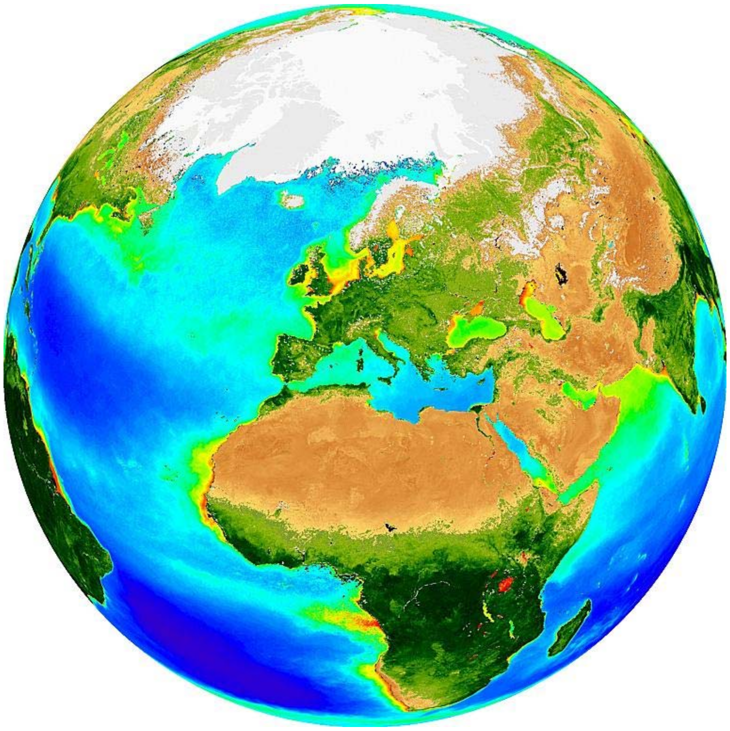
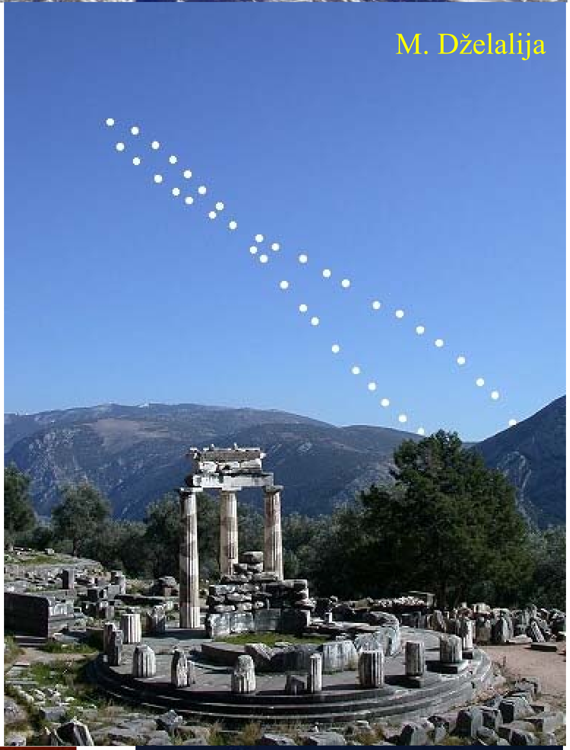
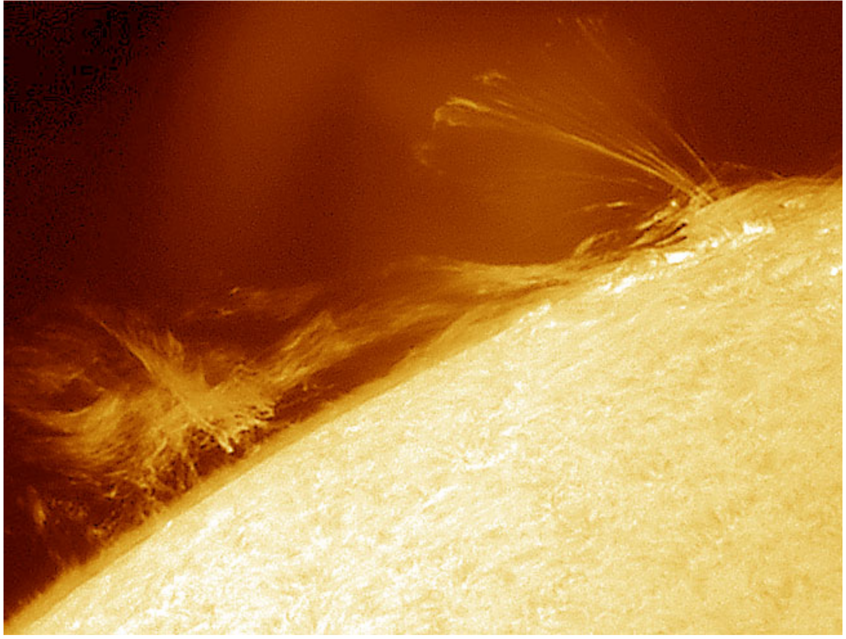
University of Molise, University of Split, Valahia University of Targoviste

ENVIRONMENTAL PHYSICS

M. Dželalija







Split, 2004

Aims and Objectives of the Course: Environmental Physics

This unit is designed to illustrate the many aspects of physics that pervade environmental processes in our everyday lives and in naturally occurring phenomena. It will be largely a descriptive course though some basic mathematical skills, necessary to gain a full understanding of some parts of the course. By the end of this course, a student will be able to:

- understand and apply the basic thermodynamics to the human environment,
- understand and apply the basic composition, structure and dynamics of the atmosphere,
- explain the workings of the hydrologic cycle and discuss the mechanisms of water transport in the atmosphere and in the ground,
- discuss specific environmental problems such as noise pollution, ozone depletion and global warming in the context of an overall understanding of the dynamics of the atmosphere,
- discuss the problems of energy demand and explain the possible contributions of renewables to energy supply, and
- understand and explain many other different topics of our environment.

Environmental Physics exam:

- Written:
  Written test consists of several conceptual and numerical questions. It will be marked and is assessed as 80 % of the course mark. Students are required to have a minimum of 60 % correct answers in written part.
- Oral:
  As the final part, oral exam consists of several conceptual questions to general problems in Environmental physics, such as:
  - Laws of Thermodynamics and the human body,
  - human environment and energy transfers,
  - noise pollution,
  - structure and composition of the atmosphere,
  - ozone in the atmosphere,
  - greenhouse effect,
  - global warming,
  - hydrosphere and hydrologic cycle,
  - water in the atmosphere and clouds,
  - cyclones and anticyclones, global convection and global wind pattern,
  - physics of ground, and
  - energy for leaving.

  This part is assessed as 20 % of the course mark.

Basic Environmental Physics Course book:

- Nigel Mason and Peter Hughes: *Introduction to Environmental Physics: Planet Earth, Life and Climate,* Taylor and Francis, 2001.

Contents

1 Introduction

Nature has amazing richness across the range of spatial and temporal scales at which processes and their interactions occur. We know from our own experience that winds blow and oceans move. Our Earth is not solid, if we define *solid* to mean forever immovable in space. The drift of continents can have the major influence on both climate and life. Except for local phenomena such earthquakes, landslides, and mountain glaciers, the time frame for major continent-scale Earth motions is thousands to millions of years. How the "solid" Earth interacts with air, water, and life is essential for understanding the Earth as a system, as knowledge of how and why the Earth system changes over geologic time allows us to calibrate our tools needed to forecast global changes.

The Earth is a marvellous place and since its formation 4.6 billion years ago both living and non-living entities have developed. In a global environment that is structured within the relationship between the land, the air, the oceans and the biosphere. However, to appreciate our environment it is necessary to understand the basic physical science that regulates its development.

In the past few decades the possible detrimental impact humanity is having on the planet has caused increasing concern. As humanity has sought to improve its so called *prosperity*, it has often done so by exploiting the Earth's abundant natural resources. The discovery of the ozone hole, the first signs of industrially induced global warming, the widespread phenomenon of acid rain and the growing evidence of health problems caused by urban pollution, have attracted world-wide attention from both social and political commentators. Debates have taken place, in the scientific and political communities, about the actual evidence for such phenomena and what actions should be taken to alleviate such impacts. The environmental problems cannot be addressed comprehensively by looking through the limited lens of only one of the traditional disciplines established in academia, such as, physics, chemistry, biology, engineering, or economics. It is hard to solve most global problems without the detailed information that those disciplines provide, but the study of Earth systems science suggests that we also need to find appropriate ways to integrate high-quality disciplinary work from several fields. To understand and assess the possible dangers to the Earth caused by the exploitation of its resources and the development of industry, a new branch of science, Environmental physics, has evaluated in the past 30 years, which is dedicated to study of 'Environmental Issues'.

Environmental physics is an interdisciplinary subject that integrates the physics processes in the following disciplines:

- the atmosphere,
- the biosphere,
- the hydrosphere, and
- the geosphere.

Environmental physics can be defined as the response of living organisms to their environment within the framework of the physics of environmental processes and issues. It is structures within the relationship between the atmosphere, the oceans (hydrosphere), land (lithosphere), soils and vegetation (biosphere). It embraces the following themes:

- human environment and survival physics,

- built environment,
- urban environment,
- renewable energy,
- remote sensing,
- weather,
- climate and climate change, and
- environmental health.

To understand how any specific environmental process evolves, it is necessary to appreciate that all these processes are interdependent. The formation and mobility of clouds, for example, illustrate just one aspect of a number of global environmental processes and require the study of:

- solar radiation transformations and the radiation balance,
- phase changes in the water cycle,
- monitoring physical phenomena,
- exchanges between the Earth, the oceans, the atmosphere and the biosphere,
- transport phenomena, especially mass and thermal energy transfer.

However, it is important to appreciate that the principles and lows of physics are in evidence in many different environments and govern how all species live on the Earth.

The environment may be defined as the medium in which any entity finds itself. For example, for a cloud, its environment may be the region of the atmosphere in which it is formed, while for a plant, it is a field in which it lies, and for a whale it is the sea in which it swims. Thus, it is informative to discuss environmental issues within the context of the surroundings in which an object finds itself. In the following chapters the applications of the principles of physics to environmental processes and problems will discussed and put in the context of current environmental issues.

2 The human environment

Living organisms have to adapt and survive in a variety of environmental conditions, including hot and cold climates. They are thermodynamic entities characterized by energy flows both within the body, and between the body and its environment. For people to survive, the core body temperature has to be maintained within a narrow temperature range of 35-40^{0}C. The rate of these energy transfers and the mechanism of thermoregulation are governed by the following laws and concepts of physics:

- Laws of thermodynamics,
- Principles of entropy, enthalpy, and the Gibbs free energy,
- Principles of conduction, convection, radiation and evaporation,
- Newton's law of cooling, and
- Wien's and Stefan-Boltzmann radiation laws.

Human beings have managed to live in all the different environments present throughout the Earth: from the wastes of the Arctic to the deserts of Mongolia, from the jungles of Africa to the coral islands of the Pacific. Mammals, including humans, have the remarkable ability to maintain a constant body temperature, in spite of dramatic changes in environmental conditions. They are called *homeotherms*. They sustain their body temperatures by adjusting the rate of energy transfer and energy production (transformation).

In contrasts, certain animal species, such as reptiles and amphibians, have core body temperatures that respond to environmental temperatures. Such animals are called *poikilotherms*. Both homeotherms and poikilotherms respond to conditions in a variety of physiological and behavioural mechanisms. In cold weather we put on 'warmer' clothing, while bears have fur. In hot weather we wear thinner clothing.

Planet Earth provides many environmental and ecological contexts for living things to survive and develop. For life to be sustained we should not only be concerned with the chemistry and biochemistry of metabolic reactions, but also with the physics of thermal processes. It is necessary to discuss the laws of thermodynamics to see how they apply to the body's energy metabolism.

2.1 Laws of thermodynamics

2.1.1 First law of thermodynamics

The general formulation of the First law of thermodynamics for an ideal gas is that

$$\mathrm{d}Q = \mathrm{d}U + \mathrm{d}W,$$

where $\mathrm{d}Q$ is the energy supplied to or extracted from a closed system, $\mathrm{d}U$ is the change in the internal energy of the system, and $\mathrm{d}W$ is the work done by the system. The First law is an expression of the principle of the conservation energy, and the internal energy refers to the total kinetic energy (chaotic motion, also rotation and vibration) of all the atoms and molecules comprising the gas and their vibrational potential energy.

Another useful concept is that of *enthalpy*. Enthalpy, H, is the heat content of a system and is a thermodynamic state function[1], which is related to the internal energy, U, the pressure, p, and volume, V, in the form:

$$H = U + pV.$$

Often it is more useful to speak of the enthalpy change, dH, of a chemical reaction. In the situation where no external work is achieved, dW = 0. Thus, dH = dU. This enthalpy change can be assessed by the amount of energy generated (or absorbed) in a reaction.

2.1.2 Second law of thermodynamics

An internal combustion engine and the human body have similarities in that they function as heat engines. A heat engine is a means of extracting useful mechanical work from a system with a temperature difference between its interior and its environment. The heat engine is, therefore, a useful analogy for our body. The operation of any heat engine is governed by the Second law of thermodynamics, originally stated by the French physicist Sadi Carnot. He proposed that in a heat engine work done by a system is obtained from the energy transferred between one body at a higher temperature and another at a lower temperature. It cannot of itself go in the opposite direction unless acted upon by an external agency. It is often expressed in terms of *efficiency*:

$$e = (T_1 - T_2)/T_1,$$

where T_1 is the higher temperature and T_2 is the lower temperature. The importance of the Second law is that it defines the direction in which thermal energy will flow.

2.1.3 Third law of thermodynamics

If a cup of tea at 60^0C is left in a room at 20^0C, it will gradually cool. The temperature of the tea will decrease from a higher to a lower level. Without any external input, it is not possible for its temperature to rise. That is, the process is irreversible. This is simple example of the Second Law of Thermodynamics. Similarly, for a human, without the external agency of food as a source of chemical energy and the impact of solar radiation, the body's temperature would fall, and with starvation, death would result. The temperature difference between our bodies and the local environment not only sustains us, but also allows us to produce useful mechanical work. Since the temperature of the body is usually greater than that of the surroundings, energy flows out of the body into the environment. The process is irreversible, and the environment gains energy, dQ, at this environmental temperature, T. This provides us with a definition of *entropy change*, dS:

$$\mathrm{d}S = \mathrm{d}Q/T.$$

The entropy change for the entire system is greater than zero, $\mathrm{d}S_{\mathrm{body}} + \mathrm{d}S_{\mathrm{environment}} > 0$.

---

[1] A thermodynamic state function is characteristic and descriptive of the thermodynamic state of a system. Examples include internal energy, temperature, entropy.

Ludwig Boltzmann defined entropy in terms of probability, W, of the number of ways in which energy distributions can be generated:

$$S = k \cdot \ln W,$$

where k is Boltzmann's constant, $k = 1.38 \cdot 10^{-23}$ J/K. W tells us that the probability of obtaining certain outcomes in a particular energy distribution depends on the number of ways it can be distributed.

The *entropy*, S, of a system can be determined if use is made of the Third law of thermodynamics, which assumes that at absolute zero temperature, 0 K, entropy is zero. For example, the absolute entropy of 1 mole of pure water as ice at 0^0C and as liquid at 0^0C is 41 J/K and 63 J/K, respectively. The entropy change, $\mathrm{d}S = (63 - 41)$ J/K = 22 J/K, we can determine also using the expression, $\mathrm{d}S = \mathrm{d}Q/T$. This is precisely the quantity of energy, $\mathrm{d}Q$, extracted from the surroundings, that brings about the change of phase, solid to liquid. Since $\mathrm{d}Q = mL$, where m is the mass of 1 mole of water, $m = 0.018$ kg, and L is the latent heat of fusion, $L = 333000$ J/kg, then the entropy change will be

$$\mathrm{d}S = \mathrm{d}Q/T = 0.018 \cdot 333000/273 \text{ J/K} = 22 \text{ J/K}.$$

In a physical sense, entropy is, therefore, a measure of the 'disorder' of a system. Most natural processes, like the cooling tea or the decreasing radioactivity resulting from a radioactivity source, are irreversible. If a process goes in its 'normal' manner, the entropy of the system increases. If it proceeds in the opposite direction, the entropy decreases.

2.2 Laws of thermodynamics and the human body

The Second law governs changes that act in the direction in which entropy increases. We will now see through a detailed examination how the laws of thermodynamics relate to the energetics of the body.

2.2.1 Energy and metabolism

Metabolism is the total of all the chemical processes that occur in the cells of a body. It consists of anabolism in which molecules are built-up and catabolism in which enzymes break down the food consumed through hydrolysis, and at the cellular level involves the process of *phosphorolysis*.

The *basal metabolic rate* (BMR) is the rate at which a fasting, sedentary body generates sufficient energy to achieve the vital functions of respiration, maintaining the body's temperature, the heart beat and production of tissue. BMR is approximately equal to the metabolic rate while sleeping, and while resting most of the energy is dissipated as thermal energy. BMR can be calculated using direct calorimetry or by use of a spirometer, which measures the oxygen consumption per unit time. In the calorimetric method, a person is placed in a chamber through which there are pipes carrying water. The amount of energy produced can be determined from the energy gained by the water passing through the pipes. In spirometry, the energy generated is related to the amount of oxygen taken in during respiration, and thus the metabolic rate measured.

For a man, BMR is about 170 $kJm^{-2}h^{-1}$, and is 155 $kJm^{-2}h^{-1}$ for woman. Thus for a man of about 1.8 m^2 surface area, this would make 7300 kJ per day or about 85 W. During the day, in addition to the basal requirements, energy will be required for mechanical work and physical exercise. Typical energy dissipations are:

- sleeping: 75 W,
- sitting: 80-100 W,
- walking: 150-450 W,
- running hard: 400-1500 W.

The average person needs an additional 4200 kJ for a 'normal' working day; thus making a total requirements of about 12000 kJ per day. Since carbohydrates provide about 17 kJ/g, proteins 38 kJ/g and fats 17 kJ/g, by adjusting the various amounts this figure can be attained.

Metabolism involves the chemical processes in the body in which energy is transferred between various chemical compounds and in which thermal energy is generated. If the rate of metabolic reactions increases, then the rate of energy generation also increases. People require certain amounts of energy to achieve certain tasks. This has implications, for example, for athletic performance and survival. A sedentary man can produce energy of the order of 0.07 $kJkg^{-1}min^{-1}$ (which is about 80 W for a 70 kg-man).

2.2.2 Thermodynamics and the human body

Humans breathe in oxygen and eat food, which is composed of carbohydrates, fats, oils and proteins. The carbohydrates are converted into glucose, the proteins into amino acids, and the fats into fatty acids. The blood then transports these, together with oxygen, to the cells, where enzymes, which are biological catalysts, convert the glucose into pyruvic acid, through the process of glycolysis. The fatty and most of the amino acids are converted into acetoacetic acid. These are changed into acetyl Co-A, and with further ocidation, produce adenosine triphosphate (ATP), carbon dioxide and water. This entire process is called the *Krebs Cycle*.

ATP generates the energy that could be potentially used by the cells. The energy is stored in the phosphate bond when adenosine diphosphate (ADP) is transformed to adenosine triphosphate, and is dissipated when ATP is converted into ADP. When the energy is released it takes the form of heat, and this is transferred by the blood, around the body. Energy is also transferred from the cells to their surroundings by conduction because of the thermal gradient created between the cells and their environment.

