A short introduction to physical oceanography

for the course and tutorial 'Introduction to Marine Microbiology', updated winter semester 2007/08, University of Bremen

Christoph Völker

Alfred-Wegener-Institut für Polar- und Meeresforschung, Bremerhaven, Germany

1 Introduction

The physical state of the oceans with its circulation patters, temperature distribution, light field, etc. influences life in the oceans in a variety of ways. The aim of this course is to explain some main features of the physics of the ocean, as much as can be done in a few lecture hours. The focus is on the large-scale features, such as the Gulf stream of the global heat flux. Some important things are definively missing; especially small scale processes, such as turbulence and waves near the ocean surface, would require a lecture of their own.

It is obvious that a short course such as this cannot proceed by starting from a few general physical principles to finally arrive at the observed patters of ocean circulation, wave motion etc., as is the aim of 'dynamical physical oceanography'. Here instead, I have aimed to mix what is usually called 'descriptive physical oceanography' with some explanation of general principles and their application. As much as possible is explained without using formulae; however in some sections (especially the section on the balance of forces) the main arguments are presented in their mathematical form as well, in order to show how the same thought can be expressed in a formula. But I have tried to avoid progressing by purely mathematical reasoning. For those who are interested in diving a bit deeper into the subject of physical oceanography, here are some suggestions for further reading:

Open university oceanography course team (1989, 2nd ed. 1991): Ocean Circulation. Pergamon Press, Oxford This book is a good first introduction into the circulation of the oceans and its driving forces.

J. Marshall, R.A. Plumb (2007): Atmosphere, Ocean and Climate Dynamics: An Introductory Text. Elsevier Academic Press, Boston A very good textbook to continue after the later part of this script! Unfortunately so new, it only exists as expensive hardback.

Keith Stowe (1979, 2nd ed. 1983): Ocean Science. John Wiley and sons, New York A very broad introduction from the seafloor to surface waves. The focus is more on geology and chemistry than physics, but contains a good non-mathematical introduction to the Coriolis force.

G. Pickard and W.J. Emery (4th ed. 1982): Descriptive physical oceanography. Pergamon Press, Oxford Introduces the physical properties of the ocean and how they are measured, and the regional distribution of the properties.

S. Pond and G. Pickard (1978): Introductory Dynamic Oceanography. Pergamon Press, Oxford On an introductory, but still mathematical level, introduces physical laws and their mathematical formulation relevant for oceanography.

Also, an increasingly large number of oceanographic lectures is placed on the web. Some of that material complements this rather short introduction well. Here is a pick of web-sites that I found useful:

'An Introduction to Physical Oceanography' by Matthias Tomczak:
`http://www.cmima.csic.es/mirror/mattom/IntroOc/newstart.html`
In my opinion the best e-learning web-site for oceanography. (the address given is the European mirror site, the original lecture is on a server in Australia)

'Physics of Atmospheres and Oceans' by John Marshall:
`http://ocw.mit.edu/Ocw-Web/Earth--Atmosphere--and-Planetary-Sciences/12-003Fall-2007/CourseHome/index.html`
Combines well with the book by Marshall and Plumb; an introduction to large-scale dynamics with many lab exercises. Part of the MIT's 'open course ware', where lectures at MIT are published. (If the link has changed, start from `http://ocw.mit.edu`)

A lecture on chemical oceanography by James W. Murray:
`http://www.ocean.washington.edu/courses/oc400/lecture_notes.html`
The first chapters of this lecture cover the subject of our section 2 in much more detail.

A lecture on Climate Change by Wolfgang H. Berger:
`http://earthguide.ucsd.edu/virtualmuseum/climatechange1/cc1syllabus.shtml`
Discusses the energy balance of the earth and how it is affected by trace gases in the atmosphere in much more detail than we do in section 4.

2 Properties of water and seawater

2.1 The molecular structure of water and its consequences

Life on earth is made possible by water. This is because water has a number of unusual properties, such as its high heat capacity, or its dissolving capability. Most of these properties are rooted in the molecular structure of water.

Some unique properties of water that are important for the existence of life on earth

Water has the highest heat capacity of all solids and liquids, except liquid NH_3. It takes about 4000 J to heat a kg of water by 1°C, and the same amount is released when water cools by 1°C. Water thus has a large thermal buffer capacity. The oceans act as a climate thermostat, and ocean currents can carry huge amounts of heat.

Water has the highest heat of evaporation of all liquids. It takes about $2.4 \cdot 10^6$ J to evaporate a kg of seawater. This additionally creates thermostating capacity. Mammals cool by sweating.

Water has an unusually high melting and boiling point, when compared to other hydrides of the group VIA from the periodic table, such as H_2S, H_2Se, *...* This causes that water on earth exists predominantly in liquid form.
Water has the highest dielectric constant of all substances except H_2O_2 and HCN. Many substances that can be split into positively and negatively charged ions easily dissolve in water, making the basic chemical reactions of life possible.

A water molecule consists of hydrogen atoms that are each bound to a central oxygen atom by sharing an electron in one of four sp^3 hybrid orbits of the oxygen atom. The two H-O bonds form an angle of 105 degrees (Figure 1), i.e. the three atoms are not aligned along a line but rather form a triangle. The charge of the molecule is not evenly distributed, i.e. it is polar and acts as a little dipole with the negative end closer to the oxygen atom and the positive end between the two H atoms.

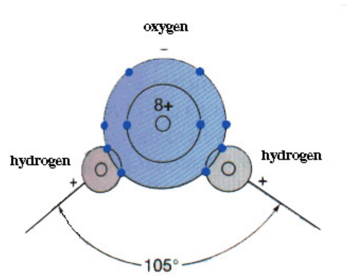


Figure 1: *Structure of a single water molecule. Figure taken from the online lecture by E. Spanier.*

Because of this polarity water is attracted both to anions and cations. When a salt dissolves in water its ions are hydrated, i.e. surrounded with water molecules. This hydration sphere is hold together by attraction between the charged ion and one of the polar ends of the water molecules.

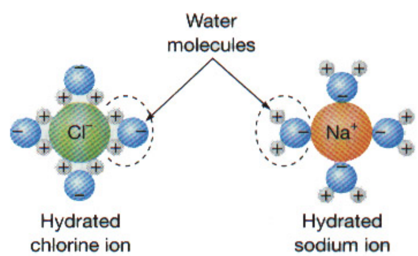


Figure 2: *Dissolved ions in water are surrounded by water molecules in a so-called 'hydration sphere'. Figure taken from the online lecture by E. Spanier.*

Also because of its polarity, water has a tendency to form hydrogen bonds between individual water molecules; even in liquid water the individual molecules act not independently, but tend to form clusters of molecules hold together by hydrogen bonds. This 'supramolecular structure of water' is responsible for most of its unusual properties, such as its high heat capacity, its high latent heat of evaporation, the fact that pure water at atmospheric pressure has its density maximum at 4°C, well above its freezing point of pure water, etc.

Figure 3: *Due to the formation of hydrogen bonds, liquid water shows a weak structure of small clusters on water molecules. Figure taken from the online lecture by E. Spanier.*

One important consequence of the hydrogen bonds between individual water molecules is that water has a much higher freezing and boiling point than would be expected from the size and weight of the molecule alone. If the hydrogen bonding would not exist, water would freeze at -110 °C and boil at -80 °C, and no liquid water would exist on earth.

2.2 Salinity

The distinct property of seawater is that it is salty. A first definition of salinity (we will learn more definitions later on) is the total mass of dissolved material per mass of seawater. It is therefore a dimensionless quantity (mass/mass).

Seawater has a typical salinity around 34.7 g kg^{-1} with most seawater having values between 33 and 36 g kg^{-1}.

Seawater contains a mixture of several dissociated salts, with NaCl as the most important, but with a significant contribution from MgCl and other salts. Except for a number of minor constituents, the ions occur in constant proportion:

Salt		Concentration (g/kg) in seawater of S=35	Percent of total salt
Cations			
Sodium	Na^+	10.71	30.62
Magnesium	Mg^{2+}	1.29	3.68
Calcium	Ca^{2+}	0.413	1.18
Potassium	K^+	0.385	1.10
Strontium	Sr^{2+}	0.007	0.02
Anions			
Chloride	Cl^-	19.27	55.07
Sulfate	SO_4^{2-}	2.70	7.72
Bicarbonate	HCO_3^-	0.140	0.40
Bromide	Br^-	0.067	0.19
Fluoride	Fl^-	0.035	0.10

Table 1: *The major constituents of seawater (besides water, of course!)*

The reason for the almost constant composition of the salts in seawater is that the *residence time* (see below) of these salts in the ocean is on the order of millions of years. This is much longer than it takes an average water molecule to travel from the surface of the ocean into the deep sea and back, which is on the order of thousands of years. On the timescales of million years the ocean is therefore well mixed.

What exactly is meant by the residence time?

Most of the constituents of sea salt enter the ocean as dissolved substances in river water. The residence time for a certain ion is calculated by dividing the total mass of that ion contained in the oceans by the influx (mass/year) of that ion into the ocean.
The residence time is an estimate for the average time that a salt ion that enters the ocean remains in the ocean before it leaves it again, e.g. by the formation of salt deposits or salt spray (the tiny droplets of water that form when waves break).
Here are some estimates for the residence time of the six most common ions; these numbers have to be seen with a grain of salt, because of uncertainties in the salt content of continental runoff:

Element	Residence time (years)
Sodium	55,000,000
Magnesium	13,000,000
Calcium	1,100,000
Potassium	12,000,000
Chloride	87,000,000
Sulfate	8,700,000

(Source: Periodic table of Elements in the Ocean, `http://www.mbari.org/chemsensor/pteo.htm`)

> Exercise I: River water contains much less salts than sea water. Why then does the ocean does not become less saltier all the time?

The only major constituent of sea salt whose concentration varies to some extent independent of salinity is bicarbonate HCO_3^-. This is because of the chemistry of carbonate: In water, bicarbonate can react with H^+ and OH^- ions to form carbon dioxide CO_2 (or its hydrated form H_2CO_3) and carbonate CO_3^{2-}. At seawater pH ≈ 8, bicarbonate is by far the most common of the three (bicarbonate $\approx 90\%$, carbonate $\approx 9\%$, carbonic acid $\approx 1\%$). Carbon is the major constituent of organic matter and of the $CaCO_3$ shells that some organicsm form. Biological activity therefore results in uptake of release of either of the abovementioned forms. Also, carbon dioxide as a gas can exchange with the atmosphere.

Concentration units

Concentrations have been given in mass of solute per mass of solution (e.g. g/kg). But they are also often given in molar units, i.e. the number of moles (one mole is N_A molecules, where N_A is Avogadro's number $6.02214199 \cdot 10^{23}$) of the solute per kg of solution. Modellers sometimes prefer molal units, i.e. the number of moles per litre of solution.
The number of moles of a substance can be obtained by dividing its mass by its molecular mass, which is defined as the mass of a mole of the respective substance.

Example: The typical concentration of Na^+ in seawater is 10.71 g/kg. The molecular mass of Na^+ is 23 g/mol. The molar concentration of Na^+ is therefore $10.71/23 = 0.466$ mol/kg.

> Exercise II: The molecular mass of water is about 18 g mol^{-1}. Approximately how many moles of water are in a kg of seawater?

The definition of salinity given above (mass of all dissolved material per mass of seawater) is not very practical to measure in practice, e.g. by evaporating the water and weighing the remainder. At low temperatures the remaining salts still contain some water in the crystal structure and at higher temperatures material gets lost and chemical reactions with the air occur.

Chlorine, however can be measured relatively easy at high accuracy (by titration with silver nitrate). Because of the constancy of the composition of sea salt, salinity can then be calculated from chlorine concentration. This was for long times the preferred method of measuring salinity.

Later, oceanographers began to use the fact that the conductivity of seawater depends on salinity (and temperature. Inverting this relation makes it possible to determine salinity from conductivity and temperature measurements. With the definition of the practical salinity scale (PSS) in 1978 salinity was then officially *defined* as a complicated function of conductivity and temperature. The definition was done in such a way that a salinity S on the PSS still corresponds to an amount of S grams of dissolved salts per kilogram of seawater.

2.3 Temperature

We usually know how to measure temperature and can relate it to our own experience of cold/warm. We also know that, when we bring two bodies into contact, heat flows from the warmer to the colder body, until they have reached equal temperature.

In some cases we will use absolute temperatures measured in Kelvin, denoted K. These are obtained from temperature on the Celsius scale (sometimes called decrees Celsius, sometimes degrees centigrade) by adding 273.155, i.e. a temperature of 0°C corresponds to 273.155 Kelvin.

There is no need here for a more scientific definition of temperature, suffice it to say here that in liquid water, absolute temperature is very nearly proportional to internal energy.

2.4 Density

Although the density of seawater varies by just a few percent, it is an extremely important variable in oceanography. This is because even small differences of density can drive strong currents.

The ocean is usually stably stratified, i.e. density increases with depth. If this is not the case, i.e. if denser water lies atop less dense water, vertical motion sets in until the stratification becomes marginally stable again. Horizontal density differences can also drive motion, because they influence the distribution of pressure and by that also the ocean currents (see Chapter 7.2 and 7.3 of this script).

The density of seawater ρ varies between 1000 kg m^{-3} in very fresh estuaries and 1070 kg m^{-3} in deep sea trenches. Most surface seawater has $\rho \approx 1025 \pm 2$ kg m^{-3}. Because the variation is only in the last two digits, oceanographers often use a quantity $\sigma = (\rho - 1000)$ for brevity. A value of $\sigma = 25$ (it should have the unit kg m^{-3}, but the unit is often left away) corresponds to a density of 1025 kg m^{-3}.

Density can be measured directly e.g. with the the help of aerometers, glass devices filled with air that are constructed such that they float; the depth of immersion is a function of fluid density. However, this is a tricky measurement and is almost never done in practice. Instead, density is calculated from salinity, temperature and pressure, using a complicated empirical formula known as the International Equation of State for Seawater that has been published by the UNESCO in 1980. $\sigma(S, T, p)$ increases with *increasing* salinity S and pressure p, and with *decreasing* temperature T.

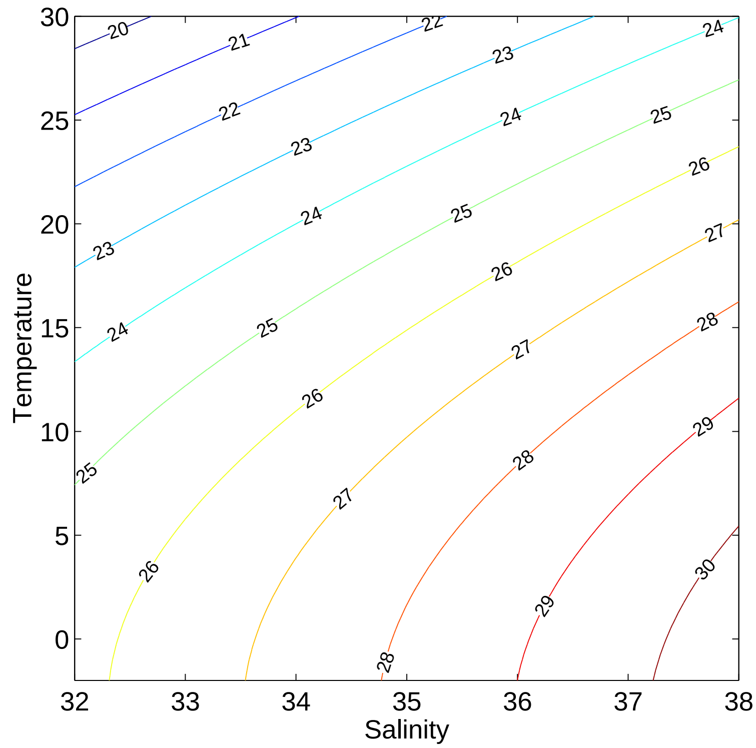


Figure 4: *Density of seawater (expressed as $\sigma(S, T, p = 0)$) at sea surface pressure.*

Figure 4 shows the variation of density with salinity and temperature at surface pressure. Density increases almost uniformly with increasing salinity, i.e. the same increase of salinity leads to almost the same increase in density irrespective of the initial salinity and temperature. In contrast, an increase in temperature leads to a different decrease in density at low temperatures than at high temperature. This nonlinearity causes that the lines of constant density in Figure 4 are curved.

Exercise III: Judging from Figure 4, does an increase in temperature by one degree lead to a stronger density decrease when the initial temperature is higher or when it is lower? How is that related to the curvature of the lines of constant density?

The increasing pressure with depth leads to a compression of the seawater. However, water, in contrast to gases, is almost incompressible, i.e. although pressure in the deep sea is several hundred times the pressure at the sea surface (approximately one atmospheric pressure per 10 meters depth, Chapter 7.1), the corresponding density increase is just a few percent.

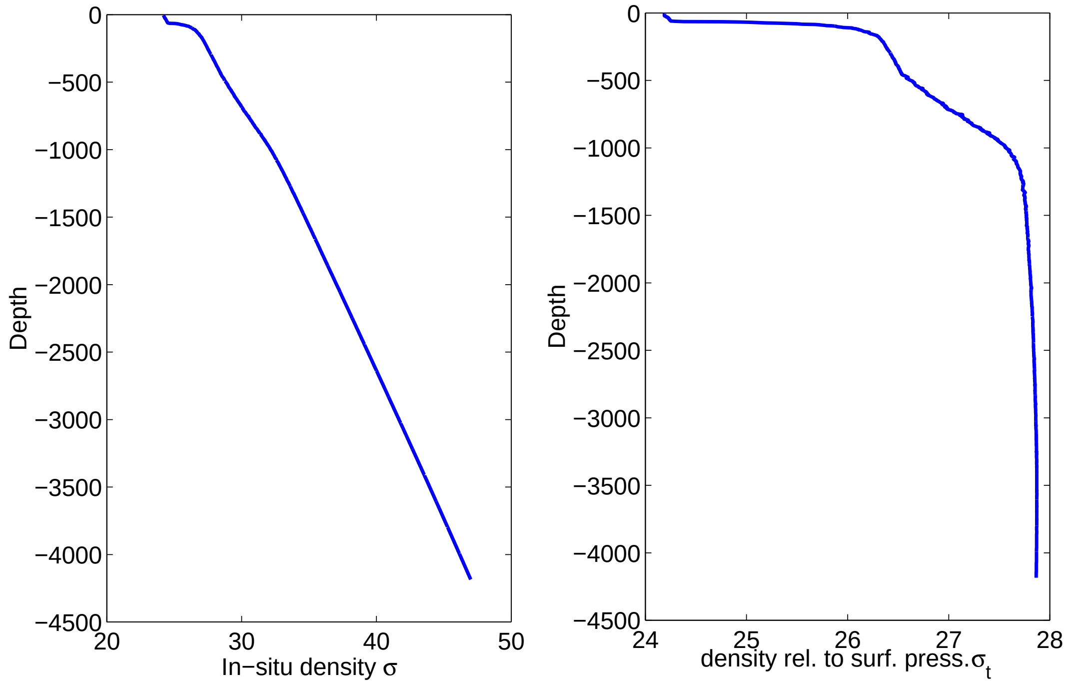


Figure 5: *Left panel: Variation of density ($\sigma(S,T,p)$) with depth for a typical location in the subtropical oceans, near Bermuda. Right panel: Potential density (referenced to atmospheric pressure) σ_0 for the same location*

The left panel in Figure 5 shows a typical profile of the variation of density σ with depth. The variation in the uppermost few hundred meters is predominantly caused by temperature and salinity differences. But the almost linear increase of density with depth in the deep ocean is caused by the compression due to increasing pressure.

