Lecture Notes on Physics of the Environment (by Georgy I. Burde)

January 25, 2004
1 Introductory remarks

Historically, the study of the atmosphere and oceans, our human environment, together with the study of the broader celestrial environment of the sun, moon and planets, provided the initial impetus to develop the field of physics. In more recent times, however, the earth sciences of meteorology, oceanography and climatology have had comparatively little influence on modern developments in physics. Now the Earth system could be rather considered as a natural laboratory, in which a variety of physical processes takes place. The purpose of these lectures is to show how basic physical concepts and principles can help us model, interpret and predict some of these processes. In general this task cannot be implemented in several lectures. Therefore, we will concentrate on the concepts that are not only crucial in the study of Earth system science but are ubiquitous and can be applied in the studies of other sciences. Those concepts are: equilibrium (or steady state), stability and feedbacks. Of course we will need some other physical principles and laws to illustrate those concepts by models of specific phenomena.

Because of complexity of the Earth system, atmosphere researches, unlike laboratory physicists, cannot perform controlled experiments on the Earth system or describe this system completely by some equations. Instead, after an atmospheric phenomenon is discovered, we develop models, which incorporate a selection of processes that we consider be significant for the phenomenon. These models are usually formulated in terms of mathematical equations and, instead of experiments in physical laboratories, the model experiments are performed by solving these equations under various conditions and interpreting the solutions in terms of physical behavior. The performance of the model (and thus the appropriateness of the hypothesized set of processes) is judged by comparing the model behavior and that of the Earth system.

Normally, a hierarchy of the models is used starting with simple models and progressing to the highly complex general circulation models which require high computer resources. The models considered in these lectures are of the simpler type. Since the simple models can usually be investigated analytically, they may provide a basic physical intuition, which can then be applied in the interpretation of the more complex model results or straight to the natural phenomena. However, because of their very simplicity, they cannot usually be expected to give accurate simulations and predictions of observed atmospheric or oceanic behavior.

2 Basic concepts with simple radiative models

2.1 Physical background: radiation laws

2.1.1 Black body radiation

The Sun’s electromagnetic radiation consists of a mixture of radiations having different wavelengths. The way in which a particular radiation interacts with the atmosphere depends on the wavelength and on the flux of radiant energy. Relationships between radiation absorbed and emitted were examined by Kirchhoff. If the absorptivity of a surface \(a_\lambda\) is defined as the fraction of incident radiation absorbed at a specific wavelength \(\lambda\), then
Kirchhoff’s law states that the ratio of the intensity of emission \( I_\lambda \) at the wavelength \( \lambda \) to the absorptivity \( a_\lambda \) of any substance does not depend on the nature of the substance. It depends only on the temperature \( T \) and the wavelength

\[
\frac{I_\lambda}{a_\lambda} = B_\lambda(T)
\]  

(1)

By definition, a black body is a perfect absorber - an object completely absorbing radiation at all wavelengths. For such an object \( a_\lambda = 1 \) for all values of \( \lambda \), and thus the function \( B_\lambda(T) \) represents the intensity of emission of the black body for a given temperature and wavelength. For any real body \( a_\lambda \) is less than one so that \( I_\lambda < B_\lambda(T) \).

If the emissivity \( \epsilon_\lambda \) is defined as the ratio of the emitted intensity to that for the Black body \( \epsilon_\lambda = I_\lambda/B_\lambda \), Kirchhoff’s law can also be expressed by

\[
\epsilon_\lambda = a_\lambda
\]

(2)

The spectral dependence of the intensity of emission for a black body radiation can be described theoretically by Planck’s function that is shown in Fig. 1 for different temperatures.

Figure 1

Despite the fact that these are idealized spectra (perfect black bodies do not exist in nature), they provide a very close description to the radiation from incandescent material under many conditions. In Fig. 2, for example, we see how closely the solar spectrum corresponds to the Planck spectrum for material at a temperature 5785 K. There is a narrow region in the near ultraviolet radiation where the solar spectrum falls below the Planck’s spectrum, and, in the far ultraviolet and X-ray regions, the solar spectrum rises
above the Planck spectrum, but overall there is a close fit between the two spectra in the visible and infrared regions where sunlight is strongest.

Figure 2

2.1.2 Stefan-Boltzmann law and Wien’s law

These laws were established before Plank’s work but they can be also derived from Planck’s function.

The **Stefan-Boltzmann law** formulates that the total energy emitted by the black body is related to its temperature $T$ by

$$B(T) = \sigma T^4$$

(3)

where $B$ (W m$^{-2}$) is the flux emitted by a unit area of the black body into an imaginary hemisphere surrounding it and $\sigma$ is the Stefan-Boltzmann constant which is independent of the material and is equal to $\sigma = 5.67 \times 10^{-8}$ W m$^{-2}$ K$^{-4}$. 
Wien’s displacement law shows how the maximum of the energy distribution of black body radiation (see Fig. 1) depends on the temperature by

\[ \lambda_{\text{max}} T = \text{const} \tag{4} \]

where the constant on the right-hand side equal to 2897 µm K.

A black body at \( T = 6000 \) K (solar spectrum) will emit maximally at \( \lambda_{\text{max}} \approx 0.5 \) µm and a black body at \( T = 300 \) K (terrestrial spectrum) will emit maximally at \( \lambda \approx 1 \) µm.

2.2 Equilibrium (steady-state) models

2.2.1 Thermal equilibrium

Virtually, all the exchange of energy between the Earth and the rest of the universe takes place by way of radiative transfer. The Earth and its atmosphere are constantly absorbing solar radiation and emitting their own radiation to space. The observation that both the temperatures of the atmosphere and the ocean have remained relatively constant implies that energy is being lost from the atmosphere, the ocean and the Earth’s surface at approximately the same rate as it is being supplied by the Sun. The only mechanism by which this heat can be lost is by the emission of electromagnetic radiation from the Earth’s atmosphere into space. Over a long period of time, the rates of absorption and emission are very nearly equal, that is to say the Earth’s system is in thermal equilibrium.

2.2.2 Radiative equilibrium temperature

It is a basic observational fact that the Earth’s mean surface temperature is about 288 K. In this section we consider whether this can be explained in simple terms given the input of solar radiation and some elementary physics.