Thermal energy loss from the body is achieved through conduction, convection, radiation and evaporation from the skin, and through respiration. In humans energy is transferred to the surroundings at the skin's interface with the air outside. Since cooling results, this implies that a temperature gradient exists between the body's core and the skin's surface. This body temperature is stable as long as the production of energy equals the energy loss.

Living organisms are also thermodynamics entities, in which thermal processes are characterized by energy flows and fluxes both within the body, and between the body and its environment. For people to survive, the core body temperature has to be maintained within a narrow temperature range of 35-40^{0}C. The normal body temperature is 37^{0}C. However, this is the core temperature. There is a temperature gradient as one moves away from the core. Hence, not only is there a temperature drop between a person and their external environment, but also there is one within the body.

What is the relevance of physics in a discussion on energy and metabolism? Physics underpins the biochemical processes that provide us with energy. Although this chapter is not concerned with biochemical processes, we will look at how physics, through the laws of thermodynamic, relates to metabolic processes.

2.2.3 First law of thermodynamics and the human body

For an energy balance, under steady-state conditions where the core body temperaure and the ambient temperature remain constant, the quantity of energy produced will equal the quantity of energy dissipated. Hence, it is possible to invoke the First law of thermodynamics to the body. The total energy produced in the body is called the *metabolic rate* ($\mathrm{d}M$). It is related to the total metabolic energy production of the body ($\mathrm{d}H$), and the external work done by the body ($\mathrm{d}W$), by the expression:

$$\mathrm{d}M = \mathrm{d}H + \mathrm{d}M.$$

There is an obvious analogy if this is compared with the expression for the First Law. $\mathrm{d}H$ varies from one person to another, and depends on the activity engaged in, and on, the body's surface area. On average, the body's surface area is about 1.84 m^2, the average male mass is 65-70 kg and the average female mass 55 kg. For a sedentary person the metabolic rate is about 100 W, and is 400 W for a person engaged in heavy physical work. Energy transfers in metabolic processes are governed by the First law of thermodynamics, and the Law can be applied to determine the quantity of energy that can be generated. If no mechanical work is done ($\mathrm{d}W = 0$), then the chemical energy input will be transferred as thermal energy, i.e. $\mathrm{d}H = \mathrm{d}U$ is the energy produced by the oxidation of the chemicals, and the total $\mathrm{d}H$ is the mass of the chemicals oxidized times $\mathrm{d}H$.

2.2.4 Second law of thermodynamics and the human body

If a metabolic process occurs in a particular direction, does it also occur in the reverse manner? The Second law helps to explain both the direction and attainment of equilibrium in metabolic processes, and now it can be seen that the entropy change can assist in the understanding of the direction that a metabolic process will take. It also tells us whether that particular process will occur.

In the oxidation of glucose amount of energy is 'wasted'[2]. Thus, the process is not 100 % efficient. The 'waste' is the production of energy as heat – a prerequisite for maintaining the core body temperature. This 'wasted' energy is the driving force for the direction in which a metabolic process should go. The idea of potential energy is a useful direction it will occur. If one drops a body, its potential energy is transformed into kinetic energy and then into heat, sound and, possible, light. As a result, the entropy of the surroundings (i.e. Universe) will increase. The change in entropy is a function of the energy transferred from the body. Thus, we have

$$\mathrm{d}S_{\mathrm{environment}} = -\,\mathrm{d}Q_{\mathrm{body}}/T,$$

---

[2] Here 'wasted' means: some of the energy is transformed into internal energy of a systems.

where the conditions, at the cellular level, are assumed to be isothermal. This equation implies that if energy is lost from a body, indicated by the negative sign, then $\mathrm{d}S_{\text{environment}}$ will increase.

If the entropy tells us the direction of a spontaneous change, it would be useful to develop the criteria, from energetic considerations, for the propensity of a system to provide ‘free energy’ to do useful work. The criterion is provided by the idea of the *Gibbs free energy*, *G*.

Since the First law o thermodynamics can be represented as

$$\mathrm{d}Q = \mathrm{d}U + p\cdot\mathrm{d}V,$$

where p and $\mathrm{d}V$ is the pressure and the change in volume, and the Second law by

$$\mathrm{d}S = \mathrm{d}Q/T,$$

then

$$T\cdot\mathrm{d}S = \mathrm{d}U + p\cdot\mathrm{d}V,$$

where $\mathrm{d}S$ is the change in entropy related to a change in energy, $\mathrm{d}Q$. Therefore, the change in internal energy is

$$\mathrm{d}U = T\cdot\mathrm{d}S - p\cdot\mathrm{d}V.$$

This is the Gibbs equation. It incorporates the idea of temperature, it embraces the Zeroth, the First and Second laws of thermodynamics. The temperature is a central characteristic of a thermodynamic system. It can be applied to any physical systems (and to the biophysical system that is the human body).

Using the definition of enthaly, $H = U + pV$, we have

$$\mathrm{d}H = \mathrm{d}U + p\cdot\mathrm{d}V + V\cdot\mathrm{d}p = T\cdot\mathrm{d}S + V\cdot\mathrm{d}p.$$

Now, we define the Gibbs free energy, G, as

$$G = U - TS + pV.$$

Thus, it is easy to see that the change in the Gibbs free energy is

$$\mathrm{d}G = \mathrm{d}H - T\cdot\mathrm{d}S.$$

The above equation gives the maximum possibility of a process achieving work. G is not free energy in the sense that it comes from nothing. It implies that it is the energy available for work. $\mathrm{d}G$ influences the possible direction of a metabolic process. If it is negative, then free energy is released and the process will occur. If it is positive, it will not.

2.3 Energy transfers

To feel warm, whether one is in the house or walking outside, is a question of energy conservation, but the underlying principle is that of an energy balance, and for this to be achieved energy exchange is necessary. Energy can be transferred from one point to another by the following mechanisms: conduction, convection, radiation, and evaporation. The physics of each of those mechanisms will be discussed in turn.

2.3.1 Conduction

Thermal conduction is the process by which energy can be transferred between two points in a material at different temperatures. In solids this is achieved in two ways: (i) through molecular vibrations transferring energy through the crystal lattice, and (ii) through the mobility of free conduction electrons throughout lattice.

In semiconductors both components contribute to energy transfer because there are less free electrons, but in insulators the above (i)-way predominates. These lattice vibrations, which are called *phonons*, generate elastic acoustic standing waves which pass through the material at the speed of sound for that material.

J. Fourier had discovered that the *rate of flow of thermal energy*, dQ/dt, through a material depends on the cross-sectional area, A, the length or thickness of the material, L, and the difference in temperature between the two sides, $\Delta T = T_1 - T_2$. This can be expressed as

$$dQ/dt = - kA \cdot \Delta T/L,$$

where k is the thermal conductivity of a material. The effectiveness of a material as an insulator can be determined by measuring its thermal conductivity. Good thermal conductors like copper have a high thermal conductivity, e.g. 380 $Wm^{-1}K^{-1}$, while poor conductors like water have a low thermal conductivity, i.e. 0.59 $Wm^{-1}K^{-1}$. The ratio of the temperature difference divided by the length is called the *temperature gradient*.

The minus sign in the above equation is significant. It shows that the flow of energy is from the region at the higher temperature to that at the lower temperature, i.e. it flows along the temperature gradient. It implies that energy flow is unidirectional. The equation is true for steady-state conditions, i.e. when the two temperatures are stable and that the thermal energy input equals the thermal out, or for the short time interval.

Example: Fourier's law of thermal conduction.

A rambler is walking up a steep hillside in January. He is wearing clothing 1 cm thick, his skin temperature is 34^0C and the exterior surface is close to freezing at 0^0C. Determine the rate of flow of energy outwards from his body, through thermal conduction, when:

- it is fine dry. Assume that the thermal conductivity for clothing, under dry conditions, is 0.042 $Wm^{-1}K^{-1}$.
- It has been raining heavily and the rambler is soaked. The thermal conductivity is now 0.64 $Wm^{-1}K^{-1}$. Assume that the walker has surface area of 1.84 m^2.

Solution:

- Apply the Fourier's law of thermal conduction.
  $dQ/dt = - kA \cdot dT/L = -0.042 \cdot 1.84 \cdot 34/0.01$ W = - 263 W.
- Applying the Fourier's law again, dQ/dt = - 4004 W.

When clothing becomes wet, it becomes a better conductor for the outward dissipation of energy because water has a higher thermal conductivity than dry clothing. This is why jeans are inappropriate trouser-ware for strenuous outdoor pursuits in wet weather.

2.3.2 Convection

Convection occurs when thermal energy is transferred by the motion of a fluid. The fluid can be either a liquid or a gas. The air in close proximity to any living body will warm up due to heat radiating from the body and expand, so becoming less dense and rise. Colder and denser air will take its place down and a convection current will be set up. A similar process occurs but on much larger scale in the Earth's atmosphere.

There are two types of convection: (i) natural (when fluids move without forcing), and (ii) forced (when the fluid is forced, such as blowing over a hot cup of tea). Of particular importance for the human environment is forced convection and the Newton's law of cooling provides a physical model of it.

2.3.2.1 Newton's law of cooling

A number of factors will influence the rate of convection from an object in a fluid, including the temperature of the object, the shape, the size, the temperature of the fluid and the type of flow relative to the object. Newton determined that the rate at which energy is lost from a body $\mathrm{d}Q/\mathrm{d}t$ is directly proportional to the difference between the body's temperature, T, and environmental temperature, T_0, i.e.

$$\mathrm{d}Q/\mathrm{d}t = -\, kA\cdot\Delta T.$$

This equation represents Newton's law of cooling, where k is a proportionality constant whose magnitude depends upon the nature and surface area of the body, and is called *convective energy transfer coefficient*. For a plate in still air k is 4.5 $\mathrm{Wm^{-2}K^{-1}}$ and is about 12 $\mathrm{Wm^{-2}K^{-1}}$ when air flows over it at 2 m/s. Strictly, this law applies to objects cooling in a draughter, such as blowing over a cup of tea. It does not apply to humans as the process of metabolism attempts to maintain the body temperature at a constant value. However, walking against a strong wind or in a wind tunnel is a reasonably good approximation of this law.

Example: Fluid flow and convection.

A student volunteers to take part in the following simulations of convective energy loss:

- he is placed in a wind tunnel in which air, at -2^0C, blows through at 40 km/h.
- He is placed in a flow of water, which is at 12^0C. The velocity of flow is 0.5 m/s.

In each case calculate: (i) the convective energy transfer coefficient, and (ii) the convective energy transfer flux, $\mathrm{d}Q/\mathrm{d}t$. Assume A = 1.8 $\mathrm{m^2}$ and the skin temperature is 31^0C. Assume convective energy transfer coefficient k = 44.8 $\mathrm{Wm^{-2}K^{-1}}$ for 40 km/h speed of flow, and k = 34.4 $\mathrm{Wm^{-2}K^{-1}}$ for 0.5 m/s.

Solution:
Applying Newton's law for cooling, we have:

- $\mathrm{d}Q/\mathrm{d}t$ = 2661 W.
- $\mathrm{d}Q/\mathrm{d}t$ = 1176 W.

2.3.3 Radiation

Radiation plays an important role in the energy balance of human beings. It is the process in which energy can be transferred in the form of electromagnetic waves from one point to another through a vacuum. All objects release energy in the form of electromagnetic waves. The best absorbers usually make the best emitters of radiation and these are called *black-body*. Human beings emit radiation in the infrared band.

There are two experimental laws that usually we use to explain radiation: Wien's and Stefan-Boltzmann laws. Wien's law tells us about the wavelength, λ_m, that the body at the temperature, T, radiates with the maximum intensity

$$\lambda_\mathrm{m}\cdot T = b,$$

where b is a constant, and for the black-body is $b = 3\cdot 10^{-3}$ m·K.

Stefan-Boltzmann law explains the total amount of radiation energy per second (or as power) from a black-body, and was discovered to be proportional to the forth power of the temperature, T, and the area, A, of the surface emitting the radiation

$$P = \sigma A T^4,$$

where σ is Stefan's constant, and for the black-body is $\sigma = 5.7\cdot 10^{-8}\ \mathrm{Wm^{-2}K^{-4}}$.

No real body is a perfect black-body radiator, and to distinguish between a perfect black-body and real bodies the idea of *emissivity*, ε, has been introduced. Thus, for the radiation power we have:

$$P = \varepsilon\, \sigma A T^4.$$

An object release energy in an environment which itself is radiating energy. Then the object will absorb energy. This means that the net rate of radiation emitted is the difference between the energy emitted by the object and the energy absorbs from its surroundings.

Example: Radiation and the Stefan-Boltzmann law.

A person sitting reading a book releases radiant energy of between 70 and 100 W. Calculate how much energy the person is radiating. Assuming that the emissivity for the human body is 0.5, with an average surface temperature of 35^0C, that room temperature is 20^0C, and the body surface area is 1.8 m^2.

Solution:
Applying the Stefan-Boltzmann law, $P = 83.6$ W.

2.3.4 Evaporation

Anyone engaged in strenuous physical activities increases their metabolic rate, and to maintain a constant core body temperature losing energy by conducting, convection and radiation is insufficient. Evaporation constitutes a fourth method for obtaining an energy balance for energy transfer. Evaporation is important in understanding weather and climate generally, and clouds in particular.

Evaporation is the process whereby a liquid can be transformed into a vapour. This implies a phase change, or change of state, and it is an example of a latent heat change, in which the evaporative energy loss depends on the mass of the liquid and the energy required to vaporize the liquid

$$Q = mL,$$

where Q is the energy extracted or supplied to bring about a phase change, m is the mass of liquid to be vaporized and L is the specific latent heat of vaporization. For pure water, $L = 2.25 \cdot 10^6$ J/kg, but sweat, which is 99 % water with sodium chloride as solute, is an electrolyte with $L = 2.43 \cdot 10^6$ J/kg. For humans and animals L depends on temperature.

As sweat vaporizes, the energy to achieve this has come from the hotter body. The net result is a cooling effect, and the surface temperature of the body decreases. The rate of evaporation depends on the surface area, the temperature difference, the humidity and (therefore, the difference in vapour pressure) the rate of sweating and the velocity of air flow.

In physics, there are several cases of transport phenomena caused by a difference in a physical variable over a length. For example, we have seen that in thermal conduction energy will flow from the hot end of a bar to the colder end because of the temperature gradient which has been set up (Fourier's law). In the case of evaporation, the mechanism operates through a difference in vapour pressure of water. Evaporation incorporates the passage of water vapour through the body's boundary layer. The rate of evaporation from the body can be expressed as:

$$\mathrm{d}Q/\mathrm{d}t = hA(p_\mathrm{S} - p_0),$$

where h is the evaporative energy transfer coefficient, A is the area of the skin's surface, p_S is the water vapour pressure adjacent to the skin and p_0 is the water vapour pressure in the surroundings air. The vapour pressure at the skin's surface depends on the environmental humidity and the sweating rate.

For non-evaporative energy loss (such as conduction and convection) the driving mechanism is the temperature gradient between a body and its surroundings, but for evaporation it is the vapour pressure gradient. In region of very high humidity, evaporative cooling becomes less efficient because this gradient changes. Evaporative energy loss occurs during perspiration and respiration. For a thermally comfortable and resting person, evaporative water loss can be about 30 g/h. In comparison, exercising hard can produce a rate of sweating of 200 g/h. Under normal conduction, energy loss by perspiration is the dominant factor.

2.4 Survival in cold climates

As we grow older our metabolic rate decreases. This means that for some people energy cannot be produced as fast as it is dissipated. Thus, in cold weather, without adequate insulation and heating, *hypothermia* can set in. In contrast, in very hot summers babies and the elderly can become vulnerable to heat stress.

Wind, like water, is a fluid, and the wind flows can either be laminar or turbulent. For natural and forced convection, the convective airflow patterns can either be laminar at lower speed, or turbulent at higher speeds. The wind can have a major impact on the human body. While walking in the hills bareheaded and with the wind blowing you may have a slight headache. This is the result of the wind facilitating the flow of thermal energy out from your body through the head. The wind-chill temperature is the temperature that prevails from a particular combination of wind speed and air temperature. As wind speed increases, wind-chill temperature is depressed. Obviously, the walker has to appreciate the impact of wind-chill and be prepared accordingly, especially in terms of clothing.

Hypothermia is the lowering of the normal body temperature and is an example of a switching mechanism from negative to positive biological-physical feedback. Under normal conditions, if a person feels cold the body readjust and in the temperature range 37-35^{0}C as the body's temperature falls, the body works hard to produce the extra energy to compensate. This is negative feedback. However, below 35^{0}C the body cannot generate sufficient energy as fast as it is being lost. The body's temperature continues to fall, and the more the body tries to supplement the body's temperature, the faster it falls. Hypothermia has several stages:

- mild hypothermia: core temperature starts to fall from 37 to 36^{0}C.
- Moderate hypothermia: temperature falls from 35 to 32^{0}C. Thermoregulation begins to breakdown.
- Severe hypothermia: temperature falls from 31 to 30^{0}C.
- Acute hypothermia: 29^{0}C, i.e. death.

In addition to wind-swept areas, the sea is another context in which people can become victims of hypothermia. Immersion in cold, swirling water can facilitate the loss of energy by forced convection. Lean people are more susceptible.

2.5 Survival in hot climates

If energy transfers from the environment into the human body, without some dissipating mechanism the body's temperature will increase, to the point of heat stress, and beyond to heat stroke and death. The balance can be brought to an equilibrium steady-state, by perspiration in humans and panting in some animals. Energy from the body is used to vaporize sweat and a cooling effect results. Working in very hot climates, such as deserts, can results in a water loss of 10-12 kg/day. In hot climates evaporative cooling becomes the dominant mechanism of energy transfer and thermoregulation.

Like hypothermia, heat stress or *hyperthermia* has stages, all of which are accompanied by continuing dehydration:

- sweating and vasodilatation: the core effect of muscular work is to destabilize the energy balance with the result that body temperature increases with vasodilation and sweating.
- Heat cramp: results from continuing dehydration and may affect certain muscles, particularly in the legs and stomach. There may also be some dizziness.
- Heat exhaustion: occurs when water and salt are not consumed to replenish those lost in sweat. Both dehydration and a continuing rise in the core body's temperature may be in evidence.
- Heat stroke: this is very serious and can occur when the body's temperature is in excess of 41^{0}C. The thermoregulatory system, especially the process of vasodilation, starts to collapse, with the result that the body cannot effectively dissipate energy and, therefore, the body's temperature continues to rise. During this time the cardiovascular system is put under increasing stress and the flow of

blood to the brain can be reduced. Unconsciousness may then result and death in the extreme case.

Questions:

- During a warm day a walker loses 1.5 kg of perspiration by evaporation. Given that the latent heat of vaporization is 2.25 MJ/kg, calculate how much thermal energy is required to achieve this. (**Answer**: $3.375 \cdot 10^6$ J)
- A person inhales in one breath $5 \cdot 10^{-4}$ m^3 of dry air at atmospheric pressure and 20^0C. The air is then warmed to the body core temperature of 37^0C in the lungs. If the person takes 12 breaths per minute, calculate the heat transferred per minute to the air from the body. Assume that there are no pressure changes in the inhaled air during respiration. (**Answer**: $1.2 \cdot 10^2$ J/min)
- A women whose mass is 55 kg has a metabolic rate of 9 W/kg when she is running up a hill at an angle 5^0 to the horizontal with a constant speed of 6 m/s. At what rate is she gaining potential energy? At what rate is she using energy? (**Answer**: 290 W; 500 W)

3 Noise pollution

Noise is not often thought of as a *pollutant*, but unwanted sound (noise) can seriously degrade the quantity of life. The acceptance of noise by people obviously depends on the individual, but it has been legislated on levels of sound acceptable to the community. The level of noise deemed to be acceptable is dependent upon:

- the type of environment: acceptable levels of surroundings noise are affected by the type of activity. A library, for example, has different requirements to those on a factory floor.
- Frequency structure: different noises contain different frequencies and some frequencies are found to be more annoying than lower frequency rumbles.
- Duration: a short period of high level noise is less likely to annoy than a long period.