To distinguish between density differences caused by temperature and salinity from those caused by pressure oceanographers often use a quantity called 'potential density'. The potential density is the density that the water *would have* at a specific reference pressure, not at the actual pressure it is experiencing. To distinguish this quantity from the *in situ* density $\sigma(S,T,p)$, it is denoted by σ_p, where p is the reference pressure. Most often, the density at surface pressure is used, denoted by σ_0.

The right panel in Figure 5 shows the profile of potential density σ_0 for the same station as above. Here it becomes evident that the temperature- and salinity-related density differences are large in the upper part of the profile, but very small in the lower part.

3 The distribution of temperature and salinity in the oceans

An important reason for studying the distributions of salinity and temperature in the oceans is that they are primarily influenced by processes occurring at the sea surface, namely the exchange of heat and water with the atmosphere. Analysis of the salinity and temperature at depth can therefore be used to infer where a certain water mass has been at the sea surface for the last time and therfore to study the paths of circulation in the deep sea. A second reason is that temperature, salinity and pressure together determine density. And since density differences are responsible for a large part of the motions in the ocean, knowledge of them helps in understanding the forces that drive motion.

3.1 Distribution of temperature at the surface

At the sea surface, the strongest variations of temperature generally occur in latitudinal direction. Temperatures decrease with increasing latitude. Variations in longitudinal direction are much smaller compared to the latitudinal variation. This results in lines of equal temperature generally being directed along latitude lines (Figure 6).

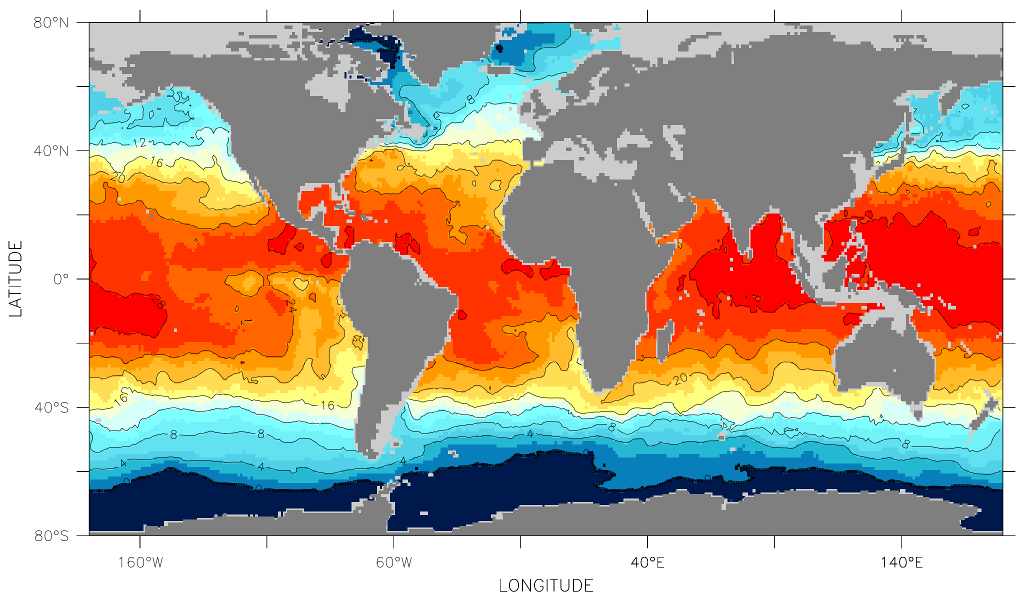
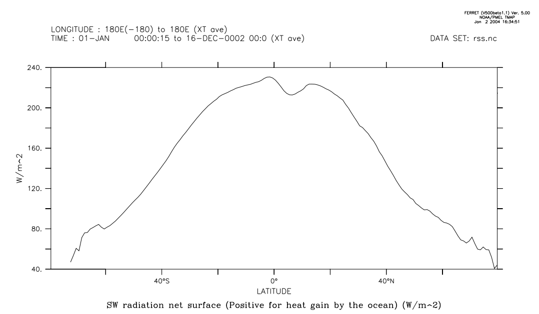


Figure 6: *Average distribution of sea surface temperature (SST) from the COARDS collection of ship observations.*

This distribution of temperature is to a large extent determined by the amount of solar radiation energy received per square meter of the surface. The incoming radiation is highest when the sun is directly overhead and decreases with increasing inclination of the surface with respect to the solar direction (i.e. with increasing latitude), because the same amount of radiation is stretched over a larger area (Figure 7).

Earths rotation axis is inclined by about 23.5° with respect to the ecliptic, the plane in which the earth moves around the sun. Without this inclination, the incoming solar radiation at the poles would be zero all year long. The inclination causes that there is no sunshine at the poles during winter, but permanent sunshine during summer. This seasonality is responsible for the fact that the annually averaged irradiance at the poles is larger than zero and therefore decreases the temperature contrast between poles and equator.



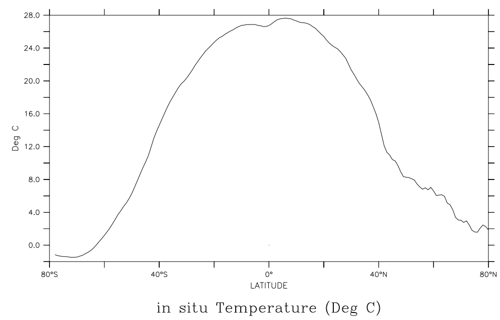


Figure 7: *Left panel: average distribution of solar irradiance at the sea surface with latitude. Right panel: average distribution of sea surface temperature with latitude.*

The correspondence between solar radiation and sea surface temperature, though, is not perfect, and this is because currents in the ocean and atmosphere can carry heat. The effects of heat transport are most notable along the continental boundaries and the equator. Figure 6 of sea surface temperature shows the influence of currents, e.g. the Gulf stream transporting warm water northward along the coast of North America in the Atlantic, see Figure 25. Off Namibia and Mauritania we can see the influence of upwelling, where cold (and often nutrient-rich) water from the deep ocean is moving upwards and reaches the sea surface.

3.2 Distribution of temperature with depth

It has been known for a long time that even in tropical regions, the ocean below a few hundred meters depth is generally cold. The average temperature of the ocean as a whole is only 3.51°C. Only the uppermost layer up to a few hundred meters depth shows seasonal fluctuations of temperature (Figure 8). Below this layer, the region of a more or less gradual decrease of temperature is called *thermocline*. Below the thermocline, temperature varies only very little with latitude or ocean basin.

Figure 9 shows a vertical section of temperature along 30 W through the whole Atlantic Ocean from 70 S (left) to 70 N. The Figure shows that there are continuous regions of low temperature extending from the surface at high latitudes to the interior of the deep equatorial ocean and beyond, and suggest that the cold water in the deep ocean has its origins at the surface in high latitudes. This is indeed the case. In fact the two main water masses filling the deep ocean have their origin in the polar and subpolar North Atlantic (North Atlantic Deep Water, NADW) and along the boundaries of the Antarctic continent (Antarctic Bottom Water, AABW).

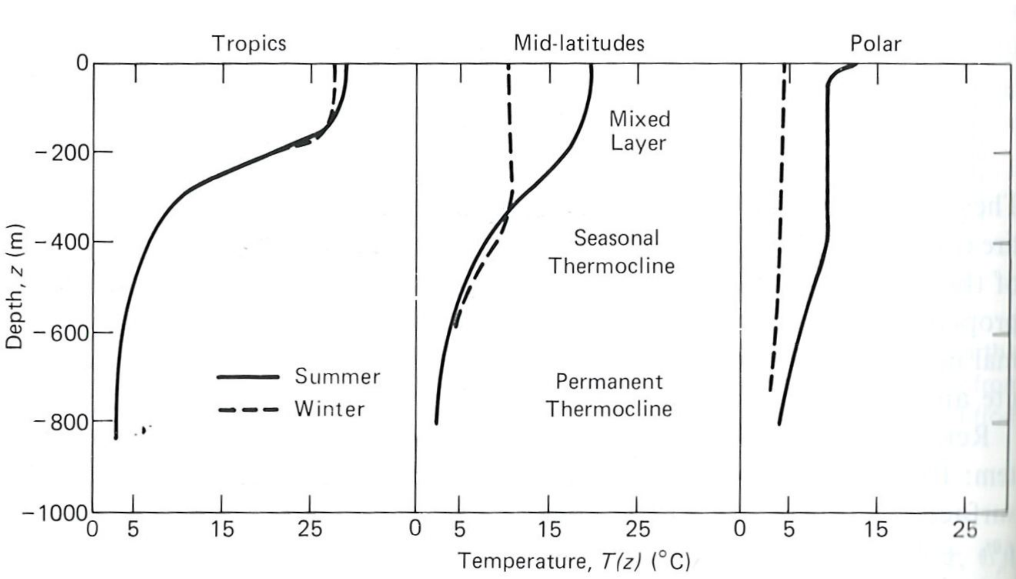


Figure 8: *typical vertical distribution of temperature within the upper km of the water column. Figure taken from Apel (1989).*

Potential temperature

It is not strictly true that water that is subducted into the deep ocean keeps its temperature as soon as it has left the surface: As we have seen in Chapter 2.4, an increase of pressure leads to a slight compression of seawater, i.e. to an increase in its density. This compression increases the internal energy of the water and thereby also its temperature. Therefore, oceanographers often study a quantity called *potential temperature*, which is the temperature a certain water parcel would have, if it was brought from its actual pressure to the surface of the ocean, preventing any heat exchange with its surroundings. The potential temperature is often denoted by the greek Θ to distinguish it from in-situ temperature T. Θ is not measured directly, but has to be calculated from the measured actual temperature and the pressure. But it has the advantage that potential temperature is not affected by pressure. Potential temperature is always a bit lower than actual temperature, but the difference is pretty small (Figure 10).

3.3 Distribution of salinity

The distribution of salinity at the sea surface (Figure 11) is also predominantly constant along latitude lines, though less so than for temperature. Here, however, the maximum values are not reached at the equator, but rather in the subtropics. The Atlantic is generally saltier than the Pacific, and there are noticeable spots of reduced salinity where large rivers such as the Amazon enter the ocean.

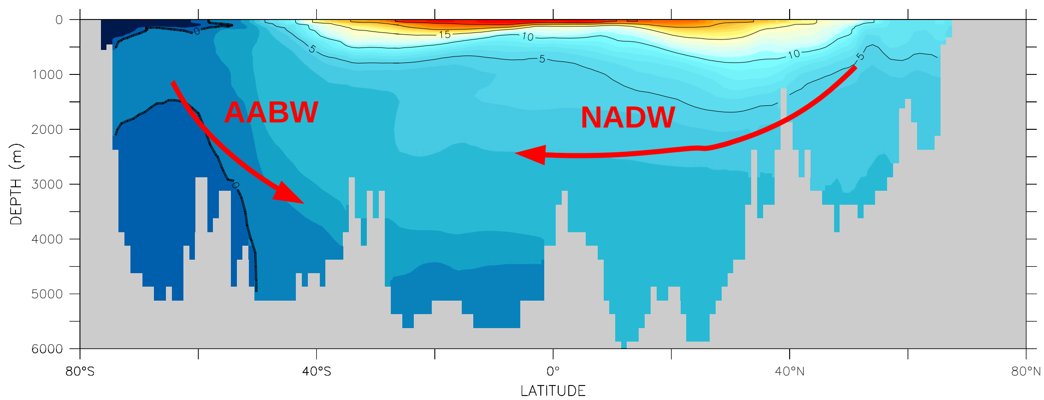


Figure 9: *Temperature along 30° W in the Atlantic*

The latitudinal distribution of salinity is explained by the distribution of evaporation and precipitation of water at the sea surface. Evaporation is highest in the subtropics, while there are bands of strong precipitation at the intertropical convergence zone near the equator and at higher latitudes. This pattern is related to the large atmospheric circulation cells (see Figure 24) with air rising in the intertropical convergence and sinking at the horse latitudes (around 30 N and 30 S).

3.4 Mixing and water mass analysis

The main process that changes salinity and (potential) temperature below the sea surface is the mixing between different water masses. Although this is a slow process, the propagation of water within the ocean is also slow so that mixing effects can be noticeable.

It is intuitively clear that mixing of two waters with different salinities will produce a water with an intermediate salinity. The exact salinity of the mixture follows from the conservation of mass both for the salt and for the total amount of solution. Assume that we mix a kg of seawater of salinity S_1 and b kg of seawater of salinity S_2. The total amounts of salt within the two water masses are $a \cdot S_1$ and $b \cdot S_2$, together therefore $a \cdot S_1 + b \cdot S_2$. The total amount of seawater after the mixing is $a + b$. The salinity of the mixture S_m is the total amount of salt, divided by the total amount of seawater, i.e. $S_m = (a \cdot S_1 + b \cdot S_2)/(a + b) = a/(a + b) \cdot S_1 + b/(a + b) \cdot S_2$. This simply means that the salinity of a mixture is equal to an average of the salinities before the mixing, weighted by their relative contributions. If we for example mix 30% of water with salinity 20 with 70% of water with salinity 30 the mixture has a salinity of $0.3 \cdot 20 + 0.7 \cdot 30 = 27$.

The effect of mixing on temperature is a priori less clear, because there is no such thing as a heat substance. Heat is one of many forms of energy that can be converted into each other.

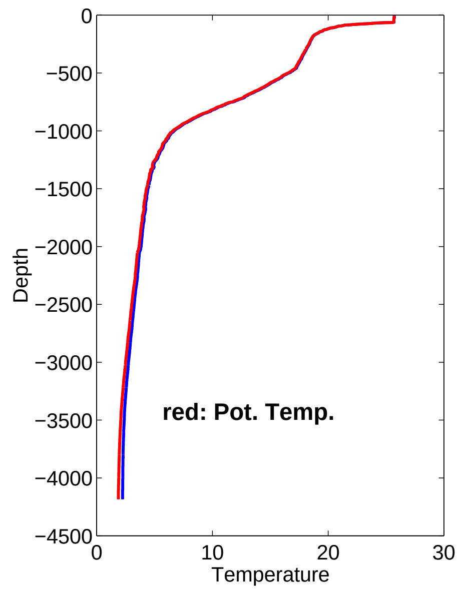


Figure 10: *Comparison of temperature and potential temperature for a location close to Bermuda in the subtropical Atlantic.*

However, in the ocean, the conversion of mechanical energy into heat contributes only very little to the heat balance; only the heating due to compression at high pressures contributes a little. Because of this, potential temperature behaves to a very good approximation as a substance, i.e. the effect of mixing is the same as for salt.

If we therefore denote two water masses by two points in a diagram of potential temperature Θ vs. salinity S , the results of mixing between the two have to lie of the straight line between the two points. If we mix three different water masses, all mixing results in a $\Theta - S$-diagram are contained in the triangle with the three original water masses as corners. This is sketched in Figure 13.

EXERCISE IV: Going back to Figure 4, what happens with density when we mix two water masses that have the same density (i.e. they lie on the same line of constant density in Figure 4) but different temperature and salinity?

Given a number of measurements of Θ and S, e.g. from a vertical profile, oceanographers therefore often plot the two against each other, and try to make sense of the resulting plot as the result of the mixing between a few distinct water masses.

This procedure is illustrated in Figure 14 that shows values of potential temperature vs. salinity obtained by sampling water in different depths at a location in the South Atlantic. The depth the measurements were taken (in units of 100 m) is shown as the small numbers beside the data points. Also shown as shaded rectangles are the properties of three source water masses that fill most of the deep South Atlantic, Antarctic Bottom Water (AABW),

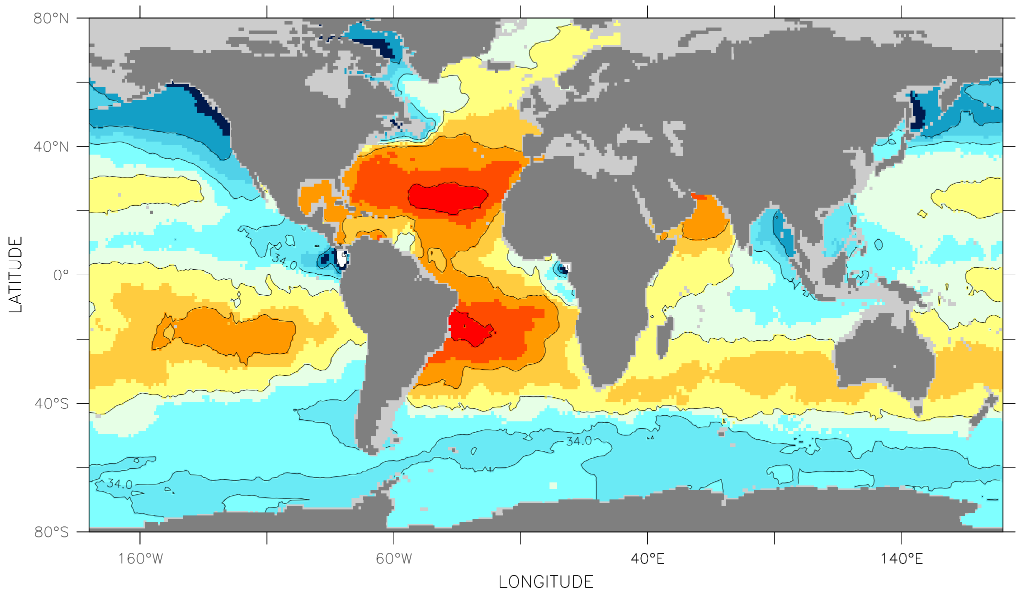


Figure 11: *Average distribution of sea surface salinity (SSS), also from the COADS data set.*

North Atlantic Deep Water (NADW), and Antarctic Intermediate Water (AAIW). From the plot it becomes obvious that the observed water mass properties below 1800 m depth are explained by a mixing between AABW and NADW, while between 1800 m and 400 m depth, three source water masses contribute, namely NADW, AAIW and a third shallower water mass that is probably Subantarctic Mode Water SAMW with properties close to the observed values at 400 m depth.

4 Energy and freshwater balances

4.1 The earth as a greenhouse

The radiation of the sun provides the energy that drives climate on earth. Because the solar surface temperature is about 5900 K, the received energy is mainly in the form of short-wave visible light with wavelengths between 0.35 and 0.7 μm (Figure 15). In this wavelength range the atmosphere is almost transparent, i.e. only very little of this energy is absorbed in the atmosphere. One part of the solar radiation that is absorbed strongly in the atmosphere is the ultraviolet radiation below about 0.3 μm wavelength.

