telephones and clocks. There are two types of LCDs: backlit and reflective. The one shown above is a reflective LCD, but we will explore both types in this section. Backlit LCDs generate light behind their displays; reflective LCDs like the one above utilize ambient light.

LCDs rely on polarization. The characters of a digital watch display consist of "digit" segments: regions that can be made dark. These segments are filled with a substance called *liquid crystal* that is optically active



Gotchas

A light wave is a transverse wave. Yes. Both of its components, an electric and a magnetic field, oscillate perpendicularly to its direction of travel. *Radio signals and light waves are fundamentally different.* Both are forms of electromagnetic radiation, so we lean toward "no" in response to this statement. The wavelength and frequency of radio transmissions and light are significantly different, and humans can see light, but not radio waves, so one could say "yes". However, both are electromagnetic waves, and both move at the speed of light. *Intensity is the same thing as average energy density.* No, intensity represents the average rate of **power** transmission of an electromagnetic wave per unit **area**, perpendicular to the direction of propagation of the wave. Average energy density represents the average amount of **energy** contained in an electromagnetic wave per unit **volume**.

Equations

Proportionality of fields

$$\frac{E}{B} = c = \frac{1}{\sqrt{\mu_0 \varepsilon_0}}$$

Poynting vector

$$\mathbf{S} = \frac{1}{\mu_0} \mathbf{E} \times \mathbf{B}$$

Intensity of electromagnetic radiation

$$I = S_{\text{avg}} = \frac{E_{\text{max}}^2}{2\mu_0 c} = \frac{E_{\text{rms}}^2}{\mu_0 c}$$

$$I = \frac{P}{4\pi r^2}$$

Energy density

$$u_{\text{E}} = \frac{\varepsilon_0 E^2}{2}, \quad u_{\text{B}} = \frac{B^2}{2\mu_0}$$

$$u_{\text{E}} = u_{\text{B}}$$

$$u = u_{\text{E}} + u_{\text{B}} = 2u_{\text{E}} = 2u_{\text{B}}$$

$$u_{\text{avg}} = \frac{\varepsilon_0 E_{\text{max}}^2}{2}$$

Momentum transferred by radiation absorption

$$\Delta p = \frac{\Delta U}{c}$$ for a blackbody

Light can refract □ change direction □ as it moves from one medium to another. For instance, if you stand at the edge of a pool and try to poke something underwater with a stick, you may misjudge the object's location. This is because the light from the object changes direction as it passes from the water to the air. You perceive the object to be closer to the surface than it actually is because you subconsciously assume that light travels in a straight line. Although refraction can cause errors like this, it can also serve many useful purposes. Optical microscopes, eyeglass lenses, and indeed the lenses in your eyes all use refraction to bend and focus light, forming images and causing objects to appear a different size or crisper than they otherwise would. Where a lens focuses light, and whether it magnifies an object, is determined by both the curvature of the lens and the material of which it is made. Scientists have developed quantitative tools to determine the nature of the images created by a lens. We will explore these tools thoroughly later, "focusing" first, so to speak, on the principle of refraction underlying them. To begin your study of refraction, try the simulation to the right. Each of your helicopters can fire a laser □ a sharp beam of light □ at any of three submarines lurking under the sea. The submarines have lasers, too, and will shoot back at your craft. Your mission is to disable the submarines before they disarm your helicopters. When you make a hit, you can shoot again. Otherwise, the submarines get their turn to shoot until they miss. You play by dragging the aiming arrow underneath any one of your helicopters. Press FIRE and the laser beam will follow the direction of this arrow until it reaches the water, where refraction will cause the beam to change direction. In addition to hitting the submarines before they get you, you can conduct some basic experiments concerning the nature of refraction. As with reflection, the angle of incidence is measured from a line normal (perpendicular) to a surface. In this case, the surface is the horizontal boundary between the water and the air. Observe how the light bends at the boundary when you shoot straight down, at a zero angle of incidence, or grazing the water, at a large angle of incidence. You can create a large angle of incidence by having the far right helicopter, for example, aim at the submarine on the far left.