Different people have different hearing sensitivities, but average values can be measured and provide a map of the sound that the human ear can detect. The *threshold of hearing* is the weakest sound that the average human hearing can detect. The threshold varies slightly with the individual, but it is remarkably low. There is also high threshold, the *threshold of pain*, which is the strongest sound that the human ear can tolerate.

Absolute measurements of sound intensity can be expressed in either $\mathrm{Wm^{-2}}$ or in sound pressure, Pa, but such units do not correspond directly to the way in which the human ear responds to sound levels. Since the human ear has a non-linear response to the energy content of sound, a logarithmic scale is used to describe the response of the ear. It is converted to *sound level* measured in decibels, dB

$$L = 10\cdot\log(I/I_0) \text{ or } L = 20\cdot\log(p/p_0),$$

where $I_0 = 10^{-12}\ \mathrm{Wm^{-2}}$ and $p_0 = 2\cdot10^{-5}$ Pa are the values for the threshold of hearing, I and p the intensity and pressure of the sound being measured. The faintest audible sound (at 1000 Hz) is rated as 0 dB. Normal speech is 50 dB, road traffic 70 dB and an aircraft engine at close range is about 120 dB.

Example: Noise pollution.

If one jet causes a sound level of 120 dB on take off, what is the sound level of three such jets taking off simultaneously?

Solution:
For one jet: L = 120 dB gives us $I = 1\ \mathrm{Wm^{-2}}$.
For three jets: $I = 3\ \mathrm{Wm^{-2}}$ giving $L = 10\cdot\log(3/10^{-12}) = 124.8$ dB.

3.1 Domestic noise and the design of partitions

The sound level in a room is determined by a number of factors. If there is no internal source of noise, then these factors are:

- direct transmission (from neighbouring rooms or from outside),
- Flanking transmission (by walls, floor or ceiling), and

- contact noise (through floors or ceilings).

The degree of direct transmission from outside is determined by the insulating properties of the separating walls. The *transmission loss* (T_L) is defined as:

$$T_L = 10 \cdot \log(P_I / P_T),$$

where P_I is the power incident on the wall and P_T is the power transmitted through it. Another, more common measure is the *noise insulation*, R. This is defined as

$$R = 10 \cdot \log(I_I / I_T),$$

where I_I is the intensity of noise incident on a wall and I_t is the intensity transmitted. These two measures can be related. The power incident on a partition of area A is $P_I = I_I A$. The power transmitted into the receiving room is equal to the rate at which energy is absorbed in the receiving room (assuming that no energy is transmitted onwards) and, so, $P_T = I_T W$, where W is the *sound absorption* of the receiving room. Thus, putting these together,

$$T_L = R + 10 \cdot \log(A/W).$$

R is usually measured for a band of frequencies. To reduce noise level, R must be large. To achieve this, it is necessary to: (i) use absorbing materials (such as foams), (ii) have hollow (cavity) walls, (iii) double glazed windows. It is necessary to avoid resonances (i.e. must not set up a standing wave in the cavity).

For a planar, non-porous, homogeneous, flexible wall it can be shown that

$$T_L = 20 \cdot \log(f \rho_A),$$

where f is the frequency and ρ_A is the mass per unit area. This is often called the mass law. This usually gives an overestimate of the transmission loss since it ignores the effect of the stiffness of the panel. The point is that the panel can support flexural waves that, above a critical frequency, can be excited by the sound waves. The incoming sound wave couples effectively to these flexural waves which transmit sound through the panel. A properly designed panel must ensure that these flexural waves (which are unavoidable) cannot couple to sound waves within hearing range. Even if problems with direct partitions are solved (sometimes by using double partition walls), there remain two other basic difficulties:

- reflected (flanking) noise. This can be reduced by:
  - covering walls with absorbing materials (tapestries) or objects that break up the wave front (pictures, china ducks etc),
  - internal cavity walls,
  - special dishes in walls which use destructive interference to remove reflected waves,
  - remove all direct paths (i.e. fit draught excluders).
- Contact noise. This can be reduced by using sprung floors. Also it helps to place vibrating equipment on shock absorbers (e.g. rubber mats).

Questions:

- A washing machine generating 90 dB of noise is turned on at the same time as a ghetto-blaster generating 100 dB is on in the room. What is the total noise level in the room? (**Answer**: 100.4 dB)

4 Atmosphere and radiation

4.1 Structure and composition of the atmosphere

The Earth's atmosphere is a gaseous envelope, retained by gravity, surrounding the Earth. The density falls rapidly with height: 90 % of the mass of the atmosphere is contained within the first 20 km, 99.9 % within the first 50 km. The atmosphere becomes increasingly tenuous with increasing height until at about 1000 km. It merges with interstellar space. Compare with the radius of the Earth (about 6400 km), 99.9 % of the Earth's atmosphere is in a ring whose thickness is 0.8 % of the radius of the Earth. This implies that almost every property of the atmosphere is strongly anisotropic. For example, the temperature decreases with height at a rate of 6 K/km (up to about 15 km), but the strongest horizontal temperature gradients (associated with warm or cold fronts) are about 0.05 K/km. Away from the surface, the isopleths[3] are nearly horizontal and the atmosphere consists of horizontal layers. In particular, the atmosphere is divided into regions characterised by their temperature. Each layer is called a *sphere* and the boundary between layers is called a *pause*. The boundary is named from the lower layer. The layers are:

- troposphere (0-10 km). This is the lowest and contains 80 % of the mass. Almost all the weather is confined to the troposphere. In particular, it contains the clouds. The temperature falls linearly with height until at the top of the troposphere the temperature is approximately -50^0C.
- Stratosphere (10-50 km). Above the tropopause the temperature begins to rise again until at about 50 km the temperature is about $+10^0$C. The upper part of the stratosphere contains ozone – an essential molecule for life on Earth since it filters out (harmful) UV radiation.
- Mesosphere (50-85 km). Above the stratopause the temperature falls rapidly to about -80^0C. This is the coldest region of the atmosphere.
- Ionosphere (100-200 km). This is an intensely ionised region of the atmosphere and the temperature rises rapidly. Solar UV ionises the molecules of the atmosphere

  $$h\nu + \mathrm{AB} \rightarrow \mathrm{AB}^+ + \mathrm{e}^-.$$

  The ionosphere reflects radio waves and is also the region of the aurorae seen in high latitudes in the northern (aurora borealis) and southern (aurora australis) hemispheres. Intense visible and UV lines caused by electron (or proton) collisions

  $$\mathrm{e}^- \text{ (energy } E_\mathrm{i}) + \mathrm{AB} \rightarrow \mathrm{AB}^* + \mathrm{e}^- \text{ (energy } E_\mathrm{f}),$$

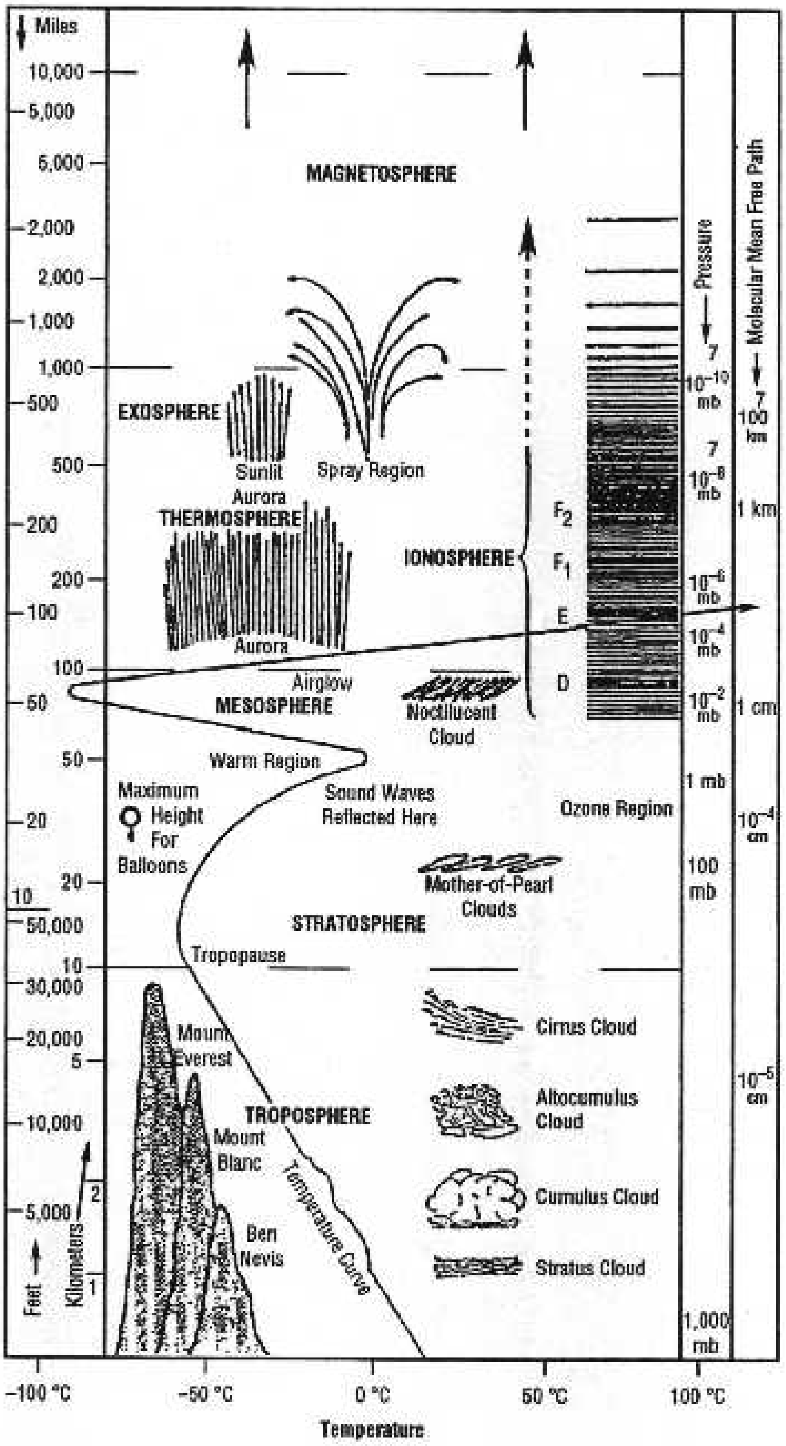
  where AB can be O_2 or N_2. The energy $\Delta E = E_\mathrm{i} - E_\mathrm{f}$ given up by the electron excites the molecule to the AB* state. This then decays back to the ground state, releasing a photon of frequency $\nu\,(h\nu = \Delta E)$.
- Thermosphere (200-500 km). Temperature rises rapidly and varies strongly with the time of day, degree of solar activity and latitude. Variation between 400^0C and 2000^0C possible. Minimum temperatures are at sunrise, and maximum at about 1400^0C. However, the pressures are very low, and there is little heat

---

[3] Lines or surfaces of equal quantity such as isobars (equal pressure) or isotherms (equal temperature).

transfer. The idea of a *temperature* is becoming increasingly meaningless (high vacuum). Better to think in terms of molecular speeds.

- Exosphere (500- about 1000 km). Atoms and molecules are sparse and can escape into space.
- Magnetosphere (above 1000 km). In this region the Earth's magnetic field interacts with the solar wind and traps charged particles (electrons and protons) in the, so called, *Van Allen belts*.



In the troposphere, stratosphere and mesosphere, mixing mechanisms ensure that there is essentially a constant chemical composition with an N_2/O_2 ratio of about 4:1. This uniform composition gives these three regions the collective name of the homosphere. Above 100 km (the mesopause) the ratio changes as a function of height. This (upper) region is called the heterosphere.

The atmosphere contains a number of trace gases whose concentration does change with height. The most important gases are:

- ozone – concentrated in the stratosphere (sometimes called the ozonosphere) at a height of 20-30 km,
- water – very variable concentration throughout the homosphere,
- carbon dioxide – important greenhouse gas in the upper troposphere (0.03 % of the atmosphere).

4.1.1 Residence time

The *residence time*, τ, is the mean lifetime of a gas molecule in the atmosphere. This is as important as the concentration. Indeed, when considering pollutants, it is the most important parameter. It is given by

$$\tau = \langle m \rangle / \langle F \rangle,$$

where $\langle m \rangle$ is the total average mass of the gas in the atmosphere and $\langle F \rangle$ is the total average influx (or out-flux). $1/\tau$ is, therefore, the rate of turnover of the gas. If τ is small, then the molecule (or atom) is in the atmosphere for only a short time. For example, it may be reactive and so will not be homogeneous throughout the atmosphere, but will be localised in particular areas (acid rain is an example). Another possibility is that the compounds may be part of a cycle (water is an example; the hydrologic cycle circulates water from sea to clouds to rain to sea again with a residence time of about 10 days). Using the residence time as a measure, we can divide the constituents of the atmosphere into three categories:

- permanent: τ is very long (order of a few million years), e.g. N_2, O_2, rare gases (CO_2 but see also later),
- semi-permanent: τ is of order months to years, e.g. CH_4, N_2O, CO, CFCs, and
- variable: τ is of order days to weeks, e.g. ozone O_3 (cycle in the stratosphere), H_2O (cycle in the troposphere), SO_2 and H_2S (acid rain), NO_2, NH_3 (car exhausts) but also part of the nitrogen cycle.

4.1.2 Photochemical pollution

Pollutants can be divided into two categories: *primary* and *secondary*. Primary pollutants are the chemical species emitted directly, while secondary are those formed from primary pollutants by local chemistry. The greatest damage often comes from the secondary pollutants rather that the primary. For example, while sulphur dioxide might be emitted from a power station its conversion into sulphuric acid is more damaging to the local environment.

The major hazardous pollutants in urban conurbations are carbon monoxide, nitric oxide, sulphur dioxide, ozone, particulate matter and smog. There is a new form of smog, caused by photochemistry. *Photochemical smog* is induced by sunlight which powers the following series of reactions:

- $NO_2 + \gamma\ (\lambda < 385\ \text{nm}) \rightarrow NO + O$
- $O + O_2$ + (any molecule) $\rightarrow O_3$ + (any molecule)
- $O_3 + NO \rightarrow O_2 + NO_2$

The molecule in the second step can be any species. It is needed to take the reaction energy away from the system. There is no overall change in the concentration of the reactants. The effect of the set of reactions is simply to convert the energy of the radiation into the kinetic energy of molecules. However, in the steady state, there will be finite concentrations of the intermediate species. In particular, the concentration of ozone is in the range 0.005-0.05 particles per million (ppm). This ozone near the surface is poisonous in itself and also reacts with olefins (carbon compounds with carbon-carbon double bonds; they are present in car exhausts) such as ethylene ($H_2C{=}CH_2$) and propylene ($H_2C{=}CH{-}CH_3$) to give products such as formaldehyde ($H_2C{=}O$), acrolein ($H_2C{=}CH{-}CH{=}O$) and (with NO_2) peroxy-acetylnitrate ($H_3C{-}(C{=}O){-}O{-}O{-}NO_2$). These irritate the eyes and nose. Since ozone is also produced from photocopiers, most people will have smelt it at one time. Single exposure of a few hours of 80-120 ozone particles per billion (ppb) have been found to cause respiratory problems.

4.1.3 Atmospheric aerosol

An aerosol contains solid or liquid particles in suspension (e.g. dust particles). In the atmospheric aerosol, particles arise from:

- combustion – forest fires or industry (soot),
- gas phase reactions (particles of sulphates or nitrates),
- dispersion of solids (wind and water erosion of rocks),
- dispersion of salts from the sea (sea-spray), and
- volcanoes.

Typical concentrations are 10^3 particles cm^{-3} (over the ocean), 10^4 particles cm^{-3} (over the country), 10^5 particles cm^{-3} (over cities). The size ranges from aggregates of a few hundred molecules (diameter about 1 nm) to the largest particles (about 10 μm). The removal of the aerosol from the atmosphere depends on the size of the particles:

- Aitken nuclei have diameters less than a 0.1 μm. They make up the majority of particles able to nucleate cloud droplets in a cloud chamber (a fact originally used to count them). Once away from where they are produced, the numbers are rapidly reduced by collision (due to Brownian motion) and coagulation to give larger particles. These nuclei make up about 20 % of the aerosol mass.
- Large nuclei (diameters 0.1-1 μm). Although much less numerous than Aitken nuclei they make up about 50 % of the aerosol mass. Fall speeds are still much lower than the ability of up-draughts to maintain them in the atmosphere, but the concentration above the cloud base is greatly reduced because they are excellent nuclei for cloud droplets.
- Giant nuclei (diameters greater than 1 μm). Although the number density is low, they comprise at least 30 % of the aerosol mass. They are mainly produced by fine dust lifted by the wind off arid surfaces. As with large nuclei, they are very effective nuclei for droplets. Also, they are efficiently swept out of the atmosphere by falling rain (the pressure wave around a falling drop is not strong enough to

push them away). Also, their falling speed is now great enough for gravitational settling (sedimentation) to be important when the up-draughts are low.

4.2 Atmospheric pressure

We know that in climbing mountains both pressure and temperature decreased with increasing height. It can be shown that the pressure (in the troposphere) decreases by the following expression:

$$p = p_0 \cdot e^{-gh/RT},$$

where p_0 is the atmospheric pressure at the surface (h = 0 m), and p the pressure at a height h.

Since the pressure falls exponentially with height, 90 % of the mass of the atmosphere is contained within the first 21 km, and 99.9 % in the first 50 km. The pressure therefore drops from 10^5 Pa at the Earth's surface to 10^4 Pa at 20 km, and to 10^2 at 50 km. The temperature also falls with altitude and it is called the *laps rate* (dT/dz).

4.3 Escape velocity

The Earth's atmosphere is retained by gravity. We recall the idea of an escape velocity. Consider a rocket of mass m launched from the surface of the Earth at a velocity fast enough so that it just escapes from the gravitational influence of the Earth. Then the work done is

$$W = m\cdot(\gamma M/R),$$

where M and R are the mass and radius of the Earth, respectively, and γ is the gravitational constant. To escape, the kinetic energy of the rocket must be balanced by the work done in liberating the rocket from the Earth's gravitational potential such that

$$\frac{1}{2}mv^2 = \gamma\frac{mM}{R}.$$

Thus,

$$v = (2\gamma M/R)^{1/2},$$

where v is the velocity of the rocket. Since $g = \gamma M/R^2$, the starting velocity (which is the escape velocity) can be written as $v = (2gR)^{1/2}$. Putting in values appropriate for Earth (g = 9.81 m/s^2; R = 6400 km), we get v = 11,2 m/s.

Since this is independent of the mass of the object; this analysis is equally appropriate for a molecule. The distribution of speeds of a molecule in the atmosphere is given by the Maxwell speed distribution. The most probable speed is given by $v_p = (2k_BT/m)^{1/2}$, where k_B is Boltzmann's constant. At the temperature of 288 K, $v_p(O_2)$ = 387 m/s, and $v_p(N_2)$ = 414 m/s. These will both be retained. But $v_p(He)$ = 1094 m/s and $v_p(H_2)$ = 1550 m/s. This is much faster. However there is a finite probability that these atoms and molecules will have speed grater than 11 km/s. The lighter the gas, the larger the proportional of atoms or molecules having such speed. Moreover, at the top of the atmosphere the temperature is much higher. These gases can leak away. Over 10^9 years, He and H_2 have been lost to space.