The incident flux, or power per unit area, of solar energy at the Earth’s mean distance from the Sun (the so-called solar constant is \( F_s = 1370 \) W m\(^{-2}\)). The solar beam is essentially parallel at the Earth, so the power that is intercepted by the Earth is contained in tube of cross-sectional area \( \pi a^2 \), where \( a \) is the Earth’s radius (see Fig. 3). The total energy received per unit area is therefore \( F_s \pi a^2 \).
We assume that the Earth-atmosphere system has an albedo $A = 0.3$, that is 30% of the incoming solar radiation is reflected back to space without being absorbed. The Earth therefore reflects $AF_s \pi a^2$ of the incoming radiation back to space. If the Earth is assumed to emit as a black body at a uniform absolute temperature $T$ then by the Stefan-Boltzmann law the power emitted per unit area is equal to $\sigma T^4$. However power is emitted in all directions from the total surface area $4\pi a^2$ so the total power emitted is $\sigma T^4 4\pi a^2$. We assume in the present model that all the power is transmitted to space with none absorbed by the atmosphere. Then, assuming that the Earth is in thermal equilibrium, the incoming and outgoing power must balance, so

$$(1 - A)F_s \pi a^2 = \sigma T^4 4\pi a^2$$  

On substituting the values of $A$, $F_s$ and $\sigma$ into this, we obtain $T = 255$ K. This value is significantly lower than the observed mean surface temperature of about 288 K. The present model is clearly lacking of some vital ingredient. We find in the next section that inclusion of the radiation-trapping effect of the atmosphere (the greenhouse effect) leads to a surface temperature that is much closer to reality.

### 2.2.3 Radiative equilibrium temperature with the greenhouse effect

The extent to which the temperature of the surface and the lower portions of the atmosphere can differ from the temperature calculated in the previous section depends on the constituents of the atmosphere. The important property is the opacity of the gas to electromagnetic radiation and particularly to the infrared radiation emitted by the ground.
In Fig. 4 we show the shape of the spectrum of incident solar radiation picking in the visible because of the high effective temperature of the Sun, and the shape of the thermal emission spectrum of the Earth, picking far out in the infrared. This is in accordance with our above estimates based on the Wien’s displacement law (4).

**Figure 4**

![Diagram of Planck curves showing the proportion of energy radiated at each wavelength by the Earth and the solar radiation incident on the Earth.](image)

*At the top are the Planck curves that show the proportion of energy radiated at each wavelength by the Earth and the solar radiation incident on the Earth. There is very little overlap between the curves, which enables us to talk of different atmospheric characteristics for short-wave solar radiation and long-wave heat radiation. At the bottom is the percentage of the radiation at each wavelength that is absorbed in the atmosphere. Absorption is strong on the average for the planetary radiation but weak on the average for the solar radiation. (After R. M. Goody, 1954, Physics of the Stratosphere, Cambridge University Press.)*

In Fig. 4, we also show, at each wavelength, the fraction of the light absorbed from the beam passing directly through the Earth’s atmosphere. We can conclude that the atmosphere is moderately transparent to the visible so that much of the solar radiation can pass through the atmosphere without being absorbed. On the other hand, minor atmospheric constituents, of which the water vapor is the most important, absorb strongly in the infrared so the atmosphere is largely opaque to the planetary heat radiation.

What happens when the atmosphere absorbs radiation emitted from the surface of the Earth? The atmosphere cannot steadily accumulate radiation since it would become heater and heater. Instead, a thermal equilibrium state is established in which the atmosphere emits radiation at the same rate as it absorbs. The radiation is emitted in all
directions and a substantial part of it is intercepted and absorbed by the surface. So the Earth’s surface is heated not only by direct sunlight but also by infrared radiation emitted by the atmosphere. For this reason, in a thermal equilibrium state, the Earth’s surface must radiate away more energy, than it receives directly from the Sun, and the surface temperature should have a temperature that exceeds the value calculated in the previous section for a transparent atmosphere.

To calculate a corrected value we will consider a simple model in which the atmosphere is modelled by a shallow layer of uniform temperature $T_a$ (it is only an approximation since in a real atmosphere the temperature varies substantially with height). The atmosphere is assumed to transmit a fraction $\tau_s$ of any incident solar (short-wave) radiation and a fraction $\tau_l$ of any incident thermal (infra-red or long-wave) radiation (these fractions are called transmittances), and to absorb the reminders. We assume that the ground is at the temperature $T_g$.

Taking account of albedo effects and the differences between the area of the emitting surface $4\pi a^2$ and the intercepted cross-sectional area $\pi a^2$ of the solar beam (see the previous section), the mean unreflected incoming solar flux at the top of the atmosphere is

$$F_0 = \frac{1}{4}(1 - A)F_s$$

or about 240 W m$^{-2}$. Of this, a proportion $\tau_s F_0$ reaches the ground, the reminder being absorbed by the atmosphere.

The ground is assumed to emit as a black body, and it therefore (see (3)) emits upwards the flux $F_g = \sigma T_g^4$, of which a proportion $\tau_l F_g$ reaches the top of the atmosphere. The atmosphere is not a black body, so that its emissivity (the ratio of the actual emitted flux to the flux that would be emitted by a black body at the same temperature), according to Kirchhoff’s law (2), equals to its absorptivity $1 - \tau_l$ (the atmosphere is assumed to emit as a grey body, which is defined as having an emissivity independent of wavelength). Therefore the atmosphere emits fluxes $F_a = (1 - \tau_l)\sigma T_a^4$ both upwards and downwards, as shown in Fig. 5.
We can now balance these fluxes assuming that the system is in equilibrium. At the top of the atmosphere we have

\[ F_0 = F_a + \tau_l F_g \]  

and at the ground

\[ F_g = F_a + \tau_s F_0 \]  

By eliminating \( F_a \) from equations (7) and (8), we obtain

\[ F_g = \sigma T_g^4 = F_0 \frac{1 + \tau_s}{1 + \tau_l} \]  

In the absence of an absorbing atmosphere, we would have \( \tau_s = \tau_l = 1 \) so \( F_g \) would equal \( F_0 \) giving \( T_g = 255 \text{ K} \) as in the previous section. However, rough values for the Earth's atmosphere are \( \tau_s = 0.9 \) (strong transmission and weak absorption of solar radiation) and \( \tau_l = 0.2 \) (weak transmission and strong absorption of thermal radiation), so that \( F_g = 1.6F_0 \) leading to a surface temperature of \( T_g = 1.6^{1/4} \times 255 \text{ K} \approx 286 \text{ K} \), which is quite close to the observed mean value of 288 K (The close agreement is partly fortuitous, however, since in reality non-radiative processes also contribute to the energy balance).

This is a simple example of the greenhouse effect: the greater temperature depends on the fact that there is less absorption (greater transmission) for solar radiation than there is for thermal radiation. Thus, the atmosphere readily transmits solar radiation but
tends to trap thermal radiation (The term "greenhouse effect" is a misnomer, however, since the elevated temperature in a greenhouse does not primarily depend on the similar radiative properties of glass but rather on suppression of convective heat loss).

We can also find the atmospheric flux and the corresponding temperature from equation (7) and (8)

\[ F_a = (1 - \tau_l)\sigma T_a^4 = F_0 \frac{1 - \tau_s \tau_l}{1 + \tau_l} \]  \hspace{1cm} (10)

and this gives the temperature of the shallow atmosphere in the model \( T_a \approx 245 \text{ K} \).

2.3 Stability assessment

2.3.1 Stability and instability concepts

Concepts of stability and instability are crucial in the study of Earth system science. Although assessing the stability of complicated systems can be a demanding task, understanding what is meant is straightforward if we discuss the concept by analogy with a mechanical system, such as an object on a hilly surface (see Fig. 6).