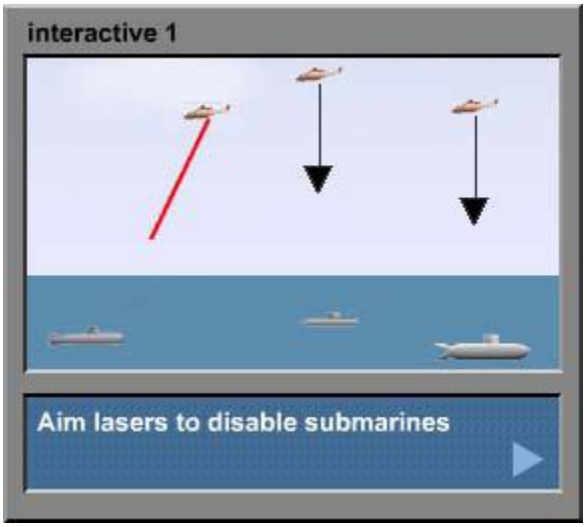
You can also observe how refraction differs when a laser beam passes from air to water (your lasers) and from water to air (the submarines' lasers). Observe the dashed normal line at each crossover point and answer the following question: Does the laser beam bend toward or away from that line as it changes media? You should notice that the laser beams of the submarines behave differently than those of the helicopters

when they change media. As a final aside: You may see that some of the laser beams of the submarines never leave the water, but reflect back from the surface between the water and the air. This is called total internal reflection.

Refraction: The change in the direction of light as it passes from one medium to another.

A material through which light travels is called a *medium* (plural: *media*). When light traveling in one medium encounters another medium, its direction can change. It can reflect back, as it would with a mirror. It can also pass into the second medium and change direction. This phenomenon, called refraction, is shown to the right. In the photo, a beam of light from a laser refracts (bends) as it passes from the air into the water. Light refracts when its speeds in the two media are different. Light travels faster through air than in water, and it changes direction as it moves from air into water, or from water into air. Although we are primarily interested in the refraction of light, all waves, including water waves, refract. Above, you see a photograph of surf wave fronts advancing parallel to a beach. Deep-ocean swells may approach a coastline from any angle, but they slow down as they encounter the shallows near the shore. The parts of a wave that encounter the shallow water earliest slow down first, and this causes the wave to refract. Sound waves can also refract. During a medical ultrasound scan, an acoustic lens can be used to focus the sound waves. The lens is made of a material in which sound travels faster than in water or body tissues.

The surface between two media, such as air and water, is called an *interface*. As with mirrors, light rays are often used to depict how light refracts when it meets an interface. Lasers are often used to demonstrate refraction because they can create thin beams of light that do not