4.4 Ozone

Ozone is only minor constituent of the Earth's atmosphere forming 0.2 % of the terrestrial atmosphere mass, such that if all the atmospheric ozone was collected at the Earth's surface it would form a ring only 3 mm thick around the Earth. However, ozone is essential to the sustaining of both plant and mammalian life. It is the presence of ozone in the atmosphere that shields the Earth's surface from harmful solar UV radiation through its ability to absorb all solar radiation with wavelengths < 293 nm. Uniquely among these molecules in the Earth's atmosphere ozone has a strong *absorption band* between 210 and 300 nm. Hence, ozone filters out the Sun's ultraviolet radiation below 300 nm preventing these wavelengths from reaching the Earth's surface.

Ozone is restricted to a thin layer in the Earth's atmosphere, the ozonosphere, with a maximum concentration between 20 and 26 km above the Earth's surface. It is formed through the combination of atomic oxygen, O, and molecular oxygen, O_2. Atomic oxygen is formed by the photo-dissociation of O_2 at around 100 km by solar radiation with $\lambda < 175$ nm, the process for which may be summarized as:

$$h\nu(\lambda < 175\ \text{nm}) + O_2 \rightarrow O + O.$$

The free oxygen atoms may then combine with oxygen molecules in a three-body collision (to conserve energy) to form ozone:

$$O + O_2 + M \rightarrow O_3 + M,$$

where M is any atom or molecule (e.g. O_2, N_2) capable of absorbing the excess energy liberated in the exothermic chemical reaction.

Most of the ozone is produced in the equatorial regions where the amount of solar UV light is maximized. Ozone formed over these latitudes is then transported towards the Poles where it accumulates. Hence, ozone concentrations show significant seasonal fluctuations and tend to be at their highest in late winter and early spring. Once formed, there are two natural destruction mechanisms for ozone:

- photo-dissociation:

$$h\nu + O_3 \rightarrow O_2 + O,$$

- and collisional dissociation:

$$O + O_3 \rightarrow 2O_2.$$

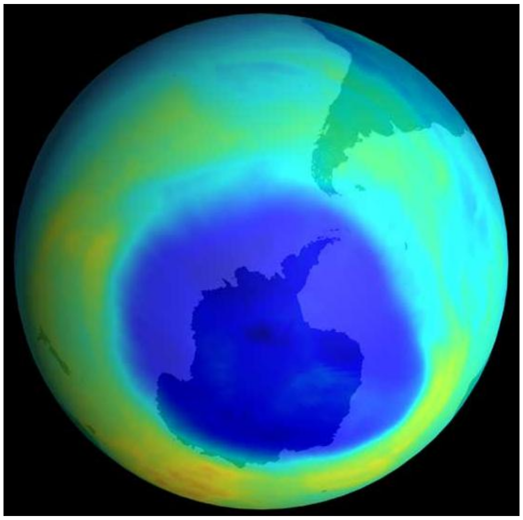
The latter being a net result of complex catalytic chemical reactions either natural or more recently man-made. Natural catalytic destruction occurs through the presence of the OH radical formed by photo-dissociation of water vapour in the atmosphere:

$$\begin{array}{c} OH + O_3 \rightarrow HO_2 + O_2 \\ \underline{HO_2 + O \rightarrow OH + O_2} \\ \text{Net: } O + O_3 \rightarrow 2O_2. \end{array}$$

The above processes explain why there is little ozone formed naturally in the troposphere where water is abundant and the flux of solar radiation (< 175 nm) is low, while lack of O_2 prevents ozone formation in the atmosphere above the stratosphere. As a result, the formation of ozone is restricted to altitudes between 10 and 50 km, where the above processes have formed a stable equilibrium producing ozone concentrations of around 10 ppm. However, it is a fragile system and introduction of new chemicals (e.g. man-made pollutants transported from the troposphere) in the stratosphere can disturb this equilibrium and lead to rapid decrease in ozone concentrations.

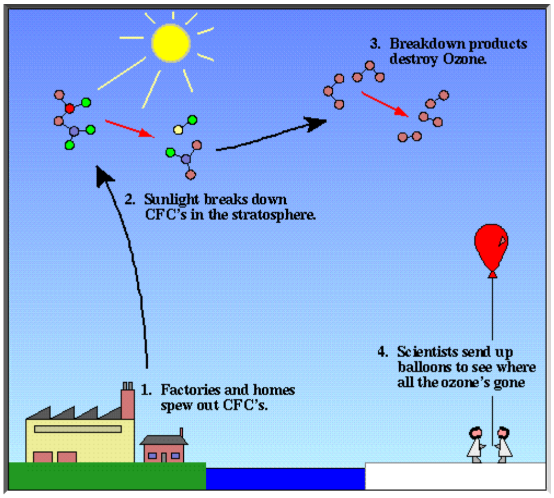
4.4.1 Ozone hole

In 1985 it was discovered a sharp reduction in ozone concentrations above Antartica; this was soon termed the *ozone hole*.



Satellite monitoring of the ozone hole has shown that it is expanding and its depth increasing.

An explanation for such dramatic ozone depletion was first proposed by Molina and Rowland in 1994. They suggested that the uncontrolled release of the chlorofluorocarbons (CFCs, CF_nCl_{4-n}, usually $n = 2$) used in refrigerators and aerosol cans into the terrestrial atmosphere would lead to the catalytic destruction of ozone.



Although chemically stable within the troposphere, in the stratosphere the CFC's can be broken down by solar radiation and release chlorine atoms which in turn form ClO catalytic species such that:

$$\begin{array}{c} ClO + O_3 \rightarrow ClO_2 + O_2 \\ \underline{ClO_2 + O \rightarrow ClO + O_2} \\ \text{Net: } O + O_3 \rightarrow 2O_2. \end{array}$$

the ClO molecules being conserved. The mechanism is far more destructive than those naturally occurring mechanisms involving the OH radical as one ClO radical may destroy several hundreds ozone molecules before it itself is removed from the ozone chemical code.

4.4.2 Ozone in polar region

We need to explain why the holes appear in the Polar Regions and in particular over Antarctica. This is due to the wind systems in these regions. The polar winter vortex is very stable due to the large land mass of Antarctica. The intense cold (183 K at heights of 15 km) produces clouds of ice crystals on which the reactions occur. These heterogeneous reactions are far more efficient than the gas-phase reactions discussed above (for a start you do not need three-body collisions; the ice crystals can carry away the excess energy). CFC's are trapped in the polar vortex and concentrated before being released in the polar spring as the solar radiation illuminates these latitudes once again. As the air warms up, the vortex breaks down. Ozone loss has now been observed at high northern latitudes. The Arctic vortex is much smaller than the Antarctic one and the mechanisms of ozone destruction are probably different. The stratospheric aerosol is expected to play a larger role and bromine compounds are more important.

One unexpected side-effect of ozone depletion is the cooling of the stratosphere. Since ozone absorbs UV radiation, the presence of ozone in the stratosphere heats it up. Thus ozone loss cools it done. This partially offsets the increased greenhouse.

4.5 Terrestrial radiation

Not all the solar radiation incident on the Earth is absorbed by the atmosphere or the ground. A large fraction is reflected back into space from the cloud tops and the ground. The fraction of radiation reflected (and hence lost to the earth) is called the *planetary albedo*, often denoted by the symbol *a*. A rough breakdown of the energy budget of the earth is as follows:

- 33 % of the solar flux is reflected back into space, mostly from cloud tops (26 %), but some from the ground (2.5 %) and from dust and aerosols in the atmosphere (4.5 %),
- 22 % are absorbed by the atmosphere (including 3 % absorbed by clouds),
- 32.5 % are scattered by the atmosphere. Of these, 28 % subsequently reach the ground and 4.5 % are scattered back into space (the dust and aerosols contribution to the albedo mentioned above). This scattering is responsible for the blueness of the sky,

- 17 % reach the ground directly. Of these, 14.5 % are absorbed and 2.5 % are reflected back into space.

Note that 45 % reach the surface of the Earth one way or another. The annual average rate of input of solar energy per unit horizontal area at the top of the atmosphere is 336 Wm^{-2}. So, the solar input to the Earth's surface is 151 Wm^{-2}. This would rapidly heat the Earth's surface to an unbearable temperature if it were all totally absorbed. The rate of absorption of solar energy must be balanced by a re-emission of this energy until a steady state is achieved. This point applies with equal force to the 19 % absorbed by the atmosphere. This re-radiation is thermal radiation (i.e. heat). The effect of the Earth is to perform a partial conversion of the solar visible and UV radiation (coming in) into the terrestrial infra-red radiation (going out).

4.6 Earth as a black body

The Sun emits as though it were a black-body with a (surface) temperature of about 6000 K, and the Earth, clouds, behave as black-body with appropriate values of temperature.

Assuming that the Earth emits terrestrial radiation as a spherical black-body of radius R_E and temperature T_E, then from the Stefan-Boltzmann law, the total power output of the planet is $P = 4\pi R_\mathrm{E}^2 \sigma T_\mathrm{E}^4$. The rate of absorption of radiation from the Sun is $S(1 - a)\pi R_\mathrm{E}^2$ where a is the albedo (in fact, the integral absorptivity of the Earth, but the symbol is conventional) and S the solar constant. These must balance, and so:

$$4\pi R_\mathrm{E}^2 \sigma T_\mathrm{E}^4 = S(1 - a)\pi R_\mathrm{E}^2$$

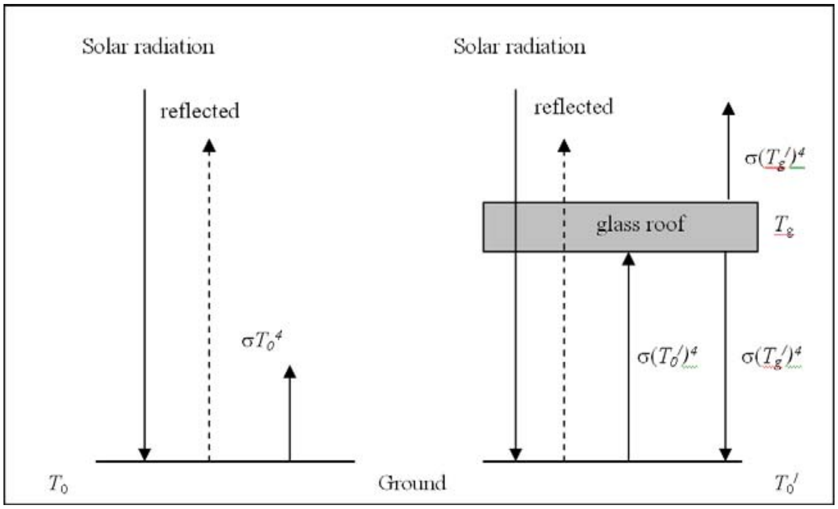
from which we can evaluate the effective temperature of the Earth as

$$T_E = \left(\frac{S(1-a)}{4\sigma} \right)^{1/4} .$$

Note that T_E is independent of the radius of the body. Given that $a = 0.31$ and $S = 1353\ \mathrm{Wm}^{-2}$, we find $T_\mathrm{E} = 255$ K (-18^{0}C). This is a bit on the cold side. The mean temperature at the surface of the Earth is 288 K (+15^{0}C). The problem is that we have neglected the atmosphere entirely. In particular we have neglected the internal energy transfers between the atmosphere and the surface. In effect we have considered the interface between the atmosphere and space and ignored the interface between the atmosphere and the ground.

4.6.1 Greenhouse effect

The elevation of the temperature at the surface of the Earth from –22^{0}C to +15^{0}C is due to the greenhouse effect. The name comes from the fact that the mechanism is in many ways analogous to the way that a greenhouse works and the atmosphere can be considered as a greenhouse roof. Consider the diagram below.



On the left is the situation that we have considered before which gives a temperature that is far too low. Now place a 'glass roof' above the ground. The glass lets in the solar radiation but prevents the infra-red radiation from radiating directly to space. This heats the roof to a characteristic temperature T_g and the roof then radiates both to the ground and out into space. The ground is therefore receiving more energy than before and its temperature will rise until a new equilibrium is achieved in which both the ground and the 'roof' emit as much as they absorb. Thus, in this state, the emission upwards from the 'glass roof' must equal the emission upwards from the ground without the 'roof' (since the balance with respect to outer space must not change. This implies that $T_0 = T_g^{/}$. If we now consider the state of the ground, then the net energy (per m^2) given to the ground directly by the sun must be σT_0^4 and the energy radiated back from the glass roof must be σT_0^4 (since $T_0 = T_g^{/}$). Thus we have

$$2\sigma T_0^4 = \sigma (T_0^{/})^4$$

from the energy balance on the right. Thus the new ground temperature must be

$$T_0^{/} = T_0\, 2^{1/4} = 298 \text{ K } (25^0\text{C}).$$

This is rather too high. However, our model of the atmosphere is still very simple. We have assumed that the 'glass roof' does nothing except pass solar radiation through and block terrestrial radiation from the surface completely. Moreover, we have assumed that we can treat the 'glass roof' as though it were another black body. If this were true, we could certainly treat the upward and downward radiation from the 'roof' in exactly the same way. However, if this is not the case (and in fact it is not) we have a more complicated balance to calculate.

4.6.1.1 Greenhouse gases

We require molecules that can absorb the infra-red radiation coming from the Earth (i.e. wavelengths in the range 5-25 μm) but not in visible region. Molecules that do this are H_2O (absorbs in bands < 4 μm, an intense band at 6.3 μm and a strong band greater than 9 μm), CO_2 (strong absorption band at 13-17 μm). O_3 absorbs in both regions, intense narrow band at 9.7 μm. O_3 is only important in the stratosphere. The essential components of the ‘roof’ are CO_2 and H_2O. The effectiveness of a gas in contributing to the greenhouse effect is measured by the global warming potential. This is defined as the added surface warming per unit molecule of the gas in the Earth’s atmosphere (referred to the effect of CO_2). It is measured in terms of number of molecules of CO_2 equivalent.

4.6.2 Global warming

The huge expansion of industrialisation has led to the prediction that man is now altering the global radiation balance by enhancing the natural greenhouse effect. Industrialisation leads to increased CO_2 emissions and hence increased CO_2 concentrations in the lower atmosphere. Between 1770 and the present, CO_2 concentrations have increased from 280 ppm to 350 ppm (i.e. an increase of 25 %). Moreover, the greatest increase has been in the last 50 years and the concentrations of CO_2 are still increasing. Much work has been done on climate modelling, modelling the atmosphere and then seeing the effect of changing parameters like CO_2 (or other greenhouse gas) concentrations, mean solar flux and so on. Moreover, if one wants to make predictions about the local climate, the calculations get even worse. The basic ingredients include:

- modelling the dynamics of the atmosphere by Newton’s equations (including such things as mass conservation, hydrostatic effects and so forth,
- equation of state of the gases in the atmosphere,
- thermodynamic effects (latent heat of water vapour for example),
- clouds and their effect on the radiation equilibrium,
- convection in the atmosphere,
- coupling atmospheric effects to the oceans (effects of ocean currents and ice caps),
- long term systems (such as El Nino in the Pacific), and
- effect of different terrains (desert, forests, etc).

Also, one must remember that CO_2 is not the only greenhouse gas although it is the largest contributor (about 70 %). Sizeable contributions also come from methane (23 %) and N_2O (7 %):

- ethane concentrations before 1800 were about 0.8 ppmV (ppm by volume) but since then the concentration has more than doubled and is increasing at about 1 %/yr. Although the concentrations of methane are much lower than CO_2 (2 ppmV versus 350 ppmV), methane is 7.5 times more absorbing than CO_2. Methane emissions come from fossil fuels (100 Mtons/yr), rice paddies (60 Mtons/yr) cows (80 Mtons/yr) forest fires and misc. (40 Mtons/yr). Methane is removed by chemical reaction with the OH radical and has a lifetime of about 11 yrs.
- Nitrous oxide (N_2O) concentrations are 0.3 ppmV rising at 0.25 % per year. The concentration is now about eight times that of the pre-industrial period. Sources are fertilisers, chemical industry (nylon production) and deforestation. Lifetime in the atmosphere about 150 yr.
- CFCs Concentrations low, but one CFC molecule is 500-10000 times more effective than CO_2. It is possible that CFCs account for up to 20 % of global warming. Effect partially offset by the destruction of ozone.
- Ozone increased smog make O_3 a tropospheric species as well as a stratospheric one. Not clear what the effect will be.

- Indirect production from CO, NO, NO_2, which form (and also destroy greenhouse gases). The chemical feedback loops possible are very complex

The relative contribution of various gases to global warming since 1800 is: CO_2 (55 %), CH_4 (15 %), CFC-12 (21 %), N_2O (4 %), O_3 in troposphere (2 %), and others (3 %).

Various computer models are beginning to converge on a common answer. As the concentration of CO_2 increases, the global average temperature will increase by a few tenths of a degree in the next few years and then by about 1.5^0C over the following 70 years. The slow start is due to the thermal inertia of the oceans. This is a global average. There are marked local variations. The northern hemisphere warms about twice as much as the southern and there are strong variations from continent to continent. The amount seems small, but it leads to major climactic changes including desertification in some areas, increased rainfall in others (for example the monsoon region will spread). This has important implications for agriculture.

As an example of the consequences of global warming let us consider the rising of the sea level. As it gets warmer, thermal expansion leads to the expanding of the oceans and hence the rising of sea levels (there is also a small effect from melting ice). The amount of thermal expansion strongly depends on the temperature of the water. Cold water expands only a little (the maximum density of water is at 4^0C; ice floats). At 5^0C a one degree rise in temperature causes a volume increase of 1 in 10^4. At 25^0C a one degree rise causes a volume change of 3 in 10^4. This corresponds to 3 cm rise in the sea level. The predictions of sea level rise over the next century are in the range 20-50 cm. This is enough to make major changes to the coastline.

Questions:

- If carbon monoxide (CO) is in the air at 1 ppm (particle per million) what are the number of CO in 1 m^3, and what is the mass of CO in 1 m^3? (**Answer**: $2.7 \cdot 10^{19}$ molecules per m^3; 1.25 mg/m^3)
- If the ratio of ^4He to ^{40}Ar entering the Earth's atmosphere from radioactive decay during the past 10^9 years is close to unity, why is the present-day ^4He measured in the Earth's atmosphere much less than that of ^{40}Ar?
- Discuss why ozone is produced mainly in the equatorial regions and hence how it is circulated around the Earth.
- Why is ozone restricted to such thin layer in the Earth's atmosphere between 20 and 50 km?
- Explain why it is imperative that there are greenhouse gases in the Earth's atmosphere.
- Calculate the density of air at the summit of Mount Everest where the pressure and temperature are $3.13 \cdot 10^4$ Pa and -38.5^0C, respectively. (**Answer**: 0.465 kg/m^3)
- More than 2/3 of the Earth's surface is covered by water. Solar radiation reaches the Earth as roughly parallel beams of mean intensity 1.4 kW/m^2. Of this radiation, on average 50 % reaches the surface. The oceans have a mean reflectivity of 7 %, while they reradiate 35 % of the energy they absorb. using these data, estimate the mean annual rainfall over the whole of the Earth's surface, expressed as a depth in milimetres. (**Answer**: 1000 mm)
- The value of the solar constant can be found using Stefan's law. Use the data: $\sigma = 5.67 \cdot 10^{-8}$ Wm^{-2} K^{-4}, $T_{\mathrm{Sun}} = 5770$ K, $R_{\mathrm{Sun}} = 6.98 \cdot 10^8$ m, $R_{\mathrm{Earth's\ orbit}} = 1496 \cdot 10^8$ m) to calculate its value. (**Answer**: 1368 W/m^2)

5 Water

The density of water is at a maximum at 4^0C. Ice I (there are at least ten phases of ice) is an open structure, held together by hydrogen bonds, and is less dense than liquid water. This is useful since the oceans freeze from the top down, not the bottom up. The latent heat of fusion (i.e. the heat required to convert 1 kg of ice to water without changing the temperature) is 0.334 kJ/kg. The latent heat of vaporisation (i.e. the heat required to convert water to vapour without a change of temperature) is 2300 kJ/kg.