Figure 6

![Stable and unstable systems](image)

Begin by considering a system in equilibrium. In such a system, the vector sum of forces is zero. Now consider what happens to the system if a small perturbation (it might be a small displacement or a small disturbing force which also results in a small displacement...
of the object). If a ball is placed in position ("state") 1 and to it a small force is applied, the ball will not keep moving: it will oscillate for a while and then settle back into the original position, position 1 is said to be stable. If the ball happens to be temporarily in position 2, a small force applied, either to the right or left, is sufficient to move the ball to another state, position 2 is said to be unstable. One can conclude, with the help of Fig. 6, that instability occurs when the unbalance of forces arising as a result of the perturbation results in a force tending to move the ball away from the original position. Correspondingly, an equilibrium state is stable if imposing a perturbation results in the force unbalance that tends to return the ball into the original position. More broadly speaking, one may say that an equilibrium state is unstable if any perturbation imposed on it grows with time and it is stable if any perturbation decays with time.

The concepts of stability and instability can be looked at in a more general way, if one distinguishes between small and finite perturbations, as it is illustrated with the help of Fig. 7. The state 1 may be stable with respect to small perturbations but unstable with respect to larger perturbations, and, as a result of the instability, the system may, after passing through the unstable state 2, find itself in a new stable state 3. Nevertheless, assessing the stability of a system should start from considering small perturbations which provides one with necessary conditions for stability.

The importance of stability concept is in that only stable states can be realized and so, in general, unstable states should not be observed. However, the time scale is also important. The bed of a river is quite stable over a few decades or perhaps a few centuries, but it assumes a new state if one allows it to evolve for a few millennia.
2.3.2 Stability of the Earth’s radiative equilibrium

Let us assess stability of the radiative equilibrium state considered in Section 2.2.2.

Initially, the atmosphere is in thermal equilibrium and therefore the solar radiation received by the Earth’s system is in balance with the emission of long-wavelength planetary radiation into space (equation (5)). It is convenient, for the following calculations, rewrite this equilibrium condition, using a somewhat different notation, in the form

\[ F_0 = \sigma T_e^4 \]  \hspace{1cm} (11)

where \( F_0 \) is the incoming (unreflected) solar flux defined by (6) and \( T_e \) denotes the equilibrium temperature. Now consider the situation where a perturbation is imposed - for example, a heat is suddenly added to the atmosphere, as the result of burning of fossil fuels or by a large thermonuclear explosion. If all the heat were liberated at the same time, then the atmospheric temperature would increase suddenly by \((\Delta T)_0\). The following arguments are aimed at investigating stability of the thermal equilibrium (11) with respect to such a perturbation.

Let us start from qualitative considerations. It is clear, that the planetary radiation emitted into the space would, after imposing the perturbation, be \( \sigma (T + (\Delta T)_0)^4 \), and therefore more radiation would be emitted than received from the Sun. This deficit implies that the Earth would lose the energy and, as the result, would cool. Thus, the perturbation of the temperature would decay with time which means that such an equilibrium state is stable.

The above qualitative arguments cannot help one in estimating the time scale of settling the Earth system back into its original state. This time scale, as a matter of fact, might give an indication of a consistency of the above qualitative arguments. If, for example, the Earth system returned to the equilibrium state slowly and thus the higher temperature conditions persisted for a long time, some physical parameters (for example, albedo), and correspondingly the amount of radiation received by the Earth, might be altered, which could result in additional thermal forcing (see more details in the next section).

Consider an atmosphere having a mass \( M_A \) per unit area and a specific heat at constant pressure \( C_p \). Then, change of the atmospheric temperature with the rate \( \frac{dT}{dt} \) would result in the corresponding change of the internal energy of the atmosphere with the rate \( M_A C_p \frac{dT}{dt} \). Since, in the situation considered, the internal energy is lost due to the difference between the emitted planetary radiation and the solar radiation, one has

\[ M_A C_p \frac{dT}{dt} = F_0 - \sigma T^4 \]  \hspace{1cm} (12)

where the atmospheric temperature \( T \) is now a function of time. In view of (11), equation (12) may be rewritten in the form

\[ M_A C_p \frac{dT}{dt} = \sigma T_e^4 - \sigma T^4 \]  \hspace{1cm} (13)

A solution of this equation is defined if the initial condition is added to (13), as

\[ T = T_e + (\Delta T)_0 \]  \hspace{1cm} at \( t = 0 \)  \hspace{1cm} (14)
Simplifications relying upon a smallness of the initial temperature perturbation \((\Delta T)_0\) can be made in (13) if a deviation \(\Delta T = T - T_e\) of the temperature from its equilibrium value is used as a variable. Then (13) and (14) can be represented in the forms (it has been taken into account that the quantity \(T_e\) defined by (11) is a constant):

\[
M_A C_p \frac{d(\Delta T)}{dt} = \sigma T_e^4 - \sigma T_e^4 \left(1 + \frac{\Delta T}{T_e}\right)^4
\]

\(\Delta T = (\Delta T)_0\) at \(t = 0\) (15)

Providing that \(\Delta T/T_e \ll 1\), then the second term on the right-hand side of (15) can be represented as

\[-\sigma T_e^4 \left(1 + \frac{\Delta T}{T_e}\right)^4 \approx -\sigma T_e^4 \left(1 + \frac{4\Delta T}{T_e}\right)
\]

and, as the result, equation (15) takes the form

\[
\frac{d(\Delta T)}{dt} = \frac{4\sigma T_e^3}{M_A C_p} \Delta T
\]

The solution of the above equation subject to the initial condition (16) is

\[
\Delta T = (\Delta T)_0 e^{-t/\tau_R}, \quad \tau_R = \frac{M_A C_p}{4\sigma T_e^3}
\]

where the constant \(\tau_R\) provides a time scale of the perturbation decay that is known as the radiation relaxation time constant.

For Earth’s atmosphere, where \(M_A = 10.3 \times 10^8\) kg m\(^{-2}\), \(C_p = 1004\) kg\(^{-1}\)K\(^{-1}\), \(T_e = 255\) K, taking \(\sigma = 5.67 \times 10^{-8}\) W m\(^{-2}\) K\(^{-4}\), the value of \(\tau_R\) can be estimated to be \(2.7 \times 10^6\) s or 32 days. From equation (19):

\[
\Delta T = (\Delta T)_0 e^{-t/\tau_R} \quad \text{at } t = \tau_R
\]

\[
\Delta T = (\Delta T)_0 e^{-2\tau_R} \quad \text{at } t = 2\tau_R
\]

\[
\Delta T = (\Delta T)_0 e^{-3\tau_R} \quad \text{at } t = 3\tau_R
\]

For an initial temperature perturbation of 1 K, equivalent to the instantaneous burning of \(240 \times 10^9\) t of coal, then after approximately 32 days, the temperature perturbation will be reduced to 0.37 K, after 64 days it will be 0.13 K and after 96 days it will be 0.05 K. For the Earth’s atmosphere, therefore, the colossal excess of heat will be lost by planetary radiation to space within 100 days and the Earth’s radiative equilibrium will be restored.
2.4 Feedbacks

2.4.1 Global warming

It is well known that the concentration of carbon dioxide $\text{CO}_2$ in the atmosphere increases in time. To understand a mechanism of potential global warming we must investigate the response of the equilibrium surface temperature to an increase in the concentration of radiatively active gases, such as carbon dioxide. (It should be emphasized that the equilibrium conditions are always meant, that is to say it is assumed that the equilibrium will be restored every time after that perturbations in the concentration of $\text{CO}_2$ have been imposed - the variables of the climate system will change to adjust to the new equilibrium state corresponding to the new carbon dioxide concentration.) It is of course an enormously complex question but a simple start can be made by using equations (7)-(9) to find how $T_g$ changes in response to small increments $\Delta \tau_l$ in the transmission of thermal radiation, $\Delta \tau_s$ in the transmission of solar radiation and $\Delta A$ in the albedo, with a fixed solar constant $F_s$.