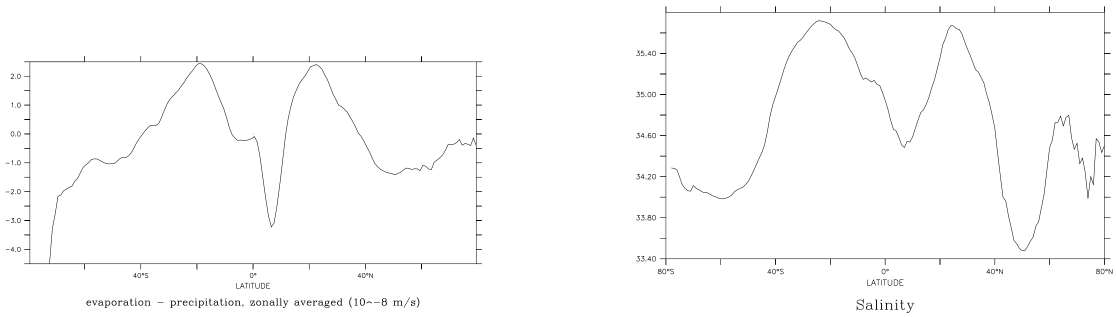


Figure 12: *Left panel: average distribution of evaporation minus precipitation at the sea surface with latitude. Right panel: average distribution of sea surface salinity with latitude.*

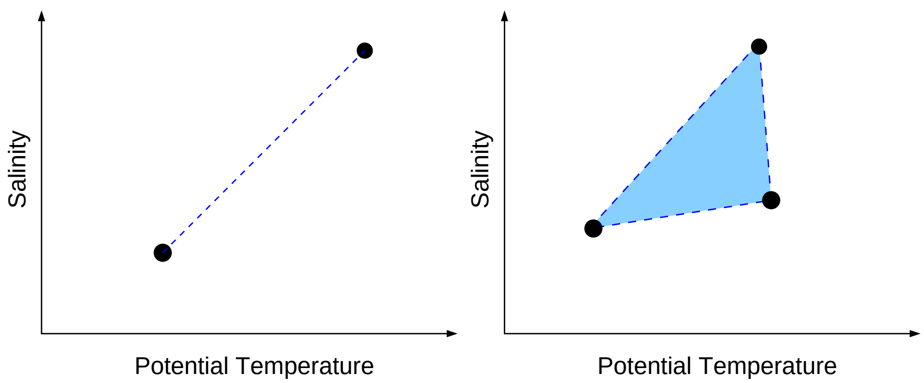


Figure 13: *Temperature-salinity-diagram showing the effect of mixing between two (left) or three (right) water masses.*

Wiens displacement law and the Stefan-Boltzmann law

Two laws important for an understanding of the energy balance of the earth are Wiens displacement law and the Stefan-Boltzmann law. Both are strictly valid only for a black body, i.e. a body that can emit and absorb light at all wavelengths equally. This is an idealisation, but the two laws still describe solar and terrestrial radiation reasonably well. Wiens law states that the wavelength λ_m at which a body emits most of its energy varies inversely with its temperature T (in K), or

$$\lambda_m = \frac{\alpha_r}{T}$$

where $\alpha_r = 2897.8\ \mu\mathrm{m}\ \mathrm{K}^{-1}$ is a constant. Qualitatively this means that the hotter a body, the shorter is the wavelength of the light it emits.
The Stefan-Boltzmann law (n.b.: the two names belong to different persons) states that the total amount of energy Q_b that a (black) body radiates from its surface varies with the fourth power of its temperature, or

$$Q_b = \sigma T^4,$$

with $\sigma = 5.7 \cdot 10^{-8}\mathrm{W\,m^{-2}\,K^{-4}}$

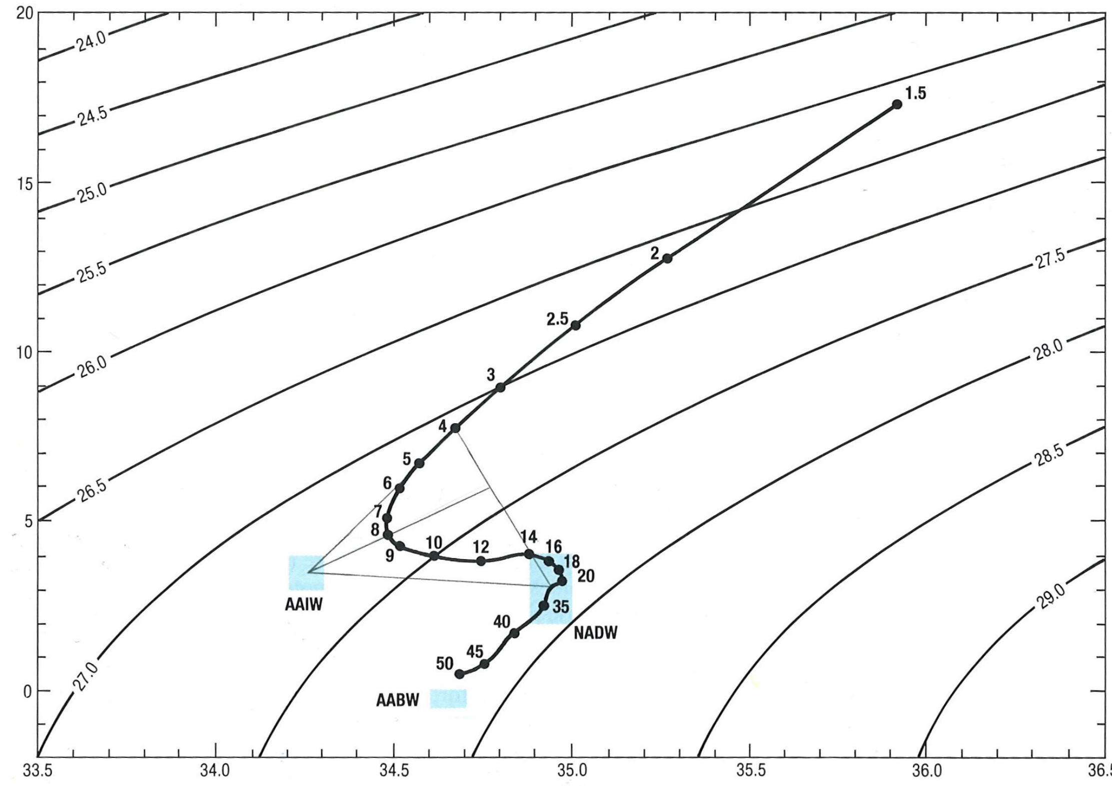


Figure 14: *Temperature-salinity-diagram showing data taken of a vertical bottle cast in the subtropical South Atlantic. Each point corresponds to a measurement at a specific depth. The depth is indicated by the numbers adjacent to the points. Figure taken from Open University (1989).*

The radiative flux of energy from the sun decreases with the square of the distance from the sun. The average flux at the mean radius of the earth's orbit is called the *solar constant* S and has the value

$$S = 1368\,\mathrm{W\,m^{-2}}.$$

Note that the name solar constant is a bit misleading, because the actual radiative flux varies by a few percent around that value, mainly because of the slight variation of the distance between earth and sun over the year.

S is the energy flux per unit area oriented perpendicular to the direction towards the sun. The earth absorbs this radiation over an area of a circle with the radius r_E of the earth, πr_E^2. This radiation is distributed (unevenly) over the surface area of the earth, which is approximately (assuming the earth to be a perfect sphere) $4\pi r_E^2$. On average, therefore, the energy arriving per unit area at the earths surface is

$$\frac{S}{4} = 344\,\mathrm{W\,m^{-2}},$$

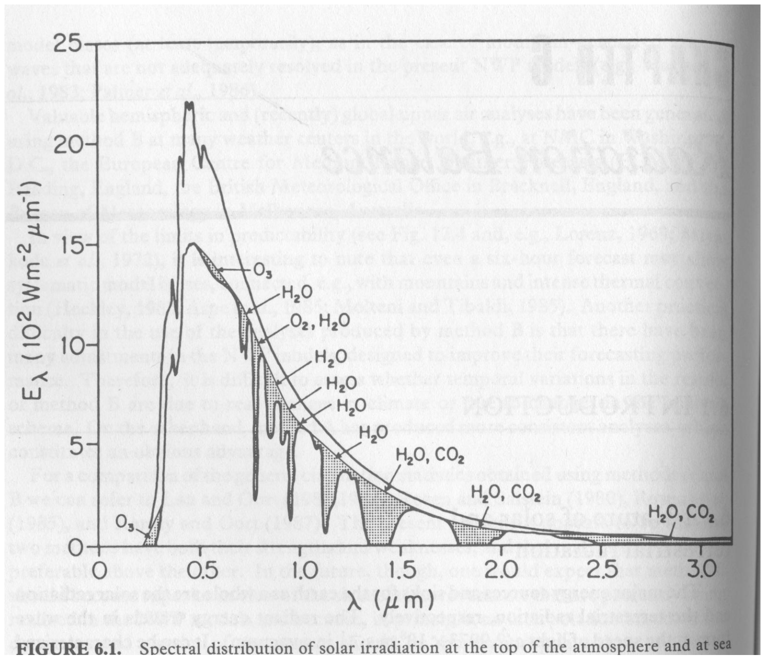


Figure 15: *Wavelength distribution of the solar radiation at the top of the atmosphere and at sea surface. The shaded areas indicate the wavelengths where different parts of the atmosphere absorb some of the incoming radiation. From Peixoto and Oort (1991)*

(one quarter is the ratio between πr_E^2 and $4\pi r_E^2$).

Some of the incoming radiation is reflected back into space. The relative amount of reflected radiation is called *albedo* α. The albedo of clouds and of snow cover is relatively high (up to 0.9), while that of the ocean surface is relatively low with values around 0.05. On average, the albedo of the earth is around 0.3, but that albedo depends on the amount of cloud cover and the extent of deserts and ice shields.

Because of the albedo, the total energy from solar radiation that is available to heat the earth is

$$Q_s = (1 - \alpha)\,\frac{S}{4}$$

This energy gain Q_s must be compensated by an equal energy loss, otherwise the earth would continually become warmer. The only way that the earth can loose energy into space is by thermal radiation Q_b. At the temperature of the earth, most of the energy in the emitted thermal radiation is contained in the infrared or long-wave domain. Q_b is therefore often called long-wave radiation, while Q_s is called short-wave radiation.

The thermal radiation of the earth depends on its temperature by the Stefan-Boltzmann law (with some small correction because earth is not a perfect black body). Let us first assume that the temperature of the earth is a constant T_0. We can then calculate it from the condition of equilibrium between the incoming and outgoing radiation $Q_s = Q_b$:

$$(1-\alpha)\,\frac{S}{4} = \sigma T_0^4 \quad \rightarrow \quad T_0 = \left(\frac{(1-\alpha)S}{4\sigma}\right)^{1/4}$$

Without atmosphere the average surface temperature of the earth would be close to this so-called radiative equilibrium temperature, which is 255 K, or -18 °C. The actual average temperature at the surface of the earth, however, is about 288 K, a bit more than 30 K higher. The main reason for this is the *natural greenhouse effect* caused by the atmosphere.

In contrast to visible light, infrared radiation is strongly absorbed in the atmosphere, mostly by the water vapour in humid air, but also by carbon dioxide and a few other gases. Due to this adsorption, the atmosphere itself is heated and emits long-wave radiation. The emission by the atmosphere is partly goes into space, but a part of it acts as an additional source of energy to the earths surface, heating it above the radiative equilibrium temperature.

A simplified picture of how this greenhouse effect works can be obtained by assuming that the atmosphere acts as one thin layer around the earth that leaves through all solar radiation Q_s, but absorbs completely the infrared radiation Q_b from the earths surface (Figure 16). As this 'atmosphere' has to be in energetic balance, too, it must also emit energy in the form of radiation. The heat emission is equal in all directions, i.e. there is an equal net flux of energy Q_a downwards as upwards.

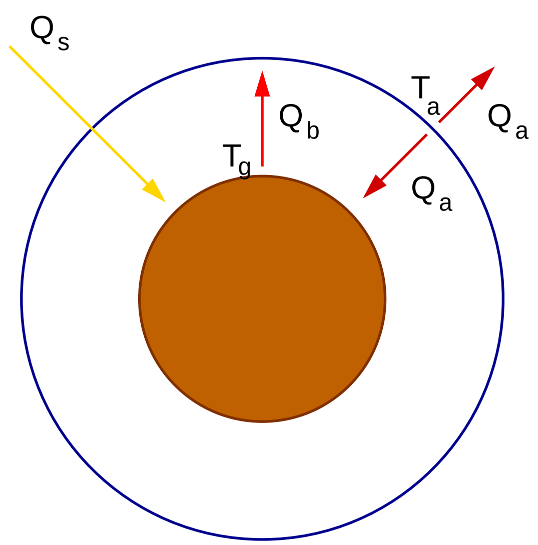


Figure 16: *A simplified model for the greenhouse effect.*

In this case now the heat balance for the earth is $Q_s + Q_a = Q_b$, i.e. there is an additional source of energy Q_a from the thermal radiation coming from the 'atmosphere'. On the other hand we have the energy balance for the 'atmosphere' $Q_b = 2Q_a$, i.e. the absorbed thermal radiation from the earth Q_b is converted into one part of the energy Q_a being lost into space and another part Q_a being radiated downwards and heating the earth. From this we can infer $Q_a = Q_b/2$ and insert this in the earths balance to obtain $Q_s + Q_b/2 = Q_b$ or $Q_s = Q_b/2$.

As before, we can now use the Stefan-Boltzmann law to infer the equilibrium temperature of the earth with this greenhouse atmosphere T_g:

$$(1-\alpha)\,\frac{S}{4} = \sigma T_g^4/2 \quad \rightarrow \quad T_g = \left(\frac{2(1-\alpha)S}{4\sigma}\right)^{1/4}$$

The final result is that the 'greenhouse temperature' T_g in our simple model is a factor of $2^{1/4}$ larger than the radiative equilibrium temperature T_0, or 303 K instead of 255 K.

Exercise V: Can you infer what must be the temperature of the 'atmosphere' in this case? The solution can be found without calculation by analogy to the earth without atmosphere.

The emission of carbon dioxide and other *greenhouse gases*, such as methane or nitrous oxide by human activities leads to additional absorption of radiation in the atmosphere and therefore to further warming.

Exercise VI: The temperature of the earth depends also on how much energy is reflected at the earths surface, i.e. the albedo. On the other hand the albedo depends on the relative amount of the earth that is covered by ice and snow, deserts, clouds, etc. Assume that the earth would cool by some amount, so that the area covered by snow increases. What would happen to the average temperature? And how would that affect the ice cover?

4.2 Heat balance of the ocean and atmosphere

Our view of the greenhouse effect has been very simple so far; a somewhat more detailed view of the fate of incoming and outgoing radiation is given in Figure 17.

The figure shows that indeed the largest part of the solar radiation reaches the earths surface, with about 30% being reflected back into space and the rest absorbed in clouds and air, corresponding to the mentioned average albedo of 0.3.

The part of radiation that reaches the ground (about 70%) leads to a temperature increase until the established temperature drives an equal energy loss. This heat loss proceeds via three main pathways:

- *radiative heat flux:* This is the loss of heat by emission of long-wave electromagnetic radiation that we already discussed in principle in the last subsection. The largest part of the long-wave radiation emitted by oceans and land is absorbed in the atmosphere, but some passes through it.

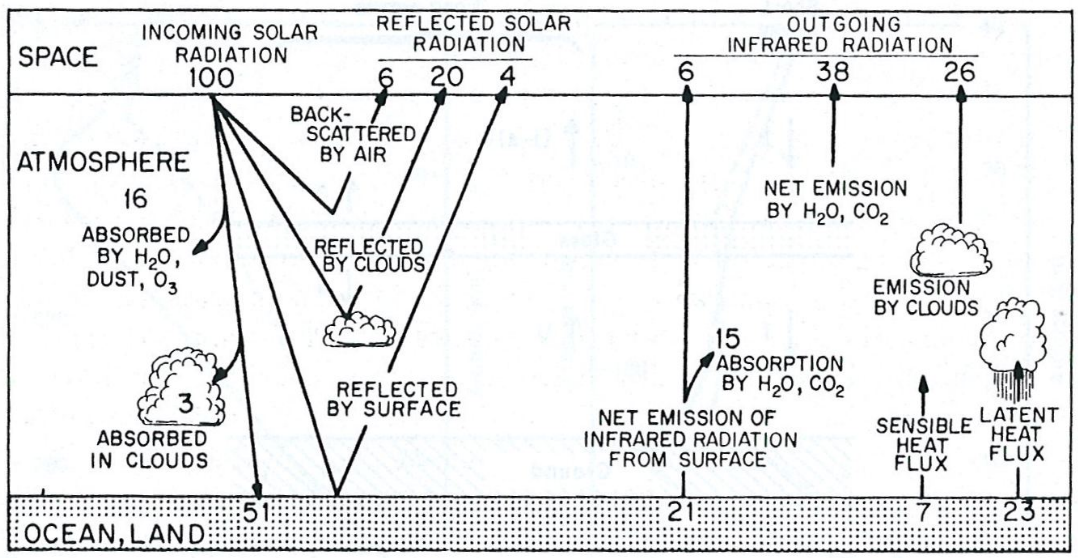


Figure 17: *A more detailed view of the energy balance of the earth and atmosphere. Energy fluxes are given as percentage of the total incoming solar radiation. Figure taken from Gill.*

- *sensible heat flux:* This is the transfer of energy by heat conduction. It requires direct contact between a colder and a warmer medium and operates via interactions between the random movements of molecules (Brownian motion) that are more vigorous the warmer a body is.
- *latent heat flux:* Water has a high latent heat of vaporisation, i.e. it takes a lot of energy (about $2.4 \cdot 10^6$ J kg^{-1}) to bring water from the liquid into the gas phase. This heat is released again when water vapour forms liquid droplets e.g. in clouds. Because the heat for evaporation is mostly taken from the water when evaporation occurs, and is released in the atmosphere when it precipitates again, this amounts to a net transfer of energy into the atmosphere.

Only a small part of the heat loss is directly lost into space, the larger part heats the atmosphere, which in turn radiates into space.

Exercise VII: Why is there sensible heat flux or latent heat flux between ocean and atmosphere, but no sensible or latent heat flux from the earth into space?

4.3 Latitudinal variation and energy transport

The amount of energy received from solar radiation per square meter varies with latitude. This is simply due to the fact that the larger the inclination of the earths surface is with respect to the direction of the sun, the larger the area that an equal amount of radiation is distributed over. This is illustrated in Figure 18.

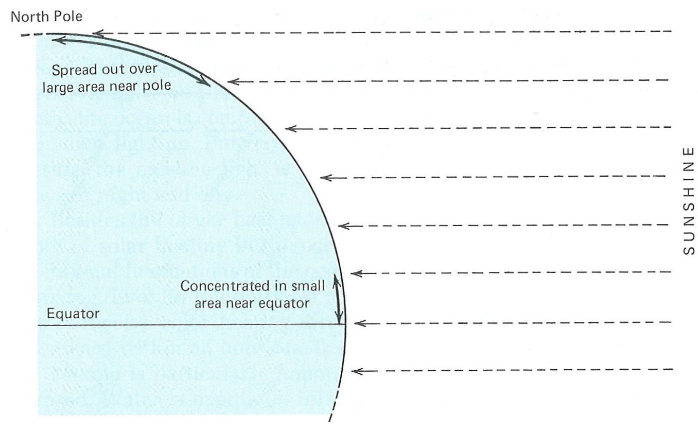


Figure 18: *The radiation energy from the sun is distributed over a larger surface area at higher latitudes than at lower latitudes, leading to a smaller received energy per square meter. Figure taken from Stowe (1981).*

Figure 18 shows the latitudinal distribution of solar radiation per square meter at the top of the atmosphere (i.e. without the effect of the distribution of albedo) and at the earths surface. The irradiance at the top of the atmosphere has a maximum at the equator and decreases uniformly with latitude. However, it does not decrease to zero at the poles.