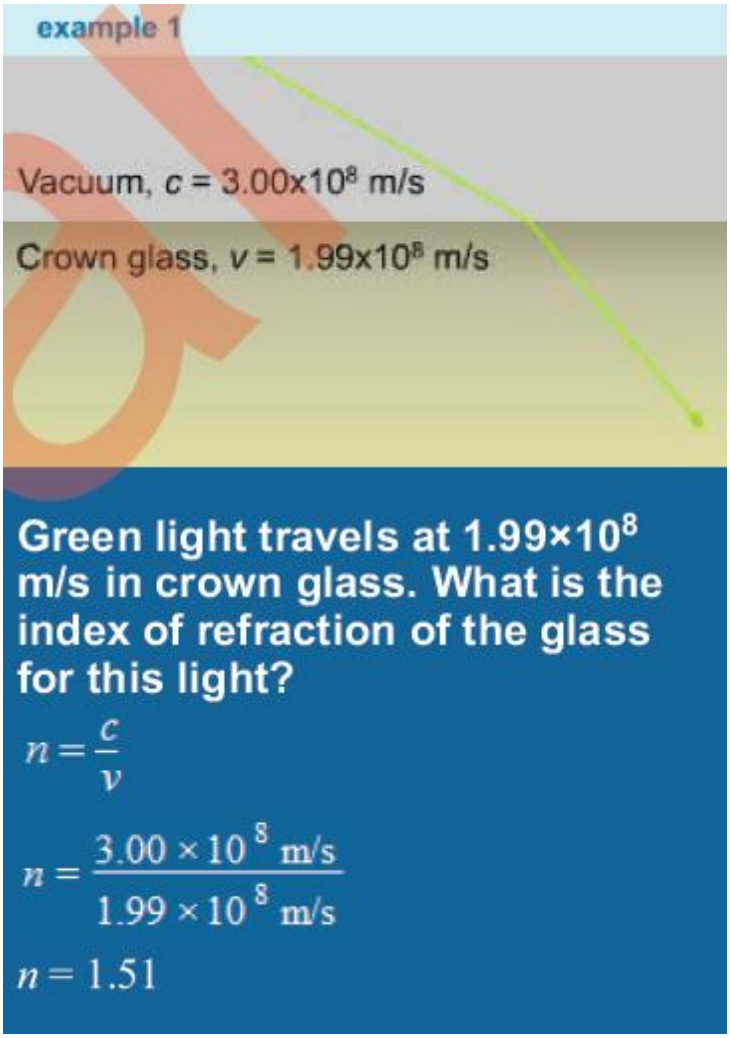


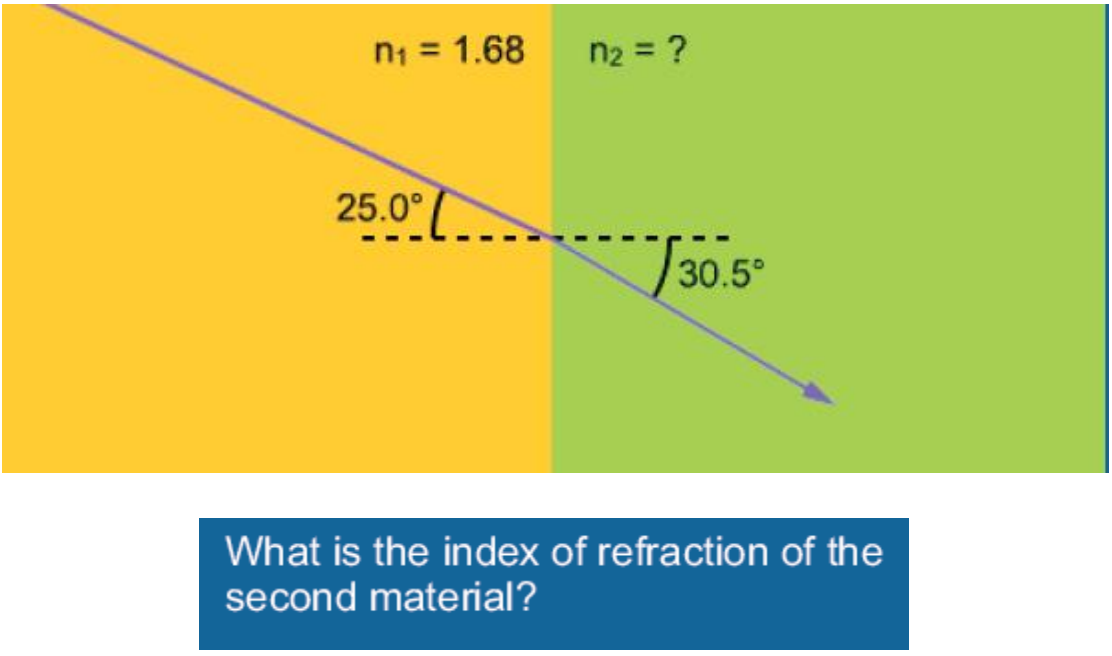
equation 2

	Index of refraction
Air	1.0003
Water	1.33
Vegetable oil	1.47
Crown glass	1.51
Salt	1.54
Flint glass	1.61
Corundum (ruby, sapphire)	1.77*
Diamond	2.42