5.1 Hydrosphere

The total volume of water on the Earth is about 1284 M km^3; 97 % of this is in the oceans. If it were spread evenly over the Earth, the planet would be covered to a depth of 2.8 km; 2.25 % is locked up in the polar ice-caps and in glaciers; about 0.75 % is in soil, lakes and rivers; 0.035 % is in the atmosphere. To give some idea of what this amount is: if all the water vapour in the atmosphere were instantly converted to rain, the total rainfall (averaged over the Earth's surface) would be about 3 cm. Yet the annual average rainfall here is 90-100 cm. There is a hydrologic cycle.

5.2 Hydrologic cycle

Water is cycled between the oceans and the atmosphere. As always, the cycle is driven by the Sun. Most of the water vapour in the atmosphere (84 %) comes from the oceans. Transpiration from plant leaves accounts for most of the rest. The Sun heats the water in the oceans (and on the land surface), giving evaporation. The warm, moist air rises, expands (under reduced pressure higher in the atmosphere) and cools. The water vapour condenses to form clouds. The winds then carry the clouds across the Earth's surface until the water is released as precipitation (rain, hail or snow) to fall on the Earth for further recycling. Most of the precipitation falls into the oceans (75 % of the surface of the planet being ocean). The rate of circulation of water within the hydrologic cycle is very rapid. Since the total mass of water in the atmosphere is constant, precipitation is balanced by evaporation. Thus, comparing throughputs shows that the average residence time of water molecules in the atmosphere is about 10 days.

5.3 Water in the atmosphere

Water greatly affects the behaviour of the atmosphere. The main areas are:

- thermodynamics (through condensation and evaporation)
- cloud formation (and thus the planetary albedo as well as precipitation effects),
- cleansing of the atmosphere by rainout (e.g. removal of substances within clouds such as hygroscopic (water-absorbing) aerosols), washout (removal of aerosols and dissolution of soluble atmospheric gases through capture by falling raindrops),
- chemistry of the atmosphere (as a solvent or participant in reactions), and
- absorption of radiation. Water is a major greenhouse gas.

The amount of water vapour in the atmosphere is limited by the saturation vapour pressure. This is the partial pressure of water vapour in equilibrium with the condensed phase (i.e. the attempt to add any more vapour will cause condensation). Note that this is the partial pressure of the water. The total pressure of the atmosphere is irrelevant. This quantity can be obtained from the Clausius-Clapeyron equation, a standard result of equilibrium thermodynamics. This states that, for a vapour in equilibrium with its condensed phase, the partial pressure, p, is given by

$$\frac{\mathrm{d}\ln p}{\mathrm{d}T} = \frac{L}{RT^2},$$

where L is the latent heat of evaporation (in kJ/mole). This is often given in J/kg, i.e. $L_S = L/M_v$, where L_S is the latent heat and M_v is the molecular weight. We can define a specific gas constant for water, $R_S(H_2O) = 1000\ R/M_v$. In the case of water, the partial pressure, p, is conventionally represented by the symbol e, and the saturated partial pressure by e_s. Setting $p = e_s$, we can easily integrate the above equation, to give

$$e_s = e_s^0 \exp\left(-\frac{L}{R}\left[\frac{1}{T} - \frac{1}{T_0}\right]\right) = e_s^0 \exp\left(-\frac{L_S}{R_S(\mathrm{H_2O})}\left[\frac{1}{T} - \frac{1}{T_0}\right]\right),$$

where e_s^0 is a constant (usually the saturated vapour pressure at S.T.P., 298.15 K and one atmosphere pressure) and T is the temperature. Note that this equation shows that e_s is strongly dependent on temperature. Warm air can contain much more water vapour than cold air. For example, in the tropics (air temperature 25^0C) the partial pressure of water vapour is 32 mb. In the polar vortex (temperature -20^0C) the partial pressure is 1.2 mb.

The other measure is the relative humidity, RH. This is the ratio of the partial pressure of water in the atmosphere to the saturated vapour pressure at that temperature, expressed as a percentage, i.e.

$$\mathrm{RH} = 100 \cdot e(T)/e_s(T).$$

5.4 Clouds

There is a general classification of clouds summarised in the list below. Terms can either be nouns (the –us form) or qualifiers (the –o form):

- cirrus (cirro). A cloud containing ice-crystals - therefore high altitude.
- Stratus (strato). A layer cloud. This may be continuous or show structure.
- Alto. A middle-height cloud.
- Cumulus (cumulo) indicates vertical circulation within the cloud. This gives the 'fluffy' shape and is quite different from the stratus form.
- Nimbus (nimbo). Cloud producing precipitation (rain, snow, hail).

Not all the possible combinations can occur even in principle, but many do. For example, a cumulonimbus cloud is a cloud with large vertical movement of air producing (usually a lot of) rain – it is a thunderstorm cloud.

Possible combinations are:

- cirrus (high[4], ice, white bands or filaments/patches with fibrous appearance),

---

[4] Low is designated from the surface to 2 km, middle 2-7 km, and high 7 km to the tropopause.

- cirrocumulus (high, ice, white patch regularly arranged in form of grains),
- cirrostratus (high, ice, whitish veil covering large area of sky),
- altocumulus (middle, water/ice, white/grey regular arrangement of small clouds),
- altostratus (middle, water/ice, greyish or bluish layer fibrous appearance covering large area of sky),
- nimbostratus (low, water/ice, grey thick layer often with snow and rain),
- stratocumulus (low, water, grey or whitish layer with dark elements, regularly arranged),
- stratus (low, water, grey layer with fairly uniform cloud base sometimes with snow and drizzle),
- cumulus (low, water, detached clouds, sharp outlines, 'cauliflower' clouds),
- cumulonimbus (5-12 km, water/ice, anvil-shaped thundercloud).

Let is consider an average cumulus cloud. The shape is (crudely) cylindrical. A typical cloud will be about 2 km in diameter and 2 km in depth. Hence the volume is about $\pi d^2 H/4 = 6.28\cdot 10^9$ m^3. A cloud contains of the order of 50-500 million water droplets per cubic metre. Each of these droplets is about 10 μm in radius. Thus the volume of a single droplet is $4\pi r^3/3$, i.e. $4.2\cdot 10^{-15}$ m^3. Given the density of water, 1000 kg/m^3, the mass of the droplet is about $4.2\cdot 10^{-12}$ kg. Thus the mass of the cloud (at the lowest reasonable density of droplets) is about $1.3\cdot 10^6$ kg. This seems a lot, but if all the water in this cloud fell at once, the depth of water would be (Volume of water)/(area of cloud base), which for the figures we gave above is $4.2\cdot 10^{-4}$ m, i.e. 0.42 mm which is not a lot. During a storm therefore, lots of 'clouds' must pass over you.

5.4.1 Physics of cloud formation

Clouds are formed as hot humid air rises and cools by adiabatic expansion. When air reaches saturation, excess water vapour condenses onto small particles called *cloud condensation nuclei*. The partial vapour pressure required to produce condensation is higher than the thermodynamic vapour pressure because the surface of a drop is curved. On the other hand, the effect of dissolving material into the drop decreases the vapour pressure. Both effects depend on the drop radius, r. The solute nuclei come from natural sources such as dust from wind erosion, sea salt, etc. They can also arise from industrial products (sulphate particles, coal dust). About 100 such particles per cm^3 are needed to form a cloud. Since there are 10^3 particles/cm^3 over the sea, 10^4 particles/cm^3 over land, and 10^5 particles/cm^3 over cities, forming clouds should not, on the face of it, be a problem. However, matters are not quite so simple.

5.4.1.1 Growing droplets in cloud

Consider a water drop at rest. Water vapour condenses onto the surface of the droplet, releasing latent heat. The temperature of the droplet rises, altering the vapour pressure, e, and eventually, a steady state is reached. There is a gradient of the partial pressure of the water vapour driving a flux of water vapour towards the drop. Also, the rate of heat loss is equal to the heat gained by the drop from the latent heat of the water. This gives an equation of the form for the growth rate of the radius r of the drop

$$r\frac{\mathrm{d}r}{\mathrm{d}t} = C\,,$$

where C is a constant. This integrates to give $r^2 = r_0^2 + 2Ct$, where r_0 is the initial droplet size. The problem is that detailed calculation shows that it would take 1-4 hours for a drop to grow from 2 to 30 μm. Yet in a typical cloud there are many droplets larger than 10 μm although the lifetime of a cloud can be as short as ten minutes. One final problem, as the cloud rises, the temperature falls and so the saturated vapour pressure falls. Thus the super-saturation in the cloud must rise. This means that more and more nuclei are capable of acting as condensation nuclei for water droplets. More of the water vapour is removed and, eventually, equilibrium is reached again. How then does the cloud grow? Clearly condensation is not enough to explain cloud growth.

Another possible growth mechanism is coalescence: two droplets collide to make a larger droplet. The probability of coalescence depends on: (i) the size of the droplets (i.e. their cross-section), and (ii) their relative velocity. The velocity of droplets depends on a balance between gravity and frictional forces. For reasonably small particles (less than 30 μm in diameter), the terminal velocity is given by Stokes Law,

$$F_{\mathrm{v}} = 6\pi\eta r v,$$

where F_{v} is the drag force, i.e. the force exerted by the viscous forces on the particle that act against the force pushing the particle through the medium (in this case gravity), η is the viscosity of the medium (air in this case), and r is the radius of the particle. When these forces are in equilibrium the droplet reaches terminal velocity. This obviously depends on the droplet size. Typical values are given in the next table

Droplet diameter (mm)	Terminal velocity (m/s)
0.01	0.003
0.1	0.025
1.0	4.03
5.0	9.09

The expression controlling the rate of growth takes the form

$$\frac{\mathrm{d}r}{\mathrm{d}t} = \frac{Ewv}{4\rho},$$

where E is the collection efficiency taking account of the fact that the drop is not a sphere, w is the volume swept out by the falling droplet, v is the droplet velocity and ρ is the density of water. Combining the two modes of growth, it is possible to see that there is a 'barrier' to growth at the 20 μm size where neither process is efficient. The mechanism that overcomes this barrier is still not understood. Possibilities include increasing the efficiency of collisions by including the effect of turbulence, invoking droplet-droplet interactions by electrical forces. When the droplet reaches a diameter of 2-3 mm they are broken apart by collisions rather than built up. Above 6 mm the droplet becomes unstable anyway. The surface tension cannot hold it together.

Once the drops grow to a size beyond the ability of the up-draughts in the cloud to maintain them, they begin to fall. Below the cloud, they are in unsaturated air and begin to evaporate again. Rain falling from a cloud need not reach the ground. Only the larger drops, those that have swept up their neighbours by coalescence, can reach the ground as rain. In the case of colder clouds, the precipitation often begins as ice or hail which melts on the way down. Sometimes, of course, it doesn't.

5.4.2 Thunderstorms

Thunderstorms occur when moist, warm air near the ground becomes buoyant and rises to form small cumulus clouds. These clouds grow and surge upwards to combine and form cumulonimbus, anvil-shaped clouds. These main extend about 1 km into the stratosphere. Within these clouds there is vigorous movement, updrafts and downdrafts at tens of metres per second. Such movements lead to a separation of electrical charges within the cloud. Positive charge gets carried to the top and negative charge to the bottom. The charge carriers are electrons, molecules, aerosol dust, hailstones and snowflakes. The process of charging is unclear but may include frictional contact, freezing, melting, and break-up of water droplets. Once charge separation has occurred, an electric field is established. The air may then become ionised. Thus a small region of the atmosphere is changed from being a good insulator to a highly conducting path. Thus, there can be (and is) an electrical discharge. This is lightning – which is simply a high-voltage spark. The average lightning discharge involves the flow of ten coulombs of charge across a potential difference of about 100 MV. Thus the energy is about 10^9 J (about 280 kWh). The main types of lightning are:

- ground discharges (thunderbolts, forked lightning). These occur between the cloud and the ground. They consist of a 'leader stroke' establishing an ionised path to the ground up which a 'return stroke' passes.
- Cloud lightning (sheet lightning, heat lightning) occurs within the cloud. It gives a diffuse illumination without a distinct channel being seen. If there are distant thunderstorms on the horizon, the sheet lightning is what will be seen.
- Air discharges (streak lightning) passes from cloud to cloud or cloud to air but does not reach the ground (i.e. no return stroke).
- Ball lightning. Unusual and still not fully explained. Usually small balls (but reported diameters up to 1 m) of electrical discharge. Move slowly in the air or along the ground. Usually disappear with a violent explosion.

Although lightning is the most spectacular effect of a thunderstorm, it is only a small part of the total energy budget of the storm.

Thunder is the sound wave produced by a lightning stroke. The sudden rise in pressure in the lightning channel produces an intense sound wave as in an explosion. Sound travels at about 330 m/s, so, the sound follows behind the lightning flash. Hence counting the interval between the flash and the thunder gives a rough estimate of the distance of a storm (about 1 km for every three seconds but there may be echoes which confuse matters).

Questions:

- How does the terminal velocity of a drop depend upon its radius in the range 10-30 μm?
- The horizontal acceleration on an air parcel of mass 1 tonne is 10^{-4} ms^{-2}. What is the net force on the air parcel? Estimate the volume of the air parcel at sea level. (**Answer**: 0.1 N; 1000 m^3)
- Given that a cumulus cloud is typically 2 km deep with a similar diameter and contains $5\cdot10^7$ water droplets per m^3 each of 10 μm radius, calculate the depth of rainfall shoul the cloud release all its water in one instant. (**Answer**: 0.42 mm)
- Calculate the space charge, assuming that there are 1000 positive ions and 900 negative ions per cm^3 in the air. (**Answer**: 16 pCm^{-3})

6 Wind

6.1 Measuring the wind

The wind is one of the more obvious features of the weather. While we have a rough and ready appreciation of what counts as a strong or gentle wind, this does not give any quantitative estimate. By about 1800, it was clear that it would be useful to categorise winds in terms of their effects on sailing ships. This was done by Francis Beaufort, whose scale remained standard until 1946 and is still that used for shipping forecasts. The table below gives a simplified modern description for use on land.

Beaufort Number	Description	Wind speed (m/s)
0	Calm: smoke rises vertically.	0-0.5
1	Light air: wind direction shown by smoke drift, not by vanes).	0.5-1.5
2	Light breeze: wind felt on face, leaves rustle, vanes move.	1.5-3.5
3	Gentle breeze: leaves and small twigs move, light flags lift, large wavelets at sea.	3.5-5.5
4	Moderate breeze: dust and loose paper lift, small branches move.	5.5-8.0
5	Fresh breeze: small leafy trees sway, moderate waves.	8.0-10.5
6	Strong breeze: large branches sway, telegraph wires whistle, umbrellas difficult to use.	10.5-13.5
7	Near gale: whole trees move, inconvenient to walk against.	13.5-17
8	Gale: small twigs break off, walking impeded, high waves and foam.	17-20.5
9	Strong gale: slight structural damage.	20.5-24.5
10	Storm: considerable structural damage, trees uprooted.	24.5-28.5
11	Violent storm: widespread damage.	28.5-32.5
12	Hurricane: at sea, visibility badly affected by driving foam and spray. Sea surface completely white.	>32.5

6.2 Physics of wind creation

The atmosphere is a gigantic heat engine. The radiation from the Sun causes convection both on a local scale and on a global scale. The difference in energy delivered at the equator and the poles creates the pressure differences that drive the major wind systems in the Earth's atmosphere. Winds are masses of air in motion. An air mass has a reasonably precise meaning in meteorology. This is a large volume of air (covering millions of square kilometres) that has reasonably constant pressure and humidity. Thus, the air mass will determine the overall weather of a region (although not the local microclimate). They come from the extensive regions of high pressure that characterise some of the areas of the Earth, the subtropical oceans (throughout the year) and the mid and high latitude continents (mainly in winter). Air spirals out from these high-pressure areas (anticyclones) to create the wind systems of the planet. Two examples of relevance to Europe are:

- the Azores anticyclone. The south westerlies blow towards the North Pole and the North-East Trades towards the equator. This warm and humid air is classified in mid-latitudes as ‘tropical maritime’.
- The polar continental. The cold dry air, in winter comes from the Eurasian.

The boundaries that separate air masses are called fronts. Hereabouts a cold front is the leading edge of a cold air mass and brings rain. The warm moist air of the tropical maritime mass it is replacing is forced upwards, cools and water precipitates out.

6.2.1 Principal forces acting on air masses

If we want to understand why the winds occur, it is necessary to consider the forces that act on the air masses in the atmosphere. To any observer stationary with respect to the surface of the Earth, there are four forces acting on a parcel of air in the atmosphere:

- gravitational,
- pressure gradient,
- Coriolis fictional force, and
- frictional force.

6.2.1.1 Gravitational force

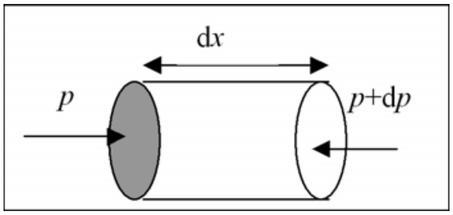
Due to the large mass of the Earth, the gravitational force is one of the strongest forces acting on the air parcel and is directed towards the centre of the Earth

$$F_g = g\rho\Delta V,$$

where g is the gravitational acceleration constant (more or less constant through the troposphere), ρ is the density of air, and ΔV is the volume of the parcel.

6.2.1.2 Pressure gradient

The pressure at the surface of an air parcel is the normal component of the force exerted by its surroundings on a unit area of surface. This force is always directed towards the parcel. The parcel will experience a net force if there is a difference between the pressures on the surfaces at different sides.



Consider the above diagram where the shaded cross-section has area dA. The net force on the parcel due to the pressure difference is

$$F_P = p\mathrm{d}A - (p + \mathrm{d}p)\mathrm{d}A = -\mathrm{d}p\mathrm{d}A.$$

If ρ is the density of air, then the force per unit mass (i.e. the acceleration) is given by

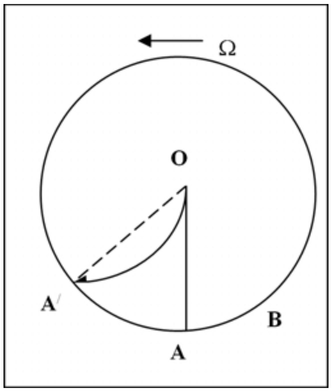
$$\frac{F_P}{\rho \mathrm{d}A\mathrm{d}x} = -\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}A}\frac{\mathrm{d}A}{\mathrm{d}x} = -\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}x}.$$

Taking the limit in the usual way and generalising to three dimensions, we get

$$\vec{F}_P = -\frac{1}{\rho}\nabla p.$$

6.2.1.3 Coriolis inertial force

Only the gravitational and pressure gradient forces can initiate motion of the air but there is a third, 'fictitious force' due to the rotation of the Earth – this is the Coriolis inertial force.