EXERCISE VIII: Can you imagine why the solar radiation per square meter does not decrease to zero at the poles?

The radiation per square meter at the earths surface is less than at the top of the atmosphere and its latitudinal dependency shows the regionally different influence of the earths albedo. This is evident e.g. in the small minimum of irradiance somewhat north of the equator that is caused by the relatively high cloud cover connected to rising air at the so-called intertropical convergence. This high cloud cover results in a higher albedo and thus in less radiation received at the surface.

EXERCISE IX: Can you estimate the albedo at the North pole from the Figure 19? Does it differ from the albedo at the equator?

Figure 19 also shows the latitudinal distribution of the emission of heat radiation at the surface. The emission is is smaller than the irradiance in low latitudes, i.e. there is a net gain of energy. Conversely, the long-wave radiation is larger than the irradiance in high latitudes, indicating a net energy loss. So, while the earth as a whole is in radiative balance, regionally there can be an imbalance. This imbalance requires that the excess energy received at low

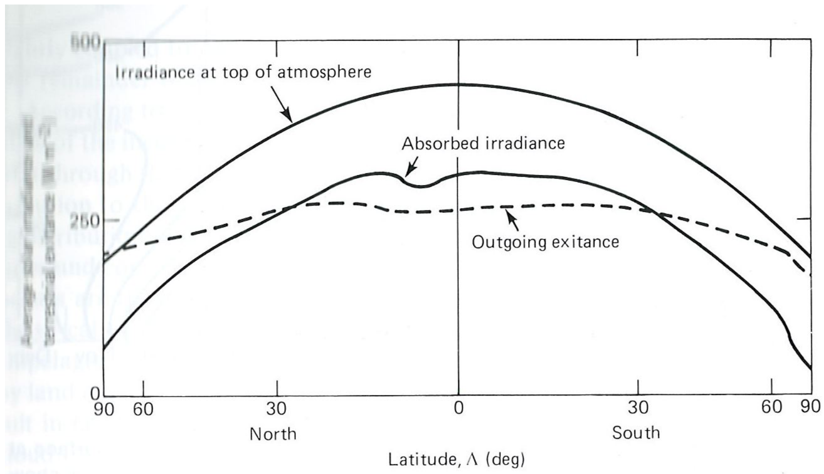


Figure 19: *Distribution of solar irradiance (solid lines) and of terrestrial long-wave radiation (dashed line) with latitude. Upper solid line: incoming solar radiation at the top of the atmosphere. Lower solid line: net solar radiation (i.e. after subtracting the reflected part) at the surface of the earth. Figure taken from Apel (1989).*

latitudes has somehow to be transported to lower latitudes to fill the energy deficit there. Otherwise the imbalance would result in perpetual warming at the equator and cooling at the poles.

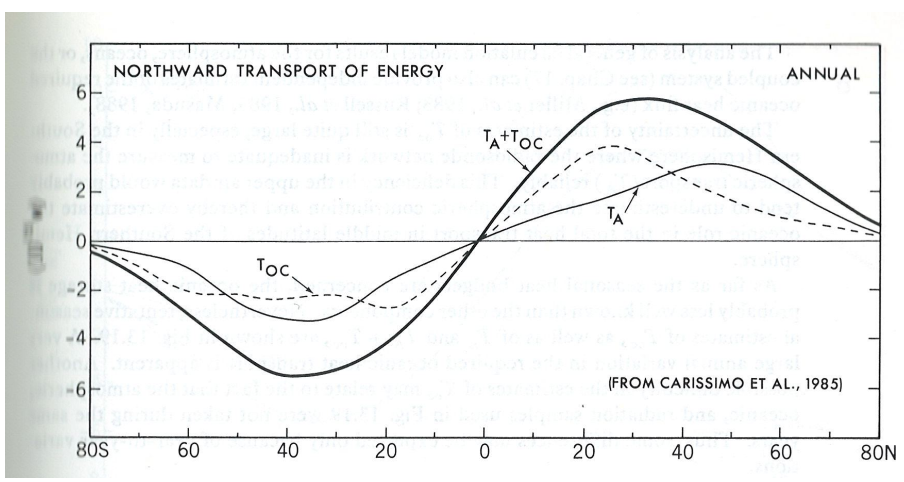


Figure 20: *Northward energy transport as a function of latitude (negative values in the southern hemisphere indicate southward transport). Figure taken from Peixoto and Oort (1991).*

This energy transport is shown in Figure 20. It is composed of an atmospheric transport (thin solid line) and an oceanic transport. The oceanic transport is driven by warmer water flowing (on average at least) predominantly poleward at the sea surface, while colder water at depth flows (on average) equatorward.

Exercise X: What happens in terms of energy when humid air formed by evaporation in low latitudes is transported poleward and the water precipitates there as rain?

4.4 The hydrologic cycle

The cycling of freshwater through the atmosphere, precipitation, river runoff etc. are together called the hydrological cycle. The hydrologic cycle is intimately coupled to the energy balance and the redistribution of heat, because evaporation and precipitation result in large amounts of heat being transferred. However, we will give only a global description of the hydrologic cycle here and not talk about its geographical distribution.

Reservoir	Percent of total
Oceans	97.96
Polar ice caps and glaciers	1.64
Ground water	0.36
Rivers and lakes	0.04
Atmosphere	0.001

Table 2: *Global distribution of water on earth over the different reservoirs at the present climate.*

The water in the climate system is distributed very unevenly over the ocean, the atmosphere, over rivers and lakes, groundwater, and ice shields (Table 2). By far the largest part is contained in the ocean, almost all the rest in the ice caps and glaciers. The fraction of water bound in ice caps depends strongly on the temperature of the earth; during the coldest stage of last ice age, the average sea level was about 120 m deeper than today, the water being bound in large ice shields. Rivers and lakes contain less that a thousandth of the total water on earth, and the atmosphere only a small fraction of that.

Although the atmosphere contains only a trace amount of the total water on earth, it acts as an important pathway for transferring water from one reservoir to another. This is because the residence time of water in the atmosphere is quite small; on average a water molecule that is evaporated stays only about 10 days in the atmosphere before it precipitates again.

Most of the water that evaporates precipitates over the ocean; only less than a third precipitates over land. Precipitation over land exceeds evaporation, the difference ultimately reaching the oceans as river and groundwater runoff. There are differences in the balance of evaporation and precipitation between the different ocean basins. There is e.g. net evaporation over large parts of the Atlantic, with part of that water precipitation over the Pacific, accounting for the lower salinity in the Pacific, compared to the Atlantic.

Exercise XI: From the numbers given in Figure 21, can you estimate a residence time for water in the ocean?

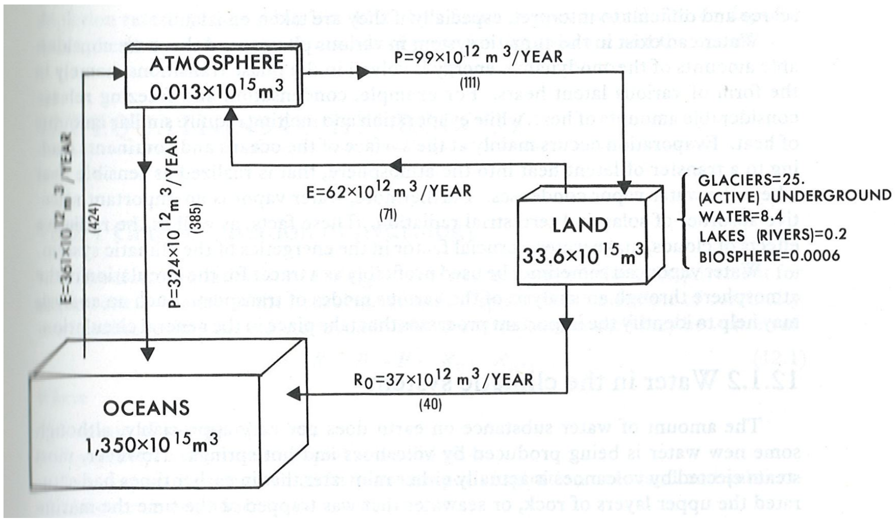


Figure 21: *Schematic diagram showing the different reservoirs and fluxes involved in the hydrologic cycle. E is evaporation, P precipitation, and R_0 river and groundwater runoff. Figure taken from Peixoto and Oort (1991).*

4.5 Marginal seas

Marginal seas are almost isolated ocean basins that exchange with the adjacent open ocean only through a strait with a sill, i.e. a ridge that blocks exchange of deeper water. Examples of marginal seas are the Baltic, the Mediterranean, the Black Sea or the Gulf of Aqaba.

The circulation within marginal seas and the exchange with the adjacent ocean are strongly influenced by whether the evaporation of water within the basin exceeds the precipitation and river runoff, or vice versa. This also has important consequences for marine life in these basins.

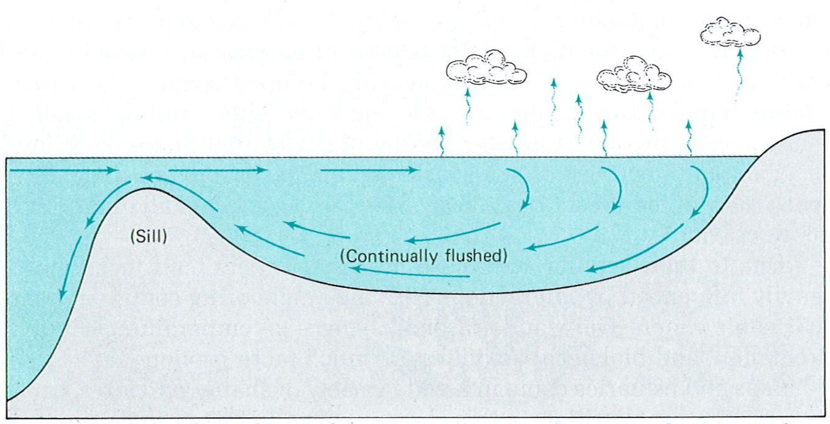


FIGURE 16.1 *Where there is large evaporation, surface water becomes saline and sinks, continually flushing out the bottom of the sea.*

Figure 22: *An evaporation basin. Figure taken from Stowe (1983).*

In marginal seas like the Mediterranean, where evaporation greatly exceeds precipitation, the net loss of water to the atmosphere leads to a gradual increase of salinity in the remaining surface water. When the surface water becomes denser than the underlying water (remember, density increases with increasing salinity), it sinks. When this dense water mass fills the deep part of the basin up to the sill height, some of that dense water flows out into the adjacent ocean (Figure 22), and is replaced by lighter, less saline water flowing in at the surface. The circulation over the sill therefore consists of an inflow at the surface and an outflow at depth.

EXERCISE XII: Are the amounts of water flowing into and out of the basin over the sill exactly equal?

Because the deep water is continually replenished with sinking of surface water, evaporation basins are usually well ventilated, i.e. the deep water contains enough oxygen to allow life of oxygen-dependent life, such as fishes.

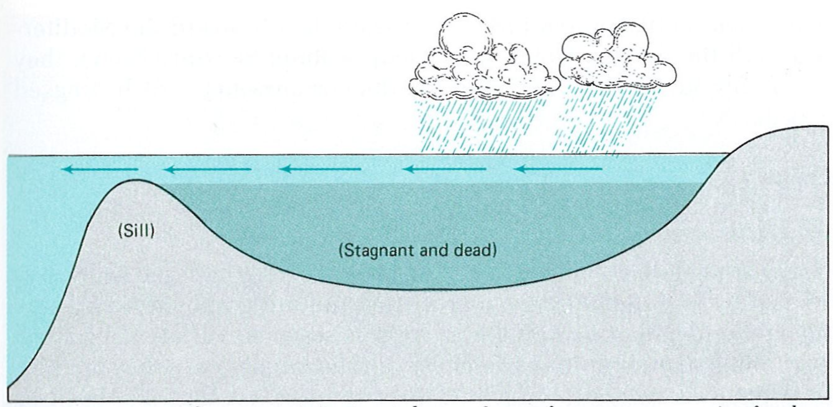


FIGURE 16.3 *Where precipitation is large, the surface water remains fresher and lighter than deeper water, and there is little vertical mixing. Once the oxygen is depleted from the deeper water, it is not replenished from the surface, and so it is lifeless.*

Figure 23: *A precipitation basin. Figure taken from Stowe (1983).*

The other extreme are marginal seas like the Baltic or the Black Sea, where precipitation and river runoff exceed evaporation. Because fresh water is relatively light it tends to accumulate at the surface, increasing the vertical density gradient and therefore increasing the vertical static stability of the water colum. The deep water in these basins is therefore only weakly ventilated and its oxygen content can get used up by bacteria that remineralise the organic matter that sinks out of the surface layer. In the Black Sea and in parts of the Baltic basically all oxygen can get used up at depth, leading to widespread anoxia and sulfate reduction.

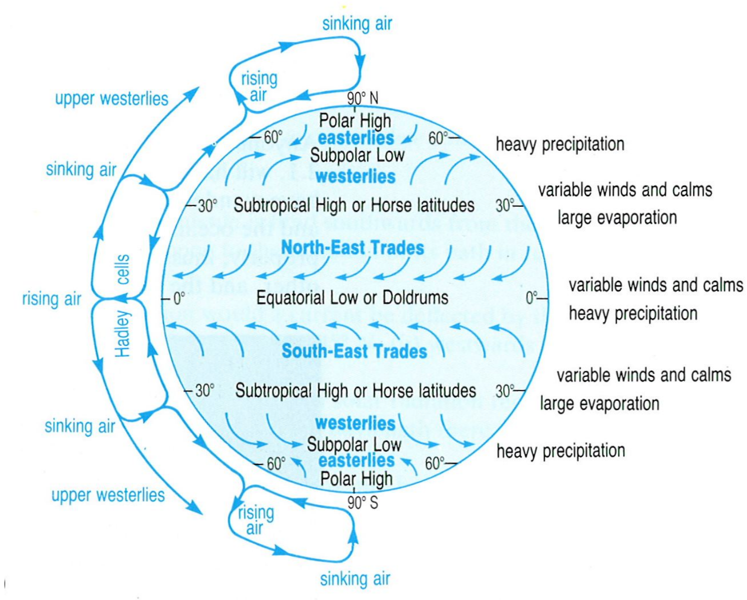


Figure 24: *Global wind patterns. Figure taken from Open University (1989).*

5 Currents and circulation in the sea

This chapter is meant as a very short overview over the large-scale circulation patters in the ocean. The large-scale circulation is usually separated into a 'wind-driven circulation' near the surface and a generally much slower 'thermohaline circulation' that ventilates the deep oceans. Although this distinction is to some degree artificial, it helps in obtaining a qualitative picture.

5.1 The wind forcing

Unlike in the ocean, the circulation in the atmosphere is relatively unconstrained in zonal (east-west) direction. The large mountain ranges deflect the winds somewhat, but to a first approximation they are independent of longitude. There is a clear pattern of westward blowing (easterly) trade winds on both sides of the equator, eastward blowing (westerly) winds in somewhat higher latitudes, and finally a belt of easterly winds encircling the polar regions.

This pattern of surface winds is related to the vertical structure of the atmospheric circulation that is characterised by three more or less closed cells of alternating rising and sinking air (Figure 24). The ultimate reason for these cells is the excess solar heating in the tropics and cooling at high latitudes. The tropical heating leads to rising of air in the 'intertropical convergence zone' near the equator, accompanied by heavy precipitation as the rising air cools.

If the earth would not rotate around its axis, the air rising near the equator would simply move pole-wards in the upper atmosphere, sink in polar latitudes, and then move equatorwards in the lower atmosphere, i.e. there would be only one vertical circulation cell in the atmosphere in every hemisphere.

However, the poleward or equatorward movements that are required in this circulation are deflected by the Coriolis force, discussed in the next section (to the right in the northern hemisphere and to the left in the southern hemisphere). This leads to the observed pattern of three rather than one circulation cells in every hemisphere and to the westward deflection of the equatorward flow in the trade winds.

5.2 The wind-driven currents

In the ocean, there is just one major current system that encircles the earth roughly along latitude lines, the 'Antarctic Circumpolar Current'. Other than that, the major surface current systems are organised in more or less closed circulation cells or gyres that are constrained to just one oceanic basin. Most prominent are the *subtropical gyres* that spin clockwise in the northern hemisphere and counterclockwise in the southern hemisphere. On the western flank of these gyres the currents are usually relatively narrow and swift and carry large amounts of heat poleward. The most prominent examples for these currents are the Florida current / Gulf stream system in the North Atlantic and the Kuroshio in the North Pacific. The equatorward currents on the eastern side of the gyres are often somewhat wider and weaker.

Poleward of these subtropical gyres we can find a *subpolar gyre* both in the northern North Atlantic and North Pacific. They rotate in the opposite direction as the subtropical gyres. Again the currents in the gyres are strongest on the western flank, where the Labrador current and the Oyashio bring cold water far southward.

The surface circulation is often also called the 'wind-driven circulation'. It is important to note, however, that the direct influence of the friction exerted by wind extends just over the uppermost few tens of meters (the so-called Ekman layer, see section 7.4), much shallower than the surface currents. The forcing of the surface current systems by the wind operates somewhat indirectly; this will be dealt with in section 7.6.

5.3 The thermohaline or overturning circulation

We have seen in section 3.2 that the deep oceans are filled with cold waters that have their origin at high latitudes, namely the *North Atlantic Deep Water* (NADW), overlying the *Antarctic Bottom Water* (AABW). These water masses are formed at high latitudes by loss of heat to the atmosphere, and by the increase of salinity connected with the formation of sea-ice, leading to comparatively high densities.