At 20° C, λ = 589 nm *Approximate value

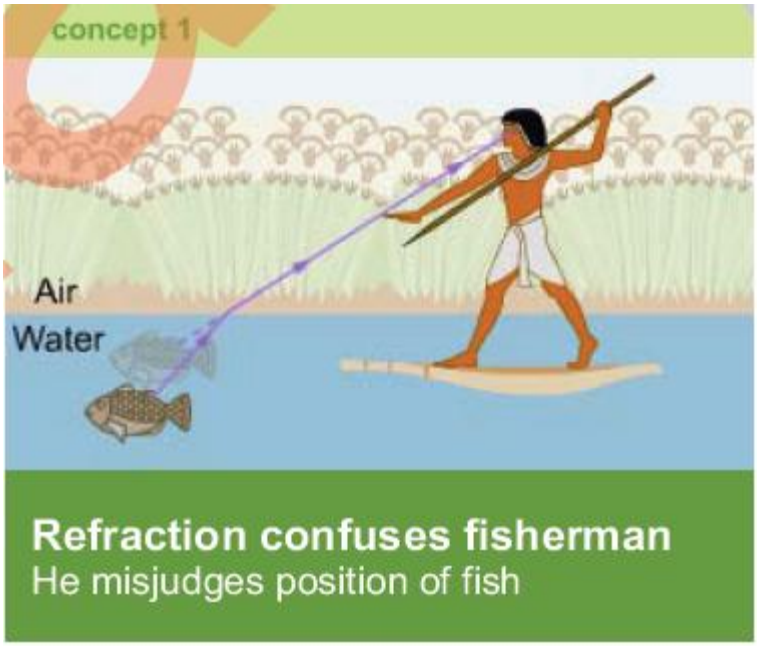
Indices of refraction

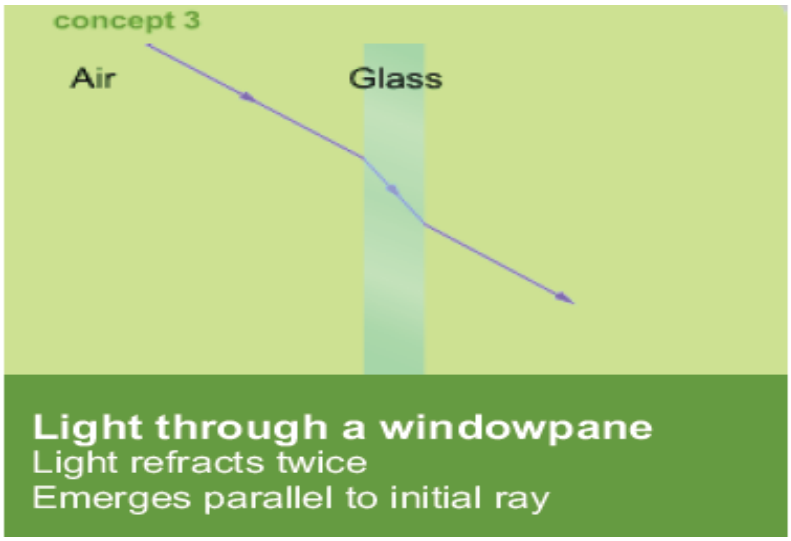




Everyday effects of refraction The refraction of light can cause interesting and sometimes confusing results as the brain interprets the position of objects it sees via

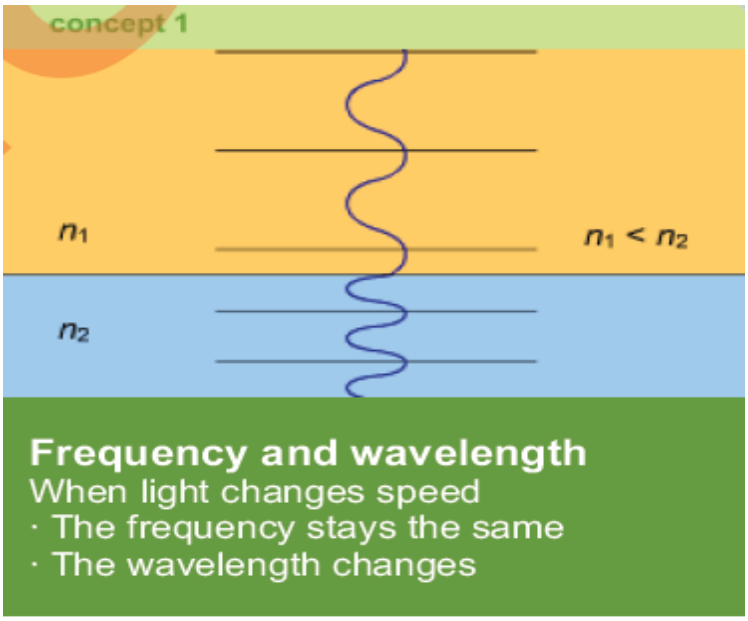
refracted light. For instance, in the upper illustration to the right, you see an ancient Egyptian fisherman trying to spear a fish. The solid line indicates the refracted path of the light traveling from the fish to his eyes. Since his brain expects light to travel in a straight line, he projects the fish to be in the position indicated at the end of the dashed line, which causes him to think the fish is nearer to the surface than it is. Experienced spear fishermen know how to compensate for this effect. The fisherman is not the only one to experience the effects of refraction. The fish does, too. The light from the world above seen by the fish also refracts. Light coming straight down will pass through the water's surface unchanged, but light at any angle will refract, in the process giving the fish a wider field of view of the world above than it would have if there were no refraction. In essence, the fish sees a compressed wide-angle view of the scene above. Certain camera lenses, appropriately called fisheye lenses, can create the same effect, as illustrated in Concept 2. Another interesting consequence of Snell's law concerns light passing through a window. The light refracts as it travels from air to glass. After passing through the glass it refracts again at the second interface. The ratio of the indices of refraction is now reversed, so the initial change in angle is cancelled out. The light ray that emerges is parallel to the initial ray, but displaced a small amount. The amount of displacement is small enough that we ordinarily do not notice "window shift." However, if you place a newspaper page on a tabletop and cover half of it with a flat pane of glass, you will be able to observe the displacement effect.