First, we will rewrite equation (9), using (6), in the form

$$ F_g = \sigma T_g^4 = \frac{1}{4} (1 - A) F_s \frac{1 + \tau_s}{1 + \tau_l} \quad (20) $$

Next, by taking logarithms and differentiation we find that the increment in surface temperature is given by

$$ 4 \frac{\Delta T_g}{T_g} = \frac{\Delta \tau_s}{1 + \tau_s} - \frac{\Delta \tau_l}{1 + \tau_l} - \frac{\Delta A}{1 - A} \quad (21) $$

If the concentration of a radiatively active gas increases, this will generally lead to decreases in the solar and thermal transmissions. A decrease in the thermal transmittance ($\Delta \tau_l < 0$) implies more trapping of thermal radiation and hence contributes to an increase in $T_g$. On the other hand, an decrease in solar transmittance means that less solar radiation reaches the ground and hence contributes to a decrease in $T_g$. Global warming will take place if the former effect prevails, or, in terms of our equation (21), the first fraction on the right-hand side of (21) is less in modulus than the second fraction.

However, the above discussion is far from being complete. The point is that, in the Earth system, there exist related effects reinforcing or, on the opposite, mitigating global warming, examples of which are given below.

Effects reinforcing global warming:

(a) Melting of ice and snow due to an increase of surface temperature will lower the albedo $A$ which, according to (21), will result in a further increase of the surface temperature $T_g$.
(b) More water vapor appearing in the air due to the enhanced evaporation processes will lead to smaller transmission $\tau_l$ for the thermal radiation which will further increase the surface temperature.
(c) More $\text{CO}_2$ leads to an increased growth of plants which lowers the albedo.
(d) A higher sea water temperature gives less $\text{CO}_2$ absorption in sea water which will cause further increase in $\text{CO}_2$. 

14
Effects mitigating global warming:
(e) Increasing long-wave radiation from the atmosphere to space due to an increase of the atmospheric temperature will generally result in losing the energy by the atmosphere and thus reducing the temperature.
(f) Decreasing the so-called adiabatic lapse rate for humid air (see Section 4 for more detailed explanation)

The reinforcing and mitigating effects ascribed above may be described as a series of feedback mechanisms.

2.4.2 Feedback concepts

The feedback mechanisms act as internal controls of the system and result from a special coupling or mutual adjustment among two or more subsystems. Part of the output returned to serve as an input again so that the net response of the system is altered; the feedback mechanisms may act either to amplify the final output (positive feedback) or to diminish it (negative feedback). There is a large number of such mechanisms operating within the various components of the climatic system or between the subsystems.

Figure 8

A simple feedback loop may be presented as in Fig.8 as a system in which the output is partially fed back to the input. A signal $V_S$ entering the loop receives an extra input $V_F$ which leads to

$$V_1 = V_S + V_F$$

(22)
reinforced by the gain function (or transfer function) $G$ as

$$V_2 = GV_1$$

(23)

The signal is picked up by $H$ (a feedback factor) and produces

$$V_F = HV_2$$

(24)

which was introduced in (22). All taken together, we have

$$V_2 = GV_1 = G(V_S + V_F) = G(V_S + HV_2)$$

(25)

which allows us to find an effective gain, which includes the feedback (see Figure (b)), so that

$$G_F = \frac{V_2}{V_S} = \frac{G}{1 - GH} = \frac{G}{1 - f}$$

(26)

where $f = GH$ is the feedback of the system. Thus

$$V_2 = G_FV_S = \frac{G}{1 - GH}V_S = \frac{G}{1 - f}V_S$$

(27)

In the case of zero feedback, $f = 0$ and $G_F = G$. For a negative feedback, $f < 0$ and $0 \leq G_F \leq G$. As $f$ becomes larger and larger negative, $G_F$ tends asymptotically to zero. For positive feedbacks with $0 \leq f < 1$, we find $G_F > G$.

If the system has several feedback mechanisms, that are independent of each other (linked in parallel), characterized by $H_i$, one could write $V_F = \sum H_iV_2$ and obtain

$$V_2 = \frac{G}{1 - \sum H_iG}V_S = \frac{G}{1 - \sum f_i}V_S$$

(28)

2.4.3 Feedbacks in the climate system

To apply the previous concepts to the climate system let us take the simplest possible atmospheric model; it is even simpler than the model considered in Section 2.2.3 as we now consider only total energy fluxes.

The net radiation (solar minus terrestrial) flux $F_{TA}$ at the top of the atmosphere will vanish under equilibrium conditions. Any external perturbations due, e.g., to changes in atmospheric carbon dioxide, water vapor, solar output, volcanic eruptions, etc, will induce an imbalance in the net radiation at the top of the atmosphere $\Delta F_{TA}$. Assume, for example, a sudden doubling of effective CO$_2$ concentration. This would lead to an effective reduction in the Earth’s long-wave radiation at the top of the atmosphere with a magnitude $\Delta I$ and consequently would cause a decrease in the flux $\Delta F_{TA} = -\Delta I$. The energy balance at the top of the atmosphere requires a constant (zero) flux $F_{TA}$. Therefore the Earth’s surface temperature should rise by $\Delta T_{SFC}$ to compensate (this effect is called thermal forcing). Thus we will regard $\Delta F_{TA}$ as the signal input $V_s$ for the climate system and $\Delta T_{SFC}$ as the output $V_2$. We can then write

$$\Delta T_{SFC} = G_F\Delta F_{TA}$$

(29)
where $G_F$ is the transfer function, so that the concepts outlined in the previous section can be applied to the climatic system. In the Fig. 9, an example of the set of positive feedback loops in the climate system is presented.

Figure 9

Example of a set of possible feedback loops in the climate system, where the input is an imbalance in the net radiation at the top of the atmosphere $\Delta F_{TA}$ and the output is the change in surface temperature of the earth. The symbols $G$, $H_1$, $H_2$, ..., and $G_F$ represent the gain, feedback factors, and effective gain of the climate system, respectively.

