



Principles of
**PLANETARY
CLIMATE**

RAYMOND T. PIERREHUMBERT

CAMBRIDGE

CAMBRIDGE

www.cambridge.org/9780521865562

the infrared as well as in the visible, and in consequence emit radiation at a much lower rate than expected from the blackbody formula. (They would make fine windows for creatures having infrared vision.) There is, in fact, a deep and important relation between absorption and emission of radiation, which will be discussed in Section 3.5.

3.3 RADIATION BALANCE OF PLANETS

As a first step in our study of the temperature of planets, let's consider the following idealized case:

- The only source of energy heating the planet is absorption of light from the planet's host star.
- The *planetary albedo*, or proportion of sunlight reflected by the planet as a whole including its atmosphere, is spatially uniform.
- The planet is spherical, and has a distinct solid or liquid surface which radiates like a perfect blackbody.
- The planet's temperature is uniform over its entire surface.
- The planet's atmosphere is perfectly transparent to the electromagnetic energy emitted by the surface.

The uniform temperature assumption presumes that the planet has an atmosphere or ocean which is so well stirred that it is able to mix heat rapidly from one place to another, smoothing out the effects of geographical fluctuations in the energy balance. The Earth conforms fairly well to this approximation. The equatorial annual mean temperature is only 4% above the global mean temperature of 286 K, while the north polar temperature is only 10% below the mean. The most extreme deviation occurs on the high Antarctic plateau, where the annual mean south polar temperature is 21% below the global mean. The surface temperature of Venus is even more uniform than that of Earth. That of Mars, which in our era has a thin atmosphere and no ocean, is less uniform. Airless, rocky bodies like the Moon and Mercury do not conform at all well to the uniform temperature approximation.

Light leaving the upper layers of the Sun and most other stars takes the form of blackbody radiation. It is isotropic, and its flux and flux spectrum conform to the blackbody law corresponding to the temperature of the photosphere, from which the light escapes. Once the light leaves the surface of the star, however, it expands through space and does not interact significantly with matter except where it is intercepted by a planet. Therefore, it is no longer blackbody radiation, though it retains the blackbody spectrum. In the typical case of interest, the planet orbits its star at a distance that is much greater than the radius of the star, and itself has a radius that is considerably smaller than the star and is hence yet smaller than the orbital distance. In this circumstance, all the rays of light which intersect the planet are very nearly parallel to the line joining the center of the planet to the center of its star; the sunlight comes in as a nearly *parallel beam*, rather than being isotropic, as would be the case for true blackbody radiation. The parallel-beam approximation is equivalent to saying that, as seen from the planet, the Sun occupies only a small portion of the sky, and as seen from the Sun the planet also occupies only a small portion of the sky. Even for Mercury, with a mean orbital distance of 58 000 000 km, the Sun (whose radius is 695 000 km) occupies an angular width in the sky of only about $2 \cdot (695\,000/58\,000\,000)$ radians, or 1.4° .

The solar flux impinging on the planet is also reduced, as compared with the solar flux leaving the photosphere of the star. The total energy per unit frequency leaving the star is $4\pi r_\odot^2(\pi B(\nu, T_\odot))$, where r_\odot is the radius of the star and T_\odot is the temperature of its photosphere. At a distance r from the star, the energy has spread uniformly over a sphere whose surface area is $4\pi r^2$; hence at this distance, the energy flux per unit frequency is $\pi B r_\odot^2 / r^2$, and the total flux is $\sigma T_\odot^4 r_\odot^2 / r^2$. The latter is the flux seen by a planet at orbital distance r , in the form of a beam of parallel rays. It is known as the solar “constant,” and will be denoted by L_\odot , or sometimes simply L where there is no risk of confusion with latent heat. The solar (or stellar) “constant” depends on a planet’s distance from its star, but the *luminosity* of the star is an intrinsic property of the star at any given stage of the star’s life. The stellar luminosity is the net power output of a star, and if the star’s emission can be represented as blackbody radiation, the luminosity is given by $\mathcal{L}_\odot = 4\pi r_\odot^2 \sigma T_\odot^4$. In this book the terms “solar constant,” “stellar constant,” “solar flux,” and “stellar flux” are all to be considered synonymous, and to refer to the parallel-beam flux from the star measured at any given position in a planet’s orbit. Most of the time, the terms will be used to refer to the average flux over the planet’s year, though in Chapter 7 we will be concerned with the seasonal variations of the fluxes for planets in non-circular orbits. The term “solar constant” as used here must be distinguished from the number L_\odot which sometimes goes by the same name, and which is an actual constant approximately equal to the Sun’s flux at the Earth’s mean orbit, averaged over the 11-year solar cycle.

We are now equipped to compute the energy balance of the planet, subject to the preceding simplifying assumptions. Let a be the planet’s radius. Since the cross-sectional area of the planet is πa^2 and the solar radiation arrives in the form of a nearly parallel beam with flux L_\odot , the energy per unit time impinging on the planet’s surface is $\pi a^2 L_\odot$; the rate of energy absorption is $(1 - \alpha)\pi a^2 L_\odot$, where α is the albedo. The planet loses energy by radiating from its entire surface, which has area $4\pi a^2$. Hence the rate of energy loss is $4\pi a^2 \sigma T^4$, where T is the temperature of the planet’s surface. In equilibrium the rate of energy loss and gain must be equal. After cancelling a few terms, this yields

$$\sigma T^4 = \frac{1}{4}(1 - \alpha)L_\odot. \quad (3.6)$$

Note that this is independent of the radius of the planet. The factor $\frac{1}{4}$ comes from the ratio of the planet’s cross-sectional area to its surface area, and reflects the fact that the planet intercepts only a disk of the incident solar beam, but radiates over its entire spherical surface. This equation can be readily solved for T . If we substitute for L_\odot in terms of the photospheric temperature, the result is

$$T = \frac{1}{\sqrt{2}}(1 - \alpha)^{1/4} \sqrt{\frac{r_\odot}{r}} T_\odot. \quad (3.7)$$

Formula 3.7 shows that the blackbody temperature of a planet is much less than that of the photosphere, so long as the orbital distance is large compared with the stellar radius. From the displacement law, it follows that the planet loses energy through emission at a distinctly lower wavenumber than that at which it receives energy from its star. This situation is illustrated in Fig. 3.3. For example, the energy received from our Sun has a median wavenumber of about $15\,000\text{ cm}^{-1}$, equivalent to a wavelength of about $0.7\ \mu\text{m}$. An isothermal planet at Mercury’s orbit would radiate to space with a median emission

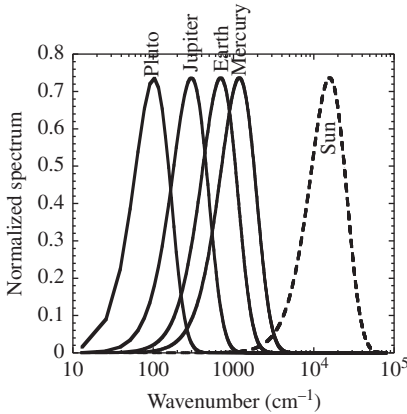


Figure 3.3 The Planck density of radiation emitted by the Sun and selected planets in radiative equilibrium with absorbed solar radiation (based on the observed short-wave albedo of the planets). The Planck densities are transformed to a logarithmic spectral coordinate, and all are normalized to unit total emission.

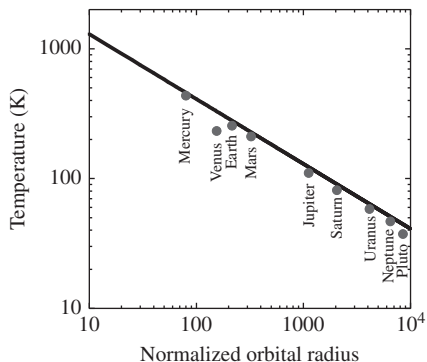
wavenumber of 1100 cm⁻¹, corresponding to a wavelength of 9 μm. An isothermal planet at the orbit of Mars would radiate with a median wavenumber of 550 cm⁻¹, corresponding to a wavelength of 18 μm.

Exercise 3.3 A planet with zero albedo is in orbit around an exotic hot star having a photospheric temperature of 100 000 K. The ratio of the planet’s orbit to the radius of the star is the same as for Earth (about 215). What is the median emission wavenumber of the star? In what part of the electromagnetic spectrum does this lie? What is the temperature of the planet? In what part of the electromagnetic spectrum does the planet radiate? Do the same if the planet is instead in orbit around a brown dwarf star with a photospheric temperature of 600 K.