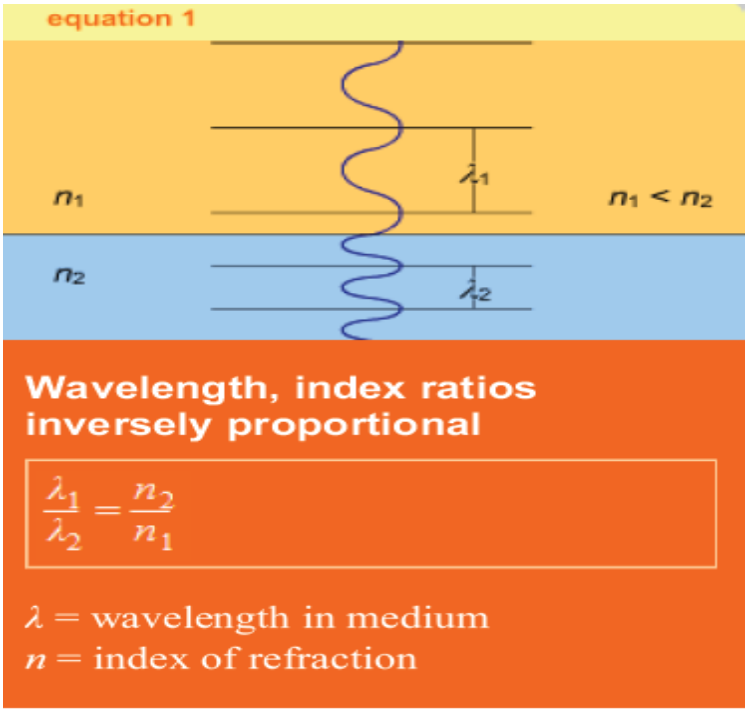




Wavelength of light in different media When light changes speed as it moves from one medium to another, its frequency stays the same but its wavelength changes. The

ratio of its wavelengths in the two media is the inverse of the ratio of the indices of refraction. We show this as an equation to the right and derive it below. Before deriving the equation, let's consider why the frequency stays the same, since this is an essential part of the derivation. The frequencies in the media must be the same, because if they were not, waves would either pile up at the interface or be destroyed. Neither occurs. You can witness this at the beach, where wave speed and wavelength may change as waves approach the beach, but the frequency of the waves does not change.





Variables

In this derivation, *c* represents the speed of light in a vacuum. For the other two media we define the variables in the following table:

	medium 1	medium 2
light speed	v_1	v_2
frequency	f_1	f_2
wavelength	λ_1	λ_2
index of refraction	n_1	n_2

Strategy

1. Use the equality of frequencies in the two media together with the wave speed equation to obtain a proportionality of the light speeds and wavelengths in the media.

2. Use the definition of the index of refraction to convert the previous proportion to one involving wavelengths and indices of refraction.

Physics principles and equations

The wave speed equation states that for any wave, the speed is the product of the wavelength and the frequency:

$$v = \lambda f$$

As a wave passes from one medium to another, its speed and wavelength may change, but its frequency must remain the same.

The definition of the index of refraction of a medium is

$$n = \frac{c}{v}$$

Step-by-step derivation

First, we explain the diagram you see above. The purple line is a light ray refracting at an interface. In the diagram, light travels more slowly in the lower medium than the upper. This could represent, for example, light passing from air into water. The gray lines perpendicular to the ray represent wave fronts. You see the wavelength labeled as □ (□i in the upper medium, □r in the lower medium). There are two right triangles in the diagram that share the hypotenuse labeled *x*. The bright yellow triangle shows

elements of a wave front that has not yet entered the lower medium. The dark orange triangle shows elements of a wave front that is now traveling in the lower medium. The angles of incidence and refraction □i and □r are also shown in the diagram. Because the wave fronts are perpendicular to the light rays, we can identify angles in each of the triangles that are equal to □i and □r. These base angles are shown in the diagram.