It should be noted that, even though the cloudiness-temperature interaction is given in Fig. 9 as an example of simple positive feedback system, clouds can lead to many
different feedback processes since they are both excellent absorbers of thermal radiation (positive feedback) and effective reflectors of solar radiation (negative feedback).

3 Atmospheric motion

3.1 Physical background: forces that effect the atmospheric motion

The atmosphere is a fluid and as such its motion may be analyzed by means of hydrodynamical principles. The basic principle is Newton’s second law which says that accelerations of a fluid element (a volume of air) are produced by forces. If no net acceleration is present the motion must be under balanced forces. Two of the most important forces will be considered in detail below.

3.1.1 Pressure gradient force

Because of variations in pressure throughout the atmosphere, both in the horizontal and vertical, there is a force acting on any given element of air. Arguments similar to those led to (40) may be used to show that the pressure force is proportional to the magnitude of the gradient, i.e., the space derivative of pressure (see Fig. 10).

In particular, the force per unit mass (acceleration due to the pressure gradient force) in the $x$ direction is

$$f_x = -\frac{1}{\rho} \frac{\partial p}{\partial x}$$  \hspace{1cm} (30)

the pressure gradient force in the $y$ direction is

$$f_y = -\frac{1}{\rho} \frac{\partial p}{\partial y}$$  \hspace{1cm} (31)
and in the z (vertical) direction it is

\[ f_z = -\frac{1}{\rho} \frac{\partial p}{\partial z} \]  \hspace{1cm} (32)

The resultant of \( \frac{\partial p}{\partial x} \), \( \frac{\partial p}{\partial y} \) and \( \frac{\partial p}{\partial z} \) is the pressure gradient vector \( \text{grad} \, p \) and in vector notation we may combine these results to express the pressure gradient force as

\[ f = -\frac{1}{\rho} \text{grad} \, p \]  \hspace{1cm} (33)

### 3.1.2 Coriolis Force

The pressure gradient force, which has just been considered, is a true force. The second force is an apparent force which results from the rotation of the earth, rather than a true force. Even though Newton’s laws are not applicable on the rotating earth, which is the frame of reference for atmospheric motions, it is convenient to use Newton’s laws in their simple form on the earth including the rotation effects as fictitious forces in the equations of motion of a particle. The Coriolis force arises because a moving object (an air parcel) tries to maintain its original direction with respect to an absolute coordinate system, but on the Earth observing platform, it undergoes an apparent deflection (see Fig. 11).

This looks like a deflecting force acts perpendicular to the right (in the northern hemisphere) of the direction of motion of the parcel at any instant. In a similar fashion, it would be deflected to the left in the southern hemisphere. It follows then that the deflecting force tends to change the direction of motion of an air parcel but not its speed.
If one uses the coordinate system with $x$ and $y$ axes, directed respectively along the latitude and meridional circles, and the corresponding $x$ and $y$ velocity components $u$ and $v$, the horizontal components of the deflecting force at latitude $\phi$ are

$$F_x = 2\omega \sin \phi v, \quad F_y = -2\omega \sin \phi u$$

(34)

where $\omega$ is the angular velocity of the earth about its polar axis. If $V$ denotes the resultant of the velocity components $u$ and $v$, the magnitude of the deflecting force acting on unit
mass of air moving over the earth’s surface is

\[ F = 2\omega \sin \phi V \]  

(35)

It is seen that the Coriolis force increases with latitude and vanishes at the equator.

### 3.2 Steady-state atmospheric motion (motion under balanced forces)

The discussions in this and next sections will mainly concern the motion of air in the northern hemisphere. The corresponding conclusions for the southern hemisphere may be easily derived by changing the deflecting force direction in all the arguments.

#### 3.2.1 The geostrophic wind

The first of the atmospheric motions to be considered is that resulting when the pressure gradient force just balances the deflecting force. A mass of air at rest in a pressure field is acted upon only by the pressure gradient force. As soon as it commences to move as a result of this force, however, the deflecting force starts to operate, and the more rapid the motion, the greater the corresponding deflecting force becomes. It is clear, therefore, that air does not flow perpendicular to the isobars, from high to low pressure, but that it is subjected to a force which continually acts to deflect it to the right. The only situation in which the two forces may balance is that where the air is moving parallel to isobars. The resulting wind is known as the geostrophic wind. The balance of forces may be seen with the aid of Fig. 12. With a suitable value of \( v \), then, the deflecting force will exactly balance the pressure gradient force and steady motion results.

The geostrophic wind equation is, therefore

\[ 2\omega \sin \phi v = \frac{1}{\rho} \frac{\partial p}{\partial x} \]  

(36)

and \( v \) is the geostrophic wind velocity. In a general case, when the isobars are parallel to neither \( x \) nor \( y \) axis, but have some intermediate orientation, the geostrophic wind velocity \( v \) may be considered as the resultant of the two component velocities \( u \) and \( v \) where \( v \) is defined by equation (36) and \( u \) by the following equation

\[ 2\omega \sin \phi u = -\frac{1}{\rho} \frac{\partial p}{\partial y} \]  

(37)
3.2.2 The effects of friction in the geostrophic motion

It is assumed in the above treatment that frictional forces are so small to be negligible. This is an accurate statement of conditions in the free atmosphere, at great heights (1 km or greater) above the surface. Charts depicting patterns in the middle and upper troposphere reveal that the wind is nearly parallel to the isobars, low pressure on the left, and that its speed increases in proportion to the pressure gradient - i.e. in the free atmosphere masses of air move according to the geostrophic wind relations (36) and (37).

At lower levels, the frictional drag of the irregularities at the earth’s surface on the moving air is appreciable and leads to a decrease in velocity. The pressure gradient force is not affected thereby, but the deflecting force is reduced. This leads to air motion across isobars, from high to low pressure. Steady motion results, with the deflecting and frictional forces balancing the pressure gradient force. The flow across isobars may be noted on any surface weather map.

3.2.3 The effects of curvature in the geostrophic motion

The geostrophic wind is defined as a hypothetical motion, parallel to the isobars, that would provide a balance between pressure gradient and Coriolis forces. Thus, the effects of friction are ignored, and more significantly, any possible curvature of the actual flow, which would lead to centripetal accelerations of the parcels, is also neglected. In other words, when the isobars are curved, the motion of the air can no longer be linear, and the differences of the pressure gradient and Coriolis forces should be in balance with a centrifugal force resulting because of the curvature of the trajectory of the parcel. The centrifugal force will have magnitude $v^2/r$, where $v$ is the velocity of the parcel, and $r$ is the radius of curvature of the motion, and acts to a direction normal to the motion at a
point on the curved trajectory. The radius of curvature of the motion must be constant if the motion is to be steady, and so the wind must blow in circular paths.

Consider a region of low pressure, having circular isobars, as indicated in Fig. 13.