The separation between absorption and emission wavenumber will prove very important when we bring a radiatively active atmosphere into the picture, since it allows the atmosphere to have a different effect on incoming vs. outgoing radiation. Since the outgoing radiation has longer wavelength than the incoming radiation, the flux of emitted outgoing radiation is often referred to as *outgoing longwave radiation*, and denoted by *OLR*. For a non-isothermal planet, the *OLR* is a function of position (e.g. latitude and longitude on an imaginary sphere tightly enclosing the planet and its atmosphere). We will also use the term to refer to the outgoing flux averaged over the surface of the sphere, even when the planet is not isothermal. As for the other major term in the planet’s energy budget, we will refer to the electromagnetic energy received from the planet’s star as the *shortwave, solar, or stellar flux*. Most stars, our Sun not excepted, have their primary output in the visible, ultraviolet, and near-infrared part of the spectrum, all of which are shorter in wavelength than the thermal infrared – the *OLR* – by which planets lose energy to space. (See Fig. 3.1 for the definitions of ultraviolet, near infrared, and thermal infrared.)

Formula 3.7 is plotted in Fig. 3.4 for a hypothetical isothermal planet with zero albedo. Because of the square-root dependence on orbital distance, the temperature varies only weakly with distance, except very near the star. Neglecting albedo and atmospheric effects, Earth would have a mean surface temperature of about 280K. Venus would be only 50K warmer than the Earth and Mars only 53 K colder. At the distant orbit of Jupiter, the black-body equilibrium temperature falls to 122K, but even at the vastly more distant orbit of Neptune the temperature is still as high as 50K. The emission from all of these planets lies

Figure 3.4 The equilibrium blackbody temperature of an isothermal spherical zero-albedo planet, as a function of distance from a Sun having a photospheric temperature of 5800 K. The orbital distance is normalized by the radius of the Sun. Dots show the equilibrium blackbody temperature of the Solar System planets, based on their actual observed albedos.



in the infrared range, though the colder planets radiate in the deeper (lower wavenumber) infrared. An exception to the strong separation between stellar and planetary temperature is provided by the “roasters” – a recently discovered class of extrasolar giant planets with $\frac{r}{r_{\odot}}$ as low as 5. Such planets can have equilibrium blackbody temperatures as much as a third that of the photosphere of the parent star. For these planets, the distinction between the behavior of incoming and outgoing radiation is less sharp.

It is instructive to compare the ideal blackbody temperature with observed surface temperature for the three Solar System bodies which have both a distinct surface and a thick enough atmosphere to enforce a roughly uniform surface temperature: Venus, Earth, and Saturn’s moon Titan. For this comparison, we calculate the blackbody temperature using the observed planetary albedos, instead of assuming a hypothetical zero albedo planet as in Fig. 3.4. Venus is covered by thick, highly reflective clouds, which raise its albedo to 0.75. The corresponding isothermal blackbody temperature is only 232 K (as compared with 330 K in the zero albedo case). This is far less than the observed surface temperature of 740 K. Clearly, the atmosphere of Venus exerts a profound warming effect on the surface. The warming arises from the influence of the atmosphere on the infrared emission of the planet, which we have not yet taken into account. Earth’s albedo is on the order of 0.3, leading to a blackbody temperature of 255 K. The observed mean surface temperature is about 285 K. Earth’s atmosphere has a considerably weaker warming effect than that of Venus, but it is nonetheless a very important warming, since it brings the planet from subfreezing temperatures where the oceans would almost certainly become ice-covered, to temperatures where liquid water can exist over most of the planet. The albedo of Titan is 0.21, and using the solar constant at Saturn’s orbit we find a blackbody temperature of 85 K. The observed surface temperature is about 95 K, whence we conclude that the infrared effects of Titan’s atmosphere moderately warm the surface.

The way energy balance determines surface temperature is illustrated graphically in Fig. 3.5. One first determines the way in which the mean infrared emission per unit area depends on the mean surface temperature T_s ; for the isothermal blackbody calculation, this curve is simply σT_s^4 . The equilibrium temperature is determined by the point at which the *OLR* curve intersects the curve giving the absorbed solar radiation (a horizontal line in the present calculation). In some sense, the whole subject of climate comes down to an ever-more sophisticated hierarchy of calculations of the curve $OLR(T_s)$; our attention will soon turn to the task of determining how the *OLR* curve is affected by an atmosphere. With increasing sophistication, we will also allow the solar absorption to vary with T_s , owing to changing clouds, ice cover, vegetation cover, and other characteristics.

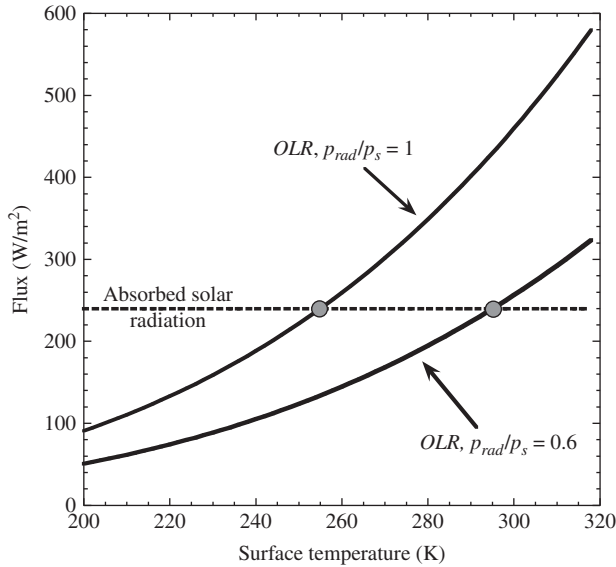


Figure 3.5 Determination of a planet’s temperature by balancing absorbed solar energy against emitted longwave radiation. The horizontal line gives the absorbed solar energy per unit surface area, based on an albedo of 0.3 and a solar constant of 1370 W/m^2 . The OLR is given as a function of surface temperature. The upper curve assumes the atmosphere has no greenhouse effect ($p_{rad} = p_s$), while the lower OLR curve assumes $p_{rad}/p_s = 0.6$, a value appropriate to the present Earth.

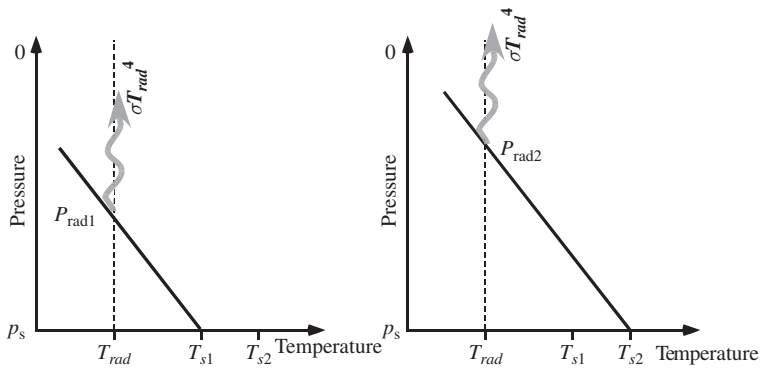


Figure 3.6 Sketch illustrating how the greenhouse effect increases the surface temperature. In equilibrium, the outgoing radiation must remain equal to the absorbed solar radiation, so T_{rad} stays constant. However, as more greenhouse gas is added to the atmosphere, p_{rad} is reduced, so one must extrapolate temperature further along the adiabat to reach the surface.

We will now consider an idealized thought experiment which illustrates the essence of the way an atmosphere affects OLR . Suppose that the atmosphere has a temperature profile $T(p)$ which decreases with altitude, according to the dry or moist adiabat. Let p_s be the surface pressure, and suppose that the ground is strongly thermally coupled to the

atmosphere by turbulent heat exchanges, so that the ground temperature cannot deviate much from that of the immediately overlying air. Thus, $T_s = T(p_s)$. If the atmosphere were transparent to infrared, as is very nearly the case for nitrogen or oxygen, the *OLR* would be σT_s^4 . Now, let's stir an additional gas into the atmosphere, and assume that it is well mixed with uniform mass concentration q . This gas is transparent to solar radiation but interacts strongly enough with infrared that when a sufficient amount is mixed into a parcel of air, it turns that parcel into an ideal blackbody. Such a gas, which is fairly transparent to the incoming shortwave stellar radiation but which interacts strongly with the outgoing (generally infrared) emitted radiation, is called a *greenhouse gas*, and the corresponding effect on planetary temperature is called the *greenhouse effect*. Carbon dioxide, water vapor, and methane are some examples of greenhouse gases, and the molecular properties that make a substance a good greenhouse gas will be discussed in Chapter 4. The mass of greenhouse gas that must be mixed into a column of atmosphere with base of 1 m^2 in order to make that column act begin to act like a blackbody is characterized by the *absorption coefficient* κ , whose units are m^2/kg . Here we'll assume κ to be independent of frequency, temperature, and pressure, though for real greenhouse gases, κ depends on all of these. Since the mass of greenhouse gas in a column of thickness Δp in pressure coordinates is $q\Delta p/g$, then the definition of κ implies that the slab acts like a blackbody when $\kappa q\Delta p/g > 1$. When $\kappa q p_s/g < 1$ then the entire mass of the atmosphere is not sufficient to act like a blackbody and the atmosphere is said to be *optically thin*. For optically thin atmospheres, infrared radiation can escape from the surface directly to space, and is only mildly attenuated by atmospheric absorption. When $\kappa q p_s/g \gg 1$, the atmosphere is said to be *optically thick*.