The circulation that leads to a continuous renewal of these waters in the deep ocean by the formation of NADW and AABW in high latitudes, and by upwelling of these water masses elsewhere is often called 'thermohaline circulation', although the term 'meridional overturning circulation' is also used. The name 'thermohaline circulation' implies that it is the increase

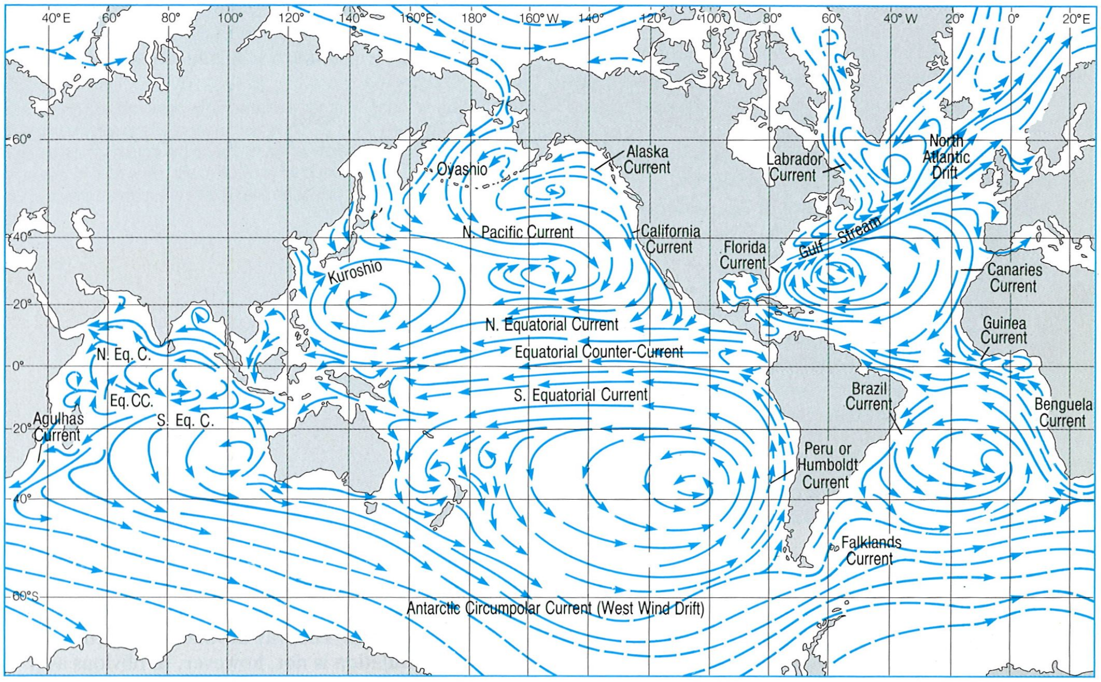


Figure 25: *Schematic picture of the global surface currents in the ocean. Figure taken from Open University (1989).*

of density connected with fluxes of heat and salt that actually *drives* the overturning, which is not necessarily true.

Here we will limit ourselves to a description of the overturning circulation as it is presently understood, without going into the discussion of what is the actual driving force (see e.g. Wunsch (2002), Science 298, pp. 1179-1180 for a discussion).

The overturning circulation has been likened to a conveyor-belt that consists of deep and bottom water formation in high latitudes, spreading through the ocean basins, upwelling of these waters, and a closure of the circulation by net transports of surface waters (Figure 26). With this conveyor belt, substances are transported into the deep ocean. Examples for that are the man-made chloro-fluoro-carbons CFC, that were brought into the atmosphere since the fifties, or radiocarbon that is produced within the atmosphere by cosmic rays, enters the ocean and decays within it. These substances can be used to estimate how long a certain water mass has been away from the surface of the oceans and therefore give information how long the overturning circulation takes for a complete renewal of the deep water. The result is that the overturning circulation takes roundabout a thousand years to complete a cycle.

The picture of one overturning circulation cell (or possibly two, connected with the formation of NADW and AABW) is a simplification, though. In reality there is a whole web of interconnected circulation cells in which the formation of intermediate waters (most prominently Antarctic Intermediate Water AAIW) also plays an important role. Figure 27 shows schematically how these cells look like in the different ocean basins and how they are connected.

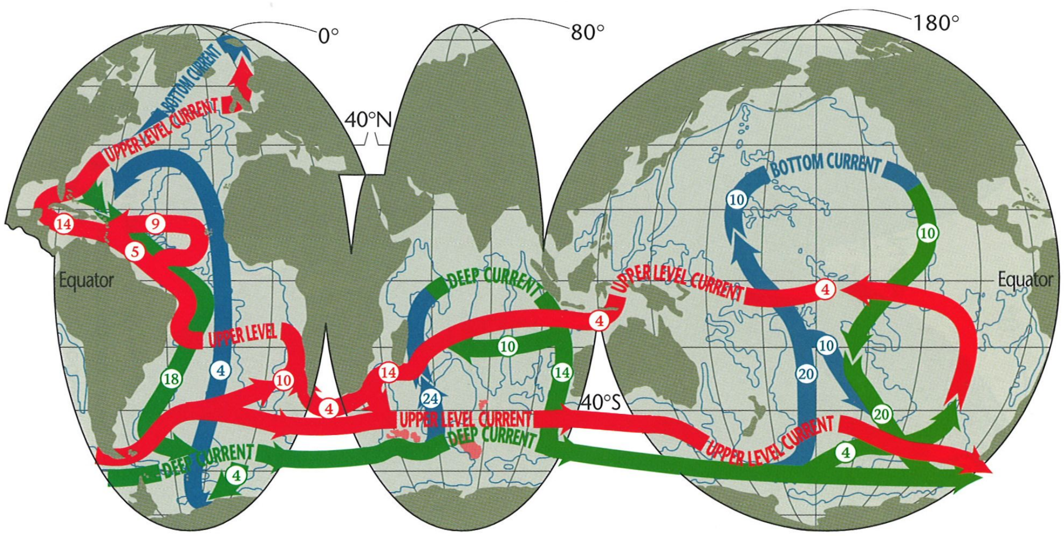


Figure 26: *A schematic diagram showing the conveyor-belt circulation. Figure taken from Schmitz (1996).*

6 Forces in the ocean

Oceanic motions are mainly driven by friction of wind at the surface and/or by density differences, causing pressure gradients. To understand them, we have to understand the forces that occur in the sea and drive currents.

6.1 Newton's second law

Newton's second law states (i) that it requires a force to change the velocity of a body, and (ii) that the force required for a specific change in velocity is proportional to the mass of the body being moved. The mathematical form of Newton's second law is

$$\mathbf{F} = m \cdot \frac{d\mathbf{V}}{dt},$$

the force $\mathbf{F}$ acting on a body is equal to the mass of the body m, times its acceleration, or rate of change of velocity $d\mathbf{V}/dt$.

Vector notation

Newton's second law is a *vector* equation, i.e. the force and the acceleration have a *direction* as well as a *magnitude*. This means that, if several forces act on a body, the resultant force is not simply the sum of the magnitudes of the individual forces. Instead it is the *vector sum*. The vector sum can be visualised by drawing each force as an arrow, where

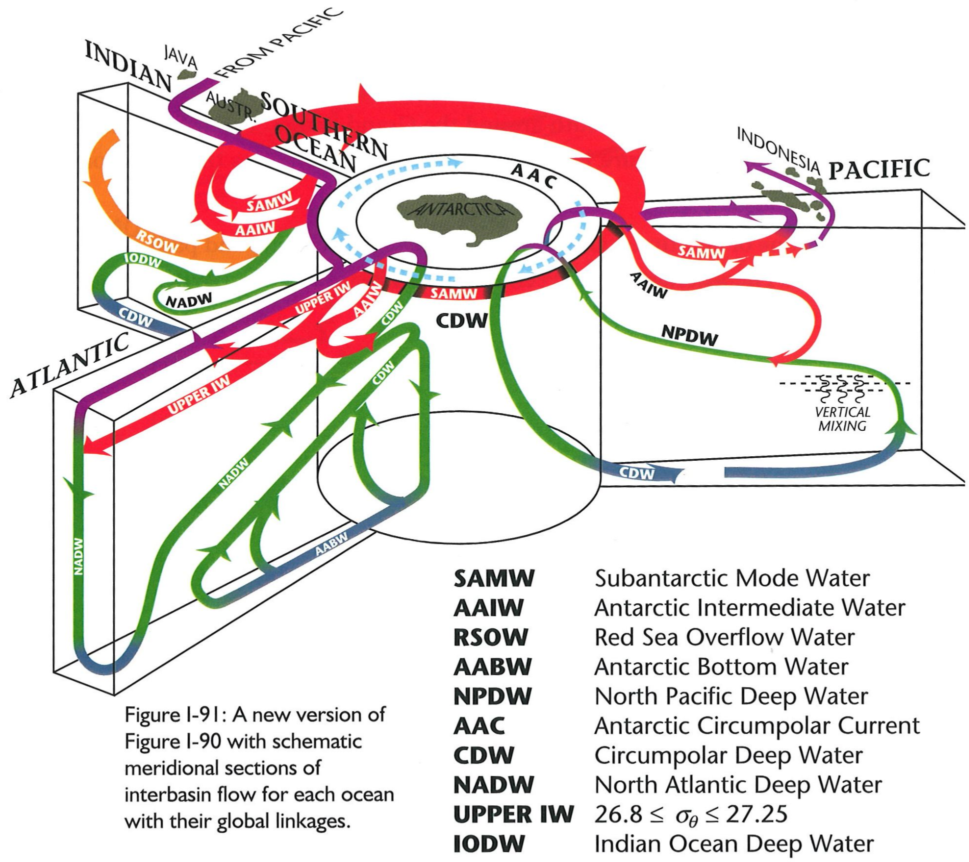


Figure 27: *A more complicated view of the thermohaline or overturning circulation separated into different ocean basins. Figure taken from Schmitz (1996).*

the length of the arrow indicates the strength of the force, and its direction the direction of the force. The sum of the forces is then symbolised by the arrow that is achieved if we let arrow 2 start at the tip of arrow 1, and connect the origin of the first with the tip of the second. The ordering of the two arrows does not matter.
Fortunately we do not have to resort to drawing to add to vectors, but instead use the *coordinate representation* of a vector. Each vector in space can uniquely be defined by its three coordinates with respect to a *orthogonal coordinate system*, which simply means three directions in space x, y and z that are mutually orthogonal.
The coordinates of a vector $\mathbf{a}$, denoted by the indices a_x, a_y and a_z are the magnitude of the orthogonal projection of the vector onto that coordinate system. If we have the coordinate representation of two vectors, we can calculate their sum by simply adding their coordinates.
In this lecture we will always use a coordinate system where x points eastwards, y northwards, and z upwards (Figure 30). In this coordinate system, Newton's law can be written as

$$
\begin{aligned}
F_x &= m\frac{du}{dt} \\
F_y &= m\frac{dv}{dt}
\end{aligned}
$$

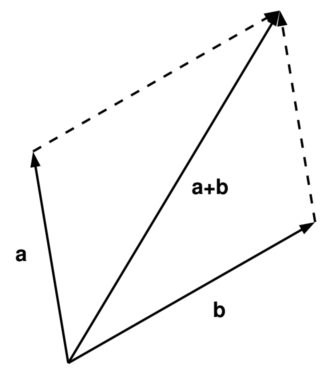


Figure 28: *Illustration of the addition of vectors* **a** *and* **b**.

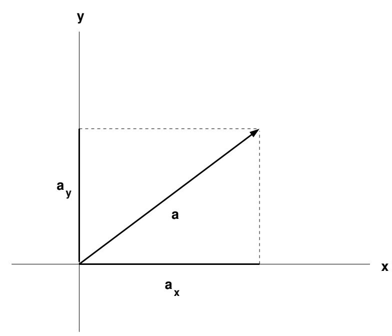


Figure 29: *Illustration of the coordinate representation of a vector* **a**.

$$F_z \quad = \quad m\frac{dw}{dt}$$

where u, v and w are the eastward, northward and upward components of velocity. It is important to keep in mind that the vector form of an equation is universally valid, while the coordinate representation of the same equation depends on the choice of the coordinate system.

The form of Newton's second law stated so far is convenient if we are dealing with distinct masses like planets or cannon balls, that interact by means of forces between them. But in the ocean we are confronted with a continuous medium and the forces acting on individual parcels of fluid. Here is is more convenient to divide Newton's by an unit of volume (a cubic meter e.g.) to obtain

$$\rho \cdot \frac{d\mathbf{V}}{dt} = \mathbf{F_v}$$

where the density ρ is the mass per unit volume, and $\mathbf{F_v}$ is the sum of all forces acting per unit volume.

6.2 Gravity and centrifugal force

Gravity is under most circumstances the strongest force acting on the ocean, and it directly or indirectly influences a whole range of phenomena in the sea.

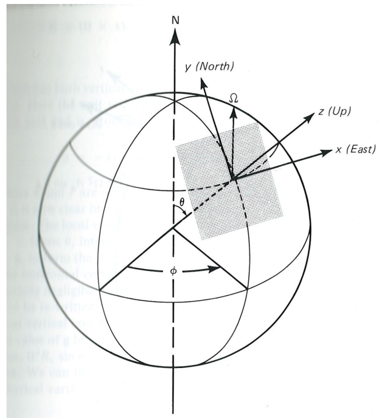


Figure 30: *Coordinate system used in this lecture.*

Every body on the surface of the earth is drawn toward the centre of the earth by the gravitational attraction between the mass of the body m and the mass of the earth M . The attractive force is proportional to the product of the two masses and inversely proportional to the square of the distance between the Centrex of mass of the two bodies.

If the earth was perfectly spherical with a homogeneously distributed mass, the gravitational force per unit volume F_{gz} (the subscript z indicating that it is only the vertical or z-component of a vector) acting at the earths surface on a fluid of density ρ would have the size

$$F_{gz} = -\frac{G\rho M}{r_E^2}$$

where G is the gravitational constant, M is the mass of the earth, and r_E its mean radius (about 6371 km). This can be written as

$$F_{gz} = -g^*\rho$$

where $g^* = GM/r_E^2$ is the gravitational acceleration. The magnitude of g^* is approximately 9.81 m s^{-2}.

The earth is not a perfect sphere, though, and the mass of the earth is not homogeneously distributed, so that the local value of the gravitational acceleration can vary by some amount. The largest deviation is due to the fact that the earth is somewhat flattened at the poles. The distance to the centre of the earth is about 20 km larger at the equator than at the poles. Since the deviations from sphericity are small compared to the mean radius of the earth, the variations of g^* are typically less than a percent.

Exercise XIII: By how much (in percent) does gravity vary between the surface of the earth (distance to the centre of the earth $\approx$ 6371 km) and at 100 km above ground?

The rotation of the earth around its axis leads to an additional force, the centrifugal force. The centrifugal force, like the Coriolis force discussed later, is an *apparent force.* It is a consequence of the fact that we are observing motions with respect to an accelerated system of reference, namely the surface of the rotating earth. Earth's rotation means that every body on its surface is moving eastward at a velocity of $u = \Omega r_E sin\phi$, where Ω is the angular frequency of Earths rotation and $r_E sin\phi$ is the distance to the axis of rotation (ϕ is the geographical latitude). Without gravitation, everything would have the tendency to continue in eastward direction in a straight line, relative to the fix stars, i.e. would fly off the earth's surface tangentially. Too keep it on the surface of the earth, the direction of the motion has to be changed continually (Figure 31). To do so, we must apply a force towards the center of rotation, the 'centripetal force'. To an observer that is fixed to the rotating earth, it seems on the contrary that the body is moved away from the surface of the earth by some force, and that he has to apply a force to keep it on the surface.

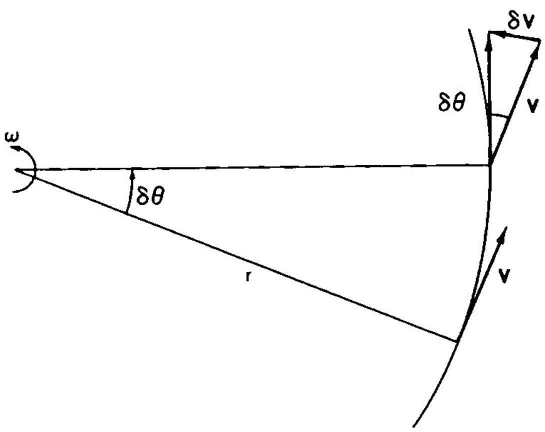


Figure 31: *To keep a body on a circular path, its velocity has to change direction continuously. This requires a force directed towards the center of rotation, the 'centripetal force'. Figure taken from J.R. Holton (1992), An Introduction to Dynamic Meteorology.*

The centrifugal force is proportional to the square of the angular frequency of rotation Ω and also proportional to the distance from the axis of rotation, $r_E \sin\phi$. Its direction is away from the axis of rotation, and its magnitude (per unit volume) is

$$F_c = \rho\Omega^2 r_E \sin\phi$$

The centrifugal force vanishes at the poles (because there the distance from the axis of rotation is zero) and is maximal it the equator. On the earth, gravity exerts the centripetal force that is required to hold things at the surface. For an observer fixed to the earth, this leads to a small apparent reduction of gravity that is largest at the equator.

Exercise XIV: How large is the centrifugal acceleration $\Omega^2 r_E \sin\phi$ at the equator? Ω is 2π divided by the rotation period (i.e. the time for one rotation of the earth around its axis (strictly speaking one has to distinguish between the time for a complete rotation with respect to the sun (solar day), and with respect to the fixed stars (sidereal day). Neglect this here!), r_E is 6371 km. Compare this to the gravitational acceleration.

The gravitational and the centrifugal acceleration are often combined into the *effective* gravitational acceleration g. Its direction is not everywhere exactly towards the centre of the earth, but very closely so (Figure 32).

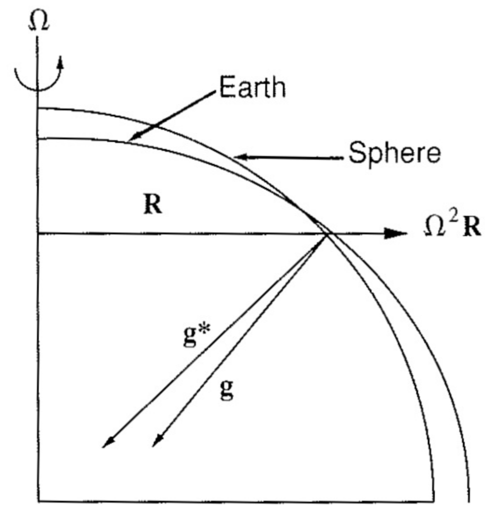


Figure 32: *The centrifugal acceleration and the gravitational attraction by the earth are often added into a resulting force, gravity. The direction of gravity slightly deviates from the direction towards the center of the earth. Figure taken from Holton (1992)*

6.3 Pressure gradient force

We usually do not think much about pressure in everyday life, because only pressure differences result in a force that we can feel. But evacuate a glass jar covered by a membrane, and there will be a force acting on the membrane that is proportional to the area of the membrane and that tries to push the membrane into the jar. The pressure is the force acting on the membrane, divided by the area of the membrane. It is an important property of pressure that is acts in all directions equally, i.e. it makes no difference in which direction we hold the jar, the force acting on the membrane remains the same. When we hold the membrane in air, there is a pressure acting on it from both sides, but there is no net force as the two pressure forces acting on either side cancel each other.

We now have to generalise this reasoning to the situation in the water, where we do not have a membrane with different pressures acting on it from either side. To do so, we use a trick that is often used in fluid dynamics, by first considering a imaginary volume in the fluid, and then letting this volume become very small.