Figure 13

Cyclone and anticyclone. (a) The geostrophic wind circulates around a low pressure centre in the counterclockwise direction in the Northern Hemisphere. This type of motion is called cyclonic motion. A low pressure system is often referred to as a 'cyclone'. (b) A high pressure system is also called an 'anticyclone'. Geostrophic wind rotates in the Northern Hemisphere in the clockwise direction around the high pressure centre. Motions of this type are called anticyclonic motions.
Since the pressure gradient force \(- (1/\rho)(\partial p/\partial r)\) always acts at right angles to the isobars, while the deflecting force and centrifugal force always act at right angles to the direction of motion, it follows that the only direction of motion which permits a balance of forces is that parallel to isobars. Theoretically either clockwise or counterclockwise motion is possible. If clockwise motion is assumed, the centrifugal force must balance the sum of the pressure gradient and deflecting forces. Under ordinary meteorological conditions, the centrifugal force is considerably smaller than either of the other two - only when \( r \) is small (\( r < 100 \) km) does the centrifugal force approach the sum of the other two in magnitude. The alternative, counterclockwise motion, as also shown in the Figure (it is known as a cyclonic case), means that the pressure gradient force just balances the sum of the deflecting and centrifugal force. No limitations of the above kind are present, and the balanced conditions may hold for any value of the radius of curvature. The balance of forces for the cyclonic case is given by

\[
\frac{1}{\rho} \frac{\partial p}{\partial r} = f r + \frac{v^2}{r}
\]

(38)

where \( f = 2\omega \sin \phi \) (see (35)) is the Coriolis parameter.

For an anticyclonic case, when the clockwise motion is around the region of high pressure, similar reasoning shows that the balance of forces may be expressed as

\[
\frac{1}{\rho} \frac{\partial p}{\partial r} = f r - \frac{v^2}{r}
\]

(39)

In this case the velocity must have a maximum value since the balance of forces is destroyed if the centrifugal force becomes greater than the deflecting force.

### 3.3 Vertical structure of the atmosphere

#### 3.3.1 Temperature structure of the atmosphere

Many properties of the atmosphere are determined by its pressure, which is highest at Earth’s surface and decrease rapidly with increasing altitude. The other crucial atmospheric property is the temperature; unlike the pressure variation, the temperature variation with height is quite complex (see Fig. 14).
The inflection points on the temperature profile are used to distinguish the different regions for study and reference. Beginning at Earth’s surface, these regions are called the troposphere, the stratosphere, the mesosphere, and the thermosphere, and their boundaries the tropopause, the stratopause and the mesopause, respectively. In our lectures, we will restrict our discussions mainly to the troposphere, which is the most important region for climate and life on Earth. The troposphere is heated from below due to the absorption of radiation coming from the sun. The troposphere is also called the lower atmosphere. It is here that most weather phenomena, such as cyclones, fronts, hurricanes, rain, snow, thunder and lightning occur.

3.3.2 Horizontal and vertical variations of pressure

The pressure of the atmosphere at a point of the earth’s surface is a measure of the weight of a column of air above this point. This pressure varies from day to day, and from one part of the earth to another. There are two causes for this pressure variation at a given point. First, as a result of the wind distribution, there may be an accumulation of air above a given point on the surface, producing a rise in pressure. A depletion of air would cause a fall in pressure. A second cause which may be in operation is the horizontal transport of air of greater or lesser density producing a rise or fall of pressure. The horizontal variations in pressure may be indicated by means of isobars. These are curves which join points of equal pressure in the same manner that lines on a contour
map join points at the same height above sea level (an example is shown in Fig. 15). The isobars are usually much more regular in their shape than contour lines, but they may also indicate "highs", "lows", "ridges", and "troughs".

Figure 15

The pressure at any point in the atmosphere (not obligatory at Earth’s surface) is, by definition, the weight of the vertical column of air of unit cross section which extends from the height to the top of the atmosphere. It follows directly then that pressure decreases with increasing height in the atmosphere. The difference in pressure between height $z$ and height $z + dz$ (see Fig. 16) may be found in the following manner.
Since unit cross section is being considered, the volume of the element of air is $dz$ and its mass is $\rho dz$, where $\rho$ denotes the average density in the height interval $dz$. Since weight is given by the product of the mass and the acceleration of gravity, and since $p$ decreases with increasing height, then

$$dp = -\rho g dz, \quad \text{or} \quad \frac{1}{\rho} \frac{dp}{dz} = -g \quad (40)$$

This equation is known as the hydrostatic equation. It states that the vertical component of the pressure gradient force (32) is exactly balanced by the acceleration of gravity $g$.

Note that even though this equation reflects a balance of vertical forces (pressure and weight) in the atmosphere with no vertical acceleration, in the atmospheric applications the hydrostatic equation is widely used not only for air in rest. The point is, that vertical accelerations are generally small as compared with these two forces, so that the force due to decrease in pressure with height is very closely in balance with the force due to gravitation. Therefore (40) is one of the basic equations in meteorology.

If the equation of state for dry air

$$p = R\rho T, \quad R = 287.04 \ \text{J kg}^{-1} \ \text{K}^{-1} \quad (41)$$
with the temperature $T$ given as a function of $z$, is introduced into equation (40), it may be integrated to obtain $p(z)$. In particular, for an isothermal layer of the atmosphere it leads to the exponential dependence

$$p = p_0 e^{-\frac{g}{R} (z-z_0)}$$

where $p_0$ is the pressure at height $z_0$.

### 3.4 General circulation over the earth

#### 3.4.1 Subtropical circulation

It is clear from the diagram shown in Fig. 17 that the equatorial region receives much more solar energy than the polar regions, and the air at the equator will become warmer than that at the poles.

Figure 17

As a result of this heating, the air will expand and rise. Intuitively, if the effects of rotation are neglected, one might expect the non-uniform heating of the atmosphere to cause the circulation shown as in Fig. 18 - rising motion in the tropics and descending motion at high latitudes (the Hadley cell).
When the rotation of the earth is considered, the Coriolis force is exerted on the moving particles of the atmosphere. If, in the circulation shown in Fig. 18, the air particle starts in the upper air toward the north pole, it is not affected by the deflecting force immediately since at the equator the force is zero (see equation (35)). But as the air advances northward, the deflecting force increases as $\phi$ increases, and by the time that it has reached latitude $20^\circ$ or $30^\circ$ N, the motion will have a marked eastward component. At about $30^\circ$ latitude the cooler denser air descends, closing what is also termed the Hadley cell, but, as distinct from the circulation assumed by Hadley, this circulation cell does not extend to the pole, but it is confined to the tropics as in Fig. 19 below.
At the latitude 30° N there is an accumulation of air which leads to a high-pressure region at the earth’s surface. At the surface on the southern side of this high pressure area the winds blow toward the equator, but again they are affected by the deflecting force. Thus they become northwest winds, rather than northerly winds, and form the northeast trade winds which can be seen in Fig. 19; see also Fig. 20 where mean meridional circulation in summer and winter is shown together with mean meridional pressure distribution.
The high-pressure belt in the vicinity of 30° N, often referred to at the horse latitudes, is called the sub-tropical high. The upward feature of the Hadley cell occurs at what is often called the *inter-tropical convergence zone* (ITCZ, see Fig. 20) where the strong upward motion of air is characterized by heavy precipitation in convective thunderstorms.