If the atmosphere is optically thick, we can slice the atmosphere up into a stack of slabs with thickness Δp_1 such that $\kappa q\Delta p_1/g = 1$. Each of these slabs radiates like an ideal blackbody with temperature approximately equal to the mean temperature of the slab. Recall, however, that another fundamental property of blackbodies is that they are perfect *absorbers* (though if they are only blackbodies in the infrared, they will only be perfect absorbers in the infrared). Hence *infrared radiation escapes to space only from the topmost slab*. The *OLR* will be determined by the temperature of this slab alone, and will be insensitive to the temperature of lower portions of the atmosphere. The pressure at the bottom of the topmost slab is Δp_1 . We can thus identify Δp_1 as the characteristic pressure level from which radiation escapes to space, which therefore will be called p_{rad} in subsequent discussions. The radiation escaping to space - the *OLR* - will then be approximately $\sigma T(p_{rad})^4$. Because temperature decreases with altitude on the adiabat the *OLR* is less than σT_s^4 to the extent that $p_{rad} < p_s$. As shown in Fig. 3.6, a greenhouse gas acts like an insulating blanket, reducing the rate of energy loss to space at any given surface temperature. All other things being equal the equilibrium surface temperature of a planet with a greenhouse gas in its atmosphere must be greater than that of a planet without a greenhouse gas, in order to radiate away energy at a sufficient rate to balance the absorbed solar radiation. The key insight to be taken from this discussion is that *the greenhouse effect only works to the extent that the atmosphere is colder at the radiating level than it is at the ground*.

For real greenhouse gases, the absorption coefficient varies greatly with frequency. Such gases act on the *OLR* by making the atmosphere very optically thick at some frequencies, less optically thick at others, and perhaps even optically thin at still other frequencies. In portions of the spectrum where the atmosphere is more optically thick, the emission to space originates in higher (and generally colder) parts of the atmosphere. In reality, then, the infrared escaping to space is a blend of radiation emitted from a range of atmospheric levels, with some admixture of radiation from the planet's surface as well. The concept

of an *effective radiating level* nonetheless has merit for real greenhouse gases. It does not represent a distinct physical layer of the atmosphere, but rather characterizes the mean depth from which infrared photons escape to space. As more greenhouse gas is added to an atmosphere, more of the lower parts of the atmosphere become opaque to infrared, preventing the escape of infrared radiation from those regions. This increases the altitude of the effective radiating level (i.e. decreases p_{rad}). Some of the implications of a frequency-dependent absorption coefficient are explored in Problem 3.29, and the subject will be taken up at great length in Chapter 4.

From an observation of the actual *OLR* emitted by a planet, one can determine an equivalent blackbody radiating temperature T_{rad} from the expression $\sigma T_{rad}^4 = OLR$. This temperature is the infrared equivalent of the Sun’s photospheric temperature; it is a kind of mean temperature of the regions from which infrared photons escape, and p_{rad} represents a mean pressure of these layers. For planets for which absorbed solar radiation is the only significant energy source, T_{rad} is equal to the ideal blackbody temperature given by Eq. (3.7). The arduous task of relating the effective radiating level to specified concentrations of real greenhouse gases is treated in Chapter 4.

Figure 3.7 illustrates the reduction of infrared emission caused by the Earth’s atmosphere. At every latitude, the observed *OLR* is much less than it would be if the planet radiated to space at its observed surface temperature. At the Equator the observed *OLR* is 238 W/m^2 , corresponding to a radiating temperature of 255 K. This is much less than the observed surface temperature of 298 K, which would radiate at a rate of 446 W/m^2 if the atmosphere did not intervene. It is interesting that the gap between observed *OLR* and the computed surface emission is less in the cold polar regions, and especially small at the winter pole. This happens partly because, at low temperatures, there is simply less infrared emission for the atmosphere to trap. However, differences in the water content of the atmosphere, and differences in the temperature profile, can also play a role. These effects will be explored in Chapter 4.

Gases are not the only atmospheric constituents which affect *OLR*. Clouds consist of particles of condensed substance small enough to stay suspended for a long time. They can profoundly influence *OLR*. Gram for gram, condensed water interacts much more strongly with infrared than does water vapor. In fact, a mere 20 grams of water in the form of liquid

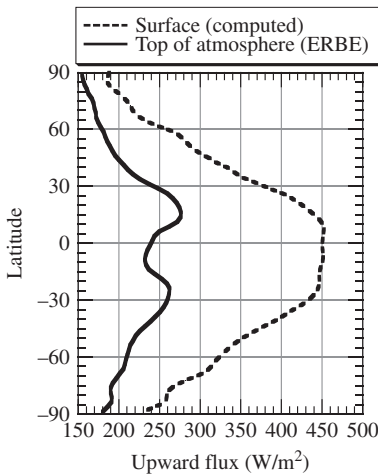


Figure 3.7 The Earth’s observed zonal-mean *OLR* for January, 1986. The observations were taken by satellite instruments during the Earth Radiation Budget Experiment (ERBE), and are averaged along latitude circles. The figure also shows the radiation that would be emitted to space by the surface (σT_s^4) if the atmosphere were transparent to infrared radiation.

droplets of a typical size is sufficient to turn a column of air 500 m thick by 1 m square into a very nearly ideal blackbody. To a much greater extent than for greenhouse gases, a water cloud layer in an otherwise infrared-transparent atmosphere really can be thought of as a discrete radiating layer. The prevalence of clouds in the high, cold regions of the tropical atmosphere accounts for the dip in *OLR* near the Equator, seen in Fig. 3.7. Clouds are unlike greenhouse gases, though, since they also strongly reflect the incoming solar radiation. It's the tendency of these two large effects to partly cancel that makes the problem of the influence of clouds on climate so challenging. Not all condensed substances absorb infrared as well as water does. Liquid methane (important on Titan) and CO₂ ice (important on present and Early Mars) are comparatively poor infrared absorbers. They affect *OLR* in a fundamentally different way, through reflection instead of absorption and emission. This will be discussed in Chapter 5.

In a nutshell, then, here is how the greenhouse effect works. From the requirement of energy balance, the absorbed solar radiation determines the effective blackbody radiating temperature T_{rad} . This is not the surface temperature; it is instead the temperature encountered at some pressure level in the atmosphere p_{rad} , which characterizes the infrared opacity of the atmosphere, specifically the typical altitude from which infrared photons escape to space. The pressure p_{rad} is determined by the greenhouse gas concentration of the atmosphere. The surface temperature is determined by starting at the fixed temperature T_{rad} and extrapolating from p_{rad} to the surface pressure p_s using the atmosphere's lapse rate, which is approximately governed by the appropriate adiabat. Since temperature decreases with altitude over much of the depth of a typical atmosphere, the surface temperature so obtained is typically greater than T_{rad} , as illustrated in Fig. 3.6. Increasing the concentration of a greenhouse gas decreases p_{rad} , and therefore increases the surface temperature because temperature is extrapolated from T_{rad} over a greater pressure range. It is very important to recognize that greenhouse warming relies on the decrease of atmospheric temperature with height, which is generally due to the adiabatic profile established by convection. The greenhouse effect works by allowing a planet to radiate at a temperature colder than the surface, but for this to be possible, there must be some cold air aloft for the greenhouse gas to work with.

For an atmosphere whose temperature profile is given by the dry adiabat, the surface temperature is

$$T_s = (p_s/p_{rad})^{R/c_p} T_{rad}. \quad (3.8)$$

With this formula, the Earth's present surface temperature can be explained by taking $p_{rad}/p_s = 0.67$, whence $p_{rad} \approx 670$ mb. Earth's actual radiating pressure is somewhat lower than this estimate, because the atmospheric temperature decays less strongly with height than the dry adiabat. The high surface temperature of Venus can be accounted for by taking $p_{rad}/p_s = 0.0095$, assuming that the temperature profile is given by the non-condensing adiabat for a pure CO₂ atmosphere. Given Venus' 93 bar surface pressure, the radiating level is 880 mb which, interestingly, is only slightly less than Earth's surface pressure. Earth radiates to space from regions quite close to its surface, whereas Venus radiates only from a thin shell near the top of the atmosphere. Note that from the observed Venusian temperature profile in Fig. 2.2, the radiating temperature (253 K) is encountered at $p = 250$ mb rather than the higher pressure we estimated. As for the Earth, our estimate of the precise value p_{rad} for Venus is off because the ideal-gas non-condensing adiabat is not a precise model of the actual temperature profile. In the case of Venus, the problem most likely comes from the ideal-gas assumption and neglect of variations in c_p , rather than condensation.