Variables

In this derivation, *x* represents the common hypotenuse of the two triangles in the diagram. For the incident and refractive media we define the variables in the following table.

	incident medium	refractive medium
angle	θ_i	θ_r
wavelength	λ_i	λ_r
index of refraction	n_i	n_r

Strategy

1. Consider the two triangles in the diagram. State the sines of their base angles □i and □r as trigonometric ratios of the triangles' sides.

2. Construct the ratio sin □i / sin □r. The common hypotenuse *x* will cancel out, leaving a ratio of wavelengths.

3. Restate the ratio of wavelengths as a ratio of indices of refraction to obtain Snell's law.

Physics principles and equations

The ratio of the wavelengths is inversely proportional to the ratio of the indices of refraction.

$$\frac{\lambda_i}{\lambda_r} = \frac{n_r}{n_i}$$

Step-by-step derivation

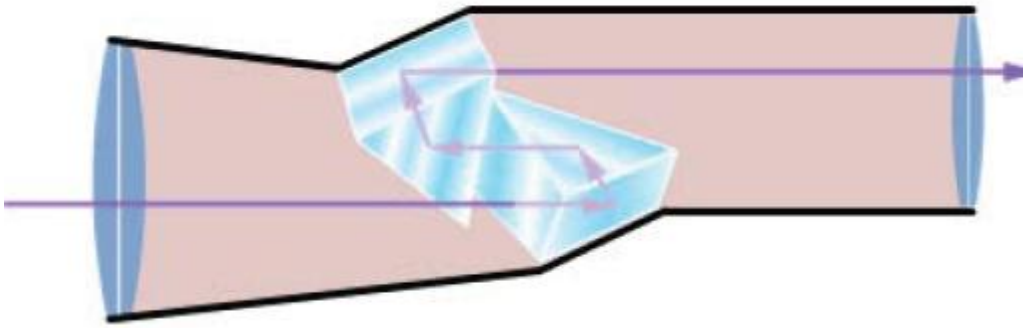
We construct the fraction sin □i / sin □r, and calculate the sines as the ratios of the sides of triangles. This leads to a ratio of wavelengths that can be replaced by a ratio of indices of refraction, yielding Snell's law.

Step		Reason
1.	$\sin\theta_i = \lambda_i / x,\ \sin\theta_r = \lambda_r / x$	definition of sine
2.	$\frac{\sin\theta_i}{\sin\theta_r} = \frac{\lambda_i / x}{\lambda_r / x}$	ratio using definition of sine
3.	$\frac{\sin\theta_i}{\sin\theta_r} = \frac{\lambda_i}{\lambda_r}$	simplify
4.	$\frac{\lambda_i}{\lambda_r} = \frac{n_r}{n_i}$	change of wavelength
5.	$\frac{\sin\theta_i}{\sin\theta_r} = \frac{n_r}{n_i}$	substitute equation 4 into equation 3

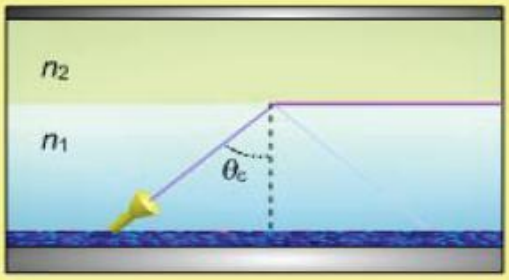
Interactive problem: helicopter and submarines

The simulation on the right is similar to the one in this chapter's introduction. As before, submarines lurk under the waves, but now you have only a single helicopter.

As before, you have a laser you can aim in an attempt to disable two submarines before they disable you. You keep shooting as long as you keep making hits. To shoot your laser, aim it by dragging the aiming arrow and fire it by clicking on FIRE. The angle of incidence is shown in an output gauge. As soon as you miss, it is the submarines' turn. Warning: The computer has been set to be far more accurate in this game than in the introductory one. Unless you are very precise with your shots, it is unlikely you will win. However, there is good news: Now you have more intellectual firepower because you have the aid of Snell's law. You are also given some assistance from an able comrade; she has computed the angles of refraction required for your laser to reach the submarines, as shown in the diagram. You should use 1.33 for the index of refraction of water, and 1.00 for the index of refraction of air. If you correctly set the angle of incidence when you aim each of your shots, you can make two straight hits and disable the submarines before they disable you.