Figure 20
3.4.2 Subpolar circulation

A similar thermally produced circulation is found in the vicinity of the pole. The air subsiding as a result of cooling in the lower layers at the pole moves toward the equator in diverging currents. These northerly winds are deviated to the right by the deflecting force as they advance southward, and thus become east winds. When the air motion is in this direction, the deflecting force and the pressure gradient force balance and the motion is steady. The air slowly becomes heated and rises, returning aloft to the pole (see Figs 19 and 20).

On the northern side of the sub-tropical high pressure belt the winds blow from west to east, giving the westerlies of the temperate zone. Near the surface, the slow movement of the air across the isobars results in a northward flow of air which transports heat of the equatorial regions toward the pole. At the northern edge of the temperate zone it meets the polar air moving westward. At the junction of these two currents a low pressure zone with strong temperature variations forms. This region of a concentrated temperature gradient that separates subtropical structure from that of the subpolar regions is often called the polar front.

3.4.3 Jet streams

The previous considerations show that the circulation, which is mainly in the meridional plane occurs only near the equator. In mid-latitudes, because of the rotation of the earth, the motion acquires the zonal component (along the latitude circles) so that both components should be present. It appears, however, that the general increase of temperature southward in the troposphere leads to the motion which is mainly zonal, and there is relatively little meridional circulation. Thus, the most pronounced feature of the atmosphere’s circulation is a broad, deep belt of westerly winds with a high-speed core that forms the great winding polar-front jet stream.

It can be shown that the horizontal gradient of temperature produces vertical gradient of the geostrophic wind which is proportional to the magnitude of the temperature gradient (the “thermal wind” concept). Because it is generally colder toward the poles in the troposphere, it is expected that the wind will have an increasing westerly component with height. The mid-latitude jet stream lies above the region in which the tropospheric temperature gradient is concentrated. Since the tropospheric temperature gradient is maintained to the top of the troposphere, the westerlies continue to increase in speed reaching a maximum at the level of the mid-latitude tropopause.

The main concept of this section is summarized in the phrase: ”The westerlies increase with height because it’s colder toward the poles”.

3.4.4 Instability of the atmospheric circulation - Cyclones and anticyclones

If the temperature distribution produced by the radiative driving from the sun would be in thermal wind balance with balance with the purely zonal flow, it would give much stronger jet stream than that observed. Such a state is not observed, however. The observed circulation contains structures representing circulations around a high pressure (anticyclone) or low pressure (cyclone) systems.
The direction and speed of rotation around high and low pressure systems were discussed in Section 3.2.3 - naturally, the air has a cyclonic rotation (i.e. counterclockwise in the northern hemisphere) around a cyclone and anticyclonic rotation around an anticyclone. The above arguments explain the balance of forces in the steady state rotation of developed cyclone or anticyclone but does not explain why the zonally symmetric flow cannot exist. The theory that explains the appearance of cyclones and anticyclones is the so called baroclinic instability theory. According to this theory, a zonally symmetric flow can be unstable to small longitude dependent disturbances. The observed circulation contains fully developed disturbances that take the form of cyclones and anticyclones. The surface cyclone is associated with weather events of interest to both forecasters and the public.

4 Effects of convection and convective stability

4.1 Effects of convection

Many of the most interesting atmospheric phenomena are associated with strong or even violent vertical motion. Grouped under the general term of convection, these include cumulus development, thunderstorms, dust devils, and possibly tornadoes. These phenomena suggest that we attempt to determine the conditions that will yield strong vertical motions.

The principle of free convection can be stated in what is almost an everyday expression: one should say that hot fluid tends to rise and cold fluid tends to fall relative to each other (as a matter of fact, we used such a semi-intuitive principle in reasoning behind the source of large-scale atmospheric motions). The principle is simple but not complete. If it would be the full story, in the atmosphere, where the air warmed by contact with the surface, the warm air could not remain below the cold air above without convection occurring, as it does in a kettle full of water that is heated from below. However, the static conditions do exist in the atmosphere. Whether or not convection will occur depends on the "lapse" rate, i.e., the rate at which the temperature of the atmosphere decreases with height. Convection will only occur when the lapse rate exceeds a certain value. This value can be calculated by considering the temperature changes of a parcel of air that moves up or down "adiabatically", i.e., without exchanging heat with the air outside the parcel.

4.2 Convective stability

To test the atmospheric stability we consider the air parcel as insulated from the surrounding air: its mass does not mix with the mass of the ambient air and no heat flows between the parcel and its surroundings. Depending on the temperature distribution with height in the air surrounding the parcel, stable or unstable conditions may occur. Consider the parcel that moves upward. If the parcel moved to the higher level finds itself to be heavier than the surrounding fluid, gravitational forces will cause the parcel to descend back toward its original level, and in this case the equilibrium is stable. If, however, the parcel is lighter than its surroundings, the equilibrium is unstable.
4.2.1 Dry adiabatic lapse rate

Since the parcel before it has been moved had the same temperature as the ambient air and it becomes cooler, when it moves, due to an expansion, the rate, with which the temperature of the ambient air decreases with height, is of the crucial importance. If it decreases faster than the temperature within the parcel, the parcel will acquire the acceleration upward (unstable conditions). In the opposite case it will be accelerated downwards (stable conditions). In the intermediate case there will be no acceleration, and parcels displaced vertically simply remain where they are left (neutral conditions).

The temperature of the parcel, that moves upwards adiabatically, decreases (as the result of its expansion) according to the relation (it may be obtained with the use of a constancy of entropy, hydrostatic equation (40) and the equation of state (41)):

\[ dT = -\Gamma_d dz \]

where \( \Gamma_d \) is called the *dry adiabatic lapse rate*. In the case of a dry atmosphere

\[ \Gamma_a = \frac{g}{c_p} \]  

\( (c_p \) is a specific heat at constant pressure), which gives for \( \Gamma_d \) approximately 10 K km\(^{-1}\).

Thus, if the atmospheric lapse rate, \( \Gamma = -\partial T/\partial z \), will differ from the adiabatic lapse rate \( \Gamma_d \) the following cases are possible.

If \( \Gamma < \Gamma_d \), an air parcel that undergoes an adiabatic displacement from its equilibrium level will be negatively buoyant when displaced vertically upward so that it will tend to return to its equilibrium level and the atmosphere is said to be *statically stable* or *stably stratified*. In this case adiabatic oscillations of an air parcel about its equilibrium level occur that are referred to as *buoyancy oscillations*. The characteristic frequency of such oscillations can be derived as shown in the Appendix.

If \( \Gamma = \Gamma_d \), no accelerating force will exist and the parcel will be in neutral equilibrium at its new level. In this case the atmosphere is said to be *statically neutral* or *neutrally stratified*.

If \( \Gamma > \Gamma_d \), a parcel displaced upward will find itself with a temperature greater than that of its environment and will continue to rise - the *statically unstable* atmosphere.