	Obs. OLR (W/m^2)	Abs. solar (W/m^2)	T_{rad} (actual)	T_{rad} (solar)
Jupiter	14.3	12.7	126 K	110 K
Saturn	4.6	3.8	95 K	81 K
Uranus	0.52	0.93	55 K	58 K
Neptune	0.61	0.38	57 K	47 K

Table 3.1 The energy balance of the gas giant planets, with inferred radiating temperature. The solar-only value of T_{rad} given in the final column is the radiating temperature that would balance the observed absorbed solar energy, in the absence of any internal heat source.

The concept of radiating level and radiating temperature also enables us to make sense of the way energy balance constrains the climates of gas giants like Jupiter and Saturn, which have no distinct surface. The essence of the calculation we have already done for rocky planets is to use the top-of-atmosphere energy budget to determine the parameters of the adiabat, and then extrapolate temperature to the surface along the adiabat. For a non-condensing adiabat, the atmospheric profile compatible with energy balance is $T(p) = T_{rad}(p/p_{rad})^{R/c_p}$. This remains the appropriate temperature profile for a (non-condensing) convecting outer layer of a gas giant, and the only difference with the previous case is that, for a gas giant, there is no surface to act as a natural lower boundary for the adiabatic region. At some depth, convection will give out and the adiabat must be matched to some other temperature model in order to determine the base of the convecting region, and to determine the temperature of deeper regions. There is no longer any distinct surface to be warmed by the greenhouse effect, but the greenhouse gas concentration of the atmosphere nonetheless affects $T(p)$ through p_{rad} . For example, adding some additional greenhouse gas to the convecting outer region of Jupiter's atmosphere would decrease p_{rad} , and therefore increase the temperature encountered at, say, the 1 bar pressure level.

The energy balance suffices to uniquely determine the temperature profile because the non-condensing adiabat is a one-parameter family of temperature profiles. The saturated adiabat for a mixture of condensing and non-condensing gases is also a one-parameter family, defined by Eq. (2.33), and can therefore be treated similarly. If the appropriate adiabat for the planet had more than one free parameter, additional information beyond the energy budget would be needed to close the problem. On the other hand, a single-component condensing atmosphere such as described by Eq. (2.27) yields a temperature profile with no free parameters that can be adjusted so as to satisfy the energy budget. The consequences of this quandary will be taken up as part of our discussion of the runaway greenhouse phenomenon, in Chapter 4.

Using infrared telescopes on Earth and in space, one can directly measure the OLR of the planets in our Solar System. In the case of the gas giants, the radiated energy is substantially in excess of the absorbed solar radiation. Table 3.1 compares the observed OLR to the absorbed solar flux for the gas giants. With the exception of Uranus, the gas giants appear to have a substantial internal energy source, which raises the radiating temperature to values considerably in excess of what it would be if the planet were heated by solar absorption alone. Uranus is anomalous, in that it appears to be emitting less energy than it receives from the Sun. Uncertainties in the observed OLR for Uranus would actually allow

the emission to be in balance with solar absorption, but would still appear to preclude any significant internal energy source. This may indicate a profound difference in the internal dynamics of Uranus. On the other hand, the unusually large tilt of Uranus' rotation axis means that Uranus has an unusually strong seasonal variation of solar heating, and it may be that the hemisphere that has been observed so far has not yet had time to come into equilibrium, which would throw off the energy balance estimate.

Because it is the home planet, Earth's radiation budget has been very closely monitored by satellites. Indirect inferences based on the rate of ocean heat uptake indicate that the top-of-atmosphere radiation budget is currently out of balance, the Earth receiving about 1 W/m^2 more from solar absorption than it emits to space as infrared.³ This is opposite from the imbalance that would be caused by an internal heating. It is a direct consequence of the rapid rise of CO_2 and other greenhouse gases, caused by the bustling activities of Earth's human inhabitants. The rapid greenhouse gas increase has cut down the *OLR*, but because of the time required to warm up the oceans and melt ice, the Earth's temperature has not yet risen enough to restore the energy balance.

Exercise 3.4 A typical well-fed human in a resting state consumes energy in the form of food at a rate of 100W, essentially all of which is put back into the surroundings in the form of heat. An astronaut is in a spherical escape pod of radius r , far beyond the orbit of Pluto, so that it receives essentially no energy from sunlight. The air in the escape pod is isothermal. The skin of the escape pod is a good conductor of heat, so that the surface temperature of the sphere is identical to the interior temperature. The surface radiates like an ideal blackbody.

Find an expression for the temperature in terms of r , and evaluate it for a few reasonable values. Is it better to have a bigger pod or a smaller pod? In designing such an escape pod, should you include an additional source of heat if you want to keep the astronaut comfortable?

How would your answer change if the pod were cylindrical instead of spherical? If the pod were cubical?

Bodies such as Mercury or the Moon represent the opposite extreme to the uniform-temperature limit. Having no atmosphere or ocean to transport heat, and a rocky surface through which heat is conducted exceedingly slowly, each bit of the planet is, to a good approximation, thermally isolated from the rest. Moreover, the rocky surface takes very little time to reach its equilibrium temperature, so the surface temperature at each point is very nearly in equilibrium with the instantaneous absorbed solar radiation, with very little day-night or seasonal averaging. In this case, averaging the energy budget over the planet's surface gives a poor estimate of the temperature, and it would be more accurate to compute the instantaneous equilibrium temperature for each patch of the planet's surface in isolation. For example, consider a point on the planet where the Sun is directly overhead at some particular instant of time. At that time, the rays of sunlight come in perpendicularly to a small patch of the ground, and the absorbed solar radiation per unit area is simply $(1 - \alpha)L_\odot$; the energy balance determining the ground temperature is then $\sigma T^4 = (1 - \alpha)L_\odot$, without the factor of $\frac{1}{4}$ we had when the energy budget was averaged over the entire surface of an isothermal planet. For Mercury, this yields a temperature of 622 K, based on the mean

³ At the time of writing, top-of-atmosphere satellite measurements are not sufficiently accurate to permit direct observation of this imbalance.

orbital distance and an albedo of 0.1. This is similar to the observed maximum temperature on Mercury, which is about 700 K (somewhat larger than the theoretical calculation because Mercury's highly elliptical orbit brings it considerably closer to the Sun than the mean orbital position). The Moon, which is essentially in the same orbit as Earth and shares its solar constant, has a predicted maximum temperature of 384 K, which is very close to the observed maximum. In contrast, the maximum surface temperature on Earth stays well short of 384 K, even at the hottest time of day in the hottest places. The atmosphere of Mars in the present epoch is thin enough that this planet behaves more like the no-atmosphere limit than the uniform-temperature limit. Based on a mean albedo of 0.25, the local maximum temperature should be 297 K, which is quite close to the observed maximum temperature.

More generally speaking, when doing energy balance calculations the temperature we have in mind is the temperature averaged over an appropriate portion of the planet and over an appropriate time interval, where what is "appropriate" depends on the response time and the efficiency of the heat transporting mechanisms of the planet under consideration. Correspondingly, the appropriate incident solar flux to use is the incident solar flux per unit of radiating surface, averaged consistently with temperature. We will denote this mean solar flux by the symbol S . The term *insolation* will be used to refer to an incident solar flux of this type, sometimes with additional qualifiers as in "surface insolation" to distinguish the flux reaching the ground from that incident at the top of the atmosphere. For an isothermal planet, $S = \frac{1}{4}L_{\odot}$, while at the opposite extreme $S = L_{\odot}$ for the instantaneous response at the *subsolar point* - the point on the planet at which the sun is directly overhead. In other circumstances it might be appropriate to average along a latitude circle, or over a hemisphere. A more complete treatment of geographical, seasonal, and diurnal temperature variations will be given in Chapter 7.

Exercise 3.5 Consider a planet which is tide-locked to its sun, so that it always shows the same face to the sun as it proceeds in its orbit (just as the Moon always shows the same face to the Earth). Estimate the mean temperature of the dayside of the planet, assuming the illuminated face to be isothermal, but assuming that no heat leaks to the nightside.