equation 1

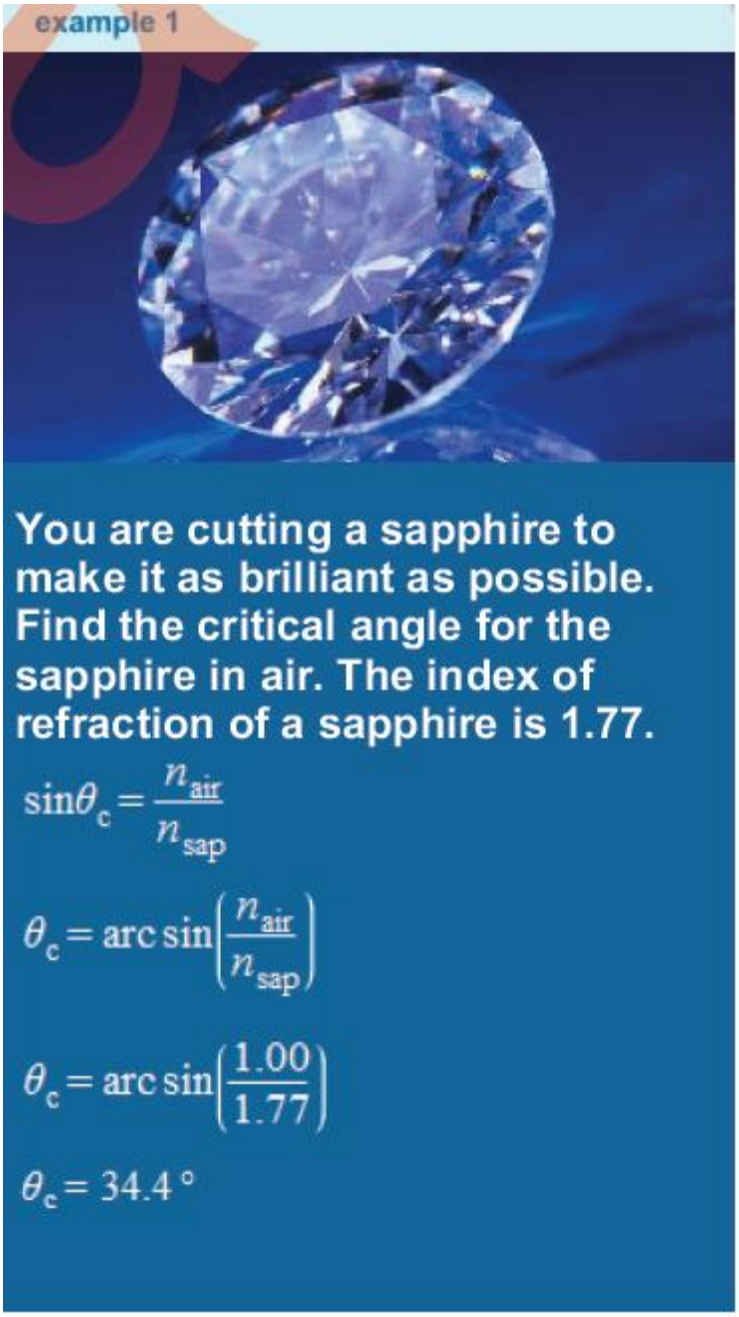


Critical angle

$$\sin\theta_c = \frac{n_2}{n_1}$$

θ_c = critical angle

n_1, n_2 = indices of refraction ($n_1 > n_2$)



Fiber optic cable is used to transfer information (in glass or plastic) using total internal reflection. Data (such as speech) is first encoded as modulations of a beam of laser light. The light remains inside the transparent cable as it travels due to its total internal reflection off the inner surfaces of the cable walls. In this way, light is transmitted through the cable with little loss. Using light instead of electricity to transfer information has several benefits. Light can be used to encode much more information than electric oscillations because of its extremely high frequency (more than 1014 Hz for red light), and fiber optic cable is immune to interference problems from nearby electrical applications

Step-by-step solution

We start by using trigonometry to establish a relationship between sin □ and sin □c. In the following steps we use Snell's law and the definition of the critical angle to find sin □ in terms of indices of refraction. Then we evaluate to find a value for the acceptance angle □.