4.2.2 Characteristic frequency of buoyancy oscillations

In order to make a calculation, we notice that, at any point, the net vertical force on the parcel is given by the Archimedian principle as the difference between the force of gravity on the parcel mass \( m_p \) and the mass \( m \) of the ambient air displaced by the parcel. Thus this force taken to be positive upward, is

\[ F = g(m - m_p) \]  

According to Newton’s second law, the acceleration of the parcel is therefore given by

\[ m_p \frac{d^2z}{dt^2} = F = g(m - m_p) \]
Division of both sides by the volume of the parcel produces the result

\[ \frac{d^2 z}{dt^2} = g \left( \frac{\rho - \rho_p}{\rho_p} \right) \]  

(47)

where \( \rho_p \) is the parcel density and \( \rho \) is the density of the ambient air. Clearly then, the parcel will accelerate upward if it is less dense than the ambient air, accelerate downward if it is more dense. Since the pressure of the air inside the parcel and that outside the parcel is the same, we may, in view of the equation of state for dry air (41), express equation (47) as

\[ \frac{d^2 z}{dt^2} = g \left( \frac{T_p - T}{T_p} \right) \]  

(48)

where \( T_p \) is the parcel temperature and \( T \) is the temperature of the surrounding air. Considering the parcel to be initially at level \( z = z_0 \) where the temperature is \( T = T_0 \) so that for a small displacement \( \delta z \) we can represent the parcel and environmental temperatures as

\[ T_p = T_0 - \Gamma_d \delta z, \quad T = T_0 - \Gamma \delta z \]  

(49)

Substituting \( z = z_0 + \delta z \) and equations (49) into (48) leads to the equation

\[ \frac{d^2 \delta z}{dt^2} = -N^2 \delta z \]  

(50)

where

\[ N^2 = \frac{g}{T} (\Gamma_d - \Gamma) \]  

(51)

is a buoyancy frequency.

5 Appendix: dimensional analysis

All physical quantities are measured through comparison with fundamental quantities (units) of one or another dimensional system. The fundamental quantities of which a physical variable is composed are variable to as its dimensions. For example, pressure is defined as a force per unit area, so its dimensions are

\[ [p] = \left[ \frac{F}{S} \right] = \frac{ML/T^2}{L^2} = \frac{M}{LT^2} \]

where the square brackets denote the operation of taking the dimensions of a physical quantity, and capital letters are used as symbols for our fundamental system of units: mass \( M \), length \( L \) and time \( T \). In thermodynamics, one more fundamental unit must be added - commonly, it is a unit of absolute temperature \( K \).

Because every physical quantity has attributes of both magnitude and dimension, any algebraic expression or a physical relationship between two or more quantities implies
that an equivalence of dimensions as well as an equality of magnitudes exists. Thus, the equation \( V = \sqrt{2gh} \) indicates not only that the velocity of free fall and the square root of twice the acceleration of gravity times the height of fall are equal in magnitude but also they are equivalent dimensionally, that is to say \([V] = [g]^{1/2}L^{1/2}\). Similarly, a physical equation composed from several terms must be dimensionally homogeneous, that is the dimensions of every term must be the same - force and energy cannot be summed.

The essence of the dimensional analysis is an endeavor to find, without going through a detailed solution of the problem, certain relations which must be satisfied by the various measurable quantities in which we are interested. The usual procedure is as follows. We first make a list of all the quantities on which the answer may be supposed to depend, we then write down the dimensions of these quantities, and then we demand that these quantities be combined into a functional relation in such a way that the relation would be dimensionally homogeneous.

We consider, as an illustration, the problem treated in Section 2.3.2: determining the time scale of the temperature perturbation decay in the atmosphere or the radiation relaxation time. The parameters, which the answer may conceivably depend on, are: a mass of the atmosphere per unit area \( M_A \), its specific heat at constant pressure \( C_p \), the mean equilibrium temperature \( T_e \) and the Stefan-Boltzmann constant \( \sigma \).

First, we will make a list of all the quantities with their dimensions.

<table>
<thead>
<tr>
<th>Name of Quantity</th>
<th>Symbol</th>
<th>Dimensions in SI</th>
<th>Dimensional Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Relaxation time</td>
<td>( \tau_R )</td>
<td>s</td>
<td>( T )</td>
</tr>
<tr>
<td>Atmospheric Mass</td>
<td>( M_A )</td>
<td>kg m(^{-2})</td>
<td>( ML^{-2} )</td>
</tr>
<tr>
<td>Specific heat</td>
<td>( C_p )</td>
<td>J kg(^{-1})K(^{-1})</td>
<td>( L^2 K^{-1} T^{-2} )</td>
</tr>
<tr>
<td>Equil. Temperature</td>
<td>( T_e )</td>
<td>K</td>
<td>K</td>
</tr>
<tr>
<td>S-B constant</td>
<td>( \sigma )</td>
<td>W m(^{-2})K(^{-4})</td>
<td>( MT^{-3} K^{-4} )</td>
</tr>
</tbody>
</table>

We are to find \( \tau_R \) as a function of \( M_A \), \( C_p \), \( T_e \) and \( \sigma \) such that the combination of \( M_A \), \( C_p \), \( T_e \) and \( \sigma \) on the right-hand side of this functional equation had the same dimensions as \( \tau_R \). In most cases, the functional equation relating the quantities appears to be algebraic. Therefore, to simplify our task, we will assume the algebraic relation from the beginning by writing it in the form

\[
\tau_R = M_A^\alpha C_p^\beta T_e^\gamma \sigma^\delta \quad (52)
\]

where \( \alpha, \beta, \gamma \) and \( \delta \) are constants to be determined. Applying the basic principle of the dimensional analysis to equation (52) yields

\[
[\tau_R] = [M_A]^\alpha [C_p]^\beta [T_e]^\gamma [\sigma]^\delta \quad (53)
\]

Substituting the dimensional formulas from the table for every unit into (53) and arranging the powers yields

\[
T = M^{\alpha+\delta} L^{-2\alpha+2\beta} T^{-2\beta+2\gamma-3\delta} K^{-\beta+\gamma-4\delta} \quad (54)
\]

36
By equating the powers for all the units on the right-hand and left-hand sides we arrive at the following system of algebraic equations for $\alpha$, $\beta$, $\gamma$ and $\delta$:

\[
\begin{align*}
\alpha + \delta &= 0 \\
-2\alpha + 2\beta &= 0 \\
-2\beta - 3\delta &= 1 \\
-\beta + \gamma - 4\delta &= 0
\end{align*}
\]  

(55)

Solving this system of equations yields

\[
\alpha = 1, \quad \beta = 1, \quad \gamma = -3, \quad \delta = -1
\]

(56)

Thus, equation (52) becomes completely defined and the desired relation is

\[
\tau_R = \frac{M_A C_p}{\sigma T^\beta}
\]

(57)

The value of $\tau_R$ determined via dimensional analysis differs from the exact value defined in (19) only by the factor $\frac{1}{4}$. 