3.4 ICE-ALBEDO FEEDBACK

Albedo is not a static quantity determined once and for all time when a planet forms. In large measure, albedo is determined by processes in the atmosphere and at the surface which are highly sensitive to the state of the climate. Clouds consist of suspended tiny particles of the liquid or solid phase of some atmospheric constituent; such particles are very effective reflectors of visible and ultraviolet light, almost regardless of what they are made of. Clouds almost entirely control the albedos of Venus, Titan, and all the gas giant planets, and also play a major role in Earth's albedo. In addition, the nature of a planet's surface can evolve over time, and many of the surface characteristics are strongly affected by the climate. Table 3.2 gives the albedo of some common surface types encountered on Earth. The proportions of the Earth covered by sea ice, snow, glaciers, desert sands, or vegetation of various types are determined by temperature and precipitation patterns. As climate changes, the surface characteristics change too, and the resulting albedo changes feed back on the state of the climate. It is not a "chicken and egg" question of whether climate causes albedo or albedo causes climate; rather it is a matter of finding a consistent state compatible with the physics of the way climate affects albedo and the way albedo

Table 3.2 Typical values of albedo for various surface types.

Surface type	Albedo
Clean new H ₂ O snow	0.85
Bare sea ice	0.5
Clean H ₂ O glacier ice	0.6
Deep water	0.1
Sahara Desert sand	0.35
Martian sand	0.15
Basalt (any planet)	0.07
Granite	0.3
Limestone	0.36
Grassland	0.2
Deciduous forest	0.14
Conifer forest	0.09
Tundra	0.2

These are only representative values. Albedo can vary considerably as a function of detailed conditions. For example, the ocean albedo depends on the angle of the solar radiation striking the surface (the value given in the table is for near-normal incidence), and the albedo of bare sea ice depends on the density of air bubbles.

affects climate. In this sense, albedo changes lead to a form of climate *feedback*. We will encounter many other kinds of feedback loops in the climate system.

Among all the albedo feedbacks, that associated with the cover of the surface by highly reflective snow or ice plays a distinguished role in thinking about the evolution of the Earth's climate. Let's consider how albedo might vary with temperature for a planet entirely covered by a water ocean – a reasonable approximation to Earth, which is $\frac{2}{3}$ ocean. We will characterize the climate by the global mean surface temperature T_s , but suppose that, like Earth, the temperature is somewhat colder than T_s at the poles and somewhat warmer than T_s at the Equator. When T_s is very large, say greater than some threshold temperature T_0 , the temperature is above freezing everywhere and there is no ice. In this temperature range, the planetary albedo reduces to the relatively low value (call it α_0) characteristic of sea water. At the other extreme, when T_s is very, very low, the whole planet is below freezing, the ocean will become ice-covered everywhere, and the albedo reduces to that of sea ice, which we shall call α_i . We suppose that this occurs for $T_s < T_i$, where T_i is the threshold temperature for a globally frozen ocean. In general T_i must be rather lower than the freezing temperature of the ocean, since when the mean temperature $T_s = T_{freeze}$ the equatorial portions of the planet will still be above freezing. Between T_i and T_0 it is reasonable to interpolate the albedo by assuming the ice cover to decrease smoothly and monotonically from 100% to zero. The phenomena we will emphasize are not particularly sensitive to the detailed form of the interpolation, but the quadratic interpolation

$$\alpha(T) = \begin{cases} \alpha_i & \text{for } T \leq T_i, \\ \alpha_0 + (\alpha_i - \alpha_0) \frac{(T - T_0)^2}{(T_i - T_0)^2} & \text{for } T_i < T < T_0 \\ \alpha_0 & \text{for } T \geq T_0 \end{cases} \quad (3.9)$$

qualitatively reproduces the shape of the albedo curve which is found in detailed calculations. In particular, the slope of albedo vs. temperature is large when the temperature is low and the planet is nearly ice-covered, because there is more area near the Equator, where ice melts first. Conversely, the slope reduces to zero as the temperature threshold for an ice-free planet is approached, because there is little area near the poles where the last ice survives; moreover, the poles receive relatively little sunlight in the course of the year, so the albedo there contributes less to the global mean than does the albedo at lower latitudes. Note that this description assumes an Earthlike planet, which on average is warmest near the Equator. As will be discussed in Chapter 7, other orbital configurations could lead to the poles being warmer, and this would call for a different shape of albedo curve.

Ice-albedo feedback of a similar sort could arise on a planet with land, through snow accumulation and glacier formation on the continents. The albedo could have a similar temperature dependence, in that glaciers are unlikely to survive where temperatures are very much above freezing, but can accumulate readily near places that are below freezing – *provided there is enough precipitation*. It is the latter requirement that makes land-based snow/ice-albedo feedback much more complicated than the oceanic case. Precipitation is determined by complex atmospheric circulation patterns that are not solely determined by local temperature. A region with no precipitation will not form glaciers no matter how cold it is made. The present state of Mars provides a good example: its small polar glaciers do not advance to the Equator, even though the daily average equatorial temperature is well below freezing. Still, for a planet like Earth with a widespread ocean to act as a source for precipitation, it may be reasonable to assume that most continental areas will eventually become ice-covered if they are located at sufficiently cold latitudes. In fairness, we should point out that even the formation of sea ice is considerably more complex than we have made it out to be, particularly since it is affected by the mixing of deep unfrozen water with surface waters which are trying to freeze.

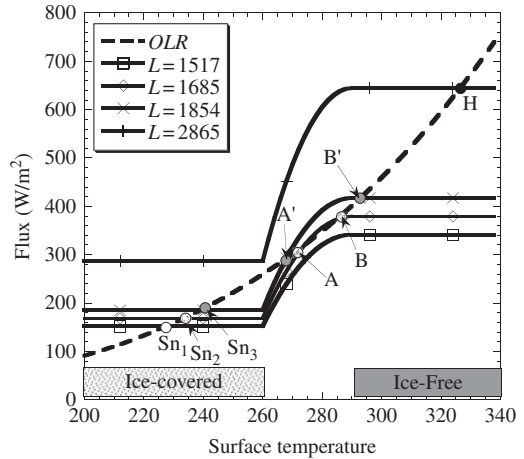
Earth is the only known planet that has an evident ice/snow-albedo feedback, but it is reasonable to inquire as to whether a planet without Earth's water-dominated climate could behave analogously. Snow is always “white” more or less regardless of the substance it is made of, since its reflectivity is due to the refractive index discontinuity between snow crystals and the ambient gas or vacuum. Therefore, a snow-albedo feedback could operate with substances other than water (e.g. nitrogen or methane). Titan presents an exotic possibility, in that its surface is bathed in a rain of tarry hydrocarbon sludge, raising the speculative possibility of “dark glacier” albedo feedbacks. Sea ice forming on Earth's ocean gets its high albedo from trapped air bubbles, which act like snowflakes in reverse. The same could happen for ices of other substances, but sea ice-albedo feedback is likely to require a water ocean. The reason is that water, alone among likely planetary materials, floats when it freezes. Ice forming on, say, a carbon dioxide or methane ocean would sink as soon as it formed, preventing it from having much effect on surface albedo.

Returning attention to an Earthlike waterworld, we write down the energy budget

$$(1 - \alpha(T_s)) \frac{L_{\odot}}{4} = OLR(T_s). \quad (3.10)$$

This determines T_s as before, with the important difference that the solar absorption on the left hand side is now a function of T_s instead of being a constant. Analogously to Fig. 3.5, the equilibrium surface temperature can be found by plotting the absorbed solar radiation and the OLR vs. T_s on the same graph. This is done in Fig. 3.8, for four different choices of L_{\odot} . In

Figure 3.8 Graphical determination of the possible equilibrium states of a planet whose albedo depends on temperature in accordance with Eq. (3.9). The OLR is computed assuming the atmosphere has no greenhouse effect, and the albedo parameters are $\alpha_0 = 0.1$, $\alpha_i = 0.6$, $T_i = 260$ K, and $T_0 = 290$ K. The solar constant for the various solar absorption curves is indicated in the key.



this plot, we have taken $OLR = \sigma T^4$, which assumes no greenhouse effect.⁴ In contrast with the fixed-albedo case, the ice-albedo feedback allows the climate system to have *multiple equilibria*: there can be more than one climate compatible with a given solar constant, and additional information is required to determine which state the planet actually settles into. The nature of the equilibria depends on L_{\odot} . When L_{\odot} is sufficiently small (as in the case $L_{\odot} = 1517 \text{ W/m}^2$ in Fig. 3.8) there is only one solution, which is a very cold globally ice-covered Snowball state, marked Sn_1 on the graph. Note that the solar constant that produces a unique Snowball state exceeds the present solar constant at Earth's orbit. Thus, were it not for the greenhouse effect, Earth would be in such a state, and would have been for its entire history. When L_{\odot} is sufficiently large (as in the case $L_{\odot} = 2865 \text{ W/m}^2$ in Fig. 3.8) there is again a unique solution, which is a very hot globally ice-free state, marked H on the graph. However, for a wide range of intermediate L_{\odot} , there are three solutions: a Snowball state (Sn_2), a partially ice covered state with a relatively large ice sheet (e.g. A), and a warmer state (e.g. B) which may have a small ice sheet or be ice-free, depending on the precise value of L_{\odot} . In the intermediate range of solar constant, the warmest state is suggestive of the present or Pleistocene climate when there is a small ice cap, and suggestive of Cretaceous-type hothouse climates when it is ice-free. In either case, the frigid Snowball state is available as an alternate possibility.