Consider the above diagram which represents a region around the North Pole (O) such that we can consider it to be a rotating disc. An air parcel starts to move horizontally away from the pole towards a point A. If no forces act on this parcel, by Newton's laws it will follow a direct path OA. However, the disc is also rotating with an angular velocity, Ω, and so it will follow the curved line OA/ with respect to the disc. The disc has moved on, and so the point A is at B by the time the parcel of air reached the edge. To an observer rotating with the Earth, it looks as though the parcel is deflected by a force away from A towards A/. This fictitious force is called the Coriolis force[5]. The Earth is a sphere and not a disc, which means that we should use the full vector notation, but the principles are the same. We must calculate the vector product between the rotation vector $\vec{\Omega}$ (directly out of the plane for the disc) and the velocity vector, $\vec{v}_g$.

Applying this to the Earth, if we consider an air parcel with velocity $\mathbf{v}_g$, and the angular velocity vector of the earth is $\vec{\Omega}$, then the Coriolis force per unit mass is given by

---

[5] A more systematic way of looking at this problem is that the simple application of Newton's Laws of the form $F = ma$ is not applicable in rotating frames of reference, however, in a case like the Earth it is much more convenient to take the Earth as stationary and consider the matter in terms of fictitious forces.

$$\vec{F} = -2\vec{\Omega} \times \vec{v}_{\mathrm{g}},$$

However, since the atmosphere is thin compared with the radius of the Earth, we know that the wind is blowing almost along the local horizontal. Thus it is useful to split the Coriolis force into two contributions with respect to the local vertical. We designate these by the unit vectors $\vec{z}$ and $\vec{y}$. If the latitude is φ, then the two components are given by: $\Omega\sin\varphi)\vec{z}$ and $(\Omega\cos\varphi)\vec{y}$, where $\vec{z}$ is the unit vector along the local vertical, **y** is in the local horizontal plane. If we assume that the wind is moving in the $\vec{y}$ direction, then the Coriolis force is given by:

$$\vec{F} = -(\Omega\sin\varphi)\vec{z} \times v_{\mathrm{g}}\vec{y} = (v_{\mathrm{g}}\sin\varphi)\vec{x},$$

i.e. there is a force in the $\vec{x}$ direction. To an Earth-bound observer, the wind is steered increasingly to the right of the initial direction of movement when viewed downward in the northern hemisphere and increasingly to the left in the southern hemisphere.

The initial movement will be down the pressure gradient (which is providing the force that creates the wind), but deflections continue until the wind is blowing perpendicularly to its original motion (and so at right-angles to the pressure gradient producing it). It cannot then be deflected any further. The Coriolis force balances the pressure gradient. Put formally

$$\left(-\frac{1}{\rho}\nabla p + f_{\mathrm{C}}v_{\mathrm{g}}\right)\vec{x} = 0; \quad f_{\mathrm{C}} = 2\Omega\sin\varphi,$$

where f_{C} is the Coriolis constant (about 10^{-4} s^{-1}), v_{g} is the air velocity at balance. This gives

$$v_{\mathrm{g}} = \frac{1}{f_{\mathrm{C}}\rho}\frac{\mathrm{d}p}{\mathrm{d}x}$$

assuming that the pressure gradient increases along the $\vec{x}$ axis.

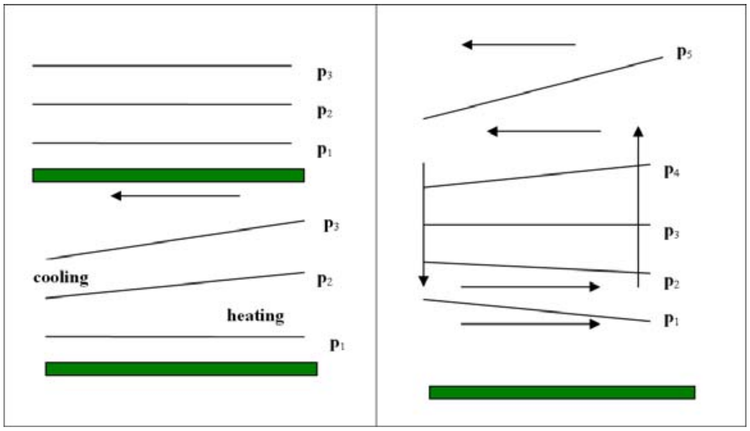
The point where the Coriolis effect balances the pressure gradient is called the geostrophic balance and the resulting wind called the *geostrophic wind*. The important point to note is that the wind direction is parallel to the isobars and (in the northern hemisphere) the wind direction is such that the lower pressure is on the left-hand side as you face downwind. Thus a low-pressure area in the northern hemisphere has the winds rotating around it in a counter-clockwise direction (but clockwise in the southern hemisphere). This type of motion is called *cyclonic*. Thus low pressure weather system is called a *cyclone*. Over a high-pressure area in the northern hemisphere, the geostrophic wind circulates in a clock-wise direction. This is called *anticyclonic* motion and therefore a high-pressure weather system is known as an *anticyclone*.

From the value of f_{C} derived above, it is clear that the Coriolis force is greatest at the poles, and decreases as one approaches the equator, where it is zero. (Note that there will also be a vertical component proportional to $\cos\varphi$ arising from the $\vec{y}$ vector. This will make a small contribution to the effective gravitational force.

6.2.1.4 Frictional force

There exists considerable frictional force between the atmosphere and the Earth's surface (for example due to mountains and hills. Lines of tall trees are planted to act as wind breaks and protect crops). Frictional forces are difficult to treat properly. The mechanism is essentially a form of viscosity (at low altitudes) and small-scale eddy mixing processes at higher altitudes. The layer where the frictional force is important is known as the *planetary boundary layer*. The thickness of the layer is very variable, from a few hundred metres in still air at night to 4-5 km over a hot surface with strong convection. This introduces the question of how the wind speed varies with height which is obviously closely tied to the question of how atmospheric pressure varies with height. We considered this question when we looked at the structure of the atmosphere. Its connection to the wind speed can be seen through the following example (thermal gradients and winds).

Pressure differences can be induced by uneven heating or cooling. Consider a region of the atmosphere with an even temperature distribution (top right with $p_1 > p_2 > p_3$). The air is now heated at one end (adding energy) and cooled at the other (extracting energy). Hot air expands and cold air contracts. Thus the air columns on the left and right hand sides will try to expand (on the right) and contract (on the left).

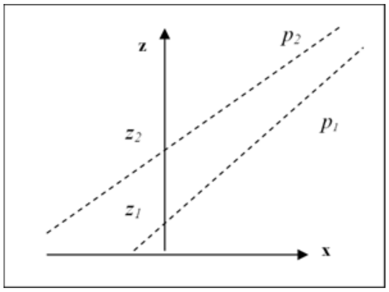


This will produce a horizontal pressure gradient. The lines joining points of equal pressure (isobars) will no longer be at the same height. There is therefore a force moving air down the pressure gradient (bottom right). This is a thermal wind. The rate of decrease of pressure at a fixed height in the warm area is equal to the mass of air flowing out above the fixed height. The rate of pressure decrease is therefore larger in the lower levels than the higher. Eventually, a steady state is reached. The pressure in the upper part of the air column will be lower in the warm area and higher in the cold areas, whereas in the lower part of the air column the pressure is higher in the cold areas and lower in the warm. This is the basis of convection in the atmosphere.

Example: A day at the seaside.

During the day, the land is heated by the Sun and the temperature rises above that of the sea (the specific heat of the land is less than the specific heat of the sea). The air over the lands is warmer than that over the sea resulting in air blowing from the sea towards the land (at low levels) – a *sea breeze*. At night the land cools below the temperature of the sea and so the situation reverses. The (low-level) wind now blows from land to sea – a *land breeze*. This simple example shows one of the basic mechanisms underlying the global circulation of the atmosphere.

We can put this argument on a more formal basis and consider the variation of the geostrophic wind with height. Let us consider this in more detail. For simplicity, let us assume as before that the pressure variation is along the ($+\vec{x}$) axis.



Then the geostrophic wind is along the (+**y**) axis and is given by:

$$v_{\mathrm{g}} = \frac{1}{f_{\mathrm{C}}\rho}\frac{\mathrm{d}p}{\mathrm{d}x}.$$

We now consider the situation in the above diagram. We take two isobars, pressure p_1 and p_2 at heights z_1 and z_2, and calculate the difference in the geostrophic wind velocity between them in the upwards direction. This is given by

$$\Delta v_{\mathrm{g}} = \frac{1}{f_{\mathrm{C}}}\left[\left(\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}x}\right)_{z_2} - \left(\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}x}\right)_{z_1}\right].$$

But we know that the hydrostatic equation states that

$$\mathrm{d}p = -\rho g \mathrm{d}z.$$

In the diagram above, the isobars are sloped, so we can measure *dp* either along the $\vec{x}$ axis or along the $\vec{z}$ axis. In the diagram above, the values of z are linked to the lines of constant pressure that intersect the **x** axis, so we can write

$$\frac{\mathrm{d}p}{\mathrm{d}x} = -\rho g \frac{\mathrm{d}z(p)}{\mathrm{d}x},$$

where we have written $z = z(p)$ to remind us of the connection with the heights and the isobars. So we can therefore write (6) as

$$\Delta v_{\mathrm{g}} = \frac{g}{f_{\mathrm{C}}}\left[\left(\frac{\mathrm{d}z(p)}{\mathrm{d}x}\right)_{z_2} - \left(\frac{\mathrm{d}z(p)}{\mathrm{d}x}\right)_{z_1}\right] = \frac{g}{f_{\mathrm{C}}}\frac{d(\Delta z)}{\mathrm{d}x}.$$

We can now use the buoyancy equation, $\mathrm{d}p = -\rho g\mathrm{d}z$, again to determine Δz. Using the ideal gas law, $p = R_S \rho T$ and rearranging, we have

$$\mathrm{d}p = \frac{pg}{R_S T}\mathrm{d}z\,.$$

Integrating between the isobars p_1 and p_2 (and therefore between the heights z_1 and z_2) we have

$$-\int_{p_1}^{p_2}\frac{R_\mathrm{S} T}{g}\frac{\mathrm{d}p}{p} = \int_{z_1}^{z_2}\mathrm{d}z\,,$$

where R_S is the specific gas constant. Integrating this and assuming that g is constant, we obtain

$$-\frac{R_S}{g}\int_{p_1}^{p_2}T(z)\mathrm{d}(\ln p) = z_2 - z_1\,.$$

If we replace the height-dependent temperature $T(z)$ by an average value, $\overline{T}$, then we obtain

$$z_2 - z_1 = \Delta z = \frac{R_\mathrm{S}\overline{T}}{g}\ln\left(\frac{p_1}{p_2}\right).$$

We can now differentiate this, remembering that we chose p_1 and p_2 to be constant

$$\frac{\mathrm{d}(\Delta z)}{\mathrm{d}x} = \frac{R_\mathrm{S}}{g}\ln\left(\frac{p_1}{p_2}\right)\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\,\boldsymbol{.}$$

For the difference in the geostrophic wind velocity, we get

$$\Delta v_g = \frac{R_\mathrm{S}}{f_\mathrm{C}}\ln\left(\frac{p_1}{p_2}\right)\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\,.$$

Finally, we can eliminate the pressure term using (12) to give the difference form of the thermal wind equation

$$\frac{\Delta v_g}{\Delta z} = \frac{g}{f_\mathrm{C}\overline{T}}\left(\frac{\mathrm{d}\overline{T}}{\mathrm{d}x}\right).$$

As Δv_g and Δz tend to zero, the left-hand side tends to the derivative and the average temperature on the right-hand side tends to that of the isobaric surface in the middle of the layer. This gives an important relationship between the vertical variation of the geostrophic wind and the horizontal temperature gradient. The isobaric slope increases with height, so the horizontal pressure gradient also increases with height. Hence the geostrophic wind should also increase with height. The increase in velocity is in a direction which is perpendicular to the temperature gradient with the cold region to the left and the warm region to the right of the wind vector increment (in the northern hemisphere). Thus the geostrophic wind rotates counter-clockwise (with altitude) when the wind blows from a cold region to a warm one, and clockwise (with altitude) when the wind blows from warm to cold.

6.3 Cyclones and anticyclones

Storms have been recorded throughout history, but the first comprehensive account of a storm is given by Daniel Defoe (the ‘great storm’ of 26th November 1703). The first observation on cyclonic systems is that of Benjamin Franklin (21st October 1743). He noted that he was prevented from seeing a lunar eclipse in Philadelphia by a N.E. gale but the eclipse was seen in Boston (300 km to the N.E.) Thus the storm itself must be moving against the direction of its constituent winds. The development of the electric telegraph allowed storms to be tracked. Admiral Fitzroy founded the U.K. Meteorological Office. He noted that cold and hot air masses were both involved in cyclonic behaviour (1840s). From 1910 to 1930, the Norwegian Meteorological Office plotted warm and cold fronts – the first use of scientific aircraft. After 1960, the use of meteorological satellites grew rapidly, enabling the plotting of cyclonic motion and ‘fronts’.

Cyclonic systems can be found in low and middle latitudes. Middle latitude cyclones are the systems responsible for the ‘bad weather’ in middle latitudes. They are usually called depressions (not because of this but because they are low pressure systems). Cyclones that arrive here form on the other side of the Atlantic (where cold polar air over the North American continent meets warm tropical air from the Western Atlantic – heated by the Gulf Stream). Similar effects occur in the North Pacific and Northern Mediterranean (alpine air meeting air from North Africa). A depression forms when a wave develops at the boundary between the two air masses. Air then starts to flow across the isobars and a low pressure area develops with cyclonic motion. This then moves in accordance with the winds in the warm sector. Since cold air tends to move faster than warm air, the cold front tends to catch up with the warm front and the warm air rises above the cold air.

Anticyclones form over Siberia, Canada and North Russia. These dominate the Asian and North American climate in winter. They give rise to (mainly) dry weather but can trap extensive low-lying clouds. Since winds are usually light, pollution levels often rise. Most anticyclones last only four or five days, but occasionally they last much longer. Such systems are called blocking highs. These block the movement of cyclones in mid-latitudes, giving stationary fronts for days and even weeks.

Tropical cyclones are hurricanes (or typhoons). These are low pressure systems of great intensity. Wind speeds of 100 m/s can be produced and heavy rainfall always occurs. The cyclonic nature of air circulation in the lower atmosphere in a hurricane or typhoon is easily seen by the nature and movement of clouds as seen from a satellite. Rising air from the equator is humid and cools on rising giving a girdle of clouds around the Earth. These give heavy rain in the equatorial regions. The cold and warm air meets at the inter-tropical convergence zone (ITCZ) and minor depressions along the edge. Some (and only some of these) develop into major storms. The reason for the early stages is not clear, but once begun, the mechanisms that give major hurricanes (or typhoons) are well understood:

- pressure starts to fall rapidly at the centre of the disturbance.
- Winds rise in a tight band, 30-60 km in radius (the central eye).
- As the storm grows it moves to the westerly in the trade winds (8-15^0 latitude) and migrates to higher latitudes.
- The maturing storm expands while the central pressure stops falling. The route depends on local surface conditions such as the surface temperature (and perhaps

the salinity) of the sea (and is unpredictable). Hurricanes only form over warm water (temperature greater than 27^0).
- The storm grows to 300 km radius or greater and then begins to decay. Decay is hastened by passing over cold water or land.
- The storm is pushed by mid-latitude westerlies in higher latitudes as it decays.

These storms are found in both hemispheres, most commonly in the North Pacific. Most frequent in late summer (when the sea surface is warmest) but they can occur all summer and autumn in the tropics. They are part of the process by which energy is transported from the equator to the poles.

6.4 Global convection

The first model to describe large-scale global convection was proposed by George Hadley in 1735. He noted that air in the lower latitudes is warmer than that in the higher (polar) latitudes due to the greater solar flux reaching the equator. Tropical air should move (rise) vertically and move northwards while the cool polar air should move southwards. As the tropical air moves north, it loses energy by radiation before descending to the ground, so replacing the southward-moving colder air. Similarly, the cold air will gain heat from the ground (which is itself heated by radiation as discussed earlier) and will therefore rise in the equatorial regions. Thus a circulation system is formed transporting thermal energy from the equator to the poles. This is the *Hadley cell*.

There are, however, significant differences between Hadley's model and the real air circulation patterns. It is true that there is a low pressure belt over the equator and a high pressure region over the pole as the Hadley model would predict. However, there is another circulating cell of air between 30^0N and 60^0N (and similarly in the southern hemisphere) in which air rises in the colder regions (i.e. at 60^0N) and descends in the warmer regions (i.e. in the opposite direction to the Hadley mechanism). This is the *Ferrel cell*. Finally, there is a third cell between 60^0N (and S) and the pole. This circulates in the same direction as the Hadley cell but is much weaker. It is known as the *Polar cell*.

6.5 Global wind patterns

Before we can describe these, we must discuss how the atmospheric pressure varies over the planet. This depends on the seasons:

- Northern winter/southern summer. There is a weak high pressure across the Arctic Ocean and low pressure centres SE of Greenland and in the North Pacific. Between 45^0N and 15^0N there is an extensive region of high pressure, stretching from the subtropical Eastern Pacific eastwards round the planet as far as India (an exception is a low-pressure region in the Mediterranean). A broad belt of weak pressure gradient encircles the planet at low latitudes. There are weak low pressure centres over the southern continents and subtropical anticyclones over the main southern oceans. At 40^0S there is a belt of rapidly changing pressure gradient (and therefore strong winds), and finally a high pressure region over Antarctica.
  - Northern summer/southern winter. The mid-latitude oceanic low centres in the North Atlantic and North Pacific weaken and the subtropical anticyclones shift towards the Pole. Extensive low pressure (rather than high pressure) is

observed over central and southern Asia (associated with the monsoon). There are also low pressure regions over the subtropical areas of Africa and North America. The reverse (from summer low to winter high) occurs over the equivalent areas in the southern hemisphere. There is a deep low pressure belt around Antarctica.

The strongest wind patterns produced are, as one would expect from the seasonal pressure variation discussed above, themselves seasonal:

- Northern winter/southern summer:
  - the mid-latitude westerlies (i.e. blowing from the west). In the Northern hemisphere these are confined by the ocean basins.
  - The north-easterlies moving out from the Asian anticyclone over the Arabian and South China seas.
  - The trade winds (so called because of their importance to sailing ships in the 16th – 19th century. The trade wind in the northern hemisphere is north-easterly (i.e. blows from the north-east). In the southern hemisphere it is south-easterly. The Inter-tropical Convergence Zone (ITCZ) is the west-east region into which the heat and moisture carried by these winds are blown.
- Northern summer/southern winter:
  - The mid-latitude westerlies are much weaker in the northern latitudes (about 65 % of their winter velocity.
  - The ICTZ over the oceans moves north (but not as far as the continental eqivalent which penetrates into sub-Saharan Africa and southern Asia). This causes a reversal of wind direction over the Arabian and South China seas which is part of the monsoon.

Such winds are not explained by Hadley's model but are due to the balance of forces set up by the pressure gradient induced by: (i) thermal convection, (ii) frictional forces of air moving across the Earth's surface, and (iii) Coriolis force.

Wind patterns higher up are much simpler. The frictional force of the continental masses is reduced. The most important seasonal winds occur between 30^0N and 50^0N (and similarly S). These are the jet-streams. An important effect in the Southern hemisphere is the polar vortex. The westerly winds high over the coasts of Antarctica flow in a more or less circular pattern and prevent heat transport (by incoming depressions) into the air above Antarctica. This increased cold is important in depleting the ozone layer and explains why the 'hole' was first observed over Antarctica.