Step	Reason
1. $\cos\beta = \sin\theta_c$	trigonometric identity
2. $\sin\beta = \sqrt{1-\cos^2\beta}$	trigonometric identity
3. $\sin\beta = \sqrt{1-\sin^2\theta_c}$	substitute equation 1 into equation 2

Step	Reason
4. $n_0 \sin\alpha = n_1 \sin\beta$	Snell's law
5. $\sin\alpha = \frac{n_1}{n_0}\sqrt{1-\sin^2\theta_c}$	substitute equation 3 into equation 4
6. $\sin\theta_c = \frac{n_2}{n_1}$	definition of critical angle
7. $\sin\alpha = \frac{n_1}{n_0}\sqrt{1-\frac{n_2^2}{n_1^1}}$	substitute equation 6 into equation 5
8. $\sin\alpha = \frac{1.60}{1.00}\sqrt{1-\frac{(1.50)^2}{(1.60)^2}}$ $\sin\alpha = 0.557$ $\alpha = 33.8°$	evaluate

Interactive problem: laser target pistol

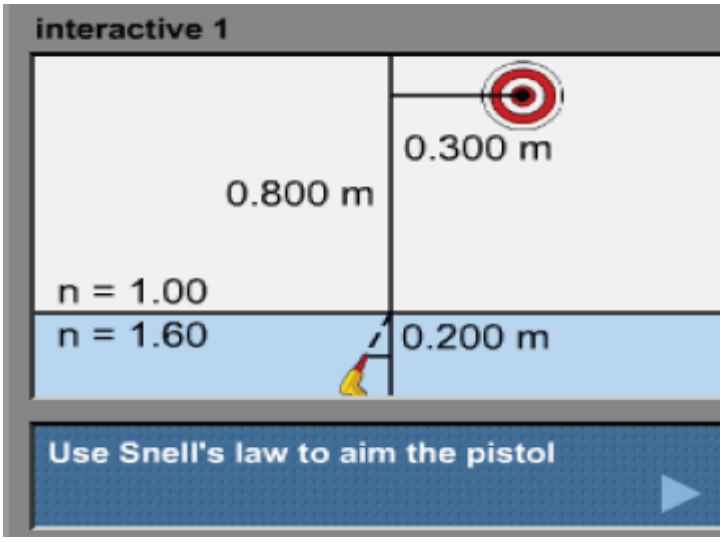
In this simulation, you fire a laser pistol at two stationary targets. You want a bullseye, a shot that passes through the central black circle of the target. You happen

to be immersed in a fluid with an index of refraction of 1.60. When you shoot the laser from within the fluid, you have to account for the refraction of your laser beam as it passes from the fluid into the air. Some portion of the beam also reflects from the interface. If the incident angle is equal to or greater than the critical angle, all of the beam will reflect.

The laser pistol is designed so its beam always hits the interface between the fluid and the air at the same point. You drag the pistol from side to side to determine its angle of incidence. An output gauge tells you the horizontal distance between the pistol and where its beam passes through the interface. You are also told how far the pistol is below the surface, a value that stays constant.

There are two targets to hit: one above the surface of the fluid and the other below the surface. You will have to use a different strategy to hit each target. For each target you know its horizontal distance from where the laser passes through the interface, and its height above the fluid. You must strike each target in its center to succeed.

You have just read a fair number of facts. It may be useful now to look at the diagram and the simulation. You will probably find the lower target easier to hit, if you just reflect a bit. You can then bend your mind toward determining how to hit the upper target. A good way to start is to apply trigonometry to calculate the necessary angle of refraction. Then, using Snell's law and some more trigonometry, you can determine where to position the pistol. When you have calculated the correct horizontal distance, drag the pistol and press FIRE to test your answer.



Light is a particle.

Many of the great scientists of the 17th and 18th centuries who made fundamental contributions to the study of optics, including Isaac Newton, thought that light consisted of a stream of "corpuscles," or particles. In the 20th century, Albert Einstein explained the photoelectric effect. His explanation, for which he was awarded the 1921 Nobel Prize, depended on the fact that light acts like a particle. This property of light led to the coining of the term "photon" for a single particle of light by the chemist Gilbert Lewis.

Light is a wave. Between the 18th and 20th centuries, physicists discovered many wave-like properties of light. They found that a number of phenomena they routinely observed with water waves they could also observe with light. For instance, the English scientist Thomas Young (1773-1829) showed that light could produce the same kinds of interference patterns that water waves produce. At the right, you see examples of interference patterns formed by light and by water waves. The similarities are striking. In this chapter, you will apply to light some of what you have studied about the interference of sound waves and traveling waves in strings. **Let there be light**. Is light a particle, a wave, or both? Perhaps an Early Authority had it right. Light is light. It is a combination of electric and magnetic fields. Trying to classify light as a particle or as a wave may be a fruitless effort better to revel in its unique properties. In this chapter, we will revel in its wave-like properties, and discuss the topic of interference. Your prior study of electromagnetic radiation modeled as a wave phenomenon will prove useful.