As the parameter L_{\odot} is increased smoothly from low values, the temperature of the Snowball state increases smoothly but at some point an additional solution discontinuously comes into being at a temperature far from the previous equilibrium, and splits into a pair as L_{\odot} is further increased. As L_{\odot} is increased further, at some point, the intermediate temperature state merges with the Snowball state, and disappears. This sort of behavior, in which the behavior of a system changes discontinuously as some control parameter is continuously varied, is an example of a *bifurcation*.

Finding the equilibria tells only part of the story. A system placed exactly at an equilibrium point will stay there forever, but what if it is made a little warmer than the equilibrium? Will it heat up yet more, perhaps aided by melting of ice, and ultimately wander far from the

⁴ Of course, this is an unrealistic assumption, since a waterworld would inevitably have at least water vapor – a good greenhouse gas – in its atmosphere.

equilibrium? Or will it cool down and move back toward the equilibrium? Similar questions apply if the state is made initially slightly cooler than an equilibrium. This leads us to the question of *stability*. In order to address stability, we must first write down an equation describing the time evolution of the system. To this end, we suppose that the mean energy storage per unit area of the planet’s surface can be written as a function of the mean temperature; let’s call this function $E(T_s)$. Changes in the energy storage could represent the energy required to heat up or cool down a layer of water of some characteristic depth, and could also include the energy needed to melt ice, or released by the freezing of sea water. For our purposes, all we need to know is that E is a monotonically increasing function of T_s . The energy balance for a time-varying system can then be written

$$\frac{dE(T_s)}{dt} = \frac{dE}{dT_s} \frac{dT_s}{dt} = G(T_s) \tag{3.11}$$

where $G = \frac{1}{4}(1 - \alpha(T_s))L_\odot - OLR(T_s)$. We can define the generalized heat capacity $\mu(T) = dE/dT$, which is positive by assumption. Thus,

$$\frac{dT_s}{dt} = \frac{G(T_s)}{\mu(T_s)}. \tag{3.12}$$

By definition, $G = 0$ at an equilibrium point T_{eq} . Suppose that the slope of G is well-defined near T_{eq} - in formal mathematical language, we say that G is continuously differentiable at T_{eq} , meaning that the derivative of G exists and is a continuous function for T_s in some neighborhood of T_{eq} . Then, if $dG/dT_s < 0$ at T_s , it will also be negative for some finite distance to the right and left of T_s . This is the case for points a and c in the net flux curve sketched in Fig. 3.9. If the temperature is made a little warmer than T_{eq} in this case, $G(T_s)$ and hence $\frac{dT_s}{dt}$ will become negative and the solution will move back toward the equilibrium. If the temperature is made a little colder than T_{eq} , $G(T_s)$ and hence $\frac{dT_s}{dt}$ will become positive, and the solution will again move back toward the equilibrium. In contrast, if $dG/dT_s > 0$ near the equilibrium, as for point b in the sketch, a temperature placed near the equilibrium moves away from it, rather than toward it. Such equilibria are *unstable*. If the slope happens to be exactly zero at an equilibrium, one must look to higher derivatives to determine stability. These are “rare” cases, which will be encountered only for very special settings of the parameters. If the d^2G/dT^2 is non-zero at the equilibrium, the curve takes the form

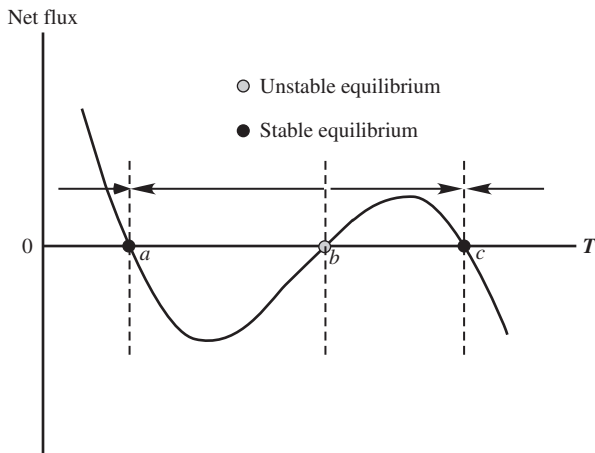


Figure 3.9 Sketch illustrating stable vs. unstable equilibrium temperatures.

of a parabola tangent to the axis at the equilibrium. If the parabola opens upwards, then the equilibrium is stable to displacements to the left of the equilibrium, but unstable to displacements to the right. If the parabola opens downwards, the equilibrium is unstable to displacements to the left but stable to displacements to the right. Similar reasoning applies to the case in which the first non-vanishing derivative is higher order, but such cases are hardly ever encountered.

Exercise 3.6 Draw a sketch illustrating the behavior near marginal equilibria with $d^2G/dT^2 > 0$ and $d^2G/dT^2 < 0$. Do the same for equilibria with $d^2G/dT^2 = 0$, having $d^3G/dT^3 > 0$ and $d^3G/dT^3 < 0$.

It is rare that one can completely characterize the behavior of a nonlinear system, but one-dimensional problems of the sort we are dealing with are exceptional. In the situation depicted in Fig. 3.9, G is positive and dT/dt is positive throughout the interval between b and c . Hence, a temperature placed anywhere in this interval will eventually approach the solution c arbitrarily closely – it will be *attracted* to that stable solution. Similarly, if T is initially between a and b , the solution will be attracted to the stable equilibrium a . The unstable equilibrium b forms the boundary between the *basins of attraction* of a and c . No matter where we start the system within the interval between a and c (and somewhat beyond, depending on the shape of the curve further out), it will wind up approaching one of the two stable equilibrium states. In mathematical terms, we are able to characterize the *global behavior* of this system, as opposed to just the *local behavior* near equilibria.

At an equilibrium point, the curve of solar absorption crosses the *OLR* curve, and the stability criterion is equivalent to stating that the equilibrium is stable if the slope of the solar curve is less than that of the *OLR* curve where the two curves intersect. Using this criterion, we see that the intermediate-temperature large ice-sheet states, labeled A and A' in Fig. 3.8, are unstable. If the temperature is made a little bit warmer than the equilibrium the climate will continue to warm until it settles into the warm state (B or B') which has a small or non-existent ice sheet. If the temperature is made a little bit colder than the equilibrium, the system will collapse into the Snowball state (Sn_2 or Sn_3). The unstable state thus defines the boundary separating the basin of attraction of the warm state from that of the Snowball state.

Moreover, if the net flux $G(T)$ is continuous and has a continuous derivative (i.e. if the curve has no “kinks” in it), then the sequence of consecutive equilibria always alternates between stable and unstable states. For the purpose of this theorem, the rare marginal states with $dG/dT = 0$ should be considered “wildcards” that can substitute for either a stable or unstable state. The basic geometrical idea leading to this property is more or less evident from Fig. 3.9, but a more formalized argument runs as follows: Let T_a and T_b be equilibria, so that $G(T_a) = G(T_b) = 0$. Suppose that the first of these is stable, so $dG/dT < 0$ at T_a , and also that the two solutions are consecutive, so that $G(T)$ does not vanish for any T between T_a and T_b . Now if $dG/dT < 0$ at T_b , then it follows that $G > 0$ just to the left of T_b . The slope near T_a similarly implies that $G < 0$ just to the right of T_a . Since G is continuous, it would follow that $G(T) = 0$ somewhere between T_a and T_b . This would contradict our assumption that the two solutions are consecutive. In consequence, $dG/dT \geq 0$ at T_b . Thus, the state T_b is either stable or marginally stable, which proves our result. The proof goes through similarly if T_a is unstable. Note that we didn’t actually need to make use of the condition

that dG/dT be continuous everywhere: it's enough that it be continuous near the equilibria, so we can actually tolerate a few kinks in the curve.