Consider a small cubic volume in a fluid with side lengths δx, δy and δz (Figure 33). The net pressure force in x-direction acting on it is the sum of the forces acting on the two opposing sides in the same direction. These forces are the product of the area $\delta y \cdot \delta z$ with the pressure

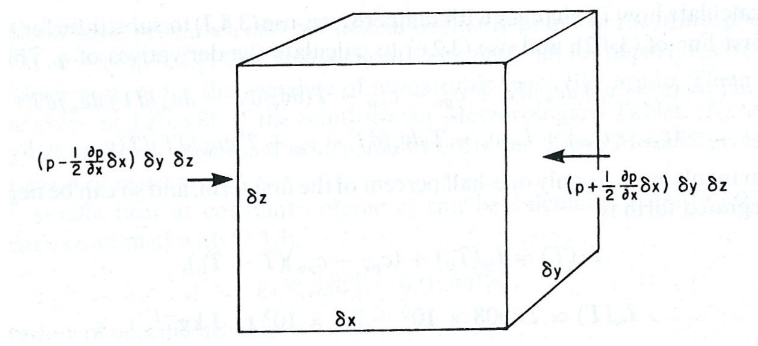


Figure 33: *Derivation of the net pressure force acting on a small volume of fluid. Figure taken from Gill (1983).*

at the respective side.

$$F_x = p(x - \delta x/2) \cdot \delta y \cdot \delta z - p(x + \delta x/2) \cdot \delta y \cdot \delta z$$

If δx is small enough we can aproximate $p(x + \delta x/2)$ by

$$p(x + \delta x/2) = p(x) + \frac{\partial p}{\partial x} \cdot \delta x/2 + \text{small rest}$$

and $p(x - \delta x/2)$ by

$$p(x - \delta x/2) = p(x) - \frac{\partial p}{\partial x} \cdot \delta x/2 + \text{small rest}$$

inserting this and dividing by the volume of the fluid in the box $\delta x \cdot \delta y \cdot \delta z$ we obtain the net pressure force (per unit volume) in x-direction:

$$F_x = -\frac{1}{\rho}\frac{\partial p}{\partial x}$$

The minus sign indicates that the force is from the region of high pressure to the region of low pressure. The same holds for the other coordinate directions, i.e. the pressure force per unit volume in y and z-directions is

$$F_y = \frac{-\partial p}{\partial y}$$
$$F_z = \frac{-\partial p}{\partial z}$$

Pressure Units

Pressure is defined as a force per surface area. Its natural unit in the SI system is the Pascal Pa = N m^{-2}, where Newton (denoted by N) is the unit for forces, which is defined as N = kg m s^{-2}. Atmospheric pressure varies around 102000 Pa or 1020 hPa (called hectopascal). Another often used unit is the bar, which is simply 100000 Pa. One bar is therefore relatively close to the average atmospheric pressure.

6.4 Coriolis force

As the centrifugal force, the Coriolis force is an apparent force, that is caused by us judging movements with respect to our environment, the rotating earth, rather that with respect to the fix stars. The Coriolis force differs from the centrifugal force by that the centrifugal force acts on us whether we move with respect to the earth or not, while the Coriolis force only appears when we move relative to the earth.

That such an apparent force must appear in a rotating environment can be visualised by children on a merry-go-round. Imagine one of the children aiming at another child and throwing a snowball at it. As soon as the snowball has left the hand it follows a straight line in the horizontal. But during its flight the two children rotate further and to them the snowball will describe a curved path. If the children do not realize that they are on a merry-go-round (as we do not sense that we are on a spinning earth even if we know it intellectually), they must ascribe the curved path of the snowball to a strange force that deflects the ball sideways (Figure 34).

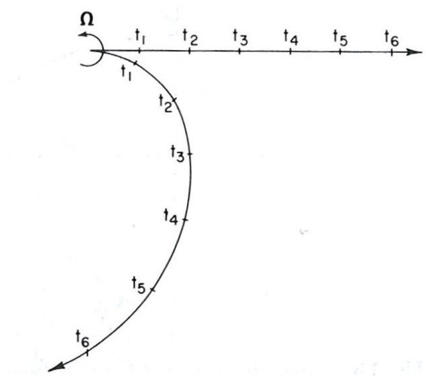


Figure 34: *Path of the snowball as seen from the outside (straight line) and from an observer on the merry-go-round. Figure taken from Holton (1992).*

On a rotating sphere rather than a rotating platform things are somewhat more complicated. A very good discussion of the effects of rotation on the spinning earth can be found in Stowe (I gave you a copy of the respective pages in the book).

Properties of the Coriolis force are:

- it is always at right angle to the direction of movement relative to the rotating earth.
- it is also always at right angle to the direction of the rotation axis of the earth.
- its magnitude is proportional to the the velocity and to the angular frequency of the earths rotation times the sine of the angle between the directions of the velocity and of the rotation axis.
- these three properties are expressed in compact form by the formula for the Coriolis

force (per unit volume)

$$\mathbf{F_c} = -2\mathbf{\Omega} \times \rho\,\mathbf{V}$$

where $\mathbf{\Omega}$ is the vector of angular frequency, $\mathbf{V}$ is the three-dimensional velocity vector relative to a fixed point on earth, and $\times$ is the vector cross-product. If you want to understand this, you have to look it up elsewhere, we do not need the formula in the following, it is here just for reference.

An important corollary from these properties is that the Coriolis force can have a vertical component even if the velocity is horizontal. This happens for example at the equator, when the velocity is horizontal along the equator. From the first two conditions it immediately follows that the Coriolis force in this case is directed upwards or downwards, depending on whether the velocity is eastwards or westwards.

It is useful to get a feeling for the order of magnitude of the Coriolis force or of the acceleration caused by it. Here is an example calculation: Assume a wind blowing with 100 km h^{-1}. The magnitude of the Coriolis acceleration resulting from that wind depends also on the angle between the velocity direction and the axis of rotation, but let us assume here that the angle is 90°, which results in the sine of the angle being one. The magnitude of the Coriolis acceleration is $a_c = 2\Omega V$, where Ω is 2π divided by the rotation period (2π / 1 day $\approx 7.3 \cdot 10^{-5}$ s^{-1}), and V is the velocity (100 km h$^{-1}\approx$ 28 m s^{-1}). The result is $a_c \approx 4 \cdot 10^{-4}$ m s^{-2}.

Therefore, the Coriolis acceleration is very small compared to the gravitational acceleration $g = 9.81$ m s^{-2}, at least for typical velocities on the surface of the earth. The Coriolis acceleration in the vertical direction can therefore be neglected to a very good approximation.

In the horizontal direction, however, there is no gravity, and Coriolis force may be important, provided all other forces are also relatively small.

This allows us to limit the discussion of Coriolis force only to its horizontal component. The properties of the horizontal component of the Coriolis force are somewhat simpler:

- The horizontal component if the Coriolis force is always at right angle to the horizontal velocity
- On the northern hemisphere it is to the right of the direction of the velocity, on the southern hemisphere to the left
- It is proportional to the velocity and to the angular frequency of the earth's rotation times the sine of the geographical latitude.

The last property means that the strength of the Coriolis force for a given velocity increases from zero at the equator to its maximal value at the poles.

The mathematical expression of the horizontal components of the Coriolis force (per unit volume) is

$$\begin{aligned} F_x &= 2\Omega \sin\phi\,\rho\,v \\ F_y &= -2\Omega \sin\phi\,\rho\,u \end{aligned}$$

where x points eastward, y northward, u is the eastward velocity, v the northward velocity, and ϕ is the geographical latitude.

$f = 2\Omega \sin\phi$ is called the *Coriolis parameter.* It is positive in the northern hemisphere, vanishes at the equator, and is negative in the southern hemisphere. It will play an important role later in section 7.8.

6.5 Friction

Friction transfers momentum from regions of high velocity to regions of low velocity, decelerating the former and accelerating the latter. A simple example is given by a layer of fluid between two parallel plates, of which the upper plate is moving while the lower is held fixed (Figure 35). We must continually push the upper plate to keep it moving because it loses its momentum through friction to the lower plate, which also has to be held fixed to keep it from accelerating.

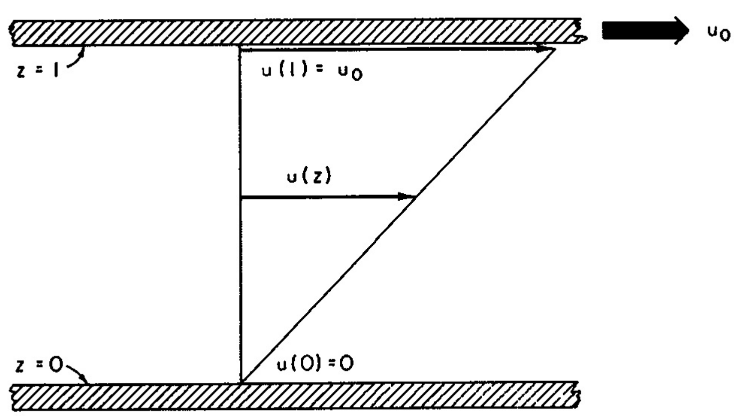


Figure 35: *Shear flow between a moving and a fixed plate. Figure from Holton (1992)*

Friction is the result of the random movement of molecules, the so-called Brownian motion. The macroscopic motion of fluid is the average over a much more irregular molecular motion of individual molecules. This irregular motion results every now and then in a molecule that comes from a region with higher velocity to enter a region with smaller velocity and to transfer its momentum to a slower molecule by collision. In this way, the region of faster average motion gets decelerated, while the region of slower motion gets accelerated.

In most fluids (these are called 'Newtonian fluids') the transfer of momentum, which physicists prefer to call a 'stress' and denote by τ, is proportional to the gradient of velocity times a quantity called viscosity. In the example in Figure 35 this is $\tau_{xz} = \mu \partial u / \partial z$ (the subscripts τ_{xz} mean that it is the transfer of the x-component of momentum into the z direction).

For the large-scale circulation in the ocean, however, the *molecular friction* is a much too weak process to be of direct importance. However, motion in the ocean is often quite turbulent. In the disordered motion that characterises turbulence, sometimes whole chunks or blobs (the noble name is 'coherent structures') of water migrate from a region of high velocity to one of lower velocity, acting in a very similar way as individual molecules in molecular friction to smooth out velocity differences (Figure 36). This process is called *turbulent friction*, and because the masses involved are greater, it is much more efficient in transferring momentum than molecular friction.

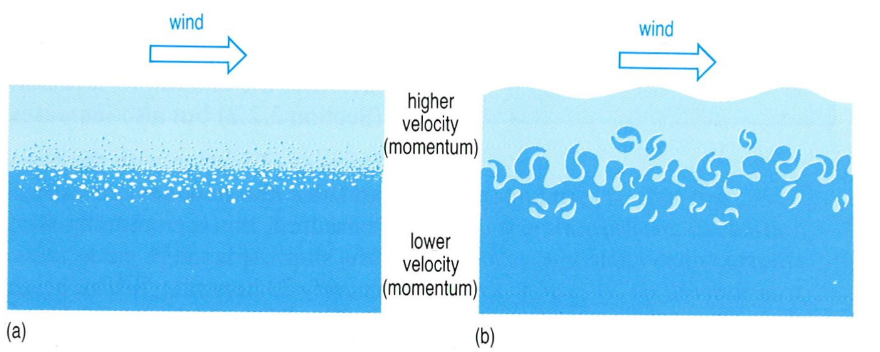


Figure 36: *Illustration of the processes of molecular vs. turbulent friction. Figure taken from Open University (1989).*

Turbulent friction is not necessarily proportional to the gradient of velocity; for simplicity, this is often assumed nevertheless, introducing a so-called 'turbulent viscosity'. However, the momentum transfer by eddies is in most cases much more effective in the horizontal than in the vertical, so one distinguishes between a horizontal and a vertical turbulent viscosity, A_h and A_v, the latter being orders of magnitude smaller than the former.

What is the relation between frictional stresses and frictional forces in a fluid? Like in the case of the pressure gradient force, we can arrive at a formulation for the net frictional force by first considering a small volume of fluid, on which frictional stresses are acting from above and below (Figure 37). By letting the sidelength of the small volume become very small, we arrive at the conclusion that only *gradients* of frictional stresses lead to a net force (per unit volume) within the fluid:

$$F_x = \frac{1}{\rho}\frac{\partial \tau_{xz}}{\partial z} = \frac{1}{\rho}\frac{\partial}{\partial z}\left(\mu\frac{\partial u}{\partial z}\right)$$

This has a simple physical interpretation: In a simple linear shear like in Figure 35 each water parcel is accelerated by the faster fluid above, but at the same time decelerated at the same rate by the slower fluid below. In total, its velocity does not change. In order for a net acceleration or deceleration to occur, the profile of velocity must be *curved*, i.e. its second derivative must be different from zero.

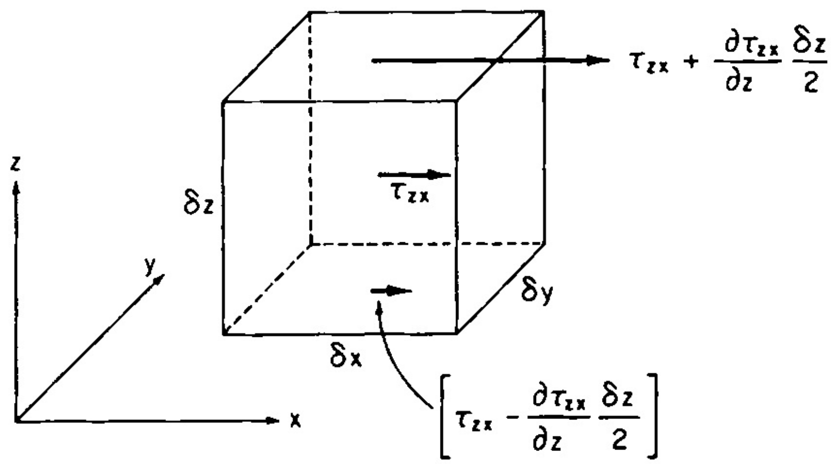


Figure 37: *The net frictional force on a small volume is the difference between the stresses at its sides. Figure from Holton (1992).*

7 The balance of forces

Although we have now discussed the four principal forces that occur in the ocean, many observed phenomena in the ocean can be explained from a balance between just two of these forces. We will consider the *hydrostatic balance* between pressure gradient and gravity, the *geostrophic balance* between pressure gradient and Coriolis force and the *Ekman balance* between friction and Coriolis force.

7.1 The hydrostatic balance

Vertical velocities and accelerations in the ocean are usually small, as well as the vertical components of friction and Coriolis force. In the vertical direction, the two remaining forces are gravity and the vertical part of the pressure gradient force. These two forces therefore have to balance each other in a steady-state situation.

The balance between these two forces *per unit mass*, i.e. accelerations can be expressed as

$$\frac{1}{\rho}\frac{\partial p}{\partial z} = -g \qquad \text{or} \qquad \frac{\partial p}{\partial z} = -g\rho \tag{1}$$

(the minus sign on the right hand side is a consequence or our choice that z increases upwards; with this convention gravity accelerates into the negative z direction).

The hydrostatic relation states that the pressure at a certain depth is equal to the weight per unit area of the water column above, plus an arbitrary constant. This can easily be seen if we first assume that density is constant with depth (which is a good first approximation). In this case dp/dz is also constant, i.e. the pressure increases linearly with depth. If we denote the atmospheric pressure at the sea surface by p_{atm} we obtain

$$p(z) = p_{atm} - g\rho z \tag{2}$$

The atmospheric pressure is roughly 1000 hPa, i.e. 10^5 N m^{-2}, where Newton N = kg m s^{-2}is shorthand for the unit of force. Taking into account that the gravitational acceleration $g \approx 9.81$ m s^{-2}, and that the density of seawater $\rho \approx 1025$ kg m^{-3}, we obtain the result that pressure increases by 10055 N m^{-2} every meter, or by about one atmospheric pressure every ten meters.

The principle that the pressure at a certain depth is equal to the weight of the layer of fluid above that depth, plus the atmospheric pressure, still holds if density increases with depth as it usually does. Let us assume first that the density is ρ_1 above a certain depth $z = \eta$ and ρ_2 below. In this case we obtain:

$$p(z) = \begin{cases} p_{atm} - g\rho_1 z & \text{if} \quad z > \eta \\ p_{atm} - g\rho_1\eta - g\rho_2(z - \eta) & \text{if} \quad z < \eta \end{cases} \tag{3}$$

In general, the fluid above the depth z is composed of many different layers of varying density. The pressure at z is then obtained by summing the density of all these layers $\rho(z)$ times the thickness dz of the layer, and multiply the sum by g. In the case that the individual layers get infinitely thin, i.e. density changes continuously with depth, this sum over an infinite number of infinitely thin layers becomes an integral

$$p(z) = p_{atm} - \int_0^z g\rho(\zeta)\, d\zeta, \tag{4}$$

which is the most general expression for the hydrostatic pressure.

7.2 Horizontal pressure gradients

Equipped with the hydrostatic balance we can now study how horizontal pressure gradients are produced within the ocean. Basically there are two possibilities: An inclination of the surface of the ocean, and an inclination of an interface between two layers of different density within the ocean.

Assume first that the density of the ocean is constant, but that the surface elevation ζ varies in an horizontal direction, say x. The height of the fluid column above a certain depth z is now $\zeta(x) - z$ (the brackets indicate that $\zeta(x)$ varies with x) and the hydrostatic pressure becomes

$$p(x, z) = p_{atm} + g\rho(\zeta(x) - z)$$

We assume here for simplicity that the atmospheric pressure does not vary with position. Nevertheless, the pressure at a constant depth then varies with x, because of the variation of ζ. The horizontal acceleration due to the pressure gradient then becomes

$$\frac{-1}{\rho}\frac{\partial p}{\partial x} = -g\frac{\partial \zeta}{\partial x}$$

i.e. it is proportional to the slope of the surface elevation.

Assume now as a second case that the surface of the ocean is flat, i.e. $\zeta = 0$ everywhere, but that the fluid consists of two layers with different densities and that the interface between these two layers, which is located at a depth η, is inclined, i.e. η again depends on horizontal position x.

The hydrostatic pressure is then given by Equations 3, where η is now assumed to depend on x. We immediately see, that the hydrostatic pressure at a certain depth depends on position only in the lower layer. The acceleration due to the pressure gradient is therefore

$$\frac{-1}{\rho}\frac{\partial p}{\partial x} = \begin{cases} 0 & \text{if} \quad z > \eta \\ g\frac{\rho_2 - \rho_1}{\rho_2}\frac{\partial \eta}{\partial x} & \text{if} \quad z < \eta \end{cases}$$

Note that the force in the lower layer is multiplied by the relative density difference $(\rho_2 - \rho_1)/\rho_2$. Because density differences in the ocean are small, an inclined density interface is much less effective in causing acceleration than an inclined surface.

These two simple cases of course almost always occur in combination, i.e. there exists both an inclined surface and inclined interfaces between layers of different density. Often the combination is such that the pressure gradient in the deep sea is very small, which requires that the inclination of the density interfaces is in the opposite direction (and much stronger to compensate for the factor involving the density difference) than the inclination of the surface. In fact, oceanographers often simply *assume* that horizontal pressure gradients below a certain depth vanish (which as we shall see below is equivalent to the assumption that motion below the same depth vanishes too). From that assumption, and the measured inclination of density interfaces within the fluid, one can then infer the direction of inclination of the surface, which is difficult to measure directly.

Exercise XV: How strong is the horizontal pressure gradient caused by a one meter elevation change of the surface of the ocean over 100 km?

7.3 Geostrophic circulation

Away from the surface or rigid boundaries, frictional forces are relatively small for the slow and relatively constant velocities that characterise the large-scale circulation of the ocean. In the horizontal direction the two dominant forces are the pressure gradient force and the Coriolis force. The balance between these two forces is called *geostrophic balance*.