Questions:

- Consider a low-presure system in the Southern-Hemisphere. In which direction does air circulate around the low-presure centre?
- Using the definition for the geostrophic wind, or otherwise, explain why hurricanes do not cross the Equator.
- Storm force northerly winds are blowing at 25 m/s over sea area at 55^0N. Calculate the horizontal pressure gradient (in Pa per 10 km) associated with this speed, assuming geostrophic balance. (**Answer**: 35.6 P per 10 km)
- Calculate the geostrophic wind speed for a pressure gradient of 0.03 mb/km, assuming that the Coriolis parameter $f = 10^{-4}$ s^{-1}. (**Answer**: 30 m/s)

7 Physics of ground

The Earth's surface is a complicated affair comprising oceans, deserts, cities, forests, tundra/savannah, ice-caps, and so on. Each of these has a different albedo, rate of water evaporation/absorption, exchange of gases with the atmosphere, and so on.

7.1 Soils

Soil is composed of both rock particles and organic matter (humus) - the remains of plants and animals in various stages of decomposition. The humus serves as food for many living organisms. Within the soil is a large population of animals, plants. These break down the humus into soluble substances that can be absorbed by the roots of larger plants. Soils are categorised by the air, water, rock and humus content. The first consideration is the size of the particles that make up the soil. A set of typical sizes for various soils is given below:

Fraction	**particle diameter (μm)**
gravel	> 2000
sand	60-2000
silt	2-60
clay	< 2

Much of the soil consists of air and water. Between 30 % and 70 % of soil is pore space. The fraction of soil that is occupied by solid is

$$\begin{aligned}\text{Fraction of soil volume occupied by solid} &= \frac{\text{Mass of soil}}{\text{Volume of soil}} \times \frac{\text{Volume of solids}}{\text{Mass of solids}} \\ &= \text{bulk density} \times \frac{1}{\text{particle density}}.\end{aligned}$$

We define the porosity as the fraction of soil volume occupied by pores:

$$\text{Porosity} = 1 - (\text{bulk density})/(\text{particle density}).$$

Typical bulk density and porosity, one can see in the following table:

Soil Texture	Bulk density (kg/m^3)	Porosity
Sandstone	2100	0.19
Sandy loam subsoil	1650	0.36
Sandy loam plough layer	1500	0.42
Clay loam subsoil	1450	0.44
Recently ploughed clay loam	1100	0.58

7.2 Soil and hydrologic cycle

The passage of water in and through the soil is an important part of the hydrologic cycle discussed above. There are two basic problems: (i) how much water is in the soil, and (ii) how does it move through the soil. The porosity gives a measure of how much water the soil can hold. In fact, this is a gross overestimate. Water in large pores and cracks (greater than 60 μm diameter) cannot be held in the soil. This can be seen from the following argument. The flow of water through a tube depends on the tube radius, the viscosity of water and the pressure gradient trying to push the water through the tube. A simple dimensional analysis gives the Hagen-Poiseuille equation (except of course for the constant in front). This states that

$$Q = \frac{\pi}{8}\frac{r^4}{\eta}\frac{\mathrm{d}p}{\mathrm{d}x},$$

where Q is the rate of flow of water (in m^3/s), r is the radius of the tube, η is the viscosity (units of Pa s) and $\mathrm{d}p/\mathrm{d}x$ is the pressure gradient. The point is the strong dependence of the flow rate on the diameter of the tube. This badly overestimates the velocity. Pores are not straight. Water velocities through pores typical of sand grains (diameter 1000 μm) are about 10^8 times those typical of clays (pore diameter about 0.1 μm). There is an upper limit to the amount of water that a soil can hold in the long term. This is the field capacity. It is the water held in pores small enough so that friction and surface tension (see below) can resist the gravitational flow. There is also a lower limit to the amount of water that can be extracted by plant roots. If all the remaining water is held in very fine pore, the plant roots cannot extract it. This lower limit is called the *Permanent Wilting Point*.

7.3 Surface tension and soils

Moreover, some of the water is tightly held in the soil. This is because of the effect of surface tension. A liquid behaves as if its surface is enclosed by a skin. A molecule in the interior of a liquid experiences forces from all directions since the liquid molecules are in all directions. A molecule at the surface experiences forces only from the lower hemisphere, since above the molecule is air or vacuum (and the effect of air molecules is negligible). It therefore costs energy to make a surface (usually expressed as energy per unit area). This can also be looked at in terms of a force, the surface tension. The surface tension is the force acting across the surface pulling the molecules into the liquid. The same applies to interfaces between liquids and solids (and liquids and other liquids) since the forces between molecules in different materials are different. If the attraction between the molecules of a liquid is less than the attraction between the molecules of the liquid and that of a solid, the liquid will wet the solid. This determines the shape of the meniscus in a tube (concave or convex). Let us consider the case where the meniscus is convex (as it almost always is for water.

Let us assume that we are trying to push the water down the pore (say by forcing air into it) with a pressure p. Then the force is the pressure multiplied by the area of the tube; i.e. $\pi r^2 p$. Opposing this is the surface tension, γ. If the angle of contact between the liquid and the wall is given by θ, then the total force is $2\pi r\gamma\cos\theta$. If the system is in equilibrium, this gives the result $p = \pi\cos\theta/r$. It is often reasonable to set $\cos\theta$ to unity, giving

$$p = 2\gamma/r.$$

In the absence of an external pressure, p represents the tendency of water to creep along a pore from the wet to the dry end, and is often called the *suction*. A case of particular importance is when a column of water is supported against gravity by surface tension. In this case, it can be shown that the height of the column is given by $h = 2\gamma/\mathrm{r\rho g}$, where ρ is the density of the liquid.

7.4 Water flow

In most cases, the two most important forces in the soil tending to move water are gravity and the changes in the suction from one area to another. It is often convenient to express these in terms of a potential. The soil-water potential, Ψ, is defined as

$$\Psi = -(\text{depth} + \text{suction}).$$

The minus sign states that the potential becomes more negative with depth or as the suction increases. This is related to the rate of water flow by Darcy's Law. The combination of gravity and surface tension sets up a potential difference in the soil. This potential difference determines the volume of water that can pass through unit cross-sectional area of soil. We write

$$Q_\mathrm{W} = -\kappa \frac{\mathrm{d}\Psi}{\mathrm{d}x},$$

where Q_W is the volume of water per unit cross-section and κ is the effective permeability of the soil (often known as the hydraulic conductivity). The value of κ depends on the soil type and structure (in particular on the size and distribution of pores). κ decreases rapidly as the soil gets drier. The sign ensures that water flows down the potential gradient. The balance of the effect of gravity and suction ensures that there will always be some moisture at the top of the soil where it is in contact with the air and can evaporate. Water is also returned to the atmosphere from the leaves of plants (transpiration).

The amount of water stored in the soil, the soil profile, S, is a balance between a number of factors

$$\Delta S = P - E_\mathrm{S} - T - D - R,$$

where the contributions are P from the precipitation, E_S from the evaporation from the soil surface, T from transpiration, D from deep drainage out of the soil layer and R from runoff. This balance is important when investigating the leaching of pollutants into the soil. The dependence of water flow rate on pore size means that the water moves faster down cracks than the surrounding small pores. This effect is hydrodynamic dispersion.

7.5 Water evaporation

Evaporation of water from the soil is an important part of the hydrologic cycle. Most of the incident solar radiation is radiated back at infrared frequencies. Of the remainder, some is conducted into the Earth, some is transferred to the overlying air and some (up to 60 %) evaporates water. The principal mechanism for removing water from the land surface is turbulent transfer. It is possible to derive an expression for the evaporation rate for a simple case such as a lake surface. The Penman equation gives

$$\text{Evaporation rate} = aR + b(c_0 + c_1 v)D,$$

where R is the net radiation flux into the soil, v is the windspeed, D is the saturation deficit of the air and a, b, c_0, c_1 are disposable coefficients. Evaporation from land is much harder to describe with a simple expression. It depends strongly on the vegetation cover.

7.6 Soil temperature

The temperature range across the Earth's surface is large. The most obvious causes of variation are latitude and season. The seasonal changes are particularly big in the centres of continents. However, variations of 10^0C in a single day are possible. Even slight topographic features affect the temperature. Low spots in fields have lower crop growth (by 33-50 %) than other parts of the same field because frost formation is more likely. Small hills, depressions, exposure to the Sun, all affect the local temperature. The differences in air temperature just above the ground only affect the topmost layers of the soil because soil is a very poor conductor. An annual temperature variation of 30^0C in air is reduced to 15^0C at a depth of 1 m and 0.5^0C at 8 m. Indeed, the heat conduction through the soil is so slow that the temperature cycle deep within the soil is the reverse of the seasonal variation at the surface. In the Northern hemisphere:

- At 3 m: minimum temperature is March/April, maximum temperature is September/October.
- At 7-11 m: minimum temperature is August, maximum temperature is February (but the variation is very small – see above).

Vast areas of Canada and Siberia (20 % of the Earth's surface) have frozen soil. Intrusion of frost into the ground is so effective that summer heat cannot thaw the ground below 1m depth. Hence rain cannot sink into the soil and such regions become swamps in summer.

Questions:

- Discuss how characteristics of the soil pore space influence the movement of water and solutes through soil.

8 Energy for living

When discussing the physics of the atmosphere, we considered a few examples changes in the environment induced by human activity. One of the most important of these is the excess 'global warming' induced by increased carbon dioxide in the atmosphere. This in turn is related to the increased burning of fossil fuels which in turn is driven by demands for economic growth. In the past, growth in the economy has been tightly linked to an increased demand for energy.

The next table shows the growth of the rate of energy use analysed in terms of the various energy sources since about 1860. To date, most energy has been derived from fossil fuels, gas, oil and coal with a little wood and bio-waste. With industrialisation, the rate of energy use has increased 30-fold. At first, the main fossil fuel was coal. Since 1950 the major growth has been in the use of oil. In 1990, estimated world consumption of energy was 8730 million tonnes of oil equivalent (toe). However, great disparities exist in the amounts of energy used per person in various parts of the world.

Energy use in tonnes of oil equivalent per person in different regions of the world (1990)

(tonnes of oil equivalent; 1 toe = 1.33 kW)

North America	7.82
Former USSR	5.01
Western Europe	3.22
Eastern Europe	2.91
Latin America	1.29
Middle East	1.17
Pacific	1.02
Africa	0.53
South Asia	0.39
World average	1.66

Another way of analysing energy is in terms of end use. Countries vary, but the pattern of end-use is reasonably typical of an industrial nation. About 40 % of the energy demand is for low-temperature heating and space cooling; about 20 % for high-temperature heating, i.e. above the boiling point of water; mainly industrial. About 30 % is used in transport. Only about 5-10 % is used for activities that require electricity, i.e. lighting, electrolysis, electronic equipment and so on. In developing countries a greater percentage of energy use goes in cooking and less in space heating but otherwise the distribution is similar. In both developing and developed countries, the average spend on energy per person is about 5 % of annual income.

If current usage continues on the present trend, by the mid 21st century resources of oil and gas will be coming under pressure. This will encourage the use of (currently) marginal resources and further exploration (Rockall, Falklands) but these will be more expensive to exploit and prices will rise. Coal stocks will last at least a couple of centuries with current reserves. Total reserves (probably) give about 1000 years. There still remain local shortages. Japan is the obvious example. Europe (apart from U.K. and Norway) has no oil and only a little gas. This means that the industrialised world is becoming increasingly reliant on imports from the developing world, above all from the Middle East. Moreover, none of this addresses the environmental problems consequent on energy use.

The OECD has offered an analysis of how these problems occur. One point is worth noting. Reference is frequently made to the fuel cycle. Most fuels are not part of a cycle (except for renewables and bio-fuels). Fossil fuels are irreversibly consumed and not replaced. With this proviso, the OECD 'fuel cycle' contains the following stages:

- exploration (e.g. geological studies, prospecting, test drillings),
- harvesting (mining, drilling and – for bio-fuels real harvesting),
- processing (extraction of the fossil fuel and any purification process needed),
- transport (fuels are rarely close to the point of consumption),
- storage (where possible; note that electricity cannot be stored as electricity but must be converted into another form of energy for storage),
- marketing, and
- end use.

One must then consider the detailed implementation of these steps. The OECD considers the result of these activities under the term residuals – not only the waste released into the environment, but also the material removed from the environment and structural changes to the environment. These are given as:

- consumption of resources needed to obtain the primary energy supply (e.g. equipment to build and maintain mines, energy to run the mine, land taken).
- Effluents (these can be material such as solids, liquids, gases or non-material such as heat, noise).
- Physical transformations (land filling, erecting buildings).
- Social/political (changes in employment, populations).

In fact this mixes up different kinds of effects. Some of them are perhaps better described as impacts. It is obvious that there is no one way of tackling a list as varied as this.

An alternative is to use a systems approach. We draw a system boundary around the components of a system that undoubtedly interact with each other and call everything outside that the system environment. So far, we are merely making explicit a judgement that we are probably making implicitly anyway. Such a distinction is necessarily a matter for judgement and may vary depending on exactly what we are interested in. A simple example is a heating system where we might wish to draw the boundary round the heater and fuel and consider the atmosphere as the system environment. This enables us to define the question of what is the effect of the system on the local system environment; the bubble of influence.

8.1 Fossil fuels

These are, and are likely to remain, the major source of energy for many years despite the increasing concerns about global warming. Thermal power stations (be they fossil or nuclear) have a heating element, a boiler and a turbine. We know that the Carnot efficiency, η, of a heat engine is given by:

$$\eta = (T_\mathrm{h} - T_\mathrm{c})/T_\mathrm{h},$$

where T_h is the temperature of the hot reservoir and T_c is the temperature of the cold reservoir. The cold reservoir is the environment (usually a river) and so, in practice, has a temperature of about 15^0C (288 K). The hot reservoir can get up to 600-700^0C (900 K or thereabouts). This gives Carnot efficiencies of the order of 70 %. Real power stations are not reversible Carnot engines and cannot reach efficiencies of anything like this. A total efficiency of 42 % would be reckoned to be good for a normal coal-fired power-station. This comes from the following:

- suitably designed boilers can reach efficiencies of 90 % in the transfer of heat to the working fluid,
- a typical power station will have three steam turbines (high pressure – about 160 bar, medium pressure –about 2 bar and low pressure – about 0.035 bar). The steam exiting the turbines is used to heat the water inlet to the boiler. The total efficiency of this setup is about 48 %.
- There are some (fairly small) mechanical losses and miscellaneous hot losses.

One device to improve the utilisation of the system is to construct a combined heat and power generator (CHP). The overall efficiency of this, η_{CHP} is defined as:

$$\eta_{\mathrm{CHP}} = \frac{\text{net power output} + \text{heat recovered}}{\text{energy input}} \times 100\% .$$

The gains are obvious: about 2/3 of the waste heat can be recovered. It works best for a fixed balance of heat and power. In practical situations the ratio required may vary. There is significant extra investment required in plant and heat pipelines. If the CHP system is also a district heating scheme, there may be problems of noise and pollution since the plant must be close to the district it serves.

8.2 Nuclear power

In principle, three methods of obtaining power from nuclear energy have been considered: (i) thermal reactors (which obtain energy from the fission of isotopes of uranium or thorium), (ii) breeder reactors (which in addition to doing this also convert the natural uranium isotope U^{238} to a fissile isotope of plutonium, Pu^{239}), and (iii) fusion reactors (which use the reaction $^2D_1 + {}^3T_1 \rightarrow {}^4He_2 + {}^1n_o$ + energy; the tritium being obtained from lithium by neutron bombardment). Only the first has been used to obtain power on a commercial scale. A few large-scale experimental breeder reactors have been built. A number of experimental fusion 'reactors' has been built. The most successful of these (the tokamak at Culham) have just about broken even, i.e. the fusion reactions inside the apparatus have delivered as much energy as that required to create the plasma needed to generate them.

A large number of designs of nuclear reactor have been proposed. All of them have the same basic features:

- The nuclear fuel. In a fission reactor, the nuclei of uranium or thorium are broken up into two approximately equal parts by neutron bombardment. One typical reaction sequence is:

$$^1_0\mathrm{n} + {}^{235}_{92}\mathrm{U} \rightarrow {}^{236}_{92}\mathrm{U} \rightarrow {}^{141}_{56}\mathrm{Ba} + {}^{92}_{36}\mathrm{Kr} + 3^1_0\mathrm{n} + 175\mathrm{MeV} .$$

The important points are: (i) the large amount of energy released, (ii) the fact that you get more neutrons back than are consumed (which gives rise to the possibility of a chain reaction), and (iii) the unavoidable production of radio-active isotopes. The neutrons produced by this reaction (prompt neutrons) are not the only ones produced. The immediate fission products also release neutrons through a beta-decay process on a timescale of seconds to minutes (delayed neutrons). An example of a reaction is

$$ {}^{87}_{35}\mathrm{Br} \rightarrow {}^{87}_{36}\mathrm{Kr} + {}_{-1}e \rightarrow {}^{86}_{36}\mathrm{Kr} + {}^{1}_{0}\mathrm{n} \, . $$

It is this fact that makes a nuclear reactor controllable. A chain reaction that relies on the prompt neutrons alone is a nuclear explosion. If maintaining the reaction relies on the existence of delayed neutrons, the process is controllable. The condition for establishing a chain reaction is a nuclear reactor is the criticality factor, *k*, defined as:

$$ k = \frac{\text{neutrons produced in the nth generation}}{\text{neutrons produced in the (n - 1)th generation}} \, . $$

Control *k* needs to be close to 1. Thus we require something to absorb enough of the neutrons to ensure that this is the case.

- The moderator. In a thermal reactor, this is done by slowing the neutrons down so that they are more likely to be absorbed by a nucleus (^{238}U or ^{235}U) rather than break it up. Neutrons are slowed down by allowing them to hit the atoms of a moderator (light nuclei that take away the initial kinetic energy of the neutrons). The neutrons are slowed to the velocities appropriate to the thermal motion of a gas. Two moderators have been used: graphite and water.
- A method of getting the heat from the reactor core. This is a (fairly) conventional piece of chemical engineering involving heat transfer circuits and boilers.
- There are two basic problems with this method of energy generation: (i) the possibility of a major release of radioactivity; the Chernobyl explosion is the clearest example, (ii) the problem of disposal of the radioactive waste. Radioactive waste is conventionally divided into three types:
  - Low level, which includes the waste produced by therapy in hospitals,
  - Medium level. An example would be the fabric, particularly the metals of a reactor.
  - High level. Medium and long-lived decay products of the nuclear reaction, including actinides produced by neutron capture rather than nuclear fission.

Most attention has been devoted to the last category; where the methods under consideration include incorporating the decay products in glass or artificial minerals and then burying them in deep repositories. However, the much larger volumes of low-level and medium-level waste are also a significant problem.

8.3 Renewable resources

These amount, in the end, to harnessing solar energy directly or indirectly (with the exception of tidal power which in the end harnesses the rotational energy of the Earth and geothermal which harnesses the internal heat of the Earth). In principle there is a lot of solar energy: about 18000 TW falls on the Earth. The basic problem is collecting it. The energy density is very low. Current usage is as follows:

- hydroelectric, 6 % of global energy requirements,
- biomass, i.e. wood-burning, 1.5 % of global requirements,
- tidal, solar, geothermal and wind, together, provide about 0.5 % of global requirements.

In other words, only hydroelectric is making a significant contribution.

8.3.1 Hydroelectric power

The main advantage of hydroelectric power is that the energy density is high. The drawbacks are serious: hydroelectric schemes require large dams. These cause large social and ecological changes.