A consequence of this result is that, if the shape of $G(T)$ is controlled continuously by some parameter like L_{\odot} , then new solutions are born in the form of a single marginal state which upon further change of L_{\odot} splits into a stable/unstable or unstable/stable pair. The first member of the pair will be unstable if there is a pre-existing stable solution immediately on the cold side of the new one, as is the case for the Snowball states S_n in Fig. 3.8. The first member will be stable if there is a pre-existing unstable state on cold side, or a pre-existing stable state on the warm side (e.g. the state H in Fig. 3.8). What we have just encountered is a very small taste of the very large and powerful subject of *bifurcation theory*.

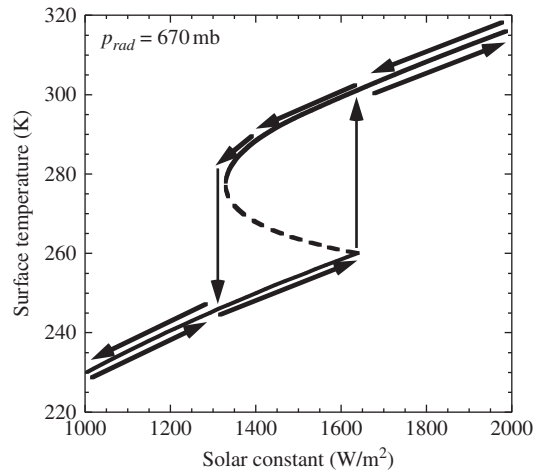
3.4.1 Faint Young Sun, Snowball Earth and hysteresis

We now have enough basic theoretical equipment to take a first quantitative look at the Faint Young Sun problem. To allow for the greenhouse effect of the Earth's atmosphere, we take $p_{rad} = 670$ mb, which gives the correct surface temperature with the observed current albedo $\alpha = 0.3$. How much colder does the Earth get if we ratchet the solar constant down to 960 W/m^2 , as it was 4.7 billion years ago when the Earth was new? As a first estimate, we can compute the new temperature from Eq. (3.8) holding p_{rad} and the albedo fixed at their present values. This yields 261 K. This is substantially colder than the present Earth. The fixed albedo assumption is unrealistic, however, since the albedo would increase for a colder and more ice-covered Earth, leading to a substantially colder temperature than we have estimated. In addition, the strength of the atmospheric greenhouse effect could have been different for the Early Earth, owing to changes in the composition of the atmosphere.

An attempt at incorporating the ice-albedo feedback can be made by using the energy balance Eq. (3.10) with the albedo parameterization given by Eq. (3.9). For this calculation, we choose constants in the albedo formula that give a somewhat more realistic Earthlike climate than those used in Fig. 3.8. Specifically, we set $\alpha_0 = 0.28$ to allow for the albedo of clouds and land, and $T_0 = 295$ K to allow a slightly bigger polar ice sheet. The position of the equilibria can be determined by drawing a graph like Fig. 3.8, or by applying a root-finding algorithm like Newton's method to Eq. (3.10). The resulting equilibria are shown as a function of L_{\odot} in Fig. 3.10, with p_{rad} held fixed at 670 mb. Some techniques for generating diagrams of this type are developed in Problem 3.34. For the modern solar constant, and $p_{rad} = 670$ mb, the system has a stable equilibrium at $T_s = 286$ K, close to the observed modern surface temperature, and is partially ice-covered. However, the system has a second stable equilibrium, which is a globally ice-covered Snowball state having $T_s = 249$ K. Even today, the Earth would stay in a Snowball state if it were somehow put there. The two stable equilibria are separated by an unstable equilibrium at $T_s = 270$ K, which defines the boundary between the set of initial conditions that go to the "modern" type state, and the set that go to a Snowball state. The attractor boundary for the modern open-ocean state is comfortably far from the present temperature, so it would not be easy to succumb to a Snowball.

Now we turn down the solar constant, and re-do the calculation. For $L_{\odot} = 960 \text{ W/m}^2$, there is only a single equilibrium point if we keep $p_{rad} = 670$ mb. This is a stable Snowball state with $T_s = 228$ K. Thus, if the Early Earth had the same atmospheric composition as today, leading to a greenhouse effect no stronger than the present one, the Earth would have inevitably been in a Snowball state. The open-ocean state only comes into being when

Figure 3.10 Hysteresis diagram obtained by varying L_{\odot} with p_{rad}/p_s fixed at 0.67. Arrows indicate path followed by the system as L_{\odot} is first increased, then decreased. The unstable solution branch is indicated by a dashed curve.



L_{\odot} is increased to 1330 W/m^2 , which was not attained until the relatively recent past. This contradicts the abundant geological evidence for prevalent open water throughout several billion years of Earth's history. Even worse, if the Earth were initially in a stable Snowball state four billion years ago, it would stay in that state until L_{\odot} increased to 1640 W/m^2 , at which point the stable Snowball state would disappear and the Earth would deglaciate. Since this far exceeds the present solar constant, the Earth would be globally glaciated today. This even more obviously contradicts the data.

The currently favored resolution to the paradox of the Faint Young Sun is the supposition that the atmospheric composition of the Early Earth must have resulted in a stronger greenhouse effect than the modern atmosphere produces. The prime candidate gases for mediating this change are CO_2 and CH_4 . The radiative basis of the idea will be elaborated further in Chapter 4, and some ideas about why the atmosphere might have adjusted over time so as to maintain an equable climate despite the brightening Sun are introduced in Chapter 8. Figure 3.11 shows how the equilibria depend on p_{rad} , with L_{\odot} fixed at 960 W/m^2 . Whichever greenhouse gas is the Earth's savior, if it is present in sufficient quantities to reduce p_{rad} to 500 mb or less, then a warm state with an open ocean exists (the upper branch in Fig. 3.11). However, for $420 \text{ mb} < p_{rad} < 500 \text{ mb}$ a stable Snowball state also exists, meaning that the climate that is actually selected depends on earlier history. If the planet had already fallen into a Snowball state for some reason, the early Earth would stay in a Snowball unless the greenhouse gases built up sufficiently to reduce p_{rad} below 420 mb at some point.

Figures 3.10 and 3.11 illustrate an important phenomenon known as *hysteresis*: the state in which a system finds itself depends not just on the value of some parameter of the system, but the history of variation of that parameter. This is possible only for systems that have multiple stable states. For example, in Fig. 3.10 suppose we start with $L_{\odot} = 1000 \text{ W/m}^2$, where the system is inevitably in a Snowball state with $T = 230 \text{ K}$. Let's now gradually increase L_{\odot} . When L_{\odot} reaches 1500 W/m^2 the system is still in a Snowball state, having $T = 254 \text{ K}$, since we have been following a stable solution branch the whole way. However, when L_{\odot} reaches 1640 W/m^2 , the Snowball solution disappears, and the system makes a sudden transition from a Snowball state with $T = 260 \text{ K}$ to the only available stable solution, which is an ice-free state having $T = 301 \text{ K}$. As L_{\odot} increases further to 2000 W/m^2 , we follow the warm, ice-free state and the temperature rises to 316 K . Now suppose we begin to gradually dim

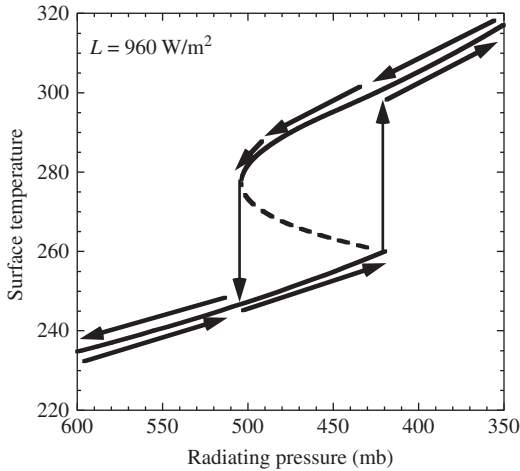


Figure 3.11 As in Fig. 3.10, but varying p_{rad} with $L_{\odot} = 960 \text{ W/m}^2$.

the Sun, perhaps by making the Solar System pass through a galactic dust cloud. Now, we follow the upper, stable branch as L_{\odot} decreases, so that when we find ourselves once more at $L_{\odot} = 1500 \text{ W/m}^2$ the temperature is 294 K and the system is in a warm, ice-free state rather than in the Snowball state we enjoyed the last time we were there. As L_{\odot} is decreased further, the warm branch disappears at $L_{\odot} = 1330 \text{ W/m}^2$ and the system drops suddenly from a temperature of 277 K into a Snowball state with a temperature of 246 K, whereafter the Snowball branch is again followed as L_{\odot} is reduced further. The trajectory of the system as L_{\odot} is increased then decreased back to its original value takes the form of an open loop, depicted in Fig. 3.10.