It should be noted here, that all horizontal forces are usually several orders of magnitude smaller than gravity, the dominant force in vertical direction; but other horizontal forces than Coriolis or pressure gradient force are still smaller than these two.

The equilibrium between pressure gradient and Coriolis force in the horizontal can be written as

$$\frac{1}{\rho}\frac{\partial p}{\partial x} = fv \tag{5}$$

$$\frac{1}{\rho}\frac{\partial p}{\partial y} = -fu \tag{6}$$

where x points eastward, y northward, u is the eastward velocity, v the northward velocity, and f is the Coriolis parameter defined in section 6.4.

The geostrophic balance is a diagnostic relation: Given the distribution of pressure it tells us the direction and strength of the horizontal flow. The relation between pressure and flow can be summarised as follows:

- The horizontal circulation is at right angle to the pressure gradient. If e.g. pressure varies in east-west direction, the flow is in north-south direction.
- On the northern hemisphere the flow direction is such that high pressure is to the right of the direction of motion, on the southern hemisphere vice versa. Thus, on the northern hemisphere, the flow around high pressure regions is clockwise, around a low pressure region anti-clockwise.
- The flow velocity is strongest where the pressure gradient is strongest. On weather maps such as in Figure 38 these regions can be found where the lines of constant pressure are the closest together.

EXERCISE XVI: Can you see the first two statements from equations 5 and 6? Assume for simplicity that pressure is constant in north-south direction and increases to the east. On the northern hemisphere f is positive, on the southern hemisphere it is negative.

Although the geostrophic balance determines the general pattern of circulation in ocean and atmosphere, it has two limitations. The first is that it cannot be valid directly at the equator, because there f and thus the horizontal Coriolis force vanishes. At the equator thus other forces, such as friction, have to balance existing pressure gradients. The second, and much more important limitation is that, being a diagnostic relation, it *does not predict the evolution* of the distribution of pressure. Indeed, if motion was always to the right angle of pressure gradients, how could a high pressure system move from one place to another? In reality this is achieved by the small deviations from geostrophy. The geostrophic balance determines the flow to first order, but it is the small deviations thereof, that are caused mainly by inertia and friction, that determine the temporal evolution of the flow. This is what makes oceanography and meteorology interesting, but we have no space to follow that here (see e.g. the book by A.E. Gill mentioned in the epilogue).

How is the geostrophic balance established? Assume that a parcel of water is initially at rest and that pressure increases in one horizontal direction, say with increasing x. Because

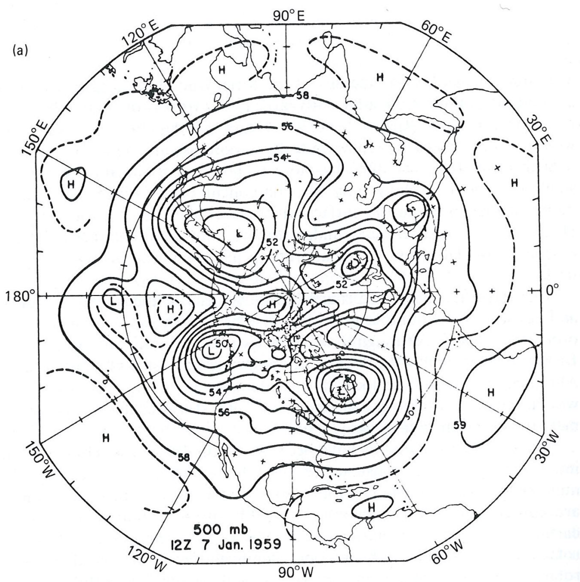


Figure 1.2.1(a) Isolines of constant pressure (isobars) at a level which is above roughly one-half the atmosphere's mass. The isobars very nearly mark the streamlines of the flow (Palmén and Newton, 1969).

Figure 38: *Map showing lines of constant pressure over the northern hemisphere for a typical weather situation. The higher pressure in the subtropics compared to the subpolar region implies a general eastward flow (i.e. westerly winds). Superimposed on that general pattern are a few high- and low-pressure systems.*

the parcel is at rest, it experiences no Coriolis acceleration, i.e. it is not in geostrophic balance. The unbalanced pressure gradient then begins to accelerate the parcel in decreasing x-direction, away from the redion of high pressure. As the parcel begins to move, it also begins to experience Coriolis acceleration; its direction of movement is deflected (to the right on the n.h.). This deflection continues until the parcel moves at right angle to the pressure gradient and the geostrophic balance is established (Figure 39).

7.4 The surface Ekman layer

During the FRAM expedition through the Arctic led by Fritjof Nansen between 1893 and 1896 it was observed that icebergs and sea ice generally do not drift in the direction of the wind but rather at an angle to the right of the wind. This was explained later (1905) by Ekman, a participant in the expedition, as resulting from a balance between the Coriolis force and friction in the uppermost layer of the ocean. Close to the ocean surface, friction cannot be neglected, because vertical shear is often high, and turbulence leads to an efficient down-gradient transport of momentum.

Ekman showed that, assuming a constant vertical turbulent viscosity, from the balance between friction and Coriolis force it follows that the flow directly at the surface must be at 45° to the right (northern hemisphere) of the wind direction, that the flow speed decreases exponentially with depth, and that the direction of the flow turns clockwise with increasing

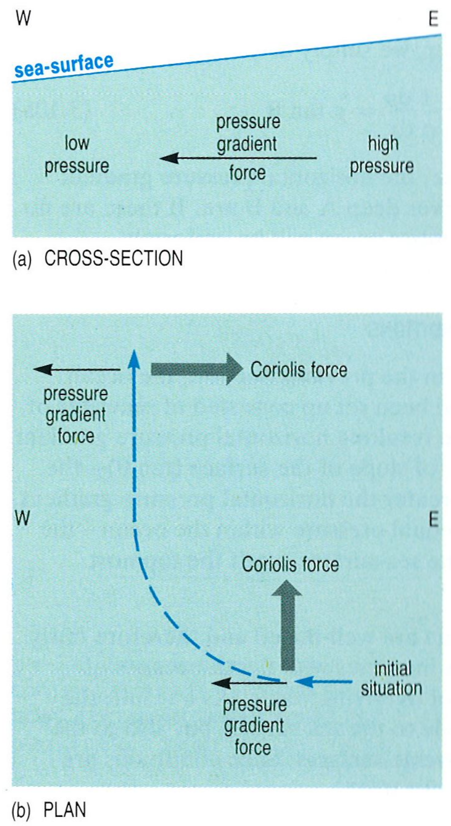


Figure 39: *Illustration of the establishment of geostrophic balance. Figure taken from Open University (1989).*

depth (Figure 40). The tip of the velocity vector therefore describes a spiral with depth.

For several reasons (unsteady winds, vertical variation of the effective viscosity, horizontal variations) this so-called Ekman spiral is, however, not observed with all of its features in nature. There are, however, some features of the Ekman spiral that are quite robust and readily observed:

- The velocity at the surface is at some angle between 0 and 90 degrees to the right of the wind (northern hemisphere). Below the surface the direction turns clockwise.
- The flow averaged over the Ekman layer is at right angle to the direction of the wind.
- The directly wind-driven flow decays rapidly with depth and is most of the times almost zero below the surface mixed layer.

The fact that the vertically averaged velocity within the Ekman layer is at right angle to the direction of the wind-stress at the surface can be understood from a force balance of the whole Ekman layer, illustrated in Figure 41: The Ekman layer experiences a wind stress at the surface in some direction (say positive x direction). At the bottom of the Ekman layer

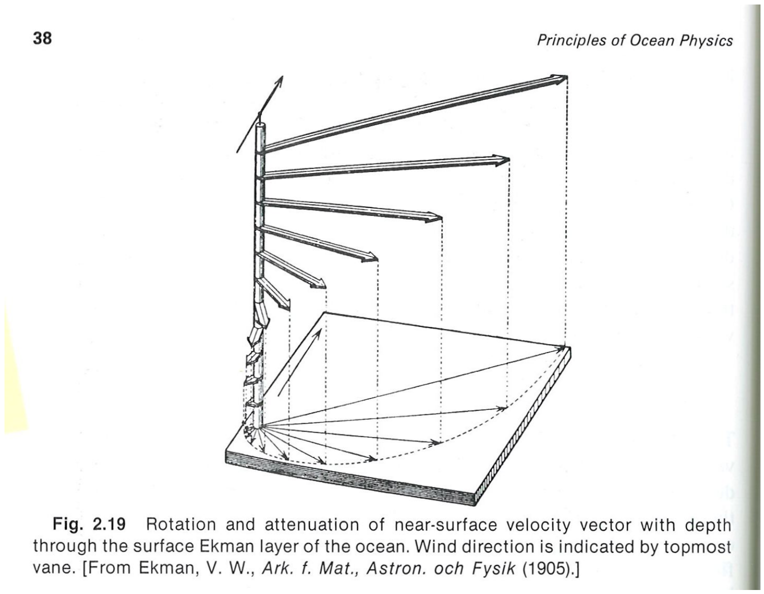


Figure 40: *The Ekman spiral in the northern hemisphere. Figure taken from Apel (1989).*

the velocity and therefore also the frictional stress vanish. The force exerted by the wind stress therefore has to be balanced by the Coriolis force averaged over the Ekman layer, that must therefore point into negative x direction. As the Coriolis force is at right angles to the direction of the flow, the vertically averaged flow must be in y direction, i.e. at right angle also to the wind stress.

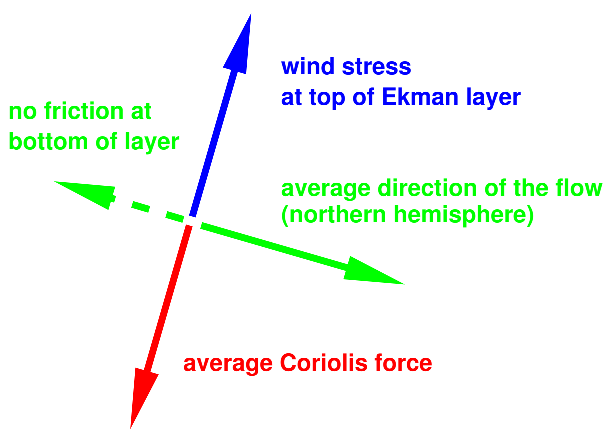


Figure 41: *Force balance averaged over the Ekman layer depth.*

EXERCISE XVII: Can you discuss in a similar way as we did for the Coriolis balance, how this balance is established? Assume that all of a sudden wind starts blowing in one direction.

7.5 Ekman convergence and divergence

The Ekman layer itself is not very deep; in most cases the Ekman flow does not reach beyond a few tens of meters depth. But the flow in the Ekman layer can enforce motion also in the deeper layers of the ocean.

The most simple example for this is wind-driven coastal up- and downwelling, illustrated in Figure 42. When the winds blow parallel to a coast, the implied Ekman transport within the Ekman layer is at right angles to the wind, i.e. either directed either towards the coast or away from it. Continuity requires that at the coast this transport has to go somewhere or come from somewhere, which can only be a deeper layer of the ocean. The convergence of Ekman transport towards the coast therefore drives downwelling, the divergence away from the coast upwelling.

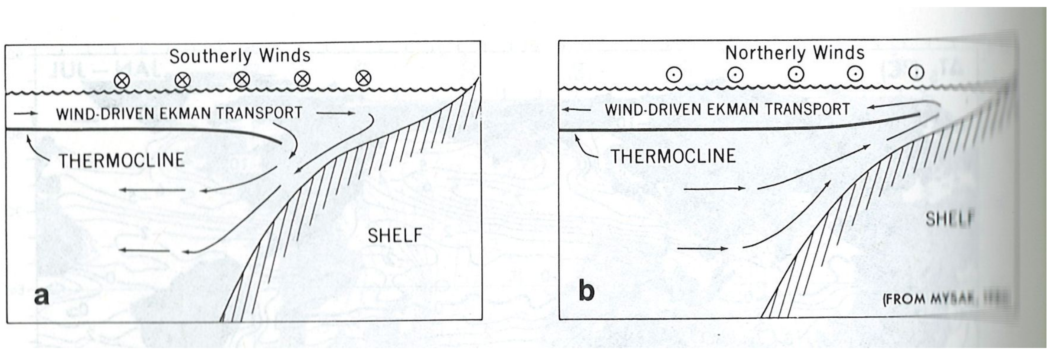


Figure 42: *Illustration of the upwelling and downwelling along a coast in the northern hemisphere. Note that southernly winds means winds blowing from south in northward direction. Figure taken from Apel (1989).*

Such coastal upwelling that brings up cold and nutrient-rich waters is behind many of the main fishing grounds, such as the coast of Mauritania, Namibia or Peru.

Exercise XVIII: What is the pattern of Ekman flow that you would expect from the trade winds along the equator? What does it imply?

7.6 The subtropical gyres, part 1: How does wind drive currents?

As already discussed in the last exercise, divergence or convergence of Ekman transports does not need a coastline to occur. An important example are the processes that lead to establishment of the subtropical gyre circulations that were described in section 5.2. Consider for example the subtropical North Atlantic. Around 15 degrees north the main direction of the trade winds is westward, and the Ekman transport (on the northern hemisphere) is therefore directed to the north. Around 45 degrees north, however, the predominant wind direction is eastward, and the implied Ekman transport to the south. In between, there is therefore a region of convergence of the Ekman transport.

This convergence leads to a piling up of water (i.e. an elevated sea surface) and a downwelling that depresses the thermocline (Figure 43).

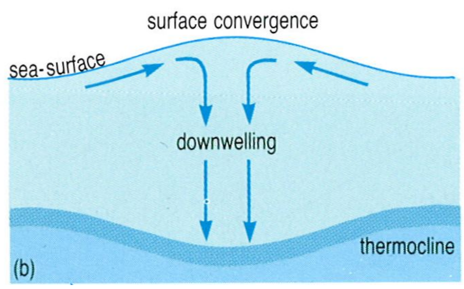


Figure 43: *Effect of a convergence of Ekman flow. Figure taken from Open University (1989).*

The elevation of the surface is difficult to measure directly (this changes with increased accuracy of sea-surface height measurements from satellite), but the depression of the thermocline becomes evident when one looks at the distribution of temperature in the upper 1000 m of the Atlantic ocean along an north-south section (Figure 44): The layer of warm water is thickest near 30° S and 30° N, in the center of the subtropical gyres, where the convergence of Ekman flow leads to downwelling of water, and thin at the equator, where the divergence of Ekman transport leads to upwelling of deep water to the surface.

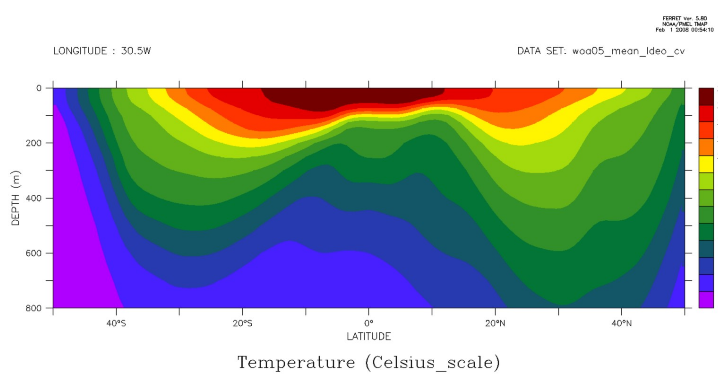


Figure 44: *Seawater temperature in the upper part of the Atlantic on a section along 30° W. Data from the World Ocean Atlas 2005.*

What is the distribution of pressure that is associated with this pattern of the surface elevation and of the thickness of the thermocline?

At a fixed depth, not too deep below the surface, the pressure is highest in the centre of the gyres, because of the weight of the surface bulge there, and becomes lower both to the north and the south of that latitude. The pressure gradient force at that depth is thus directed outwards from the convergence region, i.e. northwards north of the convergence zone, and southwards south of the centre of the gyre. This pressure gradient in turn drives a geostrophic flow. As geostrophic flow is to the right of the pressure gradient in the northern hemisphere, the direction of the flow is westward south of the convergence zone, and eastward north of it. These consequences are schematically illustrated in Figure 45.

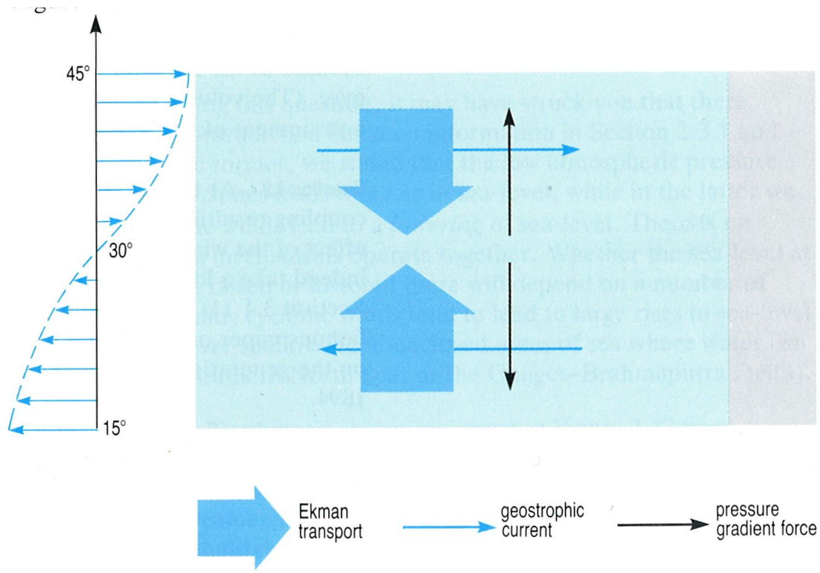


Figure 45: *Effect of a convergence of Ekman flow in an idealised North Atlantic basin. Figure taken from Open University (1989).*

Note that the geostrophic flow is in the direction of the wind, although it is *not* directly driven by it. This case is a typical illustration of how the 'wind-driven' surface circulation is in reality driven rather indirectly by the wind: The wind drives Ekman currents, the convergence of the Ekman currents creates pressure gradients, and these drive a geostrophic flow.

Deeper in the water column, the horizontal pressure gradient gets weaker because of the distribution of density caused by the downwelling: In the centre of the gyre, there is thick layer of warm and relatively light water, so that pressure increases slower with depth than at the fringes of the gyre, where denser water is closer to the surface. AS pressure gradients get weaker the deeper one goes, so does the geostrophic flow; the wind-driven current is more or less limited to a few hundred meters depth.

7.7 Part 2: Circulation in a closed basin

Section 7.6 explains how the pattern of winds in the subtropics leads to similar pattern of alternating westward and eastward geostrophic currents in the upper part of the subtropical ocean. The next thing to explain then is how these zonal currents are connected by northward and soutward flows along the continental boundaries to form closed circulation cells, the subtropical gyres seen in Figure 25. It is plausible from conservation of mass, that there must be a poleward flow along the boundary of the ocean basin (the Gulf stream in the North Atlantic, the Kuroshio in the North Pacific) and an equatorward flow near the eastern side of the basin. But how are these flows balanced, and why is the western boundary current usually so much narrower and swifter than the equatorward return flow at the eastern boundary of the Atlantic basin? This question was investigated by Henri Stommel in a classical paper (H. Stommel, 1948: The westward intensification of wind-driven ocean currents. Trans.Americ.Geophys.Un. 29, 202-206).