The basic method is simple. Water passes from a dam down a tube and through a turbine. The idea is to convert the potential energy of the water first into kinetic and then into electrical energy. If ρ is the density of water, Q is the flow-rate then P_0, the maximum power available to be generated, is given by:

$$P_0 = \rho ghQ,$$

where h is the height drop. Equivalently, one can look at the problem from the point of view of the kinetic energy. If the speed of the water is v, the power available is $\rho Qv^2/2$. We shall briefly consider one kind of turbine: the Pelton impulse turbine. Consider the case where the water from the jet is hitting the bottom cup. If the speed of the cup is v_t and the speed of the jet is v_j. Then, if we take the ideal case where the cup deflects the stream by 180^0 and there is no friction to worry about then, with respect to the frame of reference of the cup, the speed of the water jet is $(v_\mathrm{j} - v_\mathrm{t})$ both before and after the water hits the cup. Thus the change of momentum of the fluid (and thus the force exerted on the cup) is

$$F = 2\rho Q\left(v_\mathrm{j} - v_\mathrm{t}\right).$$

The power transferred is

$$P = Fv_\mathrm{t} = 2\rho Q\left(v_\mathrm{j} - v_\mathrm{t}\right)v_\mathrm{t}\,.$$

This is a maximum for $v_\mathrm{j}/v_\mathrm{t} = 0.5$ in which case the power output is the kinetic energy of the water in the jet, i.e. the turbine is 100 % efficient. Real efficiencies vary from 50 % (for small units) to 90 % for large commercial systems.

8.3.2 Tidal power

This is similar to hydroelectric power except that it is not a continuous source. In principle there is a lot of energy available but there is the problem both of energy density (how many estuaries are suitable) and of environmental problems. The largest tidal installation (and has been for many years) is at La Rance (France) with a capacity of 240 MW. The configuration of the estuary is close to ideal, but: (i) tidal barrages are expensive, (ii) it would drastically change the environment of the estuary.

The basic idea is to trap the tide behind a barrier and let the water out through a turbine at low tide. If the tidal range is R and the estuary area is A, then the mass of water trapped behind the barrier is ρAR, and the centre of gravity is $R/2$ above the low tide level. The maximum energy per tide is therefore $(\rho AR)g(R/2)$. Averaged over a tidal period of τ, this gives a mean power available of

$$<P> = \rho AR^2 g/2\tau.$$

This is too crude. It is necessary to further average the tides over a month (to allow for spring and neap tides). To get this power out requires special turbine (designed for a comparatively low head). Even so, it is not possible to get significant power out close to low tide. The total power output can be greatly increased by using the turbines as pumps close to high tide to increase the tidal head.

8.3.3 Wind power

This is not a new idea. Modern wind turbines consist of a two or three bladed propeller (33 m in diameter). The rate of power generated in a wind speed of Beaufort scale 6 (strong breeze) is 300 kW. Hence, we need a wind farm.

An example is the Fair Isle scheme in 1982, a 50 MW wind farm was built to generate electricity from winds. This provides 90 % of the usage of the island. This shows that wind-power can be a preferred choice in some cases. It does not show that it is a major contender. The main problem is that the peak wind and the peak demand are unlikely to coincide. Further, large wind farms are unpopular. They are very visible and often on sites of considerable natural beauty.

Again, the basic physics is simple. The kinetic energy in a unit volume of air is given by $\rho v^2/2$, where ρ is the air density and v the wind velocity. The volume of air passing cross-section A perpendicular to the wind velocity in time t is given by vAt (or v per unit cross-section per unit time). If the angle of the wind direction to the normal of the cross section defined by the wind turbine, is β, the volume of air passing through unit area of the turbine cross-section is $v\cos\beta$. Hence the maximum power per unit area is

$$P_0/A = (\rho u^3 \cos\beta)/2.$$

In principle, the maximum power available occurs when $\cos\beta = 1$ and then

$$P_0/A = \rho u^3/2.$$

In practice, only a small fraction of this is really available, and the coefficient of performance, C_P, is introduced.

It is possible to obtain a theoretical upper bound to C_P, the *Betz limit*. We consider airstreams at constant velocity passing through and by the turbine. The turbine rotor is considered as an actuator disc. There is a change of pressure across the turbine as energy is extracted and a decrease in the linear momentum of the wind. It can be shown that the coefficient of performance, C_P, is given by

$$C_P = 4a\,(1-a)^2,$$

where $a = (v_{\text{wind}} - v_{\text{back}})/2v_{\text{wind}}$ is the fractional decrease in the wind speed at the turbine, called *interference factor*.

The maximum value of C_P occurs for $a = 1/3$, when $C_P = 0.59$. In practice, a modern wind turbine can manage a C_P value of about 0.4. Given this factor, the generation of energy is about 95 % efficient, i.e. the efficiency of the turbine generator itself. Wind systems are of most use in niche areas where connecting to the grid is expensive. Since wind energy is variable they need a backup, i.e. a battery or the grid itself. Of all renewables, wind is the closest to being competitive with conventional fossil fuels. Most commercial projects are based on wind farms. The individual generators must be separated by about ten times the blade length and a further buffer zone round the farm is required.

8.3.4 Wave power

In principle, large amounts of energy can be obtained from waves. Most devices are designed to extract energy from deep water waves, where the mean depth of the seabed, D, is greater than half the wavelength of the wave, λ. The basic properties of such waves are:

- the surface waves are sine waves of irregular phase and direction.
- The motion of any particle of water is circular. The waves move but the water does not.
- Water on the surface stays on the surface.
- The amplitude of the motions of the water particles decreases exponentially with depth.
- The amplitude of the surface wave is independent of the wavelength or velocity.
- Wave breaks when the slope of the surface is about 1 in 7.

The power in a wave comes from the change in potential energy of the water as it rotates on the circular paths beneath the surface. It can be shown that the power carried forward by a wave is given by:

$$P = \rho g^2 \mathrm{A}^2 T/8\pi,$$

where A is the amplitude of the wave at the surface, and T is the period of the wave. Two devices intended to extract this power are the Salter duck and the oscillating column. The Salter duck consists of a cone that oscillates with the waves and is connected to a rotary pump that drives a generator. The oscillating column uses the wave to drive a trapped air column past a turbine. A number of prototypes have been tried but the economics of the power generation is not yet good enough for a full commercial trial.

8.3.5 Biomass

Second in importance to hydro (at present) is the use of biomass as a renewable fuel. The term covers domestic, industrial and agricultural dry waste material, wet waste material and crops. The essential difference between this and fossil fuels is that the biomass cycle is a true cycle provided that for each plant used as fuel a replacement is planted. Examples of bio-fuels include:

- gaseous bio-fuels are used for: (i) heating and cooking, and (ii) in engines for electricity and heat generation, and occasionally for transport. Examples include biogas (CH_4 and CO_2) from anaerobic digestion of plant and animal wastes, and producer gas (CO and H_2) from gasification of plants, wood and wastes.
- Liquid bio-fuels are used mainly for transport fuels. Examples are oils from crop

seeds (e.g. rape, sunflower), esters produced from such oil, ethanol from fermentation and distillation, and methanol from acidification and distillation of woody crops.
- Solid bio-fuels. Examples are wood from plantations, forest cuttings, timber yards and other wastes, charcoal from pyrolysis, and refuse derived fuels, e.g. compressed pellets.

A major user of biomass is Brazil: the source being waste from the sugar-cane industry. Bagasse (residue after crushing the cane), and barbojo (leaves of the cane). Perhaps 67 % of the 80 sugar-cane producing countries can use this as fuel.

8.3.6 Solar power

The simplest way of making use of energy from the Sun is to turn it into heat. A black surface directly facing full sunlight can absorb 1 kW/m^2.

Solar energy can be either direct or diffuse. Only direct radiation can be concentrated. The energy received from the Sun at a given place depends on the latitude, time of day and season. If you wish to maximise the solar energy absorbed on a surface, you must slant it so that its normal points at the Sun. For best results, the orientation should change during the day, and even correcting the angle from day to day to allow for the declination with the seasons. It is usually not worth the cost to do this. It is enough to set the surface to face the Sun at noon and fix the angle with the horizontal to this. For some applications a concentrator can be used. The maximum temperature value is about 1150 K. In practice, temperatures of 950 K are obtainable. This is high enough for efficient electricity generation. Also, it is the basis of the solar furnace.

8.3.6.1 Solar collector

The solar radiation is absorbed and the absorber is heated up to about 80^{0}C. The absorber should be painted black (absorption coefficient nearly unity). The heat is then transferred to water tubes. The system can then be run in a similar way to a standard boiler. A certain amount of heat is lost in conduction to the supports, convection and radiation. Radiation loss is the most significant. Recall the Stefan-Boltzmann law. If the temperature is 80^{0}C (353 K), then the radiated energy is 880 W/m^2. This is a substantial fraction of the total available, S = 1353 W/m^2. The main problem is therefore to overcome the radiation losses. Black chrome is used as the absorber. This has a high absorption coefficient for the wavelengths of solar (incoming) radiation, but a low absorption coefficient for the outgoing terrestrial radiation. Also, glazing is placed above the collector to reduce convection. Modern solar collectors have an area of 3 m^2. The covering glass plate has a transmission coefficient of 90 %.

Let us consider this in more detail. In the arrangement described above, of the incoming radiation flux, SA (where A is the area of the plate), a fraction t is transmitted through the glass covering and a further fraction, a is absorbed. Ignoring losses from convection, radiation and conduction, the net power absorbed, P, is therefore $SAta$. If we assume that this is all transferred to the fluid in the heat exchanged, then the heat gained per unit time by the water flowing through the heat collector is

$$P = C\frac{\mathrm{d}m}{\mathrm{d}t}\left(T_{\mathrm{out}} - T_{\mathrm{in}}\right) = C\rho Q\left(T_{\mathrm{out}} - T_{\mathrm{in}}\right),$$

where dm/dt is the mass rate of flow of the water, T_{out} is the temperature at the outlet of the collector and T_{in} is the temperature at the inlet. In the second identity ρ is the density and Q the flow rate. This is, of course, an idealisation. There are bound to be some losses. These can be expressed as an effective thermal resistance of the collector. We define this resistance R as

$$\text{Energy losses} = (T_{\mathrm{out}} - T_{\mathrm{in}})/R.$$

The capture efficiency, n, of the system is defined as the fraction of solar power impinging on the device that is converted into useful heat. The heat in the pipes is thus given by

$$taSA - (T_{\mathrm{out}} - T_{\mathrm{in}})/R = nSA$$

and we have

$$n = ta - (T_{\mathrm{out}} - T_{\mathrm{in}})/(RAS).$$

It is common to define U, the energy transfer coefficient as $1/RA$ which gives finally

$$n = ta - U(T_{\mathrm{p}} - T_{\mathrm{a}})/S.$$

This is the Hottel-Whillier-Bliss equation. The parameters (ta) and U are usually used to characterise a particular solar water heater.

Consider an example where we put in a bit more detail. A set of numbers are given in the table below:

Solar Constant	S	700 $\mathrm{Wm^{-2}}$
Temperature of the collector	T_{out}	35^0C
Temperature of the cover	T_{c}	35^0C
Temperature of the surrounding air	T_{air}	5^0C
Temperature of the sky	T_{sky}	-10^0C
Transmission and absorption coefficients of the collector	ta	0.9
Emissivity of glass	ε	0.1
Specific heat of water	C	4184 $\mathrm{Jkg^{-1}K^{-1}}$
Convection coefficient of the system	h	2.82$\mathrm{Wm^{-2}K^{-1}}$
Area of collector	A	3 $\mathrm{m^2}$
Rate of flow of water through the collector	dm/dt	0.042 kg/s
Input temperature of the water	T_{in}	30^0C

The losses due to absorption and transmission through the plate are 133 $\mathrm{Wm^{-2}}$. If the over is at the outlet temperature of the water, then the net radiation losses are given by $\varepsilon\sigma\left(T_{\mathrm{c}}^4 - T_{\mathrm{sky}}^4\right) = 23.9\,\mathrm{Wm^{-2}}$ and the convection losses are $h\left(T_{\mathrm{c}} - T_{\mathrm{air}}\right) = 84\,\mathrm{Wm^{-2}}$. This leaves 459 $\mathrm{Wm^{-2}}$ to heat the water. If there are 2 % losses in the heat exchanges, this costs another 9.2 $\mathrm{Wm^{-2}}$ leaving about 450 $\mathrm{Wm^{-2}}$. If the water is heated by 5^0C, this means that the two parameters characterising the heater are ta = 0.81 and U = (23.9 + 84 + 9.2)/5 = 23.4 $\mathrm{Wm^{-2}K^{-1}}$. A decent value would be in the range 6-8 $\mathrm{Wm^{-2}K^{-1}}$. This system could heat about 232 litres of water per hour by 5^0C.

8.3.6.2 Solar photovoltaic

Solar radiation can be converted directly into electricity by solar photovoltaic cells. Examples of use include watches/calculators, solar arrays for space craft. Practical cells made of amorphous silicon. Efficiencies are 10-20 %. Hence a panel of cells 1 m^2 facing full sunlight will give 100-200 W, i.e. large area required for significant amounts of power. In 1990, world capacity was about 50 MW. To meet WEC (World Energy Council) projections we would need 1000 times this amount by 2020.

8.4 Energy demand and conservation

The other obvious approach is to reduce the energy demand. Energy conservation is often specific to the process and is difficult to treat in a general fashion. Since a large part of the energy demand in most countries is for space heating, it makes sense to consider this as our basic example of conservation. In this case, we must consider a number of issues. There is the basic physics of heat transfer and thermal insulation. There is the trade-offs that are essential between an energy-efficient building and other conditions that must be considered.

8.4.1 Heat transfer and thermal insulation

Heat can be transferred by: (i) conduction, (ii) convection and (iii) radiation. Thermal insulation reduces the transfer of heat from one point to another, especially from the interior of a building to the outside. Effective insulation reduces the amount of heat that has to be supplied.

The effectiveness of an insulator is measured by its thermal conductivity, κ. This is defined from the Fourier heat equation. Fourier asserted that heat flow per unit cross-sectional area, J, is proportional to the temperature gradient, i.e.

$$J = -\kappa \frac{\mathrm{d}T}{\mathrm{d}x}.$$

The negative sign states that heat flows down the temperature gradient: from hot to cold. Typical values of κ ($\mathrm{Wm^{-1}K^{-1}}$) are Al (160 $\mathrm{Wm^{-1}K^{-1}}$), steel (50 $\mathrm{Wm^{-1}K^{-1}}$), brick (0.84 $\mathrm{Wm^{-1}K^{-1}}$). Air is a good insulator but should not be in motion or convection will transfer heat.

Heat transfer by convection is usually divided into: (i) natural convection (where the fluid moves without any forcing), and (ii) forced convection (where the fluid is moved by draughts). Natural convection is described by Newton's law of cooling:

$$J = k(T - T_0),$$

where T is the temperature of the object, T_0 is the ambient temperature, and k is the convection coefficient. Heat loss by radiation is given by the Stefan-Boltzmann law.

Since in all cases, we are interested in the problem of heat transfer through the parts of building, it is convenient to choose a measure that does not depend on the mechanism. Engineers use a measurement to quantify the thermal behaviour of a structural element. The U-value (or thermal transmittance coefficient) is the rate at which heat flows through an area of 1 m^2 of an element when the temperature change across it is 1^0C. Clearly, this is most easily related to the thermal conductivity. We can express the heat transfer equation as a finite difference equation,

$$J = -\kappa \frac{\Delta T}{\Delta x}.$$

Thus, from the definition of U given above,

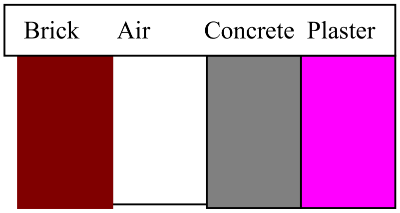
$$U = \frac{J}{\Delta T} = -\frac{\kappa}{\Delta x}.$$

Radiation losses are forced into this form. For convection (as treated above) $U = k$. The total U-value for a complex system is obtained by using Kirchoff's law to sum the resistances. We define the thermal resistance $R = \kappa^{-1}\Delta x$. There are also resistances due to the presence of interfaces. These are given by heat transfer coefficients, h. Thus, the total U-value for a complex wall with heat transfer coefficients h_{in} and h_{out} for transfer into and out of the wall respectively is given by:

$$1/U = 1/h_{\mathrm{in}} + \sum\nolimits_j R_j + 1/h_{\mathrm{out}}.$$

For example, a single window has $U = 5.7$ Wm^{2}K^{-1}. Since a double-glazed window has a 2 cm air space, U is (roughly) halved. The lower the value of U, the better the insulation.

Example: A cavity wall.



Let us assume that the thermal conductivities are:
brick: 0.8 Wm^{-1}K^{-1},
concrete, 0.2 Wm^{-1}K^{-1},
plaster: 0.17 Wm^{-1}K^{-1},
and that all the materials are 0.1 m thick.
The thermal resistances are: plaster/air interface: 0.12 m^{2}KW^{-1}, brick/air interface: 0.16 m^{2}KW^{-1}, cavity (including interfaces): 0.19 m^{2}KW^{-1}. The thermal resistances of the materials are: brick = 0.1/0.8 = 0.125, concrete = 0.1/0.2 = 0.5, plaster = 0.1/0.17 = 0.59. Hence, the U value is given by

$$U = 1/(0.125+0.5+0.59+0.19+0.06+0.12) = 0.63$$

ignoring the two interfacial transfer coefficients. Typical values for these would be h_{in} = 7 Wm^{-2}K^{-1} and h_{out} = 18 Wm^{-2}K^{-1}. This gives a final U-value of 0.56 Wm^{-2}K^{-1}.

8.4.2 Heat loss in buildings

The amount of energy lost from a particular building depends on the following loss factors;

- insulation of the building,
- area of external surfaces of the building,
- temperature difference between internal and external environments,

- air change rate for ventilation,
- degree of exposure to climatic effects, such as wind.

Each of these can be considered in terms of U-values. It is convenient to divide these into two main kinds: (i) fabric loss, and (ii) ventilation loss. Fabric loss is the heat loss through the external 'skin' of the buildings (walls, floor, ceiling, windows) and can be written as

$$P = UA(T_{\mathrm{in}} - T_{\mathrm{air}}),$$

where P is the power loss, U is the effective U-value for the building and $(T_{in} - T_{air})$ is the temperature drop across the 'skin' of the building. The largest U-values are usually for the windows, but the largest losses are usually through the walls. The effect of the much greater area dominates the effect of the U-value. Ventilation losses are the other major contribution to heat losses. Ventilation is also an essential part of building design. The average person (mass 84 kg) requires oxygen at a rate of 50 ml/min per kg of body weight of the person, i.e. for the average person, 4200 ml/min. Obviously the exact amount depends on what you are doing: 2000-3000 ml/min (at rest) to 6000 ml/min (athlete running). Heat is lost through ventilation (energy taken away by the convecting air). This is given by:

$$Q = mC_{\mathrm{P}}\Delta T,$$

where m is the mass of air, C_{P} is the specific heat at constant pressure and ΔT is the temperature difference between the inside and outside of the building. Consider a room of volume V. If it takes t seconds to replace all the air in the room, then the rate of heat loss is $V\rho C_{\mathrm{P}}\Delta T/t$ (where ρ is the density of air).

Buildings also have a number of sources of heat (apart from the explicit heating system). These include:

- solar heat through windows, walls and roof,
- body heat from inhabitants,
- heat from lighting equipment and electrical appliances (TV, fridges and so on),
- heat from cooking processes, and
- water heating.

Added together, these can be far from trivial and should be taken into account when designing the heating system required.

Questions:

- The solar constant has a value which varies from 1420 W/m^2 in December to 1330 W/m^2 in June. Suggest why the solar constant is different in December from June.