The thought experiment of varying L_{\odot} in a hysteresis loop is rather fanciful, but many atmospheric processes could act to either increase or decrease the greenhouse effect over time. For the very young Earth, with $L_{\odot} = 960 \text{ W/m}^2$, the planet falls into a Snowball when p_{rad} exceeds 500 mb, and thereafter would not deglaciate until p_{rad} was reduced to 420 mb or less (see Fig. 3.11). The boundaries of the hysteresis loop, which are the critical thresholds for entering and leaving the Snowball, depend on the solar constant. For the modern solar constant, the hysteresis loop operates between $p_{rad} = 690 \text{ mb}$ and $p_{rad} = 570 \text{ mb}$. It takes less greenhouse effect to keep out of the Snowball now than it did when the Sun was fainter, but the threshold for initiating a Snowball in modern conditions is disconcertingly close to the value of p_{rad} which reproduces the present climate.

The fact that the freeze-thaw cycle can exhibit hysteresis as atmospheric composition changes is at the heart of the Snowball Earth phenomenon. An initially warm state can fall into a globally glaciated Snowball if the atmospheric composition changes in such a way as to sufficiently weaken the greenhouse effect. Once the threshold is reached, the planet can fall into a Snowball relatively quickly – in a matter of a thousand years or less – since sea ice can form quickly. However, to deglaciate the Snowball, the greenhouse effect must be increased far beyond the threshold value at which the planet originally entered the Snowball state. Atmospheric composition must change drastically in order to achieve such a great increase, and this typically takes many millions of years. When deglaciation finally occurs, it leaves the atmosphere in a hyper-warm state, which only gradually returns to normal as the atmospheric composition evolves in such a way as to reduce the greenhouse effect. As discussed in Chapter 1, there are two periods in Earth’s past when geological evidence suggests that one or more Snowball freeze-thaw cycles may have occurred. The first is in

the Paleoproterozoic, around 2 billion years ago. At this time, $L_{\odot} \approx 1170 \text{ W/m}^2$, and the thresholds for initiating and deglaciating a Snowball are $p_{rad} = 600 \text{ mb}$ and $p_{rad} = 500 \text{ mb}$ in our simple model. For the Neoproterozoic, about 700 million years ago, $L_{\odot} \approx 1290 \text{ W/m}^2$ and the thresholds are at $p_{rad} = 650 \text{ mb}$ and $p_{rad} = 540 \text{ mb}$.

The boundaries of the hysteresis loop shift as the solar constant increases, but there is nothing obvious in the numbers to suggest why a Snowball state should have occurred in the Paleoproterozoic and Neoproterozoic but not at other times. Hysteresis associated with ice-albedo feedback has been a feature of the Earth's climate system throughout the entire history of the planet. Hysteresis will remain a possibility until the solar constant increases sufficiently to render the Snowball state impossible even in the absence of any greenhouse effect (i.e. with $p_{rad} = 1000 \text{ mb}$). Could a Snowball episode happen again in the future, or is that peril safely behind us? These issues require an understanding of the processes governing the evolution of Earth's atmosphere, a subject that will be taken up in Chapter 8.

Exercise 3.7 Assuming an ice albedo of 0.6, how high does L_{\odot} have to become to eliminate the possibility of a Snowball state? Will this happen within the next five billion years? What if you assume there is enough greenhouse gas in the atmosphere to make $p_{rad}/p_s = 0.5$?

Note: The evolution of the solar constant over time is approximately $L_{\odot}(t) = L_{\odot p} \cdot (0.7 + (t/22.975) + (t/14.563)^2)$, where t is the age of the Sun in billions of years ($t = 4.6$ being the current age) and $L_{\odot p}$ is the present solar constant. This fit is reasonably good for the first 10 billion years of solar evolution.

The “cold start” problem is a habitability crisis that applies to waterworlds in general. If a planet falls into a Snowball state early in its history, it could take billions of years to get out if one needs to wait for the Sun to brighten. The time to get out of a Snowball could be shortened if greenhouse gases built up in the atmosphere, reducing p_{rad} . How much greenhouse gas must build up to deglaciate a Snowball? How long would that take? What could cause greenhouse gases to accumulate on a Snowball planet? These important questions will be taken up in subsequent chapters.

Another general lesson to be drawn from the preceding discussion is that the state with a stable, small ice cap is very fragile, in the sense that the planetary conditions must be tuned rather precisely for the state to exist at all. For example, with the present solar constant, the stable small ice cap solution first appears when p_{rad} falls below 690 mb. However, the ice cap shrinks to zero as p_{rad} is reduced somewhat more, to 615 mb. Hence, a moderate strengthening in the greenhouse effect would, according to the simple energy balance model, eliminate the polar ice entirely and throw the Earth into an ice-free Cretaceous hothouse state. The transition to an ice-free state of this sort is continuous in the parameter being varied; unlike the collapse into a Snowball state or the recovery from a Snowball, it does not result from a bifurcation. In light of its fragility, it is a little surprising that the Earth's present small-icecap state has persisted for the past two million years, and that similar states have occurred at several other times in the past half billion years. Does the simple energy balance model exaggerate the fragility of the stable small-icecap state? Does some additional feedback process adjust the greenhouse effect so as to favor such a state while resisting the peril of the Snowball? These are largely unresolved questions. Attacks on the first question require comprehensive dynamical models of the general circulation, which we will not encounter in the present volume. We will take up, though not resolve, the second question in Chapter 8. It is worth noting that small-icecap states like those of the past two

million years appear to be relatively uncommon in the most recent half billion years of Earth's history, for which data is good enough to render a judgement about ice cover. The typical state appears to be more like the warm relatively ice-free states of the Cretaceous, and perhaps this reflects the fragility of the small-icecap state.

The simple models used above are too crude to produce very precise hysteresis boundaries. Among the many important effects left out of the story are water vapor radiative feedbacks, cloud feedbacks, the factors governing albedo of sea ice, ocean heat transports, and variations in atmospheric heat transport. The phenomena uncovered in this exposition are general, however, and can be revisited across a hierarchy of models. Indeed, the re-examination of this subject provides an unending source of amusement and enlightenment to climate scientists.

3.4.2 Climate sensitivity, radiative forcing and feedback

The simple model we have been studying affords us the opportunity to introduce the concepts of *radiative forcing*, *climate sensitivity coefficient*, and *feedback factor*. These diagnostics can be applied across the whole spectrum of climate models, from the simplest to the most comprehensive.

Suppose that the mean surface temperature depends on some parameter Λ , and we wish to know how sensitive T is to changes in that parameter. For example, this parameter might be the solar constant, or the radiating pressure. It could be some other parameter controlling the strength of the greenhouse effect, such as CO_2 concentration. Near a given Λ , the sensitivity is characterized by $dT/d\Lambda$.

Let G be the net top-of-atmosphere flux, such as used in Eq. (3.11). To allow for the fact that the terms making up the net flux depend on the parameter Λ , we write $G = G(T, \Lambda)$. If we take the derivative of the energy balance requirement $G = 0$ with respect to Λ , we find

$$0 = \frac{\partial G}{\partial T} \frac{dT}{d\Lambda} + \frac{\partial G}{\partial \Lambda} \quad (3.13)$$

so that

$$\frac{dT}{d\Lambda} = -\frac{\frac{\partial G}{\partial \Lambda}}{\frac{\partial G}{\partial T}}. \quad (3.14)$$

The numerator in this expression is a measure of the *radiative forcing* associated with changes in Λ . Specifically, changing Λ by an amount $\delta\Lambda$ will perturb the top-of-atmosphere radiative budget by $\frac{\partial G}{\partial \Lambda} \delta\Lambda$, requiring that the temperature change so as to bring the energy budget back into balance. For example, if Λ is the solar constant L , then $\frac{\partial G}{\partial \Lambda} = \frac{1}{4}(1 - \alpha)$. If Λ is the radiating pressure p_{rad} , then $\frac{\partial G}{\partial \Lambda} = -\frac{\partial OLR}{\partial p_{rad}}$. Since OLR goes down as p_{rad} is reduced, a reduction in p_{rad} yields a positive radiative forcing. This is a warming influence.

Radiative forcing is often quoted in terms of the change in flux caused by a standard change in the parameter, in place of the slope $\frac{\partial G}{\partial \Lambda}$ itself. For example, the radiative forcing due to CO_2 is typically described by the change in flux caused by doubling CO_2 from its pre-industrial value, with temperature and everything else held fixed. This is practically the same thing as $\frac{\partial G}{\partial \Lambda}$ if we take $\Lambda = \log_2 p\text{CO}_2$, where $p\text{CO}_2$ is the partial pressure of CO_2 . Similarly, the climate sensitivity is often described in terms of the temperature change caused by the standard forcing change, rather than the slope $\frac{dT}{d\Lambda}$. For example, the notation ΔT_{2x} would refer to the amount by which temperature changes when CO_2 is doubled.