Stommel looked at the problem in a very idealised setting, a rectangular basin of roughly the width of the Atlantic in zonal direction and a depth of about the depth of the wind-driven circulation. To further simplify the solution, he only looked at the vertically averaged circulation, and assumed that besides the Coriolis force and pressure gradient force there is a frictional force at the top from the wind (the wind stress) and a frictional force at the bottom, which he assumed to be proportional to the velocity.

Stommel then considered three different cases: The first is a nonrotating earth, where Coriolis force vanishes, and pressure gradients therefore have to be balanced by friction. The second is a rotating earth, but assuming that the Coriolis parameter $f = 2\Omega \sin \phi$, which is the factor between velocity and Coriolis acceleration, is constant, neglecting its dependency on latitude. The third case takes into account that f varies with latitude, so that for the same velocity the Coriolis force is weaker closer to equator.

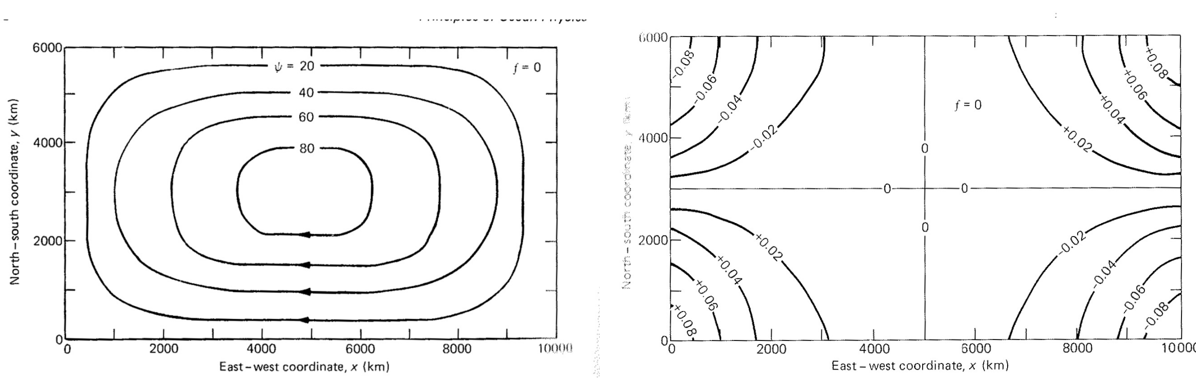


Figure 46: *Velocity streamlines (left) and surface elevation (right) in a rectangular basin driven obtained by Stommel for a nonrotating earth ($f = 0$). Figure taken from Apel (1989).*

The solutions that Stommel obtained (with the assumptions above the solutions can actually be found with pencil and paper) are shown in Figures 46 to 48. In case 1 and 2, the pattern of the circulation is identical and symmetric in east-west-direction, i.e. the northward and southward flows have equal width. But in spite of the similarity of the circulation, the elevation of the sea surface, which determines the pressure field, is completely different in the two cases: In case 1 (Figure 46), the sea surface is highest in the southwest and northeast corners of the basin and lowest in the northwest and southeast corners. This is because in the absence of rotation and thus the Coriolis force, wind and friction are balanced by pressure gradients. Wind stress is zonal, i.e. the northward and southward part of the flow must be driven by a corresponding pressure gradient. In case 2 (Figure 47), in contrast, the sea surface is highest in the center of the gyre; the pressure gradient force points outward, and the circulation is more or less (not exactly, because there is also some friction) at right angle to the pressure gradient, because of the geostrophic balance between pressure gradient force and Coriolis force.

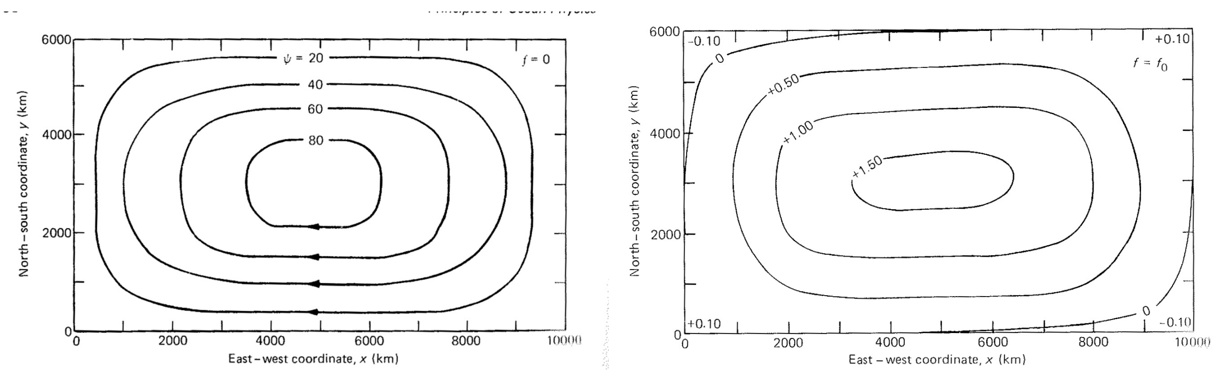


Figure 47: *Velocity streamlines (left) and surface elevation (right) in a rectangular basin driven obtained by Stommel for a uniformly rotating earth (f = const.). Figure taken from Apel (1989).*

However, the most interesting case is the third (Figure 48): In this case Stommel obtained an asymmetric circulation, with the northward flow near the western side much narrower and faster than the southward flow at the eastern side of the basin, similar to the actual patterns in the subtropical ocean gyres. Stommel thus showed that it is the *variation of the Coriolis parameter with latitude* that is responsible for the westward intensification of the gyre circulations.

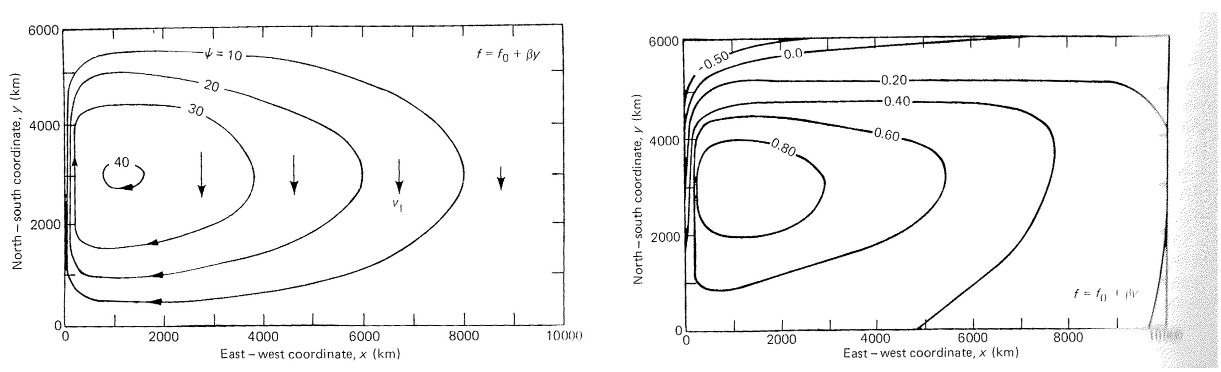


Figure 48: *Velocity streamlines (left) and surface elevation (right) in a rectangular basin driven obtained by Stommel for a vertical component of earth's rotation that varies with latitude ($f = f_0 + \beta y$). Figure taken from Apel (1989).*

7.8 Part 3: Vorticity

The results by Stommel can best be understood by introducing a new quantity, called *vorticity*. Vorticity is a generalisation of the concept of angular momentum, which is used to describe rotating rigid bodies, to a fluid. However, because in a fluid individual parts of the fluid can move relative to each other, the concept is somewhat more subtle than that of angular momentum.

We can visualize whether a two-dimensional current has vorticity with the aid of an imaginary paddle-wheel that we place in the current. If the current makes the paddle-wheel to move in an anticlockwise direction, it has positive vorticity, if the motion is anticlockwise, it has negative vorticity (Figure 49).

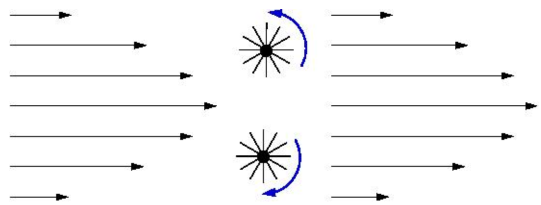


Figure 49: *A current with shear has vorticity.*

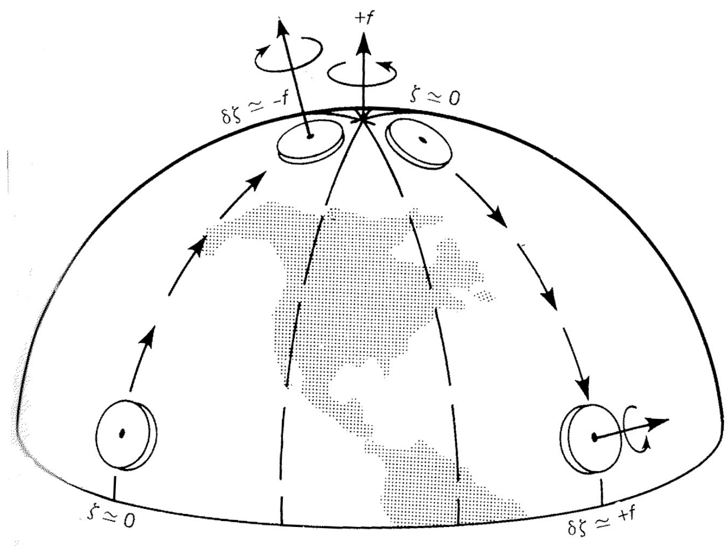


Figure 50: *A parcel of air that initially is at rest near the equator acquires negative relative vorticity when it is is moved poleward, to compensate for its gain in positive planetary vorticity. Figure taken from Apel (1989).*

7.9 Part 4: Why are western boundary currents narrow and swift?

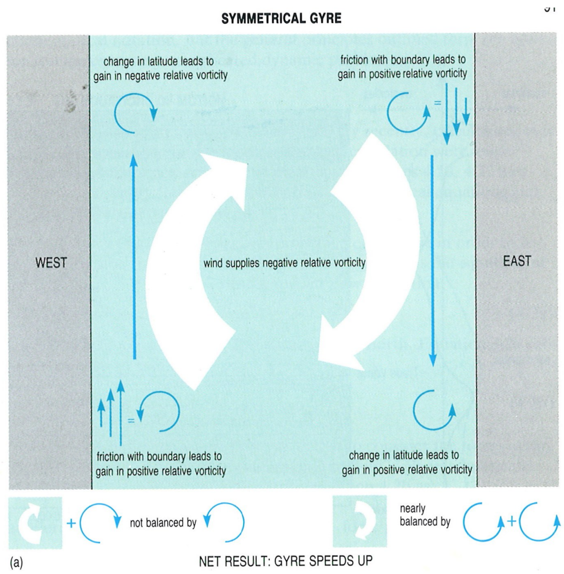


Figure 51: *Figure taken from Open University (1989).*

8 Epilogue

There are quite a number of themes that could not be treated in this course, but that are also of importance to marine biologists. Just to name a few:

- The scale-dependence of the dominant forces. One reason for the difficulty in explaining the Coriolis force is that we do not experience it in everyday life (the commonly held belief that the direction of spin of the water movement when it flows out of the bathtub depends on the hemisphere is simply wrong). Why is it that the Coriolis force is the dominant force in large-scale oceanic motion but not when we walk?
- Connected to this question is also the question what governs motion on the very small scales of microscopic organisms. For a microbe water has the properties that a very viscous fluid such as honey has to us. Why is that and does it have consequences for microbial life?
- The mixed layer at the surface of the ocean plays a large role for marine biology. What governs the dynamics of the mixed layer and motion within it?

These questions are dealt with in some depth in

K.H. Mann and J.R.N. Lazier (1996): Dynamics of Marine Ecosystems (2nd ed.). Blackwell Science, Cambridge, Massachusetts. This is not so much a book about physical oceanography, but a book that explains in much detail the role that the physics of the ocean play for marine biology, from the very small scales of the turbulence and diffusion around microplankton to the large scales of oceanic circulation gyres. The first subject is also treated in the online version of a classical lecture by Purcell:
`http://brodylab.eng.uci.edu/ jpbrody/reynolds/lowpurcell.html`

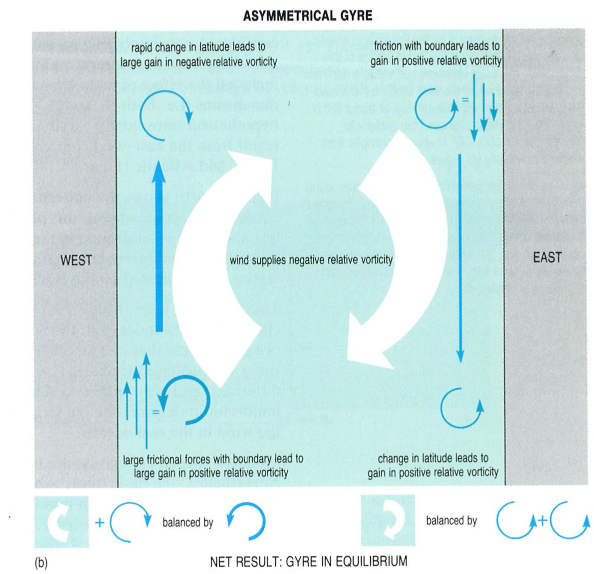


Figure 52: *Figure taken from Open University (1989).*

Also, for those who have discovered their interest in physical oceanography, and want to follow the more mathematical progression from first principles (Energy conservation, conservation of momentum etc) to the laws governing waves and circulation in the ocean (and atmosphere), two good starting points are:

J.R. Apel (1987): Principles of Ocean Physics. Academic Press, London A very broad introduction into the physical laws applied to the sea that covers also subjects that are not so often found in introductory textbooks, such as ocean optics and sound.

A.E. Gill (1982): Atmosphere-Ocean Dynamics. Academic Press, London. This book is focused more on what is usually called geophysical fluid dynamics, i.e. the explanation of dynamical features such as waves and circulation structure from physical principles.

9 Answers to Exercises

Exercise 1: Because there is also evaporation of water from the sea that removes water, but not much salt. This makes that the residence time of salts in the ocean is much larger than that of water.

Exercise 2: A kg of seawater contains about 35 g of salts, i.e. about 1000 - 35 = 965 g of water. The molar concentration of water is therefore 965 (g/kg) / 18 (g/mol) = 53.6 mol/kg.

Exercise 3: The decrease in temperature caused by a 1 degree temperature increase is lower when the initial temperature is low, it is is higher, when the initial temperature is high.

Exercise 4: The mixing product lies on the straight line between the two initial water masses. Because of the curvature of the lines of constant density, this mixing product cannot fall on the same line of constant density. Instead it is to the right of the density line, i.e. the mixing product has a higher density than the two original water masses.

Exercise 5: The outgoing atmospheric long-wave radiation Q_a is the only radiation that leaves the earth + atmosphere. It must therefore be equal to the incoming short-wave radiation. The temperature of the atmosphere T_a must therefore be equal to the radiative equilibrium temperature T_0.

Exercise 6: Cooling leads to larger areas covered by ice and snow. Larger area covered by ice and snow leads to increased albedo. Increased albedo leads to more incoming radiation reflected into space and less available for heating the earth. Less radiation heating the earth means further cooling, a positive feedback!

Exercise 7: In vacuum there can be no sensible heat flux as there is no matter to transport heat by molecular motion. There can also be no latent heat flux, as there is no water in vacuum.

Exercise 8: This is because the earths axis of rotation is tilted by about 23.5 degrees with respect to the plane in which the earth orbits the sun. The direction of the rotation axis with respect to the stars remains fixed throughout the annual movement of the earth around the sun, i.e. for a part of the year the North Pole is on the light side of the earth, while it is on the dark side during another part of the year.

Exercise 9: The irradiance at the top of the atmosphere at the North pole is about 200 W m^{-2}, the absorbed irradiance about one third of that. The difference is mostly due to reflection, with a small contribution due to absorption in the atmosphere. This means that the albedo at the North pole must be around 2/3.

Exercise 10: Evaporation in low latitudes leads to a transfer of energy from heat into latent heat. This latent heat is released as heat when the water precipitates in higher latitudes. This is a net transport of heat from low to high latitudes.

Exercise 11: The total inflow of water is $(37 + 324) \cdot 10^{12}$ $\mathrm{m}^{-3}\mathrm{year}^{-1}$, the reservoir size is $1350 \cdot 10^{15}$ m^{-3}. Dividing the reservoir size by the flux gives $1350 \cdot 10^{15}/(37+324) \cdot 10^{12} \approx 3700$

years.

Exercise 12: No; there is a net loss of water from the basin to the atmosphere equal to the difference between evaporation minus precipitation. To compensate for that, the inflow over the sill must be larger than the outflow by the same amount of water.

Exercise 13: Gravity varies inversely proportional to the square of the distance to the centre of the earth. At 100 km above ground, the distance is 6471 km. Therefore gravity is smaller by a factor of $6371^2/6471^2 \approx 0.97$, or by 3%.

Exercise 14:

$$\begin{aligned} \Omega &= \frac{2\pi}{1\,\mathrm{day}} = \frac{2\pi}{24 \cdot 3600\,\mathrm{s}} \approx 7.27 \cdot 10^{-5}\mathrm{s}^{-1} \\ \Omega^2 r_E \sin\phi &= (7.27 \cdot 10^{-5}\mathrm{s}^{-1})^2 6.371 \cdot 10^6 \mathrm{m} \sin 0^\circ \approx 0.036 \mathrm{m\,s}^{-2} \end{aligned}$$

Exercise 15: The pressure gradient is $g\rho\Delta\zeta/\Delta x$, i.e. approximately 10 m $\mathrm{s}^{-2}\times$ 1000 kg $\mathrm{m}^{-3}\times$ 1 m / 100000 m = 0.1 kg $\mathrm{m}^{-2}\mathrm{s}^{-2}$, oder 0.1 Pa m^{-1}, oder 10^{-6} bar m^{-1}.

Exercise 16: If pressure is constant in north-south direction and increases to the east, then $\partial p/\partial y = 0$ and $\partial p/\partial x > 0$. From equation 5 we then get $fv > 0$. On the northern hemisphere $f > 0$ so v must be positive, $v > 0$, and the flow is northward. On the southern hemisphere $f < 0$ i.e. $v < 0$ the flow is southward. From equation 6 we get $fu = 0$, i.e. $u = 0$.

Exercise 17: When the wind starts blowing, the water in the Ekman layer is initially at rest. The friction from the wind at the surface begins to set the water into motion in the direction of the wind. As soon as the water begins to move, it begins to experience the Coriolis force that tends to deflect the motion to the right (northern hemisphere). The percel of water therefore describes a curved trajectory being deflected more and more to the right until a balance is established between acceleration due to wind and Coriolis acceleretion.

Exercise 18: Winds near the equator generally blow in westward direction. On the northern hemisphere the Ekman flow is to the right of the wind, i.e. northwards, while it is to the left of the wind, i.e. southwards on the southern hemisphere. This divergence of the Ekman flow implies upwelling of deeper water near